PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link. http://hdl.handle.net/2066/74912

Please be advised that this information was generated on 2020-02-13 and may be subject to change.
Estimating the Knowledge-Capital Model
for Multiple Parents and Hosts:
Taking the cross-classified structure of the data into account

Maureen Lankhuizen
Eelke de Jong
Ben Pelzer

Nijmegen Center for Economics (NiCE)
Institute for Management Research
Radboud University Nijmegen

P.O. Box 9108, 6500 HK Nijmegen, The Netherlands
http://www.ru.nl/nice/workingpapers
Abstract

Data on bilateral FDI and trade flows are clustered within parent and host countries and parent-host combinations. Failure to take into account these forms of clustering will lead to biased estimates of the coefficients due to an omitted variables problem and to an underestimation of the coefficients' standard errors. Hence coefficients are incorrectly classified as significant. A cross-classified multilevel estimation procedure takes care of both problems. This paper estimates the knowledge-capital model of bilateral FDI for a sample of multiple parent and host countries by means of the cross-classified multilevel approach and compares the estimates with other techniques.

Key words: intra-class correlation, cross-classified multilevel models, foreign direct investment
Corresponding author:
Maureen Lankhuizen
Department of Spatial Economics
Free University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam, The Netherlands
Email: mlankhuizen@feweb.vu.nl

Eelke de Jong
Institute of Management Research
Department of Economics, Radboud University Nijmegen
P.O. Box 9108
6500 HC Nijmegen, The Netherlands
Email: e.dejong@fm.ru.nl

Ben Pelzer
Department of Social Science Research Methods
Radboud University Nijmegen
P.O. Box 9104
6500 HE Nijmegen, The Netherlands
Email: b.pelzer@ma.ru.nl

Acknowledgement
We thank Bruce Blonigen for providing the data.
1. Introduction
The knowledge-capital model (Markusen et al., 1996; Markusen, 1997 and 2002) is the most articulate model of bilateral foreign direct investment (FDI) tied back to microeconomic behaviour of the multinational enterprise.¹ Initial estimations of the model in Carr et al. (2001) use bilateral data for the U.S. (the U.S. is present in all country pairs, either as the parent country or as a host country). Blonigen et al. (2003) (henceforth referred to as BDH) interestingly estimate the knowledge-capital model of bilateral FDI for a sample that spans 15 OECD parent countries and 39 host countries (OECD and non-OECD). Their estimations are obtained by Ordinary Least Squares (OLS). Yet, the structure of the data on bilateral FDI is such that observations are correlated within parent countries, host countries as well as parent-host combinations. Failure to take these intra-class correlations into account will lead to biased estimates of the coefficients and to an underestimation of the coefficients’ standard errors, thereby incorrectly suggesting significant results (e.g. Snijders and Bosker, 1999).

Let us be a little more succinct. Barcikowski (1981) illustrates what happens to the true probability of making a type-I error in statistical testing when intra-class correlation is ignored. Assume that OLS results indicate that the α-value of a coefficient is 0.05. If one has a sample with N=10 observations and an intra-class correlation of 0.01, the true α is 0.06. If, however, one has 100 observations and an intra-class correlation of 0.20, the true α is 0.70. The Blonigen sample, which is also used in this paper, contains up to 2,400 observations and the correlation of two FDI values drawn from the same parent-host combination is as high as 0.78. This means that the actual significance levels are quite lower than suggested by OLS-estimates.

This paper re-estimates the knowledge-capital model for the OECD sample of BDH taking into account the cross-classified structure of data. A cross-classified multilevel model is particularly well suited for the data. A cross-classified multilevel model can account for correlation within parent countries, host countries and parent-host combinations at the same time. An alternative method for taking into account correlation in the data is using cluster-robust standard errors (e.g. Anderson and Marcouiller, 2002). Yet, cluster-robust linear regression can only take into account one cluster at a time. The multilevel model also corrects for omitted variables bias. Multilevel models account for possible unobserved parent and host country influences through random, rather than fixed, effects.²

¹ Other seminal contributions to general equilibrium models of FDI include Brainard (1993, 1997).
² Brainard (1997) reports random effects estimates. However, she uses data that are bilateral with the U.S. only (cf. Carr et al., 2001). This entails that in the U.S. outward regressions there are only host country effects to
and host countries, fixed effect estimation suffers from multicollinearity of the variables. The multilevel model on the other hand is parsimonious and avoids collinearity. In this paper we illustrate the effects of estimating the knowledge-capital model for the OECD sample using cluster-robust linear regression, fixed effect estimation and a cross-classified multilevel model.

The paper is organised as follows. Section 2 briefly describes the analysis in BDH and the data. Section 3 describes the multilevel model to be used in this paper. Section 4 presents the results of estimating the knowledge-capital model for the BDH country sample taking into account correlation of observations within parent and host countries. Section 5 summarises our main conclusions.

2. The knowledge-capital model of bilateral FDI and data
BDH estimate the knowledge-capital model for a sample of OECD countries. The knowledge-capital model integrates two motivations for FDI: to access markets in the face of trade frictions (horizontal FDI) or to exploit factor-cost differentials due to different relative factor supplies (vertical FDI). The empirical specification of the model is given by Carr et al. (2001) (henceforth CMM):

\[
FDI_{i(j)} = b_0 + b_1 \text{SUMGDP}_{i(j)} + b_2 \text{GDPDIFF}_{i(j)}^2 + b_3 \text{SKDIFF}_{i(j)} + \\
b_4 (\text{GDPDIFF}_{i(j)} \times \text{SKDIFF}_{i(j)}) + b_5 \text{INVCH}_{i(j)} + b_6 \text{TCH}_{i(j)} + \\
b_7 (\text{TCH}_{i(j)} \times \text{SKDIFF}_{i(j)}^2) + b_8 \text{TCP}_{i(j)} + b_9 \text{DIST}_{i(j)} + e_{i(j)},
\]

control for, while in the U.S. inward regressions only parent country effects are relevant. In the present case, the sample consists of multiple parent and host countries. Hence, both parent and host country effects as well as country-pair effects need to be taken into account simultaneously.

We would like to stress that the main purpose in BDH is to show that estimating a version of the knowledge-capital model with absolute size and skill differences rather than a version with simple differences, no longer supports the knowledge-capital model in favour of a horizontal model. The estimations with the OECD sample merely serve as a robustness check to strengthen this point. Accordingly, a more in-depth analysis of the estimations with the OECD sample is not given.

The knowledge-capital model by and large gives theoretical foundations for a gravity-type explanation of the pattern of FDI across countries. The empirical specification in (1) bears a gravity character with the inclusion of country size (proxied by GDP) and measures of trade and investment barriers and distance. However, contrary to common practice in gravity models, Carr et al. (2001) do not log the data, they use interactions of variables in levels to capture nonlinearities. However, the residuals from estimating equation 1 in Carr et al. “are far from white noise” (Blonigen, 2005, p. 27). This raises the issue of functional form. Blonigen and Davies (2004) report that logging the data goes a long way toward getting random residuals. The same applies to the estimations with the OECD sample, a loglinear model is more appropriate. Still, in order to be consistent with the analyses in Carr et al. (2001) and BDH we adhere to the linear form throughout the paper.
where $FDI_{n(j)}$ denotes FDI from parent $i$ to host $j$. $SUMGDP_{n(j)}$ is the sum of real GDP in both countries and captures the horizontal motives for FDI. Its coefficient is expected to be positive. $GDPDIFF_{n(j)}^2$ is the squared difference in real GDP between the parent country and the host country and is expected to have a negative influence, since theory suggests an inverted U-shaped relation to differences in country size, with a maximum at zero difference. The variables $SKDIFF_{n(j)}$ and $(GDPDIFF_{n(j)} \times SKDIFF_{n(j)})$ are the key variables that distinguish vertical FDI within the knowledge-capital model. $SKDIFF_{n(j)}$ measures the skill abundance in the parent country relative to the host country. Its coefficient is expected to be positive because the headquarters of firms are expected to be located in the skilled-labour-abundant country. The interaction term $(GDPDIFF_{n(j)} \times SKDIFF_{n(j)})$ is expected to have a negative coefficient. Relative skill abundance in the parent country is reinforced if the parent country is small ($GDPDIFF_{n(j)} < 0$) and relatively skill abundant compared to the host country ($SKDIFF_{n(j)} > 0$). The fifth and sixth variables $INVCH_{n(j)}$ and $TCH_{n(j)}$ measure the cost of investing in and exporting to the host country $j$. The coefficient of $INVCH_{n(j)}$ is expected to be negative; the cost of investing in the host country is likely to reduce FDI. The coefficient of $TCH_{n(j)}$ is expected to be positive as high trade costs will induce substitution of horizontal FDI for exports to the host market. The positive effect of $TCH_{n(j)}$ on FDI is reduced if the two countries are very dissimilar in relative endowments. If countries are dissimilar horizontal FDI will be less important. Therefore the coefficient of $(TCH_{n(j)} \times SKDIFF_{n(j)}^2)$ is expected to be negative. $TCP_{n(i)}$ is a measure of the cost of exporting to the parent country $i$, and is expected to be negatively related to $FDI_{n(j)}$ as trade costs diminish the incentive to locate plants abroad and export back to the parent country. Finally, geographical distance $DIST_{ij}$ is added to the relation. According to CMM the sign of this variable is ambiguous in theory as distance can be an element in export costs or investment and monitoring costs. In the former case, one would expect the coefficient to be positive as distance encourages the substitution of exports by FDI. In the latter case, the coefficient will be negative as investment and monitoring costs act to reduce FDI.\footnote{Notice that for most explanatory variables we use the subscript $t(ij)$ to denote the fact that these variables can change over the years for a given parent/host pair $ij$. In contrast, the variable $DIST$ (geographical distance between parent and host country) does not change over period 1982-1992 and hence the subscript $ij$ is used.}
BDH use data on outward FDI stocks in millions of real U.S. dollars. Data are from the International Direct Investment Statistics database of the OECD and cover the period 1982-1992. Data on GDP are constructed from the Penn World Tables (mark 5.6). Skilled labour endowments are proxied by the average education attainment level in number of years. Data are from the Barro and Lee dataset. A composite measure of investment openness is used which includes measures of political risk, financial risk, and other economic indicators. Investment costs are defined as 100 minus the investment openness score. In a similar vein, trade costs are defined as 100 minus a measure of trade openness, where the latter is defined as ((imports+exports)/GDP)*100. Data are from the Penn World Tables (mark 5.6). Data definitions of the indicators and descriptive statistics are presented in the Data Appendix.

BDH estimate the model using simple OLS. However, the structure of the data is such that observations on FDI are clustered for countries of origin and countries of destination. Consequently, coefficients’ estimates are biased and standard errors of the regression coefficients are underestimated when using simple OLS. In this paper we use the data set of BDH described above and re-estimate the knowledge-capital model, taking into account the structure of the data.

3. Cross-classified multilevel regression model
Multilevel problems often concern data with a hierarchical structure (Hox, 2002). Hierarchical data are found in social science research, e.g. a population consisting of schools and pupils within these schools. In this example, pupils (level 1) are nested within schools (level 2). In addition, schools could be nested in a higher level like districts or countries. If such a hierarchical data structure exists, the individual observations are generally not entirely independent, leading to intra-level correlation (Hox, 2002). However, intra-level correlation may also exist in multilevel data that are not strictly hierarchical. Cross-classified data are examples of non-hierarchical multilevel data. Cross classifications exist when several higher-level units exist next to each other, as is the case in our data sample. FDI can be grouped in three ways: by parent country, by host country and by parent-host combinations. Parent and host countries are considered to be at the same level because parent countries are not nested.

---

6 Lipsey (2001) points out that data on FDI stocks “do not purport to measure the size of multinational firms or their foreign affiliates, or their activities in their host countries. They measure only the value of the parent firms’ financial stakes in their foreign affiliates.” (p.14) However, detailed data on the activities of foreign affiliates is available for the United States, but is often sparse or unavailable for other countries. On the other hand, matters may not be so bad. Blonigen and Davis (2004, note 17) report a strong correlation between US bilateral affiliate sales and US bilateral FDI stock: 0.92 for inbound FDI activity and 0.90 for outbound FDI.
within host countries and host countries are not nested within parent countries. The crossed structure of the bilateral FDI data is illustrated in Figure 1.

In the present dataset there are two levels, years and countries. On the lower level 1 there is the FDI stock observation for each individual year \( t \); on the higher level 2 there is the mean FDI stock over the total time interval under study (1982-1992) in host country \( j \) originating in parent country \( i \). Traditional OLS regression for these data is likely to be inappropriate, since observations from the same parent and/or host country (combination) may be correlated, which will cause standard errors of regression coefficients to be underestimated. The cross-classified multilevel model that we use in this paper accounts for these correlations.

To explain the cross-classified multilevel model we depart from a baseline model that has no independent variables (predictors). We can write a baseline model expression for both levels in our data. For level-1, the baseline model for the FDI stock in a given year \( t \) in parent country \( i \) and host country \( j \) can be written as:

\[
FDI_{t(ij)} = b_0(ij) + \epsilon_t(ij),
\]

where \( \epsilon_t(ij) \sim N(0, \sigma^2_\epsilon) \).

In equation (2),

- \( b_0(ij) \) represents the 1982-1992 period mean value of the FDI stock, i.e., the mean of the yearly figures for the interval 1982-1992, in host country \( j \) originating from parent country \( i \);
- \( \epsilon_t(ij) \) represents the individual year residual, i.e. the deviation of the FDI stock in year \( t \) from mean value \( b_{0(ij)} \) for parent country \( i \) and host \( j \).

In simple OLS, instead of \( b_0(ij) \), a constant or “non-varying” intercept \( b_0 \) is used, while in multilevel models \( b_{0(ij)} \) is typically assumed to vary randomly across higher level units, which, in the present case, are host and parent countries. For level-2, the baseline model for the 1982-1992 mean FDI stock in host country \( j \) originating in parent country \( i \) is written as:
\[
\begin{align*}
    b_{0(ij)} = b_{00} + p_{0i} + h_{0j} + c_{0ij},
\end{align*}
\]  

where,

- \( b_{00} \) is the grand mean, i.e., the expected FDI amount invested by a randomly selected parent country in a randomly selected host country, in a randomly selected year.
- \( p_{0i} \) is the deviation of parent country \( i \) from grand mean \( b_{00} \), where \( p_{0i} \sim N(0, \sigma^2_{0p}) \) and \( b_{00} + p_{0i} \) is the 1982-1992 mean investment of (parent) country \( i \) in foreign countries;
- \( h_{0j} \) is the deviation of host country \( j \) from grand mean \( b_{00} \), where \( h_{0j} \sim N(0, \sigma^2_{0h}) \) and \( b_{00} + h_{0j} \) is the 1982-1992 mean investment in (host) country \( j \) by foreign countries;
- \( c_{0ij} \) is the deviation of parent-host combination \( ij \) from grand mean \( b_{00} \) that is not fully captured by the main-deviations \( p_{0i} \) and \( h_{0j} \) only, where \( c_{0ij} \sim N(0, \sigma^2_{0ph}) \).

The variances \( \sigma^2_{0p} \), \( \sigma^2_{0h} \) and \( \sigma^2_{0ph} \) represent the amount of heterogeneity in the 1982-1992 mean FDI stock values of parent countries, host countries and country pairs, respectively. By substituting equation (3) in (2), we obtain the following overall baseline model:

\[
\begin{align*}
    FDI_{t(ij)} &= b_{00} + p_{0i} + h_{0j} + c_{0ij} + e_{t(iij)},
\end{align*}
\]  

Model (4) is often called the 'random intercept' or 'variance component' model. The model decomposes the overall variance in \( FDI_{t(iij)} \) according to the four sources of variance: variance between parent countries, variance between host countries, variance between parent-host combinations and finally the variance in the yearly figures for the 1982-1992 period within each parent-host combination. In total, then, the model has five parameters, four 'random' parameters, being \( \sigma^2_{0p} \), \( \sigma^2_{0h} \), \( \sigma^2_{0ph} \) and \( \sigma^2_e \), and one 'fixed' parameter, \( b_{00} \). Since the mutual covariances of \( p_{0i} \), \( h_{0j} \) and \( e_{t(iij)} \) are assumed to be zero, the total variance in FDI is given by \( \sigma^2_{0p} + \sigma^2_{0h} + \sigma^2_{0ph} + \sigma^2_e \). Instead of (4) one may also consider a simpler cross-classified baseline model by omitting term \( c_{0ij} \) from the equation, in case the between combinations variance-source is (close to) zero.
By adding to model (4) explanatory variables we are able to, at least partially, explain the variance in FDI. Building on Carr et al. (2001), we specify the following multilevel version of the knowledge-capital model:

$$F_{DI(ij)} = b_0 + p_{0i} + h_{0j} + c_{0ij} +$$
$$b_1\text{SUMGDP}_{(i)} + b_2\text{GDPDIFF}^2_{(i)} + b_3\text{SKDIFF}_{(i)} +$$
$$b_4\text{GDPDIFF}_{(i)} \times \text{SKDIFF}_{(i)} + b_5\text{INVCH}_{(i)} + b_6\text{TCH}_{(i)} +$$
$$b_7\text{TCH}^2_{(i)} \times \text{SKDIFF}_{(i)} + b_8\text{TCP}_{(i)} + b_9\text{DIST}_{ij} +$$
$$e_{(ij)}.$$  

For a parent/host pair $ij$ the intercept of model (5) is given by the expression $b_0 + p_{0i} + h_{0j} + c_{0ij} + b_9\text{DIST}_{ij}$. The interpretation of $b_0$ has, compared to model (3), changed from 'grand mean' to 'grand intercept', with $p_{0i}$, $h_{0j}$ and $c_{0ij}$ now being parent, host and parent-host combination deviations from this 'grand intercept'. Also note that the intercept value for parent/host pair $ij$ depends on the geographical distance between the two countries.

Compared to traditional 'fixed intercept' regression, with or without correction for clustering, the multilevel approach offers the following advantages. First, by using the terms $p_{0i}$, $h_{0j}$ and $c_{0ij}$, the model accounts for possible parent and host country influences on FDI that do not appear in the set of predictors proposed by CMM or BDH. Estimation of model (5) without $p_{0i}$, $h_{0j}$ and $c_{0ij}$ could lead to a bias in the estimates of the remaining predictors' coefficients. This is known as the 'omitted variables' problem (Greene, 2000). In addition to bias, omitting relevant predictors could result in a non-normal distribution of the error terms. These problems inherent to using a traditional regression model, might be overcome by incorporating a set of dummy variables for parent and host countries (or dummies for country pairs), with one parent and host country (country pair) acting as a reference. The regression equation with dummy variables for parent and host countries is:

$$F_{DI(ij)} = b_0 + \sum_{i=2}^{I} P_i \sum_{j=2}^{J} H_j +$$
$$b_1\text{SUMGDP}_{(i)} + b_2\text{GDPDIFF}^2_{(i)} + b_3\text{SKDIFF}_{(i)} +$$
$$b_4\text{GDPDIFF}_{(i)} \times \text{SKDIFF}_{(i)} + b_5\text{INVCH}_{(i)} + b_6\text{TCH}_{(i)} +$$
$$b_7\text{TCH}^2_{(i)} \times \text{SKDIFF}_{(i)} + b_8\text{TCP}_{(i)} + b_9\text{DIST}_{ij} +$$
$$e_{(ij)}.$$  

7
Such a fixed effects estimator, could, however, cause near collinearity of the predictors, including the dummy variables. Especially if the number of FDI observations in parent/host pairs is relatively small, near collinearity can highly inflate standard errors of the coefficients and thus seriously reduce power. The gain of less biased estimators then comes with the cost of less precision and power. The multilevel approach does not suffer from collinearity, since instead of values for $p_{0i}$, $h_{0j}$ and $c_{0ij}$ for each separate country or country pair, only the variances $\sigma^2_{0p}$, $\sigma^2_{0h}$ and $\sigma^2_{0ph}$ over all countries and pairs are explicitly estimated. Hence, the multilevel model is parsimonious and at the same time avoids bias and loss of power.

A related advantage of estimating only between-parent, between-host and between parent-host combination variances $\sigma^2_{0p}$, $\sigma^2_{0h}$ and $\sigma^2_{0ph}$ is that we could have used time invariant country/country-pair attributes, e.g. geographical features. The fixed effects solution for a fixed intercept model would not allow for such predictors.

Accounting for the between-country or between-country-pair variances is the multilevel model's way to deal with the dependency in the data. Departing from baseline model (4) the correlation of two FDI values drawn from the same parent country, host country or parent-host combination can be written as $\frac{\sigma^2_{0p}}{\sigma^2_{0p} + \sigma^2_{0h} + \sigma^2_{0ph} + \sigma^2_e}$, $\frac{\sigma^2_{0h}}{\sigma^2_{0p} + \sigma^2_{0h} + \sigma^2_{0ph} + \sigma^2_e}$ and $\frac{\sigma^2_{0ph}}{\sigma^2_{0p} + \sigma^2_{0h} + \sigma^2_{0ph} + \sigma^2_e}$ respectively. As illustrated in the Introduction, ignoring these correlations even when small (<0.05) can dramatically increase the true probability of making a type-I error in statistical testing and, hence, result in incorrectly rejecting null-hypotheses.

A final advantage of the multilevel model has to do with the possibility of generalizing results to the population of parent/host countries from which our sample countries are considered to be drawn. As with all samples, generalizing makes sense if the sample can be considered representative of the population. For our data this means that results also apply to non-OECD countries for which the OECD countries in the sample can be considered representative. In contrast, generalizing to a wider population of countries is problematic with OLS and cluster-robust regression, where the countries in the sample are considered as a restrictive framework wherein the FDI data were collected and for which analysis results are to hold.
4. Estimation results

How large is the intra-correlation actually in the Blonigen dataset? As argued above, the estimates of between-country variances can be used to calculate correlation of observations within parent and host countries and parent-host combinations. Table 1 gives the estimates of the within parent-host combination variance $\sigma_{ij}^2$ as well as the variances between parent countries, host countries and parent-host combinations, $\sigma_{0p}^2$, $\sigma_{0h}^2$ and $\sigma_{0ph}^2$, respectively, for two different baseline models, one including and the other excluding the combination effect $c_{0ij}$.

Note that the model including $c_{0ij}$ has a much better fit in terms of -2LL. Based on the estimates of this model, the correlation of two FDI values drawn from the same parent country or host country is 0.09 and 0.17, respectively. The correlation of two FDI values drawn from the same parent-host combination is 0.78. All three variances are statistically significant. Recalling the illustration by Barcikowski (1981) from the introduction, we conclude that intra-class correlation is a serious issue in these data that needs to be taken into account in order to make correct inferences.

Table 2 demonstrates the effects of re-estimating the knowledge-capital model for the BDH country sample using cluster-robust linear regression, fixed effect estimation and a cross-classified multilevel model. Shaded areas in columns 2-6 indicate differences in the level of significance vis-à-vis OLS.

---

7 To test the significance of the variances we proceed as follows (see Snijders and Bosker, 1999). We compare the deviance -2LL of the model including all three effects with the corresponding -2LL value of the model without $p_{0i}$, $h_{0j}$ and $c_{0ij}$, respectively. An effect is significant if its inclusion decreases the -2LL. The difference in -2LL is Chi-squared distributed with 1 df. (In general, the number of df equals the number of parameters in an unrestricted model, including random effects, that have to be set to zero to obtain the restricted model.) In the case at hand, not including $p_{0i}$, $h_{0j}$ or $c_{0ij}$ causes the value of -2LL to increase by 25, 47 and 2118, respectively. These increases are highly significant.
Column 1 reproduces the BDH results using OLS. All coefficients, except the one on $TCH_{t(i)}$, have the correct sign. More important from the perspective of this paper, most coefficients are statistically significant. Again, $TCH_{t(i)}$ is the exception.

The easiest way to take into account the intra-class correlations is through cluster-robust linear regression. Cluster-robust linear regression adjusts the standard errors of the regression coefficients to account for the loss of independence.\(^8\) It leaves regression coefficients unaltered and thus does not correct for the omitted variables bias. Columns 2 to 4 in Table 2 present the results of OLS estimates with cluster-robust standard errors. In cluster-robust linear regression one can only use one cluster variable at a time. Hence, Table 2 presents the results of clustering by parent and host country and by country pair. As expected, the $t$-values decrease across the board. The values in columns 2-4 appropriately reflect that there is less independence in the data than is implicitly assumed by the OLS procedure. Skill differences $SKDIFF_{t(i)}$ and investment costs in the host country $INVCH_{t(j)}$ are no longer statistically significant with clustering either by parent country, host country or country pair. The variables $GDPDIFF_{t(i)}$, $TCH_{t(i)} \times SKDIFF_{t(i)}$ and $TCP_{t(i)}$ become insignificant with clustering by host country. $GDPDIFF_{t(i)}$ is also insignificant with clustering by country pair.

The results for the fixed effect estimation are listed in column 5.\(^9\) The regression in column 5 includes dummies for fourteen parent countries and thirty-eight host countries. We also ran a regression with dummies for country pairs, but such a specification failed due to multicollinearity of the predictors, including the dummies. Even in the model with parent and host-country fixed effects the tolerance for the explanatory variables is extremely low.\(^10\) This indicates that the collinearity is quite strong and coefficients are unstable. Hence, our results clearly illustrate the problems one incurs with fixed effects estimation when the number of lower level units (here: yearly FDI figures) is relatively small compared to the number of higher level units (here: countries). With fixed effects, the sign of the coefficients of $TCH_{t(i)}$.

---

\(^8\) Under Gauss-Markov conditions the variance of the OLS estimator $b$ is routinely calculated as $\sigma^2(X'X)^{-1}$. Cluster-robust standard errors use a consistent estimator $(X'X)^{-1}\sum_{j=1}^{n} e_{ij}^2 x_{ij}'(X'X)^{-1}$, where $n_c$ is the total number of clusters. The formula for the clustered estimator is that of the heteroskedasticity-consistent (White) standard errors with the individual $e_{ij} x_{ij}$'s replaced by their sums over each cluster.

\(^9\) The results for the dummy variables are available on request.

\(^10\) The tolerance for a variable is $1 - R^2$ for the regression of that variable on all the other independents, ignoring the dependent. When tolerance is close to 0 there is a high degree of multicollinearity of that variable with other independents and the coefficients will be unstable. The tolerance for the model with fixed effects ranges mostly between 0.001 to 0.195.
and $INVCH_{i(j)}$ changes vis-à-vis OLS. In the case of $TCH_{i(j)}$, the coefficient is now consistent with the predictions of the knowledge-capital model. Moreover, the effect is statistically significant. The coefficient of $INVCH_{i(j)}$ suggests that high investment costs in the host country induce FDI, which is at odds with the predictions of the knowledge-capital model. The coefficient of $INVCH_{i(j)}$ is not statistically significant though. The coefficient on $TCP_{i(j)}$ is not statistically significant in the fixed effect estimation.

The sixth column presents the results of the cross-classified multilevel estimation with parent and host country random effects. The results are qualitatively similar to the fixed effects regression. The main difference concerns the statistical significance of $SKDIFF_{i(i)}$, $TCH_{i(j)}$ and $TCP_{i(i)}$. The first two variables are significant in the fixed regression but not in the multilevel regression; $TCP_{i(i)}$ is statistically significant in the multilevel regression but not in the fixed effects regression. Similar to fixed effect estimation, the coefficient of $INVCH_{i(j)}$ is positive.

The last column in Table 2 presents the results for the cross-classified multilevel model with parent, host, and parent-host random effects. The coefficients of $SKDIFF_{i(i)}$ and $TCP_{i(i)}$ are no longer statistically significant. The coefficients on $INVCH_{i(j)}$ and $(TCH_{i(j)} \times SKDIFF_{i(i)}^2)$ are positive, i.e. contrary to the predictions of the knowledge-capital model. The coefficients are both statistically significant.

The point in this paper is that a cross-classified multilevel model is particularly well suited given the structure of our data. If we take into account the cross-classified structure of the data, there is actually less statistical support for knowledge-capital model than suggested by OLS.

To check this last result we re-ran the estimations of the knowledge-capital model for the OECD sample using alternative indicators. We use the indicators of skills, trade and investment costs used in CMM. The CMM data cover a smaller group of countries and years. The estimation sample below consists of 13 parents and 28 hosts. Details of data definitions and sources are given in the Data Appendix. Table 3 presents the results.

<insert Table 3 about here>
The results for the cross-classified multilevel model with parent, host, and parent-host random effects are given in column 7. Once again the main differences between the OLS and multilevel estimates concern the coefficients of the country specific variables. The coefficients of all three country specific variables - \( INVCH_{t(j)} \), \( TCH_{t(j)} \), and \( TCP_{t(i)} \) - change sign. Those of the host countries’ characteristics become wrongly signed whereas coefficient of parent country’s trade costs obtains its theoretically expected sign. The coefficient of \( INVCH_{t(j)} \) (wrongly signed) and \( TCP_{t(i)} \) are statistically significant. Except for the host countries’ characteristics, all coefficients have signed predicted by the knowledge-capital model.

Discussion
As expected, appropriately taking care of the cross-classified structure of the data has the most significant influence on the coefficients of the country-specific variables \( INVCH_{t(j)} \), \( TCH_{t(j)} \) and \( TCP_{t(i)} \), for which observations for a particular host or parent (in a given year) are repeated for each country pair that includes this host or parent. Hence, there is less independent information for these variables than OLS regressions assume. Cross-classified multilevel estimation frequently leads to sign changes and reduces the significance of these variables’ coefficients (as indicated by the grey areas in Tables 2 and 3).

The results reveal a striking regularity: the coefficient of \( INVCH_{t(j)} \) is positive whenever unobserved country effects are controlled for, whether through fixed or random effects. This is true for both the BDH and the CMM indicator set. These results are counter-intuitive, as they suggest that FDI from parent \( i \) to host \( j \) increases with investment costs in \( j \). We regressed FDI on investment costs in the host country for each country pair. Leaving aside whether coefficients are statistically significant, we found a positive relation in 50% of the cases.\(^{11}\) Subsequent to this finding we also estimated a specification in which we allowed the coefficient on \( INVCH_{t(j)} \) to vary randomly across parent and host countries. The results indicate that there is significant variance between host countries in the effect of \( INVCH_{t(j)} \) (see Appendix A for details). This indicates that the relation between investment costs in the host country and FDI differs across host countries and parent-host combinations.

\(^{11}\) These results are available on request.
For the knowledge capital model the coefficients of $SKDIFF_{t(ij)}$ and
$(GDPDIFF_{t(ij)} \times SKDIFF_{t(ij)})$ are important. The latter has the correct sign and remains
significant in all regressions. The sign of coefficient of the skill differences often becomes
insignificant when clustering is taken account of.

5. Conclusion
Datasets of bilateral stocks of FDI or bilateral trade flows for various parent and host
countries are characterized by correlation within countries of origin and within countries of
destination. Failure to take into account these intra-class correlations leads to biased
estimations of coefficients and to an underestimation of the corresponding standard errors of
the regression coefficients.

This paper uses the datasets of Blonigen et al. (2003) and Carr et al. (2001) to re­
estimates the knowledge-capital model of bilateral FDI taking into account the cross­
classified structure of the data. Observations are correlated within parent countries, host
countries as well as parent-host combinations. We illustrate that a cross-classified multilevel
model is particularly well suited given the structure of these two datasets. The advantage of a
cross-classified multilevel model over cluster-robust linear regression is that it can account for
correlation within parent countries, host countries and parent-host combinations at the same
time, and that it corrects for the omitted variables bias. Second, multilevel models are random
effects models. This entails that the variances over parent/host countries and country pairs are
estimated. On the one hand, accounting for the between-country or between-country-pair
variances is the multilevel model's way to deal with the dependency in the data. On the other
hand, by including random effects the model also accounts for unobserved country and
country pair influences on FDI that do not appear in the set of predictors (omitted variables).
But, different from fixed effects estimators, multilevel estimation is parsimonious and avoids
collinearity. In the two datasets used, the fixed effect estimation suffers from collinearity
between the variables.

Re-estimating the knowledge-capital model with the data as in BDH, but taking into
account the cross-classified structure of the data, we find less statistical support for the
knowledge-capital model than suggested by OLS. The coefficients on skill differences and
parent-country trade costs are statistically insignificant. The coefficients on trade and
investment costs in the host country have signs that are contradictory to the predictions of the
knowledge-capital model. Both coefficients are statistically significant. Using alternative
indicators of skills, trade and investment costs, there is relatively more support for the knowledge-capital model in the multilevel model (two indicators have an incorrect sign, one of which is statistically significant).

This paper illustrates the cross-classified multilevel model for estimating the knowledge-capital model of FDI, but the point of intra-class correlation applies in principle to any bilateral phenomena. As a result, this paper holds relevant lessons for the gravity literature in general.
References


Data Appendix
This appendix provides details of data definitions and sources used in this paper.

<table>
<thead>
<tr>
<th>Variable</th>
<th>BDH Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Real GDP constructed from Penn World Tables 5.6.</td>
</tr>
<tr>
<td>Skill abundance</td>
<td>Proxy for skilled labour endowment: average education attainment level in number of years. Barro and Lee dataset.</td>
</tr>
<tr>
<td>$INVCH_{t\mid j}$</td>
<td>FDI openness measure for host country recently obtained from Business Environment Risk Intelligence, S.A. (BERI). Includes measures of political risk, financial risk, and other economic indicators. Investment barriers are defined as 100 minus the BERI’s composite score.</td>
</tr>
<tr>
<td>$TCH_{t\mid j}$, $TCP_{t\mid i}$</td>
<td>Trade openness measure from Penn World tables, defined as $(\text{imports}+\text{exports})/\text{GDP}$. Trade costs are defined as 100 minus the trade openness measure.</td>
</tr>
<tr>
<td>$DIST_{ij}$</td>
<td>Distance between capital cities (in miles).</td>
</tr>
</tbody>
</table>

Table 1 gives the descriptive statistics for the estimation sample with the BDH indicators.
Table 1. Summary statistics for BDH estimation sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$FDI_{t(ij)}$</td>
<td>4,321.5</td>
<td>11,762.1</td>
<td>-357.1</td>
<td>176,781</td>
</tr>
<tr>
<td>$SUMGDP_{t(ij)}$</td>
<td>1674.1</td>
<td>1497.7</td>
<td>73.0</td>
<td>6,449</td>
</tr>
<tr>
<td>$GDPDIFF^2_{t(ij)}$</td>
<td>3,156,324</td>
<td>5,820,788</td>
<td>0.000122</td>
<td>2.1e+07</td>
</tr>
<tr>
<td>$SKDIFF_{t(ij)}$</td>
<td>1.65</td>
<td>2.69</td>
<td>-5.40</td>
<td>8.1</td>
</tr>
<tr>
<td>$GDPDIFF_{t(ij)} \times SKDIFF_{t(ij)}$</td>
<td>3,401</td>
<td>6,460</td>
<td>-6,996</td>
<td>31,012</td>
</tr>
<tr>
<td>$INVCH_{t(ij)}$</td>
<td>42.0</td>
<td>12.3</td>
<td>17.3</td>
<td>65</td>
</tr>
<tr>
<td>$TCH_{t(ij)}$</td>
<td>31.3</td>
<td>59.3</td>
<td>-286.2</td>
<td>87.3</td>
</tr>
<tr>
<td>$TCH_{t(ij)} \times SKDIFF_{t(ij)}$</td>
<td>422.6</td>
<td>1,050.1</td>
<td>-6,559.1</td>
<td>5,599.5</td>
</tr>
<tr>
<td>$TCP_{t(i)}$</td>
<td>52.0</td>
<td>22.3</td>
<td>-18.8</td>
<td>82.4</td>
</tr>
<tr>
<td>$DIST_{ij}$</td>
<td>6,303</td>
<td>4,792</td>
<td>174</td>
<td>18,372</td>
</tr>
</tbody>
</table>

**Variable**

Skill abundance

The sum of ISCO-68 categories 0/1 (professional, technical and kindred workers) and 2 (administrative workers) in employment divided by total employment. Data are from the Yearbook of Labour Statistics published by the International Labour Organisation.

**CMM indicators**

$INVCH_{t(ij)}$ Index ranging from 0-100 of investment impediments in the host country. Simple average of a number of indexes reported in the World Competitiveness Report. 1986, 1989-1994

$TCH_{t(ij)}, TCP_{t(i)}$ Index ranging from 0-100 of impediments to trade. Simple average of a number of indexes reported in the World Competitiveness Report. 1986, 1989-1994

Table 2 gives the descriptive statistics for the estimation sample with the CMM indicators.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDI_{r(j)}</td>
<td>5290.5</td>
<td>14049.9</td>
<td>-357.1</td>
<td>176,781</td>
</tr>
<tr>
<td>SUMGDP_{r(j)}</td>
<td>1591.1</td>
<td>1539.2</td>
<td>88.6</td>
<td>6,449</td>
</tr>
<tr>
<td>GDPDIFF_{r(j)}^2</td>
<td>3,216,244</td>
<td>6,130,371</td>
<td>0.004</td>
<td>2.1e+07</td>
</tr>
<tr>
<td>SKDIFF_{r(j)}</td>
<td>0.04</td>
<td>0.12</td>
<td>-0.26</td>
<td>0.29</td>
</tr>
<tr>
<td>GDPDIFF_{r(j)} \times SKDIFF_{r(j)}</td>
<td>57.4</td>
<td>211.6</td>
<td>-297.8</td>
<td>1,086.8</td>
</tr>
<tr>
<td>INVCH_{r(j)}</td>
<td>38.3</td>
<td>11.3</td>
<td>14.9</td>
<td>68.3</td>
</tr>
<tr>
<td>TCH_{r(j)}</td>
<td>34.7</td>
<td>12.4</td>
<td>7.86</td>
<td>81.4</td>
</tr>
<tr>
<td>TCH_{r(j)} \times SKDIFF_{r(j)}^2</td>
<td>0.56</td>
<td>0.76</td>
<td>0</td>
<td>5.50</td>
</tr>
<tr>
<td>TCP_{i}</td>
<td>34.3</td>
<td>9.46</td>
<td>14.3</td>
<td>56.6</td>
</tr>
<tr>
<td>DIST_{i}</td>
<td>6,111</td>
<td>4,963</td>
<td>174</td>
<td>18,837</td>
</tr>
</tbody>
</table>
Appendix A: Random Slopes

In this Appendix we allow the coefficient on $INVCH_{t,(j)}$ to vary randomly across parent countries, host countries and parent-host combinations. So, rather than having a ‘fixed’ parameter $b_6$ we now have:

$$b_{5(j)} = b_{50} + p_{5i} + h_{5j} + c_{5ij},$$

where,

- $b_{50}$ is the mean slope of $INVCH_{t,(j)}$ for all the host / parent combinations;
- $p_{5i}$ is the deviation of parent country $i$ from the mean slope $b_{50}$, where $p_{5i} \sim N(0, \sigma^2_{5p})$;
- $h_{5j}$ is the deviation of host country $j$ from the mean slope $b_{5j}$, where $h_{5j} \sim N(0, \sigma^2_{5h})$;
- $c_{5ij}$ is the deviation of parent-host combination $ij$ from the mean slope $b_{50}$ that is not fully captured by the main-deviations $p_{5i}$ and $h_{5j}$ only, where $c_{5ij} \sim N(0, \sigma^2_{5ph})$.

Table A1 gives the estimates of the variances.\(^{12}\) $\sigma^2_{5p}$, $\sigma^2_{5h}$ and $\sigma^2_{5ph}$ give the variance of the intercept between parent countries, host countries and parent-host combinations, respectively. $\sigma^2_{5p}$, $\sigma^2_{5h}$ and $\sigma^2_{5ph}$ give the corresponding variances in the relation between FDI and $INVCH_{t,(j)}$. Note that there is no estimate for the variance $\sigma^2_{5p}$. A model with the effect of $INVCH_{t,(j)}$ on FDI made random over parent countries, host countries and parent-host combinations failed to achieve convergence. The reason is that the random effect of $INVCH_{t,(j)}$ over parents does not exist.\(^{13}\)

Making the effect of $INVCH_{t,(j)}$ random over host countries and parent-host combinations significantly improves the fit of the model. The difference in -2LL between model (5), which only includes a random intercept and the model where the random slope is added, equals 878.7. This is highly significant for a Chi-squared distributed with 4 df.\(^{14}\)

---

\(^{12}\) Covariances of the random slopes and the random intercepts are also estimated. These are not shown.

\(^{13}\) An additional test using a model where the effect of $INVCH_{t,(j)}$ is made random over parents only confirms that the effect is insignificant.

\(^{14}\) As mentioned before, the number of df equals the number of parameters in the unrestricted model one has to set to zero to obtain the restricted model. In the present specification the number of df is four: the model with
effects are also significant individually. Not including \( h_{sj} \) or \( c_{sj} \), respectively, increases the value of \(-2LL\) by 61 and 499, respectively. These differences are highly significant for a Chi-squared distribution with 2 df. We conclude that the relation between FDI and investment costs in the host country varies significantly across host countries and parent-host combinations, both jointly and separately.

Table A1. Estimates of covariance parameters

<table>
<thead>
<tr>
<th>Level</th>
<th>Estimates model with random effects over hosts and parent-host combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>( \sigma^2_i ) 9270445</td>
</tr>
<tr>
<td>Level 2</td>
<td>( \sigma^2_{0p} ) 8158623 ( \sigma^2_{5h} ) -</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_{0h} ) 626298515 ( \sigma^2_{5h} ) 483703</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_{0ph} ) 549051076 ( \sigma^2_{5ph} ) 744857</td>
</tr>
</tbody>
</table>

Dependent variable: \( FDI_{it} \). The covariances of the random slopes and the random intercepts are not shown.

Table A2 presents the results for a model with random intercept and the effect of \( INVCH_{i(t)} \) on FDI random over host countries and parent-host combinations. The table indicates that the average relation between FDI and investment costs across all countries is 106. The coefficient is statistically insignificant.

---

random slope estimates two additional variances, \( \sigma^2_{0h} \) and \( \sigma^2_{0ph} \), and covariance of the random slope and the random intercept over host countries and parent-host combinations.
Table A2. Estimation results multilevel model with random intercept and slope

Parent, host and parent-host combinations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SUMGDP_{i(j)}$</td>
<td>12.35***</td>
<td>(19.69)</td>
</tr>
<tr>
<td>$GDPDIFF_{i(j)}$</td>
<td>-0.0009***</td>
<td>(6.36)</td>
</tr>
<tr>
<td>$SKDIFF_{i(j)}$</td>
<td>196.8</td>
<td>(0.43)</td>
</tr>
<tr>
<td>$GDPDIFF_{i(j)} \times SKDIFF_{i(j)}$</td>
<td>-1.04***</td>
<td>(6.31)</td>
</tr>
<tr>
<td>$INVCH_{i(j)}$</td>
<td>106.5</td>
<td>(0.81)</td>
</tr>
<tr>
<td>$TCH_{i(j)}$</td>
<td>-5.79</td>
<td>(0.50)</td>
</tr>
<tr>
<td>$TCH_{i(j)} \times SKDIFF_{i(j)}$</td>
<td>1.34**</td>
<td>(2.38)</td>
</tr>
<tr>
<td>$TCP_{i(i)}$</td>
<td>-43.81**</td>
<td>(2.19)</td>
</tr>
<tr>
<td>$DIST_{i}$</td>
<td>-0.64***</td>
<td>(3.91)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-88.907887*</td>
<td>(1.78)</td>
</tr>
</tbody>
</table>

Dependent variable: $FDI_{i}$

Absolute t statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%
Table 1. Estimates of covariance parameters and intra-class correlations of two baseline models

<table>
<thead>
<tr>
<th></th>
<th>Estimates of model</th>
<th>Estimates of model</th>
</tr>
</thead>
<tbody>
<tr>
<td>( FDI_{ij} )</td>
<td>( b_{0i} + p_{0i} + h_{ij} + e_{ij} )</td>
<td>( FDI_{ij} ) = ( b_{0i} + p_{0i} + h_{ij} + c_{ij} + e_{ij} )</td>
</tr>
<tr>
<td>Level 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual variance, ( \sigma^2 )</td>
<td>85393849</td>
<td>25768340</td>
</tr>
<tr>
<td>Level 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent country variance, ( \sigma^2_{0p} )</td>
<td>20392486</td>
<td>10502087</td>
</tr>
<tr>
<td>Host country variance, ( \sigma^2_{0h} )</td>
<td>38130613</td>
<td>19750797</td>
</tr>
<tr>
<td>Parent-host combination variance, ( \sigma^2_{0ph} )</td>
<td>-</td>
<td>62116747</td>
</tr>
<tr>
<td><strong>Intra-class correlation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same parent country</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Same host country</td>
<td>0.17</td>
<td>0.78</td>
</tr>
<tr>
<td>Same parent / host combination</td>
<td>-2LL</td>
<td>29423</td>
</tr>
</tbody>
</table>

Dependent variable: \( FDI_{ij} \).

Calculation of intra-class correlations:
- Same parent country: \( \frac{10502087 + 19750797 + 62116747}{25768340 + 10502087 + 19750797 + 62116747} = 0.09 \)
- Same host country: \( \frac{19750797 + 10502087 + 62116747}{25768340 + 10502087 + 19750797 + 62116747} = 0.17 \)
- Same p/h combination: \( \frac{10502087 + 19750797 + 62116747}{25768340 + 10502087 + 19750797 + 62116747} = 0.78 \)
Table 2. Estimation results BDH data set. Dependent variable $FD_{ij}$

<table>
<thead>
<tr>
<th></th>
<th>OLS (BDH)</th>
<th>Cluster-robust standard errors</th>
<th>Fixed effects model</th>
<th>Cross-class. multilevel model</th>
<th>Parent, host and parent-host combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parent (1)</td>
<td>Host (2)</td>
<td>Country pair (4)</td>
<td>Parent and host (5)</td>
<td>Parent and host (6)</td>
</tr>
<tr>
<td>$SUMGDP_{n(j)}$</td>
<td>9.28*** (25.88)</td>
<td>9.28*** (4.75)</td>
<td>9.28*** (2.33)</td>
<td>23.2*** (4.09)</td>
<td>19.1*** (24.28)</td>
</tr>
<tr>
<td>$GDPDIFF_{n(j)}$</td>
<td>-0.0007*** (7.22)</td>
<td>-0.0007** (2.03)</td>
<td>-0.0007 (1.03)</td>
<td>-0.003*** (1.43)</td>
<td>-0.002*** (21.11)</td>
</tr>
<tr>
<td>$SKDIFF_{n(j)}$</td>
<td>272.5** (2.56)</td>
<td>272.5 (1.31)</td>
<td>272.5 (1.08)</td>
<td>5,052*** (3.44)</td>
<td>435.8 (0.91)</td>
</tr>
<tr>
<td>$GDPDIFF_{n(j)} \times SKDIFF_{n(j)}$</td>
<td>-0.69*** (10.76)</td>
<td>-0.69*** (3.63)</td>
<td>-0.69* (1.92)</td>
<td>-0.83*** (13.94)</td>
<td>-0.81*** (13.61)</td>
</tr>
<tr>
<td>$INVCH_{n(j)}$</td>
<td>-46.2** (2.27)</td>
<td>-46.2 (1.16)</td>
<td>-46.2 (0.58)</td>
<td>94.3 (1.47)</td>
<td>72.7 (1.26)</td>
</tr>
<tr>
<td>$TCH_{n(j)}$</td>
<td>-4.14 (0.92)</td>
<td>-4.14 (0.61)</td>
<td>-4.14 (0.34)</td>
<td>53.2*** (2.36)</td>
<td>2.51 (0.15)</td>
</tr>
<tr>
<td>$TCH_{n(j)} \times SKDIFF_{n(j)}$</td>
<td>-1.38*** (4.78)</td>
<td>-1.38** (2.94)</td>
<td>-1.38* (1.19)</td>
<td>-1.07*** (13.30)</td>
<td>-1.22*** (13.61)</td>
</tr>
<tr>
<td>$TCP_{n(i)}$</td>
<td>-69.9*** (5.95)</td>
<td>-69.9** (2.25)</td>
<td>-69.9 (1.50)</td>
<td>-6.05 (2.44)</td>
<td>-96.0** (0.12)</td>
</tr>
<tr>
<td>$DIST_{j}$</td>
<td>-0.25*** (5.74)</td>
<td>-0.25* (1.73)</td>
<td>-0.25* (1.77)</td>
<td>-0.60*** (11.94)</td>
<td>-0.61*** (12.12)</td>
</tr>
<tr>
<td>Intercept</td>
<td>726.5 (0.77)</td>
<td>726.5 (0.27)</td>
<td>726.5 (0.19)</td>
<td>-60,279*** (7.22)</td>
<td>-9,796*** (2.79)</td>
</tr>
</tbody>
</table>

Absolute t statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Shaded areas of the columns 2-7 indicate significant differences from the BDH results. a The percentage reduction of the level-1 variance $\sigma^2_1$ in a model with explanatory variables vis-à-vis an empty model.
Table 3. Estimation results, CMM indicators of skills, trade and investment costs

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Cluster-robust standard errors</th>
<th>Fixed effects model</th>
<th>Cross-class. multilevel model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>Parent</td>
<td>Host</td>
<td>Country pair</td>
<td>Parent and host</td>
</tr>
<tr>
<td>SUMGDP&lt;sub&gt;m&lt;/sub&gt;</td>
<td>11.89***</td>
<td>11.89***</td>
<td>11.89***</td>
<td>38.51***</td>
</tr>
<tr>
<td></td>
<td>(26.48)</td>
<td>(4.95)</td>
<td>(2.33)</td>
<td>(4.19)</td>
</tr>
<tr>
<td>GDPDIFF&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-0.0016***</td>
<td>-0.0016**</td>
<td>-0.0016**</td>
<td>-0.005***</td>
</tr>
<tr>
<td></td>
<td>(13.55)</td>
<td>(2.69)</td>
<td>(1.55)</td>
<td>(2.52)</td>
</tr>
<tr>
<td></td>
<td>(3.56)</td>
<td>(2.08)</td>
<td>(0.98)</td>
<td>(1.43)</td>
</tr>
<tr>
<td>INVCH&lt;sub&gt;m&lt;/sub&gt;</td>
<td>-255.2***</td>
<td>-255.2**</td>
<td>-255.2**</td>
<td>-255.2***</td>
</tr>
<tr>
<td></td>
<td>(6.73)</td>
<td>(3.01)</td>
<td>(2.24)</td>
<td>(3.89)</td>
</tr>
<tr>
<td>TCH&lt;sub&gt;m&lt;/sub&gt;</td>
<td>64.13*</td>
<td>64.13</td>
<td>64.13</td>
<td>64.13</td>
</tr>
<tr>
<td></td>
<td>(1.89)</td>
<td>(1.36)</td>
<td>(0.79)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>TCH&lt;sub&gt;m&lt;/sub&gt; x SKDIFF&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-795.7*</td>
<td>-795.7*</td>
<td>-795.7</td>
<td>-795.7</td>
</tr>
<tr>
<td></td>
<td>(1.78)</td>
<td>(1.87)</td>
<td>(1.09)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>TCP&lt;sub&gt;m&lt;/sub&gt;</td>
<td>18.41</td>
<td>18.41</td>
<td>18.41</td>
<td>18.41</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.40)</td>
<td>(0.37)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>DIST&lt;sub&gt;g&lt;/sub&gt;</td>
<td>-0.49***</td>
<td>-0.49**</td>
<td>-0.49**</td>
<td>-0.49***</td>
</tr>
<tr>
<td></td>
<td>(7.95)</td>
<td>(2.84)</td>
<td>(2.72)</td>
<td>(2.86)</td>
</tr>
<tr>
<td>Intercept</td>
<td>2,178.13</td>
<td>2,178.13</td>
<td>2,178.13</td>
<td>2,178.13</td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(1.07)</td>
<td>(0.66)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>Observations</td>
<td>1474</td>
<td>1474</td>
<td>1474</td>
<td>1474</td>
</tr>
<tr>
<td>R²</td>
<td>0.44</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.44</td>
<td>0.70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Absolute t statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Shaded areas of the columns 2-7 indicate significant differences from the OLS results in column 1. * The percentage reduction of the level-1 variance $\sigma_e^2$ in a model with explanatory variables vis-à-vis an empty model.
Figure 1. Diagram of crossed structure

Source: adapted from Rasbash and Browne (2002).