Ruhr Economic Papers

Timo Mitze

Endogeneity in Panel Data Models with Time-Varying and Time-Fixed Regressors: To IV or not IV?

#83

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstraße 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstraße 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen)
Hohenzollernstrasse 1/3, 45128 Essen, Germany

Ruhr Economic Papers #83

Responsible Editor: Thomas K. Bauer

ISSN 1864-4872 (online) – ISBN 978-3-86788-093-0

Ruhr Economic Papers

#83

Timo Mitze

Endogeneity in Panel Data Models with Time-Varying and Time-Fixed Regressors: To IV or not IV?

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-093-0

Timo Mitze*

Endogeneity in Panel Data Models with Time-Varying and Time-Fixed Regressors: To IV or not IV?

Abstract

We analyse the problem of parameter inconsistency in panel data econometrics due to the correlation of exogenous variables with the error term. A common solution in this setting is to use Instrumental-Variable (IV) estimation in the spirit of Hausman-Taylor (1981). However, some potential shortcomings of the latter approach recently gave rise to the use of non-IV two-step estimators. Given their growing number of empirical applications, we aim to systematically compare the performance of IV and non-IV approaches in the presence of time-fixed variables and right hand side endogeneity using Monte Carlo simulations, where we explicitly control for the problem of IV selection in the Hausman-Taylor case. The simulation results show that the Hausman-Taylor model with perfect-knowledge about the underlying data structure (instrument orthogonality) has on average the smallest bias. However, compared to the empirically relevant specification with imperfect-knowledge and instruments chosen by statistical criteria, the non-IV rival performs equally well or even better especially in terms of estimating variable coefficients for time-fixed regressors. Moreover, the non-IV method tends to have a smaller root mean square error (rmse) than both Hausman-Taylor models with perfect and imperfect knowledge about the underlying correlation between r.h.s variables and residual term. This indicates that it is generally more efficient. The results are roughly robust for various combinations in the time and cross-section dimension of the data.

JEL Classification: C15, C23, C52

Keywords: Endogeneity, instrumental variables, two-step estimators, Monte Carlo simulations

January 2009

* Ruhr-University Bochum & Gesellschaft für Finanz- und Regionalanalysen (GEFRA) Münster. – The author thanks Jan Jacobs and further participants of the NAKE Research Day 2008 at Utrecht University for helpful comments and advices. – All correspondence to Timo Mitze, e-mail: Timo.Mitze@gefra-muenster.de.

1 Introduction

In contemporary panel data analysis researchers are often confronted with the problem of parameter inconsistency due to the correlation of some of the exogenous variables with the model's error term. Assuming that this correlation is typically due to unobservable individual effects (see e.g. Mundlak, 1978), a consistent approach to deal with such type of right hand side endogeneity is to apply the standard Fixed Effects Model (FEM), which uses a within-type data transformation to erase the unobserved individual effects from the model. However, one drawback of this estimator is that the within transformation also wipes out all explanatory variables that to not change in the time dimension of the model. In this case no statistical inference can be made for these variables, if they have been included in the original untransformed model based on theoretical grounds. Likewise, the Random Effect Model (REM), which rests upon the strong assumption of exogeneity of all right hand side regressors with respect to the error term, is biased for the case when endogeneity occurs.

The researcher's problem is then to find a consistent estimator, which is still capable of including time-fixed regressors in the estimation setup. A well-known example for the above sketched etimation setup in empirical work is the gravity model (of trade, capital or migration flows among other interaction effects), which assigns a prominent role given to time-fixed variables in the regression model.[1] In this paper we thus aim to focus on proper estimation strategies for Gravity type and related models, when some time-varying and -fixed right hand side regressors are correlated with the unobservable individual effects. Baltagi et al. (2003) have shown, that when there is endogeneity among the right hand side regressors the OLS and Random Effects estimators are substantially biased and both yield misleading inference. As an alternative solution the Hausman-Taylor (1981, thereafter HT) approach is typically applied. The HT estimator allows for a proper handling of data settings, where some of the the regressors are correlated with the individual effects. The estimation strategy is basically based on IV methods, where instruments are derived from internal data transformations of the variables in the model, thus no external information for model estimation is necessary. One of the advantages of the HT model is that it avoids the 'all or nothing' assumption with respect to the correlation between right hand

[1] Taking the gravity model of trade as an example, the model is a beloved playground for applied econometric work: With the recent switch from cross-section to panel data specifications, important shortcomings of earlier gravity model applications have been tackled (see e.g. Matyas, 1997, Breuss & Egger, 1999, as well as Egger, 2000), however, other methodological aspects such as the proper functional form of the Gravity equation are still subject to open debate in the recent literature (see e.g. Baldwin & Taglioni, 2006, and Henderson & Millimet, 2008, for an overview). Recently, also the time series properties of Gravity models have reached the center of academic research (see e.g. Fidrmuc, 2008, Zwinkels & Beugelsdijk, 2008).

side regressors and error components, which is made in the standard FEM and REM approaches respectively. However, for the HT model to be operable, the researcher needs to classify variables as being correlated and uncorrelated with the individual effects, which is often not a trivial task.

As a response of this drawback in empirical application of the HT approach different estimation strategies have been suggested, which strongly rely on statistical testing to reveal the underlying correlation of the variables with the model's residuals: Given the fact that the HT estimator employs variable information that in between the range of the FEM and REM, Baltagi et al. (2003) for instance suggest to use a pre-testing strategy that either converts to a FEM, REM or Hausman-Taylor type model depending on the underlying characteristics of the variable correlation in focus. The estimation strategy centers around the standard Hausman (1978) test, which has been evolved as a standard tool to judge among the use of the REM vs. FEM in panel data settings.

However, the Hausman test needs clear underlying assumptions about the consistency of estimators in comparing either the REM or HT approach with the FEM. Though the latter serves indeed as a consistent benchmark, Ahn & Low (1996) argue that the test statistics is only capable in comparing the parameter estimates of time-varying variables and not time fixed ones. The authors therefore reformulate the Hausman test in a more general framework and show that the original setup incorporates and tests only a very limited set of moment conditions among a much broader pool of IV-set candidates. The latter reformulation of the Hausman tests rests on the Sargan (1958) / Hansen (1982) statistic of testing for overidentifying restrictions. Together with the closely related C-Statistic derived by Eichenbaum et al. (1998), which allows for testing single instrument validity rather than full IV-sets, the Hansen-Sargen overidentification test may thus be seen as a more powerful tool to guide IV selection in the HT approach compared to the standard Hausman test.

As an alternative to IV estimation different 'two-step'-type estimators have been proposed recently: Plümper & Tröger (2007) for instance set up an augmented FEM model that also allows for the estimation of time-fixed parameters. Their model - labeled Fixed Effects Vector Decomposition (FEVD) - may be seen as a rival specification for the HT approach in estimating the full parameter space in the model including both time-varying and time-fixed regressors. The idea of FEM based two step estimators is thereby to first run a consistent FEM model to obtain parameter estimates of the time-varying variables. Using the regression residuals as a proxy for the unobserved individual effects in a second step this proxy is regressed against the set of time-fixed variables to obtain parameter values for the latter. Since this second step includes a 'generated regressand' (Pagan,

1984) the degrees of freedom have to be adjusted to avoid an underestimation of standard errors (see e.g. Atkinson & Cornwell, 2006, for a comparison of different bootstrapping techniques to correct standard errors in these settings). Though it is typically argued that one main advantage of these non-IV estimators is their freedom of any arbitrary classification of hand side regressors as being endogenous or exogenous, as we will show latter on two-step estimators such as the FEVD also rests upon an implicit choice that may impact upon estimator consistency and efficiency.

Giving the growing number of empirical applications of the latter non-IV approach (see e.g. Belke & Spies, 2008, Caporale et al., 2008, Etzo, 2007, and Krogstrup & Wälti, 2008, among others),[2] a systematic comparison of the HT instrumental variable approach with the non-IV FEVD is of great empirical interest. However, there are relatively few existing studies comparing the two-step estimators with the Hausman-Taylor IV approach in a Monte-Carlo simulation experiment (in particular Plümper & Tröger, 2007, as well as Alfaro, 2006), which show somewhat heterogeneous results concerning estimator superiority. Moreover, in these studies as well as the broader Monte Carlo based evidence on the Hausman-Taylor estimator (see e.g. Ahn & Low, 1996, Baltagi et al., 2003), the empirically unsatisfactory assumption is made that the true underlying correlation between right hand side variables and error term is known. Our approach therefore explicitly offsets from earlier simulation studies and allows for the existence of imperfect knowledge in the HT model estimation with IV selection based on different model/moment selection criteria (see e.g. Andrews, 1999, Andrews & Lu, 2001). The latter combines information from the Hansen-Sargan overidentification test and time-series information-criteria such as AIC/BIC. This allows for an empirical comparison of the HT and FEVD (two-step) estimators' performance, which comes much closer to the true estimation problem researchers face in applied modelling work in terms of 'To IV or not IV?'.

The remainder of the paper is organized as follows: In section 2 we briefly outline the general panel data model of interest and describe the two alternative estimation strategies. In section 3 we discuss the Hausman test and the Hansen-Sargan overidentification test as model/moment selection criteria. Section 4 presents the design and results of our Monte Carlo simulation experiment. For the Monte Carlo simulations we propose different model/moment selection algorithms for the HT model based on statistical criteria. Section 5 adds an empirical application to trade estimates in a gravity model context for German regions (NUTS1-level) within the EU27. Section 6 finally concludes.

[2]The FEVD approach is available as a user-written Stata routine upon request from the authors, which additionally facilitates a widespread empirical usage. Searching for the term "Fixed Effects Vector Decomposition"by now gives almost 200 entries in Google.

2 The model and panel data estimation techniques

We consider a general static (one-way) panel data model of the form

$$y_{it} = \beta X_{it} + \gamma Z_i + u_{it} \text{ with: } u_{it} = \mu_i + \nu_{it} \tag{1}$$

where $i = 1, 2, \ldots, N$ is the cross-section dimension and $t = 1, 2, \ldots, T$ the time dimension of the panel data. X_{it} is a vector of time-varying variables, Z_i is a vector of time invariant right hand side variables, β and γ are coefficient vectors. The error term u_{it} is composed of two error components, where μ_i is the unobservable individual effect and ν_{it} is the remainder error term. μ_i and ν_{it} are assumed to be $iid(0, \sigma_\mu)$ and $iid(0, \sigma_\nu)$ respectively.

Standard estimators for the panel data model in eq.(1), which control for the existence for individual effects, are the FEM and REM approach. To assess the main difference between the two estimators it is helpful to point out the underlying assumptions about the correlation of $X_{i,t}$ and Z_i with μ_i and $\nu_{i,t}$ respectively:

$$\begin{array}{rccr}
 & \textbf{REM} & \textbf{FEM} & \\
E(X_{i,t}\mu_j) & = 0 & \neq 0 & \forall i, j, t \quad (2) \\
E(X_{i,t}\nu_{j,s}) & = 0 & = 0 & \forall i, j, t, s \\
E(Z_i\mu_j) & = 0 & \neq 0 & \forall i, j \\
E(Z_i\nu_{j,s}) & = 0 & = 0 & \forall i, j, s
\end{array}$$

While both estimators generally assume that the all r.h.s. regressors are exogenous with respect to the remainder error term $\nu_{i,t}$, the assumption about respective variable correlation with the unobservable individual effect μ_i differs significantly: The REM assumes strict exogeneity of all regressors, while the FEM approach leaves the correlation as unknown and thus potentially different from zero.

As Boumahdi & Thomas (2008) show, based on the above stated assumptions about the underlying variable correlation with the error term most panel data estimators for eq.(1) can be written in terms of a general orthogonality (moment) condition in matrix form as:

$$E[S'(tU)] = 0 \; \frac{1}{N} S'tY = \frac{1}{N} S'tW\delta \tag{4}$$

where S is the $(NTxL)$ instrument matrix, t the $(NTxNT)$ transformation operator,

$W = (X, Z)$ and $\delta = (\beta', \gamma')$. If we define $Q = I_{NT} - B$ and $B = I_N(1/T)e_T e_T'$ as within and between matrix operators respectively, where I_{NT} and I_N are identity matrices of order NT and N respectively and e_T is a T vector of ones, we can then write the FEM and REM model in the notation of eq.(2) as:

- **FEM:** $S = X$, $T = Q$ and $\hat{\delta}_{FEM} = (X'QX)^{-1}X'QY$
- **REM:** $S = W$, $T = \Omega^{-1/2}$ and $\hat{\delta}_{REM} = (W'\Omega^{-1}W)^{-1}W'\Omega^{-1}Y$

where Ω is the variance-covariance matrix of the error term U. While the FEM uses only deviations from group means of X as instruments (the vector Z cancels out), the REM uses all available information in terms both deviations and group means as valid instrumental variables. Therefore the FEM is always a consistent estimator estimator, if there is any correlation of one of the variables with the unobservable individual effects (μ_i). On the contrary, if the strict exogeneity assumption of the REM approach is fulfilled, the latter is more efficient than the FEM since it employs more information for estimation. Testing for the validity of FEM/REM assumptions is typically done based on the Hausman (1978) exogeneity test, which we will describe more in detail in section 3.

As eq.(2) shows, choosing among the FEM and REM estimator is linked to an 'all or nothing' decision with respect to the assumed correlation of right hand side variables with the error term. However, the empirical truth may often lie in between these two extremes. This ideas motivates the specification of the Hausman–Taylor (1981) model as a hybrid version of the FEM/REM using IV techniques. HT approach therefore simply split the set of time varying variables into two subsets $X_{i,t} = [X1_{i,t}, X2_{i,t}]$ with:

$$\begin{aligned} E(X1_{i,t}\mu_j) &= 0, \quad \forall i,j,t \\ E(X1_{i,t}\nu_{j,s}) &= 0, \quad foralli,j,t,s \\ E(X2_{i,t}\mu_j) &\neq 0, \quad \forall i,j,t \\ E(X2_{i,t}\nu_{j,s}) &= 0, \quad \forall i,j,t,s \end{aligned} \tag{5}$$

$X1$ are supposed to be exogenous w.r.t μ_i and $\nu_{i,t}$. $X2$ variables are correlated with μ_i and thus endogenous w.r.t. the unobserved individual effects.[3] An analogous classification is done for the set of time–fixed variables $Z_i = [Z1_i, Z2_i]$:

$$E(Z1_i\mu_j) = 0, \quad \forall i,j \tag{6}$$

[3]Here we use the terminology of 'endogenous' and 'exogenous' to refer to variables that are either correlated with the unobserved individual effects μ_i or not. An alternative classification scheme used in the panel data literature classifies variables as either 'doubly exogenous' with respect to both error components μ_i and $\nu_{i,t}$ or 'singly exogenous' to only ν. We use these two definitions interchangeably here.

$$
\begin{aligned}
E(Z1_i \nu_{j,s}) &= 0, \quad \forall i, j, s \\
E(Z2_i \mu_j) &\neq 0, \quad \forall i, j \\
E(Z2_i \nu_{j,s}) &= 0. \quad \forall i, j, s
\end{aligned}
$$

Note that the presence of $X2$ and $Z2$ is the cause of the bias in the standard REM approach. The resulting augmented HT model can be written as:

$$
y_{i,t} = \alpha + \beta_1' X1_{i,t} + \beta_2' X2_{i,t} + \gamma_1' Z1_i + \gamma_2' Z2_i + u_{i,t}. \tag{7}
$$

The idea of HT model is to find appropriate internal instruments to estimate all model parameters. Thereby, deviations from group means of $X1$, $X2$ serve as instruments for $X1$ and $X2$ (in the logic of the FEM), $Z1$ serve as their own instruments and group means of $X1$ are used to instrument the time-fixed $Z2$. The FEM and the REM can be derived as special versions of the HT model, namely when all regressors are correlated with the individual effects the model reduces to the FEM. For the case that all variables are exogenous (in the sense of no correlation with the individual effects) the model takes the REM form.

In empirical terms the HT model is typically estimated by GLS and throughout the paper we use a generalized instrumental variable (GIV) approach, which can be summarized as:[4]

- **HT:** $S = [QX1, QX2, PX1, Z1]$, $T = \Omega^{-1/2}$ and

$$
\hat{\delta}_{HT-GIV} = [W'\Omega^{-1}S(S'\Omega^{-1}S)^{-1}S'\Omega^{-1}W]^{-1}W'\Omega^{-1}S(S'\Omega^{-1}S)^{-1}S'\Omega^{-1}y \tag{8}
$$

The GIV estimator originally proposed by White (1984) applies 2SLS to the GLS filtered model (including the instruments) as:[5]

$$
\tilde{y}_{i,t} = \tilde{\alpha} + \beta_1' \widetilde{X1}_{i,t} + \beta_2' \widetilde{X2}_{i,t} + \gamma_1' \widetilde{Z1}_i + \gamma_2' \widetilde{Z2}_i + \tilde{u}_{i,t}, \tag{9}
$$

where $\tilde{y}_{i,t}$ denotes the following transformation for a variable $\tilde{y}_{i,t} = \hat{\Omega}^{-1/2} y_{i,t}$. $\hat{\Omega}$ is the estimated variance-covariance matrix of the model defined as $\hat{\Omega} = \hat{\sigma}_\mu^2 (I_N \otimes J_T) + \hat{\sigma}_\nu^2 (I_N \otimes I_T)$ with $\hat{\sigma}_\nu^2 = [(\hat{u}'Q\hat{u})]/N(T-1)$ and $\hat{\sigma}_\mu = [(\hat{u}'P\hat{u}) - (N\hat{\sigma}_\nu^2)]/NT$ being consistent estimates of the variance terms of the error components (for details see e.g. Baltagi, 2008). Finally, the order condition for the HT estimator to exist is $k_1 \geq g_2$. That is, the total number of

[4] On has to note that the IV set is based on the interpretation of Breusch et al. (1989).

[5] One also has to note that the HT model can also be estimated based on a slightly different transformation, namely the filtered instrumental variable (FIV) estimator. The latter transforms the estimation equation by GLS but uses unfiltered instruments. However, both approaches typically yield similar parameter estimates, see Ahn & Schmidt (1999).

time-varying exogenous variables k_1 that serve as instruments has to be at least as large as the number of time invariant endogenous variables (g_2).[6] For the case that ($k_1 > g_2$) the equation is said to be overidentified and the HT estimator obtained from a 2SLS regression is generally more efficient than the within estimator (see also Baltagi, 2008).

In empirical application of the HT approach the main points of critique focus on the arbitrary IV selection in terms of $X1/X2$ and $Z1/Z2$ variable classification as well as the poor small sample properties of IV–methods when instruments are 'weak' as well as similar small sample problems of the GLS estimator. Therefore, recent two-step non-IV alternatives such as the Fixed Effects Vector Decomposition (FEVD) by Plümper & Tröger (2007) have been proposed.[7] The goal of the model is to run a consistent FEM model and still get estimates for the time-invariant variables. The intuition behind the FEVD specification is as follows: Since the unobservable individual effects capture omitted variables including the effect of time-invariant variables, it should therefore be possible to regress a proxy of the individual effects obtained from a first stage FEM regression on the time-invariant variables to obtain estimates for these variables in a second step. Finally, the number of degrees of freedom for the use of a 'generated regressand' in this second step has to be corrected (e.g. by bootstrapping methods, see Atkinson & Cornwell, 2006). We can thus sum up the FEVD estimator as:

- **FEVD:** 1.) Run a standard FEM as described above to get parameter estimates ($\hat{\beta}_{FEVD}$) of the time-varying variables.
- **FEVD:** 2.) Use the estimated group residuals as a proxy for the time-fixed individual effects $\hat{\pi}_i$ obtained from the first step as $\hat{\pi}_i = (\bar{y}_i - \hat{\beta}_{FEM}\bar{X}_i)$ to run a OLS regression of the explanatory time-invariant variables against this vector to obtain parameter estimates of the time-fixed variables:

$$\hat{\gamma}_{FEVD} = (Z'Z)^{-1}Z'\hat{\pi} \tag{10}$$

The residual term from this 2. modelling step $\hat{\eta}_i$ is composed of $\hat{\eta}_i = \zeta_i + \bar{X}_i(\hat{\beta}_{FEM} - \beta)$, where $\zeta_i = \mu_i + \bar{\nu}_i$ and the over-bar indicates the sample period mean for cross-section i e.g. $\bar{X}_i = 1/T\sum_{t=1}^{T} X_{i,t}$ (for details see Atkinson & Cornwell, 2006). Plümper & Tröger also propose a third (optional) step to control for collinearity between time-varying and time-fixed right hand side variables in a pooled OLS setup as:

[6]The total number of IVs in the HT model is $2k_1 + k_2 + g_1$ ($k_1 + k_2$ from $QX1$ and $QX2$, k_1 from $PX1$ and g_1 from $Z1$)
[7]The FEVD may be seen as an extension to an earlier model in Hsiao (2003). For details see Plümper & Tröger (2007).

$$y_{i,t} = \alpha + \beta' X_{i,t} + \gamma' Z_i + \hat{\eta}_i + e_{i,t}, \tag{11}$$

In either the 2. or 3. step also standard errors have to be corrected for $\hat{\gamma}_{FEVD}$ either asymptotically or by bootstrapping techniques (see Murphy & Topel, 1985, as well as Atkinson & Cornwell, 2006) to avoid an overestimation of t-values. To sum up, the FEVD 'decomposes' the vector of unobservable individual effects into a part explained by the time invariant variables and an error term. Since the FEVD is built on the FEM it yields unbiased and consistent estimates of the time-varying variables. According to Plümper & Tröger one major advantage of the FEVD compared to the Hausman-Taylor model is that the estimator does not require prior knowledge of correlation between the explanatory variables and the individual effects. Moreover, the specification relies on the robustness of the FEM and does not need to meet the strong orthogonality assumptions of the REM.

However, estimates of the time invariant variables are only consistent if either the time invariant variables fully account for the individual effects or the unexplained part of η_i is uncorrelated with the time-invariant variables. As Caporale et al. (2008) note, otherwise the FEVD also suffer from omitted variable bias.[8] To make this point clear we can write the FEVD model in terms of the following moment conditions:

$$\begin{aligned} E(X_{i,t}\nu_{i,t}) &= 0, \\ E(X_{i,t}\mu_i) &\neq 0, \\ E(Z_i\mu_i) &= 0. \end{aligned} \tag{12}$$

The latter orthogonality condition can be obtained from the definition of η_i above. Thus, though we are not directly confronted with the choice of classifying variables as endogenous or exogenous, the estimator itself does rely on an implicit choice: In specifying the time-varying variables the model follows the generality of the FEM approach, which assumes that these variables are possibly correlated with the unobservable individual effects (for estimation purposes deviations from group means are taken which wipe out the individual effects so that no explicit assumption about the underlying correlation needs to be stated). With respect to the time invariant variables the estimator assumes in its simple form that no time-fixed variable (Z) is correlated with the the second step error term, which is composed of the unobservable individual effects. However, if this

[8] A modification of the standard FEVD approach also allows for the possibility to estimate the second step as IV regression and thus account for endogeneity among time invariant variables and η_i. Following Atkinson & Cornwell, 2006, we can define a standard IV estimator as: $\hat{\gamma}_{FEVD} = \left(S'Z\right)^{-1} S'\hat{\pi}$, where S is the instrument set that satisfies the orthogonality condition $E(S\eta) = 0$. However, this brings back the classification problem of the HT approach, which we aim to avoid here.

implicit (and fixed) choice does not reflect the true correlation between the variables and the individual effects the estimator may in fact have lower power than the HT approach.

We carefully examine the relative performance of the two estimators in the Monte Carlo simulations. Reducing the empirical problem to the question of finding a proper instrument set (leaving the different small sample properties of IV and non-IV methods for the moment apart), the above results also advice to test IV validity before choosing among one of the rival estimators based on statistical testing. We will thus turn to statistical moment selection criteria in the next section.

3 Moment Selection Criteria

To judge among the estimators' orthogonality assumption different specification tests have been proposed. The most prominent example is the Hausman (1978) exogeneity test based on the m-statistic defined as:

$$m = \hat{q}'(\hat{Q} - \hat{V})^{-1}\hat{q}, \tag{13}$$

with $\hat{q} = \hat{\delta}_{FEM} - \hat{\delta}_{REM}$, $\hat{Q}$ and $\hat{V}$ as consistent estimates of the asymptotic covariance matrices of $\hat{\delta}_{FEM}$ and $\hat{\delta}_{REM}$ m-statistic has a χ^2-distribution with degrees of freedom equal to the number of parameter estimates. The test idea rests on the basic assumption that the REM is generally more efficient than the FEM. However, if the difference between the two estimators is large, the exogeneity assumptions in the REM are not met and the estimator is supposed to be misspecified. Thus, under the null hypothesis of the Hausman test both estimators are assumed to be consistent, but $\hat{\delta}_{REM}$ is more efficient than $\hat{\delta}_{FEM}$. Under the alternative hypothesis only $\hat{\delta}_{FEM}$ is consistent, $\hat{\delta}_{REM}$ is misspecified. In empirical application to panel data estimation, e.g. Baltagi et al. (2003) use the Hausman test to construct a pretest estimator of the following form: In the first case the pretest estimator reverts to REM if Hausman test for REM vs. FEM is not rejected. If the strong set of moment conditions of the REM is rejected, next the HT model validity is tested through a second Hausman test based on HT vs. FEM and if the HT specification is also rejected, the pretest estimator takes the FEM form.

Ahn & Low (1996) argue that the Hausman test according to eq. (13) has limited power because it only tests for the consistency and efficiency of β and not γ. Thus, the m-statistic can only detect misspecification if $\hat{\beta}_{REM}$ (as testable part of $\hat{\delta}_{REM}$) becomes substantially biased and different from $\hat{\beta}_{FEM}$, however as Ahn & Low (1996) argue moderate levels of correlation between Z_i and μ_i are unlikely to cause a significant bias in $\hat{\beta}_{REM}$. The authors therefore conclude that the Hausman test outcomes should be interpreted with caution and propose a more general test setup based on the Sargan (1958) / Hansen (1982) test for overidentification of moment conditions. In its general GMM form the J-Statistic is defined as:

$$J(\hat{\delta}_{EGMM}) = \hat{u}S'(S'\hat{\Omega}S)^{-1}S\hat{u}' \sim \chi^2(k-g) \tag{14}$$

where $\hat{u}$ are 1.step (2SLS) residuals, $(k-g)$ is the number of overidentifying restrictions. The J-Statistic is the value of the GMM objective function, evaluated at the efficient GMM estimator $\hat{\delta}_{EGMM}$. In an overidentified model it allows to test whether the model satisfies the full set of moment conditions. A rejection implies that IVs do not satisfy orthogonality conditions required for their employment. In its general form the J-Statistic allows for numerous testing setups depending on the choice of IV set and transformation operator. Thus, any test upon the J-Statistic can be regarded as a generalisation of the standard Hausman test. In fact, Ahn & Low (1996) prove that the latter tests the orthogonality conditions as $J(\hat{\delta}_{GLS})$ with $S = [QX, PX]$.

If the 'No Conditional Heteroscedasticity' NCH-condition holds, for any given IV estimator the J-Statistic coincides with the familiar Sargan (1958) statistic as:

$$J(.) = Sargan = \frac{\hat{u}'S(S'S)^{-1}S'\hat{u}}{\hat{\sigma}^2} = \frac{\hat{u}'P\hat{u}}{\hat{u}'\hat{u}/n} \sim \chi^2(k-g) \tag{15}$$

A nice fact about the Sargan (1958) Statistic is that it has a very intuitive interpretation: That is, since it has an nR_u^2 form, where R_u^2 is the uncentered R-squared and n is the total number of observations, it can be easily calculated by regressing the residuals of the IV regression on the full instrument set S. Since the model fit increases with a higher correlation of the residuals and the instrument set, this signals doubts for the validity of the model's underlying orthogonality assumptions.

Based on the J-Statistic we can also derive a test for a subset of valid moment conditions for instrument choice rather than the full IV set S. This so-called C-Statistic has been proposed by Eichenbaum et al. (1988) and tests the following hypothesis:

$$H_0 : E(S_{1i}u_i) = 0 \quad \text{and} \quad (S_{2i}u_i) = 0 \tag{16}$$

$$H_0 : E(S_{1i}u_i) = 0 \quad \text{and} \quad (S_{2i}u_i) \neq 0 \tag{17}$$

where S is divided into S_1 and S_2. The latter subset contains those instruments, for which exogeneity shall be tested. Under the null hypothesis both sub sets are orthogonal to the error term, under the alternative hypothesis only S_1 is exogenous. Numerically, the C-Statistic can be derived as the difference of two Hansen-Sargan overidentification tests with $C = J - J_1 \sim \chi^2(M - M_1)$, where M_1 is the number of instruments in S_1 and M

is the total number of IVs. We will make use of the above defined statistical criteria to guide moment condition selection in the HT case assuming imperfect knowledge about the correlation of r.h.s variables with the residuals for the Monte Carlo simulation experiment.

4 Monte Carlo Simulations

We specify a Monte Carlo simulation experiment in the spirit of Im et al. (1999) and Baltagi et al. (2003). We use a static one-way model as in eq.(1) including 4 time-varying (X) and 3 time-fixed (Z) regressors of the form:

$$\begin{aligned} y_{i,t} &= \beta_{11} x_{1_1,i,t} + \beta_{12} x_{1_2,i,t} \\ &+\beta_{21} x_{2_1,i,t} + \beta_{22} x_{2_2,i,t} \\ &+\gamma_{11} z_{1_1,i} + \gamma_{12} z_{1_2,i} \\ &+\gamma_{21} z_{2_1,i} + u_{i,t}, \\ \text{with: } u_{i,t} &= \mu_i + \nu_{i,t} \end{aligned} \tag{18}$$

where x_{1_1} and x_{1_2} are assumed to be uncorrelated with the error term, while x_{2_1} and x_{2_2} are correlated with μ_i. Analogously, z_{2_1} is correlated with the error term. The latter is composed of the unobserved individual effects (μ_i) and remainder disturbance $(\nu_{i,t})$. The time-varying regressors $x_{1_1}, x_{1_2}, x_{2_1}, x_{2_2}$ are generated by the following autoregressive process:

$$x_{n_m,i,t=1} = 0 \text{ with } n, m = 1, 2 \tag{19}$$

$$x_{1_1,i,t} = \rho_1 x1_{i,t-1} + \delta_i + \xi_{i,t} \text{ for } t = 2, \ldots, T \tag{20}$$

$$x_{1_2,i,t} = \rho_2 x2_{i,t-1} + \psi_i + \omega_{i,t} \text{ for } t = 2, \ldots, T \tag{21}$$

$$x_{2_1,i,t} = \rho_3 x3_{i,t-1} + \mu_i + \tau_{i,t} \text{ for } t = 2, \ldots, T \tag{22}$$

$$x_{2_2,i,t} = \rho_4 x4_{i,t-1} + \mu_i + \lambda_{i,t} \text{ for } t = 2, \ldots, T \tag{23}$$

For the time-fixed regressors $z_{1_1}, z_{1_2}, z_{2_1}$ we analogously define:

$$z_{1_1,i} = 1 \tag{24}$$

$$z_{1_2,i} = g_1\psi_i + g_2\delta_i + \kappa_i \tag{25}$$

$$z_{2_1,i} = \mu_i + \delta_i + \psi_i + \epsilon_i \tag{26}$$

The variable $z_{1_1,i}$ simplifies to a constant term, $z_{2_1,i}$ is the endogenous time-fixed regressor since it contains μ_i as r.h.s. variable, the weights g_1 and g_1 in the specification of

$z_{1_2,i}$ control for the degree of correlation with the time-varying variables $x_{1_1,i,t}$ and $x_{1_2,i,t}$.[9] The remainder innovations in the data generating process are defined as follows:

$$
\begin{aligned}
\nu_{i,t} &\sim N(0,\sigma_\nu^2) &(27)\\
\mu_i &\sim N(0,\sigma_\mu^2) &(28)\\
\delta_i &\sim U(-2,2) &(29)\\
\xi_{i,t} &\sim U(-2,2) &(30)\\
\psi_i &\sim U(-2,2) &(31)\\
\omega_{i,t} &\sim U(-2,2) &(32)\\
\tau_{i,t} &\sim U(-2,2) &(33)\\
\lambda_{i,t} &\sim U(-2,2) &(34)\\
\epsilon_i &\sim U(-2,2) &(35)\\
\kappa_i &\sim U(-2,2) &(36)
\end{aligned}
$$

Except μ_i and $\nu_{i,t}$, which are drawn from a normal distribution with zero mean and variance σ_μ^2 and σ_ν^2 respectively, all innovations are uniform on [-2,2]. For $\mu_i, \delta_i, \psi_i, \epsilon_i, \kappa_i$ the first observation is fixed over T. With respect to the main parameter settings in the Monte Carlo simulation experiment we set:

- $\beta_{1_1} = \beta_{1_2} = \beta_{2_1} = \beta_{2_2} = 1$
- $\gamma_{1_2} = \gamma_{2_1} = 1$
- $\rho_1 = \rho_2 = \rho_3 = \rho_4 = 0.7$

All variable coefficients are normalized to one, the specification of $\rho < 1$ assures that the time-varying variables are stationary. We also normalize σ_ν equal to one and define a load factor ξ determining the ratio of the variance terms of th e error components as $\xi = \sigma_\mu/\sigma_\nu$. ξ takes values of 2;1 and 0.5. We run simulations with different combinations in the time and cross-section dimension of the panel as $N = (100, 500, 1000)$ and $T = (5, 10)$. All Monte Carlo simulations are conducted with 500 replications for each permutation in y and u. As in Arellano & Bond (1991) we set $T = T + 10$ and cut off first 10 cross-sections, which gives a total sample size of NT observations.

[9]We vary g_1 and g_2 on the interval [-2,2]. The default is $g_1 = g_2 = 2$.

We apply the FEVD and Hausman-Taylor estimators.[10] As outlined above, one drawback in earlier Monte Carlo based comparisons between the HT model and rival non-IV candidates was the strong assumption made for IV selection in the HT case, namely that true correlation between r.h.s. variables and the error term is known. However, this may not reflect the identification and estimation problem in applied econometric work and Alfaro (2006) identifies it as one of the open questions for future investigation in Monte Carlo simulations. We therefore account for the HT variable classification problem by implementing algorithms from 'model selection criteria'-literature, which combine information from Hansen-Sargan overidentification test for moment condition selection as outlined above and time-series information-criteria. Following Andrews (1999) we define a general model selection criteria (MSC) based on IV estimation as

$$MSC_n(m) = J(m) - h(c)k_n \tag{37}$$

where n is the sample size, c as number of moment conditions selected by model m based on the Hansen-Sargan J-Statistic $J(m)$, $h(.)$ is a general function, k_n is a constant term. As eq.(37) shows, the model selection criteria centers around the J-Statistic outlined in section 3. The second part in eq.(37) defines a 'bonus' term rewarding models with more moment conditions, where the form of function $h(.)$ and the constants ($k_n :\geq 1$) are specified by the researcher. For empirical application Andrews (1999) proposes three operationalizations in analogy to model selection criteria from time series analysis:

- MSC-BIC: $J(m) - (k - g)ln\ n$
- MSC-AIC: $J(m) - 2(k - g)$
- MSC-HQIC: $J(m) - Q(k - g)ln\ ln\ n$ with $Q = 2.01$

where $(k - g)$ is the number of overidentifying restrictions, and depending on the form of the 'bonus' term, the MSC may take the BIC (Bayesian), AIC (Akaike) and HQIC (Hannan Quinn) form. [11] We apply all three information criteria in the Monte Carlo simulations motivated by the results in Andrews & Lu (2001) and Hong et al. (2003) that the superiority of one of the criteria over the others in terms of finding consistent moment conditions may vary with the sample size.[12] For each of these MSC criteria we specify the following algorithms:

[10] For the FEVD estimator we employ the Stata routine *xtfevd* written by Plümper & Tröger (2007), the HT model is implemented using the user written Stata routine *ivreg2* by Baum et al. (2003).

[11] The BIC criterion was introduced by Schwartz (1978), the AIC by Akaike (1977) and the HQIC by Hannan & Quinn (1979).

[12] Generally, the MSC-BIC criterion is found to have the best empirical performance in large samples, while the MSC-AIC outranks the other criteria in small sample settings, but performs poor otherwise.

1. Unrestricted form: For all possible IV combinations out of the full IV-set *S=(QX1, QX2, PX1, PX2, Z1, Z2)*, which satisfy the order condition $k_1 > g_2$ (giving a total number of 42 combinations), we calculate the value of the MSC criterion (for the BIC, AIC and HQIC separately) and choose that model as final HT specification, which has minimum MSC value over all candidates.

2. Restricted form: This algorithm follows the basic logic from above, but additionally puts the further restriction that only those models serves as MSC candidates for which the p-value of the J-Statistic is a above a critical value $C_{crit.}$, which we set to $C_{crit.} = 0.05$ to be sure that the selected moment conditions are true in terms of statistical pre-testing. The restricted version thus follows the advice of Andrews (1999) to ensure that the parameter space incorporates only information, which assumes that certain moment conditions are correct.

We present flow charts of the restricted and unrestricted MSC based search algorithm in figure 1.

<<< insert Figure 1 about here >>>

As Andrews (1999) argues, the above specified model selection criteria is closely related to the C-Statistic approach by Eichenbaum et al. (1988) to test whether a given subset of moment conditions is correct or not. Thus alternatively to the above described algorithms, we specify a downward testing approach based on the C-Statistic: Here we start from the HT model with with full IV set in terms of the REM moment conditions as $S_1 = (QX1, QX2, PX1, PX2, Z1, Z2)$. We calculate the value of the J-Statistic for the model with IV-set S_1 and compare its p-value with a predefined critical value $C_{crit.}$, which we set in line with the above algorithm as $C_{crit.} = 0.05$. If $P_{S_1} > C_{crit.}$ we take this model as a valid representation in terms of the underlying moment conditions. If not, we calculate the value of the C-Statistic for each single instrument in S_1 and exclude that instrument from the IV-set that has the maximum value of the C-statistic.

We then re-estimate the model based on the IV-subset S_2 net of the selected instrument with the highest C-Statistic and again calculate the J-Statistic and its respective p-value. If $P_{S_2} > C_{crit.}$ is true, we take the HT-model with S_2 as final specification and otherwise again calculate the C-Statistic for each instrument to exclude that one with the highest value. We run this downward testing algorithm for moment conditions until we find a model that satisfies $P_{S_.} > C_{crit.}$ or at the most until we reach the IV-sets S_n to S_m, where the number of overidentifying restrictions $(k - g) = 1$, since the J-Statistic is not defined for just identified models. Out of S_n to S_m we then pick the model with the lowest J-Statistic value. The C-Statistic based model selection algorithms is graphically summarized in figure 2.

<<< insert Figure 2 about here >>>

Turning to the results of the Monte Carlo simulations, detailed information about the estimators' performance for the above specified parameter settings in N, T and ξ are reported in table A.1 to table A.6.[13] Since we are interested in consistency and efficiency of the respective estimators, we compute the empirical bias, its standard deviation and the root mean square error (rmse). The bias is defined as

$$bias(\hat{\delta}) = \sum_{m=1}^{M} (\hat{\delta} - \delta_{true})/M, \tag{38}$$

where $m = (1, 2, \ldots, M)$ is the number of simulation runs. Next to the standard deviation of the estimated bias we also calculate the root mean square error, which puts a special weight on outliers, as:

$$rmse(\hat{\delta}) = \sqrt{\left(\sum_{m=1}^{M} (\hat{\delta} - \delta_{true})/M \right)^2}. \tag{39}$$

We first take a closer look at the individual parameter estimates for the parameter settings $N = 1000, T = 5$ and $\xi = 1$, which are typically assumed in the standard Panel data literature building on the large N, small T data assumption. In figure 3 we plot Kernel density distributions for all regression coefficients for the following three estimators: i.) the FEVD, ii.) the HT model with perfect knowledge about the underlying variable correlation with the error term and iii.) the HT model based on the MSC-BIC algorithm (in its restricted form). The latter shows on average the best performance among all HT estimators with imperfect knowledge about the underlying data correlation - closely followed by the C-Statistic based model selection algorithm.

For the coefficients of the two exogenous time-varying variables β_{1_1} and β_{1_2} all three estimators give almost unbiased results centering around the true parameter value of one. The standard deviation and rmse are the smallest for the HT model with perfect knowledge about the underlying data correlation, followed by the algorithm based HT estimators. The FEVD has a slightly higher standard deviation and rmse. For the estimated coefficients of the endogenous time-varying variables β_{2_1} and β_{2_2} the HT and FEVD give virtually

[13]In the following we do not present simulation results for the constant term z_{1_1} in the model. Results can be obtained from the author upon request.

identical results, while the HT based MSC-BIC in figure 1 is slightly biased for β_{2_1} but comes closer to the true parameter value for the parameter β_{2_1}.

To sum up, though there are some minor differences among the three reported estimators for the time-varying variables in figure 1, the overall empirical discrepancy is rather marginal. This picture however radically changes for the Monte Carlo simulation results of the time-fixed variable coefficients γ_{1_2} and γ_{2_1}: Here only the HT model with the ex-ante correctly specified variable correlation gives unbiased results for both the exogenous (γ_{1_2}) and endogenous variable (γ_{2_1}). Both the FEVD and HT model based on the MSC-BIC have difficulties in calculating these variable coefficients correctly, while the bias of the FEVD is somewhat lower than for the MSC-BIC Hausman-Taylor model in both cases. Especially for γ_{2_1} exclusively all HT based model selection algorithms have a large bias/standard deviation as well as a high rsme relative to the HT with perfect knowledge abot the variable correlation with the error term. The FEVD has a significant bias (approximately 50 percent higher than the standard HT) but compared to the MSC-BIC based specification a lower bias/standard deviation.

<<< insert Figure 3 about here >>>

Turning to the small sample properties for the above mentioned estimators we additionally plot Kernel density plots for the parameter settings $N = 100, T = 5, \xi = 2$. Here the results in figure 4 show that the MSC-BIC based HT model is already more biased compared to the standard HT and FEVD for the parameter estimates of the time-varying variables β_{1_1}, β_{1_2}, β_{2_1} and β_{2_2}, where in all cases the bias is the smallest for the FEVD. With respect to the rmse the smallest value for β_{1_1} and β_{1_2} is given by the C-Statistic based HT model, while FEVD and the standard HT model perform best for β_{2_1} and β_{2_2}. For the time-fixed variables again the FEVD and the MSC-BIC based HT model have a significant bias, while the HT model with perfect knowledge about the underlying variable correlation comes on average much closer to the true parameter value (in particular for γ_{1_2}). However, as already observed in Plümper & Tröger (2007) the standard deviation of the latter estimator is much higher compared to the other two estimators. This leads to the result that in terms of the rmse the FEVD performs better than the standard HT in these settings (for both γ_{1_2} and γ_{2_1}), although it shows a larger bias compared to the latter. The results in figure 4 indicate that the HT instrumental variable approach is rather inefficient in small sample settings, though the average bias is small.

<<< insert Figure 4 about here >>>

A specific problem of the MSC-BIC based HT model in small sample settings is shown in figure 5. The Kernel density plot for the coefficient γ_{2_1} of the endogenous time-fixed variables reveals a 'duality' problem for the search algorithm based estimator, which significantly increases with small values for ξ. Different from the standard HT and FEVD estimators the MSC-BIC based HT model shows a clear double peak for parameter estimates of $\hat{\gamma}_{2_1}$, with one peakaround the true coefficient value of one and a second significantly biased one. This kind of duality problem with a possibly poor MSC based estimator performance has already been addressed in Andrews (1999) for those cases where there are typically two or more selection vectors that yield MSC values close to the minimum and parameter estimates that differ noticeably from each other. As the histogram in figure 6 shows, this is indeed the problem for the MSC-BIC based HT model: Based on the Monte Carlo simulation runs with 500 reps. the algorithm tends to pick two dominant IV-sets from which one has the (inconsistent) REM form with a full instrument list, while only the second one consistently excludes $Z21$ from the instrument list. This results may be seen as a first indication that in small samples and a small proportion of the total variance of the error term due to the random individual effects (through low values of ξ), J-statistic based IV selection may have a low power and yield inconsistent results.

<<< insert Figure 5 and Figure 6 about here >>>

Turning from a comparison of single variable coefficients to an analysis of overall measures of model bias and efficiency for an aggregated parameter space, we compute NOMAD and NORMSQD values, where the NOMAD (normalized mean absolute deviation) computes the absolute deviation of each parameter estimate from the true parameter, normalizing it by the true parameter and averaging it over all parameters and replications considered. The NORMSQD computes the mean square error (mse) for each parameter, normalizing it by the square of the true parameter, averaging it over all parameters and taking its square root (for details see Baltagi & Chang, 2000). Both overall measures are thus extensions to the single parameter bias and rmse statistics defined above. We compared the FEVD model with the standard HT model and the algorithm based HT models using the C-Statistic approach, as well as the MSC-BIC1, MSC-HGIC1, MSC-AIC1 (where the index 1 denotes that all are based on the restricted specification).

We start to report surface plots for the two parameter aggregates 1.) time-varying and 2.) time-fixed variable coefficients over all different settings in our Monte Carlo simulation experiment. The choice of aggregation is motivated by the above findings that the results significantly differ with respect to time-varying and time-fixed variable coefficients. Figure 7 and figure 8 plot NOMAD and NORMSQD values for the time-varying coefficients β_{1_1}

to β_{2_2}. The figures show that in terms of bias the algorithm based HT models show a significant small sample problem, while the bias of the FEVD and standard HT is rather small for different combinations in the time and cross-section dimension of the data. In terms of the rmse all estimators (both IV and non-IV) show a significant decrease in the rmse value with increasing number of total observations NT.

With respect to the rmse the standard HT model performs best, closely followed by the FEVD. As Plümper and Tröger (2007) already expected without explicit testing, the efficiency of the HT approach significantly reduces if the underlying assumption about the variable correlation with the error term is not known. Only the MSC-BIC comes close to the FEVD as second best estimator. Another interesting finding is that figure 7 and 8 both show distinct spikes alongside the overall trend of increasing estimator power with larger sample size. These spikes (in particular for the search algorithm based HT models) are induced by low values for $\xi = 0.5$, which significantly deteriorate the estimators performance in terms of bias and rmse. This result comes close to findings in Baltagi et al. (2003), where for small values of ξ their pretest estimator reverts to the REM specification although the underlying data structure implies a correlation of r.h.s. variables with the individual effects in the sense of a HT world. Baltagi et al. (2003) report misleading inference of the pretest estimator in this case.

Turning to the NORMAD and NORMSQD values for the time-fixed regressors, figure 9 and 10 show that average bias and rmse are roughly constant over different combinations in the time and cross-section dimension of the data, where the smallest bias is obtained from the standard HT model with perfect knowledge about the underlying variable correlation. On average the bias of the FEVD is significantly higher than for the standard HT. Also the performance of the algorithm based HT models is rather poor. Contrary to the estimation bias, in terms of rmse the FEVD clearly performs best. The high rmse for HT models with imperfect knowledge about the underlying variable correlation especially in small sample settings, can be explained by the above identified duality problem of this statistical approach. The standard HT model also has small sample problems, but comes close to the rmse value of the FEVD estimator for a larger sample size.

<<< insert Figure 7 to Figure 10 about here >>>

To sum up, we finally average the NOMAD and NORMSQD values over all variable coefficients (see table 1) and plot the results in figure 11. The figure shows that the HT model with perfect knowledge about the underlying variable correlation has the lowest NOMAD value, with the FEVD having two times and algorithm based HT specification even three times higher values for the average bias over all model coefficients. For the latter the C-Statistic based model selection criteria performs slightly better than the MSC based estimators. Contrary, with respect to the NORMSQD by far the best model is the non-IV FEVD. The difference between the standard HT model and the algorithm based specification is rather low. This broad picture indicates that the HT instrumental variable model is a consistent estimator given perfect knowledge about the true underlying correlation between the r.h.s. variables and the error term. However, when one has to rely on statistical criteria to guide moment condition selection the empirical performance for the specific setup in the Monte Carlo simulation design is considerably lower. This in turn speaks in favor of using non-IV two step estimators such as the FEVD, which although it may introduce a moderate bias in estimating the time-fixed variables, has the lowest rmse due to its robust OLS estimation approach compared to the IV based rival HT estimators.

<<< insert Table 1 and Figure 11 about here >>>

5 Empirical application: Trade estimates for German regional data

Given the above findings from our Monte-Carlo simulation experiment in this section we aim to consider the empirical performance of the FEVD and HT model in an empirical application estimating gravity type models. We take up the research question in Alecke et al. (2003 & 2007) and specify trade equations for regional entities. In particular, we aim to re-estimate gravity models for export flows among German states (NUTS1-level) and its EU27 trading partners using an updated database for the period 1993-2005. We are particularly interested in quantifying the effects of time-fixed variables including geographical distance as a general proxy for trading costs as well as a set time-fixed 0/1-dummies for border regions, the East German states as well as the specific trade pattern with the CEEC countries.[14] Earlier evidence in Alecke et al. (2007) has shown that there is a considerable degree of heterogeneity for these time-fixed variables among different estimators (similar evidence is given in Belke & Spies, 2008). The empirical export model we focus on has the following form:[15]

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

[15]Results for an import equation with qualitatively similar results can be obtained from the author upon request.

$$
\begin{aligned}
log(EX_{ijt}) &= \alpha_0 + \alpha_1 log(GDP_{it}) + \alpha_2 log(POP_{it}) + \alpha_3 log(GPD_{jt}) + \alpha_4 log(POP_{jt}) \\
&\quad + \alpha_5 log(PROD_{it}) + \alpha_6 log(DIST_{ij}) + \alpha_7 SIM + \alpha_8 RLF \qquad (40) \\
&\quad + \alpha_9 EMU + \alpha_{10} EAST + \alpha_{11} BORDER + \alpha_{12} CEEC,
\end{aligned}
$$

where the index indicates German regional exports from region i to country j for time period t and imports to German state i from country j respectively. The variables in the model are defined as follows[16]:

- EX = Export flows from region i to country j
- GDP = Gross domestic product in i and j respectively
- POP = Population in i and j
- $PROD$ = Labour productivity in i and j
- $DIST$ = Geographical distance between state/national capitals
- SIM = Similarity index defined as: $log(1 - \left(\frac{GDP_{i,t}}{GDP_{i,t}+GDP_{j,t}}\right)^2 - \left(\frac{GDP_{j,t}}{GDP_{i,t}+GDP_{j,t}}\right)^2)$
- RLF = Relative factor endowments in i and j defined as: $log \left| \left(\frac{GDP_{i,t}}{POP_{i,t}}\right) - \left(\frac{GDP_{j,t}}{POP_{j,t}}\right) \right|$
- EMU = EMU membership dummy for i and j
- $EAST$ = East German state dummy for i
- $BORDER$ = Border region dummy between i and j
- $CEEC$ = CEE country dummy for j

The estimation results are shown in table 2. We particularly focus in the FEVD and HT estimates for the variable $log(DIST_{ij})$ as well as the time-fixed dummies $EAST$, $BORDER$ and $CEEC$. The HT approach rests on the C-Statistic based downward testing approach to find a consistent set of moment conditions.

Turning to the results, in line with our Monte Carlo simulation results both the FEVD and HT estimators are very close in quantifying the time-varying variables in the gravity model for German regional export activity. As expected from its theoretical foundation (see e.g. Egger, 2000, Feenstra, 2004) both home and foreign country GDP have a positive and significant influence on German export activity, indicating that trade increases with absolute higher income levels. Moreover, also home region productivity (defined as GDP per total employment) is found to be statistically significant and highly positive,

[16] Further details can be found in the data appendix in table A.7.

which in turn can be interpreted in line with recent findings based on firm-level data (see e.g. Helpman et al., 2003, or Arnold & Hussinger, 2006, for the German case) that the degree of internationalization of home firms (both trade and FDI) increases with higher productivity levels. The interpretation of the population variable in the gravity model is less clear cut: Both the FEVD and HT estimator find a positive coefficient sign for foreign population, which can be interpreted in favour of the market potential approach indicating that German export flows are higher for population intense economies. Also for the GDP based interaction variable SIM (definition see above) the two estimators show similar results.

However, as already observed throughout the Monte Carlo simulation experiment for the time-fixed variables the estimators show a considerable degree of heterogeneity. In our export model the C-Statistic based HT approach finds a coefficient for the distance variable (-1,73) that is almost twice as large as the respective coefficient in the in POLS, REM and FEVD (-0.97) case. A similar difference between FEVD and HT model results were also found in Belke & Spies (2008) for EU wide data (the authors report coefficients for the distance variable in the HT case as -1,83 compared to -1,39 in the FEVD case). Without the additional knowledge from the above Monte Carlo simulation experiment, we could hardly answer the question whether this discrepancy among estimators either indicates an upward bias of the HT model given the fact that (for national data) the parameter estimate for the distance variable typically ranges between -0.9 to -1.3 (see e.g. Disdier & Head, 2008 as well as Linders, 2005) or whether the use of smaller regional entities serves as a better proxy for geographical distance thus gives a more accurate estimate for trade costs (which may be possibly higher).

However, in the light of the Monte Carlo simulation results together the typical range of national estimates it seems plausible to rely on the FEVD estimation results, although the HT model passes the Hansen/Sargan overidentification test (treating geographical distance as endogenous). Also for the further time-fixed dummy variables in the model the FEVD estimates show more reliable coefficient signs than the HT model: That is, we would expect the border dummy to be positive as e.g. found in Lafourcade & Paluzie (2005) for European border regions. Also, German-CEEC trade was persistently found to be above its 'potential' in a couple of earlier studies such as Schumacher & Trübswetter, 2000, Buch & Piazolo, 2000, Jakab et al., 2001), Caetano et al., 2002 as well as Caetano & Galego, 2003). In both cases the FEVD estimates are thus more in line with recent empirical findings than the HT IV-estimation. For the FEVD results we may thus finally conclude that the obtained coefficients for time-fixed variables better reflect similar evidence in recent empirical work.

<<< insert Table 2 about here >>>

6 Conclusion

In this paper we have performed a Monte Carlo simulation experiment to compare the empirical performance of IV and non-IV estimators for a regression setup, which includes time-fixed variables as right hand side regressors and where endogeneity matters. We define the latter as any correlation between the r.h.s. variables with the model's error term. In specifying empirical estimators we focus on the Hausman-Taylor (1981) Instrumental Variable (IV) model both with perfect and imperfect knowledge about the underlying variable correlation with the model's residuals and non-IV two-step estimators such as the Fixed Effects Vector Decomposition (FEVD) model recently proposed by Plümper & Tröger (2007). Our results confirm earlier empirical findings (see e.g. Alfaro, 2006) that the HT (with perfect knowledge) works better for time-varying, while the FEVD for time-fixed variables. Averaging over all parameter we find the the HT model (with perfect knowledge) generally has the smallest bias, while the FEVD show to have a by far lower root mean square error (rmse) as a general efficiency measure. Especially in small sample settings our Monte Carlo simulations show that the IV-based HT model has a large standard deviation and consequently high rmse values.

Additionally, relaxing the assumption of perfect knowledge for the HT model, the empirical performance of the latter significantly worsens. We compute different algorithms to select consistent IV-sets centering around the Hansen/Sargan overidentification test (J-Statistic), however all estimates based on these algorithms generally show a weaker empirical performance than the non-IV alternative (FEVD). One major drawback of the HT models with imperfect knowledge is a 'duality' problem in small sample settings, where the estimator has difficulties to discriminate between consistent and inconsistent moment condition vectors. We may thus conclude that IV selection solely based on statistical criteria has to be treated with some caution.

Future research in this area should especially focus on possible improvements of these J-Statistic based algorithms by combining a measure for model consistency with criteria for IV-relevance and -uniqueness. Hall & Peixe (2000) as well as Jana (2005) e.g. propose to combine the J-Statistic with the Canonical correlation information criteria. For practical application Andrews & Lu (2001) additionally recommend to combine statistically driven model selection with any other information available about the correct parameter and moment vector. This would speak in favor of using statistical criteria such as the J-Statistic in terms of a fine tuning supporting the basic IV selection on grounds of 'economic intuition' as proposed by Hausman & Taylor (1981). However, the empirical performance of such a multi-level selection strategy is hardly to test in Monte Carlo simulations and has to prove its validity in practical application. An alternative choice for applied researcher are non-IV two-step estimators such as the FEVD by Plümper & Tröger (2007), which show an on average good performance in our Monte Carlo simulations and yield very plausible results for the empirical estimation of German regional trade flows using gravity type models.

References

[1] **Ahn, S.; Low, S. (1996):** ”A reformulation of the Hausman test for regression models with pooled cross-section-time series data”, in: *Journal of Econometrics*, Vol. 71, pp. 309-319.

[2] **Ahn, S.; Schmidt, P. (1999):** ”Estimation of Linear Panel Data Models Using GMM”, in: Matyas, L. (Ed.): ”Generalized methods of moments estimation”, Cambridge.

[3] **Akaike, H. (1969):** ”Fitting Autoregressive Models for Predictions”, in: *Annals of the Institute of Statistical Mathematics*, Vol. 21, pp. 243-247.

[4] **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” in: *DIW Vierteljahrshefte zur Wirtschaftsforschung*, Vol. 72, No. 4, pp. 565-578.

[5] **Alecke,B.; Mitze, M.; Untiedt, G. (2007):** ”The Trade-FDI-Nexus of German firms: Evidence from panel gravity models using regional state level data”, paper presented at the SMYE 2007, University of Hamburg.

[6] **Alfaro, R. (2006):** ”Application of the symmetrically normalized IV estimator”, Working Paper, Boston University, http://people.bu.edu/ralfaro/doc/sniv0602.pdf

[7] **Andrews, D. (1999):** ”Moment Selection Procedures for Generalized Method of Moments Estimation”, in: *Econometrica*, Vol. 67, No. 3, pp. 543-564.

[8] **Andrews, D.; Lu, B. (2001):** ”Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models”, in: *Journal of Econometrics*, Vol. 101, pp. 123-164.

[9] **Arellano, M.; Bond, S. (1991):** ”Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations”, in: *Review of Economic Studies*, Vol. 58, pp. 277-297.

[10] **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.

[11] **Atkinson, S.; Cornwell, C. (2006):** ”Inference in Two-Step Panel Data Models with Time-Invariant Regressors”, Working paper, University of Georgia.

[12] **Baldwin, R.; Taglioni, D. (2006):** ”Gravity for dummies and dummies for gravity equations”, NBER Working paper No. 12516.

[13] **Baltagi, B. (2008):** "Econometric Analysis of Panel Data", 4. Edition, Chichester: John Wiley & Sons.

[14] **Baltagi, B.; Bresson, G.; Pirotte, A. (2003):** "Fixed effects, random effects or Hausman Taylor? A pretest estimator", in: *Economic Letters*, Vol. 79, pp. 361-369.

[15] **Baltagi, B.; Chang, Y. (2000):** "Simultaneous Equations with incomplete Panels", in: *Econometric Theory*, Vol. 16, pp. 269-279.

[16] **Baum, C.; Schaffer, M.; Stillman, S. (2003):** "IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation", Boston College, Statistical Software Component No. S425401.

[17] **Belke, A.; Spies, J. (2008):** "Enlarging the EMU to the East: What Effects on Trade?", in: *Empirica*, 4/2008 , pp. 369-389.

[18] **Boumahdi, R.; Thomas, A. (2008):** "Endogenous Regressors and Correlated Effects", in: Matyas, L.; Sevestre, P. (Eds.): "The Econometrics of Panel Data", 3.edition, Berlin et al.

[19] **Breusch T.; Mizon, G.; Schmidt, P. (1989):** "Efficient Estimation using Panel Data", in: *Econometrica*, Vol. 57, pp. 695-700.

[20] **Breuss, F.: Egger, P. (1999):** "How Reliable Are Estimations of East-West Trade Potentials Based on Cross-Section Gravity Analyses?", in: *Empirica*, Vol. 26, No. 2, pp. 81-94.

[21] **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.

[22] **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.

[23] **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.

[24] **Disdier, A.; Head, K. (2008):** "The Puzzling Persistence of the Distance Effect on Bilateral Trade", in: *The Review of Economics and Statistics*, Vol. 90(1), pp. 37-48.

[25] **Egger, P. (2000):** "A note on the proper econometric specification of the gravity equation", in: *Economics Letters*, Vol. 66, pp. 25-31.

[26] **Eichenbaum, M.; Hansen, L.; Singelton, K. (1988):** "A Time Series Analysis of Representative Agent Models of Consumption and Leisure under Uncertainty", in: *Quarterly Journal of Economics*, Vol. 103, pp. 51.78,

[27] **Etzo, I.(2007):** "Determinants of interregional migration in Italy: A panel data analysis", MPRA Paper No. 5307.

[28] **Feenstra, R. (2004):** "Advanced International Trade. Theory and Evidence", Princeton.

[29] **Fidrmuc, J. (2008):** "Gravity models in integrated panels", in: Empirical Economics, forthcoming.

[30] **Hall, A.; Peixe, F. (2000):** "A Consistent Method for the Selection of Relevant Instruments", Econometric Society World Congress 2000 Contributed Papers No. 790.

[31] **Hannan, E.; Quinn, B. (1979):** "The Determination of the Order of an Autoregression", in: *Journal of the Royal Statistical Society*, Series B, Vol. 41, pp. 190-195.

[32] **Hansen, L. (1982):** "Large Sample Properties of Generalised Method of Moments Estimators", in: *Econometrica*, Vol. 50, pp. 1029-1054.

[33] **Hausman, J. (1978):** "Specification tests in econometrics", in: *Econometrica*, Vol 46, pp. 1251-1271.

[34] **Hausman, J.; Taylor, W. (1981):** "Panel data and unobservable individual effects", in: *Econometrica*, Vol. 49, pp. 1377-1399.

[35] **Helpman, E.; Melitz, M.; Yeaple, S. (2003)**: "Export versus FDI", NBER Working Paper No. 9439.

[36] **Henderson, D.; Millimet, D. (2008):** "Is Gravity Linear?", in: *Journal of Applied Econometrics*, Vol. 23, pp. 137-172.

[37] **Hong, H.; Preston, B.; Shum, M. (2003):** "Generalized Empirical Likelihood-Based Model Selection Criteria for Moment Condition Models", in: *Econometric Theory*, Vol. 19, pp. 923-943.

[38] **Hsiao, C. (2003):** "Analysis of Panel Data", 2.edition, Cambridge.

[39] **Im, K.; Ahn, S.; Schmidt, P.; Wooldridge, J. (1999):** "Efficient estimation in panel data models with strictly exogenous explanatory variables", in: *Journal of Econometrics*, Vol. 93, pp. 177-203.

[40] **Jakab, Z.; Kovcs, 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", in: *Journal of Comparative Economics*", Vol. 29, pp. 276-292.

[41] **Jana, K. (2005):** "Canonical Correlations and Instrument Selection in Econometrics", PhD-Thesis, North Carolina State University.

[42] **Krogstrup, S.; Wälti, S. (2008):** "Do fiscal rules cause budgetary outcomes?", in: *Public Choice*, Vol. 136(1), pp. 123-138.

[43] **Lafourcade, M.; Palusie, E. (2005):** "European Integration, FDI and the Internal Geography of Trade: Evidence from Western European Border Regions", paper presented at the EEA congress in Amsterdam (August 2005).

[44] **Linders, G. (2005):** "Distance decay in international trade patterns: a meta-analysis", ERSA 2005 conference paper ersa05p679.

[45] **Matyas, L. (1997):** "Proper Econometric Specification of the Gravity Model", in: *The world economy*, Vol. 20(3), pp. 363-368.

[46] **Mundlak, Y. (1978):** "On the pooling of time series and cross-section data", in: *Econometrica*, Vol. 46, pp. 69-85.

[47] **Murphy. K.; Topel, R. (1985):** "Estimation and Inference in Two-Step Econometric Models", in: *Journal of Business and Economic Statistics*, Vol. 3, pp. 88-97.

[48] **Pagan, A. (1984):** "Econometric Issues in the analysis of regressions with generated regressors", in: *International Economic Review*, Vol. 25.

[49] **Plümper, T.; Tröger, V. (2007):** "Efficient Estimation of Time Invariant and Rarely Changing Variables in Panel Data Analysis with Unit Effects", in: *Political Analysis*, Vol. 15, pp. 124-139.

[50] **Sargan, J. (1958):** "The Estimation of Economic Relationships Using Instrumental Variables", in: *Econometrica*, Vol. 26, pp. 393-415.

[51] **Schumacher, D.; Trübswetter, P. (2000):** "Volume and Comparative Advantage in East West trade", DIW Discussion Paper No. 223.

[52] **Schwartz, G. (1978):** "Estimating the Dimension of a Model", in: *Annals of Statistics*, Vol. 6, pp. 461-464.

[53] **Zwinkels, R.; Beugelsdijk, S. (2008):** "Gravity equations: workhorse or Trojan horse in explainng trade and FDI patterns across time and space?", Working paper, Erasmus School of Economics, Rotterdam.

Figure 1: MSC based model selection algorithm for HT-approach

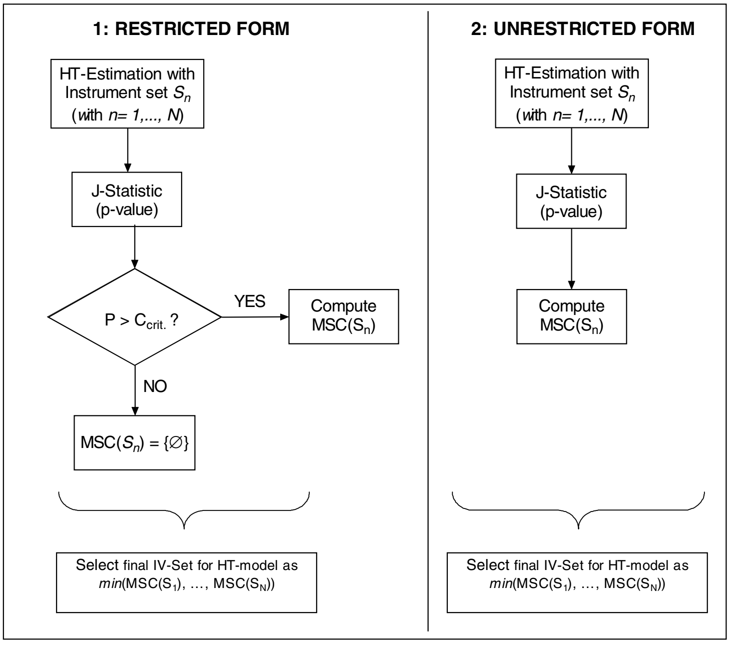

Figure 2: C-Statistic based model selection algorithm for HT-approach

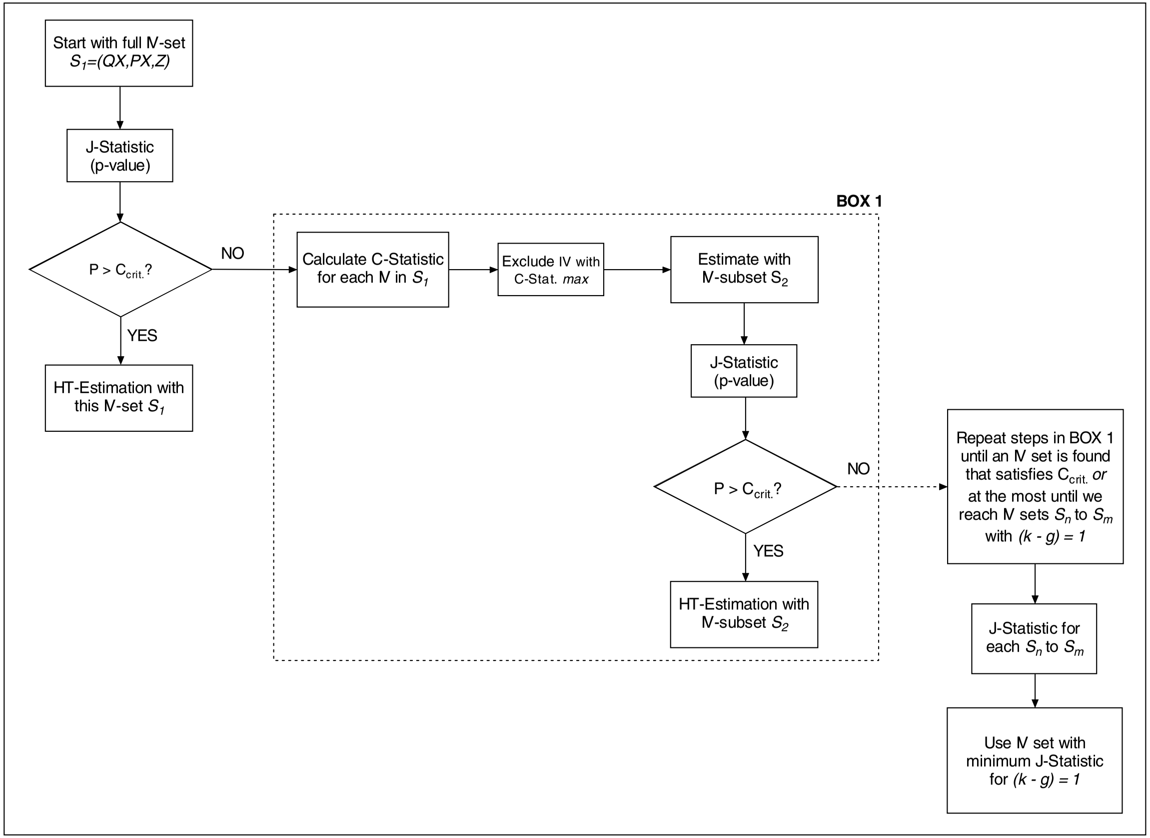

Figure 3: Kernel density plots for Monte Carlo simulation results with $N = 1000, T = 5, \xi = 1$

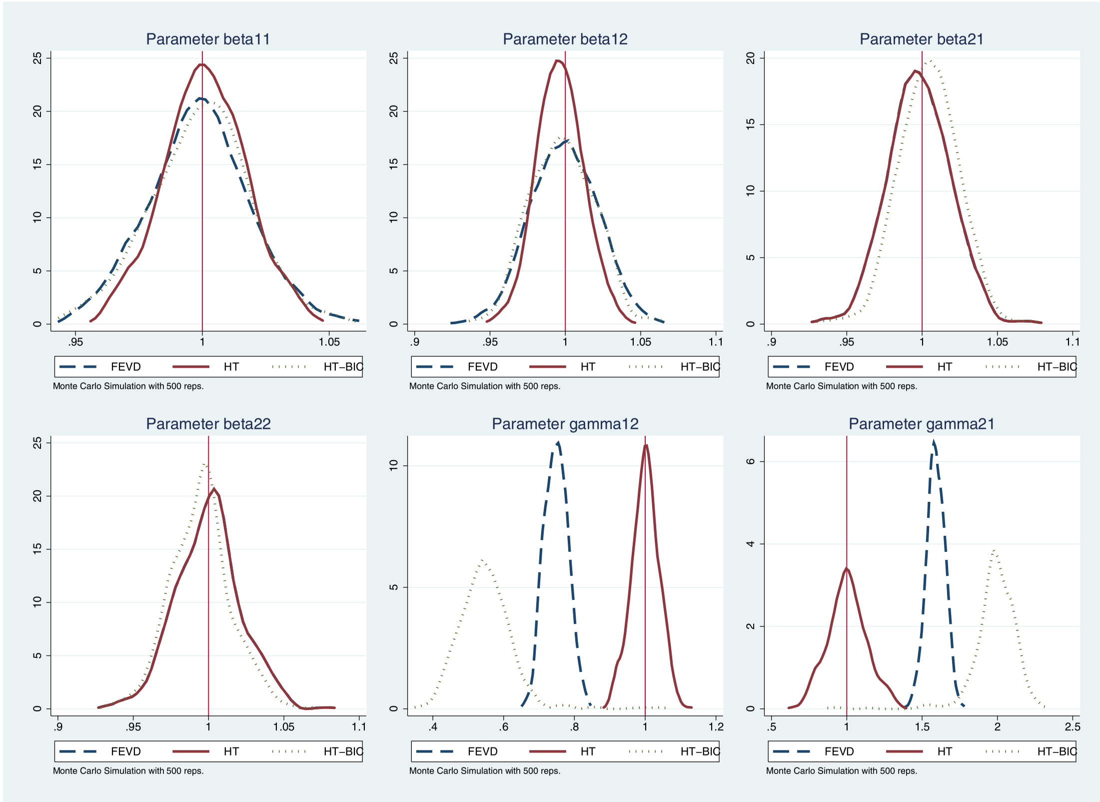

Figure 4: Kernel density plots for Monte Carlo simulation results with $N = 100, T = 5, \xi = 2$

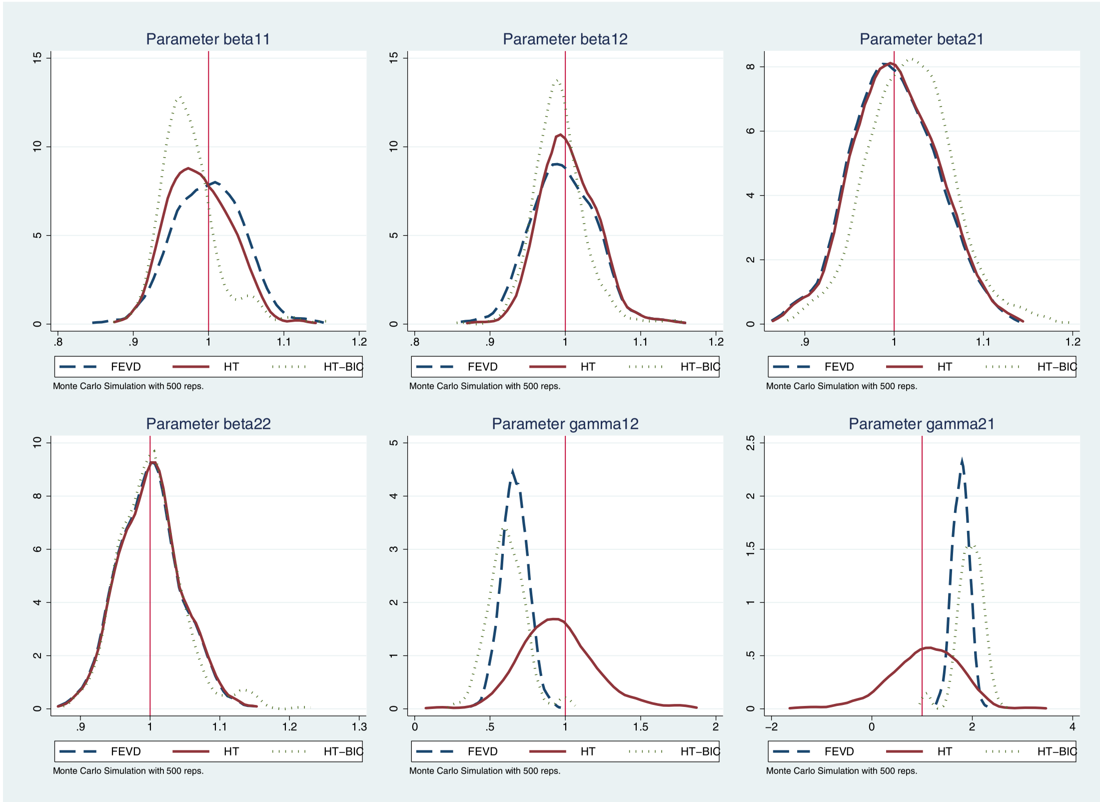

Figure 5: Kernel density plots for Monte Carlo simulation results of γ_{21} with $N = 100, T = 5, \xi = 1$

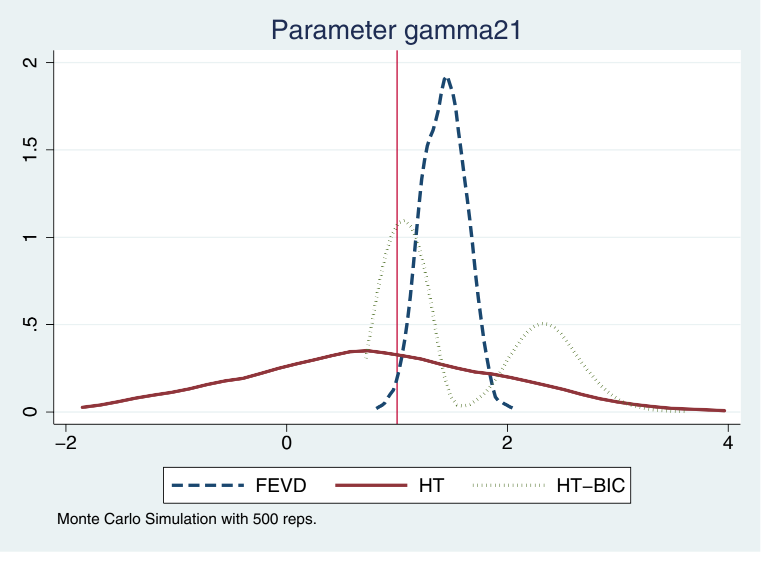

Figure 6: Histogram of selected IV-sets for Monte Carlo simulation results of γ_{21} with $N = 100, T = 5, \xi = 1$

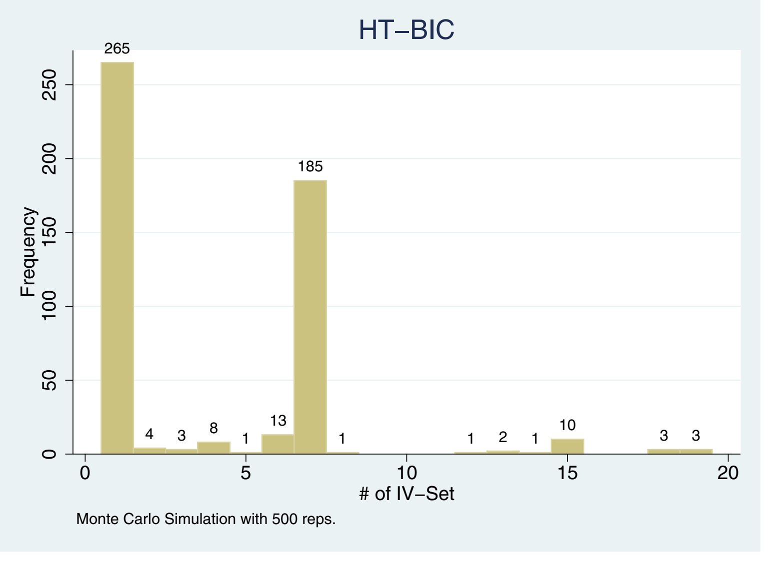

Figure 7: NOMAD surface plot for time-varying variables in MC simulations

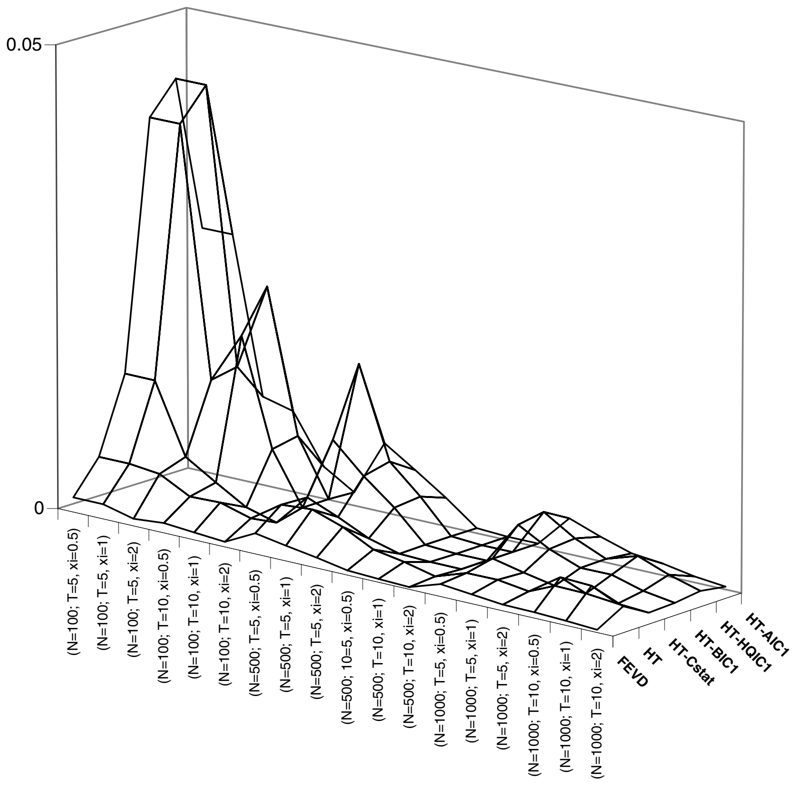

Figure 8: NORMSQD surface plot for time-varying variables in MC simulations

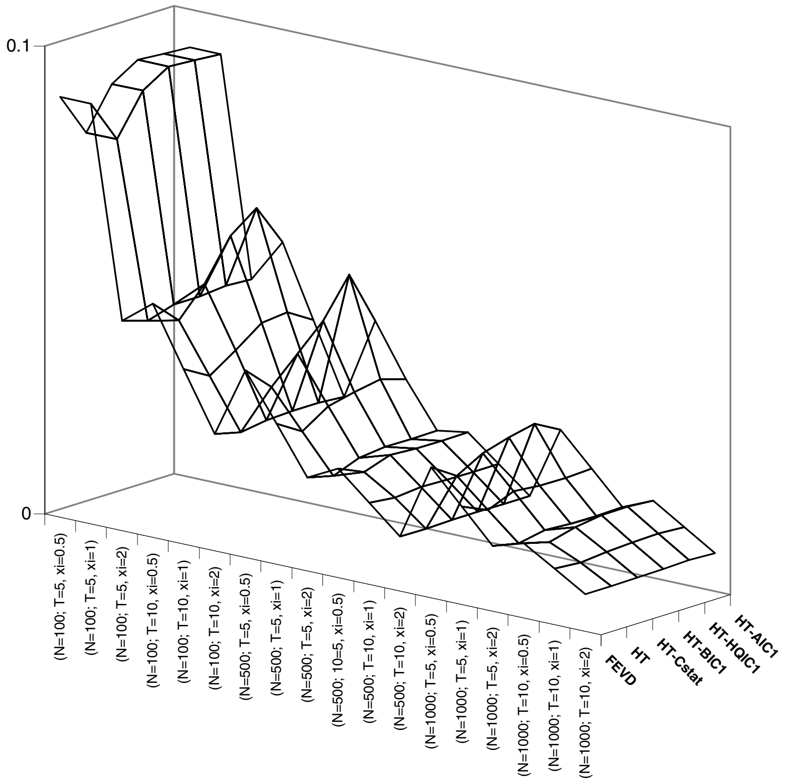

Figure 9: NOMAD surface plot for time-fixed variables in MC simulations

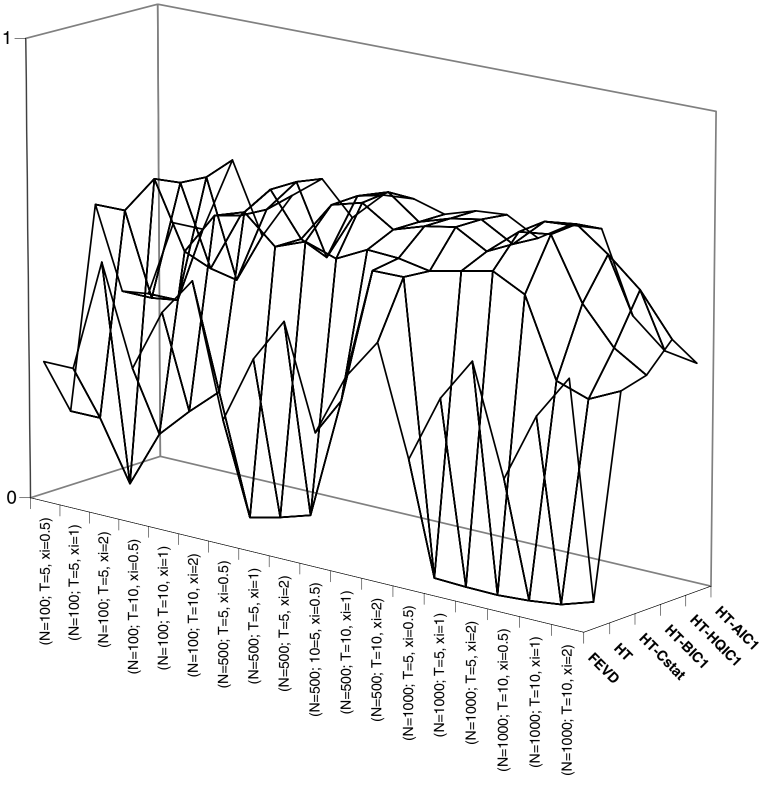

Figure 10: NORMSQD surface plot for time-fixed variables in MC simulations

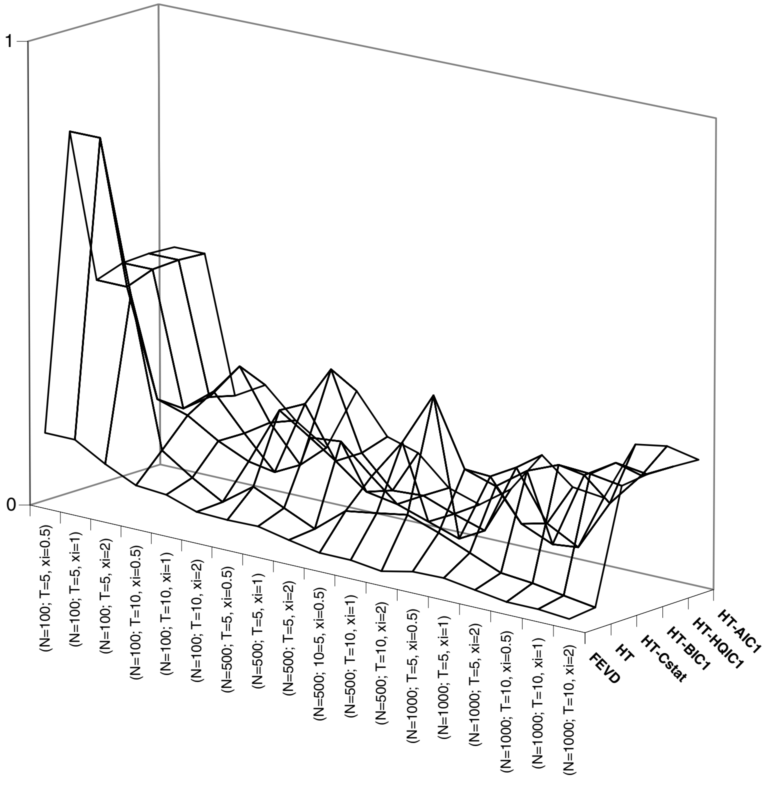

Table 1: NOMAD and NORMSQD averaged over all MC simulations

	Crit.	NOMAD	NORMSQD
Time-varying	FEVD	0.0009	0.0321
	HT	0.0029	0.0292
	HT-Cstat	0.0030	0.0321
	HT-BIC1	0.0086	0.0337
	HT-HQIC1	0.0099	0.0346
	HT-AIC1	0.0058	0.0329
Time-fixed	FEVD	0.4105	0.0672
	HT	0.1911	0.1888
	HT-Cstat	0.6009	0.2615
	HT-BIC1	0.6171	0.1990
	HT-HQIC1	0.6238	0.1952
	HT-AIC1	0.6231	0.2132
All variables	FEVD	0.2057	0.0497
	HT	0.0970	0.1090
	HT-Cstat	0.3019	0.1468
	HT-BIC1	0.3129	0.1164
	HT-HQIC1	0.3168	0.1149
	HT-AIC1	0.3144	0.1230

Note: For details about the Monte Carlo simulation setup see text.

Figure 11: NOMAD and NORMSQD values averaged over all MC simulations

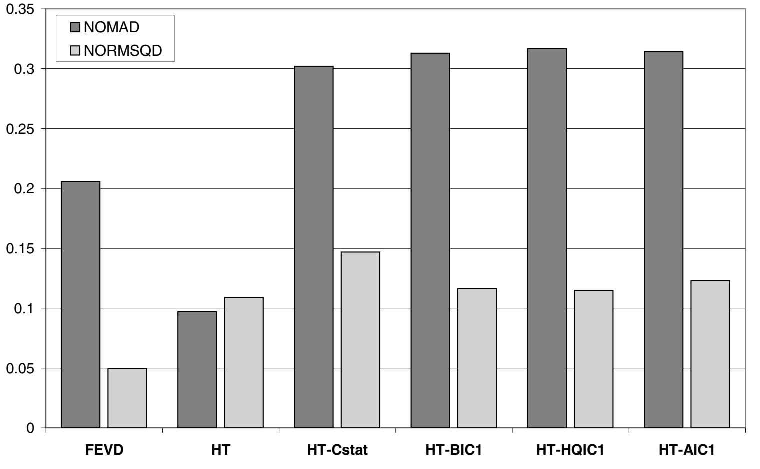

Table 2: Gravity model for EU wide Export flows for German states (NUTS1 level)

Log(EX)	**POLS**	**REM**	**FEM**	**FEVD**	**HT**$^{\$}$
Log(GDP_i)	1,04***	0,35*	0,83**	0,83***	0,87***
	(0,135)	(0,034)	(0,273)	(0,273)#	(0,271)
Log(GDP_j)	0,64***	0,31***	0,34***	0,34**	0,35***
	(0,026)	(0,034)	(0,044)	(0,044)#	(0,043)
Log(POP_i)	0,03	0.69***	-1,38***	-1,38**	0,18
	(0,132)	(0,197)	(0,398)	(0,398)#	(0,263)
Log(POP_j)	0,19***	0,48***	1,79***	1,79***	0,38***
	(0,025)	(0,041)	(0,302)	(0,302)#	(0,084)
Log($PROD_i$)	-0,15	2,11***	1,48***	1,48***	1,76***
	(0,241)	(0,228)	(0,275)	(0,275)#	(0,268)
Log($DIST_{ij}$)	-0,87***	-1,04***	(dropped)	-0,97***	-1,73***
	(0,021)	(0,052)		(0,021)#	(0,403)
SIM	-0,03***	-0,17***	-0,18***	-0,18***	-0,29***
	(0,011)	(0,052)	(0,062)	(0,048)#	(0,039)
RLF	0,01	0,03***	0,03***	0,03	0,03***
	(0,011)	(0,008)	(0,008)	(0,044)#	(0,007)
EMU	0,45***	0,36***	0,31***	0,31***	0,34***
	(0,029)	(0,019)	(0,021)	(0,054)#	(0,019)
EAST	-0,80***	-0,38***	(dropped)	-1,03***	-0,26**
	(0,039)	(0,075)		(0,043)#	(0,110)
BORDER	0,28***	0,26*	(dropped)	0,07***	-0,38
	(0,050)	(0,150)		(0,008)#	(0,438)
CEEC	0,47***	-0,20**	(dropped)	0,93***	-0,22*
	(0,055)	(0,086)		(0,063)#	(0,131)
No. of obs.	4784	4784	4784	4784	4784
No. of Groups		368	368	368	368
Time effects		yes	yes	yes	yes
Wald test (P-val.)		(0.00)	(0.00)	(0.00)	(0.00)
P-value of BP LM (POLS/REM)		0.00			
P-value of F-Test (POLS/FEM)			0.00		
Hausman m-stat.			147.2		
(REM/FEM)			(0.00)		
DWH endogeneity test					25.14
(P-value)					(0.00)
Sargan overid. test					6.25
(P-value)					(0.05)
C-Statistic for $Dist_{ij}$					14.12
(P-value)					(0.00)
Pagan-Hall IV het.test					35.9
(P-value)					(0.10)

Note: ***, **, * = denote significance levels at the 1%, 5% and 10% level respectively. Standard errors are robust to heteroskedasticity, # = corrected SEs for the FEVD estimator based on the *xtfevd* Stata routine provided by Plümper & Tröger (2007). $ = Using the C-Statistic based downward testing algorithm with group means of $X1 = [GDP_{j,t}, POP_{i,t}, RLF_{ij,t}]$ as IVs for $Z2 = [DIST_{ij}]$.

Table A.1: Monte Carlo simulation results with $N = (100, 500, 1000)$ for $T = 5$ and $\xi = 1$

Coef.	β_{11}			β_{12}			β_{21}			β_{22}			γ_{12}			γ_{21}		
Crit.	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse
$N = 100$																		
FEVD	0.000	0.091	0.091	-0.005	0.085	0.085	0.001	0.093	0.093	-0.001	0.090	0.090	-0.188	0.133	0.230	0.425	0.195	0.468
HT	-0.011	0.073	0.073	0.007	0.065	0.065	0.002	0.093	0.093	-0.001	0.090	0.090	0.095	0.412	0.422	-0.265	1.182	1.211
HT-BIC1	-0.028	0.065	0.070	-0.015	0.062	0.064	0.066	0.096	0.116	0.052	0.108	0.120	-0.229	0.293	0.372	0.590	0.668	0.891
HT-Cstat	-0.025	0.059	0.064	-0.012	0.057	0.058	0.076	0.095	0.122	0.067	0.106	0.125	-0.185	0.279	0.335	0.495	0.656	0.821
HT-BIC2	-0.032	0.061	0.068	-0.017	0.058	0.061	0.070	0.096	0.119	0.055	0.108	0.121	-0.211	0.295	0.362	0.564	0.673	0.878
HT-HQIC1	-0.024	0.077	0.081	-0.009	0.071	0.072	0.043	0.096	0.105	0.029	0.106	0.110	-0.313	0.310	0.440	0.740	0.653	0.987
HT-HQIC2	-0.024	0.077	0.081	-0.009	0.071	0.072	0.043	0.096	0.105	0.029	0.106	0.110	-0.313	0.310	0.440	0.740	0.653	0.987
HT-AIC1	-0.015	0.085	0.086	-0.006	0.078	0.078	0.023	0.094	0.097	0.009	0.095	0.095	-0.381	0.298	0.484	0.852	0.622	1.055
HT-AIC2	-0.016	0.085	0.086	-0.006	0.078	0.078	0.023	0.094	0.097	0.009	0.095	0.095	-0.381	0.298	0.484	0.852	0.622	1.055
$N = 500$																		
FEVD	-0.001	0.029	0.029	-0.002	0.031	0.031	-0.002	0.031	0.031	-0.003	0.029	0.029	-0.249	0.047	0.254	0.582	0.088	0.031
HT	-0.009	0.022	0.024	0.006	0.023	0.023	-0.001	0.031	0.031	-0.002	0.029	0.029	-0.034	0.037	0.050	0.079	0.139	0.571
HT-BIC1	-0.006	0.031	0.032	-0.002	0.032	0.032	-0.006	0.031	0.032	0.011	0.029	0.031	-0.426	0.106	0.439	0.963	0.202	0.521
HT-Cstat	-0.006	0.031	0.031	-0.001	0.032	0.032	-0.007	0.030	0.031	0.012	0.029	0.031	-0.439	0.090	0.448	0.989	0.166	0.525
HT-BIC2	-0.007	0.032	0.032	-0.003	0.033	0.033	-0.006	0.033	0.033	0.013	0.029	0.032	-0.428	0.108	0.441	0.968	0.200	0.519
HT-HQIC1	-0.005	0.030	0.031	-0.001	0.031	0.031	-0.006	0.031	0.031	0.009	0.029	0.030	-0.420	0.113	0.435	0.946	0.231	0.525
HT-HQIC2	-0.005	0.030	0.031	-0.001	0.031	0.031	-0.006	0.031	0.031	0.009	0.029	0.031	-0.420	0.113	0.435	0.946	0.230	0.525
HT-AIC1	-0.003	0.030	0.030	-0.002	0.030	0.030	-0.004	0.031	0.031	0.003	0.030	0.030	-0.390	0.140	0.414	0.875	0.304	0.528
HT-AIC2	-0.003	0.030	0.030	-0.002	0.030	0.030	-0.004	0.031	0.031	0.003	0.030	0.030	-0.390	0.140	0.414	0.875	0.304	0.528
$N = 1000$																		
FEVD	-0.001	0.020	0.020	-0.001	0.022	0.022	-0.002	0.021	0.021	0.000	0.021	0.021	-0.252	0.033	0.254	0.591	0.061	0.594
HT	0.001	0.016	0.016	-0.003	0.015	0.016	-0.002	0.021	0.021	0.000	0.021	0.021	0.001	0.039	0.039	0.004	0.128	0.128
HT-BIC1	-0.001	0.019	0.019	-0.002	0.021	0.021	0.004	0.020	0.020	-0.005	0.020	0.021	-0.451	0.077	0.458	0.988	0.141	0.998
HT-Cstat	-0.001	0.019	0.019	-0.002	0.021	0.021	0.004	0.020	0.020	-0.005	0.020	0.020	-0.457	0.062	0.461	1.001	0.105	1.006
HT-BIC2	-0.001	0.020	0.020	-0.002	0.021	0.021	0.005	0.019	0.020	-0.006	0.020	0.020	-0.455	0.068	0.460	0.997	0.115	1.003
HT-HQIC1	-0.001	0.019	0.019	-0.002	0.021	0.021	0.004	0.020	0.020	-0.005	0.021	0.021	-0.440	0.102	0.451	0.964	0.202	0.985
HT-HQIC2	-0.001	0.019	0.019	-0.002	0.021	0.021	0.004	0.020	0.020	-0.005	0.021	0.021	-0.440	0.102	0.451	0.964	0.202	0.985
HT-AIC1	-0.001	0.019	0.019	-0.001	0.022	0.022	0.001	0.021	0.021	-0.003	0.021	0.021	-0.410	0.137	0.432	0.899	0.278	0.941
HT-AIC2	-0.001	0.019	0.019	-0.001	0.022	0.022	0.001	0.021	0.021	-0.003	0.021	0.021	-0.410	0.137	0.432	0.899	0.278	0.941

Note: For details about the Monte Carlo simulation setup see text.

Table A.2: Monte Carlo simulation results with $N = (100, 500, 1000)$ for $T = 10$ and $\xi = 1$

Coef.	β_{11}			β_{12}			β_{21}			β_{22}			γ_{12}			γ_{21}		
Crit.	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse
									$N = 100$									
FEVD	-0.002	0.035	0.035	-0.001	0.038	0.038	0.000	0.040	0.040	0.000	0.037	0.037	-0.314	0.069	0.321	0.630	0.111	0.640
HT	-0.008	0.028	0.029	0.005	0.031	0.031	0.000	0.040	0.040	0.000	0.037	0.037	-0.177	0.056	0.185	0.304	0.162	0.344
HT-BIC1	-0.007	0.040	0.041	-0.005	0.043	0.043	0.012	0.042	0.044	0.004	0.041	0.041	-0.435	0.146	0.459	0.825	0.193	0.848
HT-Cstat	-0.007	0.036	0.037	-0.004	0.038	0.038	0.013	0.042	0.044	0.002	0.040	0.040	-0.433	0.137	0.454	0.864	0.195	0.886
HT-BIC2	-0.007	0.040	0.041	-0.006	0.043	0.043	0.013	0.043	0.045	0.005	0.041	0.042	-0.432	0.150	0.457	0.822	0.200	0.846
HT-HQIC1	-0.004	0.038	0.038	-0.002	0.040	0.040	0.007	0.040	0.040	0.001	0.039	0.039	-0.447	0.133	0.467	0.837	0.191	0.858
HT-HQIC2	-0.004	0.038	0.038	-0.002	0.040	0.040	0.007	0.040	0.040	0.001	0.039	0.039	-0.447	0.133	0.467	0.837	0.191	0.858
HT-AIC1	-0.002	0.036	0.036	-0.001	0.038	0.038	0.004	0.040	0.040	0.000	0.038	0.038	-0.452	0.140	0.474	0.842	0.200	0.865
HT-AIC2	-0.002	0.036	0.036	-0.001	0.038	0.038	0.004	0.040	0.040	0.000	0.038	0.038	-0.452	0.140	0.474	0.842	0.200	0.865
									$N = 500$									
FEVD	0.000	0.017	0.017	0.001	0.017	0.017	0.000	0.018	0.018	0.000	0.017	0.017	-0.257	0.029	0.258	0.589	0.051	0.591
HT	-0.002	0.015	0.016	0.003	0.016	0.016	0.000	0.018	0.018	0.000	0.017	0.017	0.362	0.064	0.367	-0.898	0.181	0.916
HT-BIC1	0.000	0.017	0.017	0.000	0.017	0.017	0.001	0.018	0.018	0.000	0.017	0.017	-0.436	0.052	0.439	0.945	0.085	0.949
HT-Cstat	0.000	0.017	0.017	0.000	0.017	0.017	0.001	0.018	0.018	0.000	0.017	0.017	-0.436	0.051	0.439	0.945	0.083	0.949
HT-BIC2	0.000	0.017	0.017	0.000	0.017	0.017	0.001	0.018	0.018	0.000	0.017	0.017	-0.436	0.051	0.439	0.945	0.083	0.949
HT-HQIC1	0.000	0.017	0.017	0.000	0.017	0.017	0.001	0.018	0.018	0.000	0.017	0.017	-0.437	0.053	0.440	0.945	0.085	0.948
HT-HQIC2	0.000	0.017	0.017	0.000	0.017	0.017	0.001	0.018	0.018	0.000	0.017	0.017	-0.437	0.053	0.440	0.945	0.085	0.948
HT-AIC1	0.000	0.017	0.017	0.001	0.017	0.017	0.000	0.018	0.018	0.000	0.017	0.017	-0.407	0.159	0.437	0.872	0.355	0.941
HT-AIC2	0.000	0.017	0.017	0.001	0.017	0.017	0.000	0.018	0.018	0.000	0.017	0.017	-0.407	0.159	0.437	0.872	0.355	0.941
									$N = 1000$									
FEVD	0.000	0.012	0.012	0.001	0.012	0.012	-0.001	0.012	0.012	0.001	0.012	0.012	-0.262	0.019	0.263	0.595	0.035	0.596
HT	-0.006	0.009	0.011	0.006	0.010	0.012	0.000	0.012	0.012	0.001	0.012	0.012	0.001	0.019	0.019	0.027	0.056	0.062
HT-BIC1	0.001	0.013	0.013	0.000	0.013	0.013	-0.001	0.012	0.012	0.002	0.013	0.013	-0.340	0.192	0.390	0.697	0.376	0.792
HT-Cstat	0.002	0.013	0.013	-0.002	0.013	0.013	-0.001	0.012	0.012	0.002	0.013	0.013	-0.489	0.042	0.491	0.998	0.060	1.000
HT-BIC2	0.001	0.013	0.013	-0.001	0.013	0.013	-0.002	0.012	0.012	0.003	0.013	0.013	-0.410	0.161	0.440	0.837	0.316	0.895
HT-HQIC1	0.000	0.012	0.012	0.000	0.013	0.013	-0.001	0.012	0.012	0.001	0.013	0.013	-0.312	0.192	0.366	0.642	0.376	0.744
HT-HQIC2	0.000	0.012	0.012	0.000	0.013	0.013	-0.001	0.012	0.012	0.001	0.013	0.013	-0.312	0.192	0.366	0.642	0.376	0.744
HT-AIC1	0.000	0.012	0.012	0.000	0.013	0.013	-0.001	0.012	0.012	0.001	0.012	0.012	-0.280	0.188	0.337	0.577	0.368	0.685
HT-AIC2	0.000	0.012	0.012	0.000	0.013	0.013	-0.001	0.012	0.012	0.001	0.012	0.012	-0.280	0.188	0.337	0.577	0.368	0.685

Note: For details about the Monte Carlo simulation setup see text.

Table A.3: Monte Carlo simulation results with $N = (100, 500, 1000)$ for $T = 5$ and $\xi = 2$

Coef.	β_{11}			β_{12}			β_{21}			β_{22}			γ_{12}			γ_{21}		
Crit.	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse
$N = 100$																		
FEVD	0.000	0.045	0.045	-0.002	0.043	0.043	0.000	0.047	0.047	-0.001	0.045	0.045	-0.336	0.089	0.347	0.769	0.164	0.786
HT	-0.012	0.041	0.043	0.004	0.038	0.038	0.002	0.047	0.047	0.001	0.045	0.045	-0.050	0.250	0.255	0.053	0.698	0.699
HT-BIC1	-0.025	0.039	0.047	-0.008	0.034	0.035	0.018	0.047	0.051	0.001	0.047	0.047	-0.389	0.124	0.409	0.959	0.263	0.995
HT-Cstat	-0.027	0.035	0.045	-0.008	0.032	0.033	0.020	0.050	0.053	0.002	0.049	0.049	-0.382	0.127	0.402	0.954	0.282	0.994
HT-BIC2	-0.026	0.039	0.047	-0.008	0.034	0.035	0.019	0.048	0.052	0.001	0.048	0.048	-0.387	0.128	0.408	0.955	0.275	0.994
HT-HQIC1	-0.019	0.045	0.049	-0.004	0.039	0.039	0.014	0.046	0.048	-0.001	0.044	0.044	-0.409	0.125	0.428	0.974	0.248	1.005
HT-HQIC2	-0.019	0.045	0.049	-0.004	0.039	0.039	0.014	0.046	0.048	-0.001	0.044	0.044	-0.409	0.125	0.428	0.974	0.248	1.005
HT-AIC1	-0.011	0.046	0.048	0.000	0.041	0.041	0.009	0.046	0.047	-0.002	0.045	0.045	-0.424	0.148	0.449	0.975	0.315	1.024
HT-AIC2	-0.011	0.046	0.048	0.000	0.041	0.041	0.009	0.046	0.047	-0.002	0.045	0.045	-0.424	0.148	0.449	0.975	0.315	1.024
$N = 500$																		
FEVD	-0.001	0.020	0.020	-0.001	0.021	0.021	-0.001	0.020	0.020	-0.002	0.019	0.019	-0.307	0.036	0.309	0.714	0.071	0.718
HT	-0.007	0.016	0.017	0.005	0.016	0.017	-0.001	0.020	0.020	-0.001	0.019	0.019	-0.047	0.025	0.053	0.106	0.092	0.140
HT-BIC1	-0.004	0.022	0.022	-0.001	0.023	0.023	-0.004	0.020	0.020	0.007	0.019	0.020	-0.444	0.071	0.450	0.982	0.134	0.991
HT-Cstat	-0.004	0.022	0.022	-0.001	0.023	0.023	-0.004	0.020	0.020	0.007	0.019	0.020	-0.451	0.061	0.455	0.997	0.103	1.002
HT-BIC2	-0.004	0.022	0.022	-0.001	0.023	0.023	-0.005	0.020	0.020	0.007	0.019	0.020	-0.448	0.064	0.453	0.991	0.112	0.997
HT-HQIC1	-0.003	0.022	0.022	-0.001	0.023	0.023	-0.004	0.020	0.021	0.005	0.020	0.021	-0.430	0.099	0.441	0.948	0.204	0.970
HT-HQIC2	-0.003	0.022	0.022	-0.001	0.023	0.023	-0.004	0.020	0.021	0.005	0.020	0.021	-0.430	0.099	0.441	0.948	0.204	0.970
HT-AIC1	-0.002	0.020	0.020	-0.001	0.021	0.021	-0.002	0.021	0.021	0.001	0.021	0.021	-0.382	0.146	0.409	0.837	0.311	0.893
HT-AIC2	-0.002	0.020	0.020	-0.001	0.021	0.021	-0.002	0.021	0.021	0.001	0.021	0.021	-0.382	0.146	0.409	0.837	0.311	0.893
$N = 1000$																		
FEVD	-0.001	0.013	0.013	0.000	0.015	0.015	-0.001	0.014	0.014	0.000	0.014	0.014	-0.307	0.025	0.308	0.720	0.049	0.722
HT	0.001	0.011	0.011	-0.003	0.012	0.012	-0.001	0.014	0.014	0.000	0.014	0.014	0.002	0.026	0.026	0.001	0.085	0.085
HT-BIC1	0.000	0.013	0.013	-0.002	0.014	0.015	0.002	0.013	0.013	-0.003	0.013	0.014	-0.463	0.057	0.466	0.995	0.104	1.000
HT-Cstat	0.000	0.013	0.013	-0.002	0.014	0.015	0.002	0.013	0.013	-0.003	0.013	0.014	-0.466	0.043	0.468	1.002	0.069	1.004
HT-BIC2	0.000	0.013	0.013	-0.002	0.014	0.015	0.003	0.013	0.013	-0.003	0.013	0.014	-0.466	0.043	0.468	1.002	0.069	1.004
HT-HQIC1	0.000	0.013	0.013	-0.002	0.014	0.014	0.002	0.014	0.014	-0.003	0.014	0.014	-0.449	0.095	0.459	0.966	0.192	0.985
HT-HQIC2	0.000	0.013	0.013	-0.002	0.014	0.014	0.002	0.014	0.014	-0.003	0.014	0.014	-0.449	0.095	0.459	0.966	0.192	0.985
HT-AIC1	-0.001	0.013	0.013	-0.001	0.015	0.015	0.000	0.014	0.014	-0.001	0.014	0.014	-0.392	0.164	0.425	0.846	0.345	0.914
HT-AIC2	-0.001	0.013	0.013	-0.001	0.015	0.015	0.000	0.014	0.014	-0.001	0.014	0.014	-0.392	0.164	0.425	0.846	0.345	0.914

Note: For details about the Monte Carlo simulation setup see text.

Table A.4: Monte Carlo simulation results with $N = (100, 500, 1000)$ for $T = 10$ and $\xi = 2$

Coef.	β_{11}			β_{12}			β_{21}			β_{22}			γ_{12}			γ_{21}		
Crit.	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse
$N = 100$																		
FEVD	-0.001	0.023	0.023	-0.001	0.025	0.025	0.000	0.027	0.027	0.000	0.025	0.024	-0.366	0.051	0.369	0.743	0.086	0.747
HT	-0.005	0.021	0.021	0.003	0.022	0.022	0.000	0.027	0.027	0.000	0.025	0.025	-0.218	0.038	0.222	0.376	0.105	0.391
HT-BIC1	-0.001	0.023	0.023	0.000	0.024	0.024	0.004	0.026	0.026	-0.001	0.025	0.025	-0.485	0.082	0.491	0.916	0.120	0.924
HT-Cstat	-0.001	0.021	0.021	0.001	0.023	0.023	0.004	0.026	0.026	-0.001	0.025	0.025	-0.484	0.072	0.489	0.928	0.109	0.934
HT-BIC2	-0.001	0.023	0.023	0.000	0.024	0.024	0.004	0.026	0.026	-0.001	0.025	0.025	-0.485	0.079	0.491	0.916	0.117	0.924
HT-HQIC1	-0.001	0.022	0.022	0.000	0.024	0.024	0.003	0.027	0.027	-0.001	0.025	0.025	-0.480	0.094	0.489	0.903	0.142	0.914
HT-HQIC2	-0.001	0.022	0.022	0.000	0.024	0.024	0.003	0.027	0.027	-0.001	0.025	0.025	-0.480	0.094	0.489	0.903	0.142	0.914
HT-AIC1	-0.001	0.022	0.022	0.000	0.024	0.024	0.001	0.027	0.027	-0.001	0.025	0.025	-0.464	0.109	0.476	0.872	0.173	0.889
HT-AIC2	-0.001	0.022	0.022	0.000	0.024	0.024	0.001	0.027	0.027	-0.001	0.025	0.025	-0.464	0.109	0.476	0.872	0.173	0.889
$N = 500$																		
FEVD	0.000	0.011	0.011	0.001	0.011	0.011	0.000	0.012	0.012	0.000	0.012	0.012	-0.308	0.021	0.309	0.711	0.041	0.712
HT	-0.001	0.011	0.011	0.001	0.011	0.011	0.000	0.012	0.012	0.000	0.012	0.012	0.637	0.073	0.641	-1.582	0.194	1.593
HT-BIC1	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.446	0.038	0.448	0.959	0.055	0.960
HT-Cstat	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.446	0.038	0.448	0.959	0.055	0.960
HT-BIC2	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.446	0.038	0.448	0.959	0.055	0.960
HT-HQIC1	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.443	0.075	0.449	0.950	0.160	0.963
HT-HQIC2	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.443	0.075	0.449	0.950	0.160	0.963
HT-AIC1	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.401	0.224	0.459	0.850	0.520	0.996
HT-AIC2	0.001	0.012	0.012	0.000	0.012	0.012	0.000	0.012	0.012	0.000	0.011	0.011	-0.401	0.224	0.459	0.850	0.520	0.996
$N = 1000$																		
FEVD	0.000	0.008	0.008	0.000	0.008	0.008	0.000	0.008	0.008	0.000	0.008	0.008	-0.321	0.015	0.321	0.724	0.028	0.725
HT	-0.004	0.007	0.008	0.003	0.008	0.008	0.000	0.008	0.008	0.001	0.008	0.008	-0.011	0.014	0.018	0.061	0.040	0.073
HT-BIC1	0.001	0.008	0.008	0.000	0.009	0.009	0.000	0.008	0.008	0.001	0.008	0.008	-0.314	0.188	0.366	0.638	0.361	0.733
HT-Cstat	0.003	0.008	0.009	-0.002	0.008	0.009	0.000	0.008	0.008	0.001	0.008	0.008	-0.498	0.029	0.499	1.002	0.039	1.002
HT-BIC2	0.001	0.008	0.009	-0.001	0.009	0.009	-0.001	0.008	0.008	0.001	0.008	0.008	-0.352	0.186	0.398	0.715	0.360	0.800
HT-HQIC1	0.000	0.008	0.008	0.000	0.009	0.009	0.000	0.008	0.008	0.000	0.008	0.008	-0.307	0.185	0.358	0.625	0.355	0.718
HT-HQIC2	0.000	0.008	0.008	0.000	0.009	0.009	0.000	0.008	0.008	0.000	0.008	0.008	-0.307	0.185	0.358	0.625	0.355	0.718
HT-AIC1	0.000	0.008	0.008	0.000	0.009	0.009	0.000	0.008	0.008	0.000	0.008	0.008	-0.303	0.184	0.355	0.618	0.352	0.711
HT-AIC2	0.000	0.008	0.008	0.000	0.009	0.009	0.000	0.008	0.008	0.000	0.008	0.008	-0.303	0.184	0.355	0.618	0.352	0.711

Note: For details about the Monte Carlo simulation setup see text.

Table A.5: Monte Carlo simulation results with $N = (100, 500, 1000)$ for $T = 5$ and $\xi = 0.5$

Coef.	β_{11}			β_{12}			β_{21}			β_{22}			γ_{12}			γ_{21}		
Crit.	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse
$N = 100$																		
FEVD	0.000	0.091	0.091	-0.005	0.085	0.085	0.001	0.093	0.093	-0.001	0.090	0.090	-0.188	0.133	0.230	0.425	0.195	0.468
HT	-0.011	0.073	0.073	0.007	0.065	0.065	0.002	0.093	0.093	-0.001	0.090	0.090	0.095	0.412	0.422	-0.265	1.182	1.211
HT-BIC1	-0.028	0.065	0.070	-0.015	0.062	0.064	0.066	0.096	0.116	0.052	0.108	0.120	-0.229	0.293	0.372	0.590	0.668	0.891
HT-Cstat	-0.025	0.059	0.064	-0.012	0.057	0.058	0.076	0.095	0.122	0.067	0.106	0.125	-0.185	0.279	0.335	0.495	0.656	0.821
HT-BIC2	-0.032	0.061	0.068	-0.017	0.058	0.061	0.070	0.096	0.119	0.055	0.108	0.121	-0.211	0.295	0.362	0.564	0.673	0.878
HT-HQIC1	-0.024	0.077	0.081	-0.009	0.071	0.072	0.043	0.096	0.105	0.029	0.106	0.110	-0.313	0.310	0.440	0.740	0.653	0.987
HT-HQIC2	-0.024	0.077	0.081	-0.009	0.071	0.072	0.043	0.096	0.105	0.029	0.106	0.110	-0.313	0.310	0.440	0.740	0.653	0.987
HT-AIC1	-0.015	0.085	0.086	-0.006	0.078	0.078	0.023	0.094	0.097	0.009	0.095	0.095	-0.381	0.298	0.484	0.852	0.622	1.055
HT-AIC2	-0.016	0.085	0.086	-0.006	0.078	0.078	0.023	0.094	0.097	0.009	0.095	0.095	-0.381	0.298	0.484	0.852	0.622	1.055
$N = 500$																		
FEVD	-0.001	0.039	0.039	-0.002	0.041	0.041	-0.002	0.041	0.041	-0.003	0.039	0.039	-0.161	0.051	0.169	0.378	0.080	0.386
HT	-0.008	0.029	0.030	0.004	0.029	0.029	-0.002	0.041	0.041	-0.003	0.039	0.039	-0.023	0.051	0.056	0.057	0.188	0.196
HT-BIC1	-0.012	0.042	0.043	-0.009	0.042	0.043	-0.001	0.048	0.048	0.018	0.044	0.047	-0.363	0.182	0.406	0.851	0.344	0.918
HT-Cstat	-0.009	0.041	0.042	-0.007	0.042	0.042	-0.003	0.048	0.048	0.020	0.043	0.047	-0.405	0.158	0.435	0.931	0.303	0.979
HT-BIC2	-0.014	0.040	0.043	-0.011	0.041	0.042	0.013	0.065	0.066	0.029	0.056	0.063	-0.327	0.213	0.390	0.775	0.427	0.885
HT-HQIC1	-0.008	0.040	0.041	-0.005	0.040	0.040	-0.005	0.045	0.045	0.012	0.041	0.043	-0.376	0.174	0.414	0.866	0.337	0.929
HT-HQIC2	-0.008	0.040	0.041	-0.005	0.040	0.040	-0.005	0.045	0.045	0.012	0.041	0.043	-0.376	0.174	0.414	0.866	0.337	0.929
HT-AIC1	-0.005	0.039	0.039	-0.003	0.039	0.039	-0.005	0.042	0.042	0.005	0.039	0.040	-0.370	0.173	0.409	0.845	0.360	0.918
HT-AIC2	-0.005	0.039	0.039	-0.003	0.039	0.039	-0.005	0.042	0.042	0.005	0.039	0.040	-0.370	0.173	0.409	0.845	0.360	0.918
$N = 1000$																		
FEVD	-0.002	0.026	0.026	-0.001	0.029	0.029	-0.002	0.028	0.028	0.000	0.028	0.028	-0.165	0.036	0.169	0.390	0.056	0.394
HT	0.000	0.021	0.021	-0.003	0.020	0.020	-0.002	0.028	0.028	0.000	0.028	0.028	-0.002	0.051	0.051	0.010	0.170	0.171
HT-BIC1	-0.004	0.028	0.029	-0.005	0.029	0.030	0.006	0.028	0.029	-0.004	0.029	0.029	-0.410	0.141	0.433	0.919	0.251	0.952
HT-Cstat	-0.002	0.027	0.027	-0.004	0.028	0.028	0.007	0.028	0.029	-0.006	0.028	0.028	-0.442	0.097	0.453	0.985	0.168	1.000
HT-BIC2	-0.005	0.029	0.030	-0.006	0.030	0.031	0.008	0.029	0.030	-0.003	0.030	0.031	-0.409	0.142	0.433	0.918	0.251	0.952
HT-HQIC1	-0.003	0.026	0.026	-0.003	0.028	0.028	0.004	0.026	0.027	-0.005	0.028	0.028	-0.408	0.144	0.433	0.911	0.269	0.949
HT-HQIC2	-0.003	0.026	0.026	-0.003	0.028	0.028	0.004	0.026	0.027	-0.005	0.028	0.028	-0.408	0.144	0.433	0.911	0.269	0.949
HT-AIC1	-0.002	0.026	0.026	-0.003	0.029	0.029	0.002	0.027	0.027	-0.003	0.028	0.029	-0.398	0.156	0.427	0.881	0.302	0.931
HT-AIC2	-0.002	0.026	0.026	-0.003	0.029	0.029	0.002	0.027	0.027	-0.003	0.028	0.029	-0.398	0.156	0.427	0.881	0.302	0.931

Note: For details about the Monte Carlo simulation setup see text.

Table A.6: Monte Carlo simulation results with $N = (100, 500, 1000)$ for $T = 10$ and $\xi = 0.5$

Coef.	β_{11}			β_{12}			β_{21}			β_{22}			γ_{12}			γ_{21}		
Crit.	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse	bias	s.d.	rmse
								$N = 100$										
FEVD	-0.003	0.047	0.047	-0.001	0.051	0.051	0.000	0.053	0.053	0.000	0.049	0.049	-0.226	0.078	0.239	0.446	0.113	0.460
HT	-0.008	0.036	0.037	0.004	0.039	0.039	0.000	0.053	0.053	0.000	0.049	0.049	-0.129	0.076	0.149	0.222	0.229	0.319
HT-BIC1	-0.016	0.050	0.053	-0.016	0.053	0.055	0.026	0.063	0.068	0.017	0.059	0.061	-0.335	0.202	0.391	0.668	0.288	0.728
HT-Cstat	-0.016	0.047	0.049	-0.015	0.050	0.052	0.031	0.061	0.069	0.016	0.060	0.062	-0.337	0.194	0.389	0.710	0.306	0.773
HT-BIC2	-0.016	0.048	0.051	-0.015	0.050	0.053	0.035	0.070	0.078	0.026	0.068	0.073	-0.304	0.221	0.376	0.616	0.345	0.706
HT-HQIC1	-0.008	0.050	0.051	-0.007	0.052	0.053	0.013	0.058	0.059	0.007	0.055	0.056	-0.375	0.188	0.419	0.703	0.262	0.750
HT-HQIC2	-0.008	0.050	0.051	-0.007	0.052	0.053	0.013	0.058	0.059	0.007	0.055	0.056	-0.375	0.188	0.419	0.703	0.262	0.750
HT-AIC1	-0.004	0.049	0.049	-0.003	0.051	0.051	0.005	0.052	0.053	0.002	0.051	0.051	-0.392	0.176	0.430	0.719	0.246	0.759
HT-AIC2	-0.004	0.049	0.049	-0.003	0.051	0.051	0.005	0.052	0.053	0.002	0.051	0.051	-0.392	0.176	0.430	0.719	0.246	0.759
								$N = 500$										
FEVD	0.000	0.023	0.023	0.001	0.023	0.023	0.000	0.025	0.025	0.000	0.023	0.023	-0.176	0.031	0.179	0.400	0.049	0.403
HT	-0.003	0.017	0.017	0.004	0.018	0.019	0.000	0.025	0.025	0.001	0.023	0.023	0.194	0.055	0.201	-0.482	0.169	0.511
HT-BIC1	0.001	0.022	0.022	0.000	0.023	0.023	0.001	0.023	0.023	0.000	0.022	0.022	-0.420	0.067	0.426	0.919	0.115	0.926
HT-Cstat	0.001	0.022	0.022	0.000	0.023	0.022	0.001	0.023	0.023	0.000	0.022	0.022	-0.420	0.067	0.425	0.919	0.115	0.926
HT-BIC2	0.001	0.022	0.022	0.000	0.023	0.023	0.001	0.023	0.023	0.000	0.022	0.022	-0.420	0.067	0.425	0.919	0.115	0.926
HT-HQIC1	0.001	0.022	0.022	0.000	0.023	0.023	0.001	0.023	0.023	0.000	0.022	0.022	-0.420	0.066	0.425	0.918	0.114	0.926
HT-HQIC2	0.001	0.022	0.022	0.000	0.023	0.023	0.001	0.023	0.023	0.000	0.022	0.022	-0.420	0.066	0.425	0.918	0.114	0.926
HT-AIC1	0.001	0.022	0.022	0.001	0.023	0.023	0.000	0.024	0.024	0.000	0.022	0.022	-0.409	0.115	0.424	0.877	0.223	0.905
HT-AIC2	0.001	0.022	0.022	0.001	0.023	0.023	0.000	0.024	0.024	0.000	0.022	0.022	-0.409	0.115	0.424	0.877	0.223	0.905
								$N = 1000$										
FEVD	0.000	0.016	0.016	0.001	0.017	0.017	-0.001	0.015	0.015	0.001	0.016	0.016	-0.172	0.021	0.173	0.392	0.033	0.394
HT	-0.006	0.011	0.012	0.006	0.013	0.014	0.000	0.015	0.015	0.001	0.016	0.016	0.005	0.024	0.024	0.007	0.073	0.073
HT-BIC1	0.001	0.016	0.016	0.000	0.017	0.017	-0.002	0.016	0.016	0.004	0.017	0.018	-0.389	0.163	0.421	0.807	0.323	0.869
HT-Cstat	0.002	0.017	0.017	-0.001	0.017	0.017	-0.002	0.016	0.016	0.004	0.017	0.018	-0.479	0.055	0.482	0.989	0.083	0.993
HT-BIC2	0.001	0.017	0.017	-0.001	0.017	0.017	-0.003	0.015	0.016	0.007	0.017	0.018	-0.445	0.113	0.459	0.922	0.217	0.947
HT-HQIC1	0.000	0.016	0.016	0.000	0.017	0.017	-0.002	0.016	0.016	0.003	0.017	0.017	-0.349	0.182	0.393	0.725	0.364	0.811
HT-HQIC2	0.000	0.016	0.016	0.000	0.017	0.017	-0.002	0.016	0.016	0.003	0.017	0.017	-0.349	0.182	0.393	0.725	0.364	0.811
HT-AIC1	0.000	0.016	0.016	0.001	0.016	0.016	-0.001	0.016	0.016	0.001	0.017	0.017	-0.293	0.188	0.348	0.612	0.377	0.719
HT-AIC2	0.000	0.016	0.016	0.001	0.016	0.016	-0.001	0.016	0.016	0.001	0.017	0.017	-0.293	0.188	0.348	0.612	0.377	0.719

Note: For details about the Monte Carlo simulation setup see text.

Table A.7: Data description and source for Export model

Variable	Description	Source
EX_{ijt}	Export volume, nominal values, in Mio.	Statistisches Bundesamt (German statistical office)
GDP_{it}	Gross Domestic Product, nominal values, in Mio.	VGR der Länder (Statistical office of the German states)
GDP_{jt}	Gross Domestic Product, nominal values, in Mio.	EUROSTAT
POP_{it}	Population, in 1000	VGR der Länder (Statistical office of the German states)
POP_{jt}	Population, in 1000	Groningen Growth & Development center (GGDC)
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}	$RLF = log\left|\left(\frac{GDP_{it}}{POP_{it}}\right) - \left(\frac{GDP_{jt}}{POP_{jt}}\right)\right|$	see above
EMP_{it}	Employment, in 1000	VGR der Länder (Statistical office of the German states)
EMP_{jt}	Employment, in 1000	AMECO database of the European Commission
$PROD_{it}$	$Prod_{it} = \left(\frac{GDP_{it}}{EMP_{it}}\right)$	see above
$PROD_{jt}$	$Prod_{jt} = \left(\frac{GDP_{jt}}{EMP_{jt}}\right)$	see above
$DIST_{ij}$	Distance between state capital for Germany and national capital for the EU27 countries, in km	Calculation based on coordinates, 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	