DISS. ETH NO. 21219

Essays in Applied Microeconomics

An Econometric Analysis of Swiss Gasoline Demand Using Panel Data

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

FABIAN MAX HEIMSCH

Master of Science in Management, Technology and Economics
ETH ZURICH

born 8th of September 1981

citizen of St. Gallen (SG)

accepted on the recommendation of

Prof. Dr. Massimo Filippini, examiner

Prof. Dr. Peter Egger, co-examiner

2013

Preface

First of all, I am especially grateful to Prof. Dr. Massimo Filippini who gave me the opportunity to write a doctoral thesis at the Centre of Energy Policy and Economics (CEPE). Without his staunch support, the challenging discussions and the exemplary way he is dedicated to academic work, this dissertation could not have been put down to paper. Most of all, it was an enriching experience to me that the task of writing a dissertation in the field of empirical economic research was not conducted in a lonely office but accompanied by a group of extraordinarily motivated people. In this environment, I was given every opportunity to plunge into academic research, which Massimo Filippini offered me in several rewarding ways: First, I was given the opportunity to work on a project with the Federal Office of Energy. The outcome of this project was a final report which served me as a very fertile base for the first essay of this dissertation. At this stage, I want to thank all people involved in the joint committee. A very special thankyou also goes to Dr. Silvia Banfi, whose deep intuition and understanding of economics proved to be essential and indispensable for finishing the project and the first essay of this dissertation successfully. A second valuable experience was the chance to collaborate in the preparation and teaching of the lectures offered by the chair. I enjoyed a high degree of independence, which also assured me of Massimo Filippini's confidence in me. Last but not least, I had the pleasure to travel to Maryland for a sojourn at the chair of Agricultural and Resource Economics at the University of Maryland. Therefore, a very special thank goes to Dr. Anna Alberini, associate professor at CEPE, whose profound understanding of econometric issues helped me a great deal in writing the last part of this thesis. Without Massimo Filippini's guidance, patience and extraordinarily high quality of support, none of this would have been possible. Thank you for everything!

I further am deeply indebted to my co-supervisor Prof. Dr. Peter Egger, who taught me the core issues of spatial econometrics, which in turn enabled me to write the second part of this dissertation. I am especially glad that he is my co-supervisor, since my research interests go in very similar directions to his. Also, I learnt a great deal in the courses on econometrics he taught and especially while reading his publications.

During the last four years, I received scientific guidance at many stages. Therefore, I want to thank Dr. Elisa Tosetti for our fruitful discussions concerning spatial econometric issues. A deep thankyou also goes to Dr. Leticia Blázquez for her invaluable support and for giving me the opportunity to co-author a publication. Also, I want to thank Giuliano Masiero for our challenging discussions on health economics and related econometric issues and the chance to co-author another publication. A special thanks of course goes to the colleagues from CEPE, Dr. Aurelio Fetz, Dr. Souvik Datta, Dr. Jelena Zoric, Dr. Andrea Horehájová, Dr. Mozghan Alaeifar, Dr. Martin Koller, Markus Bareit, Laura Di Giorgio, Florian Landis, Jan Imhof, Justin Caron, Renger van Nieuwkoop, Thomas Geissmann, Nina Boogen, Jolanda Staufer, Andrea Gossweiler and Rina Fichtl. I also want to thank Dr. Christian Amon for valuable support and for discussing issues not necessarily related to economics. Most important here, I want to thank Céline Ramseier for sharing the office with me for four years, she was a great support to me in all situations: thank you for this, Céline.

Also, I need to express esteem to my former teachers. Many of them inspired me and without teachers of this quality, I surely could not have gone to university: thank you all.

Many friends gave me support during the past four years. First of all, I would like to thank my friend and neighbour Peter Leuzinger, who unfortunately passed away last year. His different way of thinking clearly broadened my horizons, and I am proud to have known him. I also owe gratitude to his lovely parents. Clearly, I would love to include all my friends and people I appreciate in this acknowledgement, but this list could not be exhaustive.

Most importantly, I want to express my deepest gratitude and thanks to my lovely parents, to whom I am infinitely much obliged. Both of them have supported me in everything I have done and have never raised any doubts. I clearly have the best parents anyone could have. Thanks a lot for everything, Mami und Papi. I also want to address my grandparents Marianne, Franz, Ruth and Max with a big thank-you for their love, mental support and motivation.

Of course, I also want to express my gratitude to my twin brother Thomas, for being a friend, soul mate and brother who accompanied me for all my life.

Finally, I want to express my appreciation, gratitude and love to my girlfriend Prisca, who has accompanied me the past four years. It is mostly due to you that I enjoyed this period in my life so much and it was you who often brought me back to reality when I was lost in academic work. I want to thank you especially for your loving company, your incredible patience and generally for every day we have been able to share. Of course, this thankyou also applies to her lovely parents, Robert and Thérèse.

Fabian Heimsch, May 2013, Zurich

Summary

This dissertation is composed of two essays on the econometric analysis of the demand for Swiss gasoline using panel data and a third essay on the econometric properties of the Arellano Bond and Blundell Bond estimators in the context of exogenous variables with a low within variation.

The goal of the first essay is to estimate a demand function for Swiss gasoline in Switzerland's border regions and simultaneously to quantify the amount of gasoline sold to foreigners. The share of gasoline sold to foreigners is often labelled as "gasoline tourism" and is driven by the existing price differentials across the border. Therefore, the price ratio is used as an explanatory variable in an econometric model and is also weighted with the distance from the border to determine the distance range from the border within which gasoline tourism is supposed to have an impact on the amount of gasoline sold. For this reason, gasoline sales data from the five largest gasoline retailers operating in Switzerland were collected. The sales from gasoline stations were averaged at the municipal level serving as the sales from a representative gasoline station for that municipality. A static panel data model accounting for fixed and random effects was estimated. We found a significant price elasticity of Swiss gasoline demand (in the border regions) with respect to the Swiss gasoline price of -0.211 when gasoline tourism was ruled out. The average price elasticity with gasoline tourism considered is -0.65. Accordingly, the demand for Swiss gasoline is considerably affected by foreigners purchasing gasoline in the Swiss border regions. The results indicate that gasoline tourism had an impact on municipalities up to a distance of some 30 kilometres from the border, but the main part was sold at stations located within 15 kilometres. The model shows that, compared to overall sales, which accounted for some 4.5 billion litres of gasoline, gasoline tourism reached values from some 250 million litres up to 450 million litres of gasoline on average. Further, it is shown that even a small increase in the Swiss gasoline price may lead to a substantial decrease in gasoline tourism.

The overall goal of Essay 2 is to estimate the demand for Swiss gasoline at the municipal level while taking spatial effects into account. The demand for gasoline can clearly be seen as a spatial story, since the consumption of gasoline not only depends on a municipality's car fleet or population, but also on exchange traffic. We use a panel data model with spatially lagged residuals and a spatially lagged dependent variable and account for random effects. We estimate a coefficient of the spatially lagged dependent variable of 0.34 and a coefficient of the spatially lagged residuals of 0.37. This implies that an increase in gasoline demand in one municipality

by 10.0% spreads over to other municipalities and leads to an increase of 3.4% in demand, given that the regions are neighbours. As a main result, we estimate an average price elasticity of Swiss gasoline demand of -0.655 (total effect). Spatial partitioning of this value leads to a direct effect of -0.58 on average. We estimate price elasticities ranging from -0.585 to -0.855, depending on the municipalities' locations. Due to the very different approach, we find a price elasticity of Swiss gasoline demand with respect to the foreign price of 0.32 (as the average total effect), which is significantly different from that reported in Baranzini et al. (2012) but not from the value obtained in Essay 1. As the demand for gasoline in one municipality affects demand in neighbouring municipalities, those results can impose important consequences on policy makers: First, there are regions which react more sensitively to change in the gasoline price than others (e.g., this change might result from the introduction of a CO_2 tax). From the spatial analysis, we conclude that the border regions and in general the urban areas of Switzerland respond more strongly to price changes than more rural or remote regions. One explanation for this is that public transport is more readily available in urban areas and therefore serves as a substitute.

The goal of Essay 3 is to assess the estimation accuracy in terms of bias, variance and root mean squared error (RMSE) of the FD-GMM estimator and of the SYS-GMM estimator when an exogenous regressor exhibits a low within variation. For this reason, a Monte Carlo experiment is carried out. We vary the number of cross-sectional units N, the number of observations per unit T, the coefficient of the lagged dependent variable and the within variation of the exogenous regressor over a parameter range which is usually of interest or which can be met in applied empirical studies. Each experiment is replicated 1,000 times. For several parameter combinations, the bias of estimated coefficients is lower for FD-GMM. However, if the variance is preferred as the decision parameter to discriminate between the two estimators, SYS-GMM should be preferred in almost all situations. As a reasonable compromise, we use the root mean squared error, which combines variance and bias of the estimates in one measure. Using the RMSE of the estimates, the SYS-GMM estimator should generally be preferred over the FD-GMM estimator when the within variation of the exogenous regressor is low. For instance, for the situation where a small panel is used (e.g. $N = 50$ and $T = 5$), the RMSE of the short-run effects are on average almost 20% lower for the SYS-GMM estimator. For situations with a relatively low within variation compared to the between variation e.g. of only 10%, the RMSE of the SYS-GMM estimates are as much as 67% below the FD-GMM estimates. Interestingly, for a small panel with low within variation of the exogenous regressor, results in terms of bias of the short-run effect are significantly in favour of the FD-GMM estimator. However, the efficiency gain in terms of a lower variance is strongly in favour of the SYS-GMM estimator, which compensates its relatively higher bias. Similar findings can be reported for the long-run effects and the coefficient of the lagged dependent variable itself. All experiments were carried out with either only a subset of instruments or the full set of instruments used. Further, a decision matrix is created with which the researcher can decide either to use FD-GMM or SYS-GMM, dependent on how bias is weighted against variance and dependent on the within variation of the exogenous regressor, on the number of observations N and T and on the supposed persistence of the dependent variable, γ.

Zusammenfassung

Diese Doktorarbeit besteht aus vier Teilen. Der erste Teil entspricht den Vorgaben für Doktorarbeiten des Departements für Management, Technologie und Ökonomie (D-MTEC) und stellt das Einführungskapitel dar. Die restlichen Teile sind für die drei Aufsätze reserviert, welche die eigentliche Dissertation bilden.

Das Ziel des ersten Aufsatzes ist die Schätzung einer Nachfragefunktion für Benzin in den Schweizer Grenzregionen und gleichzeitig die Quantifizierung der Benzinmenge, welche an Ausländer verkauft wurde. Der Anteil an Benzin, welcher an ausländische Fahrzeughalter verkauft wird, wird oft als "Tanktourismus" bezeichnet und wird durch die Preisdifferenzen über die Grenze getrieben. Daher wird das Preisverhältnis als erklärende Variable in einem ökonometrischen Modell benutzt und wird zudem gewichtet mit der Distanz zur Grenze, um aus dem Modell dann die Reichweite zu bestimmen, in welcher Tanktourismus eine Rolle spielt. Zu diesem Zweck wurden Verkaufsdaten von Tankstellen von den fünf grössten Benzinverteiler, welche in der Schweiz operieren, erhoben. Aus diesen Verkaufsdaten wurde dann ein Durchschnittswert der Absätze pro Tankstelle in einer Gemeinde gebildet, welche dann als Absätze einer "Referenztankstelle" der jeweiligen Gemeinde dient. Ein statisches Paneldaten Modell für fixe und zufällige Effekte wurde geschätzt. Wir haben eine Elastizität der Schweizerischen Benzinnachfrage in Bezug auf den Schweizer Preis von -0.211 gefunden, solange der Tanktourismus nicht berücksichtigt wird. Die durchschnittliche Preiselastizität unter Berücksichtigung des Tanktourismus beträgt allerdings -0.65. Somit ist die Nachfrage nach Schweizer Benzin in den Grenzregionen beträchtlich durch die Nachfrage der ausländischen Fahrzeughalter beeinflusst. Die Resultate zeigen ferner, dass der Tanktourismus ungefähr innerhalb einer Reichweite von 30 Kilometern der Grenze eine Rolle spielt, der Hauptanteil dieser Verkäufe an ausländische Fahrzeughaltern stellen jedoch Tankstellen, die maximal 10 bis 15 Kilometer zur Grenze entfernt sind. Das Modell zeigt ausserdem, dass je nach Jahr 250 bis 450 Millionen Liter der totalen Benzinabsätze von 4.5 Milliarden Litern in der Schweiz dem Tanktourismus zugeschrieben werden müssen. Ausserdem wird gezeigt, dass auch nur eine kleine Erhöhung des Schweizer Benzinpreises den Tanktourismus empfindlich einbrechen lassen könnte.

Das Ziel des zweiten Aufsatzes ist die Schätzung einer Nachfragefunktion nach Schweizer Benzin auf Gemeindeebene unter Berücksichtigung von räumlicher Korrelation. Die Nachfrage nach Benzin kann auf jeden Fall als ein räumliches Phänomen verstanden werden, da die Nachfrage nach Benzin in einer Gemeinde nicht nur durch die Bevölkerung oder durch den Fahrzeugbestand in jener Gemeinde beeinflusst wird, sondern auch durch den Verkehr zwischen

den Gemeinden. Wir benutzen somit ein Paneldaten Modell mit räumlich korrelierten Residuen und einer räumlich korrelierten abhängigen Variabeln. Wir schätzen einen Koeffizienten für die räumlich korrelierte abhängige Variable von 0.34 und einen Koeffizienten für die räumlich korrelierten Residuen von 0.37. Das bedeutet, dass eine Erhöhung der Benzinnachfrage um 10.0% in einer Gemeinde zu einer 3.4%igen Erhöhung der Benzinnachfrage in benachbarten Gemeinden führt. Wir schätzen zudem eine durchschnittliche Preiselastizität der Benzinnachfrage von -0.655 (als totalen Effekt). Die räumliche Aufteilung dieses Wertes führt zu einem durchschnittlichen direkten Effekt von -0.58. Je nach geographischer Lage einer Gemeinde, können sich die berechneten Elastizitäten von -0.585 und -0.855 bewegen. Aufgrund der sehr unterschiedlichen Schätzmethoden finden wir eine Preiselastizität im Bezug auf den ausländischen Benzinpreis von 0.32 (als den durchschnittlichen totalen Effekt), welcher sich signifikant von jenem unterscheidet, der durch Baranzini et al. (2012) geschätzt wurde, aber nicht von jenem, welchen wir im ersten Aufsatz berechnet haben. Da die Nachfrage nach Benzin in einer bestimmten Gemeinde auch die Nachfrage in den benachbarten Gemeinden beeinflusst, können die erhaltenen Resultate auch energiepolitisch wichtige Konsequenzen haben: Zum Beispiel reagieren nicht alle Regionen gleich sensitiv auf Änderungen im Benzinpreis (zum Beispiel über die Einführung einer CO_2-Steuer). Aus der räumlichen Analyse können wir schliessen, dass Grenzregionen und generell städtische Gebiete sensitiver auf den Benzinpreis reagieren als Gebiete in eher abgelegenen oder alpinen Regionen. Eine mögliche Erklärung für diese Beobachtung ist, dass in städtischen Regionen unter anderem eine höhere Verfügbarkeit von öffentlichen Verkehrsmitteln und somit eine Substitutionsmöglichkeit zum privaten Verkehr besteht, was in abgelegenen Regionen häufig nicht der Fall ist.

Das Ziel des dritten Aufsatzes ist es, den Einfluss von exogenen Variablen mit kleinen Varianzen zwischen den Gruppen auf die Schätzgenauigkeit des Arellano-Bond Schätzers (in der Folge FD-GMM genannt) und des Blundell-Bond Schätzers (SYS-GMM) zu erklären. Ein oft vorkommendes Problem in empirischen Studien ist die Tatsache, dass Koeffizienten von Variablen mit sehr kleinen Varianzen innerhalb der Gruppen nur sehr ungenau geschätzt werden können. Um jedoch Vorteile wie das Berücksichtigen von individueller Heterogenität mit Paneldaten Schätzern voll auszunutzen, müssen die Variablen im Modell entsprechend transformiert werden. Dies wird je nach Schätzer unterschiedlich erreicht: Bei statischen Modellen mit fixen Effekten werden die Daten transformiert, indem gruppenspezifische Mittelwerte subtrahiert werden. Bei dynamischen Paneldaten Modellen werden überlicherweise die ersten Differenzen gebildet (Differenzen von der aktuellen Periode zur vorherigen Periode). Solche Transformationen können einen beträchtlichen Teil der Varianzen zwischen den Gruppen eliminieren. Daher weisen Variablen, welche eine kleine Varianz innerhalb der Gruppen haben, nach der Transformation auch eine kleine Varianz auf. In der angewandten empirischen Forschung tritt dieser Effekt besonders häufig auf, wenn sozio-ökonomische Daten in den Modellen verwendet werden, was oft der fall ist. Typische Variablen mit einer kleinen Varianz innerhalb der Gruppen (in Bezug auf die Varianz zwischen den Gruppen) sind Einkommensdaten, Bevölkerungsdaten oder die Anzahl der Fahrzeuge in einer Gemeinde, um einige zu nennen. Clark und Linzer (2012) haben den Einfluss von kleinen Varianzen innerhalb der Gruppen der erklärenden Variablen auf statische Paneldaten Modelle mit fixen und zufälligen Effekten analysiert.

Dementsprechend führen wir ein Monte Carlo Experiment durch, um die Fragestellung für den Fall von dynamischen Paneldaten Schätzern zu erklären. Wir variieren die Anzahl der Gruppen N, die Anzahl Beobachtungen innerhalb der Gruppen T, den Koeffizienten der zeitlich verzögerten abhängigen Variablen und die Varianz innerhalb der Gruppen eines exogenen Regressors über einen Parameterbereich, der typischerweise in empirischen Studien anzutreffen ist. Jedes Experiment wurde 1'000 mal wiederholt.

Für einige der genannten Parameter Kombinationen ist die Abweichung der geschätzten Koeffizienten vom wahren Wert signifikant geringer für FD-GMM als für SYS-GMM. Wird jedoch die Varianz der Schätzer als Qualitätskriterium hinzugezogen, so sollte der SYS-GMM Schätzer in (fast) jedem Fall vorgezogen werden. Als guter Kompromiss gilt die Betrachtung der Wurzel des mittleren quadrierten Fehlers (RMSE = root mean squared error), da dieser beide Sichtweisen, jene der Abweichung zum wahren Wert und jene der Varianz der Schätzer, einbezieht. Benützt man den RMSE als Qualitätskriterium, so ist SYS-GMM generell dann vorzuziehen, wenn die Varianz des exogenen Regressors innerhalb der Gruppen klein ist. Für eine Situation mit einem kleinen Paneldatensatz (N = 50 und T = 5), der RMSE der kurzfristigen Effekte (geschätzte Koeffizienten des exogenen Regressors) ist im Mittel 20% kleiner für den SYS-GMM Schätzer. Wenn zusätzlich nun die Varianz innerhalb der Gruppen noch klein ist, so beträgt der Unterschied bis zu 67%. Interessanterweise ist die mittelere Abweichung der Koeffizienten zum wahren Wert für die genannten Situationen im Mittel kleiner für den FD-GMM Schätzer als für den SYS-GMM Schätzer. Die Varianz des SYS-GMM Schätzers ist aber viel geringer, so dass der SYS-GMM Schätzer auch mit dem RMSE Gütekriterium zu bevorzugen ist. Ähnliche Schlussfolgerungen können auch für die langristigen Effekte des exogenen Regressors und den Koeffizienten der zeitlich verzögerten Variablen gemacht werden.

Table of Contents

List of Figures

List of Tables

I Introduction

This dissertation is composed of three essays on the demand for Swiss gasoline and on the analysis of panel data.

The purpose of the first essay is to determine the quantity of gasoline consumed by foreign car owners due to the price differentials across the border. The second essay aims to quantify the determinants for the demand of Swiss gasoline more generally and to analyse spatial correlation in the demand of Swiss gasoline using panel data. Essay 3 deals with a frequently encountered econometric problem in the analysis of panel data. The goal of Essay 3 is to analyse the impact of low within variation in exogenous regressors on the estimation accuracy of the two most frequently used dynamic panel data estimators: the first-difference-GMM and the system-GMM estimators.

This introduction will shortly outline the general issues, the methodology and the contributions of the three essays.

1 Issues and Goals

1.1 Essay 1: Gasoline Tourism in Switzerland's Border Regions

The aim of this paper is to estimate the impact of the existing price differentials across the Swiss border on Swiss gasoline demand. The reasons for estimating the amount of gasoline sold to foreigners are manifold: First of all, people living close to the border will naturally be confronted with negative externalities such as increased traffic, congestion and pollution. On the other hand, gasoline stations located in the border regions have much higher gasoline sales than those located further from the border, and hence not only have higher revenues from the sale of gasoline but also from the sales of other goods. For a small country such as Switzerland, the share of gasoline sold to foreigners as a proportion of overall sales may be substantial. The information about this quantity will enable policy makers to weigh the advantages and the disadvantages of gasoline tourism.

The goal for the first essay, therefore, is to estimate the impact of gasoline tourism in Switzerland and to incorporate two important novelties or improvements in the study: First, the data base should be extended such that sales within a wider range of the border regions can be analysed, comparing to Banfi et al. (2005). Moreover, Swiss municipalities are quite heterogeneous, and therefore the data should reflect the situation at the municipal level. Second, the distance range within which gasoline tourism is thought to occur should be estimated from an empirical model and not be implicitly assumed.

1.2 Essay 2: Spatial Panel Data Econometrics Using GMM for Static Models

In recent years, spatial econometric models have gained a substantial increase in attention from empirical researchers. From an econometric perspective, data on the location of individuals can introduce important information to a model, since the units may affect each other. Neglecting spatial correlation in the dependent variable, the regressors or the residuals, where it is actually present, might have important consequences on coefficient estimates such as bias in the first two cases and inefficiency in the case of the residuals.

There exist user-written codes in Matlab® or R based on a maximum likelihood estimation procedure. Although Kapoor et al. (2007) derived a GMM estimator to jointly estimate the coefficients of a model with spatially correlated dependent variable and residuals, the GMM estimator has not yet been incorporated in commercial software packages such as STATA®, which is probably one reason why few studies have applied a spatial GMM estimator using panel data. There are some advantages in using GMM estimation procedures: First, no a-priori assumption on the residuals' distribution is needed. Second, the incorporation of (additional) endogenous variables is possible. Third, computational issues are less present. Therefore, one goal of Essay 2 is to incorporate a spatial panel data GMM estimator into STATA®. Second, the creation of an appropriate spatial weighting matrix may be cumbersome if there are locations with missing data, which in most applications will be the case. The implementation of a code which triangulates the locations with available data should solve this problem. Third, spatial dependence is an assumption that can be tested. A variety of spatial testing procedures have been proposed, by Kelejian and Prucha (2001), Baltagi et al. (2003), Baltagi et al. (2008) or Sen and Bera (2011).

A further goal of this study is to extend the data set used in Essay 1 to the whole Swiss territory and not only the border regions. This will enable us to estimate a gasoline demand function while accounting for unobserved heterogeneity and spatial correlation. The municipalities in the data set will be assigned a price elasticity dependent on their location in space. This in turn is important information to policy makers when it comes to e.g. the introduction of a spatially graduated CO_2 tax on gasoline.

1.3 Essay 3: System GMM and Difference GMM - The Impact of Low Within Variation

In order to exploit the advantages offered by panel data estimators such as unobserved heterogeneity, it is necessary either to first-difference the data or to subtract individual means. However, such a transformation might wipe out a dominant part of the between variation of the variables in question, and therefore variables with low within variation are left with a low variation in general and so cannot be precisely estimated. Typically, this is often the case for empirical studies dealing with socio-economic data such as population, income or the stock of vehicles, to mention a few.

A trade-off between bias and variance occurs when choosing the appropriate econometric estimator. Traditionally, empirical researchers are far more concerned about bias than variance. Clark and Linzer (2012) studied the accuracy of estimated coefficients obtained by fixed and random effects estimators. They point out that empirical researchers should not only be concerned with bias but also with variance of the estimates. Alternatively, a combined measure to discriminate between the two estimators could be the root mean squared error (RMSE)[1]. Accordingly, there is good reason to be concerned also with variance and not only with bias. Clark and Linzer (2012) conclude that, in several situations, the biased estimator with the lower variance should be preferred.

For dynamic panel data models, a study comparable to Clark and Linzer (2012) has not been carried out yet. Therefore, the goal of Essay 3 is to assess the estimation accuracy of the Arellano and Bond estimator (also known as the first-difference estimator or FD-GMM estimator) and the Blundell and Bond estimator (also known as the system estimator or SYS-GMM estimator) in the context of regressors with low within variation. An additional issue is the endogeneity of the lagged dependent variable. Therefore, the FD-GMM estimator does not make use of de-meaning the data unit-wise but of first-differencing the data. However, the previously described problem persists, namely that variables with a low-within variation then cannot be precisely estimated since first-differencing reduces the within variation substantially. The SYS-GMM estimator, however, makes use not only of the regression equation in differences but also of the equation in levels and instruments it with first-differenced instruments. This in turn requires a stationarity restriction on the individual effects. It has been pointed out by Blundell and Bond (1998) that the efficiency gain can be dramatic when the SYS-GMM estimator is used. However, they analysed a purely auto-regressive process. Our focus lies on the efficiency gain (and bias) when an exogenous regressor with low within variation is present.

1 For instance, consider a situation where actually a fixed effects model would be appropriate since the exogenous regressor and the individual effects are correlated but the within variation of this regressor is very low. With the true value of the coefficient fixed to -1.0, 1,000 Monte Carlo replications of the coefficients obtained by fixed effects average to -0.99 with a standard deviation of 0.71. The estimates range from -3.47 up to 1.51. For the case of the random effects estimates, the coefficients average to -0.53 with a standard deviation of 0.20, and therefore the estimator is clearly biased. On the other hand, it only ranges from -1.19 to -0.07. Therefore, pure knowledge about the unbiasedness of a coefficient is useless when the result obtained is counter-intuitive.

2 Methodology

2.1 Essay 1: Gasoline Tourism in Switzerland's Border Regions

Following Banfi et al. (2005), we explain the demand for gasoline using the household production theory which is outlined in Deaton and Muellbauer (1980). As a main result of the theory, a demand function for gasoline should be based on explanatory variables such as the gasoline price, household income, the stock of cars and demographic and spatial characteristics. Gasoline sales data were collected from the five biggest gasoline retailers operating in Switzerland. We do not have information about the total number of stations in a municipality. Accordingly, we decided to average gasoline sales of the gasoline stations which are included in our data set for each municipality. This average value should reflect the demand for gasoline at a representative gasoline station. We further include the price ratio of the foreign gasoline price to the Swiss gasoline price in our model. In addition, we weight this ratio with the municipality's distance from the border. This enables us to find a distance threshold at which the price ratio has no impact on the demand for Swiss gasoline.

The standard econometric theory which can be used to estimate static panel data models is well described in Greene (2003), Baltagi (2005) or in Cameron and Trivedi (2009), among others. In order to account for unobserved heterogeneity, we rely on a fixed and a random effects model for the estimation of the demand equation. It is well known that the fixed effects estimator produces consistent estimates of the coefficients. If the within variation of the variables is low, the variance of the estimates might be very high, see also Clark and Linzer (2012). The data used for the present study consists in many socio-economic variables with a low within variation. One way to discriminate between the two estimators is the application of a Hausman test. Even though a Hausman test might point to the use of a fixed effects estimator, Clark and Linzer (2012) advise not to be overly concerned purely with bias but also with the precision of the estimates. A further reason for advocating the use of a random effects model lies in the goal of the study. For the prediction of gasoline tourism from the empirical model, we need to equal the price ratio equal to unity. This is a situation which is not supported by the data. Random effects models are known to perform better in out-of-sample predictions, see e.g. Baltagi (2005). Therefore, a random effects model is used.

2.2 Essay 2: Spatial Panel Data Econometrics Using GMM for Static Models

To our best knowledge, at the time when the work on the second essay was started, no GMM estimator existed in STATA® to estimate spatial econometric panel data models. Therefore, one goal of Essay 2 is to implement GMM estimators for spatial panel data models into the software such that spatial panel data models with either spatially correlated residuals, a spatially correlated dependent variable, or both could be estimated. For accurate implementation, we used the elaborations by Kelejian and Prucha (1998), (2007) and recalculated the derivations of the respective moment conditions, first, to be sure about a correct implementation and, second, for a better understanding of the procedure from our part. Generally, spatial dependence is an assumption that can be tested. Therefore, we also implemented a variety specification tests into STATA using elaborations by Kelejian and Prucha (2001) and Baltagi et al. (2003), (2007), (2008), (2009).

Once the moment conditions are available, the coefficient of the spatially correlated residuals can be estimated and the regressors can be transformed using a Cochrane-Orcutt transformation. Alternatively, feasible generalized least squares (FGLS) can be applied to estimate the coefficient of the spatially lagged dependent variable and the elasticities of interest. The procedure is described in more detail in Kapoor et al. (2007). Most importantly, the omission of spatial correlation in the dependent variable when it is actually present would result in biased coefficients. The omission of spatially correlated residuals would result not necessarily in biased but in inefficient coefficient estimates. These tests allow us to find the effects which should be accounted for.

The application of spatial econometric models requires the proper definition of a spatial weighting matrix. We used a triangulation algorithm to obtain a proper weighting matrix but confirmed robustness of results with a second weighting matrix which takes the five closest neighbours of a municipality into account. Both matrices have been maximum row-sum normalized and decrease with increasing distance.

The goal of the application in Essay 2 is to estimate Swiss gasoline demand at the municipal level taking spatial effects into account. The main interest lies in the elasticity of Swiss gasoline demand with respect to the Swiss gasoline price. We decompose the elasticity spatially. This can impose important consequences when interpreting results, since not all panel units are assigned the same elasticity, as in traditional panel data models, but one which depends on the geographical location. This in turn can be important information for policy makers.

2.3 Essay 3: System GMM and Difference GMM - The Impact of Low Within Variation

In order to assess the impact of the within variation of the exogenous regressor and the accuracy of the coefficient estimates obtained by the FD-GMM estimator and the SYS-GMM estimator, a Monte Carlo experiment is carried out. For this reason, a data generating process (DGP) is defined as

$$y_{it} = \gamma \cdot y_{i,t-1} + \beta \cdot x_{it} + u_{it}$$
$$u_{it} = \mu_i + \varepsilon_{it}$$

where γ is the coefficient of the lagged dependent variable and β is the coefficient of the exogenous regressor which is suspected to have low within variation. We are primarily interested in assessing the estimation accuracy of FD-GMM and SYS-GMM of γ and β and a coefficient resulting from the two, often labelled as the long-run effect, $\beta/(1-\gamma)$. With "accuracy", we are addressing the bias, the variance and the root mean squared error of the estimates.

We proceed in an way analogous to that of Clark and Linzer (2012). We fix the between variation of x_{it} to 1 and vary the within variation of x_{it} over $s_{w,x_{it}}$ over 0.1, 0.2, 0.5, 1, 2, 5 and 10. Note that here we relate the standard deviation to the term variation. The overall variance of the error term u_{it} is fixed to 1. Therefore, we simulate situations where the within variation of the exogenous regressor is 10%, 20%, 50% and so on relative to the between variation of the regressor. γ is varied over 0.0, 0.2, 0.4, 0.6 and 0.8 – negative values actually are seldom encountered in empirical economic applications. The coefficient β is fixed to 1. Further, it is well known that the instrument matrix of FD-GMM (and SYS-GMM) tends to become very large when the observations per unit, T, increase. Therefore, we also vary the number of observations per unit T over T = 5, 10, 20 and the number of cross-sectional units over N = 50, 250, 500. Since both estimators are known to be sensitive to a large number of instruments, we restrict the depth of lags used as instruments to 2, a procedure which is often applied in empirical studies. In addition, we also run the simulations with all available instruments. Often, the data used in empirical studies show a small number of cross-sectional units N and observations per unit, T. Accordingly, we are primarily interested in evaluating the Monte Carlo experiment for those situations, e.g. N = 50 and T = 5 or T = 10. Each experiment is replicated 1,000 times. In total, we run 1,260,000 regressions in our Monte Carlo experiment (315 experiments, each replicated 1,000 times, once evaluated for FD-GMM and once for SYS-GMM, with either a subset of instruments used or with all instruments available). We used a fixed seed and ran the experiments in STATA®.

3 Abstracts

3.1 Essay 1: Gasoline Tourism in Switzerland's Border Regions

The goal of this study is to estimate a demand function for Swiss gasoline in Switzerland's border regions and simultaneously quantify the amount of gasoline sold to foreigners. The share of gasoline sold to foreigners often is labelled as "gasoline tourism" and is driven by the price differentials across the border. Therefore, the price ratio is used as an explanatory variable in an econometric model and is also weighted with the distance from the border to determine the distance range from the border within which gasoline tourism is supposed to have an impact on the amount of gasoline sold.

For this reason, gasoline sales data from the five largest gasoline retailers operating in Switzerland were collected. The sales from gasoline stations were averaged at the municipal level serving as the sales from a representative gasoline station for that municipality. A static panel data model accounting for fixed and random effects was estimated. We found a significant price elasticity of Swiss gasoline demand in the border regions with respect to the Swiss gasoline price of -0.211 when gasoline tourism was ruled out. The average price elasticity with gasoline tourism considered is -0.65. Accordingly, the demand for Swiss gasoline is considerably affected by foreigners purchasing gasoline in the Swiss border regions.

The results indicate that gasoline tourism had an impact on municipalities up to a distance of some 30 kilometres from the border, but the main part was sold at stations located within 15 kilometres. The model shows that, compared to overall sales, which accounted for some 4.5 billion litres of gasoline, gasoline tourism reached values from some 250 million litres up to 450 million litres of gasoline on average. Further, it is shown that even a small increase in the Swiss gasoline price may lead to a substantial decrease in gasoline tourism.

Keywords: gasoline, gasoline demand, gasoline tourism, cross-border purchasing, gasoline price differential, panel data estimation, price elasticity, interaction variables, out-of-sample prediction

JEL classification: C33, Q41, R22, R41

3.2 Essay 2: Spatial Panel Data Econometrics Using GMM for Static Models

The overall goal of Essay 2 is to estimate the demand for Swiss gasoline at the municipal level while taking spatial effects into account. The demand for gasoline can clearly be seen as a spatial story, since the consumption of gasoline not only depends on a municipality's car fleet or population but also on exchange traffic. We use a panel data model with spatially lagged residuals and a spatially lagged dependent variable and account for random effects.

We estimate a coefficient of the spatially lagged dependent variable of 0.34 and a coefficient of the spatially lagged residuals of 0.37. This implies that an increase in gasoline demand in one municipality by 10.0% spreads over to other municipalities and leads to an increase of 3.4% in demand, given that the regions are neighbours. As a main result, we estimate an average price elasticity of Swiss gasoline demand of -0.655 (total effect). Spatial partitioning of this value leads to a direct effect of -0.58 (on average). We estimate price elasticities ranging from -0.585 to -0.855, dependent on the municipalities' locations. Due to the very different approach, we find a price elasticity of Swiss gasoline demand with respect to the foreign price of 0.32 (as the average total effect), which is significantly different from that reported in Baranzini et al. (2012) but not from the value obtained in Essay 1.

As the demand for gasoline in one municipality affects demand in neighbouring municipalities, those results can impose important consequences for policy makers: First, there are regions which react more sensitively to change in the gasoline price than others (e.g., this change might result from the introduction of a CO_2 tax). From the spatial analysis, we conclude that the border regions and in general the urban areas of Switzerland respond more strongly to price changes than more rural or remote regions. One explanation for this is that public transport is more readily available in urban areas and therefore serves as a substitute.

Keywords: gasoline demand, spatial effects, spatial weights, spatial triangulation, spatial dependence tests

JEL classification: C33, Q41, R22, R41

3.3 Essay 3: System GMM and Difference GMM - The Impact of Low Within Variation

The goal of Essay 3 is to assess the estimation accuracy in terms of bias, variance and root mean squared error (RMSE) of the FD-GMM estimator and of the SYS-GMM estimator when an exogenous regressor exhibits a low within variation. For this reason, a Monte Carlo experiment is carried out. We vary the number of cross-sectional units N, the number of observations per unit T, the coefficient of the lagged dependent variable and the within variation of the exogenous regressor over a parameter range which is usually is interest or which can be met in applied empirical studies. Each experiment is replicated 1,000 times.

For several parameter combinations, the bias of estimated coefficients is lower for FD-GMM. However, if the variance is preferred as the decision parameter to discriminate between the two estimators, SYS-GMM should be preferred in almost all situations. As a reasonable compromise, we use the root mean squared error, which combines variance and bias of the estimates in one measure. Using the RMSE of the estimates, the SYS-GMM estimator should generally be preferred over the FD-GMM estimator when the within variation of the exogenous regressor is low. For instance, for the situation where a small panel is used (e.g. N = 50 and T = 5), the RMSE of the short-run effects are on average almost 20% lower for the SYS-GMM estimator. For situations with a relatively low within variation compared to the between variation, e.g. of only 10%, the RMSE of the SYS-GMM estimates are as much as 67% below the FD-GMM estimates. Interestingly, for a small panel with low within variation of the exogenous regressor, results in terms of bias of the short-run effect are significantly in favour of the FD-GMM estimator. However, the efficiency gain in terms of a lower variance is strongly in favour of the SYS-GMM estimator, which compensates its relatively higher bias. Similar findings can be reported for the long-run effects and the coefficient of the lagged dependent variable itself.

All experiments were carried out with either only a subset of instruments or the full set of instruments. Further, a decision matrix is created with which the researcher can decide either to use FD-GMM or SYS-GMM, depending on how bias is weighted against variance and dependent on the within variation of the exogenous regressor, on the number of observations N and T and on the supposed persistence of the dependent variable, γ .

Keywords: Arellano Bond estimator, Blundell Bond estimator, within variation, bias, efficiency, Monte-Carlo simulation

JEL classification: B23, C15, C53

4 Contributions

Essay 1 contributes in various ways to the literature on cross-border purchasing of goods in general and on gasoline demand in particular. First, compared to previous studies, a rich database for Switzerland was collected to estimate the amount of gasoline tourism. The data were collected at the municipal level, the smallest possible resolution at which Swiss socio-economic data exist. Second, we estimated the distance within which gasoline tourism is supposed to occur using an econometric model which incorporates the interaction of the price ratio across the border with the distance from the border. Third, our model enabled us to produce counter-factual simulations of the volume of gasoline sold to foreigners in a more precise way than previous studies.

Essay 2 applies spatial econometric techniques to estimate a demand function for Swiss gasoline. To our best knowledge, this is one of the first studies applying spatial econometric methods to the demand for energy. A further contribution is the implementation of a GMM estimator for spatial panel data models into STATA®. Accordingly, technical contributions are also made: First, a GMM estimator for the Kelejian-Prucha model and also a triangulation algorithm to obtain a proper spatial weighting matrix dependent on the available data was implemented into STATA®. Second, various kinds of specification tests are embedded in the code. Third, the estimates can be bootstrapped such that the sample distribution of the spatial autoregressive coefficient (the error lag) and the sample distributions of the total effects can be obtained to calculate the standard errors.

In essay 3, a Monte Carlo experiment is conducted to assess the accuracy of FD-GMM and SYS-GMM estimates in the context of exogenous regressors with low within variation. To our best knowledge, this is the first study assessing the econometric properties of dynamic panel data models in such a context. As a second contribution, we develop a decision matrix for the applied empirical researcher serving as rule of thumb for choosing when it could be wise to trade off variance against bias and vice versa when applying dynamic panel data estimators.

II Gasoline Tourism in Switzerland's Border Regions

1 Introduction

In many countries, taxes on fuels are a major component of their retail sales price. Both the value added tax (VAT) and general taxes on gasoline are lower in Switzerland than in the neighbouring countries. The price differential across the border has encouraged the phenomenon of gasoline tourism to increase over the past few years as gasoline prices in Italy, France and Germany grew much faster than the Swiss gasoline price. From 2001 to 2008, the nominal gasoline price in Austria was some 158.9 Swiss franc cents per litre, whereas the Swiss gasoline price in the region bordering Austria was only some 150 cents. The price differences from those of France, Germany and Italy are very substantial - on average, the prices are more than 30 Swiss franc cents above the Swiss gasoline price, representing a strong incentive to foreign car owners to purchase their gasoline on the Swiss side of the border.

The aim of this paper is to estimate the impact of the existing price differentials on local gasoline demand. For this purpose, we define the term 'gasoline tourism' as the quantity of Swiss gasoline purchased by foreign car holders, where the decision to do so is solely driven by the monetary benefit and hence by the price differential[2]. The reasons for estimating the magnitude of gasoline tourism are manifold. First of all, people living close to the border will naturally be confronted with negative externalities such as increased traffic, congestion and pollution. On the other hand, gasoline stations located in the border regions have much higher gasoline sales than those located further from the border, and hence not only have higher revenues from the sale of gasoline but also from the sales of other goods. Accordingly, the density of gasoline stations is higher in the border regions and thus jobs are also created. Another positive effect of gasoline tourism is that the state receives increased tax revenues proportional to the quantity of gasoline consumed. Therefore, the estimation of the quantity of gasoline sold to foreigners allows the monetary value of increased tax revenues to the state to be identified. For a small country such as Switzerland, the share of gasoline sold to foreigners as a proportion of overall sales may be substantial. The information about this quantity will enable policy makers to weigh the advantages and the disadvantages of gasoline tourism. Moreover,

[2] One might think of a situation where foreign car owners visit Swiss municipalities for other purposes than just fuelling their cars - e.g. for an excursion or to purchase other goods as food, alcohol or cigarettes and then contemporaneously fuel their cars in Switzerland without considering the respective price differential directly.

the impact of a change in the gasoline taxes on fuel tourism and therefore on tax revenues can then be estimated.

The impact of gasoline tourism in Switzerland is analysed by estimating a demand function for gasoline using panel data from Swiss municipalities in the border regions of Austria, Italy, France and Germany from 2001 to 2008. The goal is to develop a gasoline demand model which incorporates the distance from the border so that the impact of gasoline tourism not only depends on the existing price differential but also on a municipality's distance from the border. Further, the change in gasoline tourism will be quantified for a ceteris paribus change in the Swiss gasoline price.

The structure of the paper is as follows: Section 2 summarises the literature on gasoline demand and cross-border purchasing of different goods and services. Based on the insights gained from the literature review, a model is specified in section 3 which enables the estimation of both gasoline demand and gasoline tourism at the municipal level. In section 4, the data used to estimate the specified model is analysed and characterised accurately. In section 5, the focus is laid on estimation techniques which take account of the structure of the data and the goal of estimating gasoline tourism at the municipal level. In section 6, a brief summary and conclusion are presented.

2 Literature Review

The number of studies published on cross-border fuel demand is relatively small. The literature is abundant of studies on cross-border demand for cigarettes, alcoholic beverages or lottery tickets. Generally, these cross-border demands are induced by the existing tax differentials these goods between countries.

In order to gain an overview of the literature related to the topic addressed in this paper, we need to consider three types of studies. For this reason, for the first part of this section, we discuss studies focussing on gasoline demand on a general level. Then we turn to studies which examine cross-border purchasing in general, focussing on products such as alcohol or cigarettes. These studies are of interest to us from a methodological perspective. Lastly, we discuss the most relevant studies focussing on cross-border purchasing of fuels. Finally, a tabular summary will be provided.

For the review of the studies on gasoline demand, we decided to discuss three studies on Switzerland and two studies on European countries.

Wasserfallen and Güntensberger (1988) used time-series data from 1962-1985 to estimate the price and income elasticities of gasoline consumption in the short-run. They used a partial equilibrium model to explain the demand for gasoline and the total stock of motor vehicles in the economy. The explanatory variables, annual gasoline consumption, real gasoline prices, prices for public transport, the user costs of new cars and the stock of gasoline-powered vehicles were transformed to natural logarithms and a regression in first differences was performed using ARIMA models. They found a short-run price elasticity of -0.3 and an income elasticity of some 0.5 to 0.6 for the period examined. In summary, the study emphasises the strong effect of changes in gasoline prices on gasoline consumption and on the stock of gasoline-powered cars in the Swiss economy.

Schleiniger (1995) used an error correction model analysing Swiss gasoline demand over the period 1967-1994 and found a short-run price elasticity of -0.24. The results of the cointegrating regression suggest that there is no long-run price response.

Baranzini et al. (2012) also analysed long- and short-run price and income elasticities of Swiss gasoline demand. They used quarterly data from 1970 to 2008 for the whole of Switzerland and also employed time-series cointegrating techniques to estimate a log-linear demand function for gasoline demand per capita. The independent variables were the real Swiss gasoline price, the real per capita income, the stock of cars per capita and the real foreign gasoline price in areas close to the Swiss border. They found a long-run price elasticity of gasoline demand of -0.34 and one of -0.1 for the short-run. The long-run response of demand for Swiss gasoline with respect to changes in the foreign gasoline price was estimated as 0.1, meaning a 1% increase in the foreign gasoline price increases demand for Swiss gasoline by 0.1%, whereas the effect was only 0.07 (but significant) in the short-run. Although this study did not concentrate on the phenomenon of cross-border fuelling, it highlights the fact that demand for Swiss gasoline is affected by changes in the gasoline price in adjacent countries, thus providing strong evidence for the importance of gasoline tourism in Switzerland - and, most of all, in its border regions. Even if the calculated effect of the foreign price on Swiss gasoline consumption is weak, it shows clearly that sales at the national level are influenced by foreign price levels, which indicates a certain volume of cross-border gasoline purchasing in Switzerland. The short-run elasticity of gasoline demand with respect to income was estimated to be 0.67 and to be 0.03 (insignificant) in the long-run.

Leal et al. (2009) analysed the impact of differences in automotive fuel prices among neighbouring autonomous communities in Spain with the intention of discovering whether a change in the fuel (diesel) price in one community affected sales in another. They used monthly data from January 2001 to March 2007 to estimate monthly purchases of diesel in Aragon as a function of the diesel price in Aragon, the ratio of prices between Aragon and six adjacent communities (weighted by traffic density) including Madrid, and the number of vehicles registered in each community. They estimated diesel sales in Aragon using a log-linear model with a cointegration technique. As further explanatory variables, the price ratios to the adjacent municipalities and Madrid were included, as well as the number of vehicles registered. However, neither a distance term nor a measure for population density in the adjacent regions was used. They estimated an error correction model and found a price elasticity of -2.45, which is very high compared to similar studies. This rather high elasticity is a first indicator for the presence of cross-border purchasing activities, reflecting people's opportunity to choose where they fuel their cars. Although no quantification of this phenomenon of gasoline tourism is presented, it was shown, similar to the study by Baranzini et al. (2012) that the price of fuel in adjacent regions has an impact on fuel sales in the region of interest, given a certain proximity and differences in prices.

Pock (2010) analysed data from 14 European countries over the period 1990-2004 to estimate a dynamic model specification for gasoline demand. The study emphasises that many previous studies may suffer from a bias in estimated income and price elasticities of gasoline demand due to the omission of diesel-powered cars or non-distinction between gasoline- and diesel-powered cars. The diesel share of total passenger cars has been increasing for all countries considered in the study's sample period (including Switzerland). Gasoline consumption was modelled on the basis of an average vehicle's utilisation, its average fuel efficiency and the total stock of cars in use. A two-way error component model was employed to specify a dynamic demand equation for gasoline in which gasoline consumption per gasoline-powered car was used as the dependent variable. The number of gasoline- and diesel-powered cars per driver, real per capita income and the real gasoline price were used as regressors. Nine common dynamic panel estimators were applied to the panel data set. It was found that the standard within estimator and its bias-corrected version, LSDVc, yielded reasonable estimates in terms of a positive income elasticity and a negative price effect on gasoline consumption. However, the coefficient estimates were found to be somewhat lower in absolute terms, which is partially accredited to the omitted-variable bias in other studies (e.g. the stock of diesel-powered cars). One might criticise that in a partial-adjustment model as it was employed in this study, the capital stock (the stock of cars) should not be included. However, since we are going to estimate a static gasoline demand function, we will include both the stock of gasoline- and diesel-powered cars.

The following studies have been selected for the review on studies on cross-border demand for cigarettes and other goods.

Coats (1995) estimated the effect of state cigarette taxes on cross-border sales of cigarettes for 48 contiguous states of the United States and the district of Columbia, showing the extent to which inhabitants of the border regions realise arbitrage opportunities. A model was built to estimate cigarette sales as a function of taxes on cigarettes, distance from the border, pre-tax price and several demographic variables. Two different demand relationships were analysed - demand 1, in which the prices of cigarettes in a certain state change alone, and demand 2, in which the price of cigarettes in all states change together. Comparing the price elasticities obtained from demand 1 and 2, they concluded that about four fifths of the sales response to state cigarette taxes is due to cross-border sales.

Di Matteo and Di Matteo (1996) examined cross-border shopping behaviour for the seven Canadian provinces bordering the United States. Quarterly data on same-day car trips and expenditures from 1979 to 1992 were analysed using multiple regression techniques. They underline that across the provinces, cross-border trips and expenditures can be explained by per capita income, the appreciation of the Canadian dollar, the ratio of Canadian to US gasoline prices and the general tax on goods and services. The income elasticity of same-day automobile trips ranged from 0.89 to a relatively high value of 2.98. The high elasticities of the determining variables indicate that consumers are quite mobile and quick in taking advantages of arbitrage opportunities.

Asplund et al. (2007) analysed cross-border shopping for alcohol in Swedish municipalities in order to estimate the response of alcohol sales to foreign prices and relate sensitivity to a location’s distance from the border. Monthly data on sales of spirits, wine and beer over a ten-year period were collected from each store selling alcohol in Sweden. In addition, the domestic and foreign prices of spirits, wine and beer, the number of stores per capita and income per capita were collected. A linear log-log model was used to estimate the demand for alcohol per capita as a function of the domestic and the foreign alcohol prices and the other variables mentioned above. Analogously, a parametric third-order polynomial model, with distance from the border serving as a parameter, was specified and incorporated in the model as a weight of the foreign alcohol price. It is shown that the foreign alcohol price has a diminishing effect on sales of stores which are farther from the border. The elasticity of Swedish alcohol sales with respect to foreign prices tends to be almost zero after a certain threshold distance, which is a substantial contribution to the literature. A counterfactual simulation showed that Swedish tax revues fell by some 2.2%, and in areas located within 100 kilometres from the border by almost 7.5%, due to a cut in spirit taxes in Denmark which decreased prices there by almost 27%. However, no confidence interval was calculated for this estimation.

Finally, we review three studies on cross-border demand for gasoline: One study that considers cross-border effects of gasoline price differentials was authored by Rietveld et al. (2001). In this study, the adverse effects of fuel fetching trips such as extra miles driven, congestion, pollution and losses of sales are discussed with reference to Dutch car drivers living within a distance of 30 kilometres from the borders adjacent to Belgium and Germany. In a first step, a logit model was formulated to estimate the utility which Dutch people derive from fuelling in the Netherlands compared to fuelling abroad. Using the results from a revealed and stated fuelling behaviour survey, it was found that, given only a 5 Euro Cents price differential, 30% of Dutch people living at the border would fuel in Germany. It is also shown that there is a trade-off between price differential and extra kilometres driven, but it is mentioned that the monetary gains of the trip are overrated compared to the costs (e.g. travel time, extra use of fuel) of the trip.

For our purpose, though, the most interesting study was authored by Banfi et al. (2005). They analysed gasoline tourism in the Swiss border regions for the period of 1985 – 1997. A panel data model was estimated for three border regions, namely for those adjacent to Italy, France and Germany. During the period used in the study by Banfi et al. (2005), the gasoline price in Switzerland was constantly lower than in the neighbouring countries, with an exception for the first four years for Germany, where the real gasoline price was slightly below the real Swiss gasoline price. They used the price ratio between the foreign gasoline price and the Swiss gasoline price as the driver for gasoline tourism, arguing that there is no cross-border purchasing by foreigners if the prices are equal. From the methodological point of view, they explained the household demand for gasoline using the basic framework of household production theory and accordingly, they further collected socio-demographic variables like the population for the Swiss border regions and the adjacent foreign regions, the (real) per capita income in those regions, the number of commuters being the foreign persons coming to Switzerland to work and the stock of cars in the Swiss border regions. They used sales from approximately 190 Swiss fuel stations located within 5 km of the border, implicitly assuming that after 5 km from the border, no gasoline tourism takes place. The estimated elasticity with respect to the Swiss gasoline price was found to be -1.75, which is a relatively high value, taking into account that the meta-analysis of Brons et al. (2006) reports a mean value of some -0.3 to -0.45 for the price elasticity of gasoline. However, the area observed in this study is very close to the border, and a change in the Swiss gasoline price not only affects Swiss residents but also encourages foreign car owners to change their fuelling behaviour. They calculated gasoline tourism as a percentage of total gasoline demand in the border regions and computed an average for the period concerned of some 9% of total sales. They further indicated that an increase in the Swiss gasoline price (of some 20 Swiss franc cents) would have reduced this average to a fraction of only 3% of overall sales, clearly indicating that those revenues from gasoline tourism should not be seen as a stable revenue since they are driven by the price differentials across the borders, which in turn strongly depend on foreign and domestic fiscal policy. However, no confidence interval for the reported percents of gasoline tourism was calculated, which probably would have put the seemingly high value of average gasoline tourism into perspective. These are crucial differences compared to the present study. The study by Banfi et al. (2005) only has three regions (cantons) as the panel units, which in turn makes traditional panel data approaches inappropriate and hence raises the issue of using GLS for pooled time series and cross-sectional data, as suggested by Kmenta (1997) and Greene (2003). In contrast, our study will deal with 315 municipalities being the panel units and with a time frame of eight years. Accordingly, the econometric techniques used in these studies cannot be compared.

Another study, Michaelis (2004), focuses on the potential effects of price differentials between Germany and its neighbouring countries using a cost-benefit analysis instead of an econometric model. In this context, Michaelis compares the costs of the trip to collect fuel such as additional depreciation costs of the vehicle, additional fuel costs per kilometre driven or the opportunity costs with respect to time. Those costs are compared to the potential benefit the vehicle driver may earn, which is reflected by the fuel price differential between Germany and the other country. From this comparison, it is possible to derive a critical distance up to which it is beneficial for a rational driver to undertake a fuel collecting trip. Michaelis (2004) states that people underestimate or even neglect their opportunity costs consisting in additional mileage, depreciation of the vehicle value, additional maintenance expenditures, overall travel time, and increased accident risk. Given the existing price differentials between Germany and its adjacent countries, the distances within which fuel tourism is supposed to take place vary among countries and fuel type and hence range from some 10 kilometres up to almost 60 kilometres. For Switzerland, it is reported that regions located within 23 to 48 kilometres from the border may be affected from German cross-border fuelling – provided that consumers make reasonable decisions regarding their opportunity costs. This may be an indication that the distance band chosen by Banfi et al. (2005) is probably too small.

Table II-1 summarises the abovementioned studies – the data used, methods and functional form, main findings and especially what we can learn for the present study. Based on the studies reviewed on cross-border gasoline demand, a few shortcomings can be identified:

- Use of a (too) small data sample (Banfi et al. 2005)
- Disregard of diesel-powered cars or non-distinction between gasoline- and diesel-powered cars (Banfi et al. 2005)
- Oversimplifying assumptions about the distance within which border regions are affected by cross-border purchasing (Banfi et al. 2005)
- Effects of price changes often were not clearly reported, in particular no confidence intervals concerning the effects of cross-border purchasing were calculated (Michaelis 2004), (Banfi et al. 2005), (Asplund et al. 2007)

The consequences of the issues identified may be important. Due to the omission of variables such as the stock of vehicles, income or distance to the border, the resulting estimates for elasticities with respect to prices or income may be biased. Further, the assumption about the distance from the border at which regions are no longer affected by cross-border purchasing is crucial to the calculation of a possible loss of (tax) revenues or calculations concerning the effects of domestic price changes on the whole economy.

In this empirical analysis, we consider possible ways to solve, at least partially, the problems mentioned above. For instance, as we will discuss later, our modelling approach intends to consider the impact of distance to the border on gasoline demand by foreigners in a better way. Moreover, confidence intervals for gasoline tourism will be produced.

Author(s) of the Paper	Data Sample	Model and Functional Form	Method of Estimation or Calculation	Price Elasticity	Income Elasticity	Results	Distinctions
Waserfallen and Güntensperger (1988)	Yearly time series data (1962 - 1985)	Linear log-log specification dynamic	ARIMA-models	-0.3 to -0.45 (Swiss price)	0.7 (Swiss income)	Price and income elasticity of Swiss gasoline consumption and its interference with the OPEC shock (1973)	Simultaneous simulation of stock of motor vehicles and gasoline consumption
Schleiniger (1995)	Yearly time series data (1967-1994)	Linear log-log specification	Error correction model (cointegration)	-0.24 (Swiss price)	/	Price elasticity of Swiss gasoline consumption No significant long-run price response	/
Baranzini et al. (2009)	Quarterly time series data (1970 - 2008)	Linear log-log specification static and dynamic	Error correction model (cointegration)	-0.34 (Swiss price, long run) 0.1 (foreign price, long run) -0.1 (Swiss price, short run) 0.07 (foreign price, short run)	0.67 (Swiss income, short run) 0.03 (Swiss income, long run)	Price and income elasticity of Swiss gasoline consumption in the short- and in the long-run Very weak price response in the long-run	Usage of gasoline and diesel powered cars as explanatory variables Usage of foreign gasoline price Distinction between short- and long-run, distinction between 'fuels' (gasoline and diesel) and only gasoline
Leal et al. (2009)	Monthly data (2001 - 2008) Diesel prices, stock of cars, traffic density and diesel sales in Aragon (Esp) Explanatory variables from six adjacent communities	Linear log-log specification	Error correction model (cointegration)	-2.45 (Aragon price) 1.60 (Catalonia price) 0.6 (Madrid price)	/	Own price elasticity of diesel sales in Aragon, strong response of diesel sales in Aragon with respect to price changes in adjacent communities (e.g. Catalonia and Madrid)	Usage of price ratios of 'foreign' gasoline price to own gasoline price to explain diesel sales in Aragon, weighted with traffic density Inclusion of a time trend
Pock (2010)	Yearly data (1990 - 2004) 14 European countries	Linear log-log specification dynamic	LSDVc (bias corrected version for individual fixed effects estimation)	-0.106	0.075	Short- and long-run price elasticities of gasoline demand Quantifacation of bias due to omission of diesel cars or non-distinction between diesel and gasoline powered cars	Distinction between diesel and gasoline powered cars
Coats (1995)	Cigarette sales, taxes, demographic charcteristics and time trend for 48 contiguous US-states and for the district of Columbia, yearly data (1964 - 1986)	Linear specification static	Durbin-Aitken estimation (adjustment for serial correlation in pooled OLS)	-0.806 (price changes in respective state only) -0.167 (prices change in all states together)	/	Response estimates to cigarette price changes. Four-fifths of the sales response to state cigarette taxes is due to cross-border sales	/
Di Matteo and Di Matteo (1996)	Quarterly data (1979 - 1992) on same day automobile trips from 7 Canadian provinces bordering the US to the United States	Multiple regression, linear log-log specification	OLS	-0.88 to -3.69 (real exchange rate)	0.89 - 2.98 (depending on province)	Cross-border shopping exhibits important differences across the country, 90% of variation explained through exchange rate, tax differences of goods and servises between Canada and USA, gasoline prices, income, seasonality factors and distance from the border	Usage of gasoline prices, along with cigarettes and alcohol being an important driver for cross-border purchasing Usage of distance from the border Price ratio of typical consumer basket between Canada and USA used (through exhcange rate and taxes)
Rietveld et al. (2001)	Survey on fuelling behaviour of Dutch people living close to the German border (30km), data on vehicle kilometres driven, tax receipts, sales of gas stations	Linear specification of derived utility from fuelling abroad	Logit-model	/	/	Given a 5-10 cents price differential, some 30% of people living at the border decide to fuel in Germany. Trade-off between price difference and extra kilometers driven equal to -0.93	Usage of distance as a 'disutility'-component in the utility function Size of border region (30km) intuitively defined
Asplund et al. (2007)	Monthly data (1995 - 2004) for sales and prices of spirits, beer and wine, per capita income, distance from the border (mostly to Denmark) etc. for 288 Swedish municipalities	Linear log-log specification	SUR estimation (beer, wine and spirits)	-0.24 to -1.2 (Swedish price) 0.17 to 0.24 (Danish price)	0.65 - 1.38	Elasticities for Swedish alcohol consumption depending on distance from the border. A cut in Danish taxes by 27% led to a decline in Swedish tax revenues from spirits by some 2.2%, in areas located within 100km from the border even by 7.5%	Foreign alcohol prices (Danish prices) weighted with a distance term Critical distance for cross-border purchasing endogenously calculated Confidence intervals reported
Banfi et al. (2005)	Sales of 190 gasoline stations 1985-1997 (yearly data) 3 border regions	Linear log-log specification Static	GLS for pooled time series Corrections for country-specific heteroscedasticity	-1.75 (Swiss price)	0.47 - 1.75 (Swiss income)	9% of aggregated sales in Swiss border regions is gasoline tourism Gasoline Tourism is very sensitive to domestic price changes	Border regions' size exogenously defined: d=5km Ratio of foreign to Swiss gasoline price weighted with population ratio
Michaelis (2003)	Gasoline and diesel prices of Germany and adjacent countries for 2003 Data on vehicle characteristics (4 types for each fuel)	Cost-Benefit analysis	/	/	/	Distance bands for regions affected by German gasoline and diesel tourism Minimum 9.7 km Maximum 60 km	All types of travel costs considered Distinction between bounded and unbounded rationality

Table II-1: Summary of findings of literature review

3 Model Specification

Following Banfi et al. (2005), we explain household demand for gasoline using the framework of household production theory. A detailed explanation can be found in Deaton and Muellbauer (1980)[3].

If we assume that the household combines purchased market goods and its time to produce the commodity providing utility, we write

$$U = U(S(G, \overline{CS}, T), X, D, R) \tag{II.1}$$

with S being the transport service and G denoting the quantity of gasoline consumed. $\overline{CS}$ is the (fix) capital stock (stock of cars), T is time, X is a composite good representing a commodity basket that the household consumes at unit price, while D and R represent demographic and geographic characteristics which in turn determine the household's preferences. Following Deaton and Muellbauer (1980), the household's decision process can be modelled as a two-stage optimization problem. In the first stage, the household minimises its variable costs accruing from the production of any arbitrary amount of S, say $\hat{S}$, which can be formulated as

$$\underset{G,C}{\text{Min}} \ (P_G \cdot G + P_{CS} \cdot \overline{CS})$$

subject to (II.2)

$$S = \hat{S}(G, \overline{CS})$$

The result of this optimization problem is the variable cost function

[3] We are well aware of the fact that there are different possibilities besides this to explain gasoline demand at the household level. For instance, Baltagi and Griffin (1984) specified individual gasoline demand to be the product of number of kilometres driven per car and the gasoline consumption of the average car per kilometre driven (i.e. efficiency) times the total number of cars. Accordingly, the three main factors then are the degree of car utilization, the efficiency of the car stock and the absolute level of the car stock.

$$VC = VC(P_G, \overline{CS}, \hat{S}) \tag{II.3}$$

It is homogenous in $\hat{S}$ and in factor prices, increasing in $\hat{S}$ and non-decreasing and concave in factor prices. From this cost function, the input demand function can be derived by differentiating equation (II.3) with respect to the factor price P_G.

In the second stage of the optimization problem, the household maximises its utility derived from the consumption of S and X, which is

$$\begin{gathered} \operatorname*{Max}_{S,X} \; U(S, X, D, G) \\ \text{subject to} \\ VC(P_G, \overline{CS}, S) + X \leq Y \end{gathered} \tag{II.4}$$

where Y denotes the household's income. The solution of (II.4) results in demand functions for the commodities S and X.

$$\begin{aligned} S^* &= S^*(P_G, \overline{CS}, Y, D, G) \\ X^* &= X^*(P_G, \overline{CS}, Y, D, G) \\ G^* &= E^*(P_G, \overline{CS}, S^*) = G^*(P_G, \overline{CS}, Y, D, G) \end{aligned} \tag{II.5}$$

which depends on the gasoline price P_G, the household's income Y, the stock of cars $\overline{CS}$, and the demographic D and spatial (geographic) characteristics R. The model specification stated by equation (II.5) assumes that the stock of cars is constant. Therefore, the model represents a short-run gasoline demand specification. As discussed in Baltagi and Griffin (1984), the results obtained from the estimation of a short-run gasoline demand specification such as specified by equation (II.5) with cross-sectional or panel data rather reflect long-run price and income elasticities. In this study, we seek to estimate a gasoline demand function using aggregate data at the municipal level. Unfortunately, we do not have the information on the sales of gasoline for all gasoline stations located in a municipality, neither do we have information about the total number of stations in a municipality. Accordingly, our data set includes the sales of a sample of gasoline stations per municipality, so we decided to calculate the average gasoline sales of the gasoline stations which are included in our data set for each municipality. This average value should represent the gasoline demand for a representative gasoline station. Thus, gasoline demand will of course depend on the Swiss gasoline price and the foreign gasoline price, on different socio-economic variables and on each municipality's distance from the border. For the estimation of the gasoline demand model, we collected data on gasoline station sales from the five biggest gasoline retailers in Switzerland for the period 2001-2008.

We distinguish among four border regions, being those adjacent to Italy, Austria, Germany and France. In contrast to the study by Banfi et al. (2005), we do not only consider gasoline stations located within five kilometres from the border, but all stations in the cantons[4]. Based on the studies mentioned above and on availability of the data, we specify our model as

$$G_{it} = f(PG_{CH,bt}, PG_{F,bt}, dist_i, CARSG_{CH,it}, CARSD_{CH,it}, POP_{CH,it}, Y_{CH,it}, Comm_{CH,it}, c_j) \quad \text{(II.6)}$$

where the index i refers to the municipalities ($i=1..315$), t is an index for time ($t=1..8$) and b refers to the border regions $b=1..4$, indicating whether the respective municipality belongs to a canton adjacent to France, Germany, Italy or Austria. The index j refers to the cantons ($j=1..12$). The index CH means that the variable in question refers to Switzerland, whereas the index F means that the variable refers to the adjacent country. G_{it} denotes average gasoline demand per gasoline station in a municipality, for the reasons previously mentioned.

Foreigners living close to the Swiss border encounter both their own gasoline price (PG_F) and the Swiss gasoline price (PG_{CH}). During the time period analysed, the Swiss gasoline price was constantly lower in all four border regions than in the respective adjacent countries, and consequently, it is reasonable to assume that sales of gasoline stations located closely to the border can be partially explained by the difference in levels of the Swiss gasoline price and the foreign gasoline price. According to equation (II.6), prices only vary at the border region level (index b) and not at the municipal level, since data on prices were collected by the Swiss customs authorities and were not available from the gasoline companies.

We use the distance ($dist$) of the municipality from the border since the effect of cross-border fuelling, of course, is likely to diminish over a certain distance.

In addition, the gasoline demand of households will be affected by per capita income (Y_{CH}) in Switzerland and the population (POP_{CH}), perhaps by the population abroad, the number of daily commuters ($Comm_{CH}$) and, according to Pock (2010), the stock of diesel- and gasoline-powered cars ($CARSD_{CH}$ and $CARSG_{CH}$). However, the inclusion of variables concerning the foreign regions, such as the foreign population, the foreign stock of diesel- and gasoline-powered cars and foreign per capita income is problematic. Unlike Banfi et al. (2005), we intend to analyse gasoline tourism at the lowest possible level in terms of panel units for which data can systematically be collected (the municipalities). In the study by Banfi et al. (2005), those units were the Swiss cantons, but even at that level of scope, it is difficult to compare the cantons' population with the foreign one, since the foreign regions for which reasonable data are available are far larger than the average Swiss canton, and the same applies for the stock of cars. Further, unlike most of the studies discussed above, we want to identify the distance from the border at which the phenomenon of gasoline tourism is absent and accordingly, we will focus on the smallest units available, the municipalities. Consequently, we are forced to abstain from the inclusion of the foreign demographic variables since the variation across municipalities (the between variation) of those would be too small or non-existent due to the far larger size of the foreign regions.

4 Switzerland is a federal state. A canton is a federal unit and Switzerland consists in 26 cantons of different size and demography.

We are interested in capturing the effect of a domestic price change, showing to what extent gasoline is substituted with other goods and secondly, to measure the effect of a change in the ratio of the foreign gasoline price to the Swiss gasoline price, quantifying the substitution of cross-border gasoline with Swiss gasoline. In order to achieve this, we include the real (CPI adjusted) Swiss gasoline price and the ratio between the foreign gasoline price and the Swiss gasoline price in our model. It could be argued that potential foreign customers would rather consider the pure price differential than the ratio. However, Michaelis (2004) points out that these customers face opportunity costs, such as extra kilometres driven, which are directly proportional to the foreign gasoline price. Further, it is obvious that incentives are different if one can save, say, 20 Swiss franc cents per litre of gasoline at a price of 1.50 Chf, compared to a situation where the gasoline price would be 2.00 Chf per litre. Moreover, a large number of commuters come from Germany, Austria, France and Italy every day to work in Switzerland, and we believe they have a positive influence on Swiss gasoline sales. To account for canton-specific effects, we introduce a dummy variable for each canton. For the estimation of the gasoline demand function, we decided to use a log-log specification. Therefore, our general empirical model is specified in the following form:

$$\begin{aligned}\ln(G_{it}) = {} & \alpha_0 + \alpha_1 \ln(PG_{CH,bt}) + \alpha_2 \ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right) + \alpha_3 \ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right) \ln(dist_i) + \\ & \alpha_4 \ln(CarsG_{CH,it}) + \alpha_5 \ln(CarsD_{CH,it}) + \alpha_6 \ln(POP_{CH,it}) + \\ & \alpha_7 \ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right) + \alpha_8 \ln(dist_i) + \alpha_9 \ln(Commu_{it}) + \sum_{c=1}^{11} \gamma_c g_c + \varepsilon_{it}\end{aligned} \tag{II.7}$$

Prices are treated exogenously, since both producers and consumers are assumed to be price takers. Further, gasoline can be considered as a homogenous good, and so we do not differentiate between the different gasoline companies. We expect that a ceteris paribus decrease in the Swiss gasoline price will lead to an increase in gasoline demand in the border regions. A decrease in the foreign gasoline price is expected to result in a decrease in Swiss gasoline demand, but at a smaller magnitude, since foreign car owners will demand less Swiss gasoline, but domestic residents will not respond to a change in the foreign gasoline price, at least as long that the Swiss gasoline price remains below the foreign one. From the model specified in equation (II.7), the price elasticity with respect to the Swiss gasoline price and the foreign gasoline price can be calculated as follows:

$$\varepsilon_{PG_{CH,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{CH,bt})} = \alpha_1 - \alpha_2 - \alpha_3 \ln(dist_i) \tag{II.8}$$

$$\varepsilon_{PG_{F,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{F,bt})} = \alpha_2 + \alpha_3 \ln(dist_i) \tag{II.9}$$

Similarly, an increase in the Swiss per capita income is expected to have a positive effect on Swiss gasoline demand, we obtain

$$\varepsilon_{Y_{CH,it}} = \frac{\partial \ln(G_{it})}{\partial \ln(Y_{CH,it})} = \alpha_7 \qquad \text{(II.10)}$$

From the indices notation, it can be seen that, for each municipality (i), a different domestic and foreign price elasticity is assigned, depending on its distance from the border, whereas the income elasticity of gasoline consumption does not vary across observation units. We expect a positive influence of the stock of gasoline passenger cars on gasoline consumption and the opposite for the stock of diesel-powered cars. Moreover, increasing population will increase demand for transport services and hence demand for gasoline, and in addition, it is likely that commuters have an important positive role in explaining the level of gasoline demand in the border regions, since they travel regularly to Switzerland and can take advantage of the price differential without experiencing additional opportunity costs. Further, it is likely that, other factors (like the population) being equal, with increasing distance from the border gasoline consumption may decrease, since the effect of gasoline tourism is supposed to vanish after a certain distance. Intuitively, this is the case when changes in the foreign gasoline price show no impact on Swiss gasoline consumption, meaning that the elasticity of Swiss gasoline consumption with respect to the foreign price is zero, which implies:

$$\varepsilon_{PG_{F,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{F,bt})} = \alpha_2 + \alpha_3 \ln(dist_i) = 0 \Leftrightarrow dist_{crit} = e^{-\frac{\alpha_2}{\alpha_3}} \qquad \text{(II.11)}$$

Of course, once the threshold of this critical distance is passed, the elasticity changes its sign, which is contra-intuitive, so a condition for judging the model specification to be appropriate, is that this cut be flat instead, so that the elasticities observed for municipalities after this distance are not significantly different from zero. On the other hand, we intend to report mean values for the domestic and foreign price elasticities for distance ranges and an average for the entire border regions, which will then allow for a more accurate statement concerning correct signs for the elasticities. The same problem also occurs in the study by Asplund et al. (2007). Due to the specification, it can be seen from equations (II.8) and (II.9) that the foreign price and the domestic price elasticity are closely linke, moreover, when the foreign price elasticity reaches a value of zero, the domestic price elasticity is equal to α_1, so we rely on a significant and negative sign for the coefficient of the Swiss gasoline price, or at least on one which is not significantly different from zero.

4 Data

In the first part of this section, we provide a descriptive analysis of the development of gasoline sales and gasoline prices in Switzerland and the in the adjacent countries. In the second part, we provide descriptive statistics of the variables and data used in the estimation of equation (II.7). These descriptive statistics are based on a reduced sample, since in fact some data were missing for some municipalities.

4.1 Gasoline Demand

We collected data on Swiss gasoline consumption from the Swiss Oil Association (Erdölvereinigung) for the five most important gasoline companies located in Switzerland. The data consist in yearly sales of gasoline stations for the period 2001 to 2008. These companies owned approximately 1500 out of a total of some 4140 gasoline stations in Switzerland in 2001. Moreover, these companies account for almost 60% of total Swiss gasoline sales in the period.

Figure II-1 shows the relative and absolute development of gasoline sales for Switzerland and the five companies from which we obtained gasoline station sales data[5]. It is obvious from the upper figure that aggregated gasoline demand in Switzerland (CH) decreased over the sample period by approximately 10%. From the lower figure, one can see that four gasoline companies experienced important decreases in sales, whereas one gasoline company more than doubled its sales. The company in question (B) may have had lower sales in terms of absolute values, but the relative growth with respect to 2001 was very large. Due to the inclusion of company B in the data sample, the development of the sample sales fits the development of aggregate Swiss sales (CH) surprisingly well. The gasoline stations of these five gasoline companies sell more than 55% of total Swiss gasoline sales and represent approximately 40% of all gasoline stations in Switzerland. Given these relatively high values, we are confident that the sample of stations available for this study forms a representative sample with which Swiss gasoline demand at the municipal level can be explained

---

[5] We do not distinguish between leaded and unleaded gasoline, but we can envisage that the share of unleaded gasoline is very large.

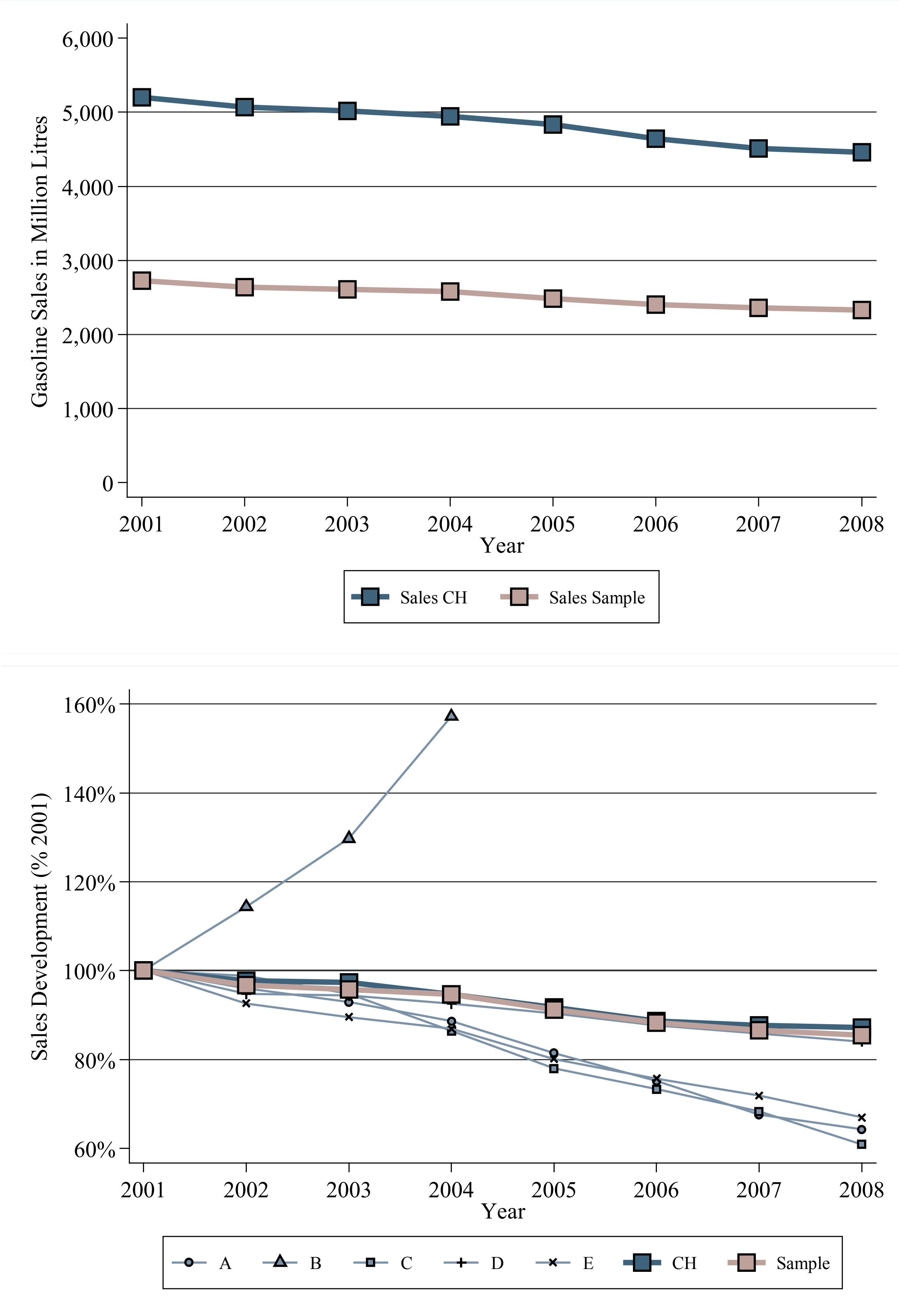
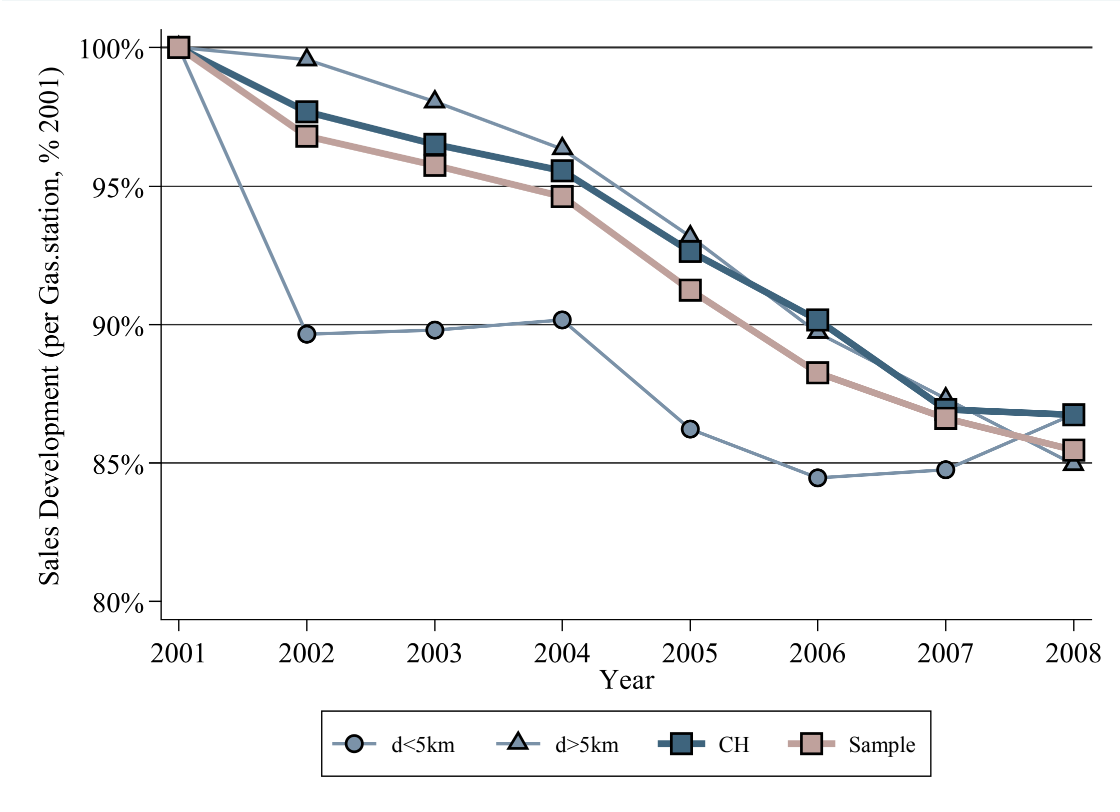


Figure II-1: Development of gasoline sales of the sample gasoline companies compared to overall Swiss sales

One should note that average gasoline sales per station for company B are nearly twice as large as average sales per station of the other companies. Further, the company exhibits a far smaller between and within variation of its sales. It is further characterised by larger gasoline stations which are typically equipped with large shops. For this reason, we adjust the model specified in equation (II.7) by adding a dummy variable for the presence of a gasoline station of company B in municipality i in year t: $D_{B,it}$ is equal to 1 if company B contributed to average gasoline sales of municipality i in year t and else equal to zero. The main effect of including the dummy for the gasoline company B is that the explanatory power of the model is increased and that several coefficients of explanatory variables of interest, such as those of the prices, increase in significance. Overall gasoline sales in Switzerland accounted for approximately 4.80 billion litres on average, of which some 2.60 billion litres per year (55%) are caught by the sample.

Further, gasoline sales developed quite differently close to the border compared to those more distant from the border. As discussed above, and as stated in equations (II.8) and (II.9), we expect gasoline stations which are close to the border to respond more intensely to changes in the Swiss gasoline price than stations far from the border. In order to visualise this, we choose a distance threshold of 5 kilometres (following Banfi et al. 2005) and compare the development of sales of gasoline stations within this band with the sales of stations out of this band. From Figure II-1 one can see that average sales per gasoline stations within 5 kilometres from the border are larger than those from more distant stations, despite the fact that the gasoline station density exceeds that one of areas located further from the border. This is another clue that gasoline tourism may play an important role in explaining gasoline demand in the Swiss border regions.

As already discussed, we distinguish between the four border regions adjacent to Austria, France, Germany and Italy. The cantons bordering these regions are quite heterogeneous in terms of culture, topology, border length, population density and so on. For instance, Switzerland shares a border length with Germany of some 292 kilometres (not including lakes), stretching over six cantons Baselland (BL), Baselstadt (BS), Thurgau (TG), Schaffhausen (SH) and Zürich (ZH). Nine cantons border France over a distance of some 480 kilometres: Baselland (BL), Baselstadt (BS), Bern (Be), Genf (GE), Jura (JU), Neuenburg (NE), Solothurn (SO), Waadt (VD) and Wallis (VS). We do not consider the cantons of Bern, Wallis and Solothurn in our sample. Wallis is hardly accessible due to topological conditions. Bern shares a very small border with France and, accordingly, people are far more likely to fuel in either Neuenburg or Jura. The canton of Solothurn is also somewhat unusual: it has a border with France, since it owns two municipality enclaves in the canton of Baselland. Accordingly, people are more likely to fuel in Basel than in Solothurn. Three cantons, Graubünden (GR), Wallis (VS) and Tessin (TI) border Italy over a length of some 675 kilometres. Here too, for topological reasons, we exclude Graubünden and Wallis in our sample. Further, the gasoline station density is much higher in Tessin than in Graubünden or Wallis. The canton St. Gallen (SG) borders Austria over a length of some 35 kilometres.



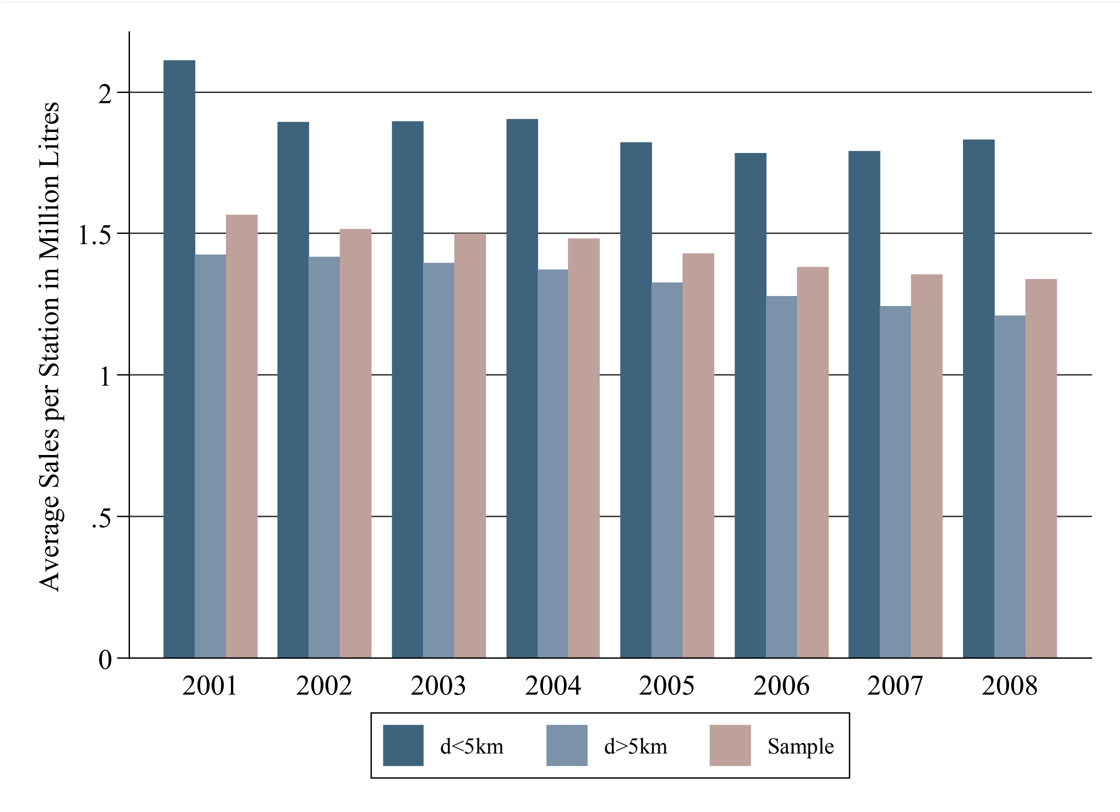


Figure II-2: Development of gasoline sales by distance from the border

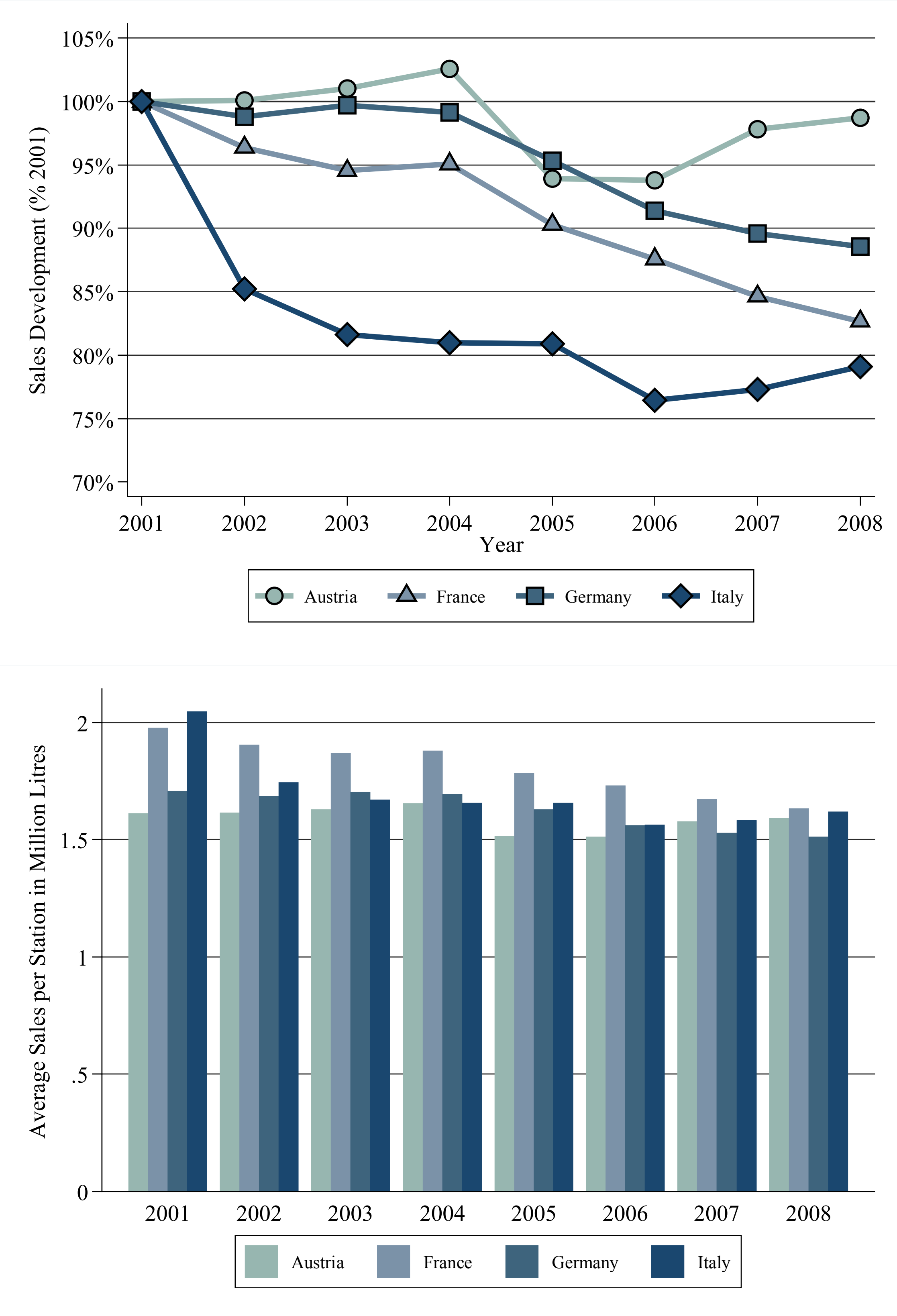


Figure II-3: Development of gasoline sales by border region

Figure II-3 shows that gasoline sales in cantons bordering Austria, France, Germany and Italy developed quite differently. Aggregate sales in cantons bordering Italy and France decreased strongly over the sample period, whereas aggregate gasoline sales in municipalities bordering Austria remained almost constant over the sample period. On average, gasoline stations located in cantons bordering France had the highest sales per station, followed by those located close to Italy, Germany and Austria. To provide a better overview, the upper figure shows the relative changes in aggregated sales over time. However, the lower figure is also of interest since it refers to the dependent variable (average gasoline sales per station) of this study.

4.2 Gasoline Prices

Data on gasoline prices were collected by the Swiss customs authorities and are available on a monthly basis. The border officers track prices for each border region in Switzerland and on each side of the four borders. On one hand, we are aware that gasoline prices are slightly higher close to the border than more distant from the border. On the other hand, we are interested in capturing the relevant price differentials (or price ratios) in order to capture the driver for gasoline tourism. Potential cross-border purchasers are presumed to live rather close to the border (according to Rietveld et al. 2001 or Asplund et al. 2007) and will compare their prices with the Swiss gasoline price. Assumed that they have no other intention than fuelling, they will consider Swiss stations located close to the border and not those very far away. Accordingly, we have price data varying over eight years for the four border regions and a corresponding foreign gasoline price. for each border region.

Figure II-4 shows the development of the nominal Swiss gasoline price and the nominal foreign gasoline price in Swiss franc cents per litre over the sample period. The third line represents the price differential between Switzerland and the foreign country. It can be seen that the price differential between Austria was the smallest and amounted to an average of approximately 7.8 Swiss franc cents per litre. The price difference to France was some 26 cents per litre and the difference to Germany almost 36 cents per litre. The highest price differential existed across the border to Italy and amounted to 38.4 cents per litre on average. The mean value of the nominal Swiss gasoline price was around 150.3 cents per litre, whereas the nominal foreign gasoline price averaged at 177.5 cents per litre. The price differential was rather high: it amounted to more than 27 cents, which allowed potential gasoline tourists to save more than 15% of their gasoline expenses, if opportunity costs are not considered. Given the development of gasoline prices depicted in Figure II-4, the potential demand for Swiss gasoline is expected to be highest at gasoline stations located in municipalities bordering Italy, and the lowest at gasoline stations in municipalities bordering Austria.

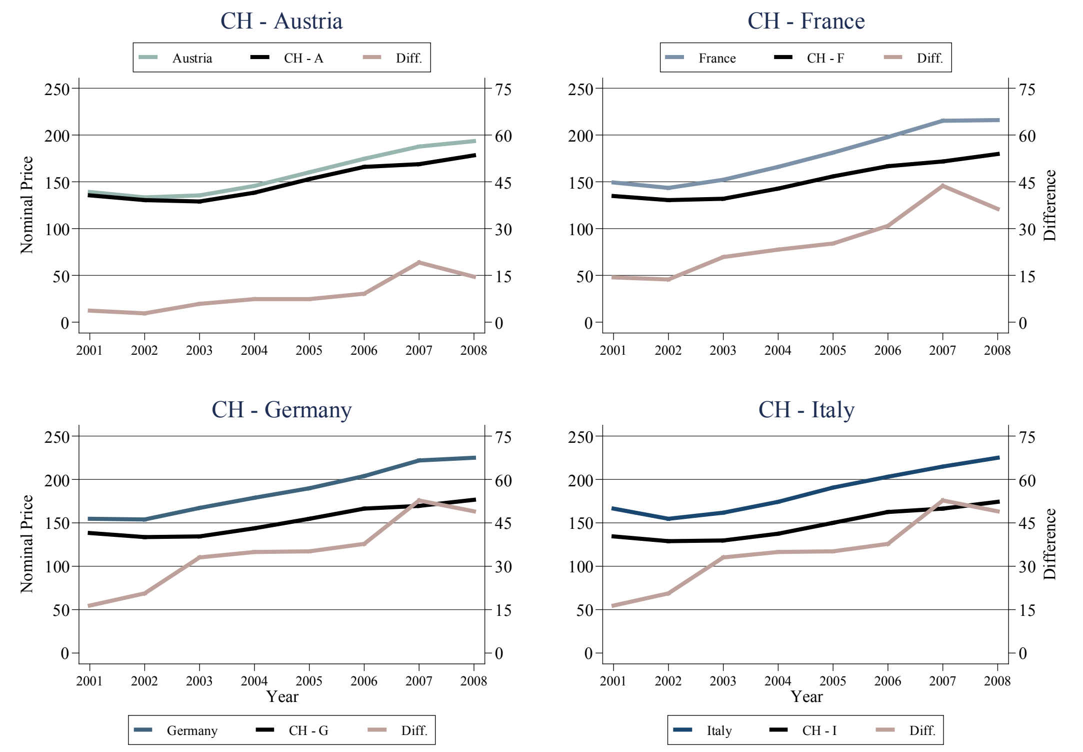


Figure II-4: Development of nominal gasoline prices in the border regions

In the cases of Germany and France, the price differentials also increased rather dramatically over the sample period, but not as much as it was the case for Austria. Therefore, the potential increase in gasoline tourism could also offset the decrease in domestic demand, partially, but not as strongly as in the case of Austria. For municipalities adjacent to Italy, the price differential has been always quite high over the sample period. For the estimation of the econometric model, we will concentrate on price ratios and not differentials; however, their evolvement is very similar to that of the differentials and the correlation coefficient is 0.96, as can be seen in Figure II-5.

The rose line (secondary axes) shows the development of the price differentials, and the blue line (primary axis) shows the development of the price ratio (foreign gasoline price to Swiss gasoline price). The border region adjacent to Italy could have the highest impact in sales from gasoline tourism. On average, the price ratio between Italy and Switzerland was the highest with a mean value of 1.26, which means that the gasoline price in Italy was 26% higher (on average) than the Swiss gasoline price in regions bordering Italy. There is no strong fluctuation, though, since both the Swiss gasoline price (adjacent to Italy) and the Italian gasoline price increased very similarly. The price ratio between Germany and the cantons bordering Germany was some 1.23 on average and experienced quite a large increase over the sample period. For France, this mean value was 1.17 and for Austria 1.05.

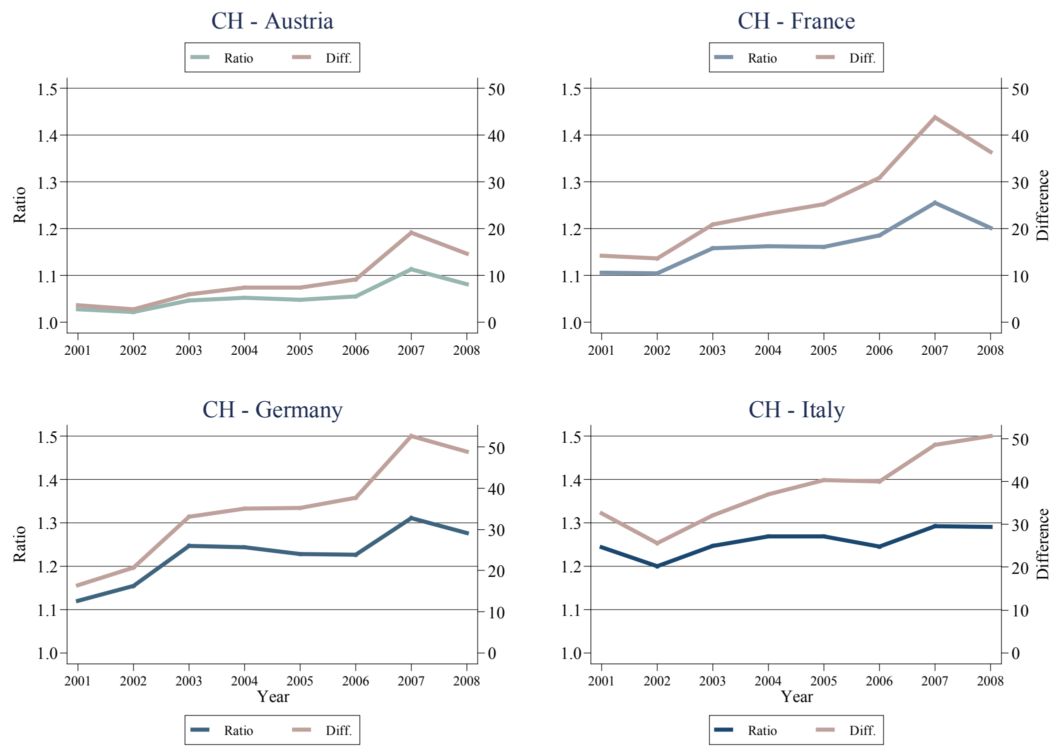


Figure II-5: Development of the nominal price difference and ratio in the border regions

4.3 Descriptive Statistics of the Variables Used for Estimation

As mentioned before, the econometric analysis only considers municipalities located in border cantons with complete information on yearly sales, prices and socio-economic variables. The sample consists of a balanced panel over eight years for 84 municipalities close to the French border, 168 close to the German border, 25 close to the Austrian border and 38 close to the Italian border, in total 315 municipalities. We note that the reduced sample's descriptive statistics concerning aggregate gasoline sales, gasoline prices and other socio-economic variables are in line with the previous results for the whole (unadjusted) sample. The descriptive statistics of the variable of interest for this reduced sample are provided below.

The development of the most relevant socio-economic data over the sample period in the four border regions is depicted in Figure II-6. We only consider municipalities for which we have gasoline sales data from the gasoline companies, so some municipalities are neglected in the present analysis of socio-economic data. All data is averaged to the municipality level.

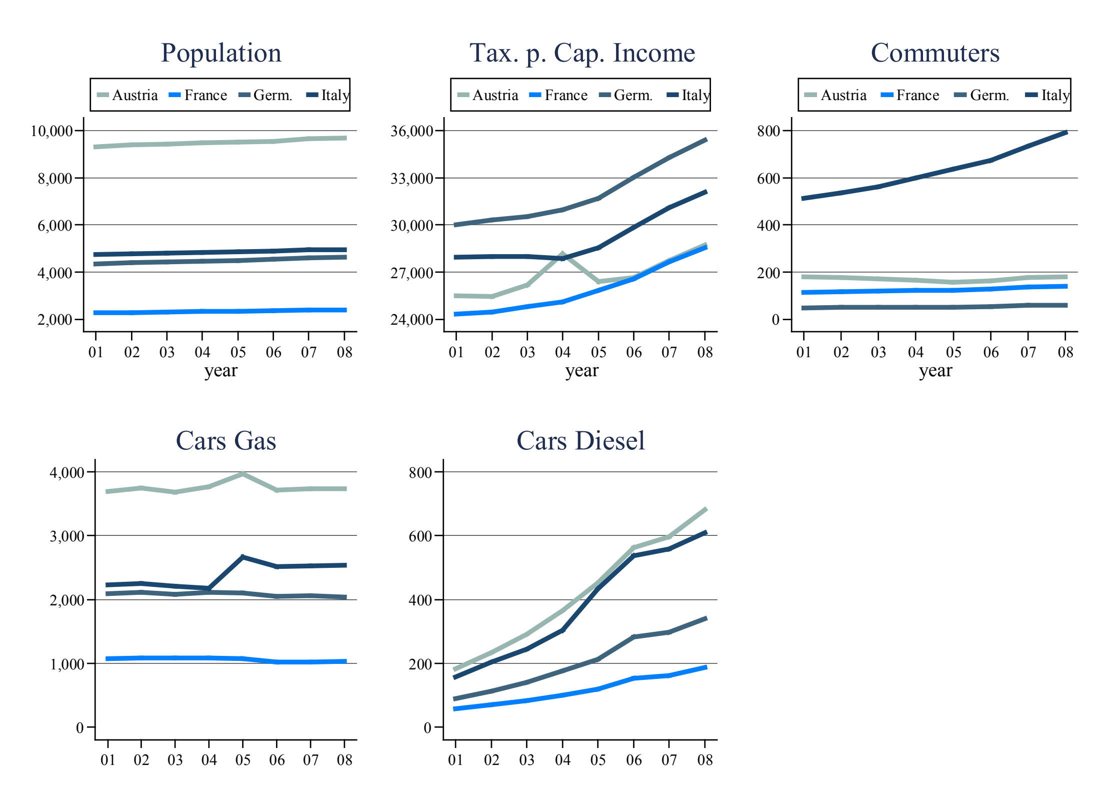


Figure II-6: Development of socio-economic variables in the border regions

According to Figure II-6, Swiss municipalities adjacent to Austria clearly have the largest average population, but those municipalities are larger compared to those adjacent to Italy or France. The population increased slightly in all border regions. Over the eight years considered, the population increased by some 3.8% in regions bordering Austria, in regions bordering France by 5.7%, Germany 6.3% and Italy 4.6%.

The taxable income per capita at the municipality level also increased in the regions bordering Austria, France, Germany and Italy over the sample period and grew in nominal terms by 12.6%, 16.7%, 17.9% and 14.8% respectively. On average, taxable income per capita was highest in municipalities bordering Germany, followed by those bordering France, Italy and Austria.

The stock of diesel-powered cars increased very strongly in all border regions. In the regions adjacent to Austria, Germany and Italy, the stock increased by some 280% and in the region bordering France by some 210%. The stock of gasoline-powered cars developed quite differently among the border regions. In the border region to Austria, there was a slight increase of 1.5% over the sample period, and in the regions adjacent to France and Germany, the stock decreased by 8.8% and 2.3% respectively. Only in the regions adjacent to Italy, there was an increase of 13%. On average, there were 1750 gasoline-powered cars registered in each municipality. The number of daily commuters remained almost constant in the regions bordering Austria, but increased in those adjacent to France and Germany by 42% and 20%

respectively. In the regions bordering Italy, there was a high increase of almost 54%. Apparently, there are many more commuters heading from Italy to Swiss municipalities than from any other foreign area. An overview of all variables that help to explain gasoline demand in the border regions with reported minimum, maximum and median values can be found in the table below.

Averaged at the municipal level				
Variable	**Measure**	**Min.**	**Max.**	**Median**
Gasoline Demand				
Border region to Austria	1'000 l / year	306	29'900	3'758
Border region to France	1'000 l / year	155	66'000	3'809
Border region to Germany	1'000 l / year	180	40'600	2'428
Border region to Italy	1'000 l / year	142	29'200	3'355
Switzerland	1'000 l / year	142	66'000	3'186
Swiss Gasoline Price				
Border region to Austria	CHF / l	1.30	1.79	1.39
Border region to France	CHF / l	1.30	1.80	1.55
Border region to Germany	CHF / l	1.30	1.79	1.53
Border region to Italy	CHF / l	1.29	1.75	1.50
Switzerland	CHF / l	1.29	1.80	1.50
Foreign Gasoline Price				
Border region to Austria	CHF / l	1.33	1.93	1.46
Border region to France	CHF / l	1.44	2.16	1.80
Border region to Germany	CHF / l	1.54	2.25	1.89
Border region to Italy	CHF / l	1.55	2.25	1.90
Switzerland	CHF / l	1.33	2.25	1.79
Swiss Per Capita Income (taxable)				
Border region to Austria	CHF / year	21'726	39'051	26'911
Border region to France	CHF / year	21'816	108'089	32'045
Border region to Germany	CHF / year	20'405	94'471	31'977
Border region to Italy	CHF / year	17'618	41'393	29'057
Switzerland	CHF / year	17'618	108'089	31'233
Commuters				
Border region to Austria	Persons / year	0	1'077	89
Border region to France	Persons / year	0	30'756	107
Border region to Germany	Persons / year	0	3'149	17
Border region to Italy	Persons / year	0	8'298	210
Switzerland	Persons / year	14	30'756	45
Swiss Population				
Border region to Austria	Persons	3'211	71'126	7'653
Border region to France	Persons	305	180'655	4'911
Border region to Germany	Persons	232	97'060	4'939
Border region to Italy	Persons	550	50'603	2'431
Switzerland	Persons	232	180'655	4'826
Stock of Cars (Gasoline)				
Border region to Austria	Cars	813	27'533	3'178
Border region to France	Cars	147	80'697	2'342
Border region to Germany	Cars	142	35'979	2'489
Border region to Italy	Cars	309	29'920	1'256
Switzerland	Cars	142	80'697	2'353
Stock of Cars (Diesel)				
Border region to Austria	Cars	46	4'735	319
Border region to France	Cars	8	10'363	251
Border region to Germany	Cars	8	5'521	232
Border region to Italy	Cars	20	7'492	213
Switzerland	Cars	8	10'363	240
Distance from Border				
Border region to Austria	km	0.70	39.55	12.37
Border region to France	km	0.50	34.86	12.34
Border region to Germany	km	0.20	44.23	17.32
Border region to Italy	km	0.70	41.68	6.12
Switzerland	km	0.70	44.23	13.50

Table II-2: Descriptive statistics of variables used in estimation

Figure II-7 depicts the geographical centre of the Swiss municipalities (grey crosses). The dark-blue shaded municipalities represent the ones in the border cantons which finally remain in our sample. We are convinced that despite the strong reduction in the data, the sample enables us to estimate gasoline tourism very well. The urban centres and the most intensely used border-crossing points are represented in the data. This is the reduced final sample with municipalities which at least have one gasoline station of the five companies. Region 2 is sparsely populated and barely accessible due to topological conditions; accordingly, we are not much concerned about too few municipalities with sales data available in that region. The regions where Germans can enter Switzerland are very well represented (region 3 and 4). From Austria, potential gasoline purchasers can only enter region 5, since region 6 would hardly allow it due to topological conditions. The same holds for the regions adjacent to Italy, where Switzerland is hardly accessible from regions 6 or 8. Region 9 is the canton Wallis, which is only accessible over mountain passes.

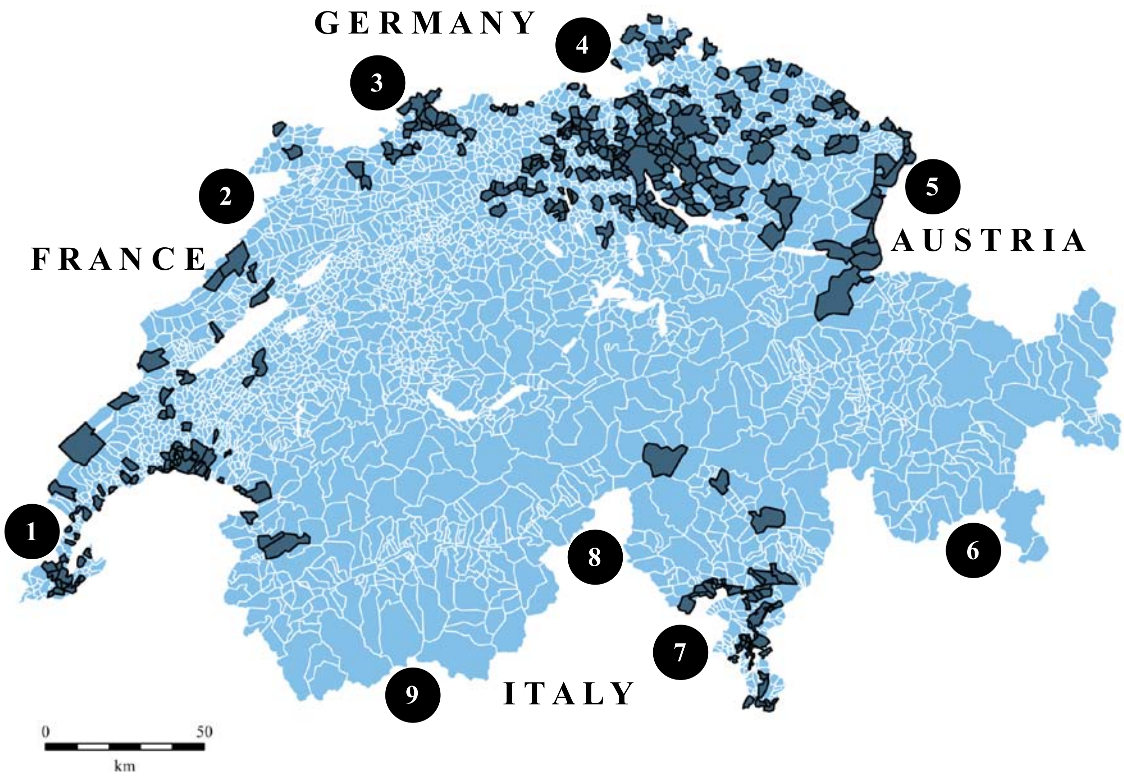


Figure II-7: Switzerland and its neighbouring countries

5 Econometric Approach and Estimation Results

5.1 Econometric Approach

When choosing an econometric approach to estimate the model stated by equation (II.7), it is important to keep in mind that the main goals of this analysis are the estimation of the price elasticity and the estimation of gasoline tourism. For the estimation of gasoline tourism, we need a model which can be used for prediction purposes. For instance, we will need to predict gasoline demand in Swiss border regions for a situation where the price differential to the foreign countries is absent.

There are different approaches to estimate models with underlying time-series cross-section data. The applicability of the various models depends mainly on the structure of the panel data set and on the final purpose of the study. With panel data, the number of observations has two dimensions, namely N the number of cross-sectional units and T the time horizon. Estimations using data sets with small N and large T can be conducted using the procedure for pooled time-series and cross-sectional data suggested by Kmenta (1997) and Greene (2003)[6]. For the cases in which T is small and N is large ($\mathrm{N} \rightarrow \infty$), there are different econometric techniques for estimation, such as pooled OLS, fixed and random effects models, and random effects autoregressive models[7].

An important advantage of panel data models is the increased precision in estimation. A further attraction of panel data is the capacity to control for unobserved individual heterogeneity. The individual-specific effects model for the dependent variable y_{it} is specified as

$$y_{it} = \alpha_i + x_{it}^T \cdot \beta + \varepsilon_{it} \tag{II.12}$$

[6] However, this approach would not be consistent for the present case, in which we only have $\mathrm{T=8}$ compared to $\mathrm{N=315}$. Banfi et al. (2005) used these estimators. In fact, the data set which they used consisted of $\mathrm{N=3}$ and $\mathrm{T=23}$ observations.

[7] The description of the econometric models for panel data is based on Cameron and Trivedi (A. Colin Cameron and Pravin K. Trivedi n.d.).

In the fixed effects model, the α_i may be correlated with the regressors x_{it}. The error term in equation (II.12) is expressed as $u_{it} = \varepsilon_{it} + \alpha_i$ and permits the regressors x_{it} to be correlated with the time-invariant component of the error, α_i. An attraction of the FE model is to obtain consistent estimates of the marginal effects β, provided the regressors are time-varying. Accordingly, fixed effects estimation is a conditional analysis, measuring the effect of x_{it} on y_{it} controlling for the individual effects α_i. But, as stated above, prediction is possible only for individuals in the particular sample being used. The drawback, on the other hand, is that, given only knowledge about β, the estimated values for y_{it} cannot be consistently predicted without consistent estimates for α_i, which is not possible for short panels. When applying a FE model, one has to be aware that the errors may be serially correlated, with the consequence that cluster-robust standard errors are required. If the error is $u_{it} = \varepsilon_{it} + \alpha_i$, then, even if it holds that ε_{it} is i.i.d., we have $Corr(u_{it}, u_{is}) \neq 0$ if α_i, meaning that the individual effect induces serial correlation.

In the random effects model, it is assumed that α_i in equation (II.12) is purely random and hence uncorrelated with the regressors, which is a stronger assumption than made for the fixed effects model. The model can be estimated using generalized least squares (GLS) which has the advantage that consistent estimates for all coefficients can be obtained, even if they are time-invariant. The disadvantage is that the model provides inconsistent estimates if the FE model is appropriate. Considering the goals of this study, a random effects model would be quite appropriate for our purpose, provided we can rely on uncorrelated error terms. In the RE model, it is assumed that α_i is i.i.d. with a variance of σ_α^2 and that ε_{it} is i.i.d. with a variance of σ_ε^2. The u_{it} has a variance of $Var(u_{it}) = \sigma_\alpha^2 + \sigma_\varepsilon^2$ and a covariance of $Cov(u_{it}, u_{is}) = \sigma_\alpha^2$. It follows for the RE model that

$$\rho_u = Corr(u_{it}, u_{is}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2} \tag{II.13}$$

This correlation is constant across i and t and is called the intra-class correlation of the error. The RE-model therefore also permits serial correlation in the model error (constant across individuals and constant for all time lags) and can approach 1 if the random effect is large relative to the idiosyncratic error.

To adjust for heteroscedasticity, one can use cluster-robust standard errors. However, micro-econometric places greater emphasis on correction for the correlation in individual errors. In contrast to the fixed effects estimation, random effects estimation is an example of a marginal analysis, as the individual effects are considered as i.i.d. random variables over the whole sample, with the effect that random effect estimator can be applied outside the sample. This means that a RE model is preferable for the present purposes.

The general FE and RE estimators are calculated on the assumption that the idiosyncratic error is $\varepsilon_{it} \sim (0, \sigma_\varepsilon^2)$. In panel applications, this assumption is often not satisfied. The panel estimators may still retain consistency provided that are independent over i, but the reported standard errors are incorrect. For short panels, it is possible to calculate cluster-robust standard errors, under the weaker assumption that the errors are independent among the individuals and that $N \to \infty$. In particular, ε_{it} may be heteroscedastic.

The choice of the most appropriate estimators should be based on the results of some statistical tests as well as on the specific goals of a study. The traditional tests are firstly to estimate the specified model by pooled OLS. Then an F test is applied to test for the presence of individual specific fixed effects $H_0 : \alpha_i = 0, \forall i$, which is very likely to be the case. Then the Breusch and Pagan Lagrangian multiplier test for random effects is applied. This tests the hypothesis that the variance of the individual effects α_i is equal to zero, which would correspond to the pooled OLS model $H_0 : Var(\alpha_i) = 0$. The next step compares the fixed and random effects model using the Hausman test. The Hausman procedure tests whether the results obtained from the random effects model are significantly different from those obtained from the fixed effects model. The test statistic is computed by

$$H = \left(\hat{\beta}_{FE} - \hat{\beta}_{RE}\right)^T \cdot Var^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{RE}) \cdot \left(\hat{\beta}_{FE} - \hat{\beta}_{RE}\right) \qquad \text{(II.14)}$$

Consequently, and as already mentioned in our discussion of the empirical strategy, a fixed effects model appears to be inappropriate for our purposes. First, the data set contains time invariant regressors such as the municipalities' distance from the border. Second, we are interested in both estimating marginal effects such as the price elasticity of gasoline demand and in being able to make predictions of estimated gasoline demand and consistent estimates for gasoline tourism. It will be seen later that the simulation to calculate the share of gasoline tourism will require the price ratio set equal to unity. This, though, is a situation which is not supported by the data and therefore represents an out-of-sample prediction. Cameron and Trivedi (2009) and Baltagi (2005) have discussed the fact that the RE model performs well in out-of-sample predictions.

5.2 Estimation Results

5.2.1 General Discussion and Choice of the Final Model

As discussed in the previous section for the estimation of the model stated by equation (II.7) (with a dummy for company B included), traditional panel data estimation approaches such as the random and the fixed effects models are appropriate, since our data set is characterised by a short sample period over a relatively large number of panel units. In Table II-4, we present the econometric results of the estimation of the model stated by equation (II.7) using a pooled OLS estimator, a FE and a RE estimator.

Most of the coefficients are significant in all models and bear the expected signs. Generally, the coefficients obtained by pooled OLS are different from those obtained by using a random effects and fixed effects model. Further, the coefficients obtained from the FE and RE specification are similar. The difference between the OLS and the RE or FE models could be due to unobserved heterogeneity bias, i.e. the coefficients obtained from pooled OLS are biased away from zero compared to those obtained from the FE or RE estimations. For this reason, we use an F-test for the presence of fixed effects to compare the FE model against the pooled OLS model and a Breusch-Pagan test to discriminate between the pooled OLS model and the RE model.

The F statistic, which tests the hypothesis that all individual effects α_i are equal to zero, is clearly rejected and hence the fixed effects model has to be preferred over the pooled OLS model. The Breusch and Pagan test for random effects tests whether the variance of the individual effects α_i is equal to zero. This hypothesis is also strongly rejected, and hence the random effects model also has to be preferred over the pooled OLS model. The Hausman test rejects the hypothesis of no systematic difference in the coefficients obtained by the FE effects model and the RE effects model (the coefficients of the FE model are consistent under the null and under the alternative, whereas the coefficients of the RE model are consistent and efficient under the alternative). However, remember that the FE estimator can be imprecise for coefficients of variables with low within variation, where "imprecise" refers to both the estimation of the variance component and the coefficient (see Cameron and Trivedi, 2009). Table II-3 summarizes the overall, between and within standard deviation of the variables used in equation (II.7).

Variable	s_o	s_b	s_w	s_w/s_b
$\ln(G_{it})$	0.739	0.708	0.215	0.304
$\ln(PG_{CH,bt})$	0.092	0.010	0.092	9.1
$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)$	0.061	0.047	0.040	0.866
$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)\ln(dist_i)$	0.260	0.237	0.108	0.456
$\ln(CarsG_{CH,it})$	0.990	0.987	0.091	0.092
$\ln(CarsD_{CH,it})$	1.077	1.009	0.379	0.375
$\ln(POP_{CH,it})$	1.058	1.059	0.032	0.030
$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.228	0.223	0.052	0.233
$\ln(dist_i)$	1.097	1.097	0.000	0.000
$\ln(Commu_{it})$	2.738	2.707	0.439	0.162
DB_{it}	0.376	0.35	0.137	0.391

Table II-3: Comparison of the overall, between and within standard deviation of the variables

All variables except the Swiss gasoline price and the price ratio show a far higher between than within variation, which may raise doubts about the precision of the fixed effects estimates (see Cameron and Trivedi, 2009). For instance, the coefficients of income per capita and population do not bear the expected sign. It should be noted, though, that the Hausman test reports a p-value of 0.003, which is low. However, we believe that the random effects model can be used even if the p-value of the Hausman test does not strongly support its application. On the other hand, the calculation of the Hausman test uses the estimated coefficients of the fixed effects model and assumes consistency of the coefficients, which is doubtful due to the low within variation of the data. Further, applied econometric works and examples (e.g. Cameron and Trivedi, 2009, Ch.8) of the Hausman test mention that the test rejects the null-hypothesis of no systematic differences in the random and fixed effects model only if (a) a fixed effects model can be trusted, which is perhaps not the case, and (b) if the assigned p-value is very low.

Moreover, the computation of the Hausman test relies on standard errors which may be considerably too small, since in this first step we adjusted neither for heteroscedasticity nor for serial correlation[8].

Based on this discussion, we can exclude the pooled OLS model from the empirical analysis and keep the FE and RE specification. As we will discuss later, for the analysis of gasoline tourism, we prefer the results obtained from the random effects specification because of the low within variation of the variables and the better performance of the RE model in out-of-sample prediction, as pointed out by Farsi and Filippini (2004) and Baltagi (2005). However, for the computation of price elasticities, we decided to illustrate the results of both econometric approaches[9].

8 The Fe and RE versions can be estimated by considering serial correlation. The Wooldrige test for autocorrelation rejects the null of no first-order autocorrelation. However, it has previously been mentioned that the random effects specification also accounts for serial correlation in the error term under the assumption that the autocorrelation coefficient is the same for all time lags and the same across all individuals. However, we extend the FE and RE models and incorporate serial correlation of order 1 in the residuals and provide respective estimation results of this FE AR-Model and RE AR-Model in the Appendix. In these models, the assumption is that the autocorrelation coefficient is the same across (315) individuals, which is doubtful in our opinion. The basic idea behind the AR models is that shocks (time variant, unobserved characteristics, e.g. the exchange rate or changes in the income of foreign car owners) affect future events and therefore should be incorporated in the model. The coefficients, except those of the unweighted and weighted price ratios, are very similar to the original model.

9 The original model has also been estimated with time dummies and with a linear time trend. The inclusion of a time trend, however, removes the explanatory power of the Swiss gasoline price and the foreign gasoline price, since those variables mainly vary over time and not across individuals. The results are provided in the Appendix.

Coeff.	Variable	Pooled OLS	FE Model	RE Model
α_0	Constant	14.59 (11.16)***	17.04 (9.17)***	13.69 (13.37)***
α_1	$\ln(PG_{CH,bt})$	-0.705 (-2.80)***	-0.1284 (-1.16)	-0.2108 (-1.95)**
α_2	$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)$	2.378 (3.72) ***	1.477 (5.18)***	1.469 (5.19)***
α_3	$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)\ln(dist_i)$	-0.899 (-4.78)***	-0.398 (-4.20)***	-0.434(-4.62)***
α_4	$\ln(CarsG_{CH,it})$	0.272 (2.35)**	0.1021 (1.97)**	0.1158 (2.24)**
α_5	$\ln(CarsD_{CH,it})$	-0.03 (-0.44)	-0.1882 (5.77)***	-0.1882 (5.87)***
α_6	$\ln(POP_{CH,it})$	-0.153 (-1.69)*	-0.1963 (-1.12)	0.1574 (2.77)***
α_7	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.212 (2.99)***	-0.0350 (-0.38)	0.0500 (0.62)
α_8	$\ln(dist_i)$	0.077 (2.08)**	-----	-0.0641 (-1.40)
α_9	$\ln(Commu_{it})$	0.040 (4.27)***	-0.01542 (-1.57)	-0.0050 (-0.55)
α_{10}	DB_{it}	0.466 (12.04)***	0.681 (22.72)***	0.667 (22.90)***
$dist_{crit}$	$e^{\frac{-\alpha_2}{\alpha_3}}$		41.0 km	29.6 km
	F test for FE		$F = 77.4^{***}$	
	B&P test for RE			$\chi^2(1) = 7'100^{***}$
	Wooldridge test			$F(1,314) = 176^{***}$
	Hausman test			$\chi^2(8) = 23.27^{***}$

Table II-4: Pooled OLS, FE and RE estimation results, dependent variable gasoline sales per station

- Number of Observations 2520 (T=8 years, N=315 municipalities)
- t-statistics are in parentheses: ***, ** and * indicate 1%, 5% and 10% significance levels respectively
- Cantonal dummies are not tabulated

Before assessing the price elasticities, as described in equations (II.8) and (II.9), we discuss the coefficients of the models.

First, the coefficient of the Swiss gasoline price is expected to be negative, since in represents the elasticity of gasoline demand in the absence of gasoline tourism. This is true for all models. However, the coefficient is not significant in the fixed effects model. Remember that α_1 can be interpreted as the elasticity of gasoline demand with respect to the Swiss gasoline price in the absence of gasoline tourism (after the critical distance). Depending on the RE model, this could mean, either, that the price elasticity is negative (-0.211) or that the price elasticity is not significantly different from zero (FE model) in the absence of gasoline tourism. This in turn would mean that Swiss car owners do not respond to price changes in the short run. Both cases are in line with the meta-analysis conducted by Brons et al. (2006). However, it is mentioned that insignificant price elasticities can seldom be found in the respective studies and, moreover, a mean value of -0.3 was found for the own-price elasticity with respect to gasoline demand, a value which is not far from that reported by the RE model. As previously mentioned, the random effects model is preferred due to the goals of the present study.

The coefficient of the unweighted price ratio α_2 is positive and significantly different from zero and moreover very similar in terms of absolute value in the FE model and the RE model.

The coefficient α_3 of the price ratio weighted with the distance from the border has the expected negative sign in both the FE and RE models, but is rather higher in terms of absolute value in the RE model.

The coefficient α_4 of the stock of gasoline-powered vehicles has the proper positive sign in both the FE and RE models. However, it is larger in absolute value (0.1158 vs. 0.1021) and has a lower standard error in the RE model. Accordingly, a ceteris paribus 10% increase of the stock of gasoline-powered cars would increase gasoline demand at the municipal level on average by 1.16% according to the RE model.

The coefficient of the stock of diesel-powered vehicles, α_5, has the proper negative sign in both models and is significant at the 1% level. In both models, it is equal to -0.1882.

The coefficient of the residential population is positive and significantly different from zero in the random effects model, but negative and insignificant in the fixed effects model. This can be explained with the help of Figure II-6 and Table II-3, where it can be seen that the population variable exhibits almost no within variation over the sample period and accordingly, the fixed effects estimator may be imprecise. In part, the same applies to the taxable income per capita: the fixed effects model assigns a negative sign to the respective coefficient but the estimated coefficient is not significantly different from zero. For the random effects model, the coefficient is positive but not significantly different from zero.

The coefficient of the distance from the border, α_8, has the expected negative sign in the random effects model but is not significant.

The coefficient of the number of commuters bears an insignificant and negative value in all models which was not expected. The dummy variable for gasoline company B has the expected sign and is significant at the 1% level in both the FE and RE models. None of the dummy variables of the cantons is significant. However, they increase the explanatory power of these models.

Summing up the preceding comments, and since an out-of-sample prediction has to be performed for the calculation of gasoline tourism, we rule out the fixed effects model. Accordingly, the random effects model in its original form is the preferred model, with which we believe we may to obtain reliable forecasts for gasoline tourism and price elasticities with respect to both the Swiss and foreign gasoline prices.

5.2.2 Swiss and Foreign Gasoline Price Elasticity

To compute the estimated price elasticity of Swiss gasoline demand with respect to Swiss and foreign gasoline prices, the results of the random and fixed effects models as depicted in Table II-4, are used and fed into equations (II.8) and (II.9). The spatial dependence of the elasticity of Swiss gasoline sales with respect to the Swiss gasoline price and foreign gasoline price is depicted in Figure II-8. The darkly shaded bars represent the average of the foreign gasoline price elasticity, and the pale bars represent the average of the Swiss gasoline price elasticity within the respective distance intervals. The red tables show the results of the fixed effects estimation of the respective models and the blue ones those from the random effects estimation. The critical distance at which the foreign gasoline price elasticity is zero is calculated by setting the elasticity with respect to the foreign gasoline price equal to zero, as in equation (II.11). For example, we obtain $dist_{crit} = e^{-\alpha_2/\alpha_3} = 29.6km$ for the original RE model and $dist_{crit} = e^{-\alpha_2/\alpha_3} = 29.6km$ for the FE model[10]. The results can also be understood graphically by considering the distance class in which the average of the foreign gasoline price elasticity is close to zero.

Beyond this distance, the elasticity of Swiss gasoline sales with respect to the foreign gasoline price is not significantly different from zero. This means that, beyond this critical distance, gasoline sales at Swiss stations are not affected by changes in the foreign gasoline price: according to the original RE model, gasoline tourism at Swiss gasoline stations occurs within a distance from approximately 30 kilometres from the border according to the original RE model. The maximum distance a municipality is located from the border in the sample is 46 km. However, more than 80% of the municipalities are located within a band of 30 kilometres from the border.

As expected, a potential increase in the Swiss gasoline price has a higher impact on Swiss gasoline sales than an increase in the foreign gasoline price would have, since then, not only foreign car owners but also domestic ones respond to the price change. However, this only holds when the coefficient α_1 is negative. If the foreign price changes, Swiss car owners are not affected by this change, as long as the Swiss price stays below the foreign price.

Overall, the development of the foreign and the Swiss gasoline price elasticity is well behaved for both the FE and RE models. According to the RE model, the average of the Swiss gasoline price elasticity over the sample distance is -0.65. The "pure" influence of the Swiss gasoline price is measured by the coefficient α_1 and is -0.211 according to the RE model. In summary, the model reports price elasticities in an acceptable range. Accordingly, if the Swiss gasoline price is, ceteris paribus, increased by 10%, then domestic sales decrease by -6.5%. However, this range consists of both effects to which a change in the domestic gasoline price might lead,

[10] For the RE AR model, a critical distance of 56km is obtained, whereas for the FE AR model, a critical distance of 336km is obtained, both are counter-intuitive. For instance, Banfi et al. (2005) assumed a critical distance of 5km, while Michaelis (2004) calculated values between 10 and 30km from the border for a choice of European countries. Basically, this would imply that gasoline sales in Switzerland in all municipalities' reference stations are affected by the foreign gasoline price.

namely the decrease in consumption by Swiss residents and the expected decrease in gasoline tourism by foreign car owners. The decrease in consumption by Swiss residents alone following by an 10% increase in the gasoline price would be around -2.1%. If the foreign gasoline price increases by 10%, then the Swiss gasoline consumption also increases as long as the absolute Swiss gasoline price is lower than the foreign gasoline price. The increase in consumption would be around 4.3% on average, according to the RE model. It is obvious from Figure II-8 that the sensitivity of Swiss gasoline demand to the foreign gasoline price is high close to the border and fades over the distance range.

These results can be compared to those obtained by Baranzini et al. (2012)and Pock (2010). Baranzini found a price elasticity of Swiss gasoline sales with respect to the domestic price of -0.1 in the short-run and of -0.34 in the long-run. The price elasticity with respect to the foreign gasoline price was reported to be 0.07 in the short-run and 0.1 in the long-run. The differences in the results may exist due to the different approach used to estimate these elasticities, as discussed in the literature review.

However, this shows that the findings are consistent with the expectation that the elasticities should be higher close to the border and secondly, that for a small country, as Switzerland is, foreign gasoline prices have a significant impact on gasoline sales. The study published by Pock (2010) analyzed gasoline demand in several European countries, including Switzerland, and found average price elasticities of gasoline demand between -0.2 and -0.5 for several European countries.

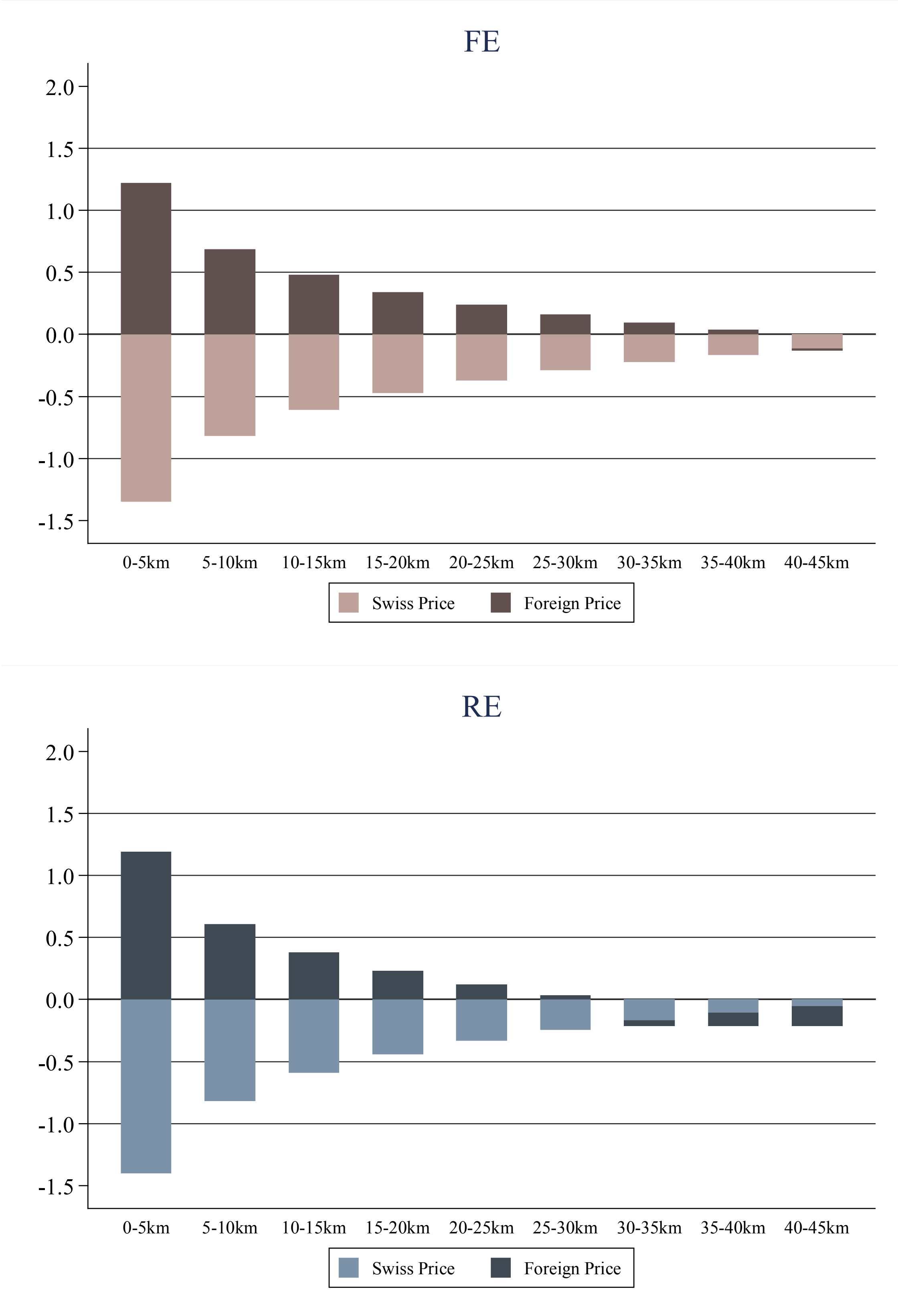


Figure II-8: Development of the Swiss and foreign gasoline price elasticity by distance from the border

5.2.3 Evaluation of Gasoline Tourism

In order to set up a scenario where Switzerland would be unaffected by foreign cross-border purchasing of gasoline, the price ratio between the foreign and the Swiss gasoline price in equation (II.7) is set equal to unity[11]. This would correspond to a situation where there is no incentive for foreigners to fuel in Switzerland. Based on the estimation results, it is possible to predict gasoline sales per station as if no price differentials across the borders existed. The difference between these simulated values and the estimated values with the actual price ratio can then be designated as estimated gasoline tourism at a reference gasoline station. The ratio between this difference and the originally estimated sales then results in percentage gasoline tourism at the reference station[12].

Municipalities located closely to the border then should display higher relative values of gasoline tourism compared to those more distant from the border. In order to evaluate the overall volume of gasoline sales to cross-border purchasers, a first approach would be to calculate a weighted average of percentage gasoline tourism for selected distance ranges from the border (e.g. 0-5km, 5-10km and so on) and then to multiply it with the average sales of gasoline stations from the sample located in the respective distance ranges. Of course, this is a simplifying assumption, since gasoline tourism is originally calculated for each municipality. In order to be able to make a projection of the overall volume, it is assumed that each municipality (i.e., each reference gasoline station) in each distance range experiences the same relative value of gasoline tourism. Then, those sales are multiplied by the actual number of gasoline stations within the distance range and summed over the distance range. Then, those sales are multiplied by the actual number of gasoline stations within the distance range and summed over the distance range. This procedure is summarized in Figure II-9.

---

[11] We use the same approach as that used by Banfi et al. (2005)

[12] Recall that sales are averaged at gasoline station level for each municipality and that this value hence represents the representative gasoline station for this municipality.

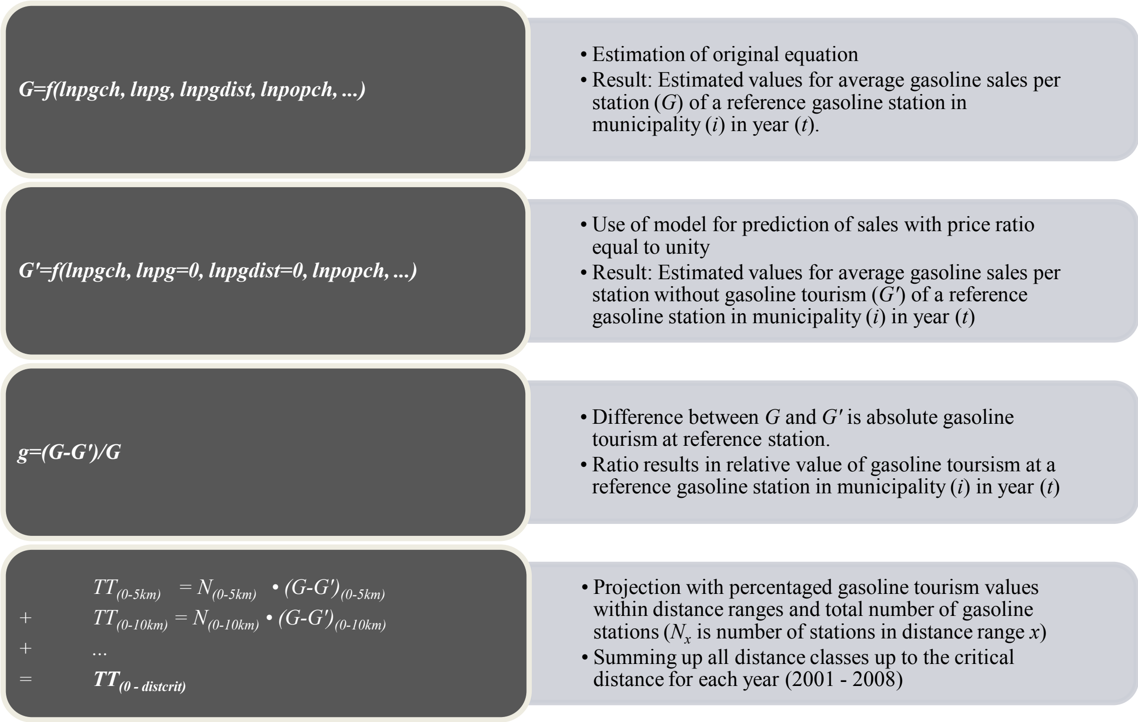


Figure II-9: Procedure to calculate relative and absolute (volume) values for gasoline tourism

As already mentioned, this is a rather crude simplification. There is no time series for the total number of stations with distance from the border over the sample period. Accordingly, the available values for the year 2008 were used.

The procedure described in Figure II-9 is applied to each border region separately. The results in terms of absolute volumes of estimated gasoline tourism are illustrated in Figure II-10 for the original random effects model. The predicted volumes were summed by distance classes of five kilometres up to the critical distance (30km) to which gasoline tourism is thought to occur. The very small share of gasoline tourism for municipalities bordering Austria can be explained by the relatively small price difference between Switzerland and Austria. In absolute terms, municipalities bordering Germany and France show the highest volumes of gasoline sold to foreigners, followed by those adjacent to Italy. Depending on the structure of the model described by equation (II.7), gasoline tourism fades the more distant a municipality is from the border. Therefore, it can be observed that approximately three quarters of absolute gasoline tourism occurs within a distance range of 15km. Roughly, a yearly average of some 4.8 billion litres of gasoline were sold at Swiss gasoline stations between 2001 and 2008.

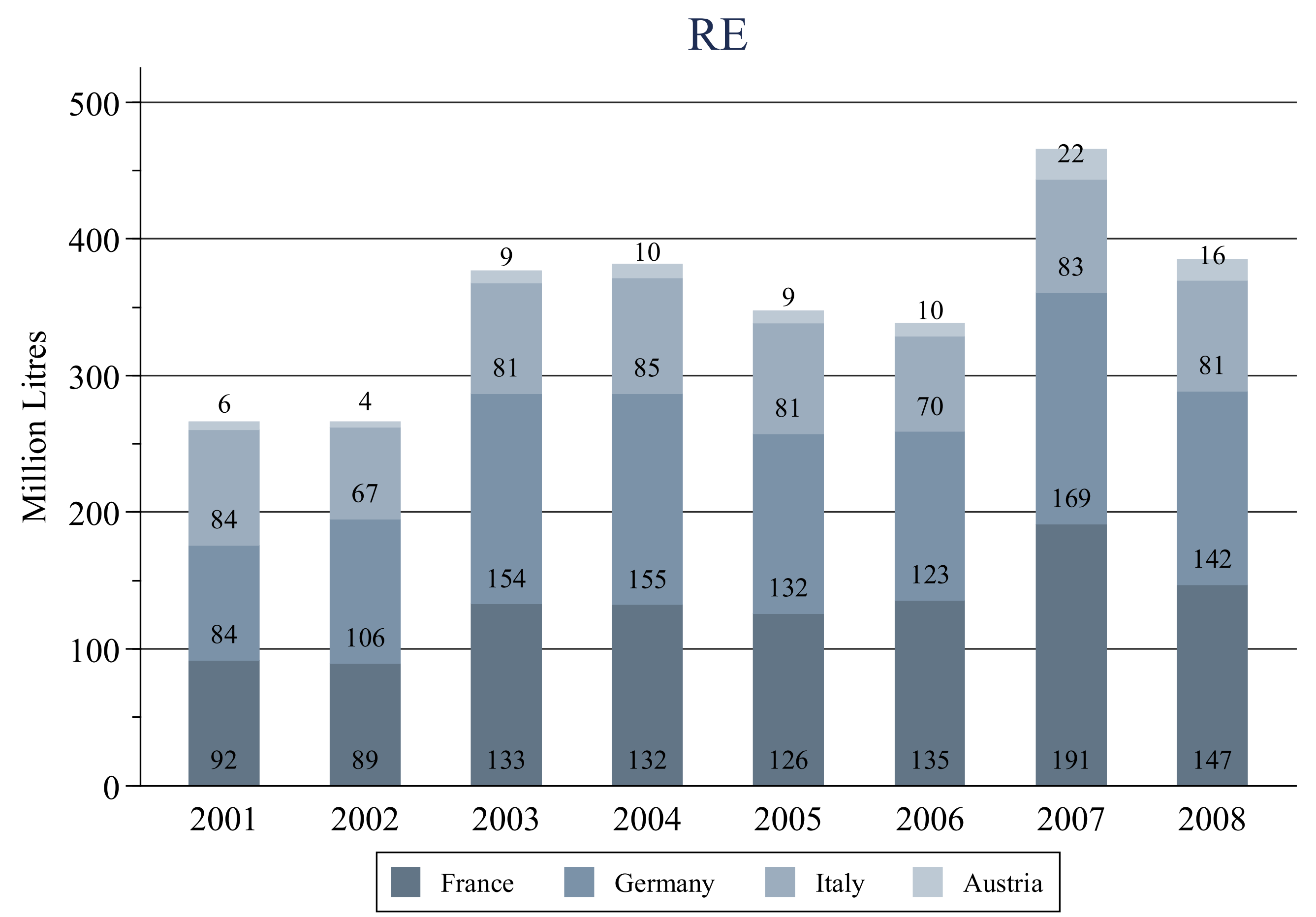


Figure II-10: Projected values of gasoline tourism for the RE model

On average, within 0 to 10 kilometres from the border, the share of gasoline tourism in overall sales approximately is twice as high as the average over the whole critical distance. However, the results illustrated in Figure II-10 only rely on the calculated means of gasoline tourism.

In order to obtain more information about the predicted ranges of the volume shares, we calculate the standard errors of the prediction (as previously described in Figure II-9). Figure II-11 depicts the development of the predicted mean, lower and upper 95% confidence intervals of relative gasoline tourism over the sample period for the RE model.

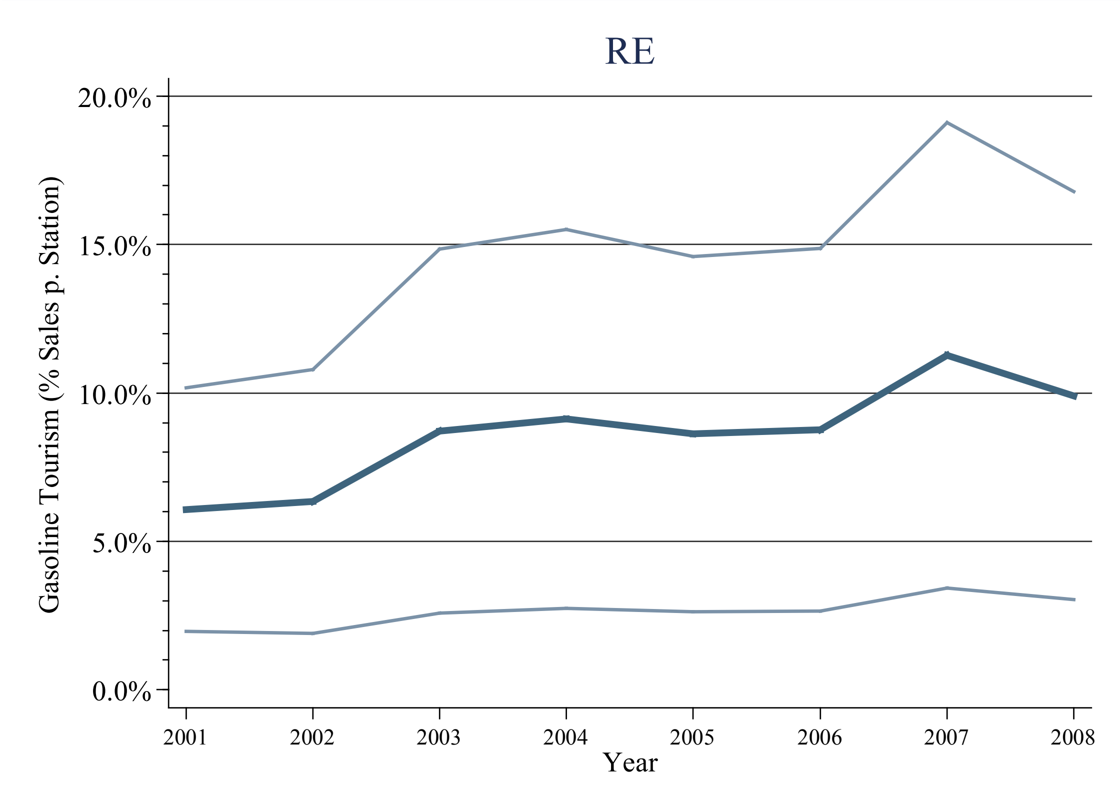


Figure II-11: 95% confidence interval of relative gasoline tourism for the RE model

The average share of gasoline tourism in the overall sales volumes was around 9% for the years 2001 - 2008[13]. The average of the lower confidence interval is 3% and the average of the upper confidence interval is 15%. In 2007, there were some 4.5 billion litres of gasoline sold on the Swiss retail market (the maximum over the sample period). The lower confidence interval for relative gasoline tourism in 2007 is 3.5%, the mean is 11% and the upper confidence interval is 19%. Accordingly, there were some 158 million - 855 million litres of gasoline sold to foreigners, with a mean value of some 495 million litres. The majority of gasoline sold to foreigners (almost three quarters) is sold within 15km from the border, being approximately the half of the critical distance.

[13] More details can be found in the report for the Federal Office of Energy at http://www.bfe.admin.ch/energie/00588/00589/00644/index.html?lang=de&msg-id=33842, Chapter 3.2.

Figure II-12 shows the spatial distribution of the estimated shares of gasoline tourism in overall Swiss gasoline sales for the year 2007 (where the existing price differentials across the borders were the highest) according to the results of the original RE model. The results look equal if depicted according to the other models but are different in values. The values of gasoline tourism obtained for the 315 municipalities in the sample were projected on the whole Swiss territory by comparing the municipalities' distances from the border and corresponding price differentials across the border. It can be seen that gasoline tourism generally reaches high values in municipalities close to the Italian, French and German borders. 91 municipalities sale more than 20% of their overall sales to foreign car owners. Out of a total of 2721 municipalities, 1783 experience a share of more than 5% of gasoline tourism. The remaining 938 municipalities hardly experience a significant share of gasoline tourism due to their distance from the border. So the distance from the border within which a significant amount of gasoline is sold to foreigners can be approximately rated to 30-40 kilometres (since up to 30km, the lower confidence interval of relative gasoline tourism is still above zero percent).

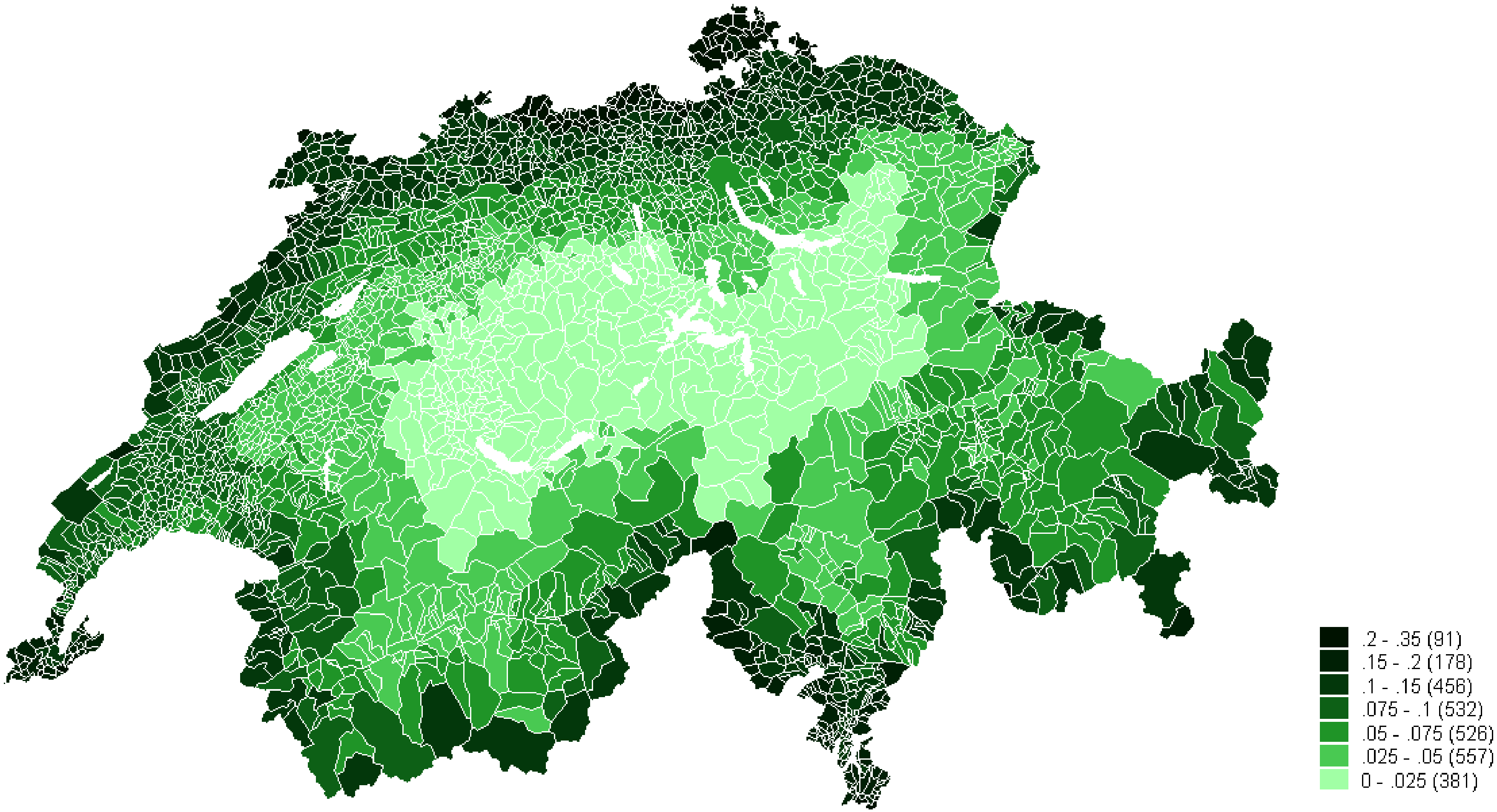


Figure II-12: Projection of percentage share of gasoline tourism to all Swiss municipalities

5.2.4 Counter-Factual Simulation: Impact of a Decrease in the Price Ratio (Induced by an Increase in the Swiss Gasoline Price)

The impact of gasoline tourism in a specific municipality depends on its distance from the border and the price ratio across the border. If the Swiss government decided to increase the retail gasoline price (e.g. due to an introduction of a carbon dioxide tax), the price ratio would decrease and gasoline stations close to the border would not only suffer from losses in sales to Swiss car owners but also of losses in sales to cross-border purchasers. The procedure described in Figure II-9 can be applied to a situation where the price ratio is not set equal to unity but simply lowered to the respective value it would have achieved given an increase in the Swiss gasoline price. The results in Figure II-13 indicate that gasoline tourism is very reactive to changes in the price ratio. An increase in the Swiss gasoline price of only 5 Swiss franc cents reduces the share of gasoline tourism in overall sales from 9% (on average) to 7.1%, an increase of 10 cents to 5.60% on average, and an increase by 20 cents reduces it to only 2.6%.

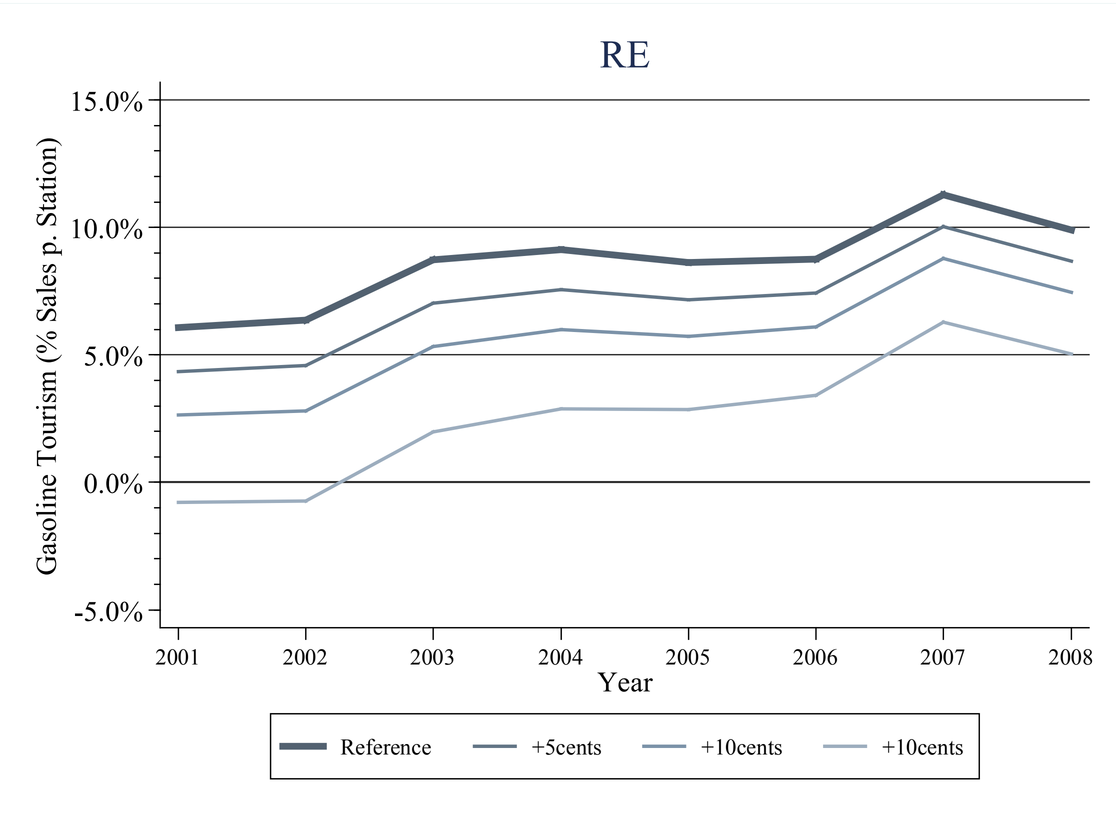


Figure II-13: Gasoline tourism as a percentage share of total sales after a hypothetical increase of the Swiss gasoline price by 5, 10 and 20 Swiss franc cents.

The border regions, however, are not equally affected by a potential increase in the Swiss gasoline price. Most municipalities would experience a decline in gasoline tourism between 45% and 50%. However, there are municipalities bordering Austria which would experience a reversed situation in which the Swiss gasoline price would exceed the Austrian gasoline price and so suffer from a decrease in gasoline tourism of more than 100% - meaning that the excess above 100% then would be gasoline imported to Switzerland by domestic car owners. The regions bordering Germany in the north, France in the west and Italy in the south experience a decline in gasoline tourism of some 40% to 60% on average, since the existing price ratio is not too strongly diminished by a price increase of 10 cents.

6 Conclusion and Outlook

The present working paper explores an ad-hoc demand model to assess gasoline demand in Switzerland at the municipal level. Further, gasoline tourism in the Swiss border regions is quantified. In order to assess gasoline tourism, the price ratio across the border is considered as the driver of cross-border purchasing of gasoline. A fixed and a random effects model is used to estimate gasoline sales per station in the municipalities and in both models, the price ratio has a significant impact in explaining gasoline sales, meaning that Swiss gasoline stations close to the border are affected by the foreign price of gasoline. From an intuitive point of view, it is obvious that gasoline tourism induced by the existing price differentials should fade with increasing distance from the border. However, there are very few studies (e.g. Michaelis 2004) which assess the phenomenon of this critical distance. Most studies made an assumption that foreign purchasing of domestic goods only takes place within e.g. 5 kilometres from the border. In the present work, we incorporate an interaction term between the price ratio across the border and the distance of the municipalities from the border. The coefficient estimated is negative and significantly different from zero and, therefore, we can conclude that gasoline tourism in Switzerland in a municipality indeed decreases with increasing distance from the border.

We found a significant price elasticity of Swiss gasoline demand (in the border regions) with respect to the Swiss gasoline price of -0.211 when gasoline tourism is not considered. The average price elasticity with gasoline tourism considered is -0.65 and significantly different from that previously mentioned. Accordingly, the demand for Swiss gasoline is considerably affected by foreigners purchasing gasoline in the Swiss border regions, so a change in the Swiss gasoline price namely not only affects Swiss consumers but also foreign consumers, since Swiss gasoline then becomes relatively less attractive to them. Moreover, the elasticity of Swiss gasoline demand in the border regions with respect to the foreign gasoline price is significantly different from zero and averages to 0.43, meaning that a decrease in the foreign price by 10% decreases demand for Swiss gasoline demand in the border regions by -4.3%.

The results indicate that gasoline tourism had an impact on municipalities up to a distance of some 30 kilometres from the border, but the main part was sold at stations located rather close to the border: within some 10 to 15 kilometres. The share of gasoline tourism in the volume sold over the sample period is not negligible. The model shows that, compared to overall sales, which accounted for some 4.5 billion litres of gasoline. Gasoline tourism reached values from some 250 million litres up to almost 450 million litres of gasoline on average. Further, it is shown that even a small increase in the Swiss gasoline price may lead to a substantial decrease in gasoline tourism.

From a policy point of view, a hypothetical CO_2 tax is likely to significantly reduce gasoline tourism in the border regions and therefore might lead to local advantages of less congestion and pollution, whereas the gasoline stations close to the border would suffer from a significant drop in sales as an adverse effect of the such a tax. Moreover, as long as the Swiss gasoline price is below the foreign gasoline price, an increase in the Swiss gasoline price has locally different effects on Swiss gasoline sales and therefore, policy makers most of all should monitor developments in the gasoline retail market in the border regions if the Swiss gasoline price is increased. In addition, if a neighbouring country decided to introduce a rebate system for residents living close to the Swiss to prevent them from fuelling in Swiss territory, sales in the adjacent Swiss border regions would decrease significantly too. In conclusion, policy makers should be aware that the price differential across the border has a significant impact on the total amount of Swiss gasoline sold and most of all on the share of gasoline sold in the border regions.

Despite the fact that the present study is fundamentally different from that conducted by Banfi et al. (2005), we were able to estimate very similar values for the share of gasoline tourism. We believe that for the present data sample of 315 municipalities observed over eight years, the estimation of a random effects model is more beneficial than a fixed effects specification, since first, we have variables with very low or no within variation; second, fixed effects is asymptotically consistent for $T \rightarrow \infty$ and, third, the estimation of gasoline tourism represents an out-of-sample prediction in the present case. Moreover, the introduction of an interaction term for the price ratio with distance from the border emerged as highly advantageous. First, the critical distance within which foreign purchasing activity is likely to take place need not be exogenously defined but is estimated by the model. Second, the specification allows for locally different patterns in gasoline sold to foreigners and price sensitivity.

Future work could investigate spatial effects with a focus not on gasoline tourism but on gasoline demand in the municipalities, using both a spatially lagged dependent variable and/or a spatial error components model.

7 Appendix

7.1 Estimation Results for FE and RE AR Models (FE and RE with Autocorrelation)

Coeff.	Variable	FE Model 2	RE Model 2
α_0	Constant	20.96 (28.27)***	14.65 (15.03)***
α_1	$\ln(PG_{CH,bt})$	-0.193 (-1.97)**	-0.217 (-2.36)**
α_2	$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)$	0.633 (2.32)**	0.664 (2.78)***
α_3	$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)\ln(dist_i)$	-0.109 (-1.08)	-0.164 (-1.89)*
α_4	$\ln(CarsG_{CH,it})$	0.034 (0.67)	0.112 (2.57)**
α_5	$\ln(CarsD_{CH,it})$	-0.081 (-1.89)*	-0.154 (-6.5)***
α_6	$\ln(POP_{CH,it})$	-0.387 (-1.55)	0.137 (2.60)***
α_7	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	-0.237 (-2.52)**	-0.022 (-0.28)
α_8	$\ln(dist_i)$	-----	-0.117 (-2.71)***
α_9	$\ln(Commu_{it})$	-0.003 (-0.34)	-0.002 (-0.27)
α_{10}	DB_{it}	0.534 (17.05)***	0.562 (19.59)***
$dist_{crit}$	$e^{\frac{-\alpha_2}{\alpha_3}}$	336.0 km	56.8 km
ρ	AR(1) disturbance	0.695	0.695
	Baltagi-WU-LBI	1.05	1.05
	F test for FE	$F = 14.2$***	
	B&P test for RE		-----
	Wooldridge test	$F(1,314) = 165$***	$F(1,314) = 167$***
	Hausman test ($\chi^2(8)$)		(35.3)***

Table II-5: FE and RE estimation results for the AR model, dependent variable gasoline sales per station

- Number of Observations 2520 (T=8 years, N=315 municipalities)
- t-statistics are in parentheses: ***, ** and * indicate 1%, 5% and 10% significance levels respectively
- Cantonal dummies are not tabulated

7.2 Estimation Results for 'Time' Models

Coeff.	Variable	FE Model dummy	RE Model dummy	FE Model time trend	RE Model time trend
α_0	Constant	12.35 (3.70)***	9.736 (3.22)***	16.15 (8.27)***	12.92 (11.96)***
α_1	$\ln(PG_{CH,bt})$	0.4796 (0.88)	0.3806 (0.703)	-0.0327 (-0.25)	-0.0527 (-0.409)
α_2	$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)$	1.93 (4.89)***	2.005 (5.10)***	1.637 (5.36)***	1.72 (5.66)***
α_3	$\ln\left(\frac{PG_{F,bt}}{PG_{CH,bt}}\right)\ln(dist_i)$	-0.416 (-4.37)***	-0.454 (-4.83)***	-0.409 (-4.31)***	-0.448(-4.77)***
α_4	$\ln(CarsG_{CH,it})$	0.090 (1.40)	0.076 (1.21)	0.06 (0.99)	0.051 (0.873)
α_5	$\ln(CarsD_{CH,it})$	-0.188 (-3.38)***	-0.157 (-2.97)***	-0.136 (-2.83)***	-0.112 (-2.41)***
α_6	$\ln(POP_{CH,it})$	-0.131 (-0.734)	0.165 (2.87)***	-0.160 (-0.905)	0.147 (2.58)***
α_7	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.062 (0.64)	0.133 (1.57)	-0.018 (-0.20)	0.069 (0.845)
α_8	$\ln(dist_i)$	-----	-0.056 (-1.21)	-----	-0.059 (-1.29)
α_9	$\ln(Commu_{it})$	-0.012 (-1.25)	-0.002 (-0.22)	-0.014 (-1.417)	-0.004 (-0.43)
α_{10}	DB_{it}	0.685 (22.82)***	0.671 (23.04)***	0.683 (22.77)***	0.669 (22.98)***
dt_1	2001	0.180 (1.35)	0.220 (1.67)*	-----	-----
dt_2	2002	0.197 (1.33)	0.224 (1.52)	-----	-----
dt_3	2003	0.171 (1.22)	0.190 (1.36)	-----	-----
dt_4	2004	0.157 (1.48)	0.175 (1.66) *	-----	-----

dt_5	2005	0.104 (1.59)	0.121 (1.88) *	-----	-----
dt_6	2006	0.070 (2.28)**	0.081 (2.67)***	-----	-----
dt_7	2007	0.004 (0165)	0.009 (0.408)	-----	-----
$time$	Time	-----	-----	-0.0148 (-1.46)	-0.022 (-2.26)**
$dist_{crit}$	$e^{\frac{-\alpha_2}{\alpha_3}}$	103.5 km	82.2 km	54.5 km	46.0 km
	F test for FE	$F = 77.3^{***}$		$F = 77.2^{***}$	
	B&P test for RE		$\chi^2(1) = 7'110^{***}$		$\chi^2(1) = 7'108^{***}$
	Hausman test ($\chi^2(8)$)		(19.75)**		(20.33)***

Table II-6: FE and RE estimation results for 'time' models, dependent variable gasoline sales per station

- Number of Observations 2520 (T=8 years, N=315 municipalities)
- t-statistics are in parentheses: ***, ** and * indicate 1%, 5% and 10% significance levels respectively
- Cantonal dummies are not tabulated

The models which incorporate time effects either as time dummies or as a time trend report a far higher value of the price ratio than the reference model (see section 5.2.1) or the AR model (see section 7.1). In the fixed and random effects models with time dummies, the time dummy dt_7 is significant at the 5% and 1% level respectively. In the random effects model, the dummies for 2001, 2004 and 2005 are also significant at the 10% level. The main effect of incorporating time dummies or a time trend is that the Swiss gasoline price coefficient α_1 and the coefficient of the stock of gasoline powered cars α_4 lose significance in explaining gasoline sales in both the random and fixed effects versions.

Further, the Hausman test for the models with time dummies (and perhaps for the models with a time trend) rather allows trusting in the random effects versions as in the reference models or the AR models). We tend to rule out the model with time dummies for the present purpose, since it is the only one where the coefficient of the Swiss gasoline price α_1 is positive (insignificantly different from zero). However, the price elasticity of Swiss gasoline demand with respect to the foreign gasoline price would then be higher in absolute value than the elasticity with respect to the Swiss gasoline price. Further, we prefer the reference model over the model with a time trend, since the coefficient of the Swiss gasoline price is significantly different from zero in the reference model, which corresponds more closely with the studies mentioned in the literature review.

III Spatial Panel Data Econometrics Using GMM for Static Models

1 Introduction

Spatial econometric models have attracted a substantial increase in attention from empirical researchers during the past years. In fact, from an econometric perspective, spatial data on the location of individuals and firms can introduce important information to the model, since they can affect each other. Neglecting these spatial effects induces spatial correlation in either the dependent variable or the residuals. In the first case, estimation results obtained using non-spatial econometric methods are biased, whereas in the latter, results would be unbiased but inefficient.

The advantages of using such models, therefore, are clear. First, if the coefficient of the spatially lagged dependent variable is positive, coefficient estimates of other variables are biased away from zero when neglecting spatial interdependence, and this bias can be substantial. Second, accounting for spatial spillovers in the residuals allows for the transmission of (e.g. economic) shocks among the individuals, as mentioned by, for instance, .Egger et al. (2005a). Neglecting spatial interdependence in the residuals generally does not lead to biased parameter estimates but to inefficient estimates. Third, the incorporation of a spatially lagged dependent variable in the respective models allows the estimated coefficients to vary among the individuals in a predefined way. As discussed in more detail below, in order to use spatial econometric methods, it is necessary to define what is termed a spatial weighting matrix that defines the spatial relation between the economic units. To define an effective weighting matrix is not always straightforward and can influence estimation results (Stakhovych & Bijmolt 2009).

As discussed by Anselin (2010) and others, research has focused more recently on the development and application of spatial econometric methods for panel data models. An additional benefit of such an approach is that the use of panel data models allows the researcher to account for unobserved heterogeneity but results in additional complications in the model specification. If the panel is unbalanced, the definition of the spatial weighting matrix is more cumbersome, since it will no more be constant over time, as observed by Egger et al. (2005a).

From the econometric point of view, two issues should be considered. The first problem is the endogeneity of the spatially lagged dependent variable and the second the possible presence of spatial correlation in the residuals. In order to deal with these two problems, two different estimation approaches have been developed: maximum likelihood estimation (ML) and instrumental variable estimation or generalised methods of moments respectively (GMM). Both approaches have their merits. On one hand, the derivation of these estimators is somewhat easier in the ML setting, whereas one is forced to rely on normally distributed error terms[14]. This is not the case when one uses a GMM procedure and moreover, sample size is less of a problem[15]. So far, a variety of empirical studies have applied the concepts of spatial econometrics to their research questions using maximum likelihood estimation. In general, the maximum likelihood procedure is more often applied in empirical studies (see Anselin 2010), since commercial software packages such as STATA® or Matlab® incorporate such models.

Spatial dependence can be the source of severe problems in the estimation when using both panel and cross-sectional data. The estimation of the coefficient of spatial lag in the dependent variable does not impose too many difficulties on the researcher. The endogeneity problem arising from the presence of a spatially lagged dependent variable, for instance, can be solved using appropriate instruments in a two-stage least squares regression (2SLS). However, the estimation of the coefficient of the spatial lag in the residuals, or the joint estimation of the coefficients of the spatial lag in the residuals and in the dependent variable, may be cumbersome. Using panel data, these problems are even more persistent, since one is also interested in considering unobserved heterogeneity. Though Kapoor et al. (2007) derived a GMM estimator to jointly estimate the coefficients of the spatial lag in the dependent variable and in the residuals under a random-effects specification, the GMM estimator has not yet been implemented in commercial software packages – which is probably one reason why few studies have applied a spatial GMM estimator using panel data. For the ML approach, one may use the Matlab® toolbox developed by LeSage and Pace (2009). Accordingly, most empirical studies using spatial econometrics in a panel data framework are based on maximum likelihood[16]. However, GMM estimation has certain advantages, as previously mentioned. First, the residuals do not necessarily have to be normally distributed, meaning they may be heteroscedastic. Second, computational power is not as great an issue as in the case of maximum likelihood (see for instance Anselin (2010). Third, recent research points to estimation of models in which more than one endogenous variable (not only the spatially lagged dependent variable) is incorporated in the model. Anselin presumes that the derivation of a maximum likelihood estimate using such a model may be highly cumbersome (Anselin 2010).

In the present paper, we want to estimate a gasoline demand function using aggregate panel data applying spatial econometric methods. We are especially interested in the presentation and

---

14 According to e.g. Lee (2004) this is not true for the quasi-maximum likelihood estimator (QMLE). However to our knowledge, these software packages incorporate the traditional ML estimator.

15 Handling a large sample size and a non-sparse weighting matrix is hardly feasible using ML. In every iteration, the inverse and the determinant of large matrices have to be calculated

16 See e.g. Coughlin & Segev (2000) or Pirotte & Madre (2011) among others. Many studies cited by Anselin (2010) use a ML estimator and not a GMM.

discussion of the GMM estimator developed by Kapoor et al. (2007) and in the preparation and implementation of the respective procedures in STATA®[17].

Further, the existence of spatial dependence is an assumption which has to be tested. Several statistical tests for spatial econometric panel data models have been developed. For instance, Kelejian and Prucha (2001) derived a Moran I test statistic to check for spatial correlation in the residuals robust to the conditional presence of a spatial lag in the dependent variable. Baltagi et al. (2003) introduced a test for serial correlation, spatial correlation and random effects and present statistics for both joint and conditional tests. The test statistics were presented to jointly test the presence of spatial lag dependence and random effects with the respective robust versions. Moreover, Sen and Bera (2011) are the first to derive a test using which the researcher can jointly and conditionally test for spatial correlation in the dependent variable and in the residuals, for random effects and for serial correlation in the residuals. Sen and Bera's statistics are very useful in the sense that the model only needs to be estimated under the null hypothesis, which is simply a pooled OLS version of the model of interest. A further goal of this paper will be to apply and compare several statistical tests on spatial dependence. Since there are various ways to define the spatial weights, we will discuss and compare the results obtained by using different spatial weight matrices.

In this paper, the estimation of a gasoline demand function for Switzerland using a GMM approach uses a panel data set covering 547 Swiss municipalities (out of a total of 2715) over the period 2001-2008. Gasoline sales were collected from the five largest gasoline companies operating in Switzerland (covering about 60% of overall sales). Since we have no information about the total number of stations in the municipalities and therefore are not able to estimate an aggregate demand function, our dependent variable is average gasoline sales per station in a municipality. Naturally this variable is strongly influenced by car traffic, which in turn is connected with spatial characteristics such as urbanisation, population density and geography. Moreover, Swiss municipalities are relatively small units, and traffic in one municipality is supposed to influence sales strongly in the neighbouring ones. Therefore, the goal is to analyse gasoline sales in the 547 Swiss municipalities considering spatial interdependences both in the dependent variable and the residuals (to account for spatial correlation in unobserved characteristics). The purpose of the application, therefore, is to apply the GMM procedure implemented in STATA® to analyse Swiss gasoline demand taking spatial interdependencies into account. From this, we will be able both to estimate the price elasticity of Swiss gasoline demand and to provide information about its spatial variation among the municipalities. We will see that this variation is quite substantial, and some municipalities are assigned a far higher elasticity compared an estimation using a non-spatial approach. This may provide interesting

[17] STATA® offers the possibility to analyse cross-sectional spatial models with either a spatial lag in the dependent variable or in the residuals – but not both together – using a ML approach. More recently, a GS2SLS (generalised spatial two-stage least squares) procedure has become available for the estimation of cross-sectional spatial models where the spatial lag in the residuals and in the dependent variable can jointly be estimated. The authors, Drukker et al. (2011), similarly describe the merits of using GMM, such as the possibility of including further endogenous right-hand side variables in a spatial model, as we have done.

information to policymakers, because some regions emerge as more elastic with respect to the gasoline price than others.

The main contributions of this paper are that, first, it offers a detailed reconstruction of the spatial GMM estimator as presented in Kapoor et al. (2007) and describes the typical issues that arise when specifying a spatial econometric model, such as the definition of a spatial weighting matrix. Second, we implement the GMM estimator, all statistical tests previously mentioned and some auxiliary codes supporting the generation of the spatial weighting matrix into STATA®. Third, we estimate a gasoline demand function for Switzerland considering spatial correlation and unobserved heterogeneity. Very few studies have used the GMM procedure developed by Kapoor et al. (2007) and, to our knowledge, none of those have applied the statistical tests mentioned at the same time. The present paper is intended to fill this gap in the literature.

This paper is structured as follows: In section 2, two papers using the GMM approach for spatial panel data models are discussed, as well as studies analysing gasoline demand. In section 3, we concentrate on procedural methods such as the specification of a spatial econometric model using panel data, the statistical tests and the proper definition of the spatial weighting matrix. Further, a reconstruction of the GMM estimator as described by Kapoor et al. (2007) is provided. In chapter 4, the model explaining gasoline demand in Swiss municipalities is specified with a dependent variable being average gasoline sales per gasoline station in a municipality. We then shortly describe the data, since it differs somewhat from that used in the first part of this dissertation. After a detailed presentation and discussion of the test statistic referring to spatial dependence, we present the results and conclude. The Appendix in Chapter 6 gives an overview of what we have done and presents a Monte Carlo simulation to show the performance of the procedures implemented.

2 Literature Review

One focus of the present paper is to provide an empirical application which aims to explain Swiss gasoline demand at the municipal level by taking spatial correlation into account. Therefore, the literature review discusses two types of paper. The first category includes studies dealing with the estimation of gasoline functions. For the second, we present papers that perform empirical analyses on the gasoline market using spatial econometric methods. Thus, the first part of this review deals with two studies on gasoline demand using panel data, in which we focus on the proper specification of gasoline demand functions rather than the econometric issues when using spatial panel data models.

In the second part of this review, we turn to two studies which analyse the wholesale gasoline market. They are particularly interesting for our topic since they provided significant theoretical contributions about spatial GMM models and applied those models to the gasoline market; the technical details which are relevant for our spatial GMM estimator are discussed in subsection 3.4. For the sake of completeness, we add one study which applies a spatial maximum likelihood estimation procedure to analyse gasoline demand.

2.1 Review of Gasoline Demand Studies

Pock (2010) analysed data from 14 European countries over the period 1990-2004 to estimate a dynamic model specification for gasoline demand. His main assertion is that many previous studies may suffer from a bias in estimated income and price elasticities of gasoline demand due to the omission of diesel-powered cars or non-distinction between gasoline- and diesel-powered cars. The diesel share of total passenger cars has been increasing for all countries considered in the study's sample period (including Switzerland). Gasoline consumption is modelled on the basis of an average vehicle's utilisation, its average fuel efficiency and the total stock of cars in use. A two-way error component model is employed to specify a dynamic demand equation for gasoline in which gasoline consumption per gasoline- powered car is used as the dependent variable and the number of gasoline- and diesel-powered cars by driver, real per capita income and the real gasoline price are used as regressors. Nine common dynamic panel estimators are applied to the panel data set. It is found that the standard within estimator and its bias-corrected version, LSDVc, yield reasonable estimates in terms of a positive income elasticity and a negative price effect on gasoline consumption. However, the coefficient estimates are found to be somewhat lower in absolute terms, which is partially accredited to the omitted-variable bias in other studies (e.g. the stock of diesel-powered cars). A substantial contribution of the study is that it shows the necessity to include the stock of diesel- powered cars when explaining the demand for gasoline. On the other hand, one could criticise the fact that, in a partial adjustment model, the capital stock, here the stock of gasoline and diesel-powered cars, should not usually be included, since it then is considered as quasi-fixed. It reports short- and long-run price and income elasticities for gasoline demand. The short-run price elasticity was found to be -0.1, the long-run -0.54, whereas the short-run income elasticity was found to be 0.23 and the long-run 1.3.

Banfi et al. (2005) performed an analysis of the gasoline demand in Switzerland using panel data covering eight years of sales of gasoline stations in 315 municipalities. The goal of this analysis, as already discussed in the first part of this dissertation, was to quantify cross-border purchasing of Swiss gasoline by foreigners induced by the price differentials across the border. The authors specified average gasoline sales per gasoline station in a municipality to depend on the weighted price ratio with the distance from the border and explanatory variables as used by Pock, such as the population, the stock of gasoline- and diesel-powered vehicles, the per capita income or the number of commuters. For the estimation of this linear log-log demand model, Banfi et al. (2005) employed static random and fixed effects estimators. The main results were that the price ratio across the border has a significant impact on Swiss gasoline demand. The inclusion of the weighted price ratio with the distance further allows the price elasticity of Swiss gasoline demand to vary with varying distance from the border. A further novelty compared to the study by Banfi et al. (2005) was that the distance within which 'gasoline tourism' takes place could be defined endogenously by the model and not by the researcher. They found a critical distance of 40km from the border and an average price elasticity of -0.65. However, this work should now be extended using data for all Swiss cantons. Subsequently, we want to analyse the demand for Swiss gasoline ruling out cross-border purchasing by foreigners and employ a spatial econometric approach. The inclusion of a spatially lagged dependent variable also will allow the coefficients estimated to vary among the municipalities, as we will see later[18]. The first paper of this dissertation is based on the analysis by Banfi et al.; the model specification is slightly different and the econometric analysis is more detailed.

[18] Dahl (2012) analysed the average price elasticity from a variety of gasoline demand studies. A test for the equality among those elasticities was strongly rejected even for 'similar' countries or regions. It further became clear from Baltagi et al. (2003) that the use of heterogeneous estimators does not necessarily improve results with respect to coefficient estimates or forecasts. Therefore, a spatial (GMM) estimator seems to be a valid alternative for the present purpose.

2.2 Review of Spatial Econometric Studies

In this subsection, we present three papers which are related to the gasoline market and apply spatial econometric methods using panel data. The first two papers apply a spatial GMM approach, while the third applies a maximum likelihood procedure. The number of studies in economics that use spatial methods for panel data is not large, but has been growing in recent years (Anselin 2010).

Kapoor (2007) uses a GMM spatial econometric approach to investigate price competition in the US wholesale gasoline industry. He argues that firms compete – at least locally – for customers and hence interact strategically. Gasoline retail prices in a region are modeled to depend on explanatory variables such as the retail price in 'close' (neighbouring) regions, on the population and on per capita income in neighbouring regions. These regions (289 in total) form a cross-sectional data-set. Prices are assumed to depend on the own region's population, per capita income and spot prices for the respective regions. For the spatial econometric analysis, Kapoor uses a weighting scheme where only the very nearest terminal was considered to be the neighbouring region[19]. The generalised spatial two-stage least squares approach suggested by Kelejian and Prucha (1998) is used to estimate the model[20]. Furthermore, it is one of the few studies which actually tests for either the presence of spatial correlation in the dependent variable or in the residuals. Kapoor applied the Moran I test statistic as described in Kelejian and Pruch (2001). The main results of the dissertation are that, first, the price of a terminal is negatively related to the quantity sold to the marginal customer of that terminal. Second, the price of a given terminal is positively correlated with the price of its neighbouring terminal and hence price competition exists so that terminals react to the prices of each other.

The study conducted by Egger et al. (2005a) applied spatial GMM estimation methods on a panel data set in order to analyse spatial tax competition for goods such as beer wine, gasoline and cigarettes among US states from 1975 to 1999. They used the new insights of Kelejian and Prucha (2007) (at that time in press) and extended the approach to estimate a spatial model for panel data with spatially correlated residuals and a spatially correlated dependent variable when the panel is unbalanced; the moment conditions then look different from those elaborated by Kelejian and Prucha. The taxation of the goods in question is estimated separately, treating the individual effects as either fixed or random. This paper argues that taxation policies for these goods follow a spatial pattern, since local jurisdictions can increase their tax base and attract cross-border shoppers. Therefore, taxation of the commodities is modelled to depend on the spatially weighted commodity tax rates of neighbours and a set of exogenous explanatory variables. The study identifies a significant and positive coefficient of the spatially lagged tax for all goods in either the FE or RE setting. A second point is that the coefficient of the spatially

---

[19] In fact, many other types of spatial weighting schemes would be applicable. On the other hand, instead of specifying the weighting matrix only with the closest neighbour, Kapoor provided a robustness test and also used weighting matrices with more neighbours than only the very closest.

[20] Kapoor also underlined the advantages of the GMM procedure over maximum likelihood estimation that it is computationally relatively easy even for large panels and that the results do not depend on the assumption of a normally distributed error term.

lagged residuals is significantly negative in all models, meaning that unobserved shocks in a state impede tax competition in others. Findings from a counterfactual simulation show that spatial effects are very strong, especially in the case of gasoline taxation. In terms of specification testing, a Moran I test is applied to test the presence of spatially correlated residuals. A Hausman test statistic discriminates between the RE and FE specification. However, the study does not present a test for the presence of spatial correlation in the dependent variable.

An interesting study for the present purpose was conducted by Pirotte and Madre (2011). They used a panel data set of 21 French regions over 17 years to analyse elasticities of car traffic. They argue that traffic (and therefore fuel quantity consumed, since the dependent variable car traffic is constructed with the fuel quantity consumed) in a region is a matter of spatial dependence. The negligence of spatial correlation in the model specification may either lead to biased coefficient estimates, if spatial correlation is present in the dependent variable, or to misleading inference, if spatial correlation is present in the residuals, or to both. The study identifies positive and significant spatial correlation in both the residuals and the dependent variable. The estimation method used is maximum likelihood and moreover, the spatial lags in the residuals and the dependent variable are not estimated jointly. The results, however, are tested against two different definitions of a spatial weighting matrix, and both approaches give similar results. As a main result, Pirotte and Madre conclude that the price elasticity of car traffic varies among the regions between -0.1 to -0.15. For us, this shows that analysing gasoline demand using a spatial econometric approach is an appropriate approach to the issue.

This short review of studies on the gasoline market using panel data and spatial econometric methods shows that few studies have applied spatial GMM procedures. Moreover, almost all of the studies that do suffer from the absence of proper statistical tests for spatial effects. The present application intends to fill this gap, first, by describing the properties of the GMM procedure developed by Kapoor et al. (2007). Second, we want to clarify about the different possibilities of using spatial weighting matrices. Third, we will discuss possible testing procedures and apply them to our data set.

3 Overview of Spatial Econometric Approaches

3.1 Different Types of Spatial Models

In the following subsection, we describe the econometric specification of spatial models for cross-section and panel data. It should be noted that this short overview cannot be exhaustive.

In the context of spatial model specification, Elhorst (2010) presents an overview of first-order spatially lagged models and addresses spatial correlation in a variety of possible manners: in the dependent variable, in the independent variable(s) or in the residuals. He starts from the most general model, the Manski model, and then discusses all possible specifications up to ordinary least squares (OLS). However, the discussion only addresses spatial dependence for models dealing with cross-sectional data. Nonetheless, the elaborations provided remain basically valid and are extendable to panel data models, which will be described in detail in subsection 3.3. Figure III-1 shows the different types of spatial models according to Elhorst.

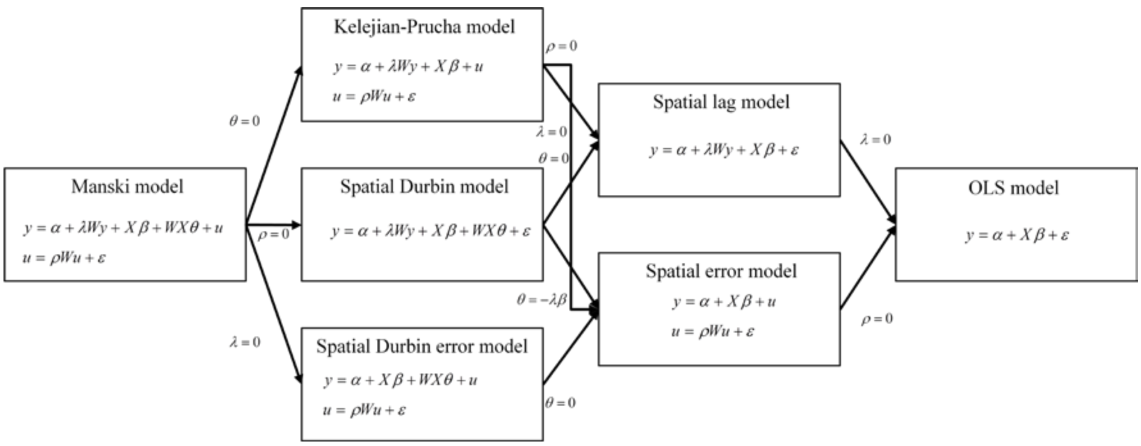


Figure III-1: The relationship of spatial dependence models for cross-sectional data (Elhorst 2010)

The most general model (for the cross-sectional case) is the Manski model, which is provided below in matrix notation:

$$\mathbf{y} = \alpha \mathbf{e}_{\mathrm{N}} + \lambda \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \mathbf{u}$$
$$\mathbf{u} = \rho \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}$$

α, λ and ρ are scalars. $\mathbf{W}$ is the spatial weighting matrix of dimension $\mathrm{N}\times\mathrm{N}$, $\mathbf{X}$ is a matrix containing exogenous regressors and is of dimension $\mathrm{N}\times\mathrm{K}$. $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are vectors of dimension $\mathrm{K}\times 1$, $\mathbf{y}$, $\mathbf{u}$ and $\boldsymbol{\varepsilon}$ are vectors of dimension $\mathrm{N}\times 1$, and $\mathbf{e}_{\mathrm{N}}$ is a $\mathrm{N}\times 1$ vector of ones. The model indicates that spatial dependence can occur due to three different types of spatial correlation, namely through an endogenous effect such that the observation y_i of a certain (e.g. geographical) location i depends on the observations of some associated locations $\mathrm{j}\neq\mathrm{i}$ where this association is defined by the matrix $\mathbf{W}$. This type of spatial correlation is usually referred to as the spatial lag (in the dependent variable) λ. The second type of spatial dependence can be explained by exogenous interactions, meaning that the observation y_i of a location can be explained by the explanatory (exogenous) variables x_{ik} of associated locations through the coefficient vector $\boldsymbol{\theta}$. Finally, spatial correlation can also appear in the error term, which means that the observation of a location depends on unobserved characteristics in associated locations, which usually is referred to as spatial autocorrelation (in the residuals) ρ. Again, it should be noted that, while these descriptions were elaborated by Elhorst for the cross-sectional case, they are equally valid for panel data. Generally, data are termed spatially lagged if one refers to data observed in neighbouring locations. The concept of neighbourhood is captured by the definition of the spatial weighting matrix $\mathbf{W}$, which is described below. Accordingly, the spatially lagged version of the dependent variable vector $\mathbf{y}$ is $\mathbf{W}\mathbf{y}$, the spatially lagged version of the independent variables' matrix $\mathbf{X}$ is $\mathbf{W}\mathbf{X}$, and the spatially lagged version of the residual vector $\mathbf{u}$ is $\mathbf{W}\mathbf{u}$. $\boldsymbol{\beta}$ is a vector of coefficients relating the observation y_i in a certain location to the independent variables x_{ik} of that location. The spatial dependence defined by the matrix $\mathbf{W}$ may be geographical, but may also refer to any kind of network, which does not necessarily have to be related to geographical terms but, for instance, to economic relations such as trade among countries. The w_{ij} element of $\mathbf{W}$ is zero if the cross-sectional unit i is not connected (i.e., spatially related) with cross-sectional unit j and non-zero if there is dependence. Dependence may be accounted for in several ways. For instance, w_{ij} is one if unit i is a neighbour of j and vice versa. Another approach would be to define neighbourhood in such a way that the units involved depend on each other, given they are located at a certain distance from each other, speaking in geographical terms. One could even say that all units are dependent on each other but that close units have a stronger impact than more distant ones. In such cases, w_{ij} typically is defined as a decaying function of some distance measure. The Manski model, the most general one, can be re-arranged so that

$$\mathbf{y} = (\mathbf{I} - \lambda \mathbf{W})^{-1}(\alpha \mathbf{e}_{N} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \mathbf{u})$$
$$\mathbf{u} = (\mathbf{I} - \rho \mathbf{W})^{-1} \cdot \boldsymbol{\varepsilon}$$

where $\mathbf{I}$ is the identity matrix of dimension $\mathrm{N}\times\mathrm{N}$. So it is obviously a necessary condition for this data generating process (DGP) that the matrices $(\mathbf{I} - \lambda\mathbf{W})$ and $(\mathbf{I} - \rho\mathbf{W})$ are non-singular. It can be shown that this condition is satisfied as long as the parameters ρ and λ lie in the interval of the inverse of the largest negative and the inverse of the largest positive real eigenvalue of $\mathbf{W}$, so that $(\lambda, \rho) \in (1/\omega_{\min}, 1/\omega_{\max})$: see, for example, Kapoor et al. (2007). According to this, the average of the row-sums of the matrices $\mathbf{W}$, $(\mathbf{I} - \rho\mathbf{W})$ and $(\mathbf{I} - \lambda\mathbf{W})$ should be bounded in absolute value as N tends to infinity. This is achieved by defining the elements of the spatial weights matrix such that they decrease with increasing distance. The approach to choosing a proper weighting matrix has been an intensively discussed topic in the literature: see, for instance, LeSage & Pace (2009) or Stakhovych & Bijmolt (2009) among others. However, a few difficulties arising from the specification of spatial interdependences should also be mentioned here. First, the researcher has to decide how to normalise the spatial weighting matrix and thus whether to incorporate distance as an absolute or as a relative measure. The normalisation in turn forces the eigenvalues of the matrix to be bounded and thus also the spatial autoregressive parameters. Second, the elements of the weighting matrix should be defined such that they decrease with increasing distance, but the functional forms available to achieve this are manifold. One could impose a regime under which the elements of the matrix are measured by $1/\sqrt{d}$, $1/d$, $1/d^2$, $1/d^k$, $e^{-k \cdot d}$ and so on. Third, another remaining question is whether to allow for spatial correlation in a way such that all units influence each other or to rather introduce a cut-off distance where it is assumed that more distant units' observations just do not depend on each other. One therefore could impose a regime where first-order spatial correlation only occurs within the three or four closest neighbours, or only within a range of some five or ten kilometres. In several Monte Carlo studies, it has been shown that if spatial dependence is strong, the costs of choosing the "wrong" spatial weighting regimes are low: see, for instance, Stakhovych and Bijmolt (2009). Moreover, one may compare the different weighting regimes by goodness-of-fit criteria after fitting the models.

Turning back to the Manski Model, Elhorst (2010) has observed that the parameter estimates cannot be interpreted as a distinct effect, since the exogenous and endogenous effects cannot be separated from each other. In such a case, the model needs to be adjusted by excluding at least one of the spatial parameters so that the model is fully identified. The best practice then is to assume the absence of spatial correlation in the error term. Even if the true data-generating process was a spatial error model, the omission of spatial correlation in the error term only affects the efficiency of the estimates, but not their consistency. On the other hand, the omission of the spatial lag in the dependent or in the independent variable would result in inconsistent estimates if the true data-generating process includes spatial correlation in other variables than the error term. Accordingly, the Spatial Durbin model (see Figure III-1) has to be preferred over the Kelejian-Prucha and the Spatial Durbin error model, since the coefficient estimates of that model are not biased or inconsistent. More advantageously, if the true DGP is the Spatial error model, the Spatial Durbin model produces correct coefficient estimates and standard errors, since the Spatial error model only represents a special case of the Spatial Durbin model, as depicted in Figure III-1.

The question arising from this consideration is whether to apply a general-to-specific or a specific-to-general approach. According to LeSage and Pace (2009), the best would be to start with the Spatial Durbin model, whereas Elhorst suggests starting with OLS estimates and then checking for spatial evidence using likelihood ratio tests. The strategy described by Elhorst is to test the null hypotheses $\mathrm{H}_0{:}\boldsymbol{\theta} = \mathbf{0}$ and $\mathrm{H}_0{:}\boldsymbol{\theta} = -\lambda\boldsymbol{\beta}$. If both hypotheses are rejected, the Spatial Durbin model should be chosen. If the first hypothesis is accepted, then the Spatial lag model is probably the appropriate to describe the data. If the second hypothesis is accepted, then the Spatial error model should be used, provided that the robust Lagrange-multiplier tests also point to the models of interest. If this is not the case, then a Spatial Durbin model should be used. On the other hand, if the estimated OLS model is not rejected in favour of the Spatial Lag, the Spatial Error or the Spatial Durbin model, it should be re-estimated with the inclusion of lagged independent variables $\mathbf{WX}$ or at least a selection of those. If a test of $\mathrm{H}_0{:}\boldsymbol{\theta} = \mathbf{0}$ cannot be rejected, then OLS is the appropriate model to describe the data. The most important message from the review above hence is that spatial models can be defined such that the residuals, the dependent variable, the independent variables or a combination of those are spatially correlated.

More recently, attention has turned to the estimation of spatial panel data models, which basically deal with the same models as depicted in Figure III-1 but allow for fixed or random individual effects. These models have already proven their wide field of applications, such as in public economics (e.g. Egger and Larch (2008) or Egger et al. (2005a)), demand for goods (as in Baltagi and Li (2006), transport economics (e.g. Pirotte and Madre (2011)), and many others. These studies all use static models. Anselin et al. (2008) describe the evolution of spatial panel data econometrics referring to the fixed and random effects estimation of a spatial lag model or a spatial error model and underline the important contributions of Elhorst (2010), Kapoor et al. (2007), and Baltagi et al. (2009), to mention only a few. However, Anselin (2010) has observed that more recent research interest has increasingly focussed on spatial-temporal models. For instance, Lee and Yu (2010) develop an estimator which allows for spatial dependence in the dependent variable or the residuals and for a time-lagged dependent variable. Monte Carlo results show that the omission of time effects can have serious consequences for the spatial parameters, namely that estimates for λ are downwardly biased and those for ρ are upwardly biased. However, spatial-temporal models are beyond the scope of this paper. We turn to a more extensive discussion of spatial panel data models in subsections 3.3 and 3.4 where the Kelejian-Prucha model is discussed in detail.

3.2 The Spatial Weighting Matrix W

3.2.1 The Concept of Neighbourhood

Consider again the Manski model describing the most general spatial data generating process

$$\mathbf{y} = \lambda \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \mathbf{u}$$
$$\mathbf{u} = \rho \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}$$

From the estimation equation of the Manski model, it is obvious that the $N \times N$ matrix $\mathbf{W}$ enables the researcher to incorporate spatial dependence in the econometric equation. However, the elements of $\mathbf{W}$ have to be exogenously defined and accordingly estimation results may reveal sensitivity to the specification of the matrix. From the equation above, it can be seen that spatial dependence in, for example, the dependent variable of an observation i with respect to the dependent variable of an observation j can be written as $\partial y_i / \partial y_j = \lambda \cdot w_{ij}$, where w_{ij} is the respective element of $\mathbf{W}$. First, the value of the spatial lag parameter λ depends on the specification of the weighting matrix. Second, the strength of spatial correlation not only depends on λ but also on the elements of the weighting matrix. Hence, a proper specification of the weighting matrix is essential. Intuitively, and as described in several econometric textbooks such as LeSage and Pace (2009), it is reasonable to assume that, provided spatial dependence exists, 'close' units reveal stronger spatial dependence than more distant ones. A question which then arises is how to define proximity among spatial units. By convention, the elements w_{ij} of $\mathbf{W}$ are zero if $i=j$, which means that no element is considered to be a neighbour of itself and that all diagonal elements are zero ($w_{ii} = 0, \ \forall i$). Further, w_{ij} is zero if element j is not 'close' to element i and strictly positive otherwise. In many applications researchers define spatial units to be close to each other if they are located within a certain distance range from each other, e.g. 5 or 10 kilometres. If spatial units' locations are two-dimensional (areas), proximity, or at least neighbourhood, can be defined among spatial units which share a common border or a common point and common border (sometimes referred to a 'queen' or 'rook' contiguity matrix). Typically, data is not entirely available over all spatial units (e.g. countries, districts or municipalities). Accordingly, a 'strict-neighbourhood' definition may impose the problem of spatial units with no neighbours, which in turn renders estimation of spatial econometric models infeasible, since the weighting matrix would be singular. As previously mentioned, a precondition for interpretation and estimation of spatial models is the invertibility of the matrices $(\mathbf{I} - \lambda\mathbf{W})$ and $(\mathbf{I} - \rho\mathbf{W})$, which is not possible for a matrix $\mathbf{W}$ of incomplete rank. If spatial units are irregularly distributed in space, the researcher could consider, for example, the three, five or ten closest neighbours to depend on the spatial unit considered. However, the problem of such a matrix specification then would be that the resulting weighting matrix would not necessarily be symmetric, since if unit i was a neighbour of j in this context, unit j would not necessarily be one of unit i. For an intuitive illustration, consider the following maps below:

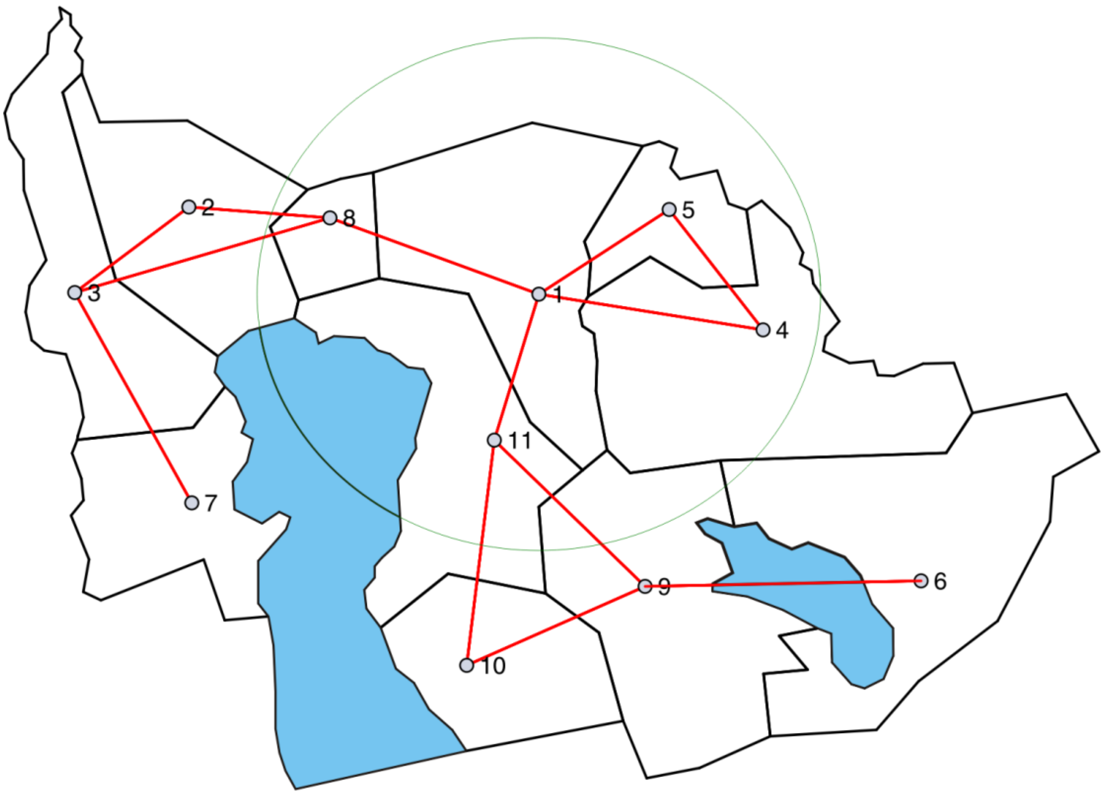


Figure III-2: Spatial dependence as a result of a maximum distance measure

Figure III-2 shows spatial dependence for spatial units on a map when proximity is defined as a measure of distance. The circle represents the maximum distance range (drawn around spatial unit 1) within which the other units are considered to be neighbours. Note that the spatial units 2 or 9 are not assigned a neighbourhood characteristic although the respective areas share a common border. If the distance threshold was increased, unit 9 would become a neighbour to unit 1, but perhaps also unit 2 or unit 10, which do not share a common border with unit 1. The distance threshold could even be increased until every spatial unit is a neighbour of all other units. Normally, the elements of the spatial weighting matrix are defined to decrease with increasing distance, and it is left to the researcher to decide whether such a specification would make sense. As previously mentioned, it is easy to see that a spatial weighting matrix representing a situation as depicted by Figure III-2 is symmetric if the elements w_{ij} are any functional form of the distance between the spatial units. The symmetry arises from the fact that if unit i depends on unit j, the opposite is the case too.

Another definition of proximity is to define, for instance, the. three closest units of any unit as neighbours. In this case, spatial dependence among units would result in a different situation.

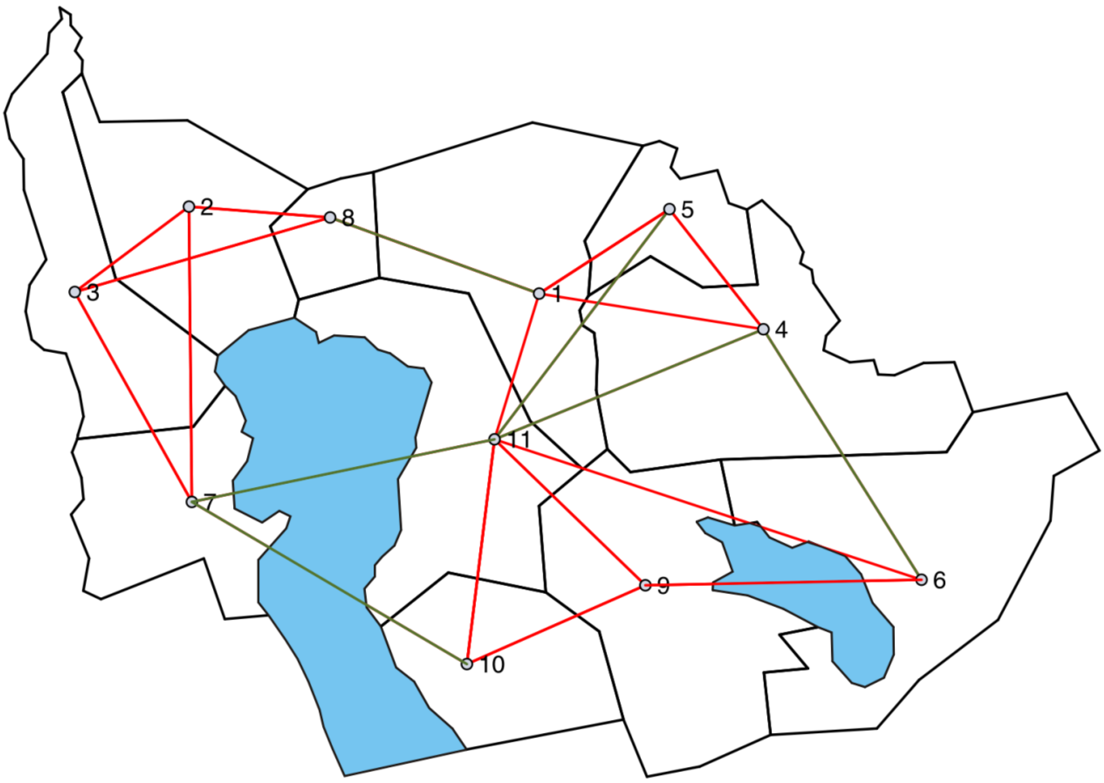


Figure III-3: Spatial dependence with three closest neighbours

Figure III-3 shows spatial dependence when the spatial units all are assigned the same number of closest neighbours (here three). This definition is justified by the fact that the strength of spatial dependence can generally be assumed to decrease with increasing distance and accordingly, only a certain number of shortest connections from a spatial unit to others are considered. It is important to note that the term neighbour can be misleading. Considering spatial unit 6, one can see that unit 11 is not a contiguous neighbour of 6, but nonetheless is defined as being connected to it. Further, some points reveal more than three connections, yet only the three closest units are considered. This is due to the fact that, if one unit is relatively close to another, meaning spatially dependent, the opposite is not necessarily true that the unit in question belongs to the subset of the three closest neighbours of the other unit (consider the green coloured lines in Figure III-3). For instance, spatial unit 4 is a neighbour to unit 11 but not vice versa; likewise, spatial unit 1 is a neighbour to unit 8 but the inverse is not the case. Defining spatial dependence in this manner may not make sense in all applications.

One useful method to create a contiguity matrix of a random allocation of spatial units on a map is what is termed the Delaunay triangulation of the space. The goal of the procedure is to create a border around each spatial unit according to their arrangement in space. Figure III-4 shows the resulting spatial dependence among units after triangulating the map. First, a convex envelope around the spatial units is created. From the borders of the envelope, a new point has to be found such that the circumference of the triangle generated only contains the points of the triangle itself and no other spatial units. The algorithm is repeated for the edges of each triangle until the space is completely triangulated. The red lines in Figure III-4 show the triangles described and simultaneously represent the spatial dependences (neighbourhoods) among the spatial units. For clarification, the perpendicular bisector of each triangle side can be drawn (represented by the thick black lines: the procedure is called Voronoi tessellation). The resulting new space has the characteristic that each polygon (sometimes termed a Thiessen polygon) builds an envelope around the spatial units without considering the original borders (which are drawn as thin black lines in Figure III-4). Spatial units are considered neighbours if the Thiessen polygons share a common border or common vertices. It can be seen in Figure III-4 from the original borders of spatial units 1 and 9 that they are neighbours. After construction of the Thiessen polygons, spatial unit 1 is a neighbour to units 4, 5, 8 and 11 but not to unit 9. From this example, one can see that the algorithm seems to realise a concept of 'closeness' or 'neighbourhood' very well.

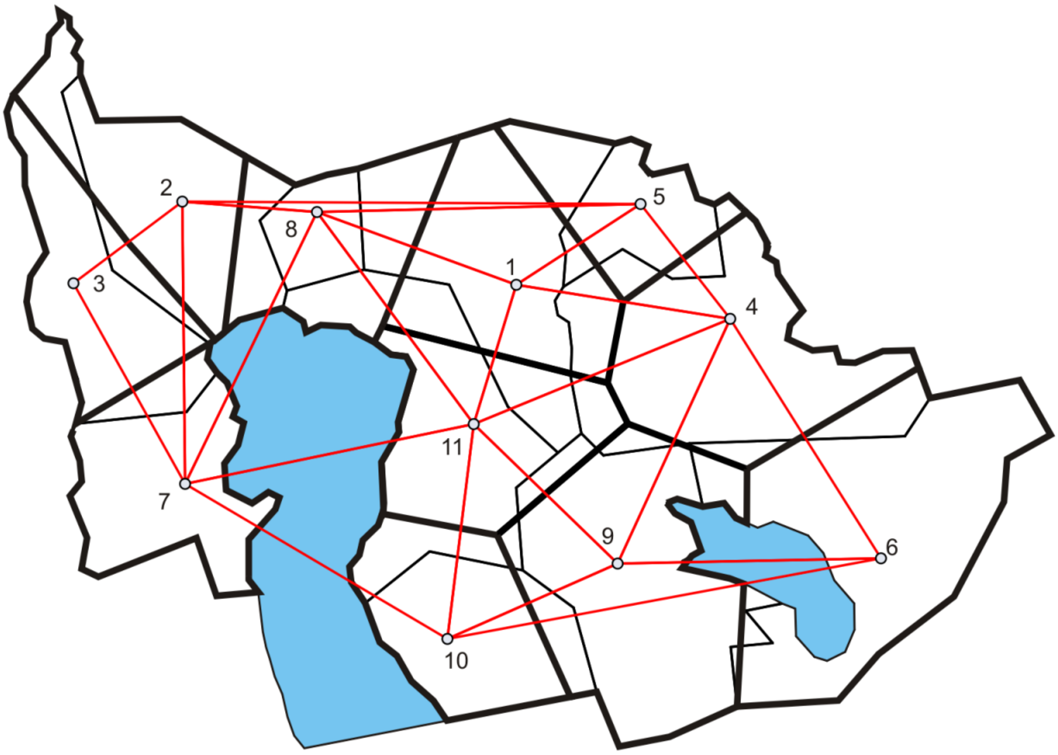


Figure III-4: Spatial dependence as a result of a Delaunay triangulation

The strength of this procedure originates from a simple fact. Very often, panel data are available for spatial entities which have a predefined border, such as a municipality. However, data are seldom available for each spatial unit. This triangulation algorithm results in an exhaustive partition of space and thus leaves no entity unassigned to data, a fact which becomes visible in Figure III-9.

A program code for the triangulation of a given space, for the construction of the Thiessen polygons and the calculation of the spatial weighting matrix has been realised in STATA® and is executed in reasonable computation time up to some 500 spatial units. Details can be found in the Appendix.

3.2.2 Elements of the Spatial Weighting Matrix

A researcher defining a concept of neighbourhood suited to the research question at hand is confronted with the task of defining the matrix entries w_{ij}, which are predefined as strictly positive and non-zero if spatial unit i is a neighbour to spatial unit j and zero otherwise.

In addition to choosing a functional form for the elements of the matrix, an appropriate way of normalising the matrix should be taken. Basically, two approaches are common. The first is called row normalisation; in this approach, the elements of w_{ij} are divided by the respective row sums of $\mathbf{W}$. In such a setting, a spatially lagged vector $\mathbf{z}_s = \mathbf{Wz}$ corresponds to a weighted average of $\mathbf{z}$ in the neighbouring regions. It is important to note that, in this row-normalised setting, even if the spatial weights w_{ij} are related to a distance measure, the new matrix entries become *relative* measures. So if a number of spatial units are very close to each other and an additional one is very distant, then row normalisation has the effect that the row sum for the distant spatial unit equals unity, no matter what the actual distance entries are. A second approach to normalisation is termed maximum row-sum normalisation. In this procedure, every element of $\mathbf{W}$ is divided by the maximum row sum of $\mathbf{W}$. In this case, a spatially lagged vector $\mathbf{z}_s = \mathbf{Wz}$ still corresponds to a weighted average of $\mathbf{z}$ in neighbouring regions – but units which are surrounded by more close neighbours than others receive higher weights. Accordingly, maximum row-sum normalisation accounts for the *absolute* distances at which the units are located from each other.

To explain this more intuitively, consider the following situation, in which three spatial units are located on a line segment and unit 3 is twice as far located from unit 2 as unit 2 from unit 1.

Figure III-5: Spatial interaction on a line segment

At first, we fill the spatial weights matrix $\mathbf{W}$ with elements that decrease with distance. One scheme would be to take the inverted distances as the appropriate measures. Moreover, w_{ij} is set equal to zero if unit i is not connected to unit j. This results in

$$\mathbf{W} = \begin{pmatrix} 0 & 1/d & 0 \\ 1/d & 0 & 1/(2d) \\ 0 & 1/(2d) & 0 \end{pmatrix}$$

Now, the row-normalisation procedure can be applied by dividing the elements of $\mathbf{W}$ by the respective row sums. Recall that the normalisation methods force the parameter values of ρ and

λ to lie within the interval of $(-1,1)$. The row-sum vector $\mathbf{rs}$ is obtained by multiplying the spatial weights matrix with a vector of ones, which yields

$$\mathbf{rs} = \begin{pmatrix} 1/d \\ 3/(2d) \\ 1/(2d) \end{pmatrix}$$

Then, the row-normalised spatial weights matrix $\mathbf{W_n}$ is

$$\mathbf{W_n} = \begin{pmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 0 & 1 & 0 \end{pmatrix}$$

Two things should be noted here. First, all rows of the transformed matrix sum up to one. Second, the spatial weights allocated to the first unit only refer to the second unit and accordingly $w_{12} = 1$. But the same weight is allocated to the third unit, which is twice as far away from the second unit than the first one. Accordingly, this is a setting where *relative* distance matters. Nonetheless, the weights allocated for the second unit are distance related and allocate more weight to unit one than to unit three, since the latter is closer. In contrast, the maximum row-sum normalised matrix $\mathbf{W_m}$ would be

$$\mathbf{W_m} = \begin{pmatrix} 0 & 2/3 & 0 \\ 2/3 & 0 & 1/3 \\ 0 & 1/3 & 0 \end{pmatrix}$$

Two points should also be noted here. First, only the second row sums up to one, since it has the maximum row sum in the original situation. But the spatial weights allocated to units 1and 3 have changed compared to the preceding situation. Since unit 1 is closer to the second unit, it receives more spatial weight than the third. Accordingly, maximum row-sum normalisation leads to a setting where *absolute* distance matters.

The interpretation of spatial dependence may be rather tricky. In the present setting, the weighting matrix was filled according to first-order dependence, meaning that the elements w_{ij} are equal to zero if and only if the spatial units concerned are directly connected with each other. Assuming there is a shock to a variable belonging to the second unit, units 1 and 3 are affected in a first stage, as described by the transformed spatial weights matrix. But of course, the shock then is fed back to the second unit again, which can be described by multiplying the (e.g. row-normalised) matrix with itself, which yields

$$\mathbf{W_n^2} = \mathbf{W_n} \cdot \mathbf{W_n} = \begin{pmatrix} 2/3 & 0 & 1/3 \\ 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \end{pmatrix} \text{ and } \mathbf{W_n^3} = \mathbf{W_n} \cdot \mathbf{W_n} \cdot \mathbf{W_n} = \begin{pmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 0 & 1 & 0 \end{pmatrix}$$

The matrix $\mathbf{W_n^2}$ describes second-order spatial dependence. For instance, it can be seen in the first row, that unit 1 is a second-order neighbour of itself and of unit 3. Due to the simplicity of the present example, the third-order dependence matrix $\mathbf{W_n^3}$ is equivalent to the first-order dependence matrix. For the maximum row-sum normalised matrix, second- and third-order dependences would be

$$\mathbf{W_m^2} = \mathbf{W_m} \cdot \mathbf{W_m} = \begin{pmatrix} 4/9 & 0 & 2/9 \\ 0 & 5/9 & 0 \\ 2/9 & 0 & 1/9 \end{pmatrix} \text{ and } \mathbf{W_m^3} = \mathbf{W_m^2} \cdot \mathbf{W_m} = \begin{pmatrix} 0 & 10/27 & 0 \\ 10/27 & 0 & 5/27 \\ 0 & 5/27 & 0 \end{pmatrix}$$

The difference from the previous case is obvious. Here too, the third-order spatial dependence matrix is of the same form as the first-order matrix (due to the arrangement of the spatial units). However, since absolute distance matters in this situation, the entries of the third-order dependence matrix are no longer the same as those of the first-order matrix. If the relation between the matrix element w_{ij} and the corresponding distance d_{ij} had been chosen such that w_{ij} decreased more strongly in d_{ij}, the elements of $\mathbf{W_m^2}$, $\mathbf{W_m^2}$ would converge to zero faster. Accordingly, the choice of the functional form and thus the choice about how fast the elements w_{ij} should decrease with the distance determines the way in which higher order spatial dependences are weighted compared to first-order dependence. For clarification, consider the maximum row-sum normalised matrix $\mathbf{W_m}$ and Figure III-5. In this case, the first row of the matrix of second-order spatial dependence, $\mathbf{W_m^2}$, measures the second-order impacts. As can be seen, spatial unit 1 (first column) has a second-order impact on itself (over unit 2) and spatial unit 3 (third column) has a second-order impact on unit 1 (and over unit 2). Taking all feedbacks into account, a power series of the spatial matrix over an infinite number of feedback loops has to be evaluated, namely

$$\lim_{n\to\infty}\left(\mathbf{I} + \lambda\mathbf{W}^1 + \lambda^2\mathbf{W}^2 + \lambda^3\mathbf{W}^3 ... + ... \lambda^n\mathbf{W}^n\right) = \left(\mathbf{I} - \lambda\mathbf{W}\right)^{-1}$$

3.2.3 Global and Local Spatial Effects

A question arising from the previous subsection is what happens if a spatial unit experiences a change in the dependent variable, in the independent variable or in unobserved characteristics (captured by the residuals of a spatial econometric model) and what the influence of those changes is on other spatial units. Consider again the Kelejian-Prucha model:

$$\begin{aligned} \mathbf{y} &= \alpha\mathbf{e_N} + \lambda\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \rho\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \end{aligned}$$

We stick to the definition by LeSage and Pace (2009) and calculate the *average* impact of such a change. The following terms are defined with this purpose in mind.

Average direct effect: The average direct effect is the average impact of changes in the independent variable x_{jk} on y_j itself. This change can be calculated as the average over all changes, being $\overline{\varepsilon}_{y_j,x_{jk}} = 1/\mathrm{N} \cdot \beta_k \cdot \mathrm{tr}\left((\mathbf{I} - \lambda \cdot \mathbf{W})^{-1}\right)$

Average total effect from or to an observation: The average total effect is the average impact on all or one y_i by changing one or all x_{jk} and can be calculated as $\overline{\varepsilon}_{y_i,x_{jk}} = 1/\mathrm{N} \cdot \mathbf{e}_\mathrm{N}{}' \cdot \beta_k \cdot \left(\mathbf{I} - \lambda\mathbf{W}\right)^{-1} \cdot \mathbf{e}_\mathrm{N}$

Average indirect effect: The average indirect impact can be defined as the difference of the average total impact and the average direct effect, which is the average change of all other spatial units' dependent variable y_i induced by a change in the independent variable of spatial unit j. Accordingly, the indirect effect is $\overline{\varepsilon}_{y_j,x_{i\neq j,k}} = \overline{\varepsilon}_{y_j,x_{jk}} - \overline{\varepsilon}_{y_j,x_{ik}}$

In applied spatial-econometric studies, the indirect impacts may be of interest. Several examples, among many, are the examination of tax-rate diffusion among European countries (Egger et al. 2005b), spatial spillovers in pollution or abatement activities (Banfi et al. 2006), and spatial spillovers in foreign direct investments (FDI) (Blonigen et al. 2007).

So far, we have seen that if a variable of a spatial unit changes, then not only does the dependent variable of the unit experience a change but also the dependent variables of all other spatial units. The first is called the direct effect or direct impact and the second the indirect effect or indirect impact. Of course, this only holds for the Kelejian-Prucha model. In case of the spatial Durbin model, the data generating process is such that a spatial unit's dependent variable depends on neighbouring units' independent variable – and thus the derivative $\partial y_j / \partial x_{ik}$ would look different and depend on the coefficient vector of the spatially lagged independent variables, $\boldsymbol{\theta}$. Accordingly, indirect effects that stem from $\boldsymbol{\theta} = \mathbf{0}$ are called local effects, whereas indirect effects which arise from $\lambda \neq 0$ are called global effects. In contrast to global effects, local effects only arise from a spatial unit's neighbourhood, provided the relevant element of the spatial weighting matrix, w_{ij}, is non-zero. Global effects also have impacts on spatial units which do not necessarily belong to the strict neighbourhood of a spatial unit from which the change stems. Although the matrix entry w_{ij} may be zero, the (i,j) element of the matrix $(\mathbf{I} - \lambda\mathbf{W})^{-1}$ in general will be non-zero. All formulas from the current subsection remain valid for the case of balanced panel data. The number of observations, however, then increases from N to NT; the identity matrix and the one vectors is also of dimension NT and the matrix $\mathbf{W}$ is replaced by $\mathbf{W} \otimes \mathbf{I}_\mathrm{T}$.

An interesting study analysing the selection of the proper weighting matrix has been conducted by Stakhovych and Bijmolt (2009). The study focuses on the 'correct' specification of a spatial weighting matrix, to determine which a Monte Carlo experiment was conducted. The experiment was designed to analyse cross-sectional data. Two types of spatial models were considered, the spatial error and the spatial lag model – but unfortunately not the Kelejian-Prucha model. The data (100 observations) were generated using either a first-order contiguity matrix, 10-nearest-neighbours matrix and a matrix with inversed distance entries (see subsection

3.2.2). The models were then estimated and tested using each of these matrices to compare results with a maximum likelihood procedure. Findings suggest that the probability of detecting the ‘true’ spatial lag model is higher than detecting the ‘true’ spatial error model, which means that the information criteria concerning the spatial lag should be considered. Recommendations point to the use of a first-order contiguity matrix or one with inversed distance entries (since a Voronoi diagram was created). However, the findings should be challenged using the Kelejian-Prucha model for panel data; in addition, the number of observations should be increased. Moreover, the recommendation of a first-order contiguity matrix might suggest the sole use of a row-normalisation procedure, which treats distances as relative, as discussed above. Therefore, both approaches to matrix normalisation should be investigated, and with different (i.e., decreasing) functions of the distance.

3.3 The General Spatial Autoregressive Model with Autoregressive Disturbances (SARAR-Model)

The purpose of this section is to show the appropriate derivation of the GMM estimator of the Kelejian-Prucha model for panel data. This is performed firstly by extending calculations presented by Kapoor et al. (2007) and secondly by underlining the assumptions and limitations of the GMM estimator. The reason for the complete and exact derivation here is that the results of the different calculation steps are needed to implement the estimator in statistical software such as STATA®. As mentioned in the introduction, one of the goals of the present study is to implement an estimation procedure for the Kelejian-Prucha in STATA® using a GMM approach. For this reason, consider the following spatial autoregressive model with autoregressive disturbances

$$\begin{aligned} \mathbf{y} &= \lambda \cdot (\mathbf{W} \otimes \mathbf{I}_{\mathrm{T}}) \cdot \mathbf{y} + \mathbf{X} \cdot \boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \rho \cdot (\mathbf{M} \otimes \mathbf{I}_{\mathrm{T}}) \cdot \mathbf{u} + \boldsymbol{\varepsilon} \end{aligned} \qquad \text{(III.1)}$$

where $\mathbf{y}$ denotes the $\mathrm{NT}\times 1$ vector of observations of the dependent variable. N denotes the number of cross-sectional units and N the length of the time period. $\mathbf{X}$ is the $\mathrm{NT}\times\mathrm{K}$ matrix of observations of the exogenous independent variables and $\boldsymbol{\beta}$ is the $\mathrm{K}\times 1$ vector of coefficients to be estimated. Furthermore, the residuals $\mathbf{u}$ of dimension $\mathrm{NT}\times 1$ should follow a spatial process too. The scalar coefficients λ and ρ represent the coefficients of the spatially lagged dependent variable and the spatially lagged residuals respectively. Equation (III.1) states that the data-generating process is such that the independent variable not only depends on the exogenous variables $\mathbf{X}$ but as well on a weighted average of the independent variables of proximate cross-sectional units. This degree of interaction is measured by the spatial weighting matrices $\mathbf{W}$ or $\mathbf{M}$ of dimension $\mathrm{N}\times\mathrm{N}$. In the current setting, the data is stacked such that i (the index referring to the cross-sectional units) is the “slow” index and $\mathbf{t}$ (the index referring to the time period) is the “fast” index.

Concerning the residuals $\boldsymbol{\varepsilon}$, the following assumptions are made[21]:

$$\begin{aligned} \boldsymbol{\varepsilon} &= \boldsymbol{\mu} \otimes \mathbf{e}_{\mathrm{T}} + \mathbf{v} \\ E(\boldsymbol{\mu}) &= \mathbf{0} \\ E(\boldsymbol{\mu} \cdot \boldsymbol{\mu}') &= \sigma_{\mu}^{2} \cdot \mathbf{I}_{\mathrm{N}} \\ E(\mathbf{v}) &= \mathbf{0} \\ E(\mathbf{v} \cdot \mathbf{v}') &= \sigma_{\mathrm{v}}^{2} \cdot \mathbf{I}_{\mathrm{NT}} \\ E(\mathbf{v} \cdot (\boldsymbol{\mu} \otimes \mathbf{e}_{\mathrm{T}})') &= \mathbf{0} \end{aligned} \qquad \text{(III.2)}$$

where $\boldsymbol{\mu}$ is a $\mathrm{N}\times 1$ vector of individual random effects with zero mean, zero covariance and variance σ_{μ}^{2}. Further, μ_i and v_{it} are i.i.d. with finite fourth moments. It is assumed that the

[21] According to the present error specification, a random effects model is discussed.

matrices $\mathbf{I}_N - \lambda\mathbf{W}$ and $\mathbf{I}_N - \rho\mathbf{W}$ are invertible (non-singular) and that λ and ρ are bounded in absolute value. Given that and the assumptions in equation (III.2), the variance-covariance matrix of $\boldsymbol{\varepsilon}$ is

$$\boldsymbol{\Omega}_{\varepsilon} = E(\boldsymbol{\varepsilon} \cdot \boldsymbol{\varepsilon}') = \sigma_{\mu}^2 \cdot (\mathbf{I}_N \otimes \mathbf{e}_T \cdot \mathbf{e}_T') + \sigma_{\nu}^2 \cdot \mathbf{I}_{NT} \tag{III.3}$$

Accordingly, the residuals are auto-correlated over time due to the presence of time-invariant disturbances $\boldsymbol{\mu}$

Further, two common matrices in panel data analysis are defined, being

$$\begin{aligned} \mathbf{Q}_0 &= \mathbf{I}_N \otimes (\mathbf{I}_T - \frac{1}{T} \cdot \mathbf{e}_T \cdot \mathbf{e}_T') \\ \mathbf{Q}_1 &= \mathbf{I}_N \otimes (\frac{1}{T} \cdot \mathbf{e}_T \cdot \mathbf{e}_T') \end{aligned} \tag{III.4}$$

Multiplication of any vector with $\mathbf{Q}_0$ results in a time-demeaned version of that vector. It is easy to see that both matrices $\mathbf{Q}_0$ and $\mathbf{Q}_1$ are symmetric ($\mathbf{Q}_i = \mathbf{Q}_i'$) and idempotent ($\mathbf{Q}_i \cdot \mathbf{Q}_i = \mathbf{Q}_i$). Moreover, they are orthogonal ($\mathbf{Q}_1 \cdot \mathbf{Q}_0 = \mathbf{0}$) to each other.

Proof:

For any convenient matrices $\mathbf{A}$ *and* $\mathbf{B}$*, it holds that*

$$(\mathbf{A} \cdot \mathbf{B})' = \mathbf{B}' \cdot \mathbf{A}',\ (\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$$

$$(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}',\ (\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$$

The matrix $\mathbf{Q}_0$ is symmetric, since

$$\mathbf{Q}_0' = \mathbf{I}_N' \otimes (\mathbf{I}_T' - \frac{1}{T}\mathbf{e}_T \cdot \mathbf{e}_T') = \mathbf{I}_N \otimes (\mathbf{I}_T - \frac{1}{T}\mathbf{e}_T \cdot \mathbf{e}_T') \equiv \mathbf{Q}_0$$

The matrix $\mathbf{Q}_1$ is symmetric, since

$$\mathbf{Q}_1' = \mathbf{I}_N' \otimes (\frac{1}{T} \cdot \mathbf{e}_T \cdot \mathbf{e}_T') \equiv \mathbf{Q}_1$$

The matrix $\mathbf{Q}_0$ is idempotent, since

$$\begin{aligned} \mathbf{Q}_0 \cdot \mathbf{Q}_0 &= \mathbf{I}_N \cdot \mathbf{I}_N \otimes (\mathbf{I}_T - \frac{1}{T}\mathbf{e}_T \cdot \mathbf{e}_T') \cdot (\mathbf{I}_T - \frac{1}{T}\mathbf{e}_T \cdot \mathbf{e}_T') \\ &= \mathbf{I}_N \otimes (\mathbf{I}_T - \frac{2}{T}\mathbf{e}_T \cdot \mathbf{e}_T' + \frac{1}{T^2} \cdot \underbrace{\mathbf{e}_T \cdot \mathbf{e}_T' \cdot \mathbf{e}_T \cdot \mathbf{e}_T'}_{T \cdot \mathbf{e}_T \cdot \mathbf{e}_T'}) \equiv \mathbf{Q}_0 \end{aligned}$$

The matrix $\mathbf{Q}_1$ is idempotent, since

$$\mathbf{Q}_1 \cdot \mathbf{Q}_1 = \mathbf{I}_\mathrm{N} \cdot \mathbf{I}_\mathrm{N} \otimes (\frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}') \cdot (\frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}') = \mathbf{I}_\mathrm{N} \otimes \frac{1}{\mathrm{T}^2} \cdot \underbrace{\mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}' \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}'}_{\mathrm{T} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}'} \equiv \mathbf{Q}_1$$

The matrices $\mathbf{Q}_0$ and $\mathbf{Q}_1$ are orthogonal, since

$$\mathbf{Q}_0 \cdot \mathbf{Q}_1 = \mathbf{I}_\mathrm{N} \cdot \mathbf{I}_\mathrm{N} \otimes (\mathbf{I}_\mathrm{T} \cdot \frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}' - \frac{1}{\mathrm{T}^2} \cdot \underbrace{\mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}' \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}'}_{\mathrm{T} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}'}) = \mathbf{0}$$

Further, the following important characteristics hold:

$$\begin{aligned} \mathbf{Q}_0 + \mathbf{Q}_1 &= \mathbf{I}_\mathrm{NT} \\ \mathrm{tr}(\mathbf{Q}_0) &= \mathrm{tr}(\mathbf{I}_\mathrm{N}) \cdot \mathrm{tr}(\mathbf{I}_\mathrm{T} - \frac{1}{\mathrm{T}} \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}') = \mathrm{N} \cdot (\mathrm{T} - \frac{1}{\mathrm{T}} \sum_{\mathrm{i}=1}^{\mathrm{T}} 1) = \mathrm{N(T\text{-}1)} \\ \mathrm{tr}(\mathbf{Q}_1) &= \mathrm{N} \end{aligned} \tag{III.5}$$

Combining equation (III.3) with equation (III.4) and using the first property of equation (III.5), the variance-covariance of the residuals can be re-written to

$$\mathbf{\Omega}_\varepsilon = \sigma_\mathrm{v}^2 \cdot \mathbf{Q}_0 + (\sigma_\mathrm{v}^2 + \mathrm{T} \cdot \sigma_\mu^2) \cdot \mathbf{Q}_1 \Rightarrow \mathbf{\Omega}_\varepsilon = \sigma_\mathrm{v}^2 \cdot \mathbf{Q}_0 + \sigma_1^2 \cdot \mathbf{Q}_1 \tag{III.6}$$

and since the matrices $\mathbf{Q}_0$ and $\mathbf{Q}_1$ are orthogonal, the inverse of the variance-covariance matrix can be written as

$$\mathbf{\Omega}_\varepsilon^{-1} = \frac{1}{\sigma_\mathrm{v}^2} \mathbf{Q}_0 + \frac{1}{\sigma_\mathrm{v}^2 + T \cdot \sigma_\mu^2} \mathbf{Q}_1 \tag{III.7}$$

For the following calculations, it is useful to see that

$$\begin{aligned} \mathbf{Q}_0 \cdot \boldsymbol{\varepsilon} &= \mathbf{I}_\mathrm{N} \otimes (\mathbf{I}_\mathrm{T} - \frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}') \cdot (\boldsymbol{\mu} \otimes \mathbf{e}_\mathrm{T} + \mathbf{v}) \\ &= \mathbf{I}_\mathrm{N} \cdot \boldsymbol{\mu} \otimes \mathbf{I}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T} + \mathbf{I}_\mathrm{N} \otimes \mathbf{I}_\mathrm{T} \cdot \mathbf{v} - \mathbf{I}_\mathrm{N} \cdot \boldsymbol{\mu} \otimes \underbrace{\frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}' \cdot \mathbf{e}_\mathrm{T}}_{\mathbf{e}_\mathrm{T}} - \mathbf{I}_\mathrm{N} \otimes \frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}' \cdot \mathbf{v} \\ &= \mathbf{I}_\mathrm{N} \otimes (\mathbf{I}_\mathrm{T} - \frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}') \cdot \mathbf{v} = \mathbf{Q}_0 \cdot \mathbf{v} \end{aligned} \tag{III.8}$$

and

$$\begin{aligned} \mathbf{Q}_1 \cdot \boldsymbol{\varepsilon} &= (\mathbf{I}_\mathrm{N} \otimes (\frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}')) \cdot (\boldsymbol{\mu} \otimes \mathbf{e}_\mathrm{T} + \mathbf{v}) \\ &= \mathbf{I}_\mathrm{N} \cdot \boldsymbol{\mu} \otimes \underbrace{\frac{1}{\mathrm{T}} \cdot \mathbf{e}_\mathrm{T} \cdot \mathbf{e}_\mathrm{T}{}' \cdot \mathbf{e}_\mathrm{T}}_{\mathbf{e}_\mathrm{T}} + \mathbf{Q}_1 \cdot \mathbf{v} \\ &= (\boldsymbol{\mu} \otimes \mathbf{e}_\mathrm{T}) + \mathbf{Q}_1 \cdot \mathbf{v} \end{aligned} \tag{III.9}$$

Equation (III.8) is very intuitive. The matrix $\mathbf{Q}_0$ is a time-demeaning matrix and, since the vector $\boldsymbol{\mu}$ only exhibits between but no within variation, the fixed effects μ_i can be eliminated from the model by multiplying with $\mathbf{Q}_0$.

The previous assumptions, namely that the spatial weighting matrices $\mathbf{W}$ and $\mathbf{M}$ only have zero as diagonal elements and that the matrices $\mathbf{I}_N - \lambda\mathbf{W}$ and $\mathbf{I}_N - \rho\mathbf{W}$ are invertible, ensure that the model stated by equation (III.1) is uniquely solvable. The non-singularity condition is needed to transform equation (III.1) to

$$\begin{aligned} \mathbf{y} &= (\mathbf{I}_{NT} - \lambda\cdot(\mathbf{W}\otimes\mathbf{I}_T))^{-1}\cdot(\mathbf{X}\boldsymbol{\beta} + (\mathbf{I}_{NT} - \rho\cdot(\mathbf{M}\otimes\mathbf{I}_T))^{-1}\cdot\boldsymbol{\varepsilon}) \\ \mathbf{u} &= (\mathbf{I}_{NT} - \rho\cdot(\mathbf{M}\otimes\mathbf{I}_T))^{-1}\cdot\boldsymbol{\varepsilon} \end{aligned} \tag{III.10}$$

This allows the derivation of an expression for the variance-covariance matrix of the spatially correlated residuals[22], namely

$$\boldsymbol{\Omega}_u = ((\mathbf{I}_N - \rho\cdot\mathbf{M})^{-1}\otimes\mathbf{I}_T)\cdot\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}}\cdot((\mathbf{I}_N - \rho\cdot\mathbf{M}')^{-1}\otimes\mathbf{I}_T) \tag{III.11}$$

and moreover

$$\begin{aligned} \mathrm{E}((\mathbf{W}\otimes\mathbf{I}_T)\mathbf{y}\cdot\mathbf{u}') &= \mathrm{E}\Big((\mathbf{W}\cdot(\mathbf{I}_N - \lambda\cdot\mathbf{W})^{-1}\otimes\mathbf{I}_T)\cdot(\mathbf{X}\boldsymbol{\beta} + \mathbf{u})\cdot\mathbf{u}'\Big) \\ &= \mathrm{E}\left((\mathbf{W}(\mathbf{I}_N - \lambda\cdot\mathbf{W})^{-1}\otimes\mathbf{I}_T)\mathbf{X}\boldsymbol{\beta}\cdot\underbrace{\mathbf{u}'}_{\mathrm{E}(\mathbf{u})=\mathbf{0}} + (\mathbf{W}(\mathbf{I}_N - \lambda\cdot\mathbf{W})^{-1}\otimes\mathbf{I}_T)\cdot\underbrace{\mathbf{u}\mathbf{u}'}_{\mathrm{E}(\mathbf{u}\cdot\mathbf{u}')=\boldsymbol{\Omega}_u}\right) \\ &= (\mathbf{W}\cdot(\mathbf{I}_N - \lambda\cdot\mathbf{W})^{-1}\otimes\mathbf{I}_T)\cdot\boldsymbol{\Omega}_{\mathbf{u}} \neq \mathbf{0} \end{aligned}$$

which makes clear that the spatially lagged dependent variable is correlated with the residuals and thus endogenous.

[22] The computational feasibility of calculating the estimator proposed depends strongly on the size of the respective matrices to be inverted. Therefore, it is advantageous to use the fact that for any convenient quadratic matrix $\mathbf{A}$, it holds that $(\mathbf{A}\otimes\mathbf{I}_T)^{-1} = \mathbf{A}^{-1}\otimes\mathbf{I}_T^{-1} = \mathbf{A}^{-1}\otimes\mathbf{I}_T$

3.4 Generalised Method of Moments - The GMM Estimator of the SARAR-Model

3.4.1 Derivation of Moment Conditions

For a proper derivation of the moment conditions, the residuals are transformed using the spatial weights matrix $\mathbf{M}$ such that

$$\begin{aligned}\bar{\mathbf{u}} &= (\mathbf{M}\otimes\mathbf{I}_T)\cdot\mathbf{u}\\ \bar{\bar{\mathbf{u}}} &= (\mathbf{M}\otimes\mathbf{I}_T)\cdot\bar{\mathbf{u}}\\ \bar{\boldsymbol{\varepsilon}} &= (\mathbf{M}\otimes\mathbf{I}_T)\cdot\boldsymbol{\varepsilon}\end{aligned} \tag{III.12}$$

where the superscript bar means that the respective variable is spatially transformed by multiplying with matrix $\mathbf{M}$.

Using the above notation results in

$$\begin{aligned}\boldsymbol{\varepsilon} &= \mathbf{u}-\rho\cdot\bar{\mathbf{u}}\\ \bar{\boldsymbol{\varepsilon}} &= \bar{\mathbf{u}}-\rho\cdot\bar{\bar{\mathbf{u}}}\end{aligned} \tag{III.13}$$

Further, let $\mathbf{A}$ be any $\mathrm{N}\times\mathrm{N}$ matrix. Then, the following properties concerning $\mathbf{A}$, $\mathbf{Q}_0$ and $\mathbf{Q}_1$ hold:

$$\begin{aligned}&(\mathbf{A}\otimes\mathbf{I}_T)\cdot\mathbf{Q}_0 = \mathbf{Q}_0\cdot(\mathbf{A}\otimes\mathbf{I}_T) \quad\text{and}\quad (\mathbf{A}\otimes\mathbf{I}_T)\cdot\mathbf{Q}_1 = \mathbf{Q}_1\cdot(\mathbf{A}\otimes\mathbf{I}_T)\\ &\mathrm{tr}(\mathbf{Q}_0\cdot(\mathbf{A}\otimes\mathbf{I}_T)) = (T-1)\cdot\mathrm{tr}(\mathbf{A}) \quad\text{and}\quad \mathrm{tr}(\mathbf{Q}_1\cdot(\mathbf{A}\otimes\mathbf{I}_T)) = \mathrm{tr}(\mathbf{A})\end{aligned} \tag{III.14}$$

The first equality can easily be conceived by substituting the expression for $\mathbf{Q}_0$ from (III.4) into the equation and checking whether the expression $(\mathbf{A}\otimes\mathbf{I}_T)\cdot\mathbf{Q}_0-\mathbf{Q}_0\cdot(\mathbf{A}\otimes\mathbf{I}_T)$ is zero, which yields

$$\begin{aligned}&(\mathbf{A}\otimes\mathbf{I}_T)\cdot(\mathbf{I}_N\otimes(\mathbf{I}_T-\frac{1}{T}\cdot\mathbf{e}_T\cdot\mathbf{e}_T'))-(\mathbf{I}_N\otimes(\mathbf{I}_T-\frac{1}{T}\cdot\mathbf{e}_T\cdot\mathbf{e}_T'))(\mathbf{A}\otimes\mathbf{I}_T) =\\ &(\mathbf{A}\cdot\mathbf{I}_N)\otimes(\mathbf{I}_T\cdot(\mathbf{I}_T-\frac{1}{T}\cdot\mathbf{e}_T\cdot\mathbf{e}_T'))-(\mathbf{I}_N\cdot\mathbf{A})\otimes((\mathbf{I}_T-\frac{1}{T}\cdot\mathbf{e}_T\cdot\mathbf{e}_T')\cdot\mathbf{I}_T) = \mathbf{0}\end{aligned}$$

In order to show that the second property of equation (III.14) holds, simply substitute $\mathbf{Q}_0$ to obtain

$$\begin{aligned}&\mathrm{tr}((\mathbf{I}_N\otimes(\mathbf{I}_T-\frac{1}{T}\mathbf{e}_T\mathbf{e}_T'))\cdot(\mathbf{A}\otimes\mathbf{I}_T)) = \mathrm{tr}((\mathbf{I}_N\cdot\mathbf{A})\otimes((\mathbf{I}_T-\frac{1}{T}\mathbf{e}_T\mathbf{e}_T')\cdot\mathbf{I}_T) =\\ &\mathrm{tr}(\mathbf{A})\cdot\mathrm{tr}(\mathbf{I}_T-\frac{1}{T}\cdot\mathbf{e}_T\cdot\mathbf{e}_T') = \mathrm{tr}(\mathbf{A})\cdot(\underbrace{\mathrm{tr}(\mathbf{I}_T)}_{T}-\frac{1}{T}\underbrace{\mathrm{tr}(\mathbf{e}_T\cdot\mathbf{e}_T')}_{T}) = \mathrm{tr}(\mathbf{A})\cdot(T-1)\end{aligned}$$

Regarding the properties concerning $\mathbf{Q}_1$ in equation (III.14), the procedure is the same.

Before turning to the moment conditions, the following properties concerning $\mathbf{Q}_0$, $\mathbf{Q}_1$ and the residuals $\boldsymbol{\varepsilon}$ and the spatially lagged residuals $\overline{\boldsymbol{\varepsilon}}$ can be stated:

$$\begin{aligned}
\mathbf{Q}_0 \cdot \boldsymbol{\varepsilon} &= \mathbf{Q}_0 \cdot \mathbf{v} \\
\mathbf{Q}_0 \cdot \overline{\boldsymbol{\varepsilon}} &= \mathbf{Q}_0 (\mathbf{M} \otimes \mathbf{I}_T) \boldsymbol{\varepsilon} = (\mathbf{M} \otimes \mathbf{I}_T) \mathbf{Q}_0 \cdot \mathbf{v} \\
\mathbf{Q}_1 \cdot \boldsymbol{\varepsilon} &= (\boldsymbol{\mu} \otimes \mathbf{e}_T) + \mathbf{Q}_1 \cdot \mathbf{v} \\
\mathbf{Q}_1 \cdot \overline{\boldsymbol{\varepsilon}} &= \mathbf{Q}_1 (\mathbf{M} \otimes \mathbf{I}_T) \boldsymbol{\varepsilon} = (\mathbf{M} \otimes \mathbf{I}_T) \mathbf{Q}_1 \cdot (\boldsymbol{\mu} \otimes \mathbf{e}_T + \mathbf{v}) \\
&= (\mathbf{M} \otimes \mathbf{I}_T) \cdot (\boldsymbol{\mu} \otimes \mathbf{e}_T + \mathbf{Q}_1 \cdot \mathbf{v})
\end{aligned} \tag{III.15}$$

The first and the third properties only reflect equations (III.8) and (III.9), whereas the second and the fourth properties have been derived with help of equations (III.8) and (III.14) and (III.9) and (III.14) respectively.

Further, for any convenient quadratic, non-stochastic matrix $\mathbf{A}$ and a convenient stochastic vector $\boldsymbol{\eta}$, it holds that

$$\mathrm{E}(\boldsymbol{\eta}' \cdot \mathbf{A} \cdot \boldsymbol{\eta}) = \mathrm{tr}(\mathbf{A} \cdot \mathrm{E}(\boldsymbol{\eta} \cdot \boldsymbol{\eta}')) \tag{III.16}$$

These properties now allow the derivation of the moment conditions.

$$\begin{aligned}
\mathrm{E}(\boldsymbol{\varepsilon}' \cdot \mathbf{Q}_0 \cdot \boldsymbol{\varepsilon}) \quad &= \mathrm{E}(\boldsymbol{\varepsilon}' \underbrace{\mathbf{Q}_0 ' \mathbf{Q}_0}_{\mathbf{Q}_0} \boldsymbol{\varepsilon}) = \mathrm{E}(\mathbf{v}' \mathbf{Q}_0 \mathbf{v}) = \mathrm{tr}(\mathbf{Q}_0 \cdot \mathrm{E}(\mathbf{v}\mathbf{v}')) \\
&= \underline{\underline{\sigma_v^2 \cdot \mathrm{N(T\text{-}1)}}}
\end{aligned} \tag{III.17}$$

$$\begin{aligned}
\mathrm{E}(\overline{\boldsymbol{\varepsilon}}' \cdot \mathbf{Q}_0 \cdot \overline{\boldsymbol{\varepsilon}}) \quad &= \mathrm{E}\left(\left[(\mathbf{M} \otimes \mathbf{I}_T) \boldsymbol{\varepsilon} \right]' \mathbf{Q}_0 ' \cdot \mathbf{Q}_0 \left[(\mathbf{M} \otimes \mathbf{I}_T) \boldsymbol{\varepsilon} \right] \right) \\
&= \mathrm{E}\left(\boldsymbol{\varepsilon}' (\mathbf{M}' \otimes \mathbf{I}_T) \mathbf{Q}_0 ' \cdot \mathbf{Q}_0 (\mathbf{M} \otimes \mathbf{I}_T) \boldsymbol{\varepsilon} \right) \\
&= \mathrm{E}\left(\boldsymbol{\varepsilon}' \mathbf{Q}_0 ' (\mathbf{M}' \otimes \mathbf{I}_T) \cdot (\mathbf{M} \otimes \mathbf{I}_T) \mathbf{Q}_0 \boldsymbol{\varepsilon} \right) \\
&= \mathrm{E}\left(\mathbf{v}' \mathbf{Q}_0 \cdot ((\mathbf{M}'\mathbf{M}) \otimes \mathbf{I}_T) \cdot \mathbf{Q}_0 \mathbf{v} \right) \\
&= \mathrm{tr}\left(\mathbf{Q}_0 ((\mathbf{M}'\mathbf{M}) \otimes \mathbf{I}_T) \mathbf{Q}_0 \cdot \mathrm{E}\left(\mathbf{v}\mathbf{v}' \right) \right) \\
&= \sigma_v^2 \cdot \mathrm{tr}\left(\mathbf{Q}_0 \cdot ((\mathbf{M}'\mathbf{M}) \otimes \mathbf{I}_T) \right) \\
&= \underline{\underline{\sigma_v^2 \cdot (\mathrm{T} - 1) \cdot \mathrm{tr}(\mathbf{M}'\mathbf{M})}}
\end{aligned} \tag{III.18}$$

$$
\begin{aligned}
\mathrm{E}(\bar{\boldsymbol{\varepsilon}}'\cdot\mathbf{Q}_0\cdot\boldsymbol{\varepsilon}) &= \mathrm{E}\left(\left[(\mathbf{M}\otimes\mathbf{I}_T)\boldsymbol{\varepsilon}\right]'\mathbf{Q}_0{}'\cdot\mathbf{Q}_0\boldsymbol{\varepsilon}\right)\\
&= \mathrm{E}\left(\boldsymbol{\varepsilon}'(\mathbf{M}'\otimes\mathbf{I}_T)\mathbf{Q}_0{}'\cdot\mathbf{Q}_0\boldsymbol{\varepsilon}\right)\\
&= \mathrm{E}\left(\boldsymbol{\varepsilon}'\mathbf{Q}_0{}'\cdot(\mathbf{M}'\otimes\mathbf{I}_T)\cdot\mathbf{Q}_0\boldsymbol{\varepsilon}\right)\\
&= \mathrm{E}\left(\mathbf{v}'\mathbf{Q}_0\cdot(\mathbf{M}'\otimes\mathbf{I}_T)\cdot\mathbf{Q}_0\mathbf{v}\right)\\
&= \mathrm{tr}\left(\mathbf{Q}_0(\mathbf{M}'\otimes\mathbf{I}_T)\mathbf{Q}_0\cdot\mathrm{E}\left(\mathbf{v}'\mathbf{v}\right)\right)\\
&= \sigma_v^2\cdot\mathrm{tr}\left(\mathbf{Q}_0\cdot(\mathbf{M}'\otimes\mathbf{I}_T)\right)\\
&= \sigma_v^2\cdot(T-1)\cdot\mathrm{tr}\left(\mathbf{M}'\right)\\
&= \underline{\underline{\mathbf{0}}}
\end{aligned}
\tag{III.19}
$$

$$
\begin{aligned}
\mathrm{E}(\boldsymbol{\varepsilon}'\cdot\mathbf{Q}_1\cdot\boldsymbol{\varepsilon}) &= \mathrm{E}\left((\boldsymbol{\mu}\otimes\mathbf{e}_T+\mathbf{v})'\mathbf{Q}_1{}'\cdot\mathbf{Q}_1(\boldsymbol{\mu}\otimes\mathbf{e}_T+\mathbf{v})\right)\\
&= \mathrm{E}\left((\boldsymbol{\mu}'\otimes\mathbf{e}_T{}'+\mathbf{v}'\mathbf{Q}_1{}')\cdot(\boldsymbol{\mu}\otimes\mathbf{e}_T+\mathbf{Q}_1\mathbf{v})\right)\\
&= \mathrm{E}\left((\boldsymbol{\mu}'\otimes\mathbf{e}_T{}'+\mathbf{v}'\mathbf{Q}_1{}')\cdot(\boldsymbol{\mu}\otimes\mathbf{e}_T+\mathbf{Q}_1\mathbf{v})\right)\\
&= \mathrm{E}\left((\boldsymbol{\mu}'\boldsymbol{\mu})\otimes(\mathbf{e}_T{}'\mathbf{e}_T)+\boldsymbol{\mu}'\otimes\mathbf{e}_T{}'\mathbf{Q}_1\mathbf{v}+\mathbf{v}'\mathbf{Q}_1{}'(\boldsymbol{\mu}\otimes\mathbf{e}_T)+\mathbf{v}'\mathbf{Q}_1{}'\mathbf{Q}_1\mathbf{v}\right)\\
&= \mathrm{E}\left((\boldsymbol{\mu}'\boldsymbol{\mu})\otimes(\mathbf{e}_T{}'\mathbf{e}_T)+\mathbf{v}'\mathbf{Q}_1\mathbf{v}\right)\\
&= \sigma_\mu^2\cdot NT+\mathrm{tr}(\mathbf{Q}_1\cdot\mathrm{E}\left(v'v\right))=\sigma_\mu^2\cdot NT+\sigma_v^2\cdot N\\
&= \underline{\underline{N\cdot\sigma_1^2}}
\end{aligned}
\tag{III.20}
$$

$$
\begin{aligned}
\mathrm{E}(\bar{\boldsymbol{\varepsilon}}'\cdot\mathbf{Q}_1\cdot\bar{\boldsymbol{\varepsilon}}) &= \mathrm{E}\begin{pmatrix}\left[(\boldsymbol{\mu}'\otimes\mathbf{e}_T{}')(\mathbf{M}'\otimes\mathbf{I}_T)+\mathbf{v}'(\mathbf{M}'\otimes\mathbf{I}_T)\right]\cdot\mathbf{Q}_1{}'\cdot\\ \mathbf{Q}_1\cdot\left[(\mathbf{M}\otimes\mathbf{I}_T)(\boldsymbol{\mu}\otimes\mathbf{e}_T)+(\mathbf{M}\otimes\mathbf{I}_T)\mathbf{v}\right]\end{pmatrix}\\
&= \mathrm{E}\begin{pmatrix}((\boldsymbol{\mu}'\mathbf{M}')\otimes\mathbf{e}_T{}')((\mathbf{M}\boldsymbol{\mu})\otimes\mathbf{e}_T)+((\boldsymbol{\mu}'\mathbf{M}')\otimes\mathbf{e}_T{}')(\mathbf{M}\otimes\mathbf{I}_T)\mathbf{Q}_1\mathbf{v}+\\ \mathbf{v}'\mathbf{Q}_1(\mathbf{M}'\otimes\mathbf{I}_T)(\mathbf{M}\otimes\mathbf{I}_T)(\boldsymbol{\mu}\otimes\mathbf{e}_T)+\\ \mathbf{v}'\mathbf{Q}_1(\mathbf{M}'\otimes\mathbf{I}_T)(\mathbf{M}\otimes\mathbf{I}_T)\mathbf{Q}_1 v\end{pmatrix}\\
&= \mathrm{E}\left(\boldsymbol{\mu}'\cdot((\mathbf{M}'\mathbf{M})\otimes(\mathbf{e}_T{}'\mathbf{e}_T))\right)\cdot\boldsymbol{\mu}+\mathrm{E}\left(\mathbf{v}'\mathbf{Q}_1\cdot((\mathbf{M}'\mathbf{M})\otimes\mathbf{I}_T)\cdot\mathbf{Q}_1\mathbf{v}\right)\\
&= \mathrm{tr}\left(((\mathbf{M}'\mathbf{M})\otimes(\mathbf{e}_T{}'\mathbf{e}_T))\cdot\mathrm{E}\left(\boldsymbol{\mu}\boldsymbol{\mu}'\right)\right)+\mathrm{tr}\left(\mathbf{Q}_1((\mathbf{M}'\mathbf{M})\otimes\mathbf{I}_T)\mathbf{Q}_1\cdot\mathrm{E}(\mathbf{v}'\mathbf{v})\right)\\
&= \sigma_\mu^2\cdot T\cdot\mathrm{tr}\left(\mathbf{M}'\mathbf{M}\right)+\sigma_v^2\cdot\mathrm{tr}\left(\mathbf{M}'\mathbf{M}\right)\\
&= \underline{\underline{\sigma_1^2\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M})}}
\end{aligned}
\tag{III.21}
$$

$$
\begin{aligned}
\mathrm{E}(\bar{\boldsymbol{\varepsilon}}'\cdot\mathbf{Q}_1\cdot\boldsymbol{\varepsilon}) &= \mathrm{E}\left(\left[(\boldsymbol{\mu}'\otimes\mathbf{e}_T{}')(\mathbf{M}'\otimes\mathbf{I}_T)+\mathbf{v}'(\mathbf{M}'\otimes\mathbf{I}_T)\right]\mathbf{Q}_1{}'\cdot\mathbf{Q}_1\left[(\boldsymbol{\mu}\otimes\mathbf{e}_T)+\mathbf{v}\right]\right)\\
&= \mathrm{E}\begin{pmatrix}((\boldsymbol{\mu}'\mathbf{M}')\otimes\mathbf{e}_T{}')(\boldsymbol{\mu}\otimes\mathbf{e}_T)+((\boldsymbol{\mu}'\mathbf{M}')\otimes\mathbf{e}_T{}')(\mathbf{M}\otimes\mathbf{I}_T)\mathbf{Q}_1\mathbf{v}+\\ \mathbf{v}'(\mathbf{M}'\otimes\mathbf{I}_T)(\boldsymbol{\mu}\otimes\mathbf{e}_T)+\mathbf{v}'\mathbf{Q}_1(\mathbf{M}'\otimes\mathbf{I}_T)\mathbf{Q}_1\mathbf{v}\end{pmatrix}\\
&= \mathrm{E}\left(\boldsymbol{\mu}'(\mathbf{M}'\otimes(\mathbf{e}_T{}'\mathbf{e}_T))\boldsymbol{\mu}\right)+\mathrm{E}\left(\mathbf{v}'\mathbf{Q}_1(\mathbf{M}'\otimes\mathbf{I}_T)\mathbf{Q}_1\mathbf{v}\right)\\
&= \mathrm{tr}\left((\mathbf{M}'\otimes\mathbf{e}_T{}'\mathbf{e}_T)\cdot\mathrm{E}\left(\boldsymbol{\mu}\boldsymbol{\mu}'\right)\right)+\mathrm{tr}\left(\mathbf{Q}_1(\mathbf{M}'\otimes\mathbf{I}_T)\mathbf{Q}_1\cdot\mathrm{E}(\mathbf{v}'\mathbf{v})\right)\\
&= \sigma_\mu^2\cdot T\cdot\mathrm{tr}\left(\mathbf{M}'\right)+\sigma_v^2\cdot\mathrm{tr}\left(\mathbf{M}'\right)\\
&= \underline{\underline{\mathbf{0}}}
\end{aligned}
\tag{III.22}
$$

Finally, the moment conditions can be rewritten and stacked into a matrix notation:

$$\mathrm{E}\begin{bmatrix} 1/(\mathrm{N(T\text{-}1)})\cdot\boldsymbol{\varepsilon}'\mathbf{Q}_0\boldsymbol{\varepsilon} \\ 1/(\mathrm{N(T\text{-}1)})\cdot\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_0\overline{\boldsymbol{\varepsilon}} \\ 1/(\mathrm{N(T\text{-}1)})\cdot\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_0\boldsymbol{\varepsilon} \\ 1/\mathrm{N}\cdot\boldsymbol{\varepsilon}'\mathbf{Q}_1\boldsymbol{\varepsilon} \\ 1/\mathrm{N}\cdot\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_1\overline{\boldsymbol{\varepsilon}} \\ 1/\mathrm{N}\cdot\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_1\boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 1/\mathrm{N}\cdot\sigma^2\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M}) \\ 0 \\ \sigma_1^2 \\ 1/\mathrm{N}\cdot\sigma_1^2\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M}) \\ 1/\mathrm{N}\cdot\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_1\boldsymbol{\varepsilon} \end{bmatrix} \tag{III.23}$$

In order to apply the moment conditions stated in equation (III.23) one needs consistent estimates $\tilde{\boldsymbol{\varepsilon}}$ for the residuals $\boldsymbol{\varepsilon}$. This in turn raises the need for consistent estimates for the spatial-autocorrelation parameter ρ, for which the moment conditions have to be evaluated. To break this vicious circle, consistent estimates $\tilde{\mathbf{u}}$ for the residuals $\mathbf{u}$ can be obtained by a two-stage ordinary least squares estimation. Once those estimates are available, the identity found in equation (III.13) can be used with the residuals $\mathbf{u}$ and $\overline{\mathbf{u}}$ replaced by the respective consistent estimates $\tilde{\mathbf{u}}$ and $\tilde{\overline{\mathbf{u}}}$. Equation (III.13) then changes to

$$\begin{aligned} \tilde{\boldsymbol{\varepsilon}} &= \tilde{\mathbf{u}} - \rho\cdot\tilde{\overline{\mathbf{u}}} \\ \tilde{\overline{\boldsymbol{\varepsilon}}} &= \tilde{\overline{\mathbf{u}}} - \rho\cdot\tilde{\overline{\overline{\mathbf{u}}}} \end{aligned} \tag{III.24}$$

With help of equation (III.13), equations (III.17)-(III.22) can be transformed as follows

$$\begin{aligned} \frac{1}{\mathrm{N(T\text{-}1)}}\mathrm{E}(\boldsymbol{\varepsilon}'\mathbf{Q}_0\boldsymbol{\varepsilon}) &= \frac{1}{\mathrm{N(T\text{-}1)}}\mathrm{E}\big((\mathbf{u}'-\rho\overline{\mathbf{u}}')\mathbf{Q}_0(\mathbf{u}-\rho\overline{\mathbf{u}})\big) \\ &= \frac{1}{\mathrm{N(T\text{-}1)}}\big(\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\mathbf{u}) - 2\rho\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\overline{\mathbf{u}}) + \rho^2\mathrm{E}(\overline{\mathbf{u}}'\mathbf{Q}_0\overline{\mathbf{u}})\big) \\ &= \sigma_{\mathrm{v}}^2 \end{aligned} \tag{III.25}$$

$$\begin{aligned} \frac{1}{\mathrm{N(T\text{-}1)}}\mathrm{E}(\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_0\overline{\boldsymbol{\varepsilon}}) &= \frac{1}{\mathrm{N(T\text{-}1)}}\mathrm{E}\big((\overline{\mathbf{u}}'-\rho\overline{\overline{\mathbf{u}}}')\mathbf{Q}_0(\overline{\mathbf{u}}-\rho\overline{\overline{\mathbf{u}}})\big) \\ &= \frac{1}{\mathrm{N(T\text{-}1)}}\big(\mathrm{E}(\overline{\mathbf{u}}'\mathbf{Q}_0\overline{\mathbf{u}}) - 2\rho\mathrm{E}(\overline{\overline{\mathbf{u}}}'\mathbf{Q}_0\overline{\mathbf{u}}) + \rho^2\mathrm{E}(\overline{\overline{\mathbf{u}}}'\mathbf{Q}_0\overline{\overline{\mathbf{u}}})\big) \\ &= \frac{\sigma_{\mathrm{v}}^2}{\mathrm{N}}\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M}) \end{aligned} \tag{III.26}$$

$$\begin{aligned} \frac{1}{\mathrm{N(T\text{-}1)}}\mathrm{E}(\overline{\boldsymbol{\varepsilon}}'\mathbf{Q}_0\boldsymbol{\varepsilon}) &= \frac{1}{\mathrm{N(T\text{-}1)}}\mathrm{E}\big((\overline{\mathbf{u}}'-\rho\overline{\overline{\mathbf{u}}}')\mathbf{Q}_0(\mathbf{u}-\rho\overline{\mathbf{u}})\big) \\ &= \frac{1}{\mathrm{N(T\text{-}1)}}\begin{pmatrix} \mathrm{E}(\mathbf{u}'\mathbf{Q}_0\overline{\mathbf{u}}) - \rho\mathrm{E}(\overline{\mathbf{u}}'\mathbf{Q}_0\overline{\mathbf{u}}) - \\ \rho\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\overline{\overline{\mathbf{u}}}) + \rho^2\mathrm{E}(\overline{\mathbf{u}}'\mathbf{Q}_0\overline{\overline{\mathbf{u}}}) \end{pmatrix} \\ &= 0 \end{aligned} \tag{III.27}$$

$$\begin{aligned}
\frac{1}{\mathrm{N}}\mathrm{E}(\boldsymbol{\varepsilon}'\mathbf{Q}_1\boldsymbol{\varepsilon}) &= \frac{1}{\mathrm{N}}\mathrm{E}\big((\mathbf{u}'-\rho\bar{\mathbf{u}}')\mathbf{Q}_1(\mathbf{u}-\rho\bar{\mathbf{u}})\big) \\
&= \frac{1}{\mathrm{N}}\big(\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\mathbf{u})-2\rho\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\bar{\mathbf{u}})+\rho^2\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\mathbf{u}})\big) \\
&= \sigma_1^2
\end{aligned} \tag{III.28}$$

$$\begin{aligned}
\frac{1}{\mathrm{N}}\mathrm{E}(\bar{\boldsymbol{\varepsilon}}'\mathbf{Q}_1\bar{\boldsymbol{\varepsilon}}) &= \frac{1}{\mathrm{N}}\mathrm{E}\big((\bar{\mathbf{u}}'-\rho\bar{\bar{\mathbf{u}}}')\mathbf{Q}_1(\bar{\mathbf{u}}-\rho\bar{\bar{\mathbf{u}}})\big) \\
&= \frac{1}{\mathrm{N}}\big(\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\mathbf{u}})-2\rho\mathrm{E}(\bar{\bar{\mathbf{u}}}'\mathbf{Q}_1\bar{\mathbf{u}})+\rho^2\mathrm{E}(\bar{\bar{\mathbf{u}}}'\mathbf{Q}_1\bar{\bar{\mathbf{u}}})\big) \\
&= \frac{\sigma_1^2}{\mathrm{N}}\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M})
\end{aligned} \tag{III.29}$$

$$\begin{aligned}
\frac{1}{\mathrm{N}}\mathrm{E}(\bar{\boldsymbol{\varepsilon}}'\mathbf{Q}_1\boldsymbol{\varepsilon}) &= \frac{1}{\mathrm{N}}\mathrm{E}\big((\bar{\mathbf{u}}'-\rho\bar{\bar{\mathbf{u}}}')\mathbf{Q}_1(\mathbf{u}-\rho\bar{\mathbf{u}})\big) \\
&= \frac{1}{\mathrm{N}}\big(\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\bar{\mathbf{u}})-\rho\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\mathbf{u}})-\rho\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\bar{\bar{\mathbf{u}}})+\rho^2\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\bar{\mathbf{u}}})\big) \\
&= 0
\end{aligned} \tag{III.30}$$

with unknown parameters ρ, ρ^2, σ_ν^2 and σ_1^2. Equations (III.25)-(III.30) can now be transformed into a matrix notation, which then results in

$$\begin{bmatrix}
2a\cdot\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\bar{\mathbf{u}}) & -a\cdot\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_0\bar{\mathbf{u}}) & 1 & 0 \\
2a\cdot\mathrm{E}(\bar{\bar{\mathbf{u}}}'\mathbf{Q}_0\bar{\mathbf{u}}) & -a\cdot\mathrm{E}(\bar{\bar{\mathbf{u}}}'\mathbf{Q}_0\bar{\bar{\mathbf{u}}}) & b\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M}) & 0 \\
a\cdot(\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\bar{\bar{\mathbf{u}}})+\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_0\bar{\mathbf{u}})) & -a\cdot\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_0\bar{\bar{\mathbf{u}}}) & 0 & 0 \\
2b\cdot\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\bar{\mathbf{u}}) & -b\cdot\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\mathbf{u}}) & 0 & 1 \\
2b\cdot\mathrm{E}(\bar{\bar{\mathbf{u}}}'\mathbf{Q}_1\bar{\mathbf{u}}) & -b\cdot\mathrm{E}(\bar{\bar{\mathbf{u}}}'\mathbf{Q}_1\bar{\bar{\mathbf{u}}}) & 0 & b\cdot\mathrm{tr}(\mathbf{M}'\mathbf{M}) \\
b\cdot(\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\bar{\bar{\mathbf{u}}})+\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\mathbf{u}})) & -b\cdot\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\bar{\mathbf{u}}}) & 0 & 0
\end{bmatrix}
\bullet
\begin{pmatrix} \rho \\ \rho^2 \\ \sigma_\nu^2 \\ \sigma_1^2 \end{pmatrix}
=
\begin{bmatrix}
a\cdot\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\mathbf{u}) \\
a\cdot\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_0\bar{\mathbf{u}}) \\
a\cdot\mathrm{E}(\mathbf{u}'\mathbf{Q}_0\bar{\mathbf{u}}) \\
b\cdot\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\mathbf{u}) \\
b\cdot\mathrm{E}(\bar{\mathbf{u}}'\mathbf{Q}_1\bar{\mathbf{u}}) \\
b\cdot\mathrm{E}(\mathbf{u}'\mathbf{Q}_1\bar{\mathbf{u}})
\end{bmatrix} \tag{III.31}$$

where $a = 1/(\mathrm{N(T\text{-}1)})$ and $b = 1/\mathrm{N}$. Since the spatial autoregressive parameter ρ appears in both linear and quadratic terms in the coefficient vector, the above equation has to be estimated using non-linear least squares using a consistent estimates of residuals $\tilde{\mathbf{u}}$ for $\mathbf{u}$, $\tilde{\bar{\mathbf{u}}}$ for $\bar{\mathbf{u}}$ and $\tilde{\bar{\bar{\mathbf{u}}}}$ for $\bar{\bar{\mathbf{u}}}$. In order to obtain consistent estimates of the residuals, the model can be estimated using a two-stage least-squares regression.

3.4.2 Obtaining Consistent Estimates of the Residuals **u**

In a first step, the spatially lagged dependent variable $\mathbf{y}_s = (\mathbf{W} \otimes \mathbf{I}_T) \cdot \mathbf{y}$ is regressed on the spatially lagged exogenous variables, being $(\mathbf{W} \otimes \mathbf{I}_T) \cdot \mathbf{X}$, $(\mathbf{W} \otimes \mathbf{I}_T)^2 \cdot \mathbf{X}$, $(\mathbf{W} \otimes \mathbf{I}_T)^3 \cdot \mathbf{X}$ and so on (more than twice spatially-lagged instruments are unlikely to be necessary). The instruments can be pooled in a matrix $\mathbf{Z} = ((\mathbf{W} \otimes \mathbf{I}_T)\mathbf{X}, (\mathbf{W} \otimes \mathbf{I}_T)^2 \mathbf{X}, ...)$. The prediction of $\mathbf{y}_s$ from this first-stage regression then is

$$\hat{\mathbf{y}}_s = \mathbf{Z} \cdot (\mathbf{Z}' \cdot \mathbf{Z})^{-1} \cdot \mathbf{Z}' \cdot \mathbf{y}_s$$

Using the result of this first-stage regression, the spatially lagged dependent variable is predicted and used in the second regression, in which the dependent variable $\mathbf{y}$ is regressed on the prediction for the spatially lagged dependent variable of the first regression ($\hat{\mathbf{y}}_w$) and on the exogenous regressors $\mathbf{X}$. This pair of regressors can be combined in a regressors' matrix $\mathbf{X}_+ = (\hat{\mathbf{y}}_s, \mathbf{X})$. From this second stage, consistent estimates for the residuals $\mathbf{u}$ are obtained:

$$\tilde{\mathbf{u}} = \mathbf{y} - \mathbf{X}_+ \cdot (\mathbf{X}_+' \cdot \mathbf{X}_+)^{-1} \cdot \mathbf{X}_+' \cdot \mathbf{y} = \left(\mathbf{I}_{NT} - \begin{bmatrix} \hat{\mathbf{y}}_s & \mathbf{X} \end{bmatrix} \cdot \left(\begin{bmatrix} \hat{\mathbf{y}}_s' \\ \mathbf{X}' \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{y}}_s & \mathbf{X} \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} \hat{\mathbf{y}}_s' \\ \mathbf{X}' \end{bmatrix} \right) \mathbf{y} \qquad \text{(III.32)}$$

Equation (III.32) can now be used to estimate ρ from equation (III.31).

3.4.3 Obtaining Consistent Estimates of the Spatial Autoregressive Parameter ρ

Kapoor et al. (2007) describe several ways to estimate equation (III.31) consistently. It can be seen that the first three rows are sufficient to estimate ρ and σ_v^2 using equal weights (e.g. $\mathbf{I}_{3\times3}$ could be such a weighting matrix). σ_1^2 then can easily be obtained from the fourth row of equation (III.31). It should be noted that it does not matter which weighting regime is chosen for the non-linear estimation. However, it is common knowledge from the literature on GMM estimators that it is optimally efficient to use the inverse of the variance covariance matrix of the sample moments.

However, a question which might arise is whether the omission of the last three rows of equation (III.31) leads to biased estimates of the parameters involved. Kapoor et al. (2007) showed that the small sample properties of the respective estimates in terms of efficiency and consistency are better when only the first three rows of equation (III.31) are used. The present paper, however, relies on a large sample and accordingly, simulations were made to see whether the conclusion of Kapoor et al. (2007) can be extended. Results can be found in the Appendix 6.2. Accordingly, the preferred estimation method is that in which only the first three rows of information in equation (III.31) are used.

3.4.4 Spatial Feasible Generalised Least Squares Estimation (S-FGLS)

The next step is quite straightforward. Once the estimates of the parameters ρ, σ_v^2 and σ_1^2 are available, the GLS estimator for the coefficients of the exogenous regressors can be written as

$$\hat{\boldsymbol{\beta}}_{\mathbf{GLS\varepsilon}} = \left[\mathbf{X}(\hat{\rho})' \cdot \boldsymbol{\Omega}\ (\hat{\sigma}_v^2, \hat{\sigma}_\mu^2)^{-1} \cdot \mathbf{X}(\hat{\rho})\right]^{-1} \cdot \mathbf{X}(\hat{\rho})' \cdot \boldsymbol{\Omega}\ (\hat{\sigma}_v^2, \hat{\sigma}_\mu^2)^{-1} \cdot \mathbf{y}(\hat{\rho}) \qquad \text{(III.33)}$$

where $\mathbf{X}(\hat{\rho}) = ((\mathbf{I}_N - \hat{\rho} \cdot \mathbf{W}) \otimes \mathbf{I}_T) \cdot \mathbf{X}$ and $\mathbf{y}(\hat{\rho}) = ((\mathbf{I}_N - \hat{\rho} \cdot \mathbf{W}) \otimes \mathbf{I}_T) \cdot \mathbf{y}$ can be viewed as a spatial Cochrane-Orcutt transformation of the model. From this, as described in Kapoor et al. (2007), spatial feasible GLS can be applied by substituting the point estimates of the parameters in equation (III.33). Alternatively, in addition to the Cochrane-Orcutt transformation, the exogenous variables and the dependent variable could be transformed once more by a standard transformation, namely $\widehat{\mathbf{X}}(\hat{\rho}) = (\mathbf{I}_{NT} - \theta \mathbf{Q}_1) \cdot \mathbf{X}(\hat{\rho})$ and $\widehat{\mathbf{y}}(\hat{\rho}) = (\mathbf{I}_{NT} - \theta \mathbf{Q}_1) \cdot \mathbf{y}(\hat{\rho})$, where $\theta = 1 - \sigma_v / \sigma_1$.

This transformation removes the equi-correlation due to the presence of individual effects in the residuals. Accordingly, the estimates obtained by equation (III.33) are consistent with the respective OLS estimates using the twice-transformed model:

$$\hat{\boldsymbol{\beta}}_{\mathbf{OLS}} = \left[\widehat{\mathbf{X}}(\hat{\rho})' \cdot \widehat{\mathbf{X}}(\hat{\rho})\right]^{-1} \cdot \widehat{\mathbf{X}}(\hat{\rho})' \cdot \widehat{\mathbf{y}}(\hat{\rho}) \qquad \text{(III.34)}$$

The proofs that $\hat{\boldsymbol{\beta}}_{\mathbf{GLS}}$ converges in probability to $\boldsymbol{\beta}$ and that $\hat{\boldsymbol{\beta}}_{\mathbf{FGLS}}$ converges in probability to $\hat{\boldsymbol{\beta}}_{\mathbf{GLS}}$ (which in turn means that both $\hat{\boldsymbol{\beta}}_{\mathbf{GLS}}$ and $\hat{\boldsymbol{\beta}}_{\mathbf{FGLS}}$ are consistent) can be found in Kapoor et al. (2007). In addition, note that fixed effects corresponds to $\theta = 1$, which means that the Cochrane-Orcutt transformation is followed by a time-demeaning transformation. Then, the coefficient estimates described in equation (III.34) corresponds to the least squares dummy variable estimator (LSDV). The fixed effects μ_i can then be derived using $\hat{\boldsymbol{\mu}} = \mathbf{Q}_1(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{OLS}})$. Again, the disadvantage of fixed effects here is that first, fixed effects are only consistent for large T and, second, time-invariant variables cancel out. Further, one has to account for the fact that the estimation of the N fixed effects changes the variance-covariance matrix of the estimated coefficients.

3.5 Specification Testing

The preceding subsection explained the implementation of the GMM estimator for the SARAR model. The underlying assumptions of this model are spatial correlation in the dependent variable and the residuals and the presence of individual effects. These are assumptions which can be tested. A large number of specification tests have recently been developed for spatial econometric models using panel data. Most publications focus on testing alternatives consisting of spatial interdependences such as the spatial lag in the dependent variable and in the residuals, random effects, and serial correlation.

Among others, several contributions can be found in Baltagi and Long (2008), Baltagi et al. (2003), Kelejian and Prucha (2001) and Sen and Bera (2011). Most recently developed tests consist of extensions of Lagrange multiplier tests (LM tests) in the sense that their specifications can both jointly and conditionally robustly tested. To our knowledge, basically three different approaches are taken in testing spatial econometric specification using panel data. The first considers a spatial error model with random effects. In the second, a spatial lag model with random effects is considered in which the presence of a spatial lag and random effects are tested in a joint and robust version of the developed test statistic. To our knowledge, the only study in which the testing of a spatial lag model with fixed effects is described was proposed by Debarsy and Ertur (2010). Pfaffermayr developed a Hausman test to discriminate between fixed and random effects specification in panel data models (Michael Pfaffermayr 2011). From the first two approaches, it becomes clear that a remaining problem is that, so far, no LM test statistic is available which combines the two approaches to test jointly and robustly for all alternatives mentioned, namely a spatial lag, spatial autocorrelation, and random effects. Sen and Bera (2011) develop Rao's score test statistics, which were intended to fill this gap. The approach is, as previously mentioned, very useful, since the model only needs to be estimated under the null hypothesis, which is a pooled OLS model. They even extend the approach by incorporating a further alternative into the specification, namely the presence of serial correlation in the residuals. Accordingly, the set of tests available can be displayed as in Figure III-6. To our knowledge, there is as yet no test statistic which robustly tests for all possible combinations in the context of the SARAR model, despite the test described by Sen and Bera (2011).

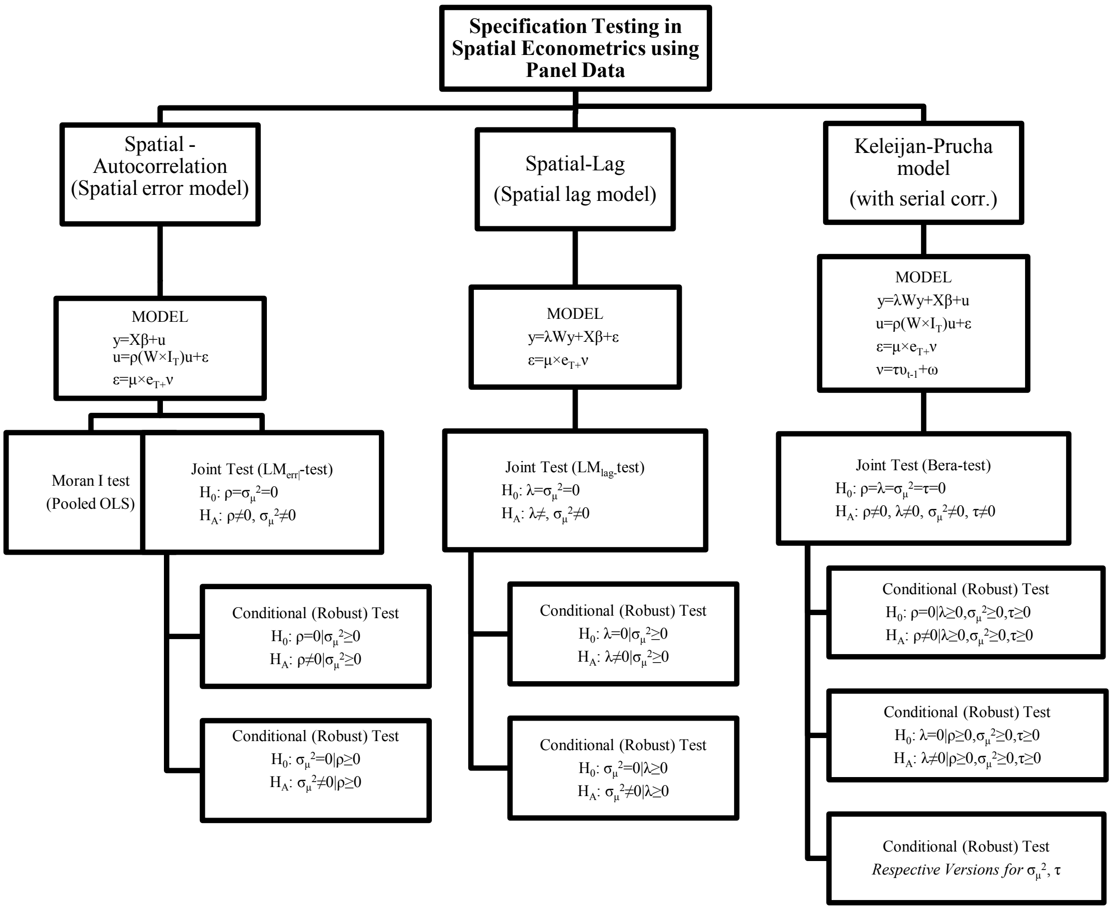


Figure III-6: Specification testing in spatial econometrics using panel data

In the following, these test statistics are described. Most of all, we focus on the conditional or robust versions of the statistics as described in Figure III-6. We do not provide the corresponding test statistics of the Bera test, since the derivation of those statistics uses a different approach with maximum likelihood residuals and is based on Rao's scores. Furthermore, we do not provide the conditional test statistics of random effects conditional on the presence of spatial spillovers, since the focus here is on the test versions concerning the spatial lag and spatial autocorrelation coefficient.

For our empirical application, we intend to use the first two branches of tests, those described for the spatial error model and those for the spatial lag model.

3.5.1 The Moran-I Test

Kelejian and Prucha (2001) derive a variety of Moran I test statistics. Most of all, they derive the asymptotic distribution of a Moran I test statistic to check for spatial auto-correlation (spatial correlation in the residuals) of a Kelejian-Prucha model. They define the Moran I statistic as:

$$I = \frac{\hat{\mathbf{u}}' \cdot \mathbf{W} \cdot \hat{\mathbf{u}}}{\sqrt{\hat{\sigma}^4 \cdot \mathrm{tr}(\mathbf{W} \cdot \mathbf{W} + \mathbf{W}' \cdot \mathbf{W}) + \hat{\sigma}^2 \cdot \hat{\mathbf{b}}' \hat{\mathbf{b}}}} \xrightarrow[d]{} \mathrm{N}(0,1) \tag{III.35}$$

where

$$\hat{\mathbf{b}} = -\mathbf{H} \cdot \mathbf{P}' \cdot \frac{1}{\mathrm{N}} \hat{\mathbf{u}}'(\mathbf{W} + \mathbf{W}')\mathbf{D}$$

and

$$\mathbf{D} = (\mathbf{Wy}, \mathbf{X})$$

$$\mathbf{H} = (\mathbf{X}, \mathbf{WX}, \mathbf{W}^2\mathbf{X}, \ldots)$$

$$\hat{\mathbf{D}} = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\mathbf{D}$$

$$\mathbf{P} = \left(\frac{1}{\mathrm{N}} \hat{\mathbf{D}}' \hat{\mathbf{D}}\right)^{-1} \cdot \frac{1}{\mathrm{N}} \mathbf{D}'\mathbf{H} \cdot \left(\frac{1}{\mathrm{N}} \mathbf{H}'\mathbf{H}\right)^{-1}$$

$$\hat{\mathbf{u}} = \mathbf{y} - \mathbf{D} \cdot (\hat{\mathbf{D}}'\hat{\mathbf{D}})^{-1} \hat{\mathbf{D}}'\mathbf{y}$$

The originally developed Moran I test statistic for cross-sectional data is defined as

$$I_c = \frac{\hat{\mathbf{u}}_{\mathrm{ols}}' \cdot \mathbf{W} \cdot \hat{\mathbf{u}}_{\mathrm{ols}}}{\hat{\mathbf{u}}_{\mathrm{ols}}' \cdot \hat{\mathbf{u}}_{\mathrm{ols}}}$$

and tests for the presence of spatial correlation, where it is not clear whether the correlation arises from the residuals, from the dependent variable or from both. Kelejian and Prucha went a step further and redefined the Moran I test statistic for the Kelejian-Prucha model (previously described) to test for spatial correlation in the residuals. Although the test is developed to test for spatial autocorrelation in cross-sectional data, it could also be used for panel data as a first indication of spatial auto-correlation[23].

The substantial contribution of the test is that the widely used Moran I test statistic is extended to check for spatial auto-correlation in a Keleina-Prucha model, which in consequence is robust to a possible presence of a spatially lagged dependent variable.

---

[23] Monte Carlo results of the test statistic used in a panel data framework can be found in the Appendix.

3.5.2 The Lagrange Multiplier Test for Spatial Autocorrelation (LM_{err}-Test)

Baltagi et al. (2003) develop a test to check for spatial auto-correlation and random effects in panel data. In a newer version of the study, a test is elaborated which jointly and robustly tests for the presence of serial correlation, spatial autocorrelation and random effects. However, we do not intend to widen the focus on serial correlation. More details can be found in Baltagi et al. (2007). The main contribution is that the test statistics are derived both as joint tests and the corresponding robust versions. The Breusch and Pagan LM test is expanded to a spatial error component model. For this reason, the following panel data model is considered:

$$\begin{aligned}
\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\
\mathbf{u} &= \rho\cdot(\mathbf{W}\otimes\mathbf{I}_T)\mathbf{u} + \boldsymbol{\varepsilon} \\
\boldsymbol{\varepsilon} &= \boldsymbol{\mu}\otimes\mathbf{e}_T + \mathbf{v}
\end{aligned} \tag{III.36}$$

Equation (III.36) represents a spatial error model in a random effects panel data framework. The dimensions of the matrices and vectors are correspondingly transformed. $\boldsymbol{\mu}$ is a vector of $N\times 1$ individual effects which may be fixed or random, $\mathbf{v}$ is a $NT\times 1$ vector of disturbances.

First, Baltagi develops a joint statistic test for the presence of random effects and spatial auto-correlation. The null hypothesis is

$$H_{0\mu}: \rho = \sigma^2 = 0$$

which is tested against the alternative that at least one component is non-zero. The two-sided test statistic is

$$LM_{\rho,\sigma_\mu^2} = \frac{NT}{2(T-1)}\left(\frac{\hat{\mathbf{u}}_{ols}'\cdot(\mathbf{I}_N\otimes\mathbf{J}_T)\cdot\hat{\mathbf{u}}_{ols}}{\hat{\mathbf{u}}_{ols}'\cdot\hat{\mathbf{u}}_{ols}}\right)^2 + \frac{N^2T}{tr(\mathbf{W}^2+\mathbf{W}'\cdot\mathbf{W})}\left(\frac{\hat{\mathbf{u}}_{ols}'\cdot(\mathbf{I}_N\otimes\mathbf{I}_T)\cdot\hat{\mathbf{u}}_{ols}}{\hat{\mathbf{u}}_{ols}'\cdot\hat{\mathbf{u}}_{ols}}\right)^2 \tag{III.37}$$

Baltagi and Long (2008) also elaborated a one-sided test statistic which we finally implemented in STATA®. The one-sided version looks similar to the one above. However, it is derived from the marginal tests. The one-sided version follows a weighted chi-square distribution and thus levels of significance can easily be obtained. The test is important, since if heterogeneity and spatial auto-correlation were neglected from a respective model, although it would describe the true data generating process, the resulting standard deviations of coefficient estimates would be biased and hence inference would be misleading.

In a second step, the conditional Lagrange multiplier test was developed. The conditional test is important, since the model is investigated for the presence of spatial auto-correlation no matter whether random effects are present or absent. The null hypothesis is

$$H_{0\mu}: \rho = 0 \mid \sigma^2 \geq 0$$

The corresponding test statistic is

$$\mathrm{LM}_{\rho|\sigma_\mu^2 \geq 0} = \frac{\hat{\mathbf{D}}(\rho)}{\sqrt{\left((\mathrm{T}-1)+\frac{\hat{\sigma}_\mathrm{v}^4}{\hat{\sigma}_1^4}\right)\mathbf{b}}} \qquad \text{(III.38)}$$

which asymptotically follows a standard normal distribution. Moreover, we have

$$\hat{\mathbf{D}}(\rho) = \frac{1}{2}\hat{\mathbf{u}}'\left[\frac{\hat{\sigma}_\mathrm{v}^4}{\hat{\sigma}_{1\mathrm{v}}^4}\cdot((\mathbf{W}'+\mathbf{W})\otimes(\frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T})) + \frac{1}{\hat{\sigma}^2}\cdot((\mathbf{W}'+\mathbf{W})\otimes(\mathbf{I}_\mathrm{T} - \frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T}))\right]\hat{\mathbf{u}}$$

$\hat{\mathbf{u}}$ is a vector of the residuals under the null-hypothesis and hence the residuals of a random effects estimation. Further,

$$\hat{\sigma}_\mathrm{v}^2 = \frac{1}{\mathrm{N(T\text{-}1)}}\cdot\hat{\mathbf{u}}'\cdot(\mathbf{I}_\mathrm{N}\otimes(\mathbf{I}_\mathrm{T} - \frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T}))\cdot\hat{\mathbf{u}}$$

$$\hat{\sigma}_1^2 = \frac{1}{\mathrm{N}}\cdot\hat{\mathbf{u}}'\cdot(\mathbf{I}_\mathrm{N}\otimes\frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T})\cdot\hat{\mathbf{u}}$$

The conditional Lagrange multiplier test guards against a misleading conclusion when interpreting group-wise correlation due to random effects such as spatial auto-correlation.

For the sake of completeness, it should be mentioned that Baltagi and Long (2008) also provide a conditional LM test statistic to test for random effects given the possible presence of spatial auto-correlation. This is the modified Breusch and Pagan test mentioned previously. We incorporated the test statistic as well into STATA® but are not presenting it here, since this conditional test statistic is not needed to test for spatial auto-correlation.

3.5.3 The Lagrange Multiplier Test for Spatial Lag Dependence (LM_{lag}-Test)

Baltagi and Long (2008) develop a further test statistic to jointly test the presence of spatial lag dependence and random effects and also develop the corresponding robust versions of the test, namely first a test for the presence of spatial lag dependence given the possible presence of random effects and second a test for random effects given the possible presence of spatial lag dependence.

The following panel data model is considered:

$$\begin{aligned} &\mathbf{y} = \lambda\cdot\boldsymbol{\beta}\mathbf{W}\boldsymbol{\varepsilon}\otimes\mathbf{I}_\mathrm{T}) + \mathbf{X} \quad + \\ &\boldsymbol{\varepsilon} = \boldsymbol{\mu}\otimes\mathbf{e}_\mathrm{T} + \mathbf{v} \end{aligned} \qquad \text{(III.39)}$$

The model stated in equation (III.39) represents the spatial lag model presented in Figure III-1 in a random effects panel data framework.

The joint Lagrange multiplier test statistic to test the null hypothesis

$$\mathrm{H}_{0\mu}\colon \lambda = \sigma^2 = 0$$

is given by

$$\mathrm{LM}_{\lambda,\sigma_\mu^2} = \frac{\mathbf{R}^2}{\mathbf{B}} + \frac{\mathrm{NT}}{2(\mathrm{T}\text{-}1)}\mathbf{G}^2 \tag{III.40}$$

where

$$\mathbf{B} = \mathrm{T}\cdot \mathrm{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W}) + \frac{1}{\hat{\sigma}_\nu^2}\hat{\boldsymbol{\beta}}_{ols}'\mathbf{X}'(\mathbf{W}'\otimes\mathbf{I}_\mathrm{T})\mathbf{M}(\mathbf{W}\otimes\mathbf{I}_\mathrm{T})\mathbf{X}\hat{\boldsymbol{\beta}}_{ols}$$

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$\mathbf{G} = \mathrm{T}\frac{\hat{\mathbf{u}}_{ols}'\cdot(\mathbf{I}_\mathrm{N}\otimes\frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T})\cdot\hat{\mathbf{u}}_{ols}}{\hat{\mathbf{u}}_{ols}'\cdot\hat{\mathbf{u}}_{ols}} - 1$$

$$\mathbf{R} = \mathrm{NT}\cdot\frac{\hat{\mathbf{u}}_{ols}'\cdot(\mathbf{W}\otimes\mathbf{I}_\mathrm{T})\cdot\mathbf{y}}{\hat{\mathbf{u}}_{ols}'\cdot\hat{\mathbf{u}}_{ols}}$$

The test statistic presented in equation (III.40) is asymptotically distributed as a chi-square distribution with two degrees of freedom. The test is important, since if heterogeneity and spatial lag dependence were neglected from a respective model, the resulting standard deviations of coefficient estimates would be biased and hence inference would be misleading. Further, the negligence of the spatially lagged dependent variable would represent an omitted variable bias and accordingly, the estimated coefficients would be biased.

However, the test presented in equation (III.40) implicitly assumes the absence of random effects under the null hypothesis. To overcome the problem, Baltagi and Long (2008) derive a conditional test statistic to test for spatial lag dependence given the possible presence of random effects. The null hypothesis is

$$\mathrm{H}_{0\mu}\colon \lambda = 0 \mid \sigma^2 \geq 0$$

and the respective conditional Lagrange multiplier test statistic is

$$\mathrm{LM}_{\lambda|\sigma_\mu^2\geq 0} = \frac{\mathbf{R}_1{}^2}{\mathbf{B}_1} \tag{III.41}$$

where

$$\mathbf{R}_1 = \frac{1}{\hat{\sigma}_{1\nu}^2}\cdot\hat{\mathbf{u}}'(\mathbf{W}\otimes\frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T})\mathbf{y} + \frac{1}{\hat{\sigma}^2}\cdot\hat{\mathbf{u}}'(\mathbf{W}\otimes(\mathbf{I}_\mathrm{T} - \frac{1}{\mathrm{T}}\mathbf{J}_\mathrm{T}))\mathbf{y}$$

$$\hat{\sigma}_{\nu}^{2} = \frac{1}{N(T\text{-}1)} \cdot \hat{\mathbf{u}}' \cdot (\mathbf{I}_N \otimes (\mathbf{I}_T - \frac{1}{T}\mathbf{J}_T)) \cdot \hat{\mathbf{u}}$$

$$\hat{\sigma}_{1}^{2} = \frac{1}{N} \cdot \hat{\mathbf{u}}' \cdot (\mathbf{I}_N \otimes \frac{1}{T}\mathbf{J}_T) \cdot \hat{\mathbf{u}}$$

$$\mathbf{B_1} = \mathrm{T} \cdot \mathrm{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W}) + \frac{1}{\hat{\sigma}_1^2} \hat{\boldsymbol{\beta}}'\mathbf{X}'(\mathbf{W}'\mathbf{W} \otimes \frac{1}{T}\mathbf{J}_T) \cdot \mathbf{X}\hat{\boldsymbol{\beta}}$$
$$+ \frac{1}{\hat{\sigma}_{\nu}^2} \hat{\boldsymbol{\beta}}'\mathbf{X}' \cdot (\mathbf{W}'\mathbf{W} \otimes (\mathbf{I}_T - \frac{1}{T}\mathbf{J}_T)) \cdot \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{A}\left[\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X}\right]^{-1}\mathbf{A}$$

$$\mathbf{A} = \left[\frac{1}{\hat{\sigma}_{1\nu}^2} \hat{\boldsymbol{\beta}}'\mathbf{X}'(\mathbf{W} \otimes \frac{1}{T}\mathbf{J}_T)\mathbf{X}\hat{\mathbf{I}} + \frac{1}{\hat{\sigma}^2} \hat{\boldsymbol{\beta}}'\mathbf{X}'(\mathbf{W} \otimes (\mathbf{I}_T - \frac{1}{T}\mathbf{J}_T))\mathbf{X} \right]$$

$$\hat{\boldsymbol{\Omega}} = \hat{\sigma}_{\mu}^{2} \cdot (\mathbf{I}_N \otimes \mathbf{J}_T) + \hat{\sigma}_{\nu}^{2} \cdot (\mathbf{I}_N \otimes \mathbf{I}_T), \; \hat{\sigma}_{\mu}^{2} = \frac{\hat{\sigma}_{1\nu}^{2} - \hat{\sigma}^{2}}{T}$$

The test statistic presented in equation (III.41) is asymptotically chi-square distributed with one degree of freedom.

Here too, a second conditional test for random effects due to the possible presence of spatial lag dependence was elaborated. We implemented this too, as all other tests, in STATA® but are not providing the test statistics here. The justification for applying the corresponding tests is the same as previously explained.

4 Gasoline Demand in Swiss Municipalities - Empirical Application

In the first paper of this dissertation, we undertook an estimation of gasoline tourism in Switzerland's border regions. In this empirical part of the second paper, we want to estimate a Swiss gasoline demand function using aggregate regional panel data and to apply the spatial econometric methods discussed in the previous sections.

The main differences between this empirical analysis and the analysis presented in the first paper are several. First, the database is extended to cover the whole Swiss territory and not only the border cantons, and second, the focus of the application changes. We analyse gasoline demand at the municipal level and are no longer interested in gasoline tourism. Third, since there are very few studies estimating a SARAR model using GMM, we want to investigate Swiss gasoline demand from a spatial perspective. For this reason, we use the statistical tests described previously to decide whether Swiss gasoline demand at the municipal level is driven by spatial correlation. Further, we estimate several spatial econometric models to arrive at information on the price and income elasticities of the Swiss gasoline demand and their spatial variation.

4.1 Determinants and Functional Form of Swiss Gasoline Demand

Based on the discussion on the choice of the model specification in the first paper and considering the spatial econometric approaches, we formulate the following gasoline demand function:

$$G_{it} = f(WG_{it}, PG_{CH,bt}, PG_{F,bt}, dist_i, CARSG_{CH,it}, CARSD_{CH,it}, POPCH_{it}, Y_{CH,it}, COMM_{CH,it}) \quad \text{(III.42)}$$

where G_{it} is the average sales per gasoline station. Further, sales depend on gasoline demand in neighbouring municipalities, $\sum_j w_{ij} G_{it}$, where w_{ij} is the (i,j) element of the spatial weight matrix $\mathbf{W}$. Sales also depend on the real Swiss gasoline price and the real foreign gasoline price, $PG_{CH,bt}$ and $PG_{F,bt}$, the municipality's distance from the border, $dist_i$, the stock of gasoline and diesel powered vehicles, $CARSG_{CH,it}$ and $CARSD_{CH,it}$, the municipality's population and taxable income, $POP_{CH,it}$ and $Y_{CH,it}$ and on the number of foreign commuters, $COMM_{CH,it}$. The meaning of the indices is the same as in the first paper. The data base was extended to cover sales from the whole Swiss territory and not only from the border cantons. Accordingly, the sample now consists in observations from i=1..547 municipalities. Considering a log-log functional form, the demand equation (III.42) can be expressed as:

$$\begin{aligned}
\ln(G_{it}) &= \alpha_0 + \lambda \cdot \sum_j w_{ij} G_{it} + \alpha_1 \ln(PG_{CH,bt}) + \alpha_2 \ln(PG_{F,bt}) + \alpha_3 \ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right) + \\
&\quad \alpha_4 \ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right) + \alpha_5 \ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right) + \alpha_6 \ln(Commu_{it}) + \alpha_7 \ln(dist_i) + \alpha_8 DB_{it} + u_{it} \qquad \text{(III.43)} \\
u_{it} &= \rho \cdot \sum_j w_{ij} \cdot u_{jt} + \varepsilon_{it} \\
\varepsilon_{it} &= \mu_i + \nu_{it}
\end{aligned}$$

First, the general empirical model as stated in equation (III.43) contains a spatially lagged dependent variable and spatially lagged residuals. Second, we decided to include the Swiss gasoline price and the foreign gasoline price and not the price ratio of the Swiss and foreign gasoline price. In the first paper, this price ratio was included in order to identify gasoline tourism by setting the ratio of the two prices equal to one. Actually, this simulation would be possible here as well if one equalised the Swiss gasoline price with the foreign gasoline price. The reason to include the price ratio, however, was that it can be influenced directly in the simulation while leaving the level of the Swiss gasoline price unaffected. Consequently, in this paper, we prefer not to include this ratio, since we are no longer interested in estimating gasoline tourism. It has been mentioned previously that it can be assumed that the foreign gasoline price shows a significant impact on domestic demand given the small size of Switzerland.

Moreover, the number of gasoline and diesel powered cars are standardised with the domestic population. The fact that we include the number of cars could suggest that we are estimating rather a short-run gasoline demand model. However, we tend to think that our model specification reflects more a long-run equilibrium situation for the following reasons: First we are not considering the level of energy efficiency of the cars. Of course, this is an important factor in the optimization process of the consumers. Second, as suggested by Griffin and Baltagi (1984), OLS, between and random effects models tend to reflect rather the long-run optimization process than the short-run one.

Apart from this, the model specification looks very similar to that used the first part. However, cantonal dummies to account for canton-specific heterogeneity are not incorporated in the model. The price elasticity of gasoline demand with respect to the Swiss gasoline price can be calculated as:

$$\varepsilon_{PG_{CH,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{CH,bt})} = \alpha_1 \tag{III.44}$$

The price elasticity of gasoline demand with respect to the foreign gasoline price then is

$$\varepsilon_{PG_{F,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{F,bt})} = \alpha_2 \tag{III.45}$$

The income elasticity of gasoline demand is

$$\varepsilon_{Y_{CH,it}} = \frac{\partial \ln(G_{it})}{\partial \ln(Y_{CH,it})} = \alpha_5 \tag{III.46}$$

These equations represent the direct effects of a change in the price or in the income. The coefficients can be interpreted directly as elasticities due to the log-log functional form of the model. The coefficient α_1 is expected to be negative since a ceteris paribus increase in the Swiss gasoline price will decrease domestic demand. Further, α_1 should take a value which is similar to the one stated in the first part of the dissertation where the average elasticity of Swiss gasoline demand was calculated for the municipalities depending on the distance from the border. The same holds for the coefficient α_2, which should take a positive value but similar to the average over all distance classes as reported in the first part of the dissertation. Similarly, the elasticity with respect to per-capita income, α_5, should be positive. Recall that this very simple definition may be misleading. As previously mentioned, where direct, indirect and total effects are explained, a ceteris paribus change in a municipality’s independent variable has a direct effect on that municipality, stated by the coefficients above, which are nothing else than the first-order spatial effect, but are superposed by the spatial effects described. To account for this superposition, the model stated in equation (III.43) can be rewritten to

$$
\begin{aligned}
\ln(G_{it}) &= \sum_j (\mathbf{I}_{\mathrm{NT}} - \lambda \cdot \mathbf{W} \otimes I_T)^{-1}{}_j \cdot \quad \left(\alpha_0 + \alpha_1 \ln(PG_{CH,bt}) + \alpha_2 \ln(PG_{F,bt}) + \alpha_3 \ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right)\right. \\
&\qquad + \alpha_4 \ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right) + \alpha_5 \ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right) + \alpha_6 \ln(Commu_{it}) \\
&\qquad \left. + \alpha_7 \ln(dist_i) + \alpha_8 DB_{it} + u_{it}\right) \qquad \text{(III.47)} \\
u_{it} &= \rho \cdot \sum_j w_{ij} \cdot u_{jt} + \varepsilon_{it} \\
\varepsilon_{it} &= \mu_i + v_{it}
\end{aligned}
$$

Accordingly, the elasticities can be rewritten to

$$
\varepsilon_{i,PG_{CH,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{CH,bt})} = \alpha_1 \cdot \sum_j (\mathbf{I}_{\mathrm{NT}} - \lambda \cdot \mathbf{W} \otimes I_T)^{-1}{}_{ij}
$$

$$
\varepsilon_{i,PG_{F,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(PG_{F,bt})} = \alpha_2 \cdot \sum_j (\mathbf{I}_{\mathrm{NT}} - \lambda \cdot \mathbf{W} \otimes I_T)^{-1}{}_{ij}
$$

$$
\varepsilon_{i,Y_{CH,bt}} = \frac{\partial \ln(G_{it})}{\partial \ln(Y_{CH,bt})} = \alpha_5 \cdot \sum_j (\mathbf{I}_{\mathrm{NT}} - \lambda \cdot \mathbf{W} \otimes I_T)^{-1}{}_{ij}
$$

And thus every spatial unit, or municipality, is assigned a different elasticity due to the presence of spatial spillovers. Of course, it does not make sense to report elasticities for all municipalities separately, but they will be reported as the previously stated average total, average direct and average indirect effects.

4.2 Specification of the Spatial Weighting Matrix **W**

The definition of the spatial weighting matrix is somewhat arbitrary, as seen previously. Since our goal is to analyse gasoline demand at the municipal level, on the available data and the size of the municipalities, we believe that a spatial weighting scheme in which absolute distance, in terms of geographical distance, matters is appropriate. However, the term geographical distance or proximity needs to be defined more exactly.

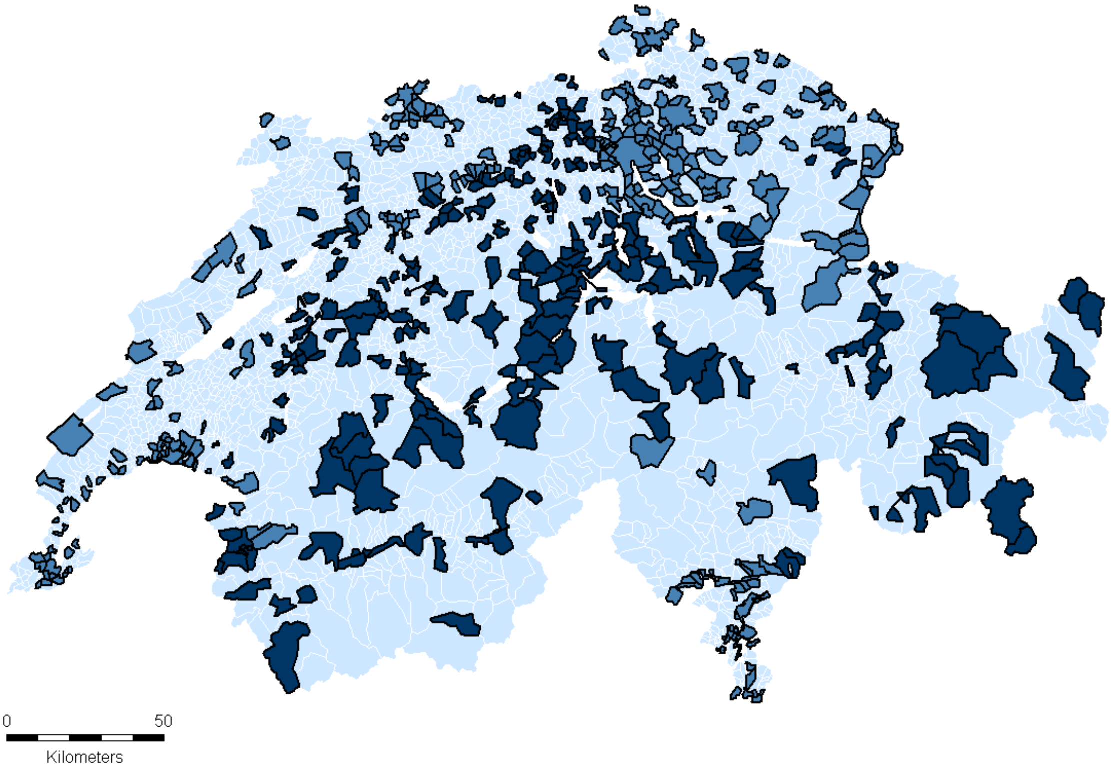


Figure III-7: Swiss municipalities with balanced gasoline sales data

Figure III-7 shows the Swiss map at the municipal level. The bright blue shaded municipalities denote those with no observation on gasoline sales available. The darker shaded ones denote those for which gasoline sales data were available for the first study of the dissertation (315 municipalities in the 12 border cantons excluding Wallis, Bern and Solothurn). The dark shaded ones are those municipalities which are considered in addition in the present study to estimate gasoline demand in Switzerland. In total, 547 municipalities are available over eight years (2001 – 2008). It is easy to see that, for the present case, the term proximity or neighbourhood in terms of the definition of a spatial weighting matrix is not easy to define, since there are municipalities on the map which do not have neighbouring municipalities. Following subsection 3.2.1 and according to the comments by Stakhovych and Bijmolt (2009), two different concepts of neighbourhood will be considered:

The first is that we define a municipality to be a neighbour (of degree one) to another one if it belongs to its five closest surrounding municipalities in terms of geographical distance. One could consider the three, four, six or even ten closest neighbours. For clarification, consider the Swiss map depicted in Figure III-8. Every municipality with available data (central coordinates depicted as points) is connected with its five closest neighbours. However, as previously mentioned, it is not necessary that, if municipality i is a neighbour to municipality j, the opposite is the case too. Accordingly, the resulting spatial matrix would not be symmetric. Increasing the number of neighbours would result in very large distances in some instances, while decreasing it might lead to a network which is represented by sub-regions which are not connected with each other.

In order to define neighbourhood in more intuitive terms, the Swiss map depicted in Figure III-7 can be triangulated according to the Delaunay triangulation described previously. The result is shown in Figure III-9. It can be seen that the spatial units are no longer defined by the original municipality borders but by Thiessen polygons (Voronoi diagram). With these polygons, (first-order degree) neighbourhood can be defined if the polygons of two municipalities share a common point or edge. It has already been seen in subsection 3.2.1 that this concept of neighbourhood usually represents the data structure quite well. Further, the Delaunay triangulation connects municipalities which are close to each other in the mathematical sense and not by a predefined concept, such as considering all municipalities within a distance range of some kilometers or only considering, for instance, the five closest 'neighbours'.

Figure III-8: Spatial dependence among Swiss municipalities with five closest neighbours

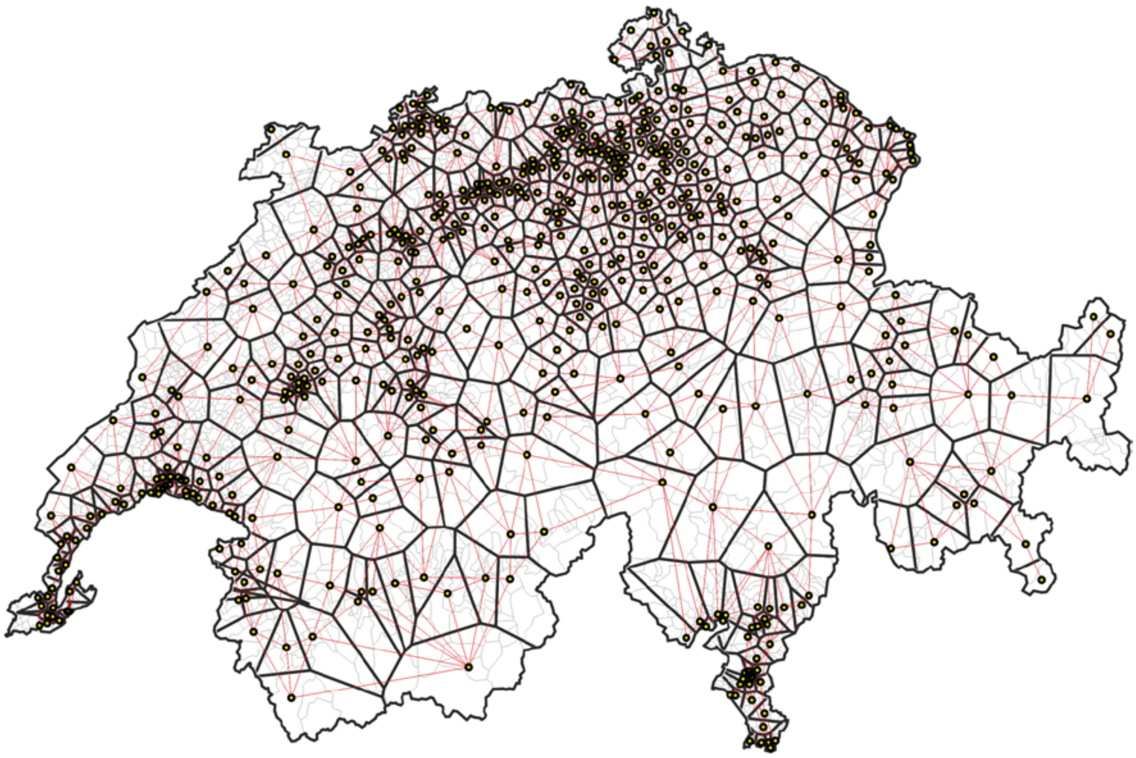


Figure III-9: Spatial dependence among Swiss municipalities as a result of a Delaunay triangulation

Moreover, both spatial matrices discussed here show similar characteristics (except symmetry). The first one has exactly five neighbours allocated to each municipality in the data, while the second has almost exactly six neighbours on average for each municipality (ranging from 1 to 11 neighbours). Both matrices are typically sparse matrices, meaning that very few elements – here around 1.1% - are non-zero. It can easily be seen from Figure III-10 that the matrix according to the Delaunay-triangulated Swiss map is sparsely filled and symmetric. The black dots in the picture below mark non-zero elements and hence indicate whether spatial unit i is a neighbour to unit j.

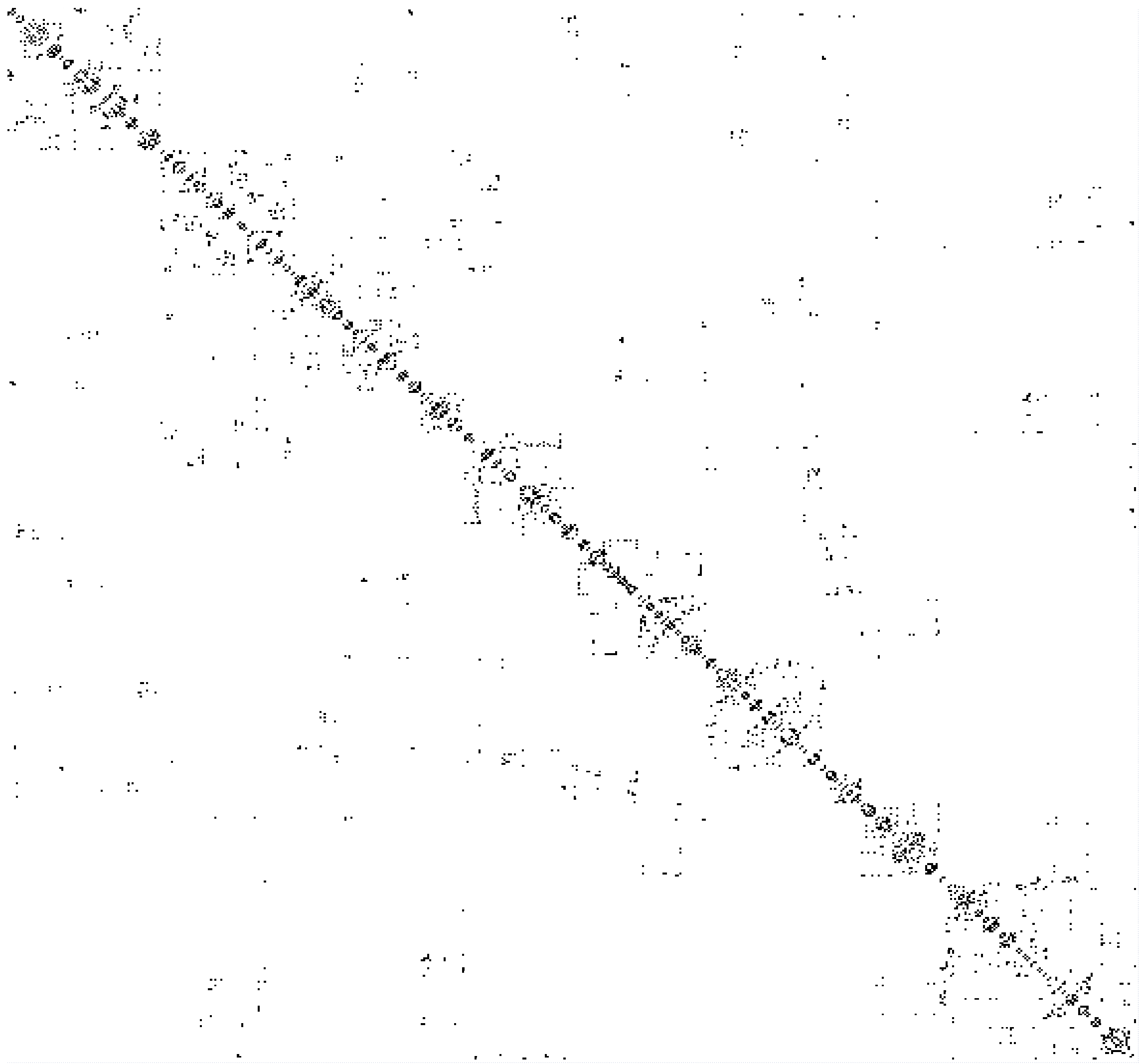


Figure III-10: Structure of the spatial weighting matrix W (547 rows and columns) according to Figure III-9.

For the matrix with five neighbours, we will use an exponentially decreasing scheme such that the elements of the matrix can be defined as $w_{ij} = e^{-\omega \cdot d_{ij}}$ where ω is a strictly positive and predefined parameter. For the second matrix the spatial weights are decreasing more moderately such as $w_{ij} = d_{ij}^{-\omega}$. In empirical studies, the parameter ω is often set equal to one.

4.3 Data

Comparing to the first part of the dissertation, the data were enriched such that the total area of Switzerland is represented and not only the 12 border cantons where gasoline tourism in Switzerland's border regions was analysed.

The sample now consists of observations of 547 out of a total of 2721 municipalities (formerly 315) from 2001 to 2008. Data on gasoline sales were collected from the five largest gasoline companies operating in Switzerland. The sale volumes of their stations were aggregated at the municipal level and averaged with the number of gasoline stations available. Gasoline sales per station in a municipality shall be described formally with the Swiss gasoline price, the foreign gasoline price, the population, the number of gasoline and diesel powered cars, the commuters, the per capita income, the distance from the border and a dummy variable of a gasoline company with extraordinarily high sales volumes.

Averaged at the municipal level

Variable	Measure	Minimum	Maximum	Median	SD (overall)	SD (within)	SD (between)	Ratio within/between
Aggreg. Gasoline Sales	1'000 l / year	142	122'600	2'309	6'763	1'173	6'666	17.6%
Number of Stations (with av. Data)	Stations	1	60	2	3.19	0.427	3.156	13.5%
Total Number of Stations	Stations	1	75	3	5.2	0.119	5.19	2.3%
Swiss Gasoline Price	CHF / l	1.29	1.8	1.44	0.165	0.161	0.036	447.2%
Foreign Gasoline Price	CHF / l	1.33	2.25	1.74	0.264	0.246	0.095	258.9%
Swiss per Capita Income (taxable)	CHF / year	7'818	108'089	29'388	9'627	2'634	9'267	28.4%
Commuters	Persons / year	0	30'756	7	1'582	200	1'571	12.7%
Swiss Population	Persons	172	358'540	4'340	20'674	374	20'687	1.8%
Stock of Cars (Gasoline)	Cars	81	138'331	2'097	7'882	602	7'865	7.7%
Stock of Cars (Diesel)	Cars	2	21'391	211	937	387	854	45.3%
Distance from Border	km	0.15	75.5	23.2	18.54	0	18.57	0.0%
Dummy Company "B"	---	0	1	0.187	0.39	0.104	0.376	27.7%

Table III-1: Descriptive statistics of variables used in the model

Table III-1 provides descriptive statistics of the variables used in the model. It can be seen that the majority of the variables except the prices exhibit a relatively higher between variation than that within. As discussed in the first paper, the low within variation of most of the variables can be an argument in favor of the use of a random effects model. The prices show a very small between variation, since they only vary at the border region level. For the municipalities in Central Switzerland, the price of the border region with the smallest distance to the municipality in question has been allocated. Since a municipality's distance from the border is a time-invariant variable, it cannot be identified by a fixed effects model. There are only four different prices for Switzerland, since the data were collected from the Swiss customs authorities, who track prices monthly at gasoline stations at the border (unfortunately, gasoline prices were not available from the gasoline companies). The descriptive statistics of the variables remained almost identical compared to the description for the first part of the dissertation, where only the border cantons were considered. This is an indicator that both sub-samples are representative for the whole of Switzerland. Naturally, some statistics differ more strongly than others; for instance, regions in Central Switzerland have lower per capita income on average than regions close to the border. They are less densely populated and, accordingly, there are also fewer cars and, of course, almost no commuters.

4.4 Estimation Results

The estimation strategy is the following. First, the model described by equation (III.43) is estimated, excluding the spatial effects both in the dependent variable and in the residuals applying pooled OLS. Then, an F-test and a Breusch-Pagan test is applied to test for individual (fixed and random) effects which may account for potential heterogeneity among the spatial units. The Hausman test statistic then is applied to check for systematic differences between the fixed and the random effects specification[24]. The goal of this first part of the econometric analysis is to identify the model which should be used in the spatial econometric analysis. Therefore, we will not discuss the results in terms of coefficients.

In a second step, two spatial regimes are considered. The first one is described by a spatial weighting matrix with the five closest neighbours to each municipality as it was depicted in Figure III-8. For this setting, the first spatial weighting matrix is maximum-row normalised and weights are defined to be exponentially decaying $w_{ij} = e^{-1 \cdot d_{ij}}$, which is common in spatial econometric studies. A second matrix will be maximum-row normalised too and defined according to the Delaunay triangulation as depicted in Figure III-9. The weights will be chosen in a less strongly decaying form as a function of inversed distances ($w_{ij} = d_{ij}^{-1}$). Subsequently, spatial interdependence assumptions will be tested using the Lagrange multiplier tests for spatial lag dependence and their robust versions, the Lagrange multiplier tests for spatial auto-correlation and the modified Moran-I test statistic which tests for spatial auto-correlation in the residuals. Depending on the outcome, a non-spatial version of the model stated in equation(III.43) ($\lambda = 0, \rho = 0$), a spatial lag version ($\lambda \neq 0, \rho = 0$), a spatial error version ($\lambda = 0, \rho \neq 0$) and a spatial lag spatial error version (Kelejian-Prucha Model) ($\lambda \neq 0, \rho \neq 0$) will be estimated.

---

24 As previously mentioned, the Hausman test statistic assumes consistency of the fixed effects specification, which is not necessarily true if the within variation of certain variables used in the estimation is low. Further, we did not implement the Hausman statistic for spatial models developed by Pfaffermayr (2011).

4.4.1 Estimation of the Non-Spatial Model

The estimation results of the non-spatial form of the model described by equation (III.43) using pooled OLS, FE and RE models are listed in Table III-2. As already mentioned, the variables are included in the model in a logarithmized form, which enables interpretation of the coefficients directly as elasticities.

Coefficient	Variable	Pooled OLS	FE–Model	RE–Model
R^2_{within}		-----	0.167	0.165
$R^2_{between}$		-----	0.075	0.194
$R^2_{overall}$		0.228	0.082	0.191
α_0	Constant	9.90 (11.8)***	15.77 (9.17)***	15.02 (13.37)***
α_1	$\ln(PG_{CH,bt})$	-0.169 (-0.62)	-0.703 (-4.87)*	-0.665 (-4.66)**
α_2	$\ln\left(PG_{F,bt}\right)$	0.004 (0.02)	0.458 (3.70)***	0.382 (3.15)***
α_3	$\ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right)$	0.273 (4.01)***	0.121 (-3.28)***	0.126 (3.46)***
α_4	$\ln\left(\frac{CARD_{CH,it}}{POP_{CH,it}}\right)$	-0.171 (-5.42)***	-0.136 (-7.78)***	-0.144 (-8.32)**
α_5	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.449 (10.1)***	-0.095 (-1.54)	0.037 (0.68)
α_6	$\ln(Commu_{it})$	0.076 (11.9)***	0.004 (0.37)	0.0254 (2.79)***
α_7	$\ln(dist_i)$	-0.067 (-4.83)***	-----	-0.160 (-5.59)***
α_8	DB_{it}	0.534 (20.4)***	0.667 (21.4)***	0.652 (24.00)***
F test for FE			$F(546, 3822) = 70.8$	
B&P test for RE				$\chi^2(1) = 7'810$***
Hausman test				$\chi^2(7) = 42.4$***

Table III-2: Pooled OLS, FE- and RE-estimations results of the non-spatial model according to equation (III.43)

- Number of observations 4376 (T=8 years, N=547 municipalities)
- t-statistics are in parentheses: ***, ** and * indicate 1%, 5% and 10% significance levels respectively

The coefficients of the OLS model are relatively different from the ones obtained from the FE and RE models, whereas the coefficients of the RE and FE models are very similar. For the choice between the FE and RE model, we should consider the result of the Hausman test, the level of within variation of the variable, the efficiency of the estimates and the difference of the coefficients. We choose the RE specification, since the explanatory power of the model is higher, the coefficient of the number of commuters is positive and significant, and some of the other variables, except the prices, have a lower standard error. Moreover, the coefficients of the prices are closer to the average elasticities we estimated in the first paper, in which we reported an average elasticity of -0.65 of the Swiss gasoline price and one of the foreign gasoline price of 0.43. Since we considered the whole Swiss territory in the present sample, we think that a lower estimate of the elasticity of the foreign gasoline price makes more sense.

4.4.2 Estimation of the Spatial Model(s)

As previously mentioned, two spatial weighting regimes are considered. Model 1 is a spatial random effects specification with the spatial weighting matrix with exponentially decaying matrix entries and the five closest neighbours to each municipality considered. Model 2 only differs in terms of the weighting matrix, namely based on the Delaunay triangulation of the municipalities available in the data with inverse distance entries. Prior to turning towards the estimation results, spatial dependence assumptions, as described in section 3.5, are tested. The results of these tests are reported in Table III-3. We choose to use three tests to check the validity of the introduction of a spatially lagged dependent variable in the model (1-3). The first test (1) tests jointly for random effects and a spatially lagged dependent variable, the second for the presence of a spatially lagged dependent variable conditional on random effects, and the third for random effects conditional on the presence of a spatially lagged dependent variable. Further, we use five tests (4-7) to check the model's validity for the presence of spatially lagged residuals: The first tests jointly for the presence of random effects and spatially lagged residuals. The second tests for spatially lagged residuals conditional on random effects, and the third for spatially lagged residuals conditional on random effects and a spatially lagged dependent variable. The fourth tests for the presence of random effects conditional on spatially lagged residuals, and the last is the Moran I test for pooled OLS.

		Test	Null Hypothesis	Test Results Model 1	Test Results Model 2
Spatial Lag	1	Joint test: $LM_{\lambda,\sigma_\mu^2}$	$\lambda = 0, \sigma_\mu^2 = 0$	$\chi^2(2) = 12'280$***	$\chi^2(2) = 12'281$***
	2	Cond. Test: $LM_{\lambda\|\sigma_\mu^2\geq 0}$	$\lambda = 0 \| \sigma_\mu^2 \geq 0$	$N(0,1) = 26.3$***	$N(0,1) = 42.9$***
	3	Cond. Test: $LM_{\sigma_\mu^2\|\lambda\geq 0}$	$\sigma_\mu^2 = 0 \| \lambda \geq 0$	$\chi^2(1) = 12'156$***	$\chi^2(1) = 12'107$***
Spatial Error Lag	4	Joint test: LM_{ρ,σ_μ^2}	$\rho = 0, \sigma_\mu^2 = 0$	$\chi^2(.) = 12'194$***	$\chi^2(.) = 12'176$***
	5a	Cond. Test: $LM_{\rho\|\sigma_\mu^2\geq 0}$	$\rho = 0 \| \sigma_\mu^2 \geq 0$	$N(0,1) = 3.63$***	$N(0,1) = 5.76$***
	5b	Modified Cond. Test: $LM_{\rho\|\lambda,\sigma_\mu^2\geq 0}$	$\rho = 0 \| \lambda, \sigma_\mu^2 \geq 0$	$N(0,1) = 2.99$***	$N(0,1) = 4.56$***
	6	Cond. Test: $LM_{\sigma_\mu^2\|\rho\geq 0}$	$\sigma_\mu^2 = 0 \| \rho \geq 0$	$\chi^2(1) = 2670$***	$\chi^2(1) = 1320$***
	7	Moran I (Pooled OLS)	$\rho = 0$	$N(0,1) = 2.43$**	$N(0,1) = 2.87$***

Table III-3: Test Results of spatial dependence hypotheses

- ***, ** and * indicate 1%, 5% and 10% significance levels respectively

The results of these tests show that for both specifications (depending on the spatial weighting matrices), all tests of no spatial correlation, either in the dependent variable or in the residuals, are rejected. Most of all, the conditional tests of spatial spillovers (tests 2 and 5) given the possible presence of random effects (heterogeneity) are rejected. The test statistics are far below the joint test statistics, which underlines the importance of conditional testing discussed by Baltagi, see Baltagi et al. (2003) and Baltagi et al. (2008). Accordingly, it can be concluded that, first, the presence of a spatial lag in the dependent variable is not due to the presence of random effects (and vice versa), since all three test statistics (1-3) significantly reject the null hypothesis. The same holds for tests 4-6 for spatial correlation in the residuals. The test in (5b) represents a modified version of the test presented by Baltagi et al. (2003) and described in subsection 3.5.2. Instead of taking the residuals of a random effects model, we alter the method such that the residuals are taken from a respective spatial lag model. The reason is that a Monte Carlo experiment – which is presented in the Appendix (6.2.1.1) – shows that the test statistic testing for the spatial lag is not sensitive to the possible presence of spatial autocorrelation, but the test for spatial autocorrelation seems to depend heavily on the presence of a spatial lag. Therefore, we believe that this modification gives more convincing results and the coverage (type I error) and power (type II error) behave more meaningfully for the modified version, as can be seen in the Appendix. Of course, this modified version is only an ad-hoc version of a robust test for spatially lagged residuals conditional on a spatially lagged dependent variable. A closer investigation of this issue would be the subject of future work.

Moreover, the Moran I statistic as well is significant in both model specifications, namely at the 5% level for the model with a spatial weighting matrix with five neighbours and exponentially decaying entries and at the 1% level for the case with the Delaunay inverse distance matrix. For completeness, the test statistics suggested by Sen and Bera (2011) were calculated as well, and the results point in the same direction as the other test statistics. In addition, the tests by Sen and Bera would suggest accounting for serial correlation in the residuals.

Thus, the specification as is written in equation (III.43) seems to be justified both with respect to the presence of potential spatial spillovers in the dependent variables, in the residuals and with respect to random effects.

The estimations are carried out as follows. From the general model, the two spatial weighting schemes are implemented and estimated in three different versions, one being a specification including only a spatial lag in the dependent variable $(\lambda \neq 0, \rho = 0)$, one only including spatial autocorrelation (spatial correlation in the residuals) $(\lambda = 0, \rho \neq 0)$ and one which allows for both spatial correlation in the dependent variable and in the residuals $(\lambda \neq 0, \rho \neq 0)$ as depicted in Figure III-11.

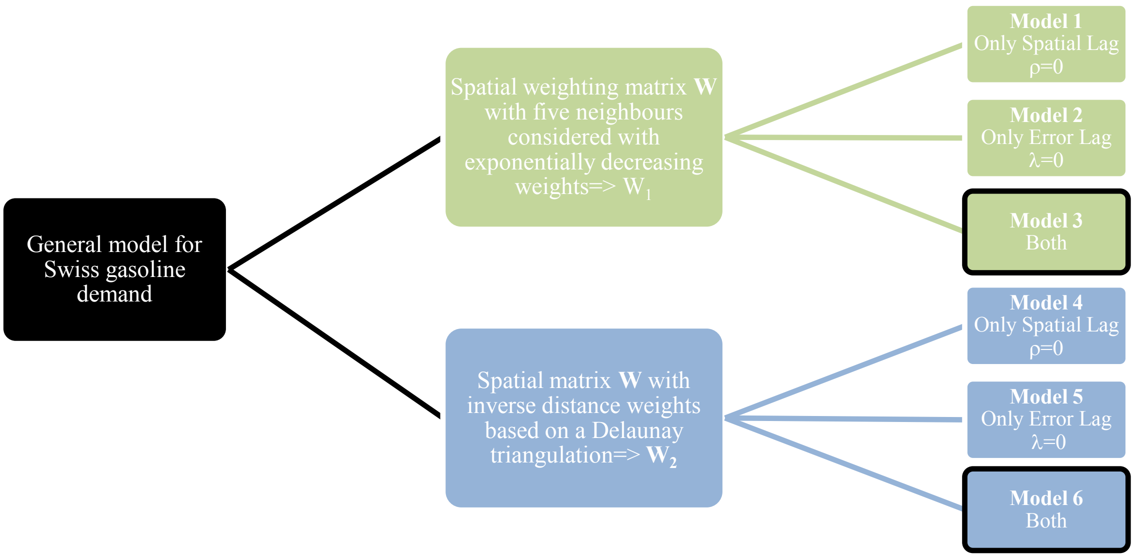


Figure III-11: Spatial estimation strategy

Coeff.	Variable	Reference (non-spatial)	Model 1 (Spatial Lag)	Model 2 (Spatial Error)	Model 3 (both)
R^2_{within}		0.165	0.17	0.16	**0.16**
$R^2_{between}$		0.194	0.22	0.51	**0.59**
$R^2_{overall}$		0.191	0.21	0.49	**0.57**
α_0	Constant	15.02 (13.37)***	9.10 (2.76)***	13.3 (17.8)***	**11.38 (12.6)*** **
α_1	$\ln(PG_{CH,bt})$	-0.665 (-4.66)**	-0.578 (-4.03)***	-0.643 (-4.29)***	**-0.609 (-4.08)*** **
α_2	$\ln\left(PG_{F,bt}\right)$	0.382 (3.15)***	0.329 (2.71)***	0.361 (2.86)***	**0.336 (2.66)*** **
α_3	$\ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right)$	0.126 (3.46)***	0.105 (2.85)***	0.106 (2.91)***	**0.095 (2.58)** **
α_4	$\ln\left(\frac{CARD_{CH,it}}{POP_{CH,it}}\right)$	-0.144 (-8.32)**	-0.131 (-7.55)***	-0.137 (-7.74)***	**-0.131 (-7.36)*** **
α_5	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.037 (0.68)	0.010 (0.18)	0.021 (0.38)	**0.010 (0.18)**
α_6	$\ln(Commu_{it})$	0.0254 (2.79)***	0.019 (2.12)**	0.020 (2.18)**	**0.019 (2.11)** **
α_7	$\ln(dist_i)$	-0.160 (-5.6)***	-0.149 (-5.19)***	-0.158 (-5.44)***	**-0.155 (-5.33)*** **
α_8	DB_{it}	0.652 (24.00)***	0.662 (23.05)***	0.656 (22.98)***	**0.658 (23.04)*** **
λ	Spatial Lag	-----	0.791 (3.91)***	-----	**0.429 (3.69)*** **
ρ	Error Lag	-----	-----	0.463 (5.3) p-value: 0.120	**0.443 (4.71) p-value: 0.131**

Table III-4: Spatial FGLS estimation results of Swiss gasoline demand according to spatial weighting matrix $\mathbf{W_1}$

- Number of observations: 4376 (T=8 years, N=547 municipalities)
- t-statistics are in parentheses: ***, * * and * indicate 1%, 5% and 10% significance levels, respectively
- Standard error for ρ according to non-linear least squares.

Coeff.	Variable	Reference (non-spatial)	Model 4 (Spatial Lag)	Model 5 (Spatial Error)	Model 6 (both)
R^2_{within}		0.165	0.18	0.15	**0.16**
$R^2_{between}$		0.194	0.23	0.62	**0.63**
$R^2_{overall}$		0.191	0.23	0.60	**0.62**
α_0	Constant	15.02 (13.37)***	7.95 (2.90)***	13.42 (18.06)***	**11.8 (13.34)*** **
α_1	$\ln(PG_{CH,bt})$	-0.665 (-4.66)**	-0.490 (-3.39)***	-0.642 (-3.94)***	**-0.580 (-3.56)*** **
α_2	$\ln\left(PG_{F,bt}\right)$	0.382 (3.15)***	0.275 (2.26)**	0.359 (2.63)***	**0.321 (2.36)** **
α_3	$\ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right)$	0.126 (3.46)***	0.079 (2.13)**	0.096 (2.60)***	**0.080 (2.13)** **
α_4	$\ln\left(\frac{CARD_{CH,it}}{POP_{CH,it}}\right)$	-0.144 (-8.32)**	-0.112 (-6.25)***	-0.132 (-6.97)***	**-0.120 (-6.26)*** **
α_5	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.037 (0.68)	-0.009 (-0.17)	-0.005 (-0.08)	**-0.011 (-0.20)**
α_6	$\ln(Commu_{it})$	0.0254 (2.79)***	0.019 (2.11)**	0.021 (2.24)**	**0.022 (2.33)** **
α_7	$\ln(dist_i)$	-0.160 (-5.6)***	-0.133 (-4.58)***	-0.166 (-5.47)***	**-0.159 (-5.25)** **
α_8	DB_{it}	0.652 (24.00)***	0.656 (23.0)***	0.650 (22.86)***	**0.655 (22.9)*** **
λ	Spatial Lag	-----	0.710 (5.02)***	-----	**0.339 (3.13)*** **
ρ	Error Lag	-----	-----	0.394 (32.34)** p-value: 0.021	**0.365 (24.8)**** **p-value: 0.026**

Table III-5: Spatial FGLS estimation results of Swiss gasoline demand according to spatial weighting matrix $\mathbf{W_2}$

- Number of observations: 4376 (T=8 years, N=547 municipalities)
- t-statistics are in parentheses: ***, * * and * indicate 1%, 5% and 10% significance levels, respectively
- Standard error for ρ according to non-linear least squares. A bootstrap exercise for model 6 was carried out. In this, the standard deviation of ρ in Model 6 after 250 replications is 0.107. Further, the mean of the 250 replications is 0.407. Accordingly, the bootstrap estimate is not significantly different from the point estimate of 0.365. For further details, see Appendix 6.2.2).

The estimation results of the six models are tabulated in Table III-4 and Table III-5. For the purposes of comparison, the results of a general random effects model are tabulated too. All coefficients in all models, except the per capita income, have the expected sign and are significant.

The price elasticity of Swiss gasoline demand with respect to the domestic price is stable among all six models and ranges from -0.49 to -0.643. In general, all six values are similar to that obtained in the reference model, but are slightly smaller in absolute value. This holds for almost all coefficients, since the omission of a (positive) spatial lag in the dependent variable will generally bias the coefficients away from zero. Similarly, accounting for spatial autocorrelation (spatial correlation in the residuals) tends to lower the standard errors in absolute values, which protects against misleading inference.

There is a strong improvement of the goodness of fit at the cross-sectional level when spatial dependence is considered in the model (compare models 2 and 3 to model 1 and the reference model in Table III-4). Given the test statistics reported in Table III-3 , the following discussion is based on the coefficients obtained using model 6. The further reasons are that, first, the test results for this model point more strongly towards a SARAR model than in model 3. Second, we believe that the spatial regime constructed with the Voronoi diagram, as depicted in Figure III-4, is more appropriate when analysing gasoline demand at the municipal level. Nonetheless, it has to be noted that the results from model 3 and 6 are very similar and therefore policy implications would be the same.

As in the other models, we observe a positive coefficient of the spatially lagged dependent variable. Second, the coefficient of the spatially lagged residuals is positive and significantly different from zero. The elasticity of Swiss gasoline price with respect to the domestic price is -0.58 and is significantly smaller compared to that tabulated in the reference model[25]. As mentioned, this is a very important characteristic of spatial models: the parameter estimates are biased away from zero when omitting a positive spatial lag in the dependent variable. A 10% increase in the Swiss gasoline price in a municipality is expected, therefore, ceteris paribus, to decrease gasoline consumption by -5.8 % in that municipality. A 10% increase in the foreign gasoline price is expected, ceteris paribus, to increase domestic consumption by 3.2%. The coefficients of the gasoline and diesel car fleet (per capita) are also lower in absolute values compared to the reference model. A ceteris paribus 10% increase in the gasoline car fleet per capita therefore is expected to increase domestic gasoline consumption by 0.8%, whereas a ceteris paribus 10% increase in the diesel car fleet per capita would decrease gasoline consumption by 1.2%. The income per capita is also insignificant in this model. The coefficient of the daily commuters is significantly different from zero: A 10% increase in the number of daily commuters, therefore, would increase gasoline consumption by 0.2

[25] Recall that the direct interpretation of the estimated coefficient may be misleading in spatial models, and we refer to the calculation of the direct, indirect and total impacts of the variables involved. Nonetheless, a comparison among the models has to be made to identify potential biases when omitting spatial correlation.

Table III-6 gives the averages of total, direct and indirect impacts of these variables and are calculated as described in subsection 3.2.3. Recall that the indirect impacts would be zero if a spatial lag was absent and that the total impact would be equal to the direct impact and equal to the coefficients of the estimated (non-spatial) model. As previously explained in subsections 3.2.1 and 3.2.2, the average total impact is the change which a spatial unit experiences in its dependent variable following a change in one of its independent variables after the shock has gone through an infinite number of feedback loops. Accordingly, the average indirect impact is the change which other spatial units experience in their independent variable after a shock has occurred in any particular spatial unit. The total impact then is the sum of the direct and the indirect impact.

Coeff.	**Variable**	**Reference Model (non-spatial)**	**Av. Total Impact**	**Av. Direct Impact**	**Av. Indirect Impact**
α_1	$\ln(PG_{CH,bt})$	-0.665	-0.6550	-0.5817	-0.0734
α_2	$\ln\left(PG_{F,bt}\right)$	0.382	0.3630	0.3223	0.0407
α_3	$\ln\left(\frac{CARG_{CH,it}}{POP_{CH,it}}\right)$	0.126	0.0905	0.0803	0.0102
α_4	$\ln\left(\frac{CARD_{CH,it}}{POP_{CH,it}}\right)$	-0.144	-0.1357	-0.1205	-0.0152
α_5	$\ln\left(\frac{Y_{CH,it}}{POP_{CH,it}}\right)$	0.037	-0.0128	-0.0114	-0.0014
α_6	$\ln(Commu_{it})$	0.0254	0.0244	0.0217	0.0027
α_7	$\ln(dist_i)$	-0.160	-0.1792	-0.1591	-0.0201
α_8	DB_{it}	0.652	0.7398	0.6569	0.0829
λ	Spatial Lag	---	0.3822	0.3393	0.0429

Table III-6: Calculation of total, direct and indirect effects according to model 6 (cumulative effects)

Finally, we want to visualise the total effect of the Swiss gasoline price in a map. This is to show the spatial differentiation of the Swiss gasoline price elasticity. Figure III-12 shows the total effect of the Swiss gasoline price on Swiss gasoline demand at the municipal level. For municipalities which are not covered by our data, we assigned the value of the corresponding Thiessen polygon it belongs to (compare with Figure III-4) in order to assign elasticities to the whole Swiss territory. The values of the elasticities and the corresponding color are shown in the legend.

Figure III-12: Elasticity of Swiss gasoline demand with respect to the Swiss gasoline price (total effect)

The lowest quintile is represented by municipalities with an elasticity between -0.61 and -0.585. The highest quintile is represented by municipalities with elasticities between 0.662 and 0.855. It should be recalled that we reported an average direct effect of the Swiss gasoline price on Swiss gasoline demand of -0.655, but there are municipalities which have far higher elasticites assigned, up to -0.855. The map shows that urban areas are assigned the highest elasticities (regions 1, 6 and 9). The southern part of Switzerland is also assigned high elasticities. Moreover, compared to the first part of the dissertation, in which the elasticities were calculated with the distance from the border, the picture is not so different. For some areas, we also have high price elasticities close to the border, given that the regions in question are urban (regions 2,7, 9). The more remote areas are assigned the lowest quintile of the elasticities (the Alpine regions 3 and 5). One possible explanation for this might be that public transport is much more readily available in the urban regions than in the remote areas. Due to the lack of substitution possibilities, then, elasticity in the rural regions is relatively smaller.

5 Conclusion

As already seen in the literature review in the first paper and in this one, many studies analyse gasoline demand using aggregate panel data, but almost none of them applies spatial econometric methods. As observed by Pirotte et al. (2011), the main difficulty when explaining gasoline demand (or road traffic) in small regions using a spatial econometric approach is that the data are only available at the panel unit level (here the municipalities). On the other hand, it is clear that road traffic and hence gasoline demand is not only dependent on a municipality's car fleet or population, but also on exchange traffic. Accordingly, the smaller the spatial units are, the stronger is the potential for spatial interaction.

Therefore, the analysis of gasoline demand might necessitate the incorporation of a spatially lagged dependent variable. The advantages of doing so are clear-cut. First, the disregard of a spatially lagged dependent variable on the right-hand side of a gasoline demand model would result in biased coefficient estimates if the true data generating process were actually described by such a spatial lag. Second, traditional panel data approaches estimate the same coefficients for all panel units, a hypothesis which can be tested and which is seldom valid. The incorporation of a spatial lag, however, allows the coefficient estimates to vary among the panel units in a predefined way through the definition of the spatial weighting matrix $\mathbf{W}$.

From a methodical point of view, one goal of this paper was to implement the GMM estimator developed by Kelejian and Prucha (2007) into STATA® and to describe the procedure properly. One advantage when using GMM rather than a maximum likelihood approach is that sample size is less of a problem. Second, there is no need for a distributional assumption about the error terms. Third, the estimation of additional endogenous variables besides the spatially lagged dependent variable is easier when using the GMM approach. We further described how to define the spatial weighting matrix and implemented several routines in STATA®. Among others, we implemented a routine for the Voronoi tessellation of spatial units, and we have seen that the maximum row-sum normalised weighting matrix with inverse distance entries describes the spatial dependence in Swiss gasoline demand in the best way.

Naturally, the assumptions about the presence of spatially lagged residuals or a spatially lagged dependent variable should be tested, an issue which has rather been neglected in the spatial econometric studies discussed. For this reason, the test statistics developed by Baltagi et al.

(2003) and (2008) have been implemented and used for the empirical application. We used different weighting matrices - one in which the five closest municipalities are considered as neighbours and one with the Voronoi tessellation we previously described. Both weighting schemes indicate the presence of both a spatially lagged dependent variable and spatially lagged residuals, but the effects seem to be more distinct for the second weighting matrix.

We estimated a spatial lag in Swiss gasoline demand of 0.34 and a spatial error lag of 0.37. Both values have proper signs and are similar in terms of magnitude compared to similar studies, see e.g. Pirotte et al. (2011) or Pennerstorfer (2008). This implies that an increase in gasoline consumption by 10% in a municipality (e.g. due to increased number of commuters) spreads to other municipalities and leads to an increase in gasoline consumption of 3.4% (given that the regions are neighbours).

A main result is that the elasticity of gasoline demand with respect to the domestic gasoline price is not significantly different to the average elasticity reported in the first part of this dissertation. If the average total impact of the domestic price on gasoline sales is considered, a 10% increase in the domestic gasoline price should decrease gasoline demand by 6.55%. However, spatial partitioning of this value reveals that the average direct impact is a decrease of 5.8%. The difference is due to spatial spillovers and therefore, the elasticity of gasoline sales with respect to the gasoline price in that municipality is only -0.58, and accordingly somewhat lower than presented in the first part of the dissertation. The reasoning holds too for partial effects of other variables included in the model and leads to price elasticities ranging from -0.585 to -0.855, depending on the geographical location of the municipality. This is probably the core aspect of spatial econometric models. Through the incorporation of the spatially lagged dependent variable, all other parameters vary among the spatial units in a predetermined way. This variation, as we have seen, can be substantial and significantly different from parameter estimates where spatial spillovers were not considered. The price elasticity of the non-spatial version of the present model is -0.665. Given the range into which the price elasticities of the spatial model fall, we see that the deviation ranges from -11% up to 30% compared to the non-spatial model. Probably due to the very different approach and different data, we find a price elasticity of Swiss gasoline demand with respect to the foreign gasoline price of 0.32, which is very different from that reported by Baranzini (2012), but not radically different from that reported in the first part of the dissertation. Considering the total effect of the foreign price on Swiss gasoline demand, the elasticity ranges from 0.32 to 0.472.

From a methodological point of view, we conclude that accounting for spatial spillovers when analysing gasoline demand with aggregated panel data can be very important. From a policy-making point of view, the fact that gasoline sales in one municipality affects sales in neighbouring municipalities may have important consequences: there are regions which react more sensitively to a price change (e.g. induced by an introduction of a CO_2 tax on gasoline). From our analysis, we conclude that the urban areas react more strongly to price changes but the rural areas less severely.

6 Appendix

In the present section we provide an overview on the different codes which were implemented into STATA®.

6.1 Spatial Weighting Matrices

6.1.1 The 'voronoi'-Command

The command creates the Delaunay-type spatial weighting matrix. The only thing the program needs as input are three variables: “**idmu**” is a (unique) identifier for the spatial entities, “**xcoor**” refers to a real-number coordinate, and “**ycoor**” defines the second dimension. The outputs of the program are the variables “**triangle1**”, “**triangle2**” and “**triangle3**” where the entity identifier and the coordinates of the Delaunay triangle vertices are stored. Further, “**location1**”, “**location2**” and **location3**” are variables defining the spatial boundaries resulting from the triangulation. Further, a spatial matrix, “**WW**”, is stored which includes distance measures based on the coordinates. The command requires a matrix size of 5000 by default.

The program is executed rapidly when there are few spatial units such as 50 to 100, and it typically uses around 1-5min. For more spatial units, such as 500, the calculation time can increase significantly.

Example

We generate 20 units which are randomly distributed in space and run the triangulation algorithm.

6.1.2 The 'normalize'-Command

Once a spatial weighting matrix "**WW**" is available, it has to be normalised. The command uses the weighting matrix and transforms it to a rowsum-normalised matrix "**WWnorm**" and a maximum-rowsum-normalised matrix "**WWmax**"

If the matrix "**WW**" contains distance entries, the weights must first be altered such that they decrease with increasing distance.

Example

6.1.3 The 'WWcontmake'-Command

Often, data are collected for predefined entities such as administrative border or similar. Depending on the variables to be analysed (e.g. population or income), data could be available for all spatial units. Therefore, the researcher is confronted with the definition of a spatial weighting matrix of the respective area, for instance, a country.

The basic input in the command is a shape file which defines the borders of the entities with an identifier and x and y coordinates. A distinct data source for shape files is e.g. http://www.diva-gis.org/gData. One might use the user written command to convert the shape files.

Example

For illustration, let us consider a small country like Belgium. The shape file is of the following form

For Switzerland, the shape file gives the boundaries of the administrative units as pictured in Figure III-13. Once the shape file is loaded into STATA®, the command produces three matrices: **WWcontig** and **WWcontigdist** and **WWdist**, in which the first matrix contains (0,1) entries and the second one the respective distances. The third matrix contains distances among all units and therefore is not referring to contiguity of the units. Further, the **unique** matrix stores the centre coordinates of the spatial units. Moreover, a graph is given as output for visual control; the result is depicted on the right. The computational time for the command is high. Shape files usually contain some 10^6 data entries, which means the program runs for some 8 hours. Accordingly, future work on could focus on automatically decreasing the size of the files.

Figure III-13: Shapefile for Switzerland (municipal level)

The result of the command is depicted below in Figure III-14.

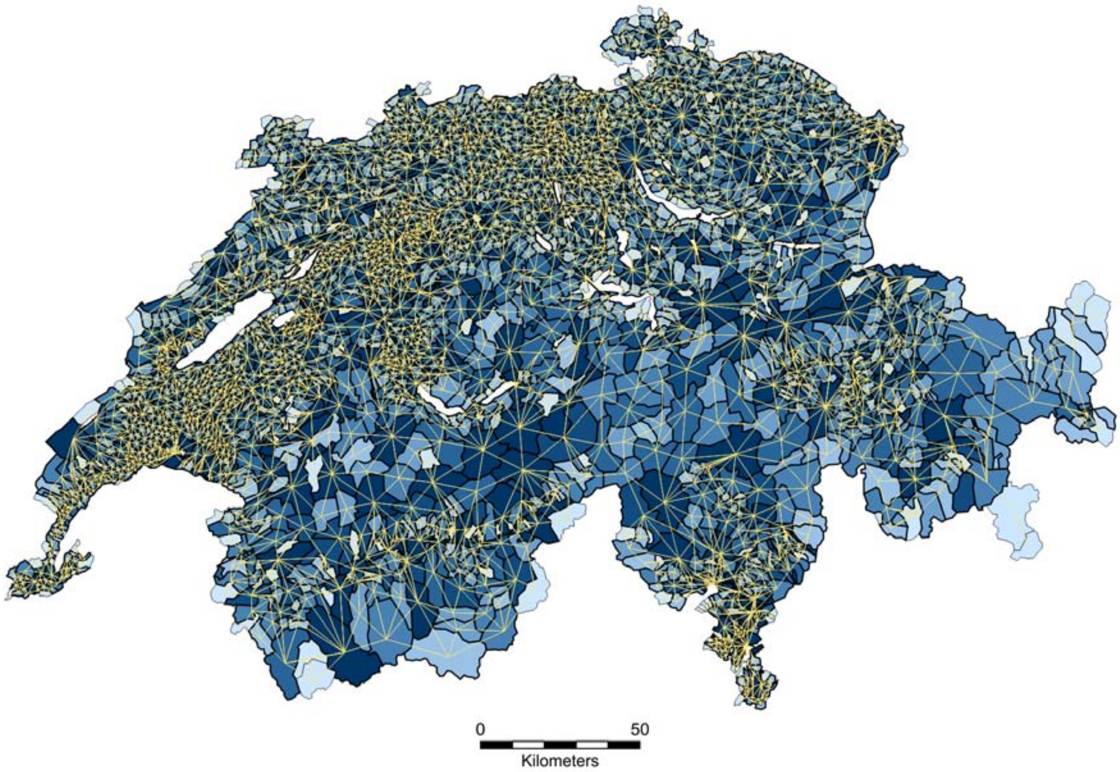


Figure III-14: Visualization of the contiguity matrix for Switzerland at the municipal level

6.2 Estimation

6.2.1 The 'xtspatreg', 'xtspatregrelag' and 'xtspatregreerror' Commands

Those commands were written to implement the Kelejian-Prucha (with both spatial lag and error lag, only spatial lag, only error lag) model in STATA®. As input, it needs a data file such as one would use to run a fixed or random effects model. It is crucial that the panel unit is labeled “idmu” and the time variable “time”. Moreover, a normalised spatial weighting matrix named “**WW**” has to be saved.

The output consists in a fixed and random effects estimation and a corresponding Hausman test. Depending on the model one would like to implement, the syntax may be

For the spatial Lag model

For the spatial Error model

For the Kelejian-Prucha model

The command was extended to the command, by which the matrix of instruments can be influenced and the number of instruments can be chosen. Further comments refer to the command.

Most of the output is stored in matrix and scalar form. In order to assess parameters and estimates , one types "mat dir" or "sca dir"

The procedure was implemented as described in section 3.4. It has to be noted that the estimator stores every data in matrix form. For this reason, matsize in STATA® has to be set to 5000 (or more). The maximum number of observations one can handle, therefore, is currently 11,000. Nonetheless, the estimator was implemented handling the matrices as efficiently as possible in terms of inversion and so on. Therefore, an estimation with 547 spatial units over eight years and some 10 explanatory variables only requires 1-2min of computational time.

Further, we conducted a small experiment to check whether the implemented estimator behaves well.

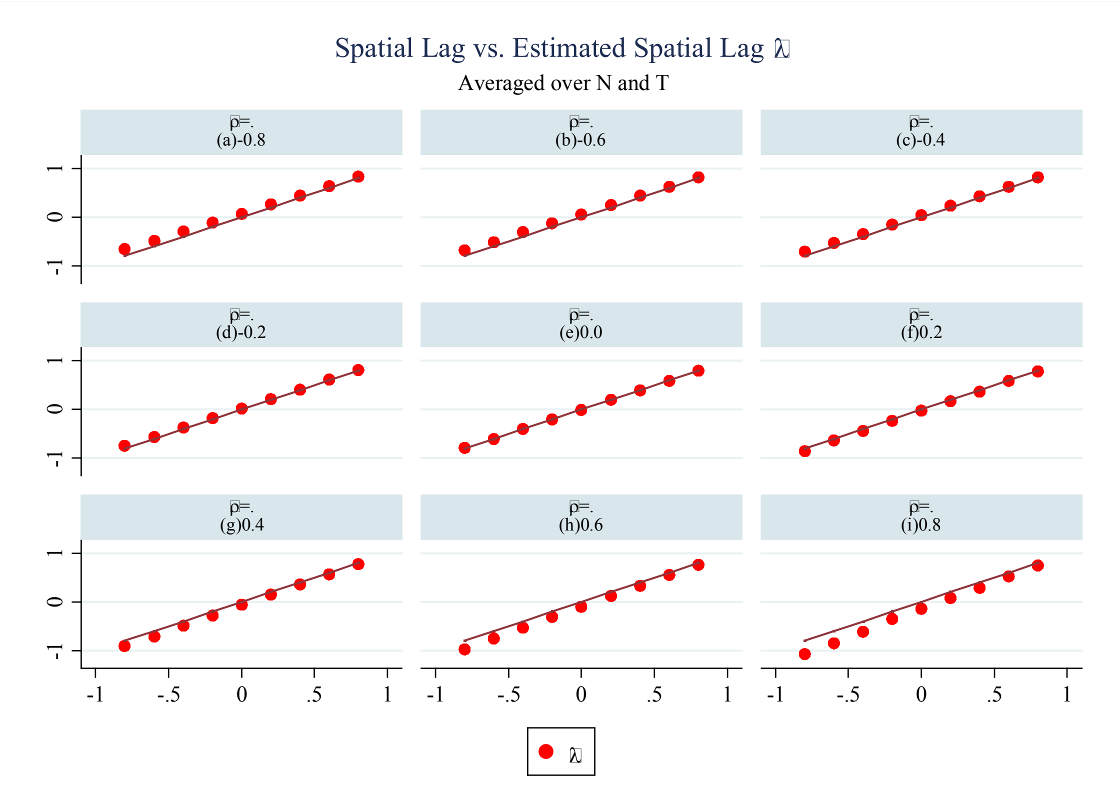
6.2.1.1 Monte Carlo Results

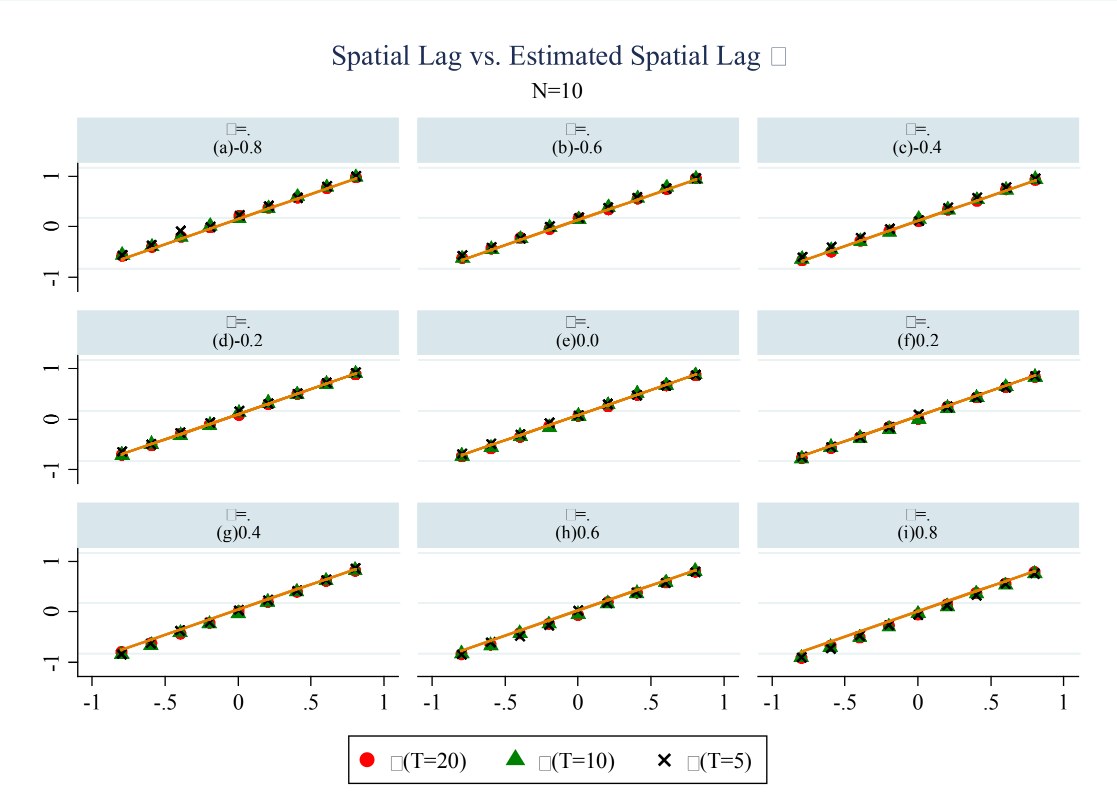
The design of the Monte Carlo study was the following:

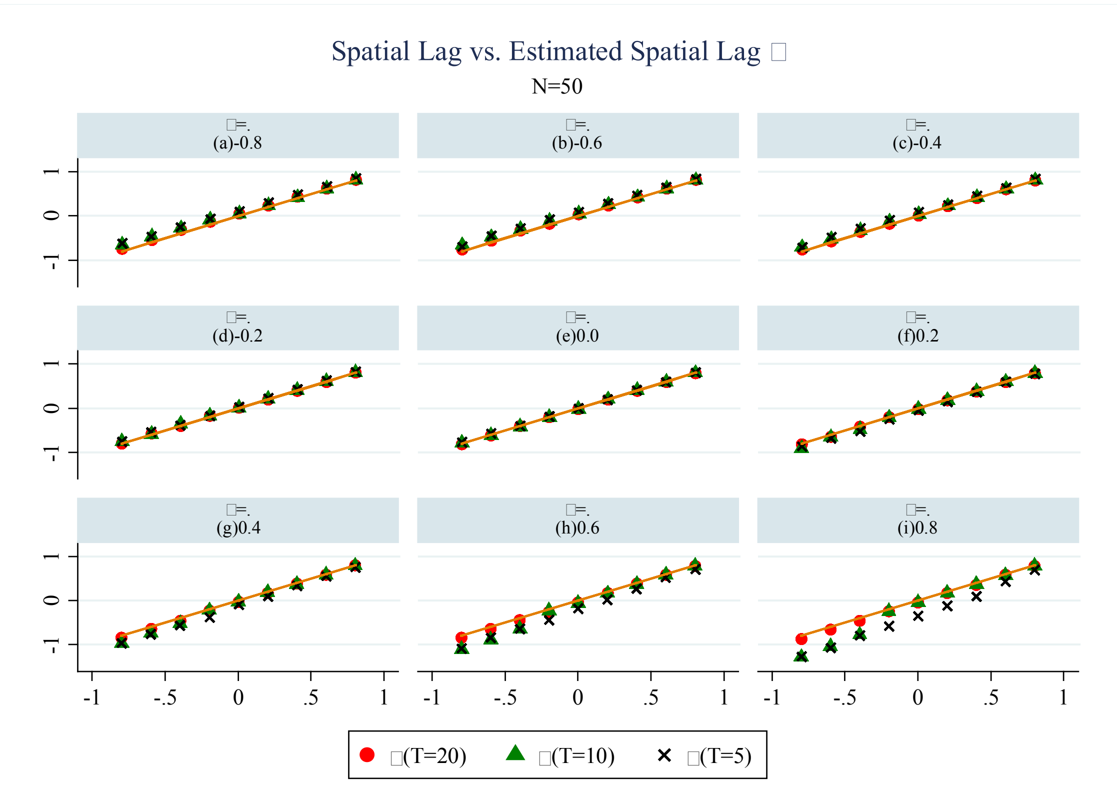
A panel data set is created with N spatial entities and T time periods. Further, three random explanatory variables x_1 (with mean 5, between variation 1, within variation 1), x_2 (with mean 3, between variation 1, within variation 1) and x_3 (with mean -5, between variation 1, within variation 1) were created. The error term was specified according to equation (III.2) with $\sigma_{\mu}^2 = 1$ and $\sigma_{v}^2 = 1$. The coefficients of the explanatory variables were fixed at $b_1 = b_2 = 2,\ b_3 = 1$

The number of spatial units were varied over 10, 50, and 100, and the number of time periods over 5,10, and 20. Both the spatial lag λ and the error lag ρ were varied over -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, and 0.8. For each experiment, 100 replications were made, which results in a total of 729,000 spatial regressions. These required a computational time of some 80 hours.

Figure III-15 depicts the true values of the spatial lag versus the estimated values of the spatial lag. The graphs are grouped by the true value of the spatial error. The upper left graph shows the situation averaged over all observations. For high positive values of the spatial error, the estimated spatial lags are biased slightly downward (e.g. graph (i)). For high negative values of the spatial error, the estimates are biased slightly upward (see e.g. graph (a)). The other three graphs show the situations for the distinct numbers of observations and time periods. The higher the fraction of $\mathrm{N/T}$ and the higher the absolute value of the spatial error is, the more biased is the spatial lag (see e.g. graph (i) for $\mathrm{N{=}100}$ and compare to graph (i) for $\mathrm{N{=}10}$). For positive values of spatial lag estimates, the value of the spatial error shows no influence. Considering the upper left graph again, the bias is only present for $|\rho > 0.4|$ and slightly affects the estimates of (negative) spatial lags.




Spatial Lag vs. Estimated Spatial Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
(T=20)
(T=10)
(T=5)


Spatial Lag vs. Estimated Spatial Lag
N=50
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
(T=20)
(T=10)
(T=5)

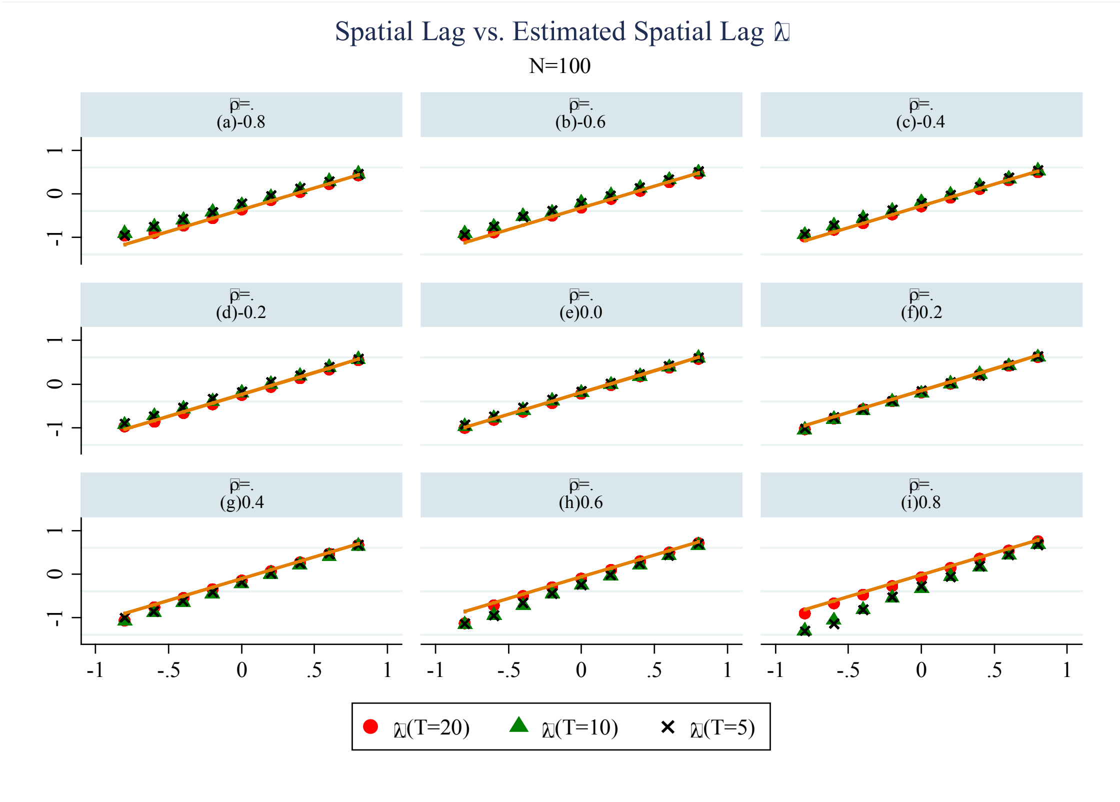
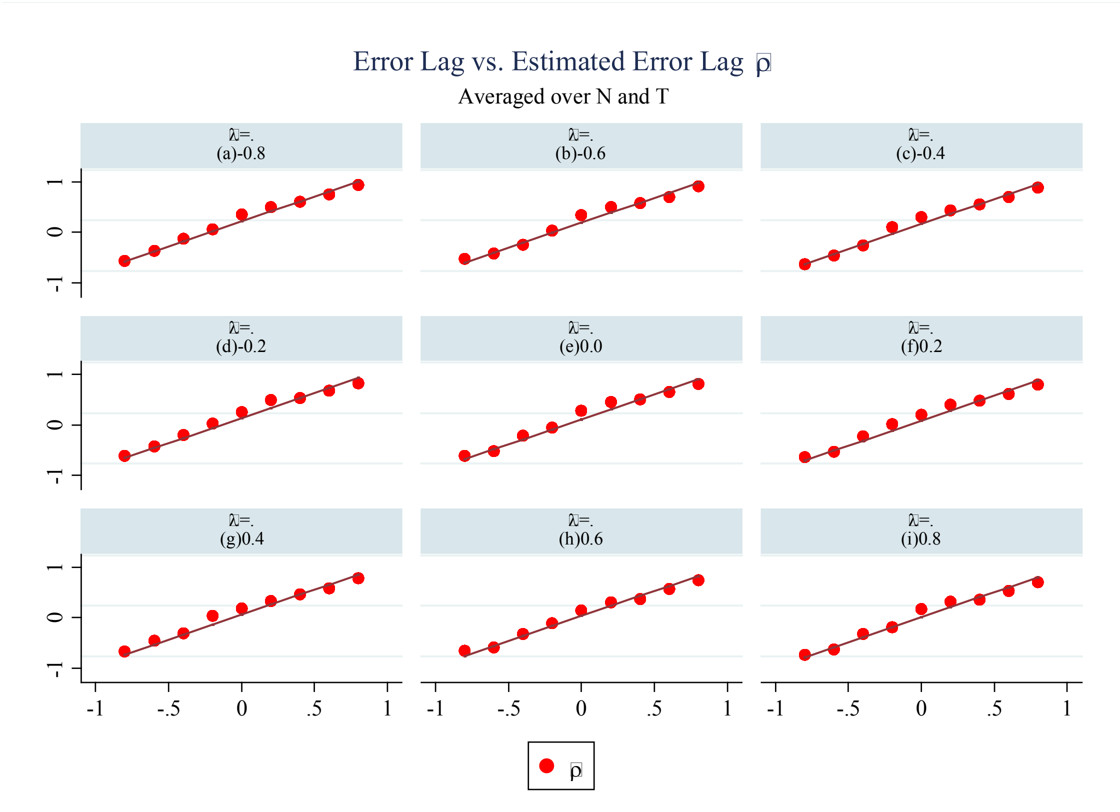
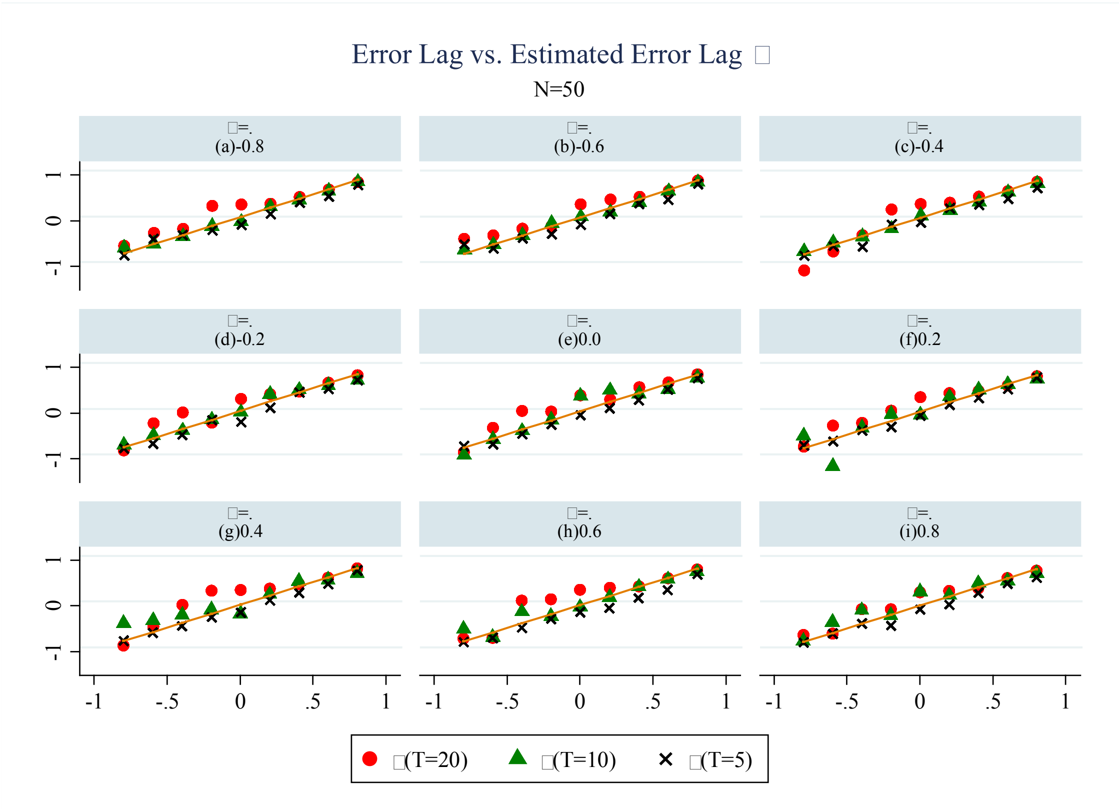


Figure III-15: True values of the spatial lag vs. the estimated spatial lag λ, over ρ




Error Lag vs. Estimated Error Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
(T=20)
(T=10)
(T=5)


Error Lag vs. Estimated Error Lag
N=50
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
(T=20)
(T=10)
(T=5)

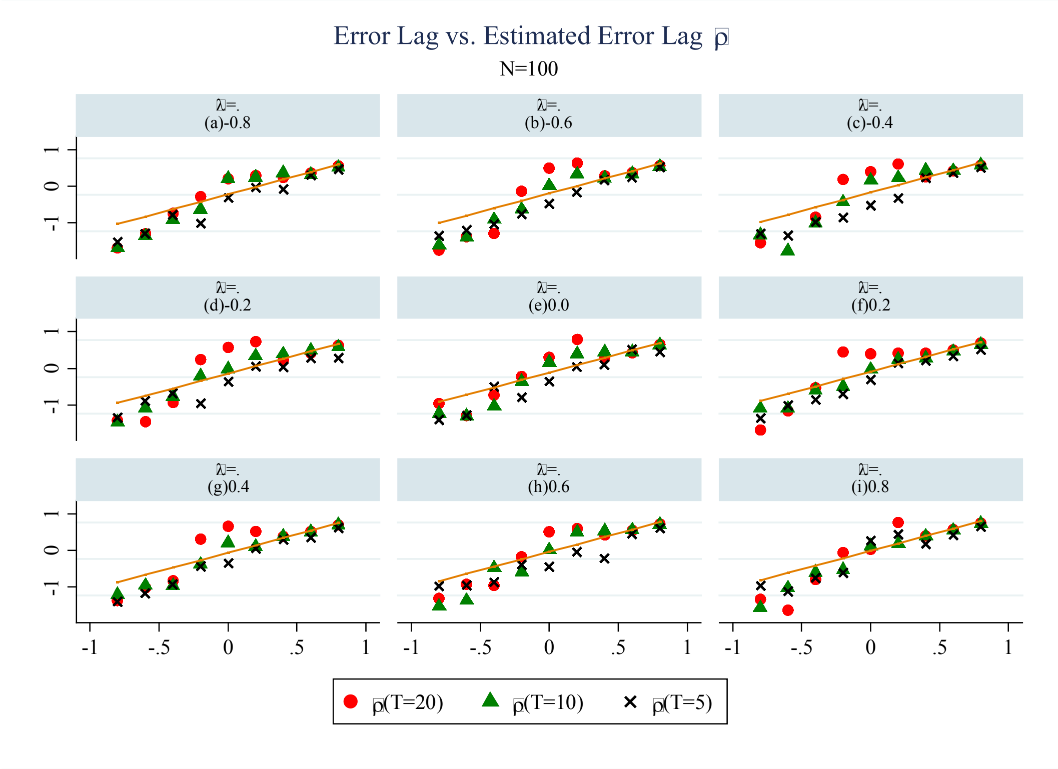
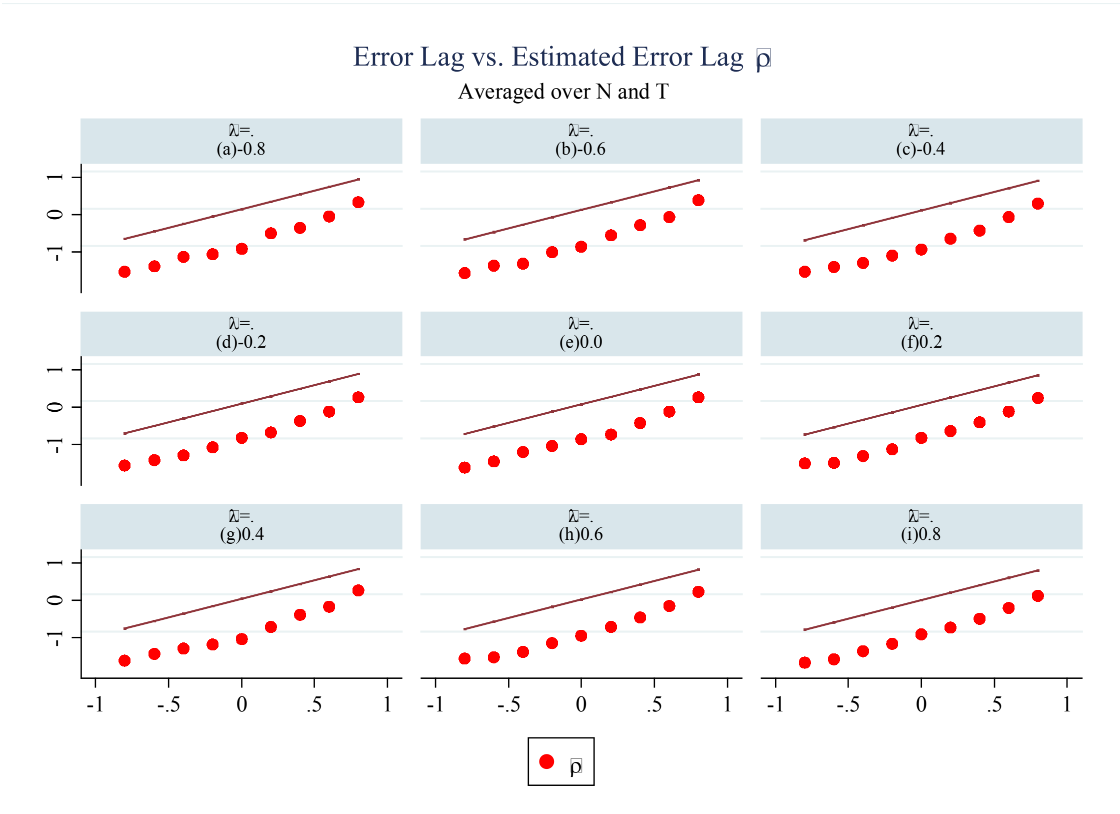
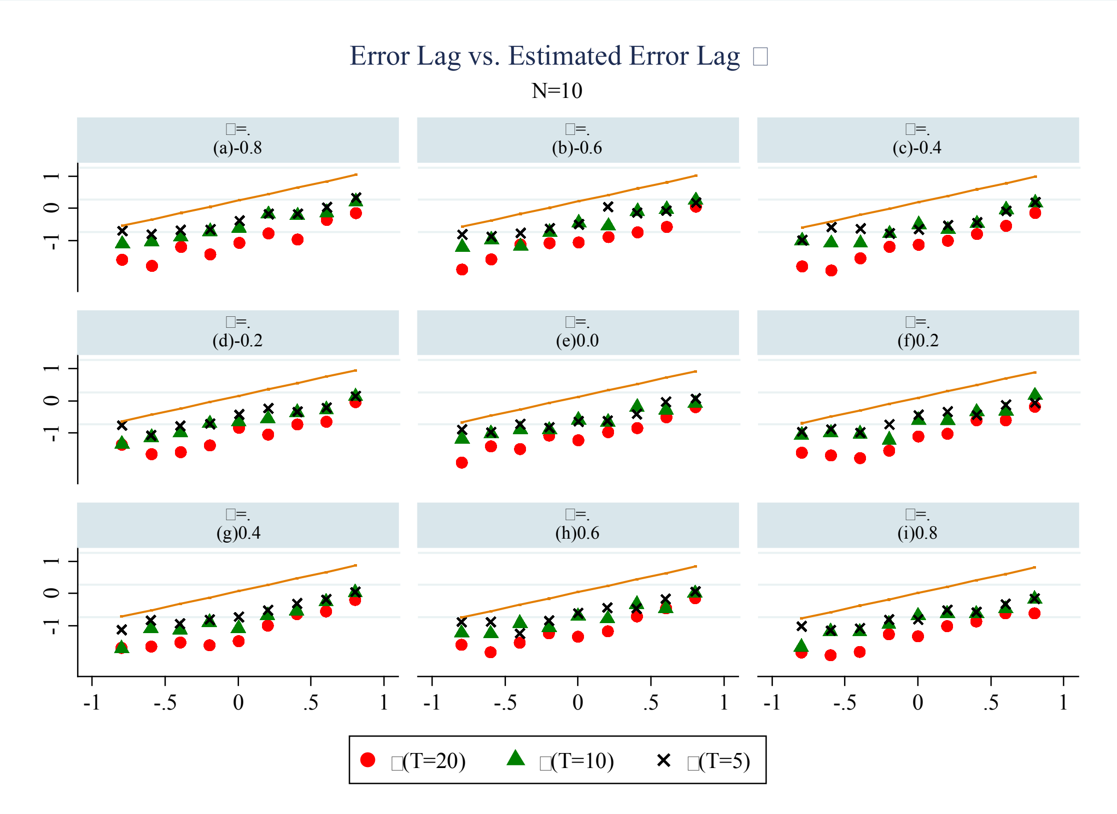
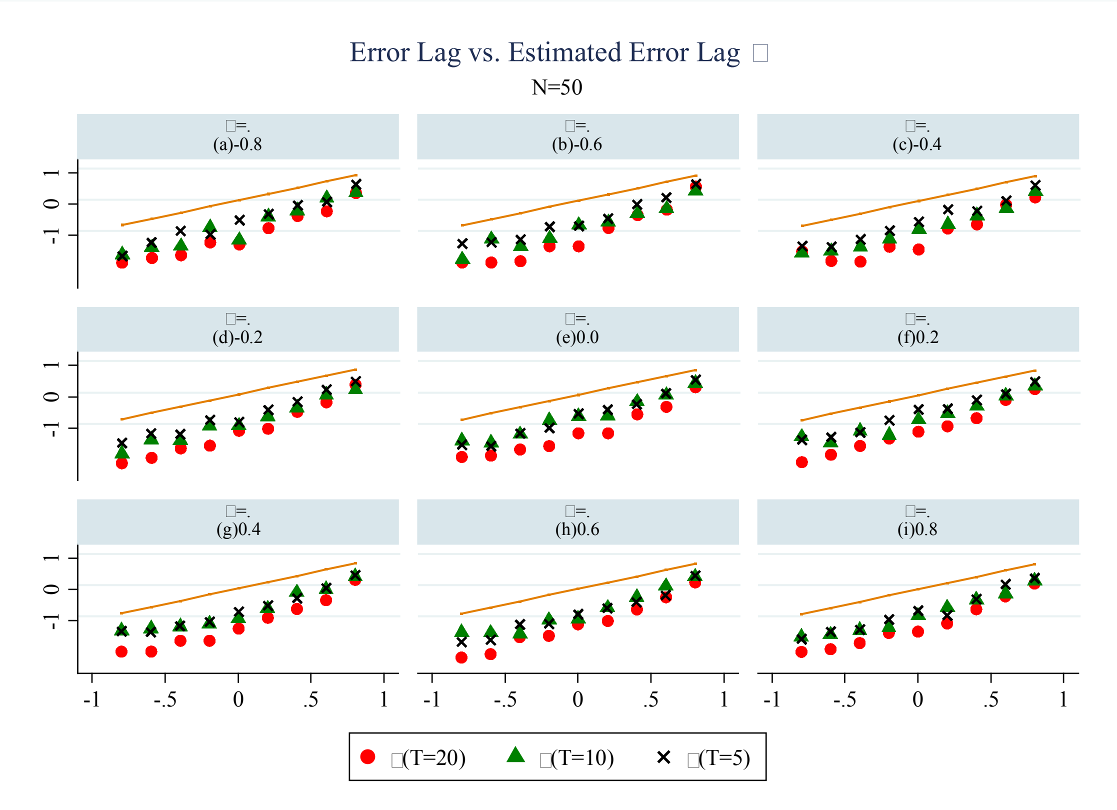


Figure III-16: True values of the error lag vs. the estimated error lag ρ (3 moment conditions), over λ




Error Lag vs. Estimated Error Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
(T=20)
(T=10)
(T=5)


Error Lag vs. Estimated Error Lag
N=50
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
(T=20)
(T=10)
(T=5)

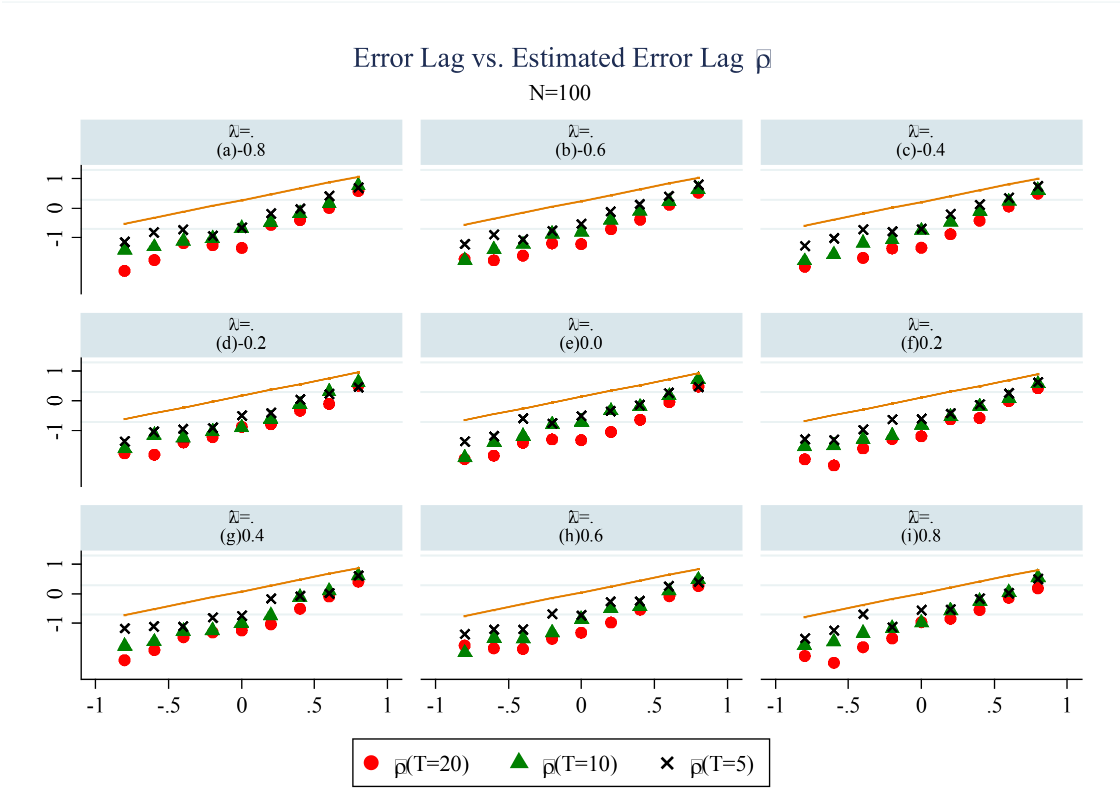


Figure III-17: True values of the error lag vs. the estimated error lag ρ (6 moment conditions), over λ

Figure III-16 depicts the true values of the spatial error versus the estimated values of the spatial error when only the first three moment conditions of equation (III.31) in subsection 3.4 are considered to estimate the spatial error. The graphs are grouped by the true value of the spatial lag. The upper left graph shows the situation averaged over all observations and shows that there is no significant bias on the spatial error estimates influenced by the spatial lag.

If the number of observations is small ($N=10$), then estimates of the spatial error are upward biased the larger the absolute value of the spatial lag is. For larger N ($N=50$, $N=100$), the accurate of an estimation of a spatial error close to zero (say $|\rho| < 0.3$) becomes less accurate. The effect is more dominant the larger the absolute value of the spatial lag is.

Comparing this result with the upper left graph of Figure III-17, in which all moment conditions are considered, it is easy to recognise that the estimated error lag is biased downward and, therefore, we confirm that the use of the first three moment conditions results in much better estimates.

We do not tabulate the coefficients here, since the short-term estimates are highly insensitive to variation in both the spatial lag and the spatial error.

We do not tabulate the coefficients here – since the short-term estimates are highly insensitive to variation in both the spatial lag and the spatial error.

6.2.2 The 'spatialbootstrap'-Command

Since the estimation of the spatial error uses a non-linear least square regression on the moment conditions, the resulting standard error is not the true standard error. For this reason, the command was implemented to obtain the standard error by bootstrapping the estimation. The procedure estimates the model of interest assuming a spatial lag model and then assigns the estimated residuals randomly to the spatial units. Then, a Kelejian-Prucha model is estimated and all coefficients and the value of the spatial error are stored in a matrix, “**param**” (parameters). Similar to the command, a normalised spatial weighting matrix, “**WW**”, has to be stored. The syntax is simply

By default, there are 250 bootstrap estimations. The value can be lowered by assessing the ado file. The program requires as much computational time as 1 hour for small panels and up to 0.5 days for large panels ($NT>5000$).

The program calculates a set of summary statistics which are the mean (mean), the 25% percentile (p25), the 50% percentile (p50), the 75% percentile (p75), the standard deviation and a corresponding t-statistic.

We ran the program to obtain bootstrap values for the coefficient of the spatial error ρ according to Model 6. We conducted 250 estimations. The result is depicted below.

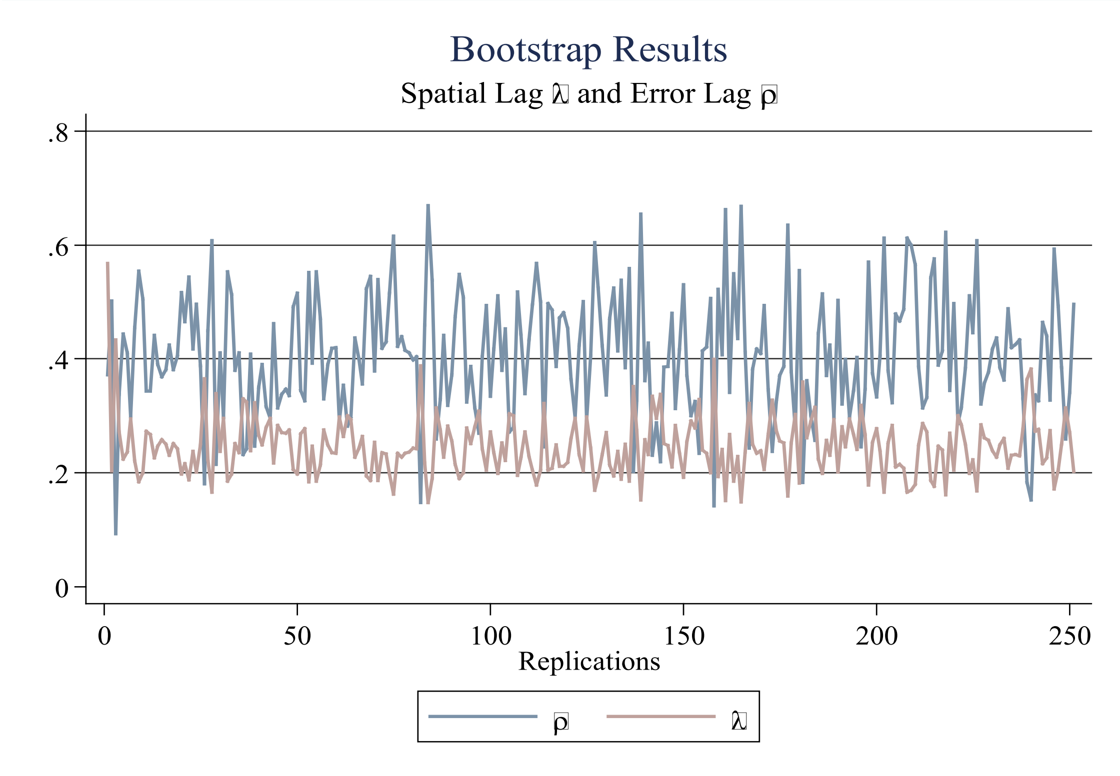


Figure III-18: Bootstrap results of the spatial parameters ρ and λ according to Model 6

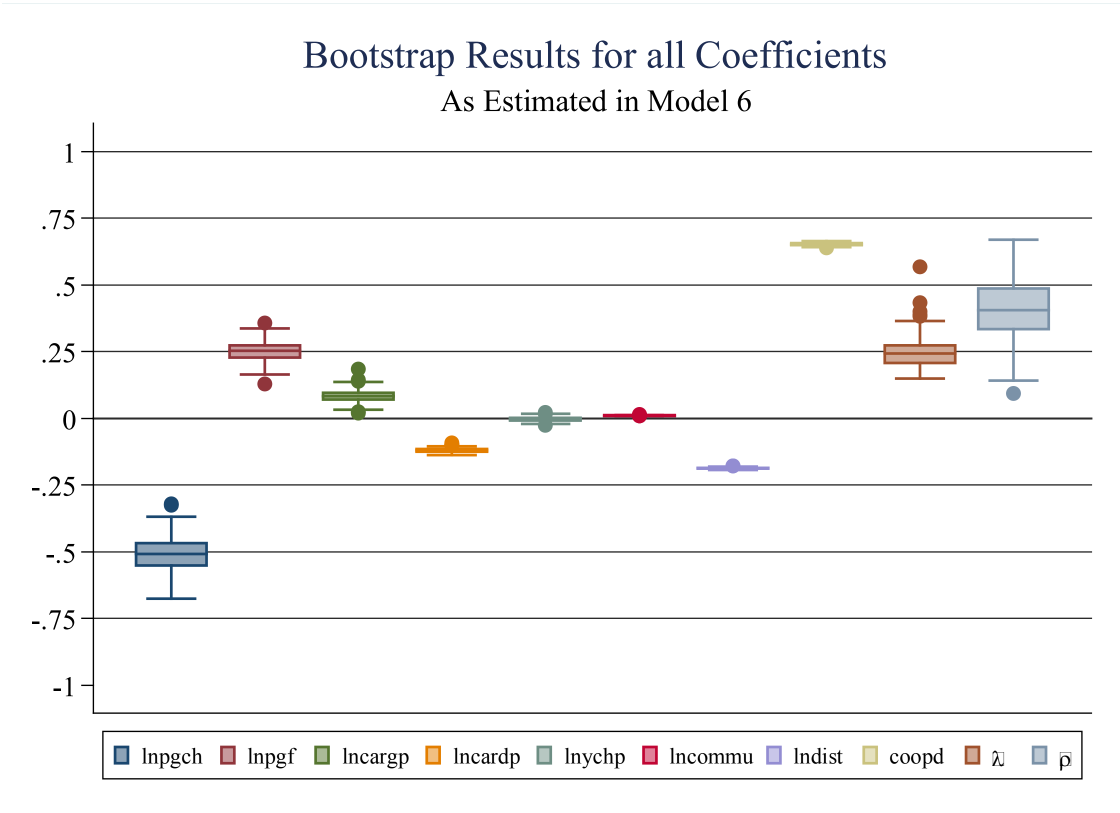


Figure III-19: Bootstrap results for all coefficients according to Model 6

The mean of ρ obtained equals 0.407, and the respective standard deviation from the bootstrap regressions is 0.107. Therefore, we conclude that the parameter is statistically significant at the 1% level according to the bootstrap. One can calculate the bootstrap-adjusted standard deviation, which is $sd_b = \sqrt{sd^2 + (IQR/1.35)^2}$. According to this, the standard deviation increases to 0.156.

Similarly, the other parameters estimated in Model 6 are bootstrapped as well. For completeness, the spatial lag is pictured too and can be compared to the estimation results in Model 6 (the bootstrap slightly underestimates the value obtained from the estimation). The other parameters could be depicted in a similar way. The point estimates of the bootstrap regressions and the standard deviations obtained are close to that obtained in the estimation of Model 6.

6.3 Specification Testing

In this subsection, we present the Monte Carlo results for the specification tests which were presented in subsections 3.5.1 - 3.5.3. The command

provides test statistics and p-values of the Lagrange multiplier test for spatial lag dependence and its robust versions, as explained in subsection 3.5.3. The command

calculates test statistics and p-values of the Lagrange multiplier test for error lag dependence and the respective robust versions according to subsection 3.5.3. The command

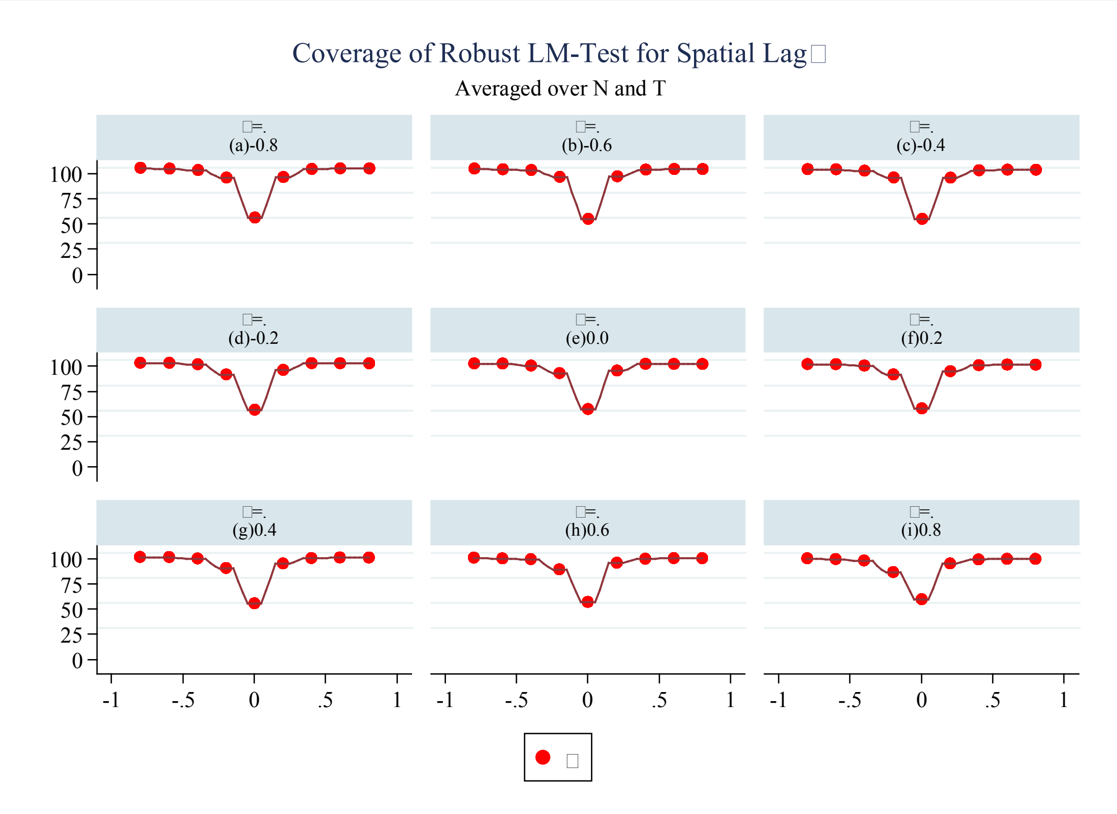
calculates the Moran-I test statistic and the corresponding p-value according to subsection 3.5.1. For all three commands, a spatial weighting matrix named WW must be stored before executing the command. We shortly present the power of the robust versions of the tests. In the Monte Carlo experiment described in subsection 6.2.1.1, we calculate the test statistics in each regression, and from this we derive the coverage and the power of the respective tests. The coverage of a test is represented by the percentage at which the test rejects the respective null hypothesis when it is actually wrong. The power of a test is represented by the percentage at which the test accepts the alternative when the null is wrong. For the results, a 95% significance level was chosen.

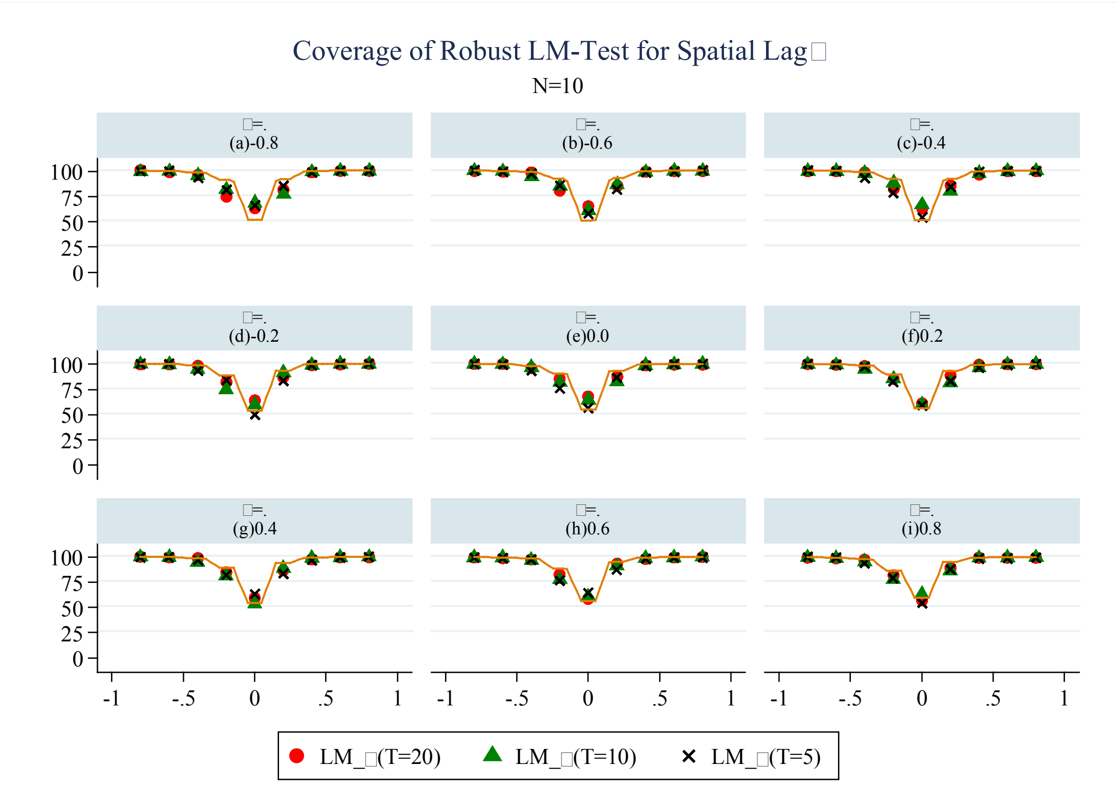
6.3.1 Coverage and Power of the Lagrange Multiplier Test for Spatial Lag Dependence (LM_{lag}-Test)

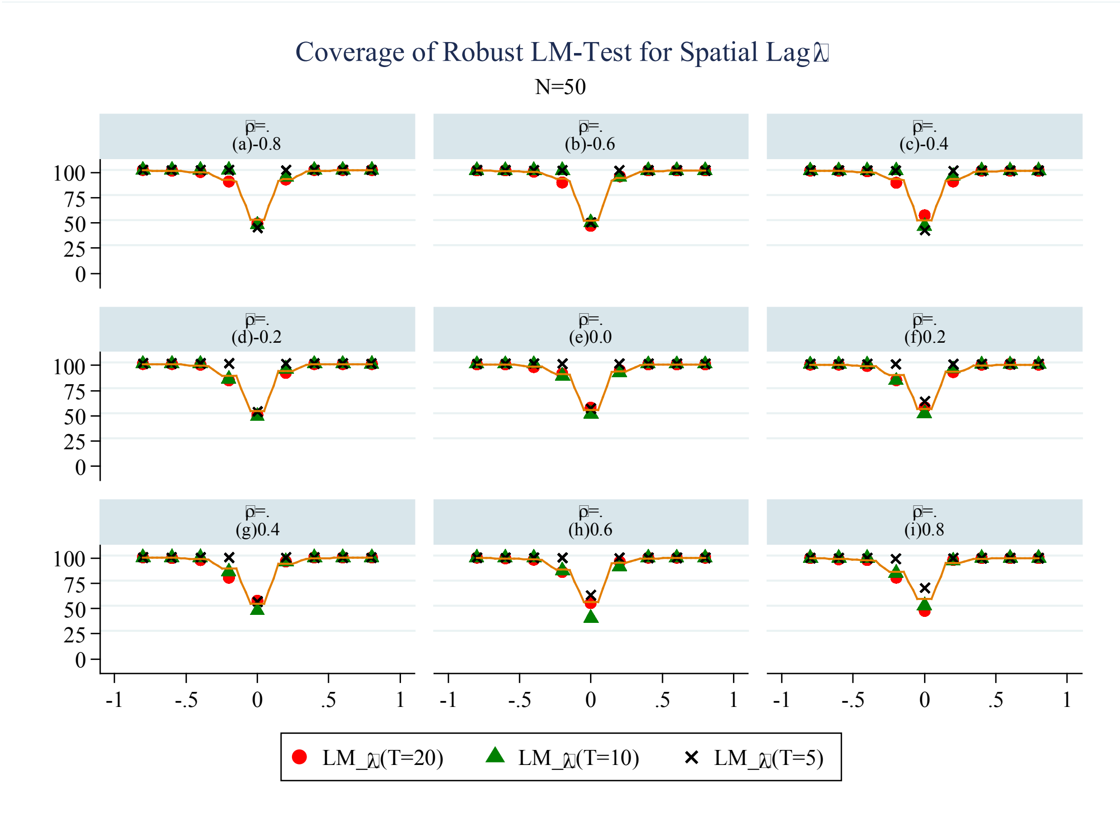
Figure III-20 depicts the coverage of the robust LM-test for spatial lag dependence. The upper left graph shows the situation averaged over N and T. It is clear that the coverage of the test is not affected by the spatial error lag ρ. For small values of the spatial lag, say $|\lambda|< 0.2$, the coverage is below 90% but above 95% for higher values of λ. This also holds for the situations in which N and T are varied. It is clear that the higher the fraction of N/T is, the better is the coverage of the test.

Considering the power of the test in Figure III-21, this becomes even more obvious: The upper left graph shows the power of the test averaged over N and T. Here too, the test is 10% and rises to some 70% when the spatial lag increases to 0.2. For higher values, the power is over 90%.

For the power, we also conclude that the higher the ratio of N/T, the higher is the power of the test. We conclude that the implemented LM-test for spatial lag dependence such as described in subsection 3.5.3 performs well in terms of coverage and power. A crucial point is that the ratio of N/T is large. In our case, where we had $N=547$ and $T=8$, this was clearly so.


Coverage of Robust LM-Test for Spatial Lag
Averaged over N and T
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8


Coverage of Robust LM-Test for Spatial Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
LM_(T=20)
LM_(T=10)
LM_(T=5)



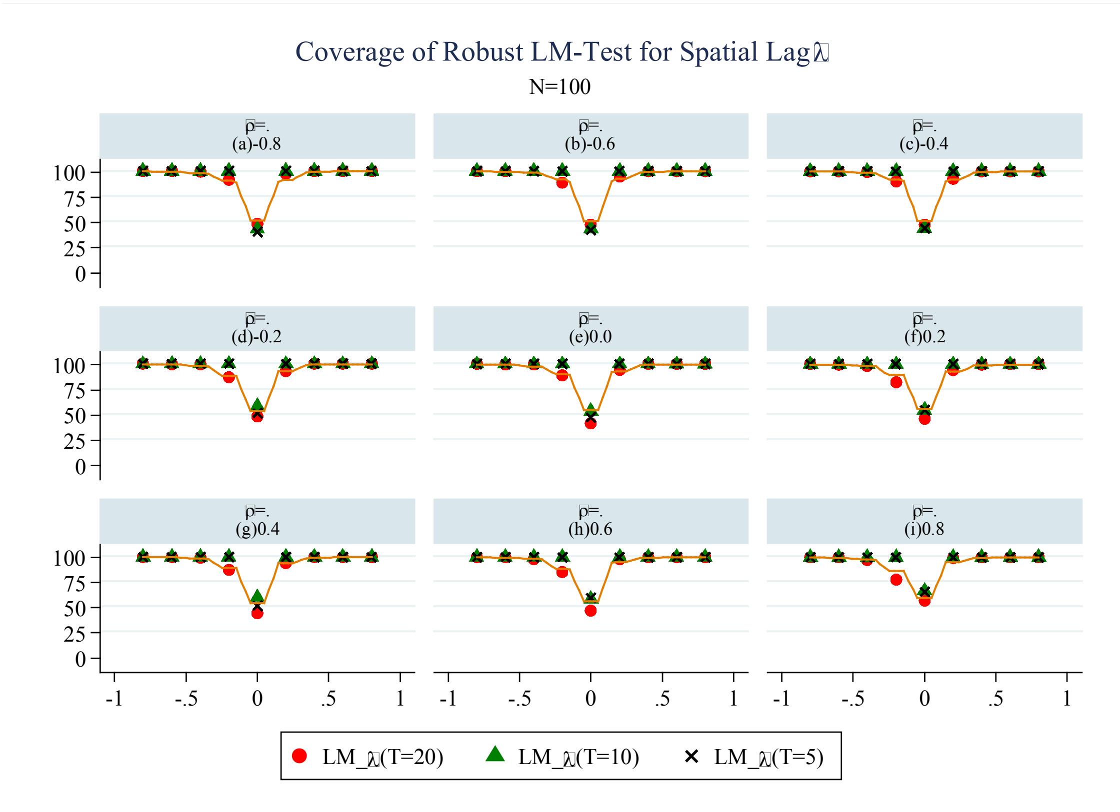
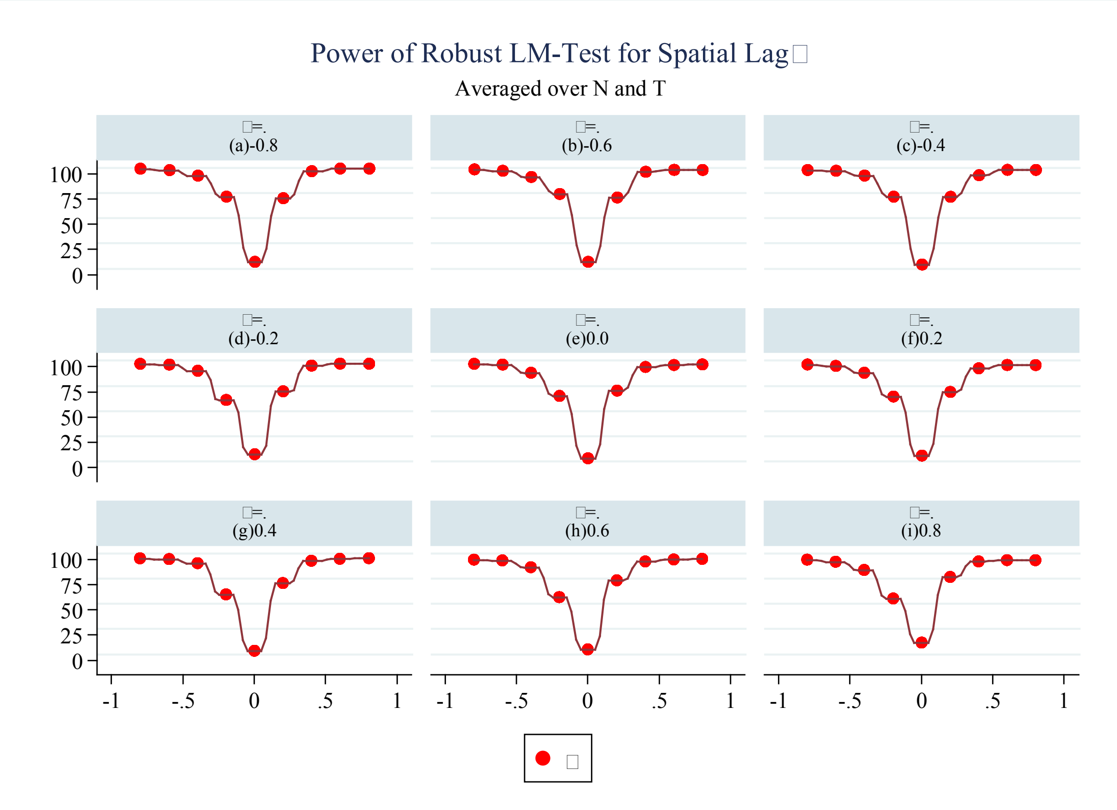
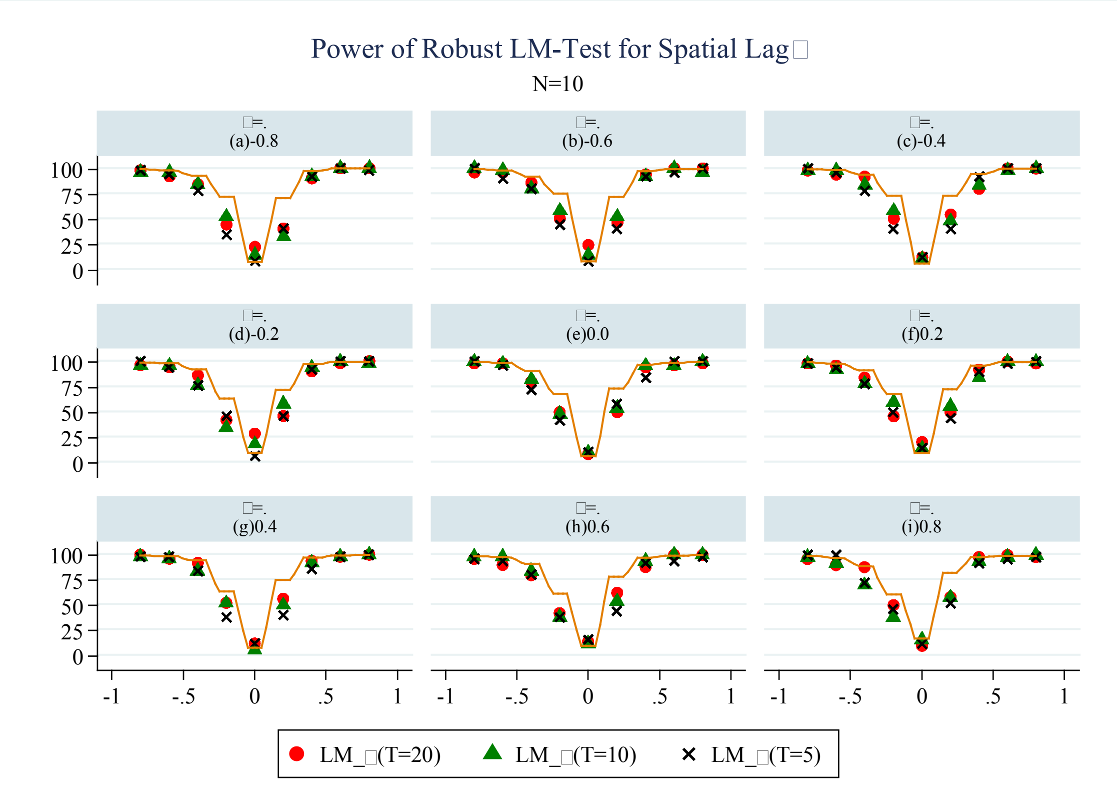
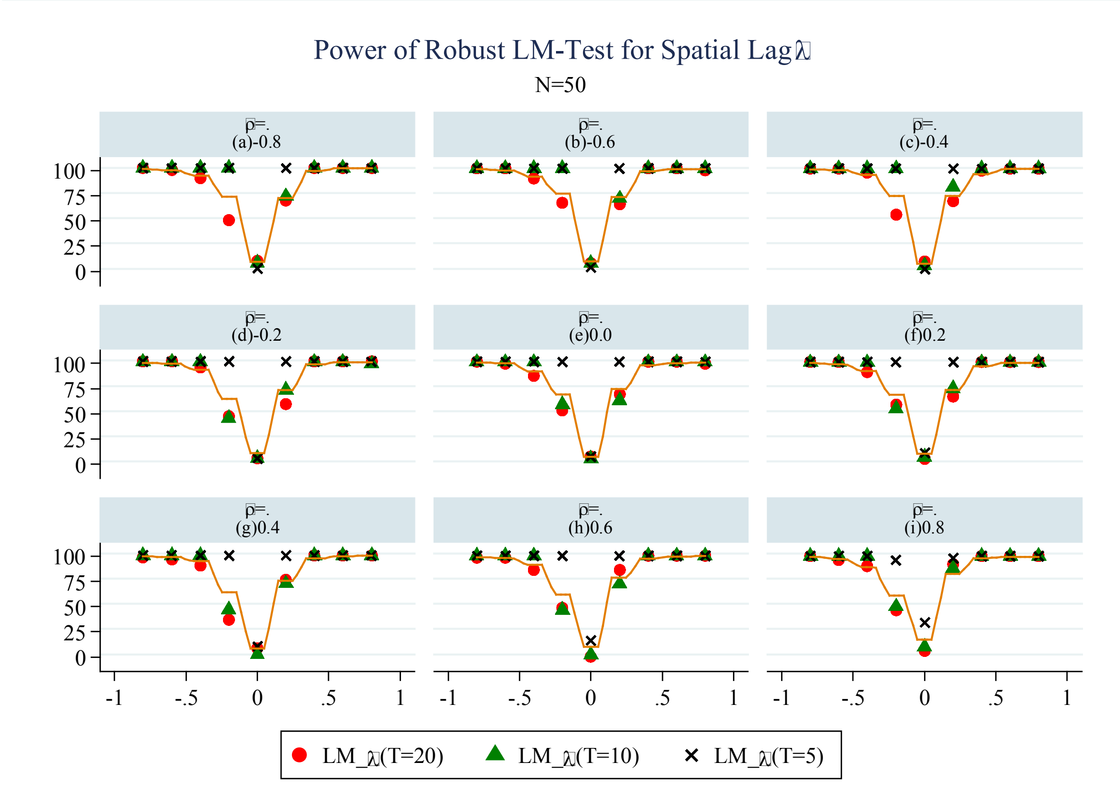


Figure III-20: Coverage of the robust lagrange multiplier test for spatial lag dependence, over ρ


Power of Robust LM-Test for Spatial Lag
Averaged over N and T
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1


Power of Robust LM-Test for Spatial Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1
LM_(T=20)
LM_(T=10)
LM_(T=5)



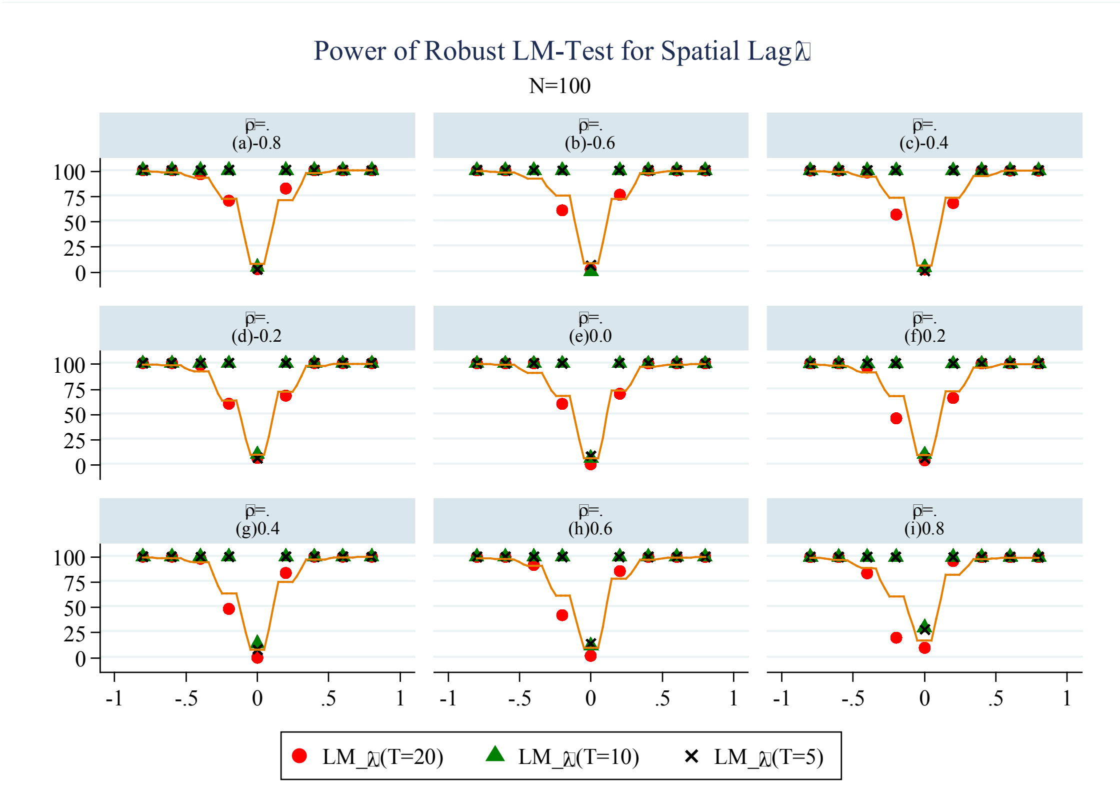


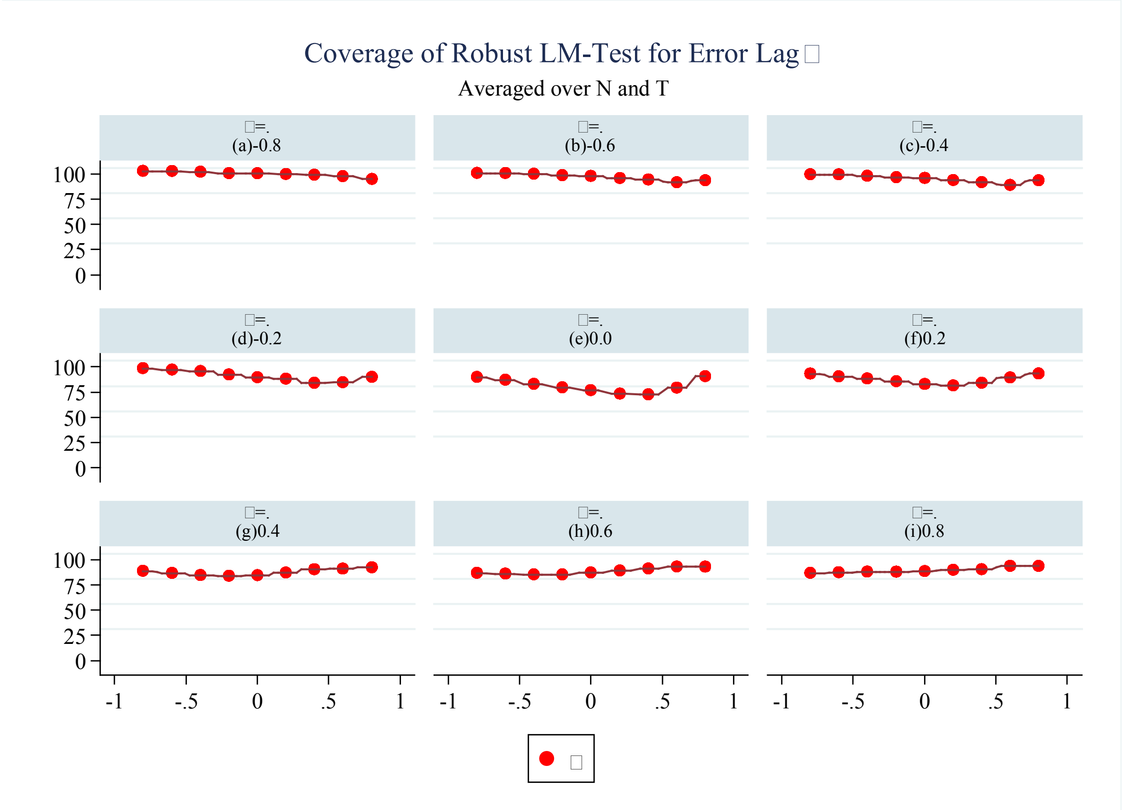
Figure III-21: Power of the robust lagrange multiplier test for spatial lag dependence, over ρ

6.3.2 Coverage and Power of the Lagrange Multiplier Test for Spatial Autocorrelation (LM_{err}-Test)

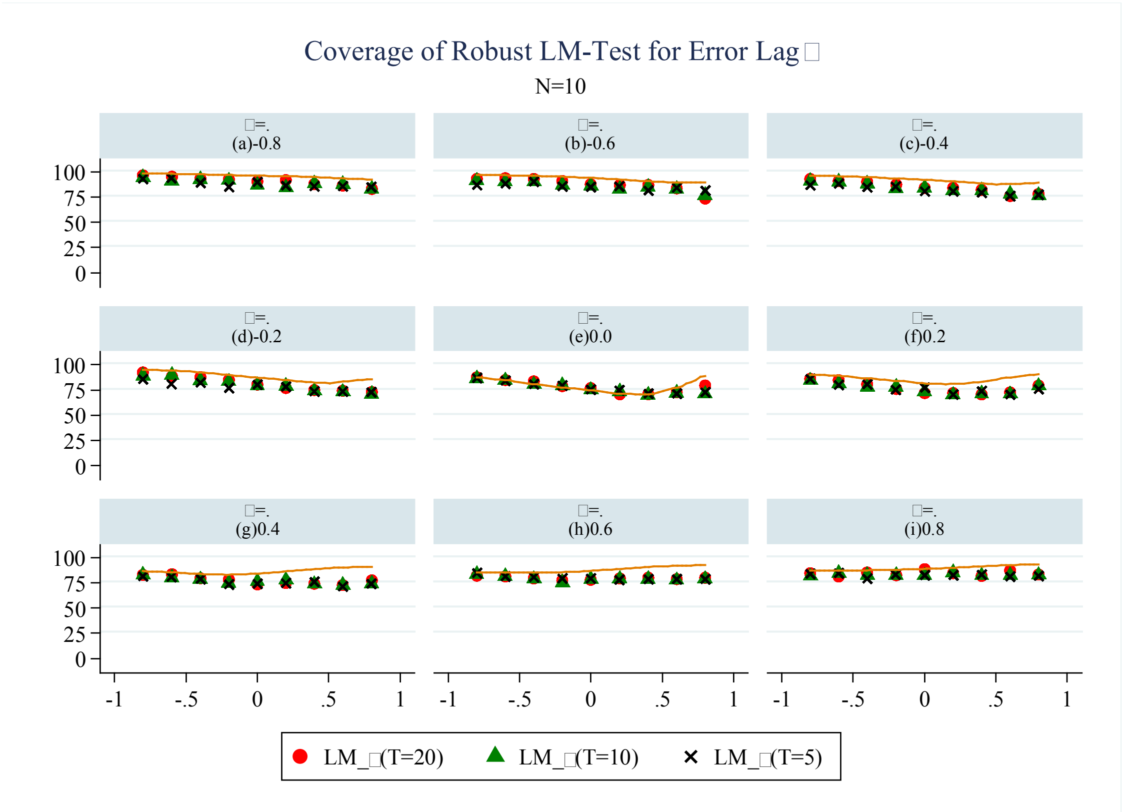
Figure III-22 and Figure III-23 depict the coverage and power of the robust LM-test for spatial auto-correlation. From the upper left graph in Figure III-22, it can be seen that the coverage does not vary strongly over different values of the spatial lag λ. However, for small values of λ, the coverage is much too high, which is actually true for all possible N, T combinations shown. Considering the power of the test in the upper left graph of Figure III-23, it can be seen that the power of the test is very small in general and almost constant over different levels of ρ except for $\lambda = 0$. We think, therefore, that the test is very sensitive to the possible presence of a spatially lagged dependent variable and, since the specification of the test is not robust with respect to that, this was much less the case for the LM-test for spatial lag dependence. Further, it can be seen in Figure III-22 and Figure III-23 that for small N (e.g. $\mathrm{N}=10$), the test seems to fail completely. For the large panel case with e.g. $\mathrm{N}=100$, the results for coverage and power only make sense if $\lambda=0$. If the spatial lag is different from zero, coverage and power of the test are each always above 90%, no matter what the value of ρ actually is. As the time dimension grows, power and coverage of the test decrease significantly.

Accordingly for our case where we have the presence of spatial dependence in the dependent variable and the number of units N is much higher than T, it is not surprising that the test rejects the null of no spatial error dependence. On the other hand, it should be noted that the Monte Carlo design was more focused on the spatial lag in the dependent variable since, for instance, the variance of the residuals was not varied, and the results of this subsection might change if one is doing so. In fact, the variance of the dependent variable predicted in the experiment by design always exceeds the variance of the residuals (meaning the regressions show a high R^2). This is not true for our empirical example. In the original work by Baltagi et al. (2003), the variance of the overall error component exceeded the variance of the dependent variable (by far), and results for the coverage and power of the test were much better than tabulated here. Moreover, it was stated that if $\sigma_\mu^2 / (\sigma_\mu^2 + \sigma_v^2)$ is large (which is the case in our example) and the overall error variance itself is large, then the test characteristics improve significantly. Therefore, we should have tabulated the characteristics of this robust test for spatial error dependence with respect to varying variance parameters and probably, the characteristics of the LM_{lag} tests then might more strongly depend on the possible presence of spatial auto-correlation. However, we conclude here that the test statistic is very sensitive to the presence of a spatially lagged dependent variable, which was not the case for the LM_{lag} test. Therefore, we intend to modify the LM_{err} test such that it might be "robust" with respect to the possible presence of a spatially lagged dependent variable.

In the next subsection, we therefore describe a modified version of the test statistic described in 3.5.2 by taking the residuals from a spatial lag model. Then, we believe, the test statistic will be robust against the possible presence of a spatially lagged dependent variable.



Coverage of Robust LM-Test for Error Lag
Averaged over N and T
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1



Coverage of Robust LM-Test for Error Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1
LM_(T=20)
LM_(T=10)
LM_(T=5)

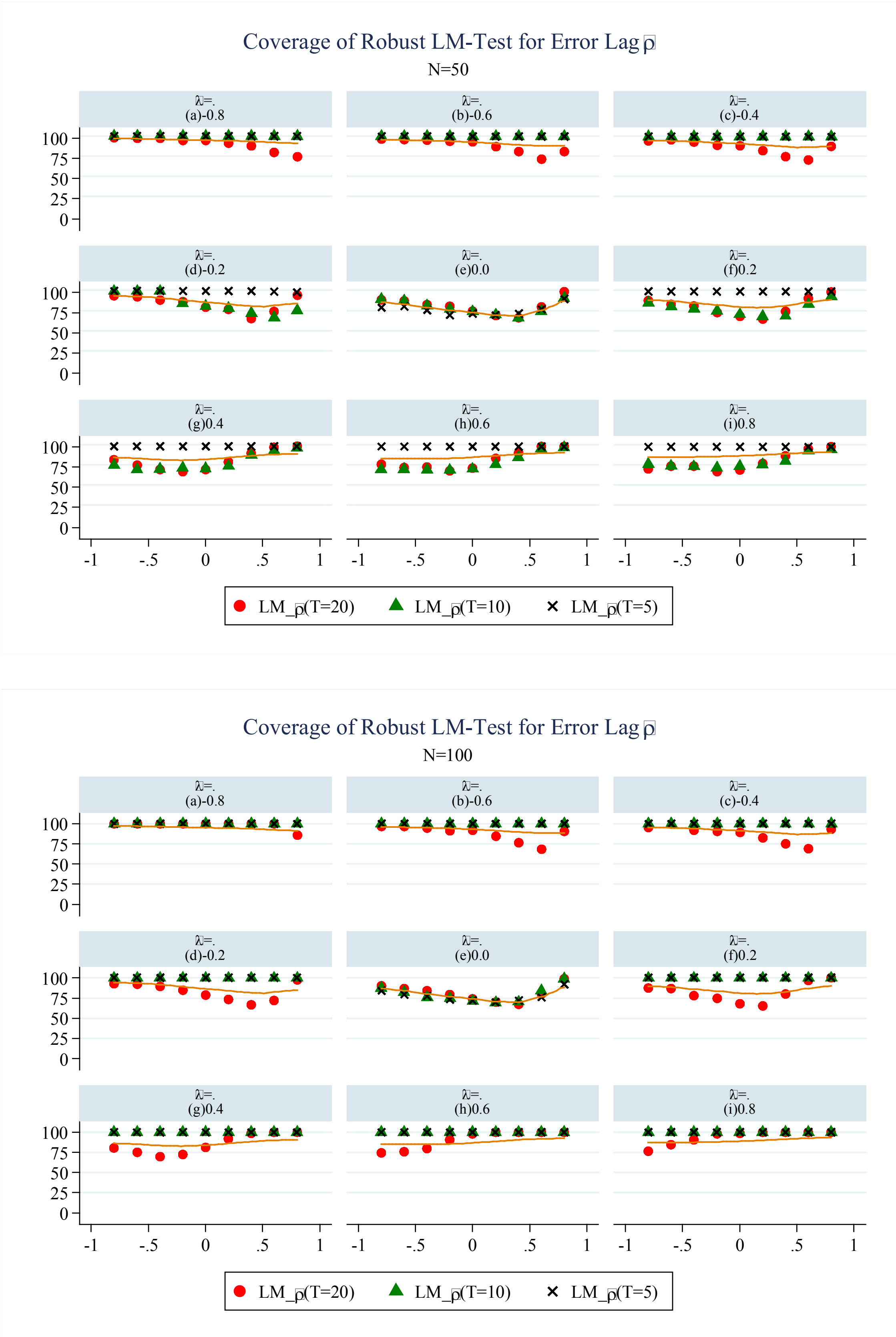
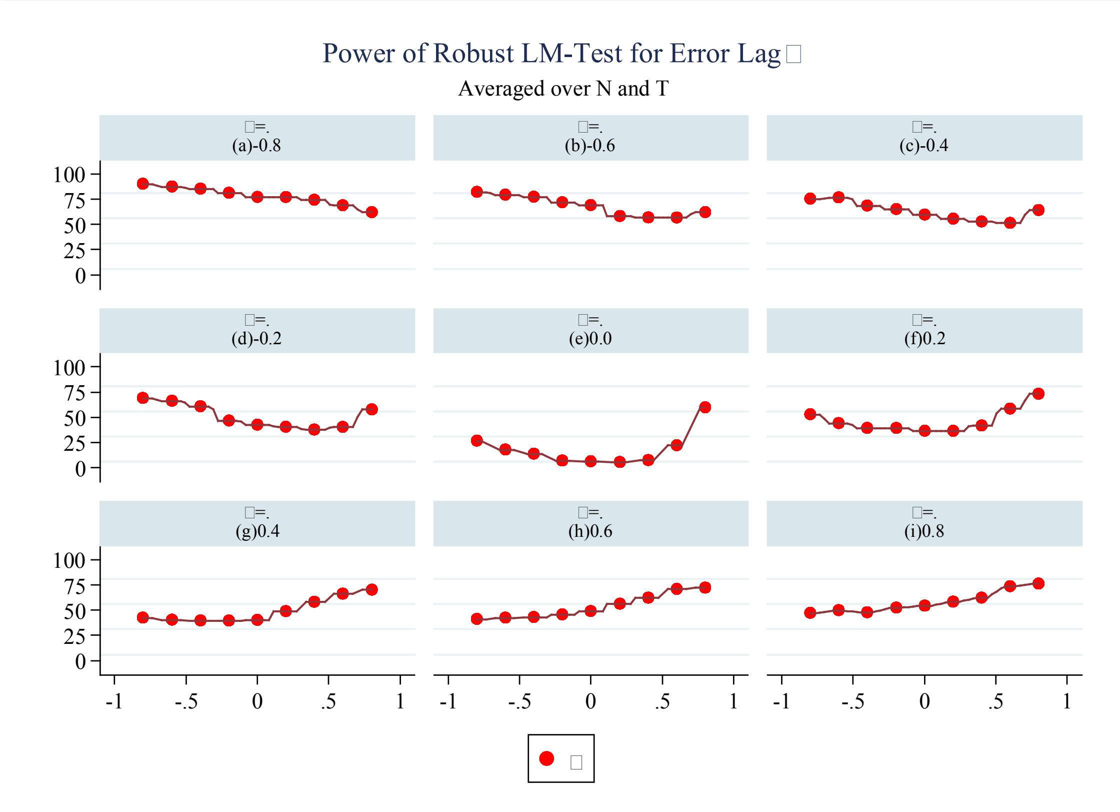
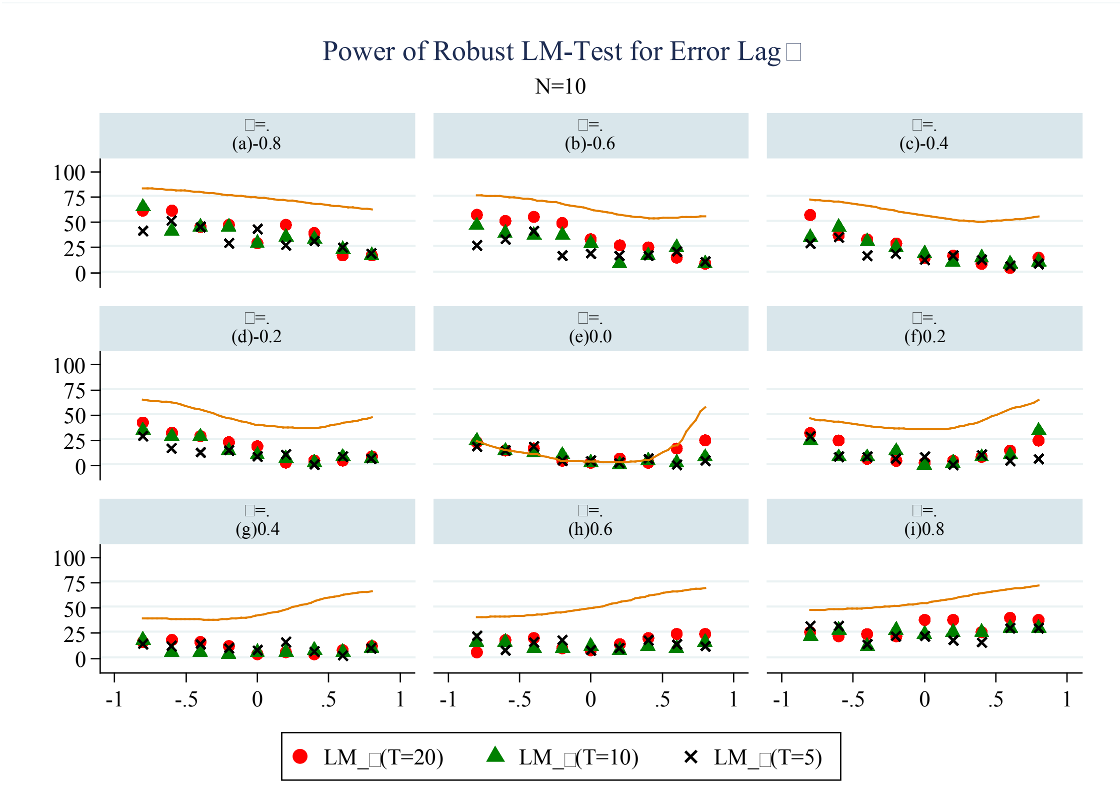
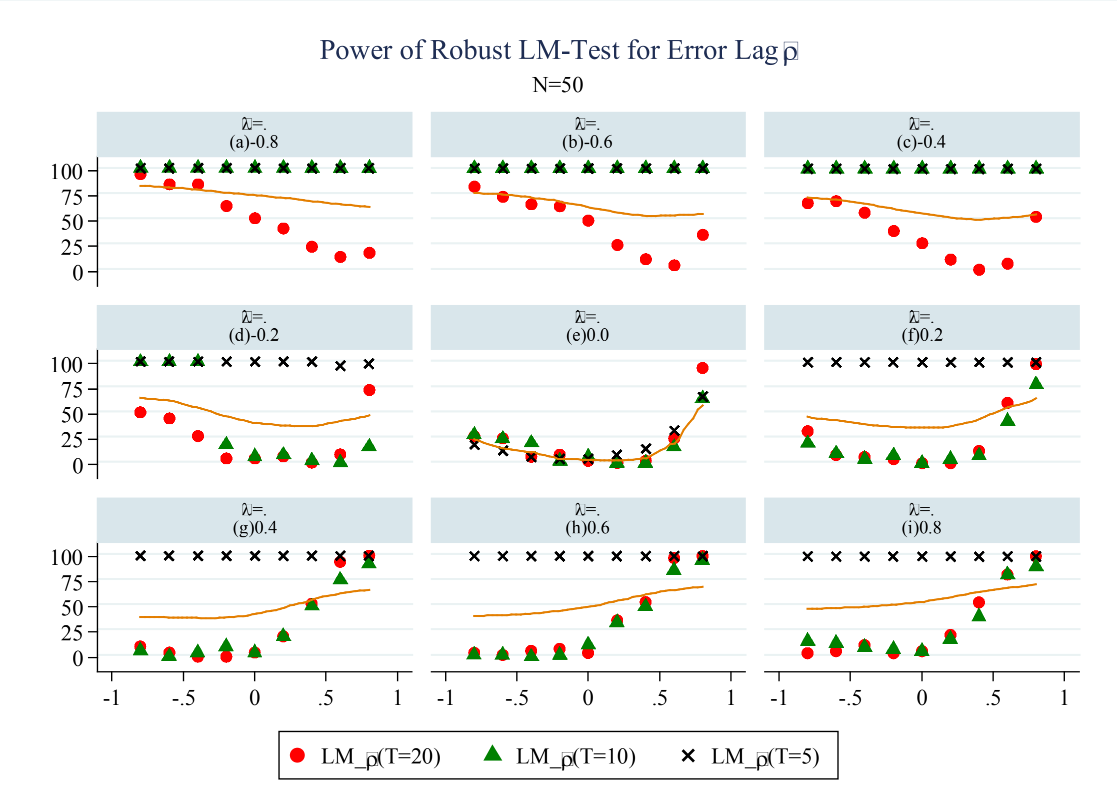


Figure III-22: Coverage of the robust lagrange multiplier test for spatial error lag dependence, over λ


Power of Robust LM-Test for Error Lag
Averaged over N and T
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1


Power of Robust LM-Test for Error Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1
LM_(T=20)
LM_(T=10)
LM_(T=5)



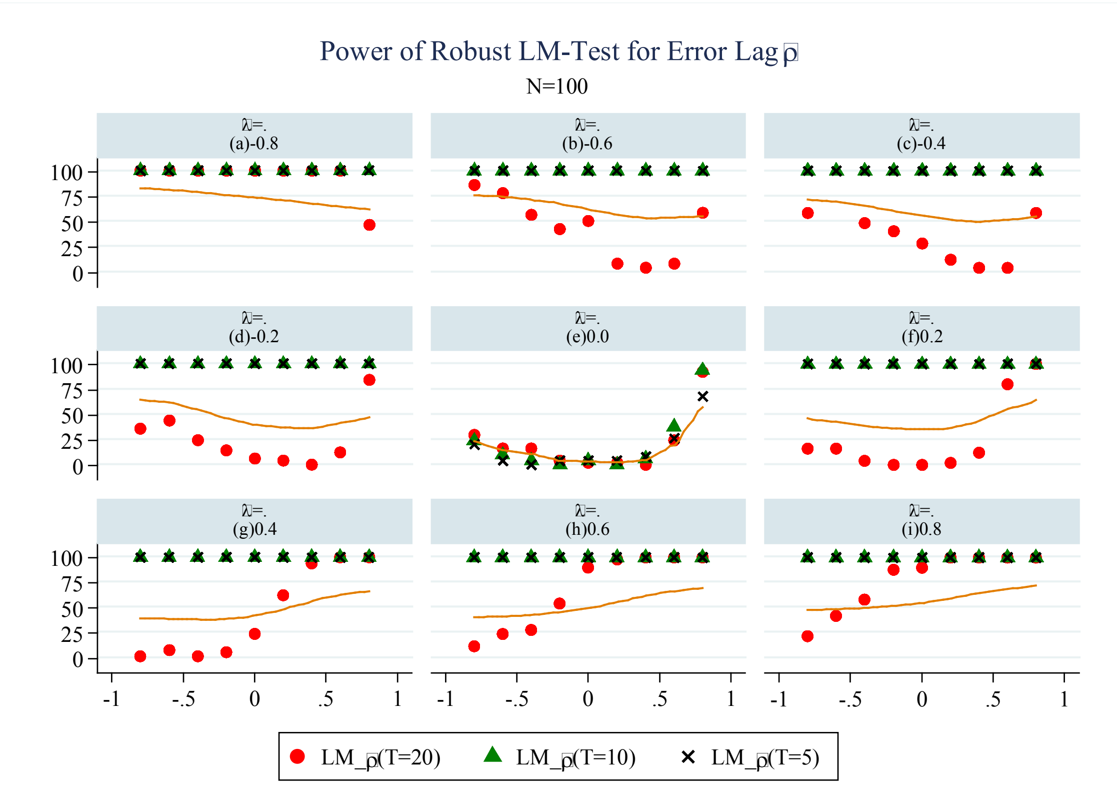


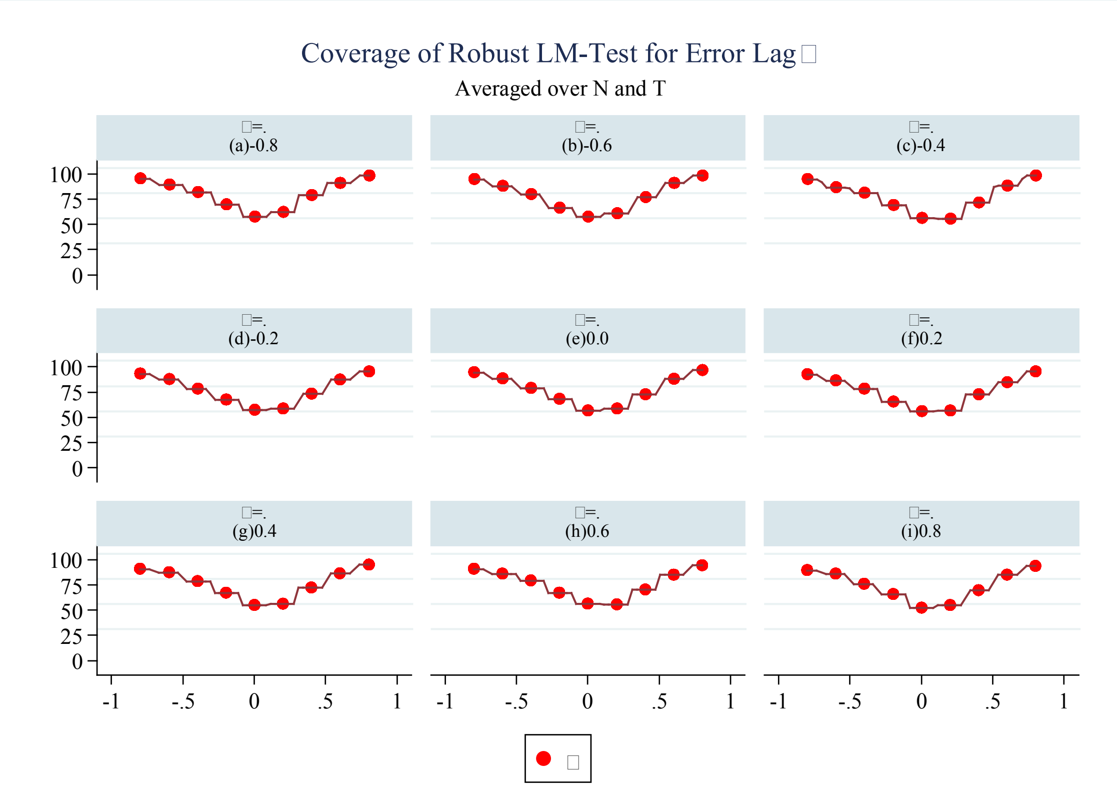
Figure III-23: Power of the robust lagrange multiplier test for spatial error lag dependence, over λ

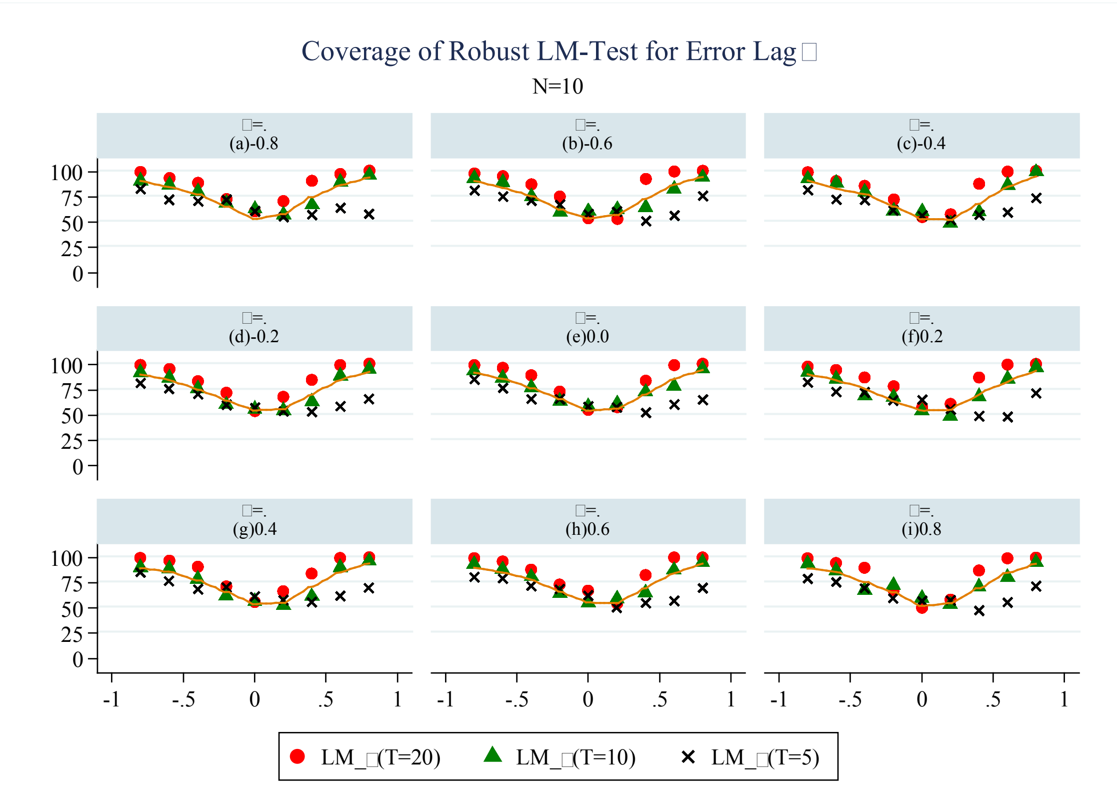
6.3.3 Coverage and Power of the Modified Lagrange Multiplier Test for Spatial Autocorrelation (LM_{err}-Test)

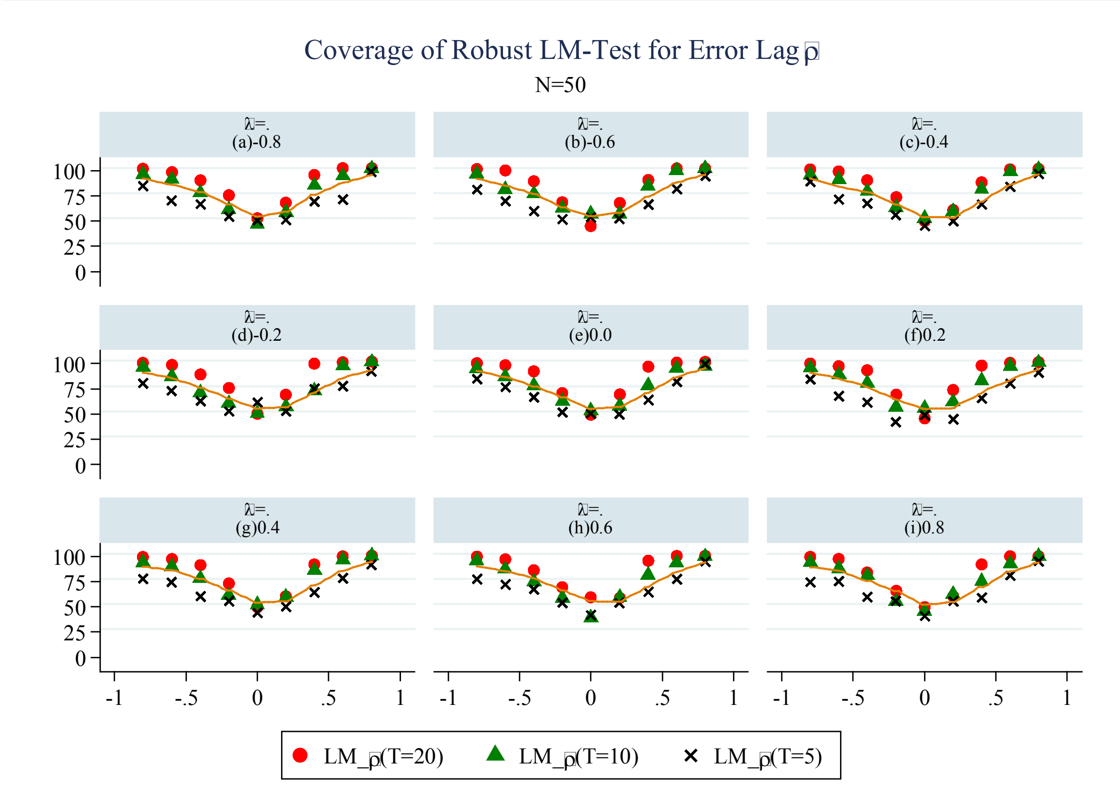
Figure III-24 and Figure III-25 show the coverage and power of the modified robust LM test statistic for spatial auto-correlation. Now the performance of the test statistic is much better when we use the residuals from a spatial lag model to calculate the test statistic.

From the upper left graph in Figure III-24, we see that coverage of the test (averaged over all N, T) is quasi-independent over the different values of λ. It can be seen that on average, the coverage is still not very good in terms of significance. Actually, the significance level against the type I error (coverage) only reaches 90% if the absolute value of the coefficient of the spatial error exceeds 0.6, which is very high, given that the value ranges from -1 to 1. Considering the other graphs in Figure III-24, one can see that the coverage of the statistic is not dependent on the values of the coefficient λ of the spatial lag. Moreover, the coverage increases the more panel units N are available and the more time periods T are available in the data. It can be seen in Figure III-24 that for $N=100$ and $T=20$, the coverage of the test already approaches the 90% level when the (true) absolute value of ρ is between 0.2 and 0.4.

The same comments hold for the power of the test depicted in Figure III-25. There, it becomes clearer that the characteristics of the test statistics improve when many time periods are available in the data, e.g. $T=20$. Generally, we conclude that the power and the coverage of the test statistic are rather low, which might be due to the fact that the variance of the residuals is not varied in our Monte Carlo experiment. We therefore conclude that the null of no spatial auto-correlation could be rejected even at the 10% level and not only at the 5% level.


Coverage of Robust LM-Test for Error Lag
Averaged over N and T
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1


Coverage of Robust LM-Test for Error Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1
LM_(T=20)
LM_(T=10)
LM_(T=5)



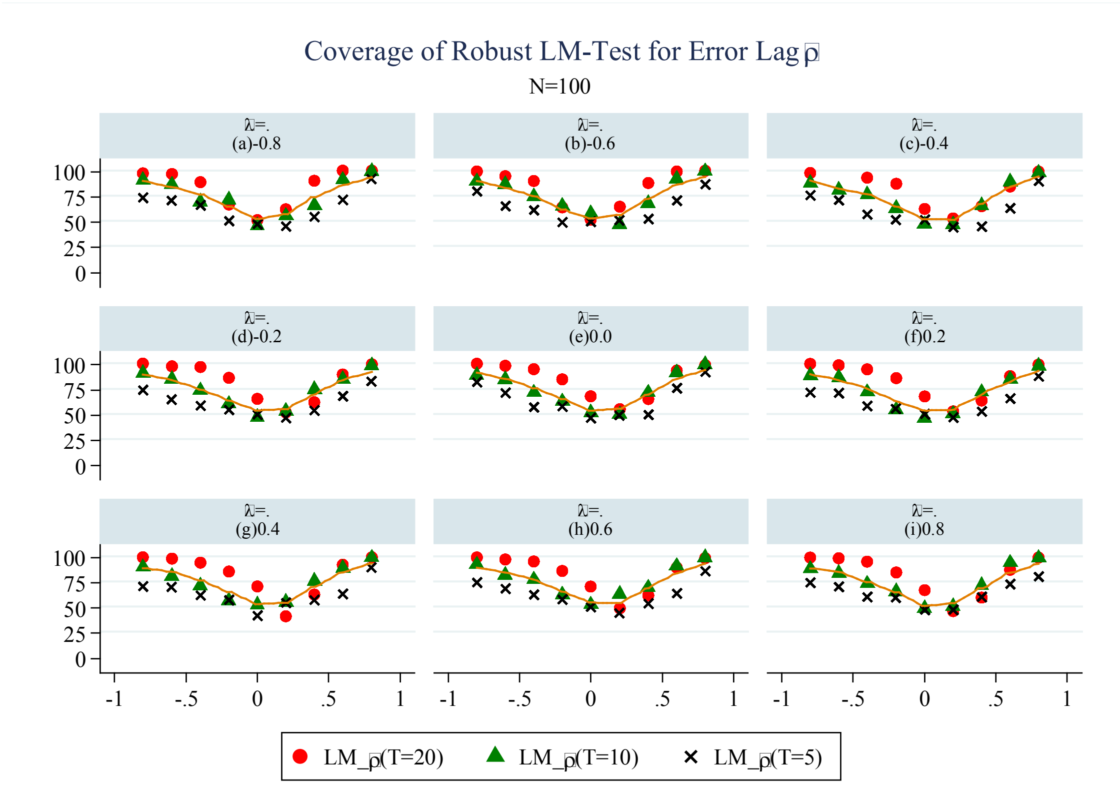
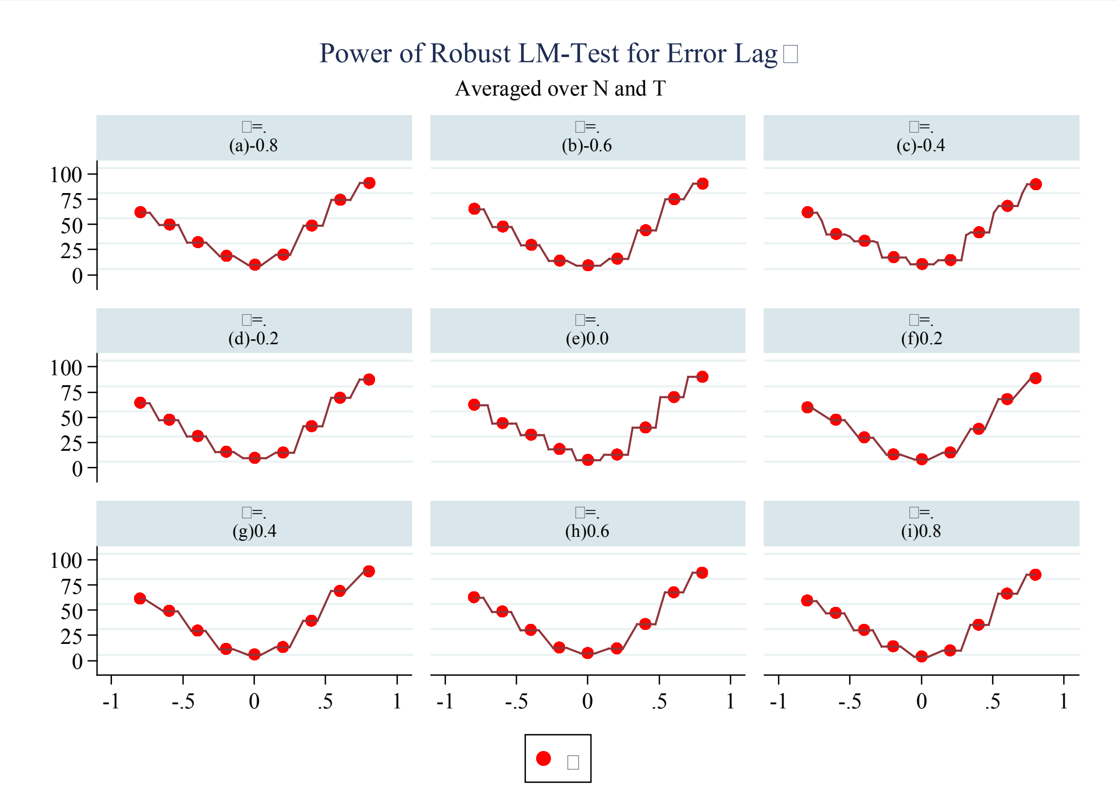
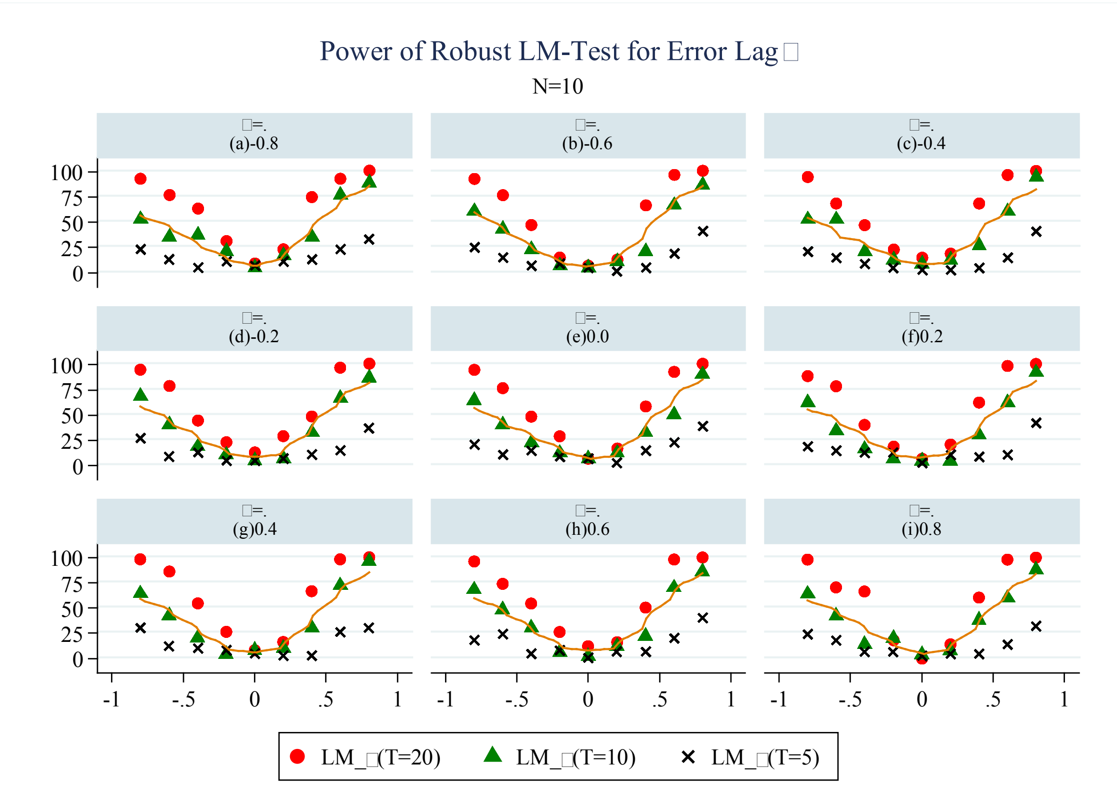
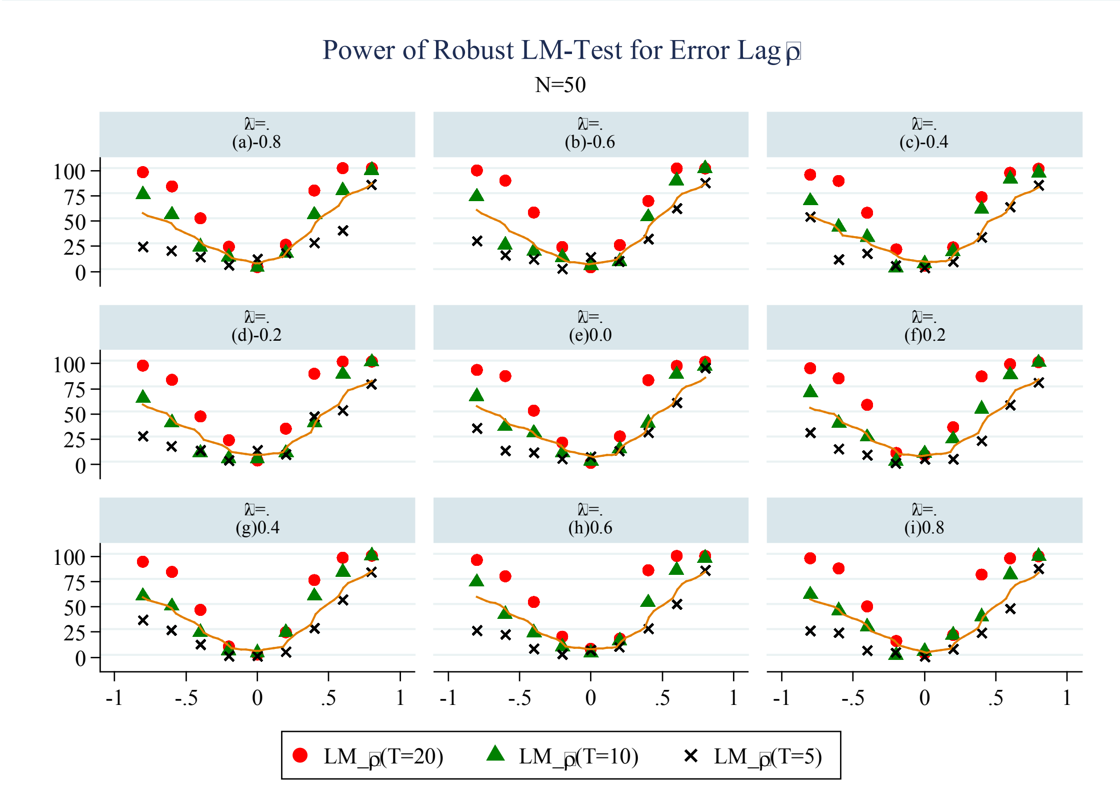


Figure III-24: Coverage of the modified robust lagrange multiplier test for spatial error lag dependence, over λ


Power of Robust LM-Test for Error Lag
Averaged over N and T
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1


Power of Robust LM-Test for Error Lag
N=10
(a)-0.8
(b)-0.6
(c)-0.4
(d)-0.2
(e)0.0
(f)0.2
(g)0.4
(h)0.6
(i)0.8
100
75
50
25
0
-1
-.5
0
.5
1
LM_(T=20)
LM_(T=10)
LM_(T=5)



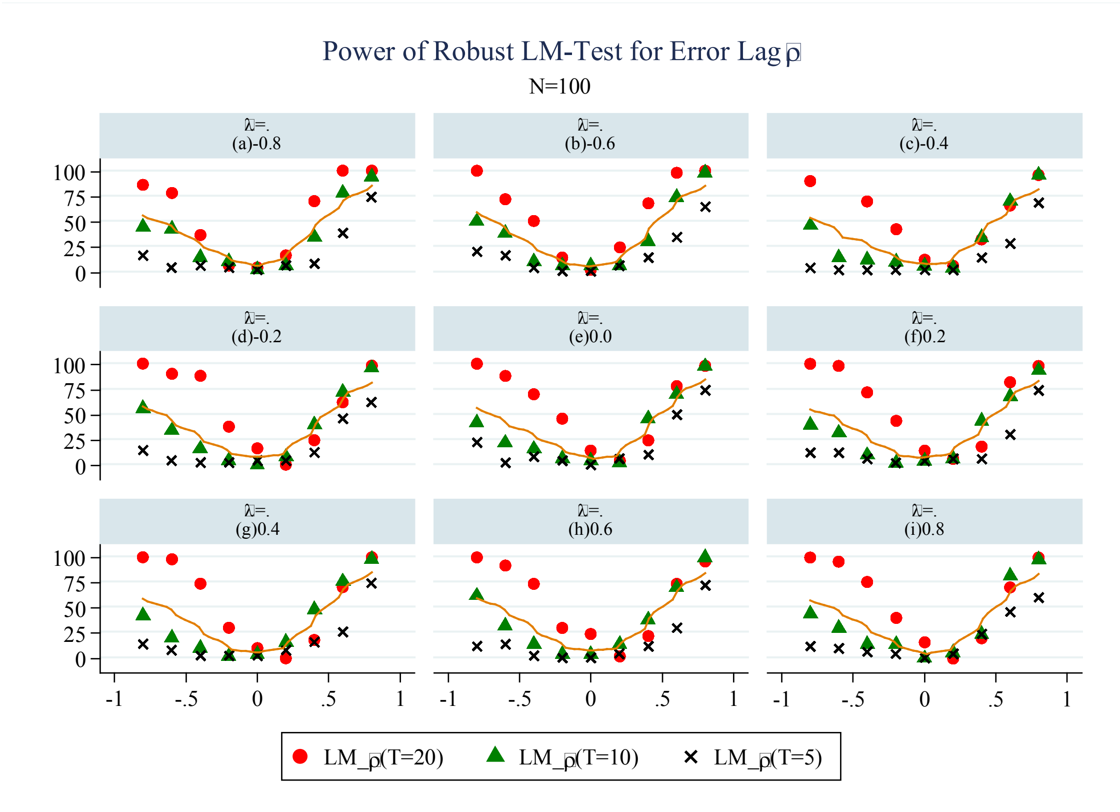


Figure III-25: Power of the modified robust lagrange multiplier test for spatial error lag dependence, over λ

IV System GMM and Difference GMM - The Impact of Low Within Variation

1 Introduction

In this paper, we want to address the impact of low within variation of exogenous variables in dynamic panel data models.

Panel data are a combination of cross-sectional and time-series observations. After the pioneering study by Balestra and Nerlove (1966), the number of studies published either dealing with estimation techniques for panel data or their application on a wide range of topics has increased enormously. Hsiao (2007) identifies three reasons for this rapid growth: First, data availability has generally increased over the past decades. Second, panel data offer more insights into the complexity of economic problems, for instance, than cross-sectional data. Third, panel data allow for simplified computation and statistical inference, under certain conditions.

The most prominent advantage of panel data is that the time dimension of the data can be exploited. Multiple observations for each cross-sectional unit allow researchers to control for unobserved (time-invariant) effects, which are often referred to as unobserved heterogeneity. Put simply, panel data allow controlling for the individuality of the cross-sectional units. Due to the two-dimensional character of panel data, the variation of the data can be described over both time (within variation) and cross-sectional units (between variation). This distinction has important consequences on the estimation of panel data models. Further, the increase in observations results in a higher degree of freedom of the models and therefore allow for more accurate inference of estimated parameters. For instance, we can have a panel data set with N = 100 units observed over a time length T = 7. Typically, microeconomic panels consist of socio-economic variables which exhibit very low variation over time. Clearly, the data set is a panel data set, but nonetheless not necessarily allows exploiting the advantages of panel data due to the low within variation.

The two most frequently used models to capture time-invariant unobserved heterogeneity in a static setting are the fixed effects and the random effects models. In the fixed effects model, unobserved heterogeneity is allowed to be correlated with the exogenous regressors. This requires the elimination of the unobserved heterogeneity from the model, which is either achieved by first-differencing or de-meaning[26] the data. In case of static models, the latter is

[26] De-meaning in this context means subtracting *individual* means.

usually applied. A significant disadvantage resulting from this "within transformation" of the data is that variables with zero within variation (constant over time) cannot be estimated and those with little within variation will be imprecisely estimated; see Cameron and Trivedi (2009). On the other hand, the most popular counterpart to the fixed effects specification is the random effects model. In the random effects model, unobserved heterogeneity is treated as purely random and therefore assumed to be uncorrelated with the exogenous regressors. Estimation is then possible by feasible generalised least squares (FGLS), which does not require first-differencing or within-transforming the data. The advantage of this is that variables with low or no within-variation can also be generally estimated with precision. The disadvantage is that the estimates are biased if the assumption of uncorrelated individual effects is violated.

The econometric literature offers many recommendations on the usage of either fixed or random effects models. The majority of the literature gives rise to the belief that applied researchers are much more sensitive to bias than variance of the estimated coefficients of an econometric model, as has been noted by Clark and Linzer (2012).

In general, the term bias denotes the difference from an evaluated statistical parameter to its true value. In a Monte Carlo experiment, bias usually refers to the difference between the mean of the replicated parameters and the true parameter value. The term variance, on the other hand, is the average of squared differences from each parameter replication to the mean of all replications and therefore a measure of dispersion about the mean. The standard deviation is simply the square root of the variance. Departing from this, a statistical parameter is said to be efficient over another if its variance (or standard deviation) is below the variance or standard deviation of the other parameter. The root mean squared error refers to the root of mean squared differences between the parameter replications to the true value of the parameter and therefore is a measure which combines both variance and bias in one measure, see e.g. Cameron and Trivedi (2009).

However, the trade-off between variance and bias of the estimated coefficients can impose serious consequences on the parameter estimates –a low bias at any price might expand the variance of the coefficient dramatically (most of all if the within variation of the regressor is low) and thus render inference on the coefficient estimate meaningless. Clark and Linzer (2012) analysed the impact of the presence of a regressor with low within variation on random and fixed effects models. As already mentioned, it is well known that many estimation techniques in either the dynamic or static panel data context make use of first-differencing or de-meaning of the variables in order to eliminate unobserved heterogeneity bias. However, doing so largely removes the between variation of the variables and thus variables with a small within variation are subject to a high variance in their coefficient estimates. Clark and Linzer (2012) conclude that it is not wise to minimise the bias of a parameter estimate at any costs, but rather to balance the trade-off between variance and the parameter estimate by considering the respective root mean squared error – a measure which incorporates both aspects, efficiency and bias. They further conclude that, for situations in which the random effects estimates of a panel data model are susceptible to bias (due to a sufficiently high correlation of the regressors with the unobserved effects), the gain in efficiency obtained by the random effects model overcompensates the potentially higher bias of those estimates.

The problem of low within variation of regressors in a panel data model, as elaborated by Clark and Linzer, is not present in the context of dynamic panel data models. Therefore, the goal of the Monte Carlo simulation in this paper is to discriminate the two most well-known dynamic panel data estimators - the Arellano-Bond estimator (1991) and the Blundell-Bond estimator (1998) in the context of bias, efficiency and root mean squared error of the parameter estimate when the exogenous regressor is subject to a relatively low within variation. On one hand, the estimator proposed by Arellano and Bond (1991) is suspected to have a higher variance of coefficient estimate than the Blundell-Bond estimator (1998). The reason is that, in addition to the first-differenced equation, the Blundell.Bond estimator uses the equation in levels and therefore also more observations. However the moment conditions derived from the equations in levels are in turn only valid for certain initial conditions, which the data generating process (DGP) has to satisfy[27]. If the initial conditions are not satisfied - which might dominantly be the case in applied econometric work - then the estimator may be biased. The goal of the present work is deliberately not to quantify this bias, but to weight it against the gain in efficiency and compare the performance to that of the Arellano-Bond estimator.

During the past years, the use of dynamic panel data models has become increasingly common in applied econometric research. According to Roodman (2007), the reason for this development lies in the abundance of specification possibilities of those models: First, the data generating process may be dynamic, such that previous realisations of the dependent variable influence current ones. Second, it can account for any form of unobserved heterogeneity. Third, some regressors may be endogenous or predetermined. As we will see later, the common GMM estimators also put some restrictions on the situations where they can be applied - e.g., they were designed for panels where the time frame (T) is short compared to the number of cross-sectional units (N). Moreover, the idiosyncratic error needs to be uncorrelated across individuals. However, many applications require the analysis of panel data with rather few individuals (N). For instance, Filippini and Alberini (2011) analysed dynamic residential electricity demand in the U.S. using aggregate data from N = 48 states over a period of T = 12 years. In fact, there are many applications where the number of panel units is not much larger than the length of the observed period, and therefore the properties of the estimators should also be analysed for designs that are typical for applied research.

---

[27] See Blundell and Bond (1998), Baltagi (2005) or Roodman (2007). The initial conditions impose an additional restriction on the data generating process (DGP). If satisfied, the Arellano Bond procedure where the first differenced equation is instrumented with lagged levels is augmented with an equation in levels that is instrumented with lagged differences. This might have important consequences when the coefficient of the lagged dependent variable is close to one, since then the instruments in levels become less and less informative for the equation in differences. Simply speaking, the initial conditions of the DGP should ensure that the instruments in difference control for the individual effects present in the equation in levels. Accordingly, the differenced instruments can be considered valid if the fixed effects are uncorrelated with the first-differenced dependent variable of the second time period (in period 1, the first-differenced dependent variable is not observable).

So far, a variety of Monte Carlo experiments have been conducted to compare the performance of the dynamic estimators. For instance, Arellano and Bond (1991) compare the performance of their first-difference GMM estimator (which will be discussed in the next section) with the OLS estimator, a fixed effects estimator and with the Anderson-Hsiao estimator. Bias and variance of the parameters of interest are best for first-difference GMM in almost all settings. These results were also confirmed later by Ahn and Schmidt (1995). To increase the efficiency of the first-difference GMM estimator, Blundell and Bond (1998) encouraged the use of additional moment conditions (as will be discussed later) and therefore call their estimator system GMM. Monte Carlo experiments have shown that this estimator indeed has a significantly lower variance over a wide range of parameter choices - at least for the purely autoregressive process which they analysed; see Blundell and Bond (1998).

The main goal of the present paper is to evaluate the impact of low within variation of exogenous regressors on the parameter estimates of the first-difference and system GMM estimators.

2 Econometric Description of FD-GMM and SYS-GMM

The general and well-known dynamic panel data model is characterised by the presence of a time-lagged dependent variable and exogenous[28] or at least predetermined[29] regressors. In the forthcoming Monte-Carlo simulation, we will use a dynamic panel data model with one explanatory exogenous regressor, namely:

$$\begin{aligned} y_{it} &= \gamma \cdot y_{i,t-1} + \beta \cdot x_{it} + u_{it} \\ u_{it} &= \mu_i + \varepsilon_{it} \end{aligned} \tag{IV.1}$$

where γ is a scalar to which we will simply refer as the 'temporal lag', β is a scalar which represents the coefficient of the regressor x_{it} ,he residuals u_{it} , which are the sum of individual specific effects, μ_i, and an idiosyncratic error term, ε_{it}. Further, it is assumed that the individual specific effects are independently and identically distributed with variance σ^2_μ and σ^2_ε respectively. According to Baltagi (2005), there are two sources of persistence over time: the presence of a lagged dependent variable and the individual specific effects, being the unobserved heterogeneity among the panel units i .

Since y_{it} depends on the individual specific μ_i, $y_{i,t-1}$ also depends on μ_i, and therefore represents a regressor which is correlated with the residuals u_{it} and therefore is endogenous. One possibility to eliminate the individual specific effects is the within transformation. The dynamic panel data mode described by equation (IV.1) is then transformed to

---

28 Exogenous means that $\mathrm{E}(x_{it}\varepsilon_{it}) = 0$. In the literature, this often is referred to as 'strictly' exogenous.

29 Predetermined means that $\mathrm{E}(x_{it}\varepsilon_{is}) = 0,\ s > t$ which implies that the regressors are uncorrelated with future realisations of the residuals but not with contemporaneous and past ones. Allowing for predetermined regressors in the regression equation enables shocks to the dependent variable (captured by the residuals) to be correlated with future realisations of those regressors. Drukker (2008) gives two examples: If one wants to estimate a model explaining crime with police force, then police force should be treated as a predetermined regressor. The reason is that (e.g. positive) shocks to the crime rate will affect future levels of police force. Similarly, if we regressed per capita income on a dummy variable indicating whether the person in question is married, then the dummy variable "marriage" is predetermined since (e.g. negative) shocks to income may cause divorces in the future.

$$
\begin{aligned}
&y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it} = \gamma \cdot \left(y_{i,t-1} - \frac{1}{T-1}\sum_{t=2}^{T} y_{i,t-1} \right) + \beta \cdot \left(x_{it} - \frac{1}{T}\sum_{t=1}^{T} x_{it} \right) + \left(u_{it} - \frac{1}{T}\sum_{t=1}^{T} u_{it} \right) \\
&u_{it} - \frac{1}{T}\sum_{t=1}^{T} u_{it} = \underbrace{\left(\mu_i - \frac{1}{T}\sum_{t=1}^{T} \mu_i \right)}_{0} + \left(\varepsilon_{it} - \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it} \right)
\end{aligned} \tag{IV.2}
$$

It can easily be seen that the μ_i are eliminated. However, by construction, $y_{i,t-1}$ is still correlated with the time average of the residuals since this also contains the summand $\varepsilon_{i,t-1}$. However, the magnitude of this correlation decreases with T and therefore, the consistency of an OLS estimator for equation (IV.2) will improve as T increases. Accordingly, the within estimator is biased, but the bias can be quantified and therefore be corrected, which is thoroughly discussed in the study by Kiviet (1995). This estimator is known as the LSDVc estimator (bias-corrected least squares dummy variable estimator). Moreover, from the previously stated facts, it is clear that the GLS transformation of equation (IV.1) does not improve the situation, since the data is de-meaned by the time averages weighted with θ. Since θ measures the fraction of variance due to μ_i in the overall variance of the residuals u_{it}, we have $0 < \theta < 1$, and therefore the correlation of the lagged de-meaned dependent variable is only weakened but not eliminated.

Another technique to eliminate the individual specific effects is to take first differences (FD). The disadvantage of this technique, clearly, is the loss in observations. The advantage, on the other hand, is that the first-differenced lagged dependent variable is uncorrelated with the first-differenced residuals under certain conditions.

2.1 FD-GMM

Anderson and Hsiao (1981) proposed taking first differences of equation (IV.1). Then, the equation is transformed to

$$
\begin{aligned}
&y_{it} - y_{i,t-1} = \gamma \cdot \left(y_{i,t-1} - y_{i,t-2} \right) + \beta \cdot \left(x_{it} - x_{i,t-1} \right) + \left(u_{it} - u_{i,t-1} \right) \\
&\left(u_{it} - u_{i,t-1} \right) = \left(\mu_i - \mu_i \right) + \left(\varepsilon_{it} - \varepsilon_{i,t-1} \right)
\end{aligned} \tag{IV.3}
$$

or more simply to

$$
\Delta y_{it} = \gamma \cdot \Delta y_{i,t-1} + \beta \cdot \Delta x_{it} + \Delta \varepsilon_{it} \tag{IV.4}
$$

Of course, $\Delta y_{i,t-1}$ is still correlated with $\Delta\varepsilon_{it}$, since $y_{i,t-1}$ is a function of $\varepsilon_{i,t-1}$ in $\Delta\varepsilon_{it}$. Anderson and Hsiao propose to instrument $\Delta y_{i,t-1}$ with past first differences, namely $\Delta y_{i,t-2}, \Delta y_{i,t-3}, \ldots$ or simply with past levels, $y_{i,t-2}, y_{i,t-3} \ldots$. Earlier Monte Carlo simulations, e.g. by Ahn and Schmidt (1995), show that the instruments in levels perform much better than the first-differenced instruments in terms of efficiency of the parameter estimates. However, the estimator does not exploit all available moment conditions and conditionally performs badly in terms of efficiency and bias over a significant parameter range. Of course, the validity of those instruments depends heavily on the assumption that the first-differenced residuals $\Delta\varepsilon_{it}$ are not serially correlated, which in turn implies that the levels of ε_{it} must not be second-order serially correlated.

If the x_{it} are exogenous, then, according to Arellano and Bond (1991), x_{it}, $t = 1..\mathrm{T}$ are valid instruments for the first-differenced lagged dependent variable in equation (IV.4) in addition to the past levels of $y_{i,t-2}$. If, on the other hand, the x_{it} are predetermined, then only $x_{i1}...x_{i,t-1}$ are valid instruments for the first-differenced lagged dependent variable at period t in equation (IV.4). According to Baltagi (2005), a combination of both predetermined and exogenous variables is more likely to occur in empirical studies than only one of those extreme cases. However, the researcher can adapt the matrix of appropriate instruments according to the previous comments. Moreover, some of the x_{it} may be uncorrelated with μ_i. This case allows for additional moment restrictions for the equation in levels (IV.1), namely that $\mathrm{E}(x_{it} \cdot u_{it}) = 0$ for $t = 2...\mathrm{T}$ and that $\mathrm{E}(u_{i2} \cdot x_{i1}) = 0$ for $t = 1$; see also Arellano and Bond (1991). Further, other moment restrictions from the equation in levels are collinear with the moment conditions for the first-differenced equation and thus redundant. In sum, we retain the definition as presented in Baltagi (2005) and define the matrix of instruments, $\mathbf{W}_{\text{ex,i}}$ for panel unit i, for the case of strictly exogenous regressors for the equation in levels:

$$\mathbf{W}_{\text{ex,i}} = \begin{bmatrix} [y_{i1}, x_{i1}...x_{iT}] & & 0 \\ & [y_{i1}, y_{i2}, x_{i1}...x_{iT}] & \\ & & ... \\ 0 & & [y_{i1}, y_{i2}...y_{i,T-2}, x_{i1}...x_{iT}] \end{bmatrix} \tag{IV.5}$$

In the case of predetermined regressors x_{it}, the instrument matrix becomes

$$\mathbf{W}_{\text{pred,i}} = \begin{bmatrix} [y_{i1}, x_{i1}, x_{i2}] & & 0 \\ & [y_{i1}, y_{i2}, x_{i1}, x_{i2}, x_{i3}] & \\ & & ... \\ 0 & & [y_{i1}, y_{i2}...y_{i,T-2}, x_{i1}...x_{i,T-1}] \end{bmatrix} \tag{IV.6}$$

In addition, if the x_{it} are uncorrelated with the individual effects μ_i, $x_{i1}...x_{iT}$ become instruments available for the equation in levels and can be as a diagonal sub-matrix to $\mathbf{W}_{\text{pred,i}}$ or $\mathbf{W}_{\text{ex,i}}$, which results in an augmented instrument matrix, $\mathbf{W}^{+}$.

Given the index notation of the instrument matrix, we rewrite equation (IV.4) to

$$\mathbf{\Delta y}_i = \gamma \cdot \mathbf{\Delta y}_{i,-1} + \beta \cdot \mathbf{\Delta x}_i + \mathbf{\Delta \varepsilon}_i \tag{IV.7}$$

or more general to

$$\mathbf{\Delta y} = \gamma \cdot \mathbf{\Delta y}_{-1} + \beta \cdot \mathbf{\Delta x} + \mathbf{\Delta \varepsilon} \tag{IV.8}$$

Departing from equation (IV.7), the form of the variance-covariance matrix of the $\mathbf{\Delta\varepsilon}_i = [\varepsilon_{i3} - \varepsilon_{i2}, ..., \varepsilon_{iT} - \varepsilon_{i,T-1}]$ is known and results in

$$\mathrm{E}(\mathbf{\Delta\varepsilon}_i \cdot \mathbf{\Delta\varepsilon}_i') = \sigma_\varepsilon^2 \cdot \begin{bmatrix} 2 & -1 & . & 0 \\ -1 & 2 & -1 & . \\ . & -1 & 2 & -1 \\ 0 & . & -1 & 2 \end{bmatrix} = \sigma_\varepsilon^2 \cdot \mathbf{G}_{(\mathrm{T}\text{-}2)\times(\mathrm{T}\text{-}2)} \Leftrightarrow \mathrm{E}(\mathbf{\Delta\varepsilon} \cdot \mathbf{\Delta\varepsilon}') = \sigma_\varepsilon^2 \cdot [\mathbf{I}_\mathrm{N} \otimes \mathbf{G}] \qquad \text{(IV.9)}$$

given that the ε_{it} are homoscedastic and serially uncorrelated. This 'known' form of the variance-covariance matrix allows the application of GLS by multiplying the first-differenced equation (IV.8) with the transposed matrix of instruments. Generally, the GLS estimator is given by

$$\begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \end{pmatrix} = \left([\mathbf{\Delta y}_{-1}, \mathbf{\Delta x}]' \cdot \mathbf{W}\hat{\mathbf{V}}^{-1}\mathbf{W}' \cdot [\mathbf{\Delta y}_{-1}, \mathbf{\Delta x}] \right)^{-1} \cdot \left([\mathbf{\Delta y}_{-1}, \mathbf{\Delta x}]' \cdot \mathbf{W}\hat{\mathbf{V}}^{-1}\mathbf{W}' \cdot \mathbf{\Delta y}_{-1} \right) \qquad \text{(IV.10)}$$

where $\hat{\mathbf{V}}$ is the variance-covariance matrix of the GLS-transformed first-differenced equation. In a first step, $\hat{\mathbf{V}}$ can be set to $\hat{\mathbf{V}} = \mathbf{W}'(\mathbf{I}_\mathrm{N} \otimes \mathbf{G})\mathbf{W}$ to obtain consistent but not necessarily efficient parameter estimates according to (IV.10) which we refer to as the one-step estimator. In a second step, the one-step estimator can be used to obtain consistent estimates for the residuals ε_{it}, which in turn allows us to obtain a consistent estimate of the variance-covariance matrix of the residuals as described by equation (IV.9). It can easily be shown that the variance-covariance matrix of the parameter estimates of this "two-step" estimator is given by the first term in equation (IV.10).

2.2 SYS-GMM

Blundell and Bond (1998) considered exploiting additional moment conditions as is the case when applying FD-GMM. They considered an autoregressive panel data model (with no exogenous regressors). Accordingly, the data generating process in focus is

$$\begin{aligned} y_{it} &= \gamma \cdot y_{i,t-1} + u_{it} \\ u_{it} &= \mu_i + \varepsilon_{it} \end{aligned} \qquad \text{(IV.11)}$$

Blundell and Bond (1998) initially also restricted the data to the case where $\mathrm{T=3}$, and therefore, there is only one moment condition according to the elaborations by Arellano and Bond (1991), namely that $\mathrm{E}(y_{i3} \cdot \Delta\varepsilon_{i3}) = \mathrm{E}(y_{i3} \cdot (\varepsilon_{i3} - \varepsilon_{i2})) = 0$. With one moment condition, the coefficient in question, γ, is just identified. We already know from section 2.1 that the FD-GMM estimator instruments the differenced equation with lagged levels of the variables. Doing so, we can evaluate equation (IV.11) at time $\mathrm{t=2}$ and subtract y_{i1} on both sides to obtain

$$\Delta y_{i2} = (\gamma - 1) \cdot y_{i1} + \mu_i + \varepsilon_{i2} \qquad \text{(IV.12)}$$

Accordingly, the first step estimator for this case would be to regress y_{i1} on Δy_{i2}. But there are two problems with that: The first is that for high values of γ (close to one), y_{i1} is only weakly correlated with Δy_{i2} and therefore, y_{i1} is a weak instrument. Second, the least-squares estimator of the equation above will generally be upwardly biased since it can be expected that $\mathrm{E}\left(y_{i1} \cdot \mu_i\right) > 0$ [30]. The situation becomes worse with increasing values of γ and with an increasing fraction of variance due to the unit effects.

Similarly to the moment conditions proposed by Arellano and Bond (1991), Blundell and Bond (1998) suggest using additional moment conditions to overcome the problems of weak instruments, which affect efficiency, and bias. The basic idea is, in addition to instrument the differenced equation with instruments in levels, to instrument the equation in levels with differenced instruments. Instead of transforming the equation to eliminate the unit effects, they difference the instruments to render them uncorrelated with the unit effects. Suppose there is a valid instrument z_{it} whose changes over time are uncorrelated with the fixed effects such that $\mathrm{E}\left(\Delta z_{it} \cdot \mu_i\right) = 0$. If this is true, then, by construction, $\Delta z_{i,t-1}$ is a valid instrument for the equation in levels, since it will no longer be correlated with the overall error term u_{it}. Accordingly, it is suggested that $y_{i,t-1}$ be additionally instrumented with $\Delta y_{i,t-1}$. The problem here is that, according to equation (IV.12), the instrument contains the unit effects, which makes it counterintuitive that $\mathrm{E}\left(\Delta y_{i,t-1} \cdot u_{it}\right) = 0$ holds. The moment restriction can hold, but only if $\mathrm{E}\left(y_{it} \mid \mu_i\right) = \mathrm{E}\left(y_{i,t-1} \mid \mu_i\right)$. Given that $y_{it} = \gamma \cdot y_{i,t-1} + \mu_i + \varepsilon_{it}$, this means that y_{it} converges to $\mu_i / (1-\gamma)$, which is often referred to as the "initial conditions".

In order to exploit the new moment conditions, Blundell and Bond stack the instrument matrix as described by equation (IV.5) (e.g. without exogenous regressors x_{it}) with the differenced instruments as described previously. Accordingly, the first-differenced equation is augmented with the equation in levels and the instrument matrix $\mathbf{W}$ used for FD-GMM becomes

$$\mathbf{W}_{\mathrm{SYS}} = \begin{bmatrix} \mathbf{W}_{\mathrm{L}} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{\mathrm{D}} \end{bmatrix}$$

where the index L refers to instruments in levels as described in (IV.5) and the index D refers to first-differenced instruments. Accordingly, the first-differenced equation is instrumented with levels and the equation in levels is instrumented with first differences. As described in Roodman (2007), using all available lags for the instrument matrix $\mathbf{W}_{\mathrm{D}}$ renders some of the moment conditions redundant. Accordingly, the instrument matrix is

---

[30] Without loss of generality, we can assume that $\mathrm{E}\left(y_{i1}\mu_i\right) = \mathrm{E}\left(y_{i2}\mu_i\right)$. Evaluating equation (IV.11) at time t=2, one obtains $\mathrm{E}\left(y_{i1}\mu_i\right) = (1-\gamma)\sigma^2 > 0$. Working from this, it can be shown that $\widehat{\gamma - 1}_{\mathrm{OLS}} = \left(y_1{}' \cdot y_1\right)^{-1} \cdot y_1 \cdot \Delta y_2 = \gamma - 1 + \left(y_1{}' \cdot y_1\right)^{-1} \cdot y_1{}' \cdot \mu$. Taking expectation on both sides leads to the insight that the estimated parameter $\gamma - 1$ is upwardly biased. By further evaluating the previously obtained expression, one also can see that the bias grows as the fraction of variance due to μ_i increases.

$$\mathbf{W}_{\mathrm{D,i}} = \begin{bmatrix} 0 & 0 & 0 & \dots \\ \Delta y_{i2} & 0 & 0 & \dots \\ 0 & \Delta y_{i3} & 0 & \dots \\ \dots & \dots & \Delta y_{i4} & \dots \end{bmatrix} \qquad \text{(IV.13)}$$

In contrast to FD-GMM, time-invariant regressors can also be estimated with SYS-GMM. According to Roodman (2007), this should not affect coefficient estimates of time-varying variables, since first, time-invariant variables cancel out in the differenced equation and second, all variables used as instruments in levels are assumed to be orthogonal to the unit effects, and therefore also to all time-invariant variables. Most of all, statistical software packages such as STATA® handle system GMM as ONE equation. The consequences of doing so are, at least as shown by Blundell and Bond (1998) for an autoregressive process, dramatic efficiency gains in the estimation of γ.

2.3 Monte Carlo Simulations

Blundell and Bond (1998) showed that several shortcomings of the FD-GMM estimator can at least be lessened when one is willing to accept a stationarity restriction on the data generation process (DGP). Doing so enables the usage of SYS-GMM, which instruments the equation in levels with first-differenced instruments in addition to the FD-GMM where the equation in differences is instrumented with lagged levels. In a Monte Carlo analysis, they found dramatic efficiency gains and also lower finite sample bias of the coefficient of the lagged dependent variable. Given a variance ratio between the unit effects and the idiosyncratic error equal to 1, $\mathrm{N}=100$ and $\mathrm{T}=4$, the FD-GMM estimator for γ has a 54% higher variance than the respective SYS-GMM estimator if the true parameter is $\gamma = 0$. With increasing values for γ, the efficiency gain becomes increasingly dramatic and for $\gamma = 0.8$, the variance of the coefficient estimate provided by FD-GMM exceeds that from SYS-GMM by a factor of almost 26. This showed that the SYS-GMM estimator indeed handles the weak instruments problem occurring in FD-GMM well. Generally, the efficiency gains are highest for small numbers of observations within units (T) The reason is that, with increasing T, more instruments become available, as can be seen from construction of the instrument matrix described by equation (IV.5). The number of instruments increases quadratically with T. According to Baltagi (2005), there might be desirable asymptotical efficiency gains when exploiting all available moment conditions, but this might be infeasible or impractical in applied research. Ziliak (1997) shows that a tradeoff between bias and efficiency exists when all moment conditions are exploited[31].

[31] Of course, the number of moment conditions not only increases quadratically with the time dimension but also proportionally with the number of exogenous regressors in the model.

In the present Monte Carlo simulation, we intend to extend the analysis provided by Blundell and Bond (1998) and incorporate exogenous regressors in the data-generating process as described by equation (IV.11). Moreover, we are particularly interested in describing the estimation qualities in terms of efficiency (variance) and bias of the FD-GMM and the SYS-GMM estimators with respect to the within variation of this exogenous variable. The reason is that, often in applied (e.g. energy) economic studies, the explanatory variables used in such models typically exhibit a small ratio of within to between variation. For instance, the within to between variance ratio of per capita income of the 547 Swiss municipalities considered in the empirical application of the second part of this dissertation is only 0.08. Similarly, this ratio equals 0.006 and 0.20 for the stock of gasoline and diesel powered cars respectively, 0.016 for the number of commuters and only 0.0003 for the municipalities' populations. Other examples can easily be found. For instance, Alberini and Filippini (2011) analysed the price elasticity of residential electricity demand in the US. Typically, the price variable exhibits a very small within variation. The problem with the estimation of such variable with small within variation is that, if one wants to exploit the advantages offered by panel data estimation techniques (such as efficiency gain or accounting for unobserved heterogeneity), the data needs to be manipulated in different ways. A fixed effects model, for instance, requires within transformation of the data (in order to eliminate the unobserved unit effects), rendering the variable in question with zero between variation. Other techniques, such as FD-GMM and SYS-GMM, require first-differencing of the data, for use either as explanatory variables or as instruments. This also significantly diminishes the between variation of the variable in question. As a consequence, for instance, fixed effects estimates or FD-GMM estimates of coefficients of variables with low within variation exhibit a large variance (and moreover, the estimation of coefficients of time invariant variables is not possible). On the other hand, SYS-GMM augments the differenced equation with equation in levels but in turn instruments with first differences again. Therefore, it is a priori not clear what the consequences of the incorporation of variables with low within variation in dynamic panel data models are.

3 The Monte-Carlo Experiment

3.1 Experimental Design

Monte Carlo experiments are primarily useful for detecting stochastic properties of parameters of interest if the properties in question either cannot be analytically derived or to prove the correctness of the theoretical properties of the parameters. For instance, suppose we want to analyse the properties of coefficients estimated from an ordinary least squares model with homoscedastic disturbances, N observations and K (exogenous) regressors. The coefficient matrix then is obtained by calculating $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. The variance of those estimates can be analytically obtained and calculated as $\mathrm{Var}(\mathbf{b}) = 1/(\mathrm{N}-\mathrm{K}) \cdot (\mathbf{y}-\mathbf{Xb})'(\mathbf{y}-\mathbf{Xb}) \cdot (\mathbf{X}'\mathbf{X})^{-1}$ and therefore also depends on the number of observations N and number of regressors K. If we wanted to prove the correctness of this identity, we could run a Monte Carlo experiment where we vary N over e.g. the range 10, 20, 50, 100 and K over 1,2,5. In each experiment, we randomly generate the regressors and the dependent variable and then regress y on X. Each experiment is replicated a number of times, typically 1,000, 5,000 or 10,000 times. The reason is that, according to the central limit theorem, as replications go towards infinity, the replicates of the coefficients converge to the "true" distribution of the coefficients. Therefore with many replicates, the mean of the replicated coefficient should converge to the "true" mean and the variance of the replicated coefficients should converge to the "true" variance of the coefficients. Proceeding in this manner, the simulated means and variances of the coefficients can be compared to the analytical values.

As previously discussed, we want to analyse the impact of the level of the within variation of an exogenous regressor on the parameter estimates in the context of dynamic panel data. For this purpose, we perform a variety of Monte Carlo experiments and use a data generation process (DGP) as stated in equation (IV.1), namely

$$y_{it} = \gamma \cdot y_{i,t-1} + \beta \cdot x_{it} + u_{it}$$
$$u_{it} = \mu_i + \varepsilon_{it}$$

We are primarily interested in studying one effect, the impact of low within variation of exogenous variables on the parameter estimates $\hat{\gamma}$ and $\hat{\beta}$. For this reason, we vary the within standard deviation $s_{w,x_{it}}$ of x_{it} over the range 0.1, 0.2, 0.5, 1, 2, 5, 10 and keep the between variation $s_{b,x_{it}}$ of x_{it} fixed at 1. We vary γ over 0, 0.2, 0.4, 0.6, 0.8 and keep the coefficient β fixed at 1. Finally, we also vary the number of groups, N, over 50, 250, 500 and the number of observations per group, T, over 5, 10, 20. We replicate each experiment 1,000 times and estimate the econometric model as stated by equation (IV.1) first with FD-GMM and second with SYS-GMM. In a first setting, we also varied other parameters such as the correlation between the unit effects and exogenous regressors, the overall variance of the residuals, and the fraction of variance due to the unit effects. However, the variation of those parameters did not alter the main insights of the present simulation, and therefore we report the results for a simulation as reported in Table IV-1 below. From the previous discussion, we are aware that using the full set of instruments when more observations per unit become available might worsen the estimation quality of GMM estimators. Therefore, we restrict the number of lags to a depth of two but report results with respect to the full set of instruments in order to compare FD-GMM and SYS-GMM with respect to this issue as well in the Appendix.

Parameter	Description	Range of Variation	Number of Outcomes
N	Cross-sectional units	N= {50, 250, 500}	3
T	Time	T= {5, 10, 20}	3
γ	Coefficient of lagged dependent variable	γ= {0.0, 0.2, 0.4, 0.6, 0.8}	5
β_{xit}	Coefficient of exogenous variable	β_{xit}= {1}	1
sw_{xit}	Within standard deviation of x_{it}, between standard deviation fixed equal to 1	sw_{xit}= {0.1, 0.2, 0.5, 1, 2, 5, 10}	7
$\rho_{xit,ui}$	Correlation between fixed effects and x_{it}	$\rho_{xit,ui}$= {0.5}	1
VAR	Overall variance of (μ_i+ε_{it})	VAR= {1}	1
θ	Fraction of variance due to μ_i	θ= {0.25}	1
STEP	1-step or 2-step estimator	STEP= {2}	1
MAXLAG	Depth of Lags for Instruments	MAXLAG= {2}	1
MODEL	FD-GMM or SYS-GMM	MODEL= {FD, SYS}	2
TOT. Designs	3·3·5·1·7·1·1·1·1·1·2		630
Replications	1,000 per experiment		630,000

Table IV-1: Manipulated and constant parameters in the Monte Carlo experiments

As elaborated by Clark and Linzer (2012), empirical modeling decisions involve a trade-off between bias and variance. For instance, the previous subsection has shown, first-differencing the data largely eliminates the between variation of a variable, and if the within variation is low, then coefficient estimates obtained by FD-GMM are on one hand not biased but will exhibit a large variance, which renders post-estimation inference on those coefficients obsolete. On the other hand, the willingness to accept a (potentially) small bias while gaining efficiency might improve the situation as a whole. Although rarely discussed, the tradeoff between variance and bias of estimated coefficients is very important in economic application. A demonstrably unbiased coefficient estimate with a high variance can be quite far from the "true" parameter - so far that even the expected sign is not the right one and results are rendered counter-intuitive. However, the data which is subject of the econometric analysis in question are a single draw, and therefore knowledge about the unbiasedness of the coefficient is useless, since the experiment will not be repeated. Coefficients with a smaller variance and a potentially small bias might therefore sometimes be preferable.

As it is frequently done in Monte Carlo experiments, three common measures will be used to assess the estimation quality of FD-GMM and SYS-GMM: the bias between the average of the replicated coefficients and the true coefficient, the standard deviation of the replicated coefficients, and the root mean square error of the replicated coefficients; see e.g. Arellano and Bond (1991) or Blundell and Bond (1998). The purpose of reporting the root mean square error (RMSE) of the replications is the following. Suppose a replication of coefficient estimates $\hat{\beta}$ with mean $\overline{\beta}$ and true value β. Then, we define the average bias in N replications as

$$\text{bias} \quad = \quad \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\beta}_i - \beta\right) \tag{IV.14}$$

The standard deviation within these replications equals

$$\text{SD} \quad = \quad \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\beta}_i - \overline{\beta}\right)^2} \tag{IV.15}$$

Finally, the RMSE of these replications equals

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\beta}_i - \beta\right)^2} \tag{IV.16}$$

For large values of N (the typical number of replications in Monte Carlo experiments lies between 500 to 5'000), equation (IV.16) can be rearranged to

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\beta}_i - \beta\right)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\beta}_i - \overline{\beta}\right)^2 + \left(\overline{\beta} - \beta\right)^2} \approx \sqrt{\text{SD}^2 + \text{bias}^2} \tag{IV.17}$$

Accordingly, the RMSE[32] is a convenient measure to assess the estimation quality of a replicated coefficient since it incorporates both the bias and the variance of estimator in question and therefore contributes to the previously stated comments concerning the trade-off between variance and bias. Therefore, following Egger et al. (2005b) or Clark and Linzer (2012), the RMSE is our preferred measure to assess the accuracy of the estimators.

---

[32] Sometimes, the RMSE is defined as $\sqrt{\text{bias}^2 + (\text{IQR}/1.35)^2}$ which yields a numerically identical statistic to the one defined by equation (IV.17) if the underlying distribution is a normal distribution (normally distributed variables have a 1.35 times higher inter quartile range than standard deviation).

3.2 Results

In this subsection, we report the results obtained from the Monte Carlo experiment. In order to restrict the number of tables reporting the results, we decided to present only the tables showing the RMSE of the replicated parameters in this main part of the paper. Results with respect to bias and standard deviation can be found in the Appendix. We start with the presentation of the estimation accuracy (in terms of RMSE) of γ, then with the results for the exogenous short-run effect β and finally with the combined long-run effect $\beta^* = \beta / (1-\gamma)$. The reason for doing so is that in applied economic research using panel data, the researcher typically is interested in the short-run effects (most often elasticities) and the long-run effects. For instance, several studies in the energy economic literature estimate dynamic panel data models for the demand of electricity to obtain short- and long-run price elasticities, see e.g. Alberini and Filippini (2011).

3.2.1 RMSE of γ

We begin the evaluation of the Monte Carlo experiment with the coefficient of the lagged dependent variable, γ. The focus lies on the estimation accuracy of the parameter with respect to the variation of the true parameter and the within variation of the exogenous variable. As already explained, we use three measures to assess the estimation quality: the bias, the standard deviation, and the root mean squared error within the replications. Concerning bias, standard deviation and RMSE of γ, we tabulate results as a ratio to the true value of the parameter.

Table IV-2 shows the RMSE of the estimates of γ relative to the true value of γ when the number of instruments is restricted to a depth of two[33]. The fields of the table are colored red if the relative RMSE of the first-difference GMM estimator (FD estimator) is below the relative RMSE of the system GMM estimator (SYS estimator) and green colored if the opposite is the case. Moreover, the values are tabulated in grey font if the difference reported by the two estimators is not significant (at 5% level of significance). First, the RMSE of the estimates obtained by the SYS-GMM estimator is significantly lower than the RMSE obtained by the FD-GMM estimator for almost all parameter combinations unless $\gamma = 0.8$ and T = 20. For N = 50 and T = 5, the RMSE obtained by SYS-GMM is 33% lower than those obtained by FD-GMM. On average, the RMSE of γ is 7% smaller for the SYS-GMM estimates. The efficiency gain is highest for N = 50 and T = 5, a situation which is often encountered in empirical research. For both estimators, the RMSE of γ decreases with increasing within variation of the exogenous regressor and strongly increases with the simulated values of γ.

Summing up these results with respect to the estimation of γ, we conclude that SYS-GMM generally has to be preferred.

[33] In a preliminary analysis we also made use of all instruments. The results of this analysis are reported in several tables included in the Appendix: Table IV-14 shows the RMSE of γ when the full set of instruments is used. Table IV-12 and Table IV-13 show the relative bias and relative standard deviation of the estimates when the full set of instruments is used.

Since the RMSE is a combined measure of bias and variance, we also want shortly to describe the results of the Monte Carlo analysis using the bias and standard deviation as a measure of performance (see Table IV-6 and Table IV-7 in the Appendix). Most of all, if the within variation of the exogenous regressor is small relative to the variation in the dependent variable (say, below one) and if N>>T, a situation which is often encountered in applications, SYS-GMM is superior to FD-GMM with respect to both bias and efficiency. For higher values of the within variation, and a smaller ratio of N/T, this superiority is no more that substantial and for high values of γ , FD-GMM performs better in term of bias and efficiency, but the difference to SYS-GMM is not substantial on average.

					T														
					5					10					20				
					γ					γ					γ				
N					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	50	sw	0.1	FD	0.0999	0.8795	0.7630	0.6370	0.4375	0.0677	0.9360	0.9240	0.7905	0.1438	0.0445	0.9770	1.0085	1.0535	0.1238
				SYS	0.0802	0.8165	0.6488	0.4302	0.2083	0.0605	0.9040	0.8315	0.6017	0.1415	0.0405	0.9325	0.9200	0.8150	0.1913
			0.2	FD	0.1006	0.8925	0.7280	0.5865	0.3823	0.0672	0.8965	0.8653	0.7907	0.1365	0.0441	0.9405	0.9690	1.0345	0.1219
				SYS	0.0812	0.7700	0.5648	0.3677	0.2013	0.0600	0.8445	0.7590	0.5680	0.1714	0.0403	0.8550	0.8428	0.7445	0.2141
			0.5	FD	0.0965	0.7930	0.6238	0.4418	0.3244	0.0574	0.6785	0.6260	0.6227	0.1334	0.0391	0.7670	0.7913	0.8727	0.1008
				SYS	0.0738	0.5965	0.4193	0.2898	0.1761	0.0538	0.6260	0.5618	0.4688	0.2109	0.0361	0.6830	0.6923	0.6420	0.2380
			1	FD	0.0822	0.6080	0.4080	0.2667	0.2788	0.0423	0.4030	0.3375	0.3090	0.1481	0.0280	0.4610	0.4695	0.5162	0.0801
				SYS	0.0589	0.4000	0.2733	0.2003	0.1484	0.0398	0.3655	0.2980	0.2808	0.2074	0.0259	0.4170	0.4308	0.4508	0.2376
			2	FD	0.0547	0.3460	0.2018	0.1653	0.1906	0.0247	0.1810	0.1198	0.0965	0.0685	0.0180	0.1880	0.1795	0.1780	0.1203
				SYS	0.0414	0.2400	0.1525	0.1090	0.0986	0.0232	0.1540	0.0958	0.0830	0.0984	0.0164	0.1655	0.1595	0.1755	0.1670
			5	FD	0.0228	0.1420	0.0855	0.0743	0.0779	0.0108	0.0535	0.0290	0.0185	0.0140	0.0070	0.0470	0.0358	0.0293	0.0220
				SYS	0.0193	0.1125	0.0620	0.0423	0.0356	0.0105	0.0510	0.0255	0.0155	0.0179	0.0066	0.0420	0.0318	0.0303	0.0344
			10	FD	0.0110	0.0635	0.0390	0.0335	0.0313	0.0051	0.0235	0.0115	0.0068	0.0056	0.0035	0.0200	0.0115	0.0085	0.0054
				SYS	0.0101	0.0550	0.0313	0.0223	0.0168	0.0051	0.0240	0.0110	0.0065	0.0055	0.0033	0.0180	0.0110	0.0090	0.0074
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	sw	0.1	FD	0.0440	0.6545	0.5938	0.4515	0.1858	0.0294	0.8280	0.8840	0.7255	0.1889	0.0198	0.9290	0.9803	1.0063	0.1870
				SYS	0.0334	0.6615	0.5473	0.3602	0.1703	0.0260	0.8360	0.8025	0.5860	0.1033	0.0188	0.8960	0.8945	0.8458	0.2004
			0.2	FD	0.0422	0.6470	0.5558	0.3960	0.1730	0.0277	0.7825	0.8253	0.7400	0.1870	0.0186	0.8880	0.9450	0.9932	0.1893
				SYS	0.0327	0.6325	0.4955	0.3223	0.1624	0.0243	0.7835	0.7350	0.5627	0.1400	0.0182	0.8490	0.8515	0.8128	0.2336
			0.5	FD	0.0413	0.5490	0.4458	0.2743	0.2028	0.0242	0.5940	0.5910	0.5875	0.1618	0.0169	0.7125	0.7610	0.8348	0.1749
				SYS	0.0311	0.4895	0.3715	0.2540	0.1520	0.0223	0.5925	0.5413	0.4580	0.1930	0.0161	0.6865	0.7008	0.6930	0.2620
			1	FD	0.0353	0.3935	0.2745	0.1345	0.2073	0.0182	0.3360	0.3055	0.2813	0.0546	0.0126	0.4265	0.4515	0.4908	0.0771
				SYS	0.0254	0.3385	0.2455	0.1803	0.1268	0.0169	0.3220	0.2713	0.2477	0.1799	0.0124	0.4120	0.4238	0.4455	0.2410
			2	FD	0.0255	0.1840	0.0985	0.0812	0.1583	0.0104	0.1305	0.1070	0.0852	0.0500	0.0073	0.1625	0.1675	0.1667	0.0849
				SYS	0.0172	0.1825	0.1340	0.0990	0.0849	0.0099	0.1185	0.0838	0.0678	0.0775	0.0072	0.1515	0.1465	0.1563	0.1470
			5	FD	0.0104	0.0625	0.0470	0.0573	0.0680	0.0044	0.0280	0.0190	0.0135	0.0085	0.0031	0.0310	0.0293	0.0250	0.0154
				SYS	0.0082	0.0650	0.0443	0.0330	0.0319	0.0043	0.0280	0.0153	0.0100	0.0140	0.0031	0.0300	0.0248	0.0248	0.0269
			10	FD	0.0049	0.0305	0.0238	0.0257	0.0266	0.0022	0.0105	0.0058	0.0038	0.0026	0.0016	0.0100	0.0080	0.0060	0.0026
				SYS	0.0044	0.0280	0.0170	0.0132	0.0106	0.0022	0.0110	0.0055	0.0033	0.0035	0.0016	0.0100	0.0073	0.0067	0.0054
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw	0.1	FD	0.0303	0.6200	0.5700	0.4377	0.1316	0.0201	0.8160	0.8628	0.7083	0.1898	0.0135	0.9240	0.9815	1.0053	0.1903
				SYS	0.0236	0.6380	0.5368	0.3612	0.1695	0.0179	0.8275	0.7900	0.5877	0.1040	0.0127	0.8970	0.8960	0.8617	0.2053
			0.2	FD	0.0296	0.5965	0.5490	0.3865	0.1466	0.0190	0.7665	0.8095	0.7148	0.1903	0.0139	0.8915	0.9498	0.9908	0.1926
				SYS	0.0236	0.6070	0.4920	0.3190	0.1633	0.0170	0.7800	0.7310	0.5673	0.1433	0.0133	0.8555	0.8573	0.8300	0.2411
			0.5	FD	0.0286	0.5195	0.4288	0.2630	0.1981	0.0162	0.5770	0.5790	0.5722	0.1649	0.0120	0.7100	0.7598	0.8357	0.1835
				SYS	0.0216	0.4805	0.3625	0.2522	0.1518	0.0152	0.5920	0.5418	0.4640	0.1979	0.0114	0.6890	0.7048	0.7058	0.2706
			1	FD	0.0255	0.3660	0.2620	0.1197	0.2043	0.0122	0.3265	0.3058	0.2802	0.0418	0.0087	0.4225	0.4458	0.4847	0.0876
				SYS	0.0178	0.3210	0.2403	0.1813	0.1280	0.0112	0.3250	0.2775	0.2497	0.1806	0.0085	0.4090	0.4158	0.4420	0.2426
			2	FD	0.0173	0.1520	0.0788	0.0657	0.1586	0.0074	0.1225	0.1063	0.0853	0.0490	0.0053	0.1595	0.1655	0.1632	0.0806
				SYS	0.0121	0.1705	0.1298	0.0957	0.0841	0.0069	0.1150	0.0850	0.0693	0.0774	0.0052	0.1500	0.1428	0.1508	0.1429
			5	FD	0.0070	0.0435	0.0388	0.0552	0.0684	0.0031	0.0225	0.0178	0.0130	0.0075	0.0021	0.0290	0.0288	0.0238	0.0144
				SYS	0.0057	0.0550	0.0418	0.0317	0.0308	0.0030	0.0240	0.0145	0.0095	0.0136	0.0021	0.0285	0.0243	0.0233	0.0256
			10	FD	0.0035	0.0220	0.0210	0.0240	0.0264	0.0015	0.0080	0.0050	0.0033	0.0019	0.0011	0.0090	0.0073	0.0055	0.0020
				SYS	0.0032	0.0215	0.0150	0.0118	0.0093	0.0015	0.0090	0.0048	0.0027	0.0030	0.0011	0.0090	0.0065	0.0060	0.0048

Table IV-2: Relative RMSE of $\hat{\gamma}$ over N, T, sw and γ

3.2.2 RMSE of β (Short-Run Effects)

Table IV-3 shows the relative RMSE of the estimated coefficients, $\hat{\beta}$ [34]. Remember that the true coefficient was fixed to one. First of all, it is obvious that SYS-GMM outperforms FD-GMM in terms of a lower RMSE of the short-run effects estimates if T is small (e.g. T = 5) and if the within variation of the exogenous regressor is below 1. In particular, the RMSE of the SYS-GMM estimates is on average 67% below the RMSE obtained by the FD-GMM estimates if N = 50, T = 5 and $s_{w,x_{it}} = 0.1$. For higher values of the within variation, e.g. 0.2, the efficiency gain still is very large, on average 45% below the RMSE of FD-GMM. For a within variation of 0.5, the efficiency gain still amounts to some 20%. When the within variation is equal to the between variation, the efficiency gain of SYS-GMM is no more substantial but still amounts to some 5% lower RMSE. As the within variation exceeds one, which corresponds to a situation where the within variation dominates the between variation, situations occur in which FD-GMM is more efficient than SYS-GMM, but the differences are generally small unless $\gamma = 0.8$. For longer periods of time T available, SYS-GMM on average still exhibits a lower RMSE than FD-GMM, but the differences are no longer as substantial and are often insignificant. On average, the RMSE of the SYS-GMM estimates are 19% below the FD-GMM estimates. We further note that the efficiency gains are most substantial for small within variation of the exogenous regressor and for situations with short time periods (e.g. T = 5) and rather few observations (e.g. N = 50, 250).

In applied empirical research, the short-run effects are of particular interest, and often the panel consists of rather few observations (e.g. N = 50, T = 5). Moreover, socio-demographic regressors are often used and typically exhibit a very low within variation compared to the between variation (e.g., consider the ratio of within to between variation of some socio-demographic regressors such as the domestic population or stock of cars in Table III-1). Alberini and Filippini (2011) analysed price elasticities of residential electricity demand in the U.S. over 13 years (N = 50, T = 12). Similarly, Blázquez et al. (2013) analysed Spanish residential electricity demand using aggregate data from N = 47 provinces over T = 9 years. Typically used exogenous regressors to explain residential electricity demand are heating and cooling degree days or per capita income, among others, which exhibit a very small within variation compared to the between variation. (For instance, for the latter study., this ratio amounts to 0.23 for the heating degree days and 0.37 for the cooling degree days respectively. For the per capita income, the ratio even is below 0.1.) For those and similar situations, SYS-GMM is to be preferred to accurately estimate the coefficients of regressors with low within variation.

Again, since the RMSE combines both bias and variance, we shortly want to refer to Table IV-8 and Table IV-9 in the Appendix to show bias and variance of the short-run effects. FD-GMM exhibits a lower bias in the short-run estimates mostly for the cases where T = 5. On average, the bias of SYS-GMM is below the one of FD-GMM, but for situations where the data is

[34] Table IV-17 shows the RMSE of the short-run effects when the full set of instruments is used.

characterised by large N and small T, the researcher in some way has to trade off bias against variance. The standard deviations of the estimates are predominantly in favour of the SYS-GMM estimates. On average, the standard deviation of SYS-GMM estimates are some 45% below those obtained by FD-GMM. This efficiency gain is predominantly present in cases where the within variation of the exogenous regressor is low. In sum, the gain in efficiency from SYS-GMM when the within variation is low and can outweigh the higher bias in terms of a lower RMSE of the short-run effects when one uses SYS-GMM.

				T														
				5					10					20				
				γ					γ					γ				
N				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	50	sw	0.1 FD	1.0343	1.0536	1.1210	1.2430	1.7313	0.5692	0.5881	0.5818	0.8313	0.9835	0.3263	0.3494	0.4049	0.5196	0.5122
			0.1 SYS	0.3659	0.4035	0.4232	0.3597	0.3920	0.2577	0.2958	0.3485	0.3019	0.7771	0.2366	0.3087	0.4134	0.4489	0.6312
			0.2 FD	0.4835	0.5286	0.5409	0.6364	0.8967	0.2865	0.2925	0.3086	0.4711	0.5055	0.1745	0.1927	0.2711	0.4101	0.2783
			0.2 SYS	0.2992	0.3243	0.3227	0.2943	0.3686	0.1996	0.2201	0.2249	0.2554	0.6435	0.1507	0.1628	0.1769	0.1848	0.4292
			0.5 FD	0.1979	0.2026	0.2278	0.3139	0.3829	0.1138	0.1222	0.1409	0.2367	0.2276	0.0677	0.1080	0.1820	0.2965	0.1060
			0.5 SYS	0.1662	0.1753	0.1838	0.2500	0.2923	0.1046	0.1061	0.1138	0.2062	0.3790	0.0641	0.0760	0.1085	0.1576	0.1724
			1 FD	0.0986	0.1026	0.1189	0.1489	0.1975	0.0571	0.0605	0.0774	0.1150	0.1550	0.0363	0.0631	0.1074	0.1750	0.0758
			1 SYS	0.0911	0.0924	0.1150	0.1457	0.1917	0.0548	0.0577	0.0730	0.1211	0.2171	0.0341	0.0511	0.0830	0.1291	0.1143
			2 FD	0.0473	0.0496	0.0602	0.0689	0.0973	0.0286	0.0299	0.0367	0.0497	0.0675	0.0187	0.0276	0.0450	0.0686	0.0694
			2 SYS	0.0441	0.0481	0.0620	0.0804	0.1108	0.0273	0.0280	0.0339	0.0491	0.0835	0.0178	0.0226	0.0364	0.0549	0.0708
			5 FD	0.0177	0.0194	0.0231	0.0272	0.0353	0.0106	0.0112	0.0137	0.0157	0.0192	0.0074	0.0085	0.0112	0.0153	0.0182
			5 SYS	0.0172	0.0192	0.0243	0.0309	0.0423	0.0103	0.0102	0.0113	0.0134	0.0180	0.0069	0.0076	0.0094	0.0125	0.0170
			10 FD	0.0091	0.0096	0.0110	0.0131	0.0172	0.0051	0.0054	0.0060	0.0071	0.0084	0.0036	0.0040	0.0045	0.0056	0.0063
			10 SYS	0.0090	0.0095	0.0116	0.0150	0.0195	0.0050	0.0051	0.0052	0.0058	0.0070	0.0034	0.0036	0.0039	0.0048	0.0057
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	250	sw	0.1 FD	0.4501	0.4870	0.4858	0.7213	1.0538	0.2473	0.2576	0.2655	0.5829	0.5312	0.1464	0.1713	0.2542	0.4081	0.2523
			0.1 SYS	0.1292	0.1628	0.1746	0.1404	0.2209	0.1116	0.1738	0.2417	0.2378	0.8827	0.0943	0.1178	0.1671	0.1224	0.9083
			0.2 FD	0.2102	0.2356	0.2346	0.3747	0.5287	0.1198	0.1201	0.1481	0.3531	0.2638	0.0736	0.1170	0.2120	0.3502	0.1268
			0.2 SYS	0.1127	0.1417	0.1413	0.1475	0.2422	0.0876	0.1184	0.1187	0.2593	0.7146	0.0633	0.0671	0.0908	0.1688	0.5739
			0.5 FD	0.0863	0.0875	0.1038	0.1939	0.1827	0.0478	0.0508	0.0829	0.1955	0.1158	0.0307	0.0834	0.1676	0.2758	0.0496
			0.5 SYS	0.0713	0.0742	0.0834	0.1498	0.2125	0.0442	0.0453	0.0707	0.2039	0.3909	0.0297	0.0654	0.1269	0.2006	0.2047
			1 FD	0.0419	0.0437	0.0605	0.1025	0.0866	0.0229	0.0280	0.0512	0.1009	0.1240	0.0160	0.0508	0.1016	0.1659	0.0267
			1 SYS	0.0384	0.0394	0.0603	0.1051	0.1520	0.0220	0.0246	0.0463	0.1114	0.2091	0.0159	0.0420	0.0812	0.1349	0.1143
			2 FD	0.0204	0.0223	0.0327	0.0416	0.0458	0.0115	0.0143	0.0245	0.0416	0.0634	0.0076	0.0208	0.0414	0.0652	0.0580
			2 SYS	0.0189	0.0216	0.0379	0.0628	0.0846	0.0109	0.0118	0.0197	0.0388	0.0740	0.0074	0.0159	0.0288	0.0474	0.0643
			5 FD	0.0077	0.0090	0.0113	0.0121	0.0189	0.0045	0.0054	0.0087	0.0133	0.0172	0.0032	0.0050	0.0088	0.0141	0.0183
			5 SYS	0.0073	0.0085	0.0137	0.0220	0.0332	0.0044	0.0043	0.0056	0.0092	0.0135	0.0032	0.0040	0.0056	0.0086	0.0133
			10 FD	0.0038	0.0042	0.0050	0.0058	0.0086	0.0021	0.0026	0.0039	0.0058	0.0068	0.0016	0.0019	0.0029	0.0045	0.0063
			10 SYS	0.0037	0.0041	0.0057	0.0086	0.0125	0.0020	0.0022	0.0024	0.0033	0.0042	0.0016	0.0018	0.0020	0.0027	0.0037
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw	0.1 FD	0.3056	0.3618	0.3209	0.6196	0.9299	0.1627	0.1762	0.2020	0.5273	0.4374	0.1035	0.1288	0.2324	0.3853	0.1879
			0.1 SYS	0.0918	0.1420	0.1526	0.1041	0.1952	0.0739	0.1316	0.1787	0.2808	0.9227	0.0683	0.1024	0.1290	0.1079	0.9783
			0.2 FD	0.1565	0.1606	0.1701	0.3698	0.4429	0.0827	0.0846	0.1208	0.3278	0.2144	0.0498	0.0991	0.2055	0.3407	0.0910
			0.2 SYS	0.0828	0.1222	0.1131	0.1322	0.2212	0.0587	0.0859	0.0797	0.2906	0.7431	0.0436	0.0503	0.0885	0.1992	0.6030
			0.5 FD	0.0608	0.0630	0.0808	0.1809	0.1418	0.0330	0.0364	0.0767	0.1895	0.0987	0.0216	0.0790	0.1627	0.2741	0.0377
			0.5 SYS	0.0500	0.0573	0.0641	0.1372	0.2120	0.0296	0.0316	0.0675	0.2109	0.4019	0.0208	0.0609	0.1267	0.2098	0.2118
			1 FD	0.0309	0.0311	0.0514	0.0955	0.0612	0.0164	0.0219	0.0473	0.0990	0.1205	0.0109	0.0496	0.0980	0.1623	0.0186
			1 SYS	0.0278	0.0294	0.0527	0.1008	0.1477	0.0157	0.0183	0.0433	0.1107	0.2108	0.0108	0.0395	0.0779	0.1334	0.1170
			2 FD	0.0147	0.0165	0.0291	0.0360	0.0379	0.0082	0.0113	0.0225	0.0414	0.0638	0.0055	0.0197	0.0404	0.0634	0.0571
			2 SYS	0.0135	0.0166	0.0340	0.0587	0.0836	0.0078	0.0087	0.0173	0.0380	0.0739	0.0054	0.0143	0.0272	0.0453	0.0628
			5 FD	0.0053	0.0067	0.0094	0.0084	0.0164	0.0030	0.0044	0.0081	0.0129	0.0168	0.0021	0.0045	0.0086	0.0141	0.0180
			5 SYS	0.0050	0.0061	0.0115	0.0199	0.0320	0.0029	0.0031	0.0045	0.0084	0.0130	0.0021	0.0033	0.0051	0.0079	0.0124
			10 FD	0.0026	0.0030	0.0037	0.0042	0.0069	0.0015	0.0020	0.0035	0.0053	0.0067	0.0011	0.0015	0.0027	0.0044	0.0064
			10 SYS	0.0025	0.0028	0.0044	0.0076	0.0114	0.0014	0.0015	0.0018	0.0027	0.0039	0.0011	0.0012	0.0017	0.0023	0.0033

Table IV-3: Relative RMSE of $\hat{\beta}$ over N, T, sw and γ

3.2.3 RMSE of β^* (Long-Run Effects)

In this subsection, we combine the results of the previous two. In a dynamic setting such as stated in equation (IV.1), a one-unit increase in the exogenous regressor has an effect of size β on the dependent variable in the same period. Through the dependence of the dependent variable on past periods, the effect on the following period therefore is $\gamma \cdot \beta$ and for the subsequent period $\gamma^2 \cdot \beta$, for the following $\gamma^3 \cdot \beta$ and so on. Summing up all those effects results in the long-run effect of a change in x_{it} which equals $\beta / (1-\gamma)$. β therefore is often labeled the short-run effect of x_{it} on y_{it}. In this subsection, we discuss the bias and standard deviation of estimated long-run effects ($\hat{\beta}^*$) of the exogenous regressor on the dependent variable. Since we fixed β equal one and varied γ over 0.0, 0.2, 0.4, 0.6, 0.8, the true values of the long-run effects are equal to 1, 1.25, 1.67, 2.5, 5.0.

Table IV-4 shows the RMSE of the long-run effects for both estimators, relative to the true values. First, despite a few situations, SYS-GMM exhibits a much lower RMSE of the long-run estimates compared to FD-GMM. For the cases in which FD-GMM has a lower RMSE, the difference is generally small. For certain parameter choices, mostly when $\gamma = 0.8$, the RMSE of the FD-GMM estimates is implausibly high. The reason is that the long-run effects are composed of the short-run effect and the coefficient of the lagged dependent variable. If the latter is sufficiently close to one, say e.g. $|\gamma - 1| < 0.05$, the long-run effects will increase by 20, since the short-run effects are fixed to one. Such situations tend to inflate the variance of the estimates strongly, but the variances in turn are incorporated in the RMSE and therefore inflates this as well. Out of the 315,000 Monte Carlo experiments in which the FD-GMM estimator was used, 8,540 point estimates for γ fall in the range mentioned above, but only 2 point estimates of the SYS-GMM estimator. Even if we rule out results where the simulated value of γ equals 0.8, which corresponds to the situation where FD-GMM estimates of γ approach one, the RMSE of long-run effects estimated by the SYS-GMM estimator is 35% below the RMSE of FD-GMM estimates. For a small within variation of the exogenous regressor, the RMSE of SYS-GMM is 40% lower than for the estimates obtained by FD-GMM. For within variations larger than one, the efficiency gain still amounts to some 25% on average. Let us consider an extreme case in which N = 50, T = 10, $s_{w,x_{it}} = 0.1$ and $\gamma = 0.8$. As previously said, the reported RMSE is so high because the standard deviation of the estimates is very high; see also Table IV-11. Investigating the case more closely, we find that 16.2% of the replications fall in the range $|\gamma - 1| < 0.05$, and 2.8% of the replications in the range $|\gamma - 1| < 0.01$. If we rule out those outliers, the relative RMSE for FD-GMM results in 1.66 and that for SYS-GMM in 0.84, which still is almost 50% lower. Analogous conclusions could be drawn for the other cases with such extreme values for the RMSE of FD-GMM estimates. We could argue that results of a Monte Carlo experiment should be considered without dropping outliers, even if the coefficient of the lagged dependent variable is very close to one, say 0.9999. On the other hand, an empirical researcher would not use results like this but rather try to re-specify the model in some sense. Either way, the results stay qualitatively the same, namely that the long-run effects estimated by

SYS-GMM exhibit a significantly lower RMSE than those obtained by FD-GMM, and most of all for situations where the within variation of the exogenous regressor is below one.

Table IV-10 and Table IV-11 in the Appendix show the relative bias and standard deviation of the long-run estimates. Rather as it holds for the bias of γ and the bias of the short-run effects β, FD-GMM exhibits a lower bias than SYS-GMM estimates for certain parameter ranges. However, the much lower standard deviation (or variance) can offset this advantage of FD-GMM when considering the combined measure of accuracy, the RMSE.

				T														
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	50	sw	0.1 FD	1.0030	0.9084	0.8622	1.2323	263.6558	0.5610	0.5013	0.5565	0.8560	1273.6803	0.3296	0.3761	0.5610	0.7750	4.8921
			SYS	0.3449	0.2842	0.2859	0.4057	0.5584	0.2585	0.1998	0.2513	0.4994	0.8486	0.2370	0.1514	0.1993	0.3855	0.7759
			0.2 FD	0.4717	0.4448	0.5022	1.6801	21.7597	0.2794	0.2915	0.4541	0.7089	11.0784	0.1839	0.2926	0.5267	0.7549	4.2984
			SYS	0.2865	0.2286	0.2859	0.4251	0.5604	0.1951	0.1823	0.3006	0.5306	0.7750	0.1610	0.1481	0.3053	0.4946	0.6733
			0.5 FD	0.1890	0.2178	0.3392	2.2331	23.2170	0.1196	0.1866	0.3502	0.5792	33.5227	0.0861	0.2346	0.4544	0.6894	15.5485
			SYS	0.1669	0.1738	0.2728	0.4142	0.5333	0.1085	0.1600	0.3079	0.5126	0.6490	0.0817	0.1850	0.3744	0.5595	0.5597
			1 FD	0.1058	0.1463	0.2599	3.4182	15.1152	0.0654	0.1150	0.2221	0.3767	1.6657	0.0525	0.1514	0.3141	0.5297	0.3500
			SYS	0.1042	0.1257	0.2126	0.3086	0.4573	0.0615	0.1070	0.2048	0.3683	0.5591	0.0485	0.1305	0.2804	0.4743	0.5353
			2 FD	0.0642	0.0910	0.1505	1.0218	33.3721	0.0353	0.0547	0.0923	0.1561	0.2471	0.0308	0.0662	0.1427	0.2605	0.3564
			SYS	0.0619	0.0810	0.1300	0.1895	0.3375	0.0339	0.0486	0.0781	0.1435	0.3297	0.0284	0.0559	0.1243	0.2471	0.4361
			5 FD	0.0264	0.0403	0.0690	0.1494	0.9804	0.0145	0.0171	0.0252	0.0349	0.0610	0.0119	0.0173	0.0316	0.0541	0.0930
			SYS	0.0260	0.0371	0.0564	0.0817	0.1510	0.0141	0.0158	0.0205	0.0289	0.0773	0.0112	0.0152	0.0270	0.0526	0.1318
			10 FD	0.0140	0.0194	0.0319	0.0614	0.1695	0.0068	0.0075	0.0093	0.0124	0.0237	0.0059	0.0073	0.0102	0.0162	0.0245
			SYS	0.0142	0.0185	0.0285	0.0437	0.0790	0.0068	0.0072	0.0083	0.0109	0.0240	0.0055	0.0066	0.0092	0.0160	0.0319
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	sw	0.1 FD	0.4357	0.3799	0.4421	0.7734	5.1180	0.2447	0.2442	0.4402	0.7772	38.0350	0.1479	0.2816	0.5288	0.7546	12.1601
			SYS	0.1249	0.1157	0.2086	0.3698	0.5059	0.1132	0.1070	0.2372	0.5671	0.9154	0.0968	0.1469	0.3029	0.5520	0.9477
			0.2 FD	0.2029	0.1984	0.3197	0.5604	121.8817	0.1176	0.1825	0.4129	0.6808	33.5318	0.0768	0.2620	0.5095	0.7356	9.1530
			SYS	0.1093	0.1112	0.2186	0.3822	0.5180	0.0878	0.1217	0.3013	0.5868	0.8141	0.0668	0.1958	0.3966	0.6176	0.7772
			0.5 FD	0.0816	0.1285	0.2662	0.4004	76.5192	0.0515	0.1464	0.3305	0.5662	1406.5060	0.0381	0.2179	0.4457	0.6770	3.6675
			SYS	0.0699	0.1103	0.2271	0.3686	0.4984	0.0473	0.1311	0.3040	0.5233	0.6531	0.0361	0.1970	0.4022	0.6063	0.6084
			1 FD	0.0435	0.0919	0.1840	0.2398	70.6051	0.0266	0.0916	0.2053	0.3634	0.2556	0.0235	0.1402	0.3080	0.5183	0.5203
			SYS	0.0403	0.0900	0.1787	0.2869	0.4292	0.0253	0.0839	0.1862	0.3488	0.5371	0.0231	0.1290	0.2818	0.4801	0.5469
			2 FD	0.0276	0.0491	0.0785	0.1556	47.0362	0.0150	0.0390	0.0858	0.1474	0.2116	0.0123	0.0578	0.1367	0.2512	0.2925
			SYS	0.0253	0.0547	0.1112	0.1790	0.3109	0.0143	0.0330	0.0676	0.1253	0.2906	0.0120	0.0502	0.1141	0.2272	0.4093
			5 FD	0.0117	0.0174	0.0318	0.0997	0.4304	0.0061	0.0097	0.0191	0.0310	0.0456	0.0053	0.0118	0.0271	0.0491	0.0738
			SYS	0.0110	0.0208	0.0395	0.0660	0.1394	0.0059	0.0081	0.0132	0.0217	0.0645	0.0053	0.0104	0.0207	0.0436	0.1082
			10 FD	0.0060	0.0087	0.0165	0.0424	0.1294	0.0029	0.0038	0.0065	0.0100	0.0137	0.0027	0.0039	0.0076	0.0130	0.0155
			SYS	0.0059	0.0092	0.0155	0.0266	0.0509	0.0029	0.0034	0.0045	0.0065	0.0168	0.0027	0.0036	0.0061	0.0120	0.0241
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw	0.1 FD	0.2955	0.2686	0.3499	0.7309	6.1023	0.1609	0.1857	0.4306	0.7586	8.8544	0.1048	0.2635	0.5237	0.7498	3.3833
			SYS	0.0883	0.0758	0.1866	0.3653	0.5020	0.0747	0.0970	0.2577	0.6072	0.9448	0.0701	0.1377	0.3163	0.5907	0.9874
			0.2 FD	0.1497	0.1442	0.3093	0.5768	10.9334	0.0816	0.1718	0.4057	0.6691	161.1206	0.0519	0.2538	0.5097	0.7330	4.0418
			SYS	0.0788	0.0796	0.2108	0.3863	0.5167	0.0586	0.1235	0.3140	0.6121	0.8353	0.0459	0.1912	0.4091	0.6402	0.7966
			0.5 FD	0.0563	0.1145	0.2598	0.3972	76.3006	0.0342	0.1405	0.3277	0.5608	6.9862	0.0262	0.2158	0.4430	0.6772	3.2755
			SYS	0.0491	0.0961	0.2211	0.3656	0.5032	0.0310	0.1298	0.3085	0.5325	0.6645	0.0250	0.1961	0.4043	0.6154	0.6200
			1 FD	0.0317	0.0879	0.1806	0.2164	212.4910	0.0190	0.0890	0.2054	0.3635	0.2133	0.0160	0.1394	0.3040	0.5143	0.5801
			SYS	0.0299	0.0845	0.1761	0.2886	0.4319	0.0178	0.0837	0.1896	0.3515	0.5404	0.0157	0.1274	0.2771	0.4782	0.5508
			2 FD	0.0193	0.0433	0.0683	0.1131	1245.7633	0.0102	0.0374	0.0853	0.1486	0.2129	0.0092	0.0566	0.1353	0.2471	0.2847
			SYS	0.0179	0.0517	0.1079	0.1744	0.3115	0.0096	0.0321	0.0680	0.1274	0.2913	0.0089	0.0490	0.1110	0.2208	0.4030
			5 FD	0.0082	0.0124	0.0247	0.0922	0.4168	0.0041	0.0086	0.0187	0.0309	0.0435	0.0036	0.0113	0.0269	0.0479	0.0707
			SYS	0.0078	0.0172	0.0370	0.0632	0.1365	0.0040	0.0071	0.0126	0.0213	0.0633	0.0035	0.0097	0.0203	0.0411	0.1037
			10 FD	0.0042	0.0059	0.0137	0.0382	0.1261	0.0020	0.0031	0.0061	0.0096	0.0117	0.0018	0.0034	0.0072	0.0122	0.0133
			SYS	0.0041	0.0069	0.0133	0.0242	0.0458	0.0020	0.0026	0.0040	0.0058	0.0149	0.0018	0.0030	0.0056	0.0108	0.0218

Table IV-4: Relative RMSE of $\hat{\beta}^*$ over N, T, sw and γ

4 Conclusion

Quite often in applied economic research, the exogenous variables are characterised by a relatively low within variation compared to the between variation. A large group of panel data estimators solve the problem of time-invariant unobserved heterogeneity bias by either first-differencing or group-wise de-meaning the variables. This may leave variables of interest with a small variation, rendering the coefficient estimates highly volatile and thus having a large variance. Clark and Linzer (2012) recently discussed the problem in the context of static panel data models using whichthey compared estimation accuracy in terms of bias and efficiency for the fixed and random effects estimator for the cases of low-within variation of the exogenous regressors. Importantly, they underline that the general tendency of applied researchers is to avoid bias in estimated coefficients while allowing for a high variance in the parameter estimate, thusrendering inference about the coefficients highly inaccurate.

In order to analyse the effect of a small within variation of the exogenous regressor on the parameter estimates in context of dynamic panel data models, we conduct a Monte-Carlo experiment where, similarly as it was done by Clark and Linzer (2012), we vary the number of panel units N, the number of observations per unit T, the within variation of the exogenous regressor and the coefficient of the lagged dependent variable. We keep the variance of the error term of the underlying DGP fixed to 1 and also fix the between variation of the exogenous regressor to 1.

Accounting for both bias and standard deviation of the parameter estimate, one should consider the root mean squared error of the replications. First of all, if the within variation of the exogenous regressor is small, SYS-GMM exhibits a significantly lower RMSE than FD-GMM. There are parameter combinations where FD-GMM has a lower RMSE than SYS-GMM, but the difference is never significant unless $\gamma = 0.8$ or higher and unless the within variation of the exogenous regressor is above 1. Quite often in applied economic research, panel data sets are used for estimations where N and T are rather small (e.g., there are N=48 onshore US states or N=47 onshore Spanish provinces, N=26 federal districts for Switzerland, N=34 states in the OECD or N=27 states in the EU and so on). When estimating the short-run marginal effect of the exogenous regressors in such a situation, it is wise to use the SYS-GMM estimator if the researches focuses on the variance and the bias of the parameter estimate jointly. Most of all, this is true as long as the within variation of the exogenous regressor is small. If either the within variation is above one and or the number of panel units is as large as N=500, the application of either FD-GMM and SYS-GMM will not result in a significant difference and the researcher should apply the estimator more suitable to the specific research question (although still differences exist but they are not as substantial as other cases).

In applied research, the long-run marginal effects of the regressors are the subject of interest. This requires inference using the estimates β and γ and thus an accurate estimate in both terms of bias and variance of γ. If the within variation of the exogenous regressor is relatively low (below one), then the SYS-GMM estimator of γ exhibits a significantly lower RMSE than the FD-GMM estimator. This holds for almost all parameter combinations. however, the difference is no longer significantly in favour of SYS-GMM when T approaches 20. There are a few situations where FD-GMM has a significantly lower RMSE than SYS-GMM if the within variation of the exogenous regressor is above one, but the difference is not big.

Table IV-5 gives a résumé of the Monte Carlo experiment and the different situations that were simulated[35]. For a complete overview, we also summarize results in terms of bias and variance. We were basically interested in the analysis of the impact of low within variation of an exogenous regressor on the estimation quality of the coefficients β and γ (bold in the table). Moreover, the focus lies on data characteristics often encountered in empirical studies using dynamic panel data methods: where either N is low and T is low or where N is large and T is low. If a researcher's interest is to obtain unbiased estimates of γ and β, FD-GMM should be preferred. On the other hand, if the researcher is willing to trade off bias against variance, or to accept a combination of the two, the RMSE, then the SYS-GMM estimator should be preferred. Often, the goal of a study using dynamic panel data is to estimate long-run effects of a certain variable. Our simulations show that unless the within variation is large and T is large, the SYS-GMM estimator should be preferred in terms of RMSE in these cases.

				Researcher's Focus / Interest								
				γ			β			β*=β/(1−γ)		
				Bias	Variance	RMSE	Bias	Variance	RMSE	Bias	Variance	RMSE
sw low	N low	T low	γ low	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
		T large	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	FD-GMM	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM
	N large	T low	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM
		T large	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	FD-GMM	FD-GMM	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM
sw large	N low	T low	γ low	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM	FD-GMM	SYS-GMM	SYS-GMM
		T large	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM
	N large	T low	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM
		T large	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM

Table IV-5: Decision rules according to the Monte Carlo experiment

35 sw low corresponds to the situation where the within variation of the exogenous regressor is below 1. N low corresponds to N = 50, whereas N = 250, 500 is denoted as N large. T low corresponds to T = 5, 10 and T = 20 denote the cases where T is large. γ large denotes the cases where γ = 0.8

5 Appendix

5.1 Results with Restricted Instruments to a Depth of 2 Lags

5.1.1 Relative Bias and Standard Deviation of γ

				T														
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	50	0.1	FD	-0.0179	-0.6560	-0.6565	-0.5315	-0.0708	-0.0055	-0.8475	-0.8928	-0.7493	0.1041	-0.0034	-0.9450	-0.9995	-1.0468	0.1106
			SYS	-0.0115	-0.6935	-0.6018	-0.4007	-0.1731	0.0062	-0.8365	-0.8133	-0.5847	-0.1263	0.0031	-0.9065	-0.9128	-0.8103	-0.1843
		0.2	FD	-0.0038	-0.6670	-0.6243	-0.4702	-0.0448	-0.0102	-0.8230	-0.8350	-0.7580	0.0893	-0.0030	-0.9115	-0.9603	-1.0278	0.1070
			SYS	0.0002	-0.6490	-0.5235	-0.3448	-0.1735	0.0029	-0.7865	-0.7403	-0.5537	-0.1609	0.0072	-0.8295	-0.8358	-0.7398	-0.2088
		0.5	FD	-0.0144	-0.5885	-0.5248	-0.3192	0.0618	-0.0077	-0.6065	-0.6008	-0.6003	0.0378	-0.0051	-0.7385	-0.7825	-0.8672	0.0765
			SYS	0.0015	-0.4725	-0.3828	-0.2707	-0.1621	0.0039	-0.5700	-0.5430	-0.4575	-0.2046	0.0040	-0.6575	-0.6848	-0.6372	-0.2339
		1	FD	-0.0056	-0.3920	-0.3020	-0.1297	0.1201	-0.0037	-0.3420	-0.3163	-0.2955	-0.0973	-0.0033	-0.4360	-0.4620	-0.5115	-0.0320
			SYS	0.0063	-0.2780	-0.2353	-0.1832	-0.1384	0.0014	-0.3080	-0.2803	-0.2720	-0.2013	0.0010	-0.3925	-0.4233	-0.4468	-0.2339
		2	FD	-0.0008	-0.1350	-0.0740	0.0213	0.1193	-0.0026	-0.1260	-0.1023	-0.0875	-0.0511	-0.0010	-0.1650	-0.1725	-0.1743	-0.1128
			SYS	0.0031	-0.1415	-0.1190	-0.0912	-0.0901	-0.0003	-0.1015	-0.0790	-0.0757	-0.0931	0.0001	-0.1430	-0.1528	-0.1720	-0.1644
		5	FD	-0.0015	-0.0015	0.0215	0.0370	0.0568	-0.0008	-0.0185	-0.0170	-0.0118	-0.0060	-0.0004	-0.0270	-0.0305	-0.0268	-0.0191
			SYS	-0.0001	-0.0360	-0.0330	-0.0247	-0.0258	-0.0002	-0.0175	-0.0123	-0.0083	-0.0139	0.0000	-0.0220	-0.0260	-0.0280	-0.0328
		10	FD	0.0000	0.0100	0.0140	0.0195	0.0221	-0.0001	-0.0040	-0.0030	-0.0020	-0.0004	-0.0002	-0.0070	-0.0073	-0.0063	-0.0030
			SYS	0.0005	-0.0085	-0.0103	-0.0087	-0.0054	0.0002	-0.0040	-0.0023	-0.0013	-0.0015	-0.0001	-0.0055	-0.0065	-0.0070	-0.0058
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	0.1	FD	-0.0001	-0.6065	-0.5735	-0.4348	-0.0295	-0.0003	-0.8110	-0.8775	-0.7163	0.1851	0.0004	-0.9225	-0.9783	-1.0050	0.1856
			SYS	0.0007	-0.6400	-0.5395	-0.3555	-0.1635	0.0029	-0.8250	-0.7983	-0.5823	-0.0990	0.0029	-0.8910	-0.8930	-0.8448	-0.1990
		0.2	FD	-0.0014	-0.5975	-0.5345	-0.3793	0.0236	-0.0016	-0.7665	-0.8190	-0.7327	0.1833	0.0006	-0.8820	-0.9430	-0.9918	0.1880
			SYS	0.0013	-0.6100	-0.4885	-0.3178	-0.1571	0.0026	-0.7735	-0.7313	-0.5593	-0.1375	0.0035	-0.8435	-0.8498	-0.8118	-0.2325
		0.5	FD	-0.0020	-0.4965	-0.4260	-0.2527	0.1321	-0.0007	-0.5795	-0.5858	-0.5823	0.1531	0.0003	-0.7075	-0.7590	-0.8337	0.1733
			SYS	0.0028	-0.4680	-0.3640	-0.2502	-0.1494	0.0019	-0.5820	-0.5375	-0.4552	-0.1915	0.0020	-0.6815	-0.6990	-0.6922	-0.2609
N	250	sw 1	FD	-0.0018	-0.3400	-0.2485	-0.0915	0.1743	-0.0013	-0.3225	-0.3013	-0.2783	-0.0221	0.0000	-0.4210	-0.4498	-0.4897	0.0700
			SYS	0.0018	-0.3150	-0.2390	-0.1770	-0.1248	0.0005	-0.3100	-0.2675	-0.2457	-0.1785	0.0008	-0.4065	-0.4220	-0.4445	-0.2401
		2	FD	-0.0019	-0.1140	-0.0500	0.0355	0.1450	-0.0005	-0.1175	-0.1035	-0.0830	-0.0458	0.0002	-0.1570	-0.1660	-0.1658	-0.0829
			SYS	-0.0002	-0.1620	-0.1275	-0.0958	-0.0831	0.0001	-0.1070	-0.0805	-0.0662	-0.0765	0.0005	-0.1460	-0.1453	-0.1555	-0.1465
		5	FD	-0.0001	0.0065	0.0303	0.0498	0.0640	-0.0002	-0.0170	-0.0165	-0.0120	-0.0063	-0.0001	-0.0265	-0.0280	-0.0243	-0.0146
			SYS	0.0001	-0.0475	-0.0385	-0.0295	-0.0303	0.0000	-0.0180	-0.0123	-0.0080	-0.0133	0.0000	-0.0255	-0.0233	-0.0243	-0.0265
		10	FD	0.0000	0.0110	0.0180	0.0227	0.0248	-0.0001	-0.0035	-0.0035	-0.0027	-0.0009	0.0000	-0.0065	-0.0068	-0.0055	-0.0019
			SYS	0.0001	-0.0140	-0.0118	-0.0098	-0.0083	-0.0001	-0.0050	-0.0030	-0.0017	-0.0028	0.0000	-0.0065	-0.0060	-0.0062	-0.0050
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	0.1	FD	-0.0007	-0.5955	-0.5593	-0.4285	0.0111	-0.0005	-0.8070	-0.8593	-0.7035	0.1881	0.0002	-0.9210	-0.9805	-1.0047	0.1898
			SYS	0.0000	-0.6270	-0.5330	-0.3585	-0.1663	0.0014	-0.8220	-0.7883	-0.5858	-0.1020	0.0019	-0.8945	-0.8953	-0.8612	-0.2045
		0.2	FD	0.0003	-0.5715	-0.5390	-0.3778	0.0709	-0.0006	-0.7585	-0.8068	-0.7110	0.1883	-0.0007	-0.8885	-0.9488	-0.9902	0.1920
			SYS	0.0001	-0.5955	-0.4888	-0.3167	-0.1608	0.0013	-0.7750	-0.7293	-0.5657	-0.1420	0.0013	-0.8535	-0.8565	-0.8295	-0.2404
		0.5	FD	-0.0006	-0.4930	-0.4185	-0.2518	0.1645	-0.0008	-0.5700	-0.5763	-0.5700	0.1610	-0.0001	-0.7075	-0.7590	-0.8350	0.1829
			SYS	0.0013	-0.4690	-0.3590	-0.2503	-0.1504	0.0003	-0.5870	-0.5398	-0.4627	-0.1970	0.0012	-0.6865	-0.7040	-0.7055	-0.2701
		1	FD	0.0000	-0.3380	-0.2500	-0.0970	0.1888	-0.0008	-0.3200	-0.3038	-0.2787	-0.0160	-0.0001	-0.4195	-0.4450	-0.4842	0.0848
			SYS	0.0012	-0.3095	-0.2370	-0.1795	-0.1269	0.0001	-0.3195	-0.2758	-0.2487	-0.1800	0.0004	-0.4065	-0.4150	-0.4415	-0.2421
		2	FD	-0.0004	-0.1155	-0.0488	0.0395	0.1515	-0.0001	-0.1165	-0.1045	-0.0842	-0.0469	0.0000	-0.1570	-0.1648	-0.1628	-0.0795
			SYS	0.0008	-0.1600	-0.1263	-0.0942	-0.0834	0.0002	-0.1095	-0.0833	-0.0683	-0.0768	0.0003	-0.1475	-0.1420	-0.1505	-0.1426
		5	FD	-0.0001	0.0080	0.0290	0.0510	0.0661	0.0001	-0.0175	-0.0163	-0.0123	-0.0064	-0.0001	-0.0270	-0.0283	-0.0235	-0.0140
			SYS	0.0001	-0.0450	-0.0385	-0.0300	-0.0299	0.0002	-0.0195	-0.0128	-0.0087	-0.0133	0.0000	-0.0260	-0.0235	-0.0230	-0.0254
		10	FD	0.0001	0.0100	0.0180	0.0223	0.0255	0.0000	-0.0035	-0.0038	-0.0027	-0.0008	-0.0001	-0.0070	-0.0068	-0.0052	-0.0015
			SYS	0.0001	-0.0140	-0.0120	-0.0102	-0.0080	0.0000	-0.0050	-0.0033	-0.0018	-0.0025	0.0000	-0.0065	-0.0060	-0.0057	-0.0046

Table IV-6: Relative bias $\hat{\gamma} - \gamma$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	50	sw	0.1 FD	0.0983	0.5855	0.3893	0.3512	0.4320	0.0676	0.3975	0.2390	0.2518	0.0991	0.0444	0.2470	0.1358	0.1180	0.0556
			0.1 SYS	0.0794	0.4310	0.2428	0.1565	0.1159	0.0602	0.3435	0.1733	0.1417	0.0641	0.0404	0.2185	0.1138	0.0868	0.0513
			0.2 FD	0.1006	0.5930	0.3743	0.3507	0.3798	0.0665	0.3560	0.2275	0.2250	0.1033	0.0440	0.2330	0.1298	0.1165	0.0584
			0.2 SYS	0.0812	0.4140	0.2120	0.1275	0.1019	0.0600	0.3075	0.1675	0.1272	0.0590	0.0397	0.2080	0.1090	0.0840	0.0480
			0.5 FD	0.0954	0.5315	0.3375	0.3057	0.3186	0.0569	0.3045	0.1765	0.1653	0.1280	0.0387	0.2085	0.1173	0.0973	0.0655
			0.5 SYS	0.0739	0.3645	0.1713	0.1037	0.0690	0.0536	0.2585	0.1440	0.1025	0.0510	0.0359	0.1845	0.1010	0.0788	0.0444
			1 FD	0.0821	0.4650	0.2745	0.2332	0.2516	0.0421	0.2135	0.1180	0.0902	0.1118	0.0279	0.1505	0.0840	0.0697	0.0735
			1 SYS	0.0586	0.2880	0.1393	0.0813	0.0536	0.0398	0.1975	0.1013	0.0705	0.0503	0.0259	0.1410	0.0805	0.0608	0.0421
			2 FD	0.0547	0.3190	0.1878	0.1640	0.1488	0.0245	0.1300	0.0623	0.0408	0.0458	0.0180	0.0905	0.0495	0.0360	0.0419
			2 SYS	0.0413	0.1940	0.0953	0.0598	0.0401	0.0232	0.1160	0.0543	0.0340	0.0316	0.0164	0.0840	0.0463	0.0345	0.0296
			5 FD	0.0228	0.1420	0.0828	0.0645	0.0533	0.0108	0.0505	0.0235	0.0142	0.0126	0.0070	0.0380	0.0185	0.0117	0.0109
			5 SYS	0.0193	0.1065	0.0525	0.0345	0.0246	0.0105	0.0480	0.0223	0.0130	0.0113	0.0066	0.0360	0.0180	0.0115	0.0106
			10 FD	0.0110	0.0630	0.0365	0.0273	0.0223	0.0051	0.0235	0.0110	0.0065	0.0056	0.0035	0.0185	0.0090	0.0055	0.0045
			10 SYS	0.0101	0.0545	0.0295	0.0207	0.0159	0.0051	0.0235	0.0108	0.0063	0.0054	0.0033	0.0175	0.0088	0.0057	0.0046
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	sw	0.1 FD	0.0441	0.2465	0.1535	0.1217	0.1835	0.0294	0.1660	0.1078	0.1150	0.0369	0.0198	0.1085	0.0613	0.0518	0.0220
			0.1 SYS	0.0334	0.1675	0.0918	0.0580	0.0475	0.0258	0.1350	0.0810	0.0652	0.0291	0.0185	0.0965	0.0520	0.0403	0.0234
			0.2 FD	0.0422	0.2485	0.1518	0.1132	0.1715	0.0277	0.1560	0.1015	0.1037	0.0374	0.0186	0.1045	0.0610	0.0527	0.0214
			0.2 SYS	0.0327	0.1670	0.0820	0.0532	0.0410	0.0242	0.1245	0.0743	0.0618	0.0264	0.0178	0.0945	0.0523	0.0410	0.0234
			0.5 FD	0.0413	0.2340	0.1318	0.1072	0.1538	0.0242	0.1300	0.0798	0.0773	0.0524	0.0169	0.0880	0.0533	0.0463	0.0238
			0.5 SYS	0.0310	0.1440	0.0733	0.0437	0.0283	0.0222	0.1115	0.0643	0.0515	0.0243	0.0160	0.0820	0.0470	0.0360	0.0235
			1 FD	0.0353	0.1980	0.1168	0.0985	0.1124	0.0181	0.0945	0.0500	0.0405	0.0500	0.0126	0.0690	0.0395	0.0335	0.0323
			1 SYS	0.0254	0.1235	0.0558	0.0342	0.0225	0.0169	0.0865	0.0450	0.0310	0.0219	0.0124	0.0670	0.0373	0.0302	0.0204
			2 FD	0.0254	0.1445	0.0848	0.0730	0.0635	0.0104	0.0565	0.0270	0.0190	0.0200	0.0073	0.0420	0.0215	0.0177	0.0186
			2 SYS	0.0172	0.0850	0.0415	0.0247	0.0171	0.0099	0.0505	0.0238	0.0152	0.0126	0.0072	0.0405	0.0203	0.0155	0.0126
			5 FD	0.0104	0.0625	0.0360	0.0287	0.0230	0.0044	0.0220	0.0095	0.0063	0.0058	0.0031	0.0165	0.0085	0.0057	0.0046
			5 SYS	0.0082	0.0445	0.0220	0.0148	0.0103	0.0043	0.0210	0.0090	0.0058	0.0046	0.0031	0.0165	0.0083	0.0053	0.0043
			10 FD	0.0049	0.0285	0.0155	0.0120	0.0098	0.0022	0.0100	0.0045	0.0028	0.0024	0.0016	0.0080	0.0040	0.0025	0.0018
			10 SYS	0.0044	0.0245	0.0123	0.0088	0.0066	0.0022	0.0100	0.0045	0.0028	0.0023	0.0016	0.0080	0.0040	0.0025	0.0018
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw	0.1 FD	0.0303	0.1710	0.1093	0.0888	0.1313	0.0201	0.1205	0.0768	0.0827	0.0251	0.0135	0.0725	0.0428	0.0362	0.0144
			0.1 SYS	0.0236	0.1200	0.0640	0.0440	0.0330	0.0179	0.0945	0.0553	0.0468	0.0200	0.0126	0.0650	0.0360	0.0287	0.0180
			0.2 FD	0.0296	0.1715	0.1035	0.0818	0.1284	0.0190	0.1110	0.0675	0.0737	0.0270	0.0139	0.0685	0.0418	0.0372	0.0150
			0.2 SYS	0.0236	0.1170	0.0573	0.0377	0.0288	0.0169	0.0875	0.0518	0.0447	0.0191	0.0133	0.0610	0.0355	0.0288	0.0183
			0.5 FD	0.0286	0.1640	0.0933	0.0758	0.1105	0.0162	0.0910	0.0555	0.0512	0.0356	0.0120	0.0625	0.0375	0.0313	0.0154
			0.5 SYS	0.0216	0.1030	0.0505	0.0298	0.0210	0.0152	0.0780	0.0445	0.0348	0.0180	0.0113	0.0575	0.0340	0.0263	0.0168
			1 FD	0.0255	0.1405	0.0780	0.0700	0.0779	0.0122	0.0660	0.0343	0.0298	0.0385	0.0087	0.0490	0.0275	0.0228	0.0225
			1 SYS	0.0178	0.0855	0.0388	0.0262	0.0164	0.0112	0.0595	0.0305	0.0222	0.0153	0.0085	0.0465	0.0268	0.0205	0.0146
			2 FD	0.0173	0.0990	0.0620	0.0525	0.0470	0.0074	0.0380	0.0200	0.0140	0.0144	0.0053	0.0285	0.0160	0.0115	0.0131
			2 SYS	0.0121	0.0600	0.0295	0.0173	0.0115	0.0069	0.0345	0.0173	0.0110	0.0093	0.0052	0.0280	0.0150	0.0107	0.0088
			5 FD	0.0070	0.0425	0.0255	0.0208	0.0174	0.0031	0.0145	0.0073	0.0043	0.0040	0.0021	0.0115	0.0060	0.0040	0.0034
			5 SYS	0.0057	0.0320	0.0160	0.0102	0.0076	0.0030	0.0135	0.0070	0.0040	0.0033	0.0021	0.0115	0.0058	0.0037	0.0030
			10 FD	0.0035	0.0195	0.0110	0.0085	0.0069	0.0015	0.0070	0.0035	0.0022	0.0018	0.0011	0.0055	0.0028	0.0017	0.0013
			10 SYS	0.0032	0.0165	0.0088	0.0060	0.0048	0.0015	0.0075	0.0035	0.0020	0.0016	0.0011	0.0055	0.0028	0.0017	0.0013

Table IV-7: Relative standard deviation of $\hat{\gamma}$ over N, T, sw and γ

5.1.2 Relative Bias and Standard Deviation of β (Short-Run Effects)

				T														
				5					10					20				
				γ					γ					γ				
N	sw			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
50	sw	0.1	FD	0.0703	0.1047	0.0574	-0.4923	-0.7561	0.0615	0.0352	-0.0682	-0.5819	-0.4945	0.0126	-0.0731	-0.2001	-0.3963	-0.0832
			SYS	0.1022	0.1692	0.1841	0.0352	-0.2060	0.0509	0.1585	0.2350	-0.0174	-0.7431	0.1508	0.2473	0.3583	0.3890	-0.5898
		0.2	FD	-0.0206	0.0739	-0.0265	-0.3080	-0.4294	-0.0062	0.0086	-0.0874	-0.3544	-0.2724	-0.0079	-0.0831	-0.2068	-0.3734	-0.0544
			SYS	0.0443	0.1376	0.0889	-0.0693	-0.2267	0.0197	0.0951	0.0893	-0.1175	-0.6157	0.0759	0.0935	0.1042	0.0781	-0.3895
		0.5	FD	0.0026	0.0112	-0.0417	-0.1806	-0.1415	-0.0031	-0.0207	-0.0754	-0.2020	-0.1352	-0.0060	-0.0802	-0.1664	-0.2869	-0.0196
			SYS	0.0209	0.0280	-0.0315	-0.1404	-0.2194	-0.0010	-0.0012	-0.0436	-0.1743	-0.3608	0.0148	-0.0368	-0.0841	-0.1379	-0.1486
		1	FD	0.0004	-0.0010	-0.0459	-0.0783	-0.0227	-0.0038	-0.0180	-0.0480	-0.0975	-0.1307	-0.0048	-0.0501	-0.1001	-0.1696	-0.0489
			SYS	0.0062	-0.0070	-0.0542	-0.0968	-0.1521	-0.0046	-0.0176	-0.0467	-0.1076	-0.2083	0.0004	-0.0359	-0.0741	-0.1220	-0.1032
		2	FD	0.0009	-0.0068	-0.0243	-0.0253	0.0102	-0.0010	-0.0092	-0.0199	-0.0382	-0.0579	-0.0013	-0.0199	-0.0401	-0.0650	-0.0639
			SYS	0.0028	-0.0085	-0.0315	-0.0560	-0.0880	-0.0014	-0.0086	-0.0196	-0.0399	-0.0775	0.0005	-0.0140	-0.0307	-0.0507	-0.0664
		5	FD	-0.0001	-0.0041	-0.0066	-0.0033	0.0093	-0.0003	-0.0025	-0.0077	-0.0107	-0.0148	-0.0003	-0.0041	-0.0084	-0.0129	-0.0153
			SYS	0.0002	-0.0034	-0.0095	-0.0173	-0.0308	-0.0003	-0.0014	-0.0051	-0.0087	-0.0143	0.0002	-0.0028	-0.0060	-0.0097	-0.0141
		10	FD	0.0001	-0.0014	-0.0019	-0.0012	0.0038	-0.0001	-0.0015	-0.0030	-0.0046	-0.0061	0.0000	-0.0011	-0.0023	-0.0036	-0.0043
			SYS	0.0001	-0.0008	-0.0027	-0.0068	-0.0097	-0.0002	-0.0007	-0.0018	-0.0029	-0.0045	0.0000	-0.0006	-0.0015	-0.0025	-0.0035
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
250	sw	0.1	FD	0.0020	0.1853	0.0031	-0.5171	-0.8122	0.0113	0.0502	-0.0642	-0.5119	-0.3578	0.0005	-0.0810	-0.2098	-0.3793	-0.0818
			SYS	0.0039	0.0998	0.1056	-0.0190	-0.1776	0.0259	0.1336	0.1951	-0.1778	-0.8744	0.0175	0.0700	0.1250	0.0213	-0.9020
		0.2	FD	-0.0043	0.0733	-0.0121	-0.2755	-0.3922	-0.0030	0.0095	-0.0808	-0.3256	-0.1709	-0.0065	-0.0899	-0.1979	-0.3414	-0.0225
			SYS	-0.0025	0.0830	0.0618	-0.0748	-0.2104	0.0078	0.0782	0.0518	-0.2356	-0.7086	0.0024	-0.0163	-0.0502	-0.1503	-0.5682
		0.5	FD	0.0051	0.0088	-0.0443	-0.1635	-0.0934	-0.0005	-0.0142	-0.0663	-0.1879	-0.0793	-0.0006	-0.0770	-0.1645	-0.2734	0.0108
			SYS	0.0028	0.0176	-0.0275	-0.1250	-0.1964	0.0012	0.0038	-0.0514	-0.1974	-0.3877	0.0001	-0.0572	-0.1226	-0.1976	-0.2003
		1	FD	0.0007	-0.0005	-0.0381	-0.0857	-0.0039	-0.0006	-0.0140	-0.0436	-0.0968	-0.1186	-0.0008	-0.0478	-0.0998	-0.1647	-0.0100
			SYS	-0.0006	-0.0063	-0.0408	-0.0930	-0.1431	-0.0003	-0.0083	-0.0390	-0.1080	-0.2073	-0.0011	-0.0384	-0.0790	-0.1335	-0.1122
		2	FD	0.0012	-0.0061	-0.0230	-0.0291	0.0217	-0.0006	-0.0070	-0.0209	-0.0395	-0.0617	-0.0006	-0.0190	-0.0403	-0.0644	-0.0566
			SYS	0.0006	-0.0080	-0.0302	-0.0566	-0.0797	-0.0005	-0.0033	-0.0157	-0.0370	-0.0728	-0.0007	-0.0136	-0.0274	-0.0464	-0.0633
		5	FD	-0.0003	-0.0041	-0.0067	-0.0020	0.0117	0.0000	-0.0028	-0.0071	-0.0122	-0.0163	-0.0002	-0.0037	-0.0081	-0.0135	-0.0177
			SYS	-0.0003	-0.0032	-0.0100	-0.0188	-0.0304	0.0001	-0.0007	-0.0033	-0.0080	-0.0127	-0.0002	-0.0023	-0.0044	-0.0079	-0.0127
		10	FD	0.0001	-0.0015	-0.0020	-0.0001	0.0047	-0.0001	-0.0014	-0.0031	-0.0052	-0.0063	0.0000	-0.0010	-0.0023	-0.0041	-0.0060
			SYS	0.0001	-0.0010	-0.0032	-0.0062	-0.0102	-0.0001	-0.0003	-0.0010	-0.0024	-0.0035	0.0000	-0.0006	-0.0011	-0.0021	-0.0032
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
500	sw	0.1	FD	0.0282	0.1711	0.0085	-0.4987	-0.7841	-0.0064	0.0526	-0.0791	-0.4915	-0.3353	-0.0080	-0.0768	-0.2057	-0.3702	-0.0767
			SYS	0.0213	0.1143	0.1181	-0.0159	-0.1673	-0.0058	0.1107	0.1453	-0.2565	-0.9187	0.0180	0.0719	0.0987	-0.0592	-0.9749
		0.2	FD	-0.0001	0.0591	-0.0276	-0.3192	-0.3514	-0.0020	0.0031	-0.0797	-0.3143	-0.1617	0.0011	-0.0829	-0.1979	-0.3360	-0.0151
			SYS	0.0162	0.0939	0.0611	-0.0869	-0.2035	-0.0037	0.0596	0.0254	-0.2811	-0.7402	0.0096	-0.0130	-0.0692	-0.1917	-0.6001
		0.5	FD	-0.0009	0.0108	-0.0466	-0.1648	-0.0769	-0.0011	-0.0139	-0.0680	-0.1854	-0.0789	-0.0021	-0.0757	-0.1609	-0.2730	0.0158
			SYS	0.0050	0.0242	-0.0279	-0.1240	-0.2034	-0.0023	0.0018	-0.0582	-0.2077	-0.4002	-0.0002	-0.0567	-0.1243	-0.2084	-0.2097
		1	FD	-0.0002	-0.0032	-0.0392	-0.0875	0.0056	-0.0012	-0.0137	-0.0435	-0.0972	-0.1174	-0.0009	-0.0481	-0.0972	-0.1616	-0.0067
			SYS	0.0004	-0.0073	-0.0425	-0.0948	-0.1427	-0.0018	-0.0083	-0.0395	-0.1092	-0.2098	-0.0007	-0.0378	-0.0768	-0.1326	-0.1161
		2	FD	0.0003	-0.0081	-0.0240	-0.0285	0.0232	-0.0003	-0.0076	-0.0207	-0.0403	-0.0628	-0.0004	-0.0187	-0.0398	-0.0630	-0.0564
			SYS	0.0004	-0.0098	-0.0299	-0.0555	-0.0809	-0.0005	-0.0037	-0.0152	-0.0371	-0.0732	-0.0003	-0.0131	-0.0264	-0.0447	-0.0624
		5	FD	0.0001	-0.0034	-0.0065	-0.0021	0.0124	0.0001	-0.0031	-0.0073	-0.0124	-0.0164	0.0000	-0.0038	-0.0082	-0.0139	-0.0177
			SYS	0.0000	-0.0024	-0.0092	-0.0184	-0.0305	0.0001	-0.0009	-0.0033	-0.0078	-0.0126	0.0000	-0.0023	-0.0044	-0.0075	-0.0120
		10	FD	0.0000	-0.0014	-0.0020	-0.0003	0.0049	0.0000	-0.0012	-0.0031	-0.0051	-0.0065	0.0000	-0.0010	-0.0024	-0.0042	-0.0062
			SYS	0.0000	-0.0008	-0.0029	-0.0063	-0.0103	0.0000	-0.0002	-0.0009	-0.0022	-0.0035	0.0000	-0.0005	-0.0012	-0.0019	-0.0030

Table IV-8: Relative bias $\hat{\beta} - \beta$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
50	sw	0.1	FD	1.0324	1.0489	1.1201	1.1419	1.5582	0.5661	0.5874	0.5781	0.5940	0.8506	0.3262	0.3418	0.3522	0.3362	0.5056
			SYS	0.3515	0.3664	0.3813	0.3581	0.3336	0.2528	0.2499	0.2574	0.3015	0.2277	0.1824	0.1849	0.2064	0.2243	0.2251
		0.2	FD	0.4833	0.5237	0.5406	0.5572	0.7876	0.2865	0.2925	0.2961	0.3105	0.4260	0.1744	0.1739	0.1755	0.1697	0.2731
			SYS	0.2960	0.2938	0.3104	0.2862	0.2907	0.1987	0.1985	0.2066	0.2268	0.1870	0.1303	0.1334	0.1430	0.1676	0.1802
		0.5	FD	0.1980	0.2024	0.2241	0.2569	0.3559	0.1138	0.1205	0.1191	0.1235	0.1832	0.0675	0.0724	0.0737	0.0749	0.1042
			SYS	0.1649	0.1731	0.1812	0.2070	0.1932	0.1046	0.1061	0.1052	0.1102	0.1161	0.0624	0.0665	0.0686	0.0763	0.0875
		1	FD	0.0986	0.1026	0.1098	0.1268	0.1963	0.0570	0.0577	0.0608	0.0610	0.0834	0.0360	0.0384	0.0391	0.0432	0.0579
			SYS	0.0909	0.0922	0.1015	0.1089	0.1166	0.0546	0.0549	0.0561	0.0555	0.0614	0.0341	0.0364	0.0373	0.0424	0.0493
		2	FD	0.0473	0.0491	0.0551	0.0641	0.0968	0.0286	0.0285	0.0309	0.0318	0.0348	0.0186	0.0191	0.0204	0.0220	0.0272
			SYS	0.0441	0.0474	0.0535	0.0576	0.0673	0.0273	0.0267	0.0277	0.0286	0.0311	0.0178	0.0178	0.0196	0.0210	0.0244
		5	FD	0.0177	0.0190	0.0221	0.0271	0.0341	0.0106	0.0109	0.0113	0.0115	0.0123	0.0074	0.0074	0.0074	0.0083	0.0099
			SYS	0.0173	0.0189	0.0224	0.0256	0.0291	0.0103	0.0101	0.0101	0.0101	0.0110	0.0069	0.0070	0.0073	0.0080	0.0094
		10	FD	0.0091	0.0095	0.0108	0.0131	0.0168	0.0051	0.0052	0.0052	0.0054	0.0058	0.0036	0.0038	0.0038	0.0042	0.0046
			SYS	0.0090	0.0095	0.0113	0.0134	0.0169	0.0050	0.0050	0.0049	0.0050	0.0054	0.0034	0.0036	0.0037	0.0041	0.0046
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
250	sw	0.1	FD	0.4503	0.4506	0.4861	0.5031	0.6718	0.2471	0.2528	0.2578	0.2789	0.3927	0.1465	0.1510	0.1436	0.1507	0.2388
			SYS	0.1292	0.1288	0.1391	0.1392	0.1314	0.1086	0.1112	0.1429	0.1580	0.1206	0.0927	0.0948	0.1110	0.1205	0.1073
		0.2	FD	0.2103	0.2240	0.2345	0.2541	0.3547	0.1199	0.1198	0.1242	0.1366	0.2010	0.0734	0.0750	0.0759	0.0781	0.1248
			SYS	0.1127	0.1149	0.1271	0.1271	0.1200	0.0873	0.0890	0.1068	0.1083	0.0926	0.0632	0.0651	0.0757	0.0770	0.0808
		0.5	FD	0.0861	0.0871	0.0939	0.1044	0.1571	0.0479	0.0488	0.0498	0.0542	0.0844	0.0307	0.0320	0.0326	0.0370	0.0484
			SYS	0.0713	0.0721	0.0788	0.0827	0.0811	0.0442	0.0452	0.0486	0.0511	0.0503	0.0297	0.0316	0.0327	0.0348	0.0421
		1	FD	0.0419	0.0437	0.0471	0.0563	0.0865	0.0229	0.0242	0.0269	0.0283	0.0364	0.0160	0.0171	0.0188	0.0201	0.0248
			SYS	0.0384	0.0389	0.0444	0.0489	0.0512	0.0220	0.0232	0.0251	0.0272	0.0268	0.0158	0.0171	0.0187	0.0194	0.0217
		2	FD	0.0203	0.0215	0.0233	0.0297	0.0404	0.0115	0.0124	0.0129	0.0130	0.0147	0.0076	0.0086	0.0095	0.0103	0.0128
			SYS	0.0189	0.0200	0.0230	0.0272	0.0285	0.0109	0.0114	0.0118	0.0117	0.0132	0.0074	0.0082	0.0089	0.0096	0.0112
		5	FD	0.0077	0.0080	0.0091	0.0119	0.0149	0.0045	0.0046	0.0050	0.0053	0.0054	0.0032	0.0034	0.0035	0.0039	0.0045
			SYS	0.0073	0.0079	0.0094	0.0116	0.0134	0.0044	0.0042	0.0045	0.0045	0.0046	0.0032	0.0033	0.0034	0.0036	0.0041
		10	FD	0.0038	0.0040	0.0046	0.0058	0.0072	0.0021	0.0022	0.0023	0.0025	0.0025	0.0016	0.0016	0.0017	0.0018	0.0020
			SYS	0.0037	0.0040	0.0048	0.0059	0.0072	0.0020	0.0022	0.0022	0.0023	0.0022	0.0016	0.0017	0.0016	0.0018	0.0019
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
500	sw	0.1	FD	0.3045	0.3190	0.3210	0.3679	0.5002	0.1626	0.1683	0.1860	0.1912	0.2810	0.1033	0.1034	0.1082	0.1068	0.1716
			SYS	0.0894	0.0843	0.0967	0.1030	0.1005	0.0738	0.0712	0.1041	0.1143	0.0860	0.0659	0.0729	0.0831	0.0902	0.0818
		0.2	FD	0.1565	0.1494	0.1679	0.1866	0.2697	0.0827	0.0846	0.0908	0.0930	0.1408	0.0498	0.0543	0.0553	0.0561	0.0898
			SYS	0.0813	0.0782	0.0952	0.0997	0.0867	0.0586	0.0619	0.0756	0.0738	0.0656	0.0426	0.0486	0.0552	0.0542	0.0591
		0.5	FD	0.0608	0.0621	0.0660	0.0747	0.1192	0.0330	0.0336	0.0356	0.0393	0.0594	0.0215	0.0226	0.0240	0.0250	0.0343
			SYS	0.0498	0.0520	0.0578	0.0587	0.0597	0.0295	0.0315	0.0342	0.0363	0.0368	0.0208	0.0223	0.0249	0.0245	0.0299
		1	FD	0.0309	0.0309	0.0332	0.0382	0.0610	0.0164	0.0171	0.0184	0.0188	0.0273	0.0109	0.0119	0.0131	0.0148	0.0174
			SYS	0.0278	0.0285	0.0313	0.0343	0.0381	0.0156	0.0163	0.0178	0.0181	0.0199	0.0107	0.0116	0.0129	0.0143	0.0150
		2	FD	0.0147	0.0144	0.0164	0.0220	0.0300	0.0082	0.0084	0.0090	0.0096	0.0112	0.0055	0.0059	0.0067	0.0075	0.0090
			SYS	0.0134	0.0134	0.0162	0.0192	0.0210	0.0078	0.0079	0.0081	0.0083	0.0101	0.0054	0.0058	0.0065	0.0070	0.0075
		5	FD	0.0053	0.0058	0.0068	0.0081	0.0107	0.0030	0.0032	0.0035	0.0035	0.0038	0.0021	0.0024	0.0025	0.0027	0.0031
			SYS	0.0050	0.0056	0.0070	0.0077	0.0095	0.0029	0.0030	0.0031	0.0031	0.0033	0.0021	0.0024	0.0024	0.0024	0.0029
		10	FD	0.0026	0.0027	0.0031	0.0042	0.0049	0.0015	0.0015	0.0016	0.0017	0.0018	0.0011	0.0011	0.0012	0.0013	0.0014
			SYS	0.0025	0.0027	0.0033	0.0042	0.0050	0.0014	0.0015	0.0015	0.0015	0.0016	0.0011	0.0011	0.0012	0.0013	0.0013

Table IV-9: Relative standard deviation of $\hat{\beta}$ over N, T, sw and γ

5.1.3 Relative Bias and Standard Deviation of β* (Long-Run Effects)

				T														
				5					10					20				
				γ					γ					γ				
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	50	sw 0.1	FD	0.0278	-0.0734	-0.2864	-0.7476	9.7540	0.0532	-0.1467	-0.4138	-0.8040	######	0.0113	-0.2473	-0.5180	-0.7629	0.9364
			SYS	0.0891	-0.0058	-0.1612	-0.3554	-0.5300	0.0585	-0.0415	-0.1995	-0.4800	-0.8355	0.1548	0.0178	-0.1551	-0.3720	-0.7655
		0.2	FD	-0.0283	-0.0846	-0.3078	-0.5258	-2.2780	-0.0159	-0.1617	-0.4094	-0.6900	0.5212	-0.0083	-0.2509	-0.5141	-0.7514	1.1210
			SYS	0.0446	-0.0231	-0.1927	-0.3845	-0.5394	0.0234	-0.0838	-0.2697	-0.5173	-0.7672	0.0852	-0.0929	-0.2893	-0.4870	-0.6663
		0.5	FD	-0.0089	-0.1111	-0.2735	-0.3381	-1.0284	-0.0084	-0.1464	-0.3355	-0.5718	1.3370	-0.0089	-0.2214	-0.4498	-0.6878	0.4587
			SYS	0.0251	-0.0775	-0.2222	-0.3813	-0.5161	0.0048	-0.1234	-0.2948	-0.5066	-0.6449	0.0207	-0.1710	-0.3692	-0.5569	-0.5565
		1	FD	-0.0012	-0.0821	-0.1836	0.0078	0.3338	-0.0061	-0.0933	-0.2105	-0.3689	-0.2422	-0.0070	-0.1421	-0.3103	-0.5278	-0.0785
			SYS	0.0154	-0.0674	-0.1760	-0.2828	-0.4416	-0.0020	-0.0861	-0.1943	-0.3620	-0.5553	0.0024	-0.1208	-0.2762	-0.4722	-0.5326
		2	FD	0.0025	-0.0339	-0.0552	0.1394	-1.0466	-0.0030	-0.0385	-0.0811	-0.1474	-0.1962	-0.0019	-0.0582	-0.1382	-0.2573	-0.3448
			SYS	0.0077	-0.0398	-0.0985	-0.1628	-0.3173	-0.0012	-0.0325	-0.0677	-0.1360	-0.3215	0.0010	-0.0475	-0.1194	-0.2440	-0.4334
		5	FD	-0.0011	-0.0032	0.0115	0.0684	0.4834	-0.0011	-0.0069	-0.0185	-0.0276	-0.0356	-0.0006	-0.0107	-0.0280	-0.0508	-0.0837
			SYS	0.0005	-0.0116	-0.0292	-0.0493	-0.1126	-0.0004	-0.0057	-0.0131	-0.0206	-0.0643	0.0003	-0.0082	-0.0227	-0.0493	-0.1268
		10	FD	0.0002	0.0014	0.0082	0.0310	0.1131	-0.0001	-0.0024	-0.0050	-0.0076	-0.0069	-0.0002	-0.0029	-0.0071	-0.0130	-0.0158
			SYS	0.0008	-0.0027	-0.0091	-0.0184	-0.0261	0.0000	-0.0018	-0.0032	-0.0049	-0.0102	0.0000	-0.0019	-0.0057	-0.0128	-0.0257
	250	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
		0.1	FD	-0.0044	0.0252	-0.2780	-0.7096	-0.8487	0.0106	-0.1269	-0.4095	-0.7655	2.3167	0.0013	-0.2528	-0.5213	-0.7520	3.3137
			SYS	0.0046	-0.0521	-0.1873	-0.3600	-0.5023	0.0293	-0.0602	-0.2202	-0.5620	-0.9114	0.0208	-0.1247	-0.2946	-0.5494	-0.9458
		0.2	FD	-0.0065	-0.0668	-0.2707	-0.5335	0.4931	-0.0045	-0.1525	-0.4048	-0.6774	1.8487	-0.0054	-0.2538	-0.5071	-0.7348	4.0238
			SYS	-0.0010	-0.0603	-0.1991	-0.3728	-0.5137	0.0106	-0.0962	-0.2926	-0.5840	-0.8119	0.0063	-0.1872	-0.3934	-0.6165	-0.7761
		0.5	FD	0.0035	-0.1013	-0.2530	-0.3839	-2.8544	-0.0008	-0.1384	-0.3275	-0.5648	######	0.0001	-0.2153	-0.4448	-0.6766	2.7874
			SYS	0.0060	-0.0884	-0.2164	-0.3621	-0.4950	0.0035	-0.1232	-0.3008	-0.5220	-0.6522	0.0024	-0.1942	-0.4012	-0.6059	-0.6077
		1	FD	-0.0004	-0.0774	-0.1711	-0.1774	-4.5217	-0.0017	-0.0872	-0.2030	-0.3616	-0.1551	-0.0005	-0.1382	-0.3071	-0.5178	0.4260
			SYS	0.0016	-0.0782	-0.1716	-0.2818	-0.4258	0.0004	-0.0794	-0.1838	-0.3473	-0.5362	-0.0001	-0.1268	-0.2809	-0.4796	-0.5463
		2	FD	-0.0002	-0.0325	-0.0515	0.0421	1.1042	-0.0010	-0.0351	-0.0838	-0.1454	-0.2031	-0.0003	-0.0558	-0.1357	-0.2503	-0.2889
			SYS	0.0008	-0.0461	-0.1054	-0.1740	-0.3069	-0.0003	-0.0292	-0.0656	-0.1236	-0.2888	-0.0001	-0.0482	-0.1130	-0.2265	-0.4087
		5	FD	-0.0003	-0.0022	0.0144	0.0811	0.3834	-0.0002	-0.0070	-0.0178	-0.0295	-0.0399	-0.0002	-0.0102	-0.0262	-0.0482	-0.0716
			SYS	-0.0001	-0.0148	-0.0345	-0.0598	-0.1334	0.0001	-0.0051	-0.0114	-0.0198	-0.0622	-0.0003	-0.0086	-0.0196	-0.0427	-0.1071
		10	FD	0.0001	0.0013	0.0103	0.0354	0.1173	-0.0002	-0.0023	-0.0055	-0.0091	-0.0096	0.0000	-0.0026	-0.0068	-0.0123	-0.0136
			SYS	0.0002	-0.0044	-0.0109	-0.0203	-0.0411	-0.0002	-0.0015	-0.0031	-0.0049	-0.0141	0.0000	-0.0022	-0.0052	-0.0112	-0.0229
	500	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
		0.1	FD	0.0249	0.0175	-0.2672	-0.6952	-0.7114	-0.0070	-0.1242	-0.4145	-0.7531	3.0029	-0.0076	-0.2494	-0.5194	-0.7486	3.0804
			SYS	0.0213	-0.0367	-0.1752	-0.3599	-0.4996	-0.0042	-0.0786	-0.2492	-0.6046	-0.9429	0.0201	-0.1238	-0.3118	-0.5894	-0.9864
		0.2	FD	-0.0004	-0.0735	-0.2842	-0.5635	0.2432	-0.0025	-0.1566	-0.4012	-0.6675	2.7307	0.0006	-0.2494	-0.5084	-0.7326	3.6082
			SYS	0.0162	-0.0478	-0.1997	-0.3807	-0.5146	-0.0022	-0.1123	-0.3098	-0.6109	-0.8342	0.0111	-0.1864	-0.4073	-0.6397	-0.7961
		0.5	FD	-0.0013	-0.0995	-0.2533	-0.3892	-3.3357	-0.0018	-0.1366	-0.3262	-0.5601	2.2484	-0.0020	-0.2144	-0.4426	-0.6770	3.0339
			SYS	0.0065	-0.0830	-0.2151	-0.3624	-0.5013	-0.0019	-0.1261	-0.3071	-0.5318	-0.6640	0.0012	-0.1947	-0.4038	-0.6152	-0.6196
		1	FD	0.0001	-0.0802	-0.1749	-0.1956	-7.3288	-0.0018	-0.0865	-0.2043	-0.3626	-0.1493	-0.0009	-0.1384	-0.3035	-0.5141	0.5330
			SYS	0.0019	-0.0782	-0.1727	-0.2859	-0.4300	-0.0017	-0.0814	-0.1884	-0.3508	-0.5400	-0.0002	-0.1264	-0.2767	-0.4779	-0.5505
		2	FD	0.0001	-0.0354	-0.0532	0.0413	-42.5470	-0.0004	-0.0356	-0.0842	-0.1476	-0.2087	-0.0004	-0.0558	-0.1348	-0.2467	-0.2828
			SYS	0.0014	-0.0476	-0.1049	-0.1718	-0.3096	-0.0003	-0.0302	-0.0668	-0.1265	-0.2904	0.0001	-0.0481	-0.1105	-0.2205	-0.4027
		5	FD	0.0000	-0.0013	0.0135	0.0820	0.3902	0.0003	-0.0074	-0.0179	-0.0302	-0.0406	-0.0001	-0.0104	-0.0265	-0.0475	-0.0696
			SYS	0.0002	-0.0134	-0.0339	-0.0604	-0.1332	0.0003	-0.0057	-0.0117	-0.0204	-0.0621	0.0000	-0.0088	-0.0198	-0.0407	-0.1031
		10	FD	0.0001	0.0011	0.0101	0.0346	0.1201	0.0000	-0.0022	-0.0055	-0.0090	-0.0094	-0.0001	-0.0027	-0.0068	-0.0119	-0.0121
			SYS	0.0002	-0.0042	-0.0107	-0.0211	-0.0407	0.0000	-0.0014	-0.0031	-0.0049	-0.0133	-0.0001	-0.0022	-0.0051	-0.0104	-0.0211

Table IV-10: Relative bias $\hat{\beta}^* - \beta^*$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	50	0.1	FD	1.0031	0.9059	0.8136	0.9802	263.6072	0.5588	0.4796	0.3723	0.2940	1273.4063	0.3296	0.2835	0.2158	0.1363	4.8041
			SYS	0.3334	0.2843	0.2362	0.1957	0.1762	0.2519	0.1955	0.1529	0.1381	0.1484	0.1796	0.1504	0.1253	0.1009	0.1268
		0.2	FD	0.4711	0.4369	0.3970	1.5965	21.6510	0.2791	0.2427	0.1966	0.1630	11.0716	0.1838	0.1505	0.1150	0.0732	4.1517
			SYS	0.2832	0.2276	0.2114	0.1814	0.1520	0.1938	0.1621	0.1328	0.1177	0.1100	0.1367	0.1154	0.0974	0.0860	0.0969
		0.5	FD	0.1889	0.1874	0.2007	2.2085	23.2058	0.1194	0.1159	0.1008	0.0925	33.5128	0.0857	0.0777	0.0649	0.0454	15.5495
			SYS	0.1651	0.1556	0.1583	0.1617	0.1343	0.1085	0.1018	0.0890	0.0775	0.0728	0.0791	0.0702	0.0622	0.0541	0.0597
		1	FD	0.1058	0.1212	0.1840	3.4198	15.1191	0.0652	0.0671	0.0709	0.0762	1.6488	0.0520	0.0524	0.0487	0.0445	0.3413
			SYS	0.1031	0.1061	0.1193	0.1236	0.1188	0.0615	0.0637	0.0647	0.0676	0.0649	0.0485	0.0494	0.0484	0.0448	0.0528
		2	FD	0.0642	0.0846	0.1401	1.0127	33.3723	0.0352	0.0389	0.0442	0.0513	0.1503	0.0307	0.0316	0.0358	0.0409	0.0902
			SYS	0.0615	0.0706	0.0849	0.0969	0.1152	0.0339	0.0362	0.0388	0.0457	0.0731	0.0284	0.0295	0.0343	0.0392	0.0481
		5	FD	0.0264	0.0402	0.0681	0.1330	0.8533	0.0144	0.0156	0.0170	0.0214	0.0496	0.0119	0.0135	0.0148	0.0186	0.0405
			SYS	0.0260	0.0353	0.0483	0.0652	0.1007	0.0141	0.0147	0.0158	0.0203	0.0428	0.0112	0.0128	0.0145	0.0182	0.0359
		10	FD	0.0140	0.0194	0.0308	0.0530	0.1263	0.0068	0.0071	0.0079	0.0098	0.0227	0.0059	0.0066	0.0074	0.0096	0.0187
			SYS	0.0142	0.0182	0.0270	0.0396	0.0746	0.0068	0.0070	0.0077	0.0098	0.0217	0.0055	0.0062	0.0072	0.0096	0.0190
	250	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
		0.1	FD	0.4359	0.3793	0.3440	0.3078	5.0496	0.2446	0.2088	0.1616	0.1341	37.9834	0.1480	0.1242	0.0888	0.0616	11.7058
			SYS	0.1248	0.1033	0.0920	0.0844	0.0609	0.1094	0.0885	0.0881	0.0760	0.0861	0.0946	0.0775	0.0704	0.0530	0.0597
		0.2	FD	0.2029	0.1870	0.1701	0.1714	121.9416	0.1175	0.1004	0.0814	0.0674	33.4976	0.0766	0.0650	0.0496	0.0343	8.2253
			SYS	0.1094	0.0934	0.0903	0.0842	0.0669	0.0872	0.0746	0.0720	0.0576	0.0589	0.0666	0.0573	0.0508	0.0374	0.0416
		0.5	FD	0.0816	0.0791	0.0829	0.1141	76.5042	0.0515	0.0479	0.0449	0.0402	1406.5752	0.0381	0.0338	0.0283	0.0227	2.3847
			SYS	0.0697	0.0659	0.0691	0.0690	0.0582	0.0472	0.0450	0.0440	0.0373	0.0344	0.0360	0.0331	0.0286	0.0225	0.0293
		1	FD	0.0435	0.0494	0.0677	0.1614	70.4954	0.0266	0.0279	0.0308	0.0352	0.2033	0.0235	0.0239	0.0233	0.0220	0.2989
			SYS	0.0403	0.0447	0.0497	0.0538	0.0534	0.0253	0.0273	0.0298	0.0328	0.0315	0.0231	0.0237	0.0232	0.0219	0.0246
		2	FD	0.0276	0.0368	0.0592	0.1498	47.0467	0.0149	0.0169	0.0184	0.0240	0.0592	0.0123	0.0147	0.0161	0.0204	0.0459
			SYS	0.0253	0.0294	0.0356	0.0422	0.0499	0.0143	0.0154	0.0163	0.0207	0.0318	0.0120	0.0140	0.0152	0.0186	0.0225
		5	FD	0.0117	0.0173	0.0284	0.0580	0.1957	0.0061	0.0067	0.0070	0.0095	0.0222	0.0053	0.0059	0.0070	0.0090	0.0178
			SYS	0.0110	0.0146	0.0194	0.0280	0.0403	0.0059	0.0062	0.0067	0.0090	0.0171	0.0053	0.0058	0.0067	0.0084	0.0149
		10	FD	0.0060	0.0086	0.0129	0.0233	0.0546	0.0029	0.0031	0.0034	0.0042	0.0097	0.0027	0.0029	0.0033	0.0042	0.0076
			SYS	0.0059	0.0082	0.0112	0.0171	0.0300	0.0029	0.0030	0.0033	0.0042	0.0091	0.0027	0.0030	0.0032	0.0041	0.0075
	500	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
		0.1	FD	0.2946	0.2682	0.2260	0.2256	6.0637	0.1608	0.1382	0.1168	0.0913	8.3338	0.1045	0.0853	0.0670	0.0434	1.4000
			SYS	0.0857	0.0664	0.0642	0.0624	0.0484	0.0746	0.0570	0.0654	0.0565	0.0607	0.0672	0.0601	0.0529	0.0397	0.0450
		0.2	FD	0.1498	0.1241	0.1222	0.1231	10.9361	0.0816	0.0710	0.0598	0.0469	161.1780	0.0519	0.0469	0.0362	0.0242	1.8222
			SYS	0.0772	0.0636	0.0677	0.0656	0.0467	0.0586	0.0514	0.0512	0.0392	0.0417	0.0445	0.0425	0.0376	0.0256	0.0298
		0.5	FD	0.0563	0.0566	0.0579	0.0792	76.2657	0.0342	0.0329	0.0311	0.0286	6.6179	0.0261	0.0240	0.0208	0.0153	1.2353
			SYS	0.0487	0.0485	0.0513	0.0486	0.0445	0.0310	0.0308	0.0299	0.0262	0.0249	0.0250	0.0234	0.0215	0.0159	0.0201
		1	FD	0.0317	0.0360	0.0449	0.0925	212.4709	0.0189	0.0207	0.0215	0.0245	0.1524	0.0160	0.0167	0.0163	0.0155	0.2291
			SYS	0.0298	0.0319	0.0345	0.0396	0.0404	0.0178	0.0192	0.0210	0.0226	0.0220	0.0158	0.0162	0.0165	0.0153	0.0172
		2	FD	0.0193	0.0248	0.0428	0.1054	1245.6595	0.0102	0.0118	0.0137	0.0176	0.0419	0.0092	0.0100	0.0116	0.0133	0.0328
			SYS	0.0179	0.0200	0.0251	0.0297	0.0348	0.0096	0.0109	0.0123	0.0149	0.0234	0.0089	0.0098	0.0110	0.0128	0.0157
		5	FD	0.0082	0.0123	0.0206	0.0420	0.1466	0.0041	0.0045	0.0052	0.0065	0.0155	0.0036	0.0042	0.0049	0.0062	0.0127
			SYS	0.0078	0.0108	0.0146	0.0187	0.0297	0.0040	0.0042	0.0047	0.0061	0.0126	0.0035	0.0042	0.0046	0.0056	0.0106
		10	FD	0.0042	0.0058	0.0091	0.0163	0.0385	0.0020	0.0022	0.0025	0.0031	0.0069	0.0018	0.0020	0.0023	0.0028	0.0055
			SYS	0.0041	0.0054	0.0079	0.0118	0.0211	0.0020	0.0022	0.0025	0.0031	0.0068	0.0018	0.0021	0.0023	0.0028	0.0053

Table IV-11: Relative standard deviation of $\hat{\beta}^*$ over N, T, sw and γ

5.2 Results with All Instruments

5.2.1 Relative Bias, Standard Deviation and RMSE of γ

					T														
					5					10					20				
					γ					γ					γ				
N	50	sw			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
			0.1	FD	-0.0183	-0.6920	-0.5770	-0.4047	-0.0418	-0.0183	-0.8735	-0.7898	-0.5663	-0.1085	-0.0267	-1.0355	-0.9680	-0.7953	-0.1723
				SYS	-0.0073	-0.6930	-0.5558	-0.3595	-0.1503	-0.0095	-0.8820	-0.7848	-0.5412	-0.1588	-0.0136	-0.9780	-0.9133	-0.7528	-0.2213
			0.2	FD	-0.0181	-0.6615	-0.5498	-0.3567	-0.0254	-0.0167	-0.8735	-0.7860	-0.5845	-0.1184	-0.0223	-0.9975	-0.9460	-0.7968	-0.1788
				SYS	-0.0057	-0.6230	-0.5045	-0.3187	-0.1445	-0.0065	-0.8455	-0.7473	-0.5328	-0.1833	-0.0083	-0.8920	-0.8548	-0.7118	-0.2374
			0.5	FD	-0.0096	-0.5850	-0.4645	-0.2688	0.0361	-0.0116	-0.6655	-0.6215	-0.5462	-0.1484	-0.0188	-0.8045	-0.7850	-0.7227	-0.1998
				SYS	0.0052	-0.4435	-0.3543	-0.2552	-0.1494	-0.0014	-0.6230	-0.5818	-0.4845	-0.2266	-0.0052	-0.7060	-0.6975	-0.6288	-0.2511
			1	FD	-0.0065	-0.3870	-0.3043	-0.1365	0.0728	-0.0095	-0.3535	-0.3110	-0.2752	-0.1373	-0.0098	-0.4590	-0.4718	-0.4808	-0.2061
				SYS	0.0067	-0.2760	-0.2228	-0.1768	-0.1274	-0.0038	-0.3140	-0.2943	-0.2890	-0.2146	-0.0037	-0.3990	-0.4235	-0.4397	-0.2355
			2	FD	-0.0040	-0.1700	-0.1070	-0.0140	0.0753	-0.0031	-0.1220	-0.0898	-0.0727	-0.0429	-0.0037	-0.1790	-0.1805	-0.1830	-0.1381
				SYS	0.0030	-0.1455	-0.1188	-0.0878	-0.0820	-0.0012	-0.1035	-0.0768	-0.0783	-0.0953	-0.0016	-0.1495	-0.1538	-0.1720	-0.1554
			5	FD	-0.0014	-0.0085	0.0025	0.0248	0.0405	-0.0007	-0.0205	-0.0140	-0.0093	-0.0029	-0.0011	-0.0345	-0.0335	-0.0318	-0.0298
				SYS	0.0002	-0.0355	-0.0350	-0.0270	-0.0246	-0.0005	-0.0180	-0.0105	-0.0073	-0.0124	-0.0008	-0.0265	-0.0263	-0.0290	-0.0350
			10	FD	0.0000	0.0025	0.0108	0.0150	0.0164	-0.0003	-0.0055	-0.0035	-0.0018	0.0000	-0.0003	-0.0070	-0.0085	-0.0082	-0.0070
				SYS	0.0008	-0.0110	-0.0085	-0.0072	-0.0061	-0.0002	-0.0045	-0.0025	-0.0012	-0.0010	-0.0002	-0.0040	-0.0065	-0.0075	-0.0073
	250	sw			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
			0.1	FD	-0.0039	-0.5435	-0.4573	-0.2988	-0.0045	-0.0044	-0.8045	-0.7843	-0.5422	-0.0516	-0.0049	-0.9245	-0.9025	-0.7435	-0.1535
				SYS	-0.0013	-0.5955	-0.4928	-0.3155	-0.1318	-0.0016	-0.8145	-0.7570	-0.5107	-0.1144	-0.0018	-0.9145	-0.8888	-0.7493	-0.2366
			0.2	FD	-0.0031	-0.5435	-0.4420	-0.2765	0.0238	-0.0053	-0.7815	-0.7630	-0.5775	-0.0643	-0.0038	-0.8990	-0.8868	-0.7468	-0.1604
				SYS	0.0002	-0.5710	-0.4485	-0.2862	-0.1333	-0.0016	-0.7580	-0.6920	-0.5073	-0.1470	0.0001	-0.8730	-0.8558	-0.7388	-0.2565
			0.5	FD	-0.0021	-0.4725	-0.3650	-0.2030	0.0490	-0.0036	-0.6015	-0.5828	-0.5193	-0.0945	-0.0039	-0.7355	-0.7490	-0.6828	-0.1771
				SYS	0.0015	-0.4460	-0.3385	-0.2312	-0.1360	0.0009	-0.5735	-0.5228	-0.4388	-0.1958	-0.0010	-0.7020	-0.7175	-0.6643	-0.2709
			1	FD	-0.0036	-0.3370	-0.2433	-0.1092	0.0885	-0.0003	-0.3030	-0.2590	-0.2222	-0.0910	-0.0021	-0.4450	-0.4605	-0.4577	-0.1773
				SYS	0.0000	-0.3020	-0.2263	-0.1673	-0.1174	0.0018	-0.3150	-0.2648	-0.2390	-0.1750	-0.0006	-0.4205	-0.4420	-0.4555	-0.2420
			2	FD	-0.0011	-0.1275	-0.0780	-0.0013	0.0905	-0.0004	-0.0945	-0.0705	-0.0492	-0.0219	-0.0005	-0.1720	-0.1763	-0.1767	-0.1203
				SYS	0.0003	-0.1555	-0.1230	-0.0897	-0.0760	0.0005	-0.1105	-0.0795	-0.0663	-0.0723	-0.0001	-0.1565	-0.1615	-0.1778	-0.1548
			5	FD	0.0002	0.0015	0.0135	0.0300	0.0474	0.0000	-0.0150	-0.0095	-0.0058	-0.0003	-0.0002	-0.0310	-0.0320	-0.0307	-0.0261
				SYS	0.0004	-0.0455	-0.0393	-0.0295	-0.0276	0.0001	-0.0195	-0.0120	-0.0085	-0.0138	-0.0001	-0.0260	-0.0270	-0.0307	-0.0344
			10	FD	0.0000	0.0095	0.0118	0.0162	0.0185	0.0001	-0.0040	-0.0023	-0.0010	0.0005	0.0000	-0.0085	-0.0085	-0.0078	-0.0064
				SYS	0.0001	-0.0130	-0.0125	-0.0097	-0.0085	0.0002	-0.0050	-0.0033	-0.0017	-0.0031	0.0000	-0.0065	-0.0073	-0.0080	-0.0075
	500	sw			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
			0.1	FD	-0.0021	-0.5465	-0.4670	-0.2970	0.0171	-0.0026	-0.7905	-0.7660	-0.5197	-0.0453	-0.0014	-0.9205	-0.9155	-0.7778	-0.1531
				SYS	-0.0009	-0.6045	-0.4988	-0.3168	-0.1338	-0.0008	-0.8005	-0.7405	-0.4965	-0.1083	0.0004	-0.8950	-0.8865	-0.7773	-0.2441
			0.2	FD	-0.0003	-0.5310	-0.4435	-0.2733	0.0349	-0.0016	-0.7590	-0.7523	-0.5573	-0.0573	-0.0018	-0.8885	-0.8978	-0.7747	-0.1604
				SYS	0.0006	-0.5785	-0.4493	-0.2853	-0.1353	0.0000	-0.7415	-0.6835	-0.5007	-0.1416	0.0010	-0.8450	-0.8520	-0.7617	-0.2639
			0.5	FD	-0.0006	-0.4685	-0.3655	-0.2032	0.0681	-0.0013	-0.5900	-0.5720	-0.5065	-0.0860	-0.0019	-0.7315	-0.7508	-0.6927	-0.1760
				SYS	0.0017	-0.4450	-0.3373	-0.2310	-0.1391	0.0002	-0.5745	-0.5200	-0.4345	-0.1908	0.0005	-0.6945	-0.7110	-0.6715	-0.2766
			1	FD	-0.0020	-0.3390	-0.2425	-0.1132	0.0961	-0.0012	-0.3030	-0.2560	-0.2170	-0.0853	-0.0009	-0.4360	-0.4488	-0.4435	-0.1730
				SYS	0.0012	-0.3090	-0.2278	-0.1683	-0.1180	-0.0003	-0.3305	-0.2678	-0.2347	-0.1711	0.0003	-0.4250	-0.4350	-0.4440	-0.2399
			2	FD	-0.0008	-0.1270	-0.0770	-0.0025	0.0970	-0.0008	-0.0920	-0.0700	-0.0497	-0.0211	-0.0004	-0.1660	-0.1635	-0.1638	-0.1116
				SYS	0.0004	-0.1615	-0.1228	-0.0897	-0.0745	-0.0003	-0.1160	-0.0840	-0.0683	-0.0719	0.0000	-0.1645	-0.1568	-0.1692	-0.1478
			5	FD	0.0001	0.0010	0.0160	0.0322	0.0486	-0.0001	-0.0145	-0.0098	-0.0060	-0.0004	0.0000	-0.0300	-0.0300	-0.0283	-0.0235
				SYS	0.0004	-0.0450	-0.0378	-0.0295	-0.0278	-0.0001	-0.0205	-0.0135	-0.0095	-0.0140	0.0001	-0.0290	-0.0273	-0.0290	-0.0325
			10	FD	0.0000	0.0080	0.0125	0.0165	0.0194	0.0000	-0.0040	-0.0025	-0.0012	0.0005	0.0000	-0.0080	-0.0078	-0.0072	-0.0056
				SYS	0.0001	-0.0130	-0.0120	-0.0097	-0.0084	0.0000	-0.0055	-0.0035	-0.0020	-0.0031	0.0000	-0.0070	-0.0070	-0.0075	-0.0069

Table IV-12: Relative bias $\hat{\gamma} - \gamma$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N	50	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
			0.1 FD	0.1007	0.5630	0.3528	0.2590	0.2348	0.0587	0.3155	0.1883	0.1340	0.0510	0.0415	0.2265	0.1250	0.0973	0.0319
			0.1 SYS	0.0837	0.4340	0.2410	0.1597	0.1225	0.0539	0.2900	0.1618	0.1130	0.0509	0.0427	0.2235	0.1168	0.0858	0.0390
			0.2 FD	0.0971	0.5600	0.3375	0.2562	0.2358	0.0586	0.3305	0.1865	0.1383	0.0515	0.0428	0.2195	0.1188	0.0933	0.0345
			0.2 SYS	0.0807	0.4220	0.2155	0.1393	0.0999	0.0542	0.2980	0.1575	0.1063	0.0480	0.0424	0.2035	0.1115	0.0805	0.0390
			0.5 FD	0.0932	0.5290	0.3045	0.2238	0.2193	0.0516	0.2595	0.1515	0.1287	0.0604	0.0350	0.1900	0.1025	0.0785	0.0343
			0.5 SYS	0.0738	0.3615	0.1750	0.1072	0.0680	0.0479	0.2395	0.1243	0.0982	0.0474	0.0352	0.1855	0.0980	0.0713	0.0403
			1 FD	0.0811	0.4575	0.2378	0.1878	0.1778	0.0400	0.1930	0.1008	0.0722	0.0603	0.0261	0.1445	0.0765	0.0568	0.0333
			1 SYS	0.0606	0.2825	0.1280	0.0793	0.0511	0.0369	0.1790	0.0945	0.0647	0.0470	0.0267	0.1490	0.0800	0.0567	0.0358
			2 FD	0.0537	0.3045	0.1703	0.1357	0.1221	0.0236	0.1165	0.0523	0.0358	0.0338	0.0157	0.0810	0.0430	0.0308	0.0228
			2 SYS	0.0387	0.1950	0.0975	0.0597	0.0380	0.0228	0.1125	0.0500	0.0342	0.0309	0.0160	0.0855	0.0445	0.0312	0.0238
			5 FD	0.0219	0.1300	0.0750	0.0617	0.0451	0.0102	0.0465	0.0205	0.0132	0.0116	0.0065	0.0360	0.0170	0.0113	0.0091
			5 SYS	0.0190	0.1005	0.0525	0.0365	0.0243	0.0096	0.0445	0.0198	0.0130	0.0113	0.0065	0.0355	0.0170	0.0120	0.0098
			10 FD	0.0107	0.0615	0.0343	0.0270	0.0205	0.0050	0.0230	0.0103	0.0063	0.0051	0.0033	0.0180	0.0088	0.0055	0.0043
			10 SYS	0.0099	0.0550	0.0283	0.0205	0.0156	0.0048	0.0220	0.0100	0.0063	0.0053	0.0034	0.0175	0.0085	0.0057	0.0044
N	250	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
			0.1 FD	0.0408	0.2355	0.1320	0.0893	0.1014	0.0283	0.1485	0.0973	0.0742	0.0238	0.0172	0.0955	0.0518	0.0400	0.0118
			0.1 SYS	0.0325	0.1735	0.0943	0.0647	0.0528	0.0249	0.1245	0.0800	0.0610	0.0231	0.0169	0.0880	0.0470	0.0348	0.0159
			0.2 FD	0.0413	0.2250	0.1260	0.0868	0.1013	0.0275	0.1520	0.0953	0.0725	0.0243	0.0178	0.0915	0.0510	0.0387	0.0116
			0.2 SYS	0.0325	0.1585	0.0873	0.0558	0.0443	0.0241	0.1240	0.0718	0.0532	0.0218	0.0167	0.0870	0.0458	0.0345	0.0154
			0.5 FD	0.0388	0.2150	0.1180	0.0755	0.0911	0.0245	0.1275	0.0753	0.0607	0.0306	0.0151	0.0795	0.0428	0.0357	0.0124
			0.5 SYS	0.0299	0.1430	0.0728	0.0415	0.0300	0.0214	0.1045	0.0588	0.0427	0.0236	0.0144	0.0765	0.0398	0.0310	0.0151
			1 FD	0.0353	0.1890	0.0920	0.0707	0.0779	0.0177	0.0920	0.0490	0.0375	0.0294	0.0121	0.0640	0.0330	0.0235	0.0121
			1 SYS	0.0251	0.1220	0.0565	0.0340	0.0229	0.0161	0.0830	0.0410	0.0290	0.0206	0.0116	0.0615	0.0328	0.0238	0.0141
			2 FD	0.0238	0.1355	0.0720	0.0610	0.0518	0.0103	0.0495	0.0243	0.0163	0.0164	0.0071	0.0375	0.0190	0.0127	0.0091
			2 SYS	0.0166	0.0830	0.0410	0.0257	0.0180	0.0100	0.0485	0.0230	0.0140	0.0113	0.0071	0.0370	0.0195	0.0135	0.0100
			5 FD	0.0102	0.0595	0.0333	0.0265	0.0210	0.0043	0.0200	0.0095	0.0058	0.0049	0.0032	0.0160	0.0078	0.0052	0.0040
			5 SYS	0.0082	0.0435	0.0225	0.0148	0.0103	0.0043	0.0200	0.0095	0.0057	0.0043	0.0031	0.0160	0.0078	0.0053	0.0043
			10 FD	0.0050	0.0270	0.0160	0.0115	0.0093	0.0022	0.0105	0.0048	0.0027	0.0024	0.0016	0.0080	0.0040	0.0025	0.0019
			10 SYS	0.0046	0.0240	0.0130	0.0087	0.0066	0.0022	0.0105	0.0048	0.0027	0.0023	0.0015	0.0075	0.0038	0.0025	0.0020
N	500	sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
			0.1 FD	0.0286	0.1655	0.0978	0.0623	0.0800	0.0188	0.1045	0.0650	0.0505	0.0161	0.0128	0.0675	0.0388	0.0287	0.0089
			0.1 SYS	0.0221	0.1225	0.0673	0.0443	0.0363	0.0159	0.0850	0.0543	0.0415	0.0159	0.0116	0.0605	0.0348	0.0257	0.0123
			0.2 FD	0.0295	0.1675	0.0953	0.0607	0.0800	0.0177	0.1045	0.0678	0.0517	0.0168	0.0131	0.0645	0.0373	0.0292	0.0093
			0.2 SYS	0.0229	0.1175	0.0628	0.0398	0.0303	0.0150	0.0805	0.0518	0.0392	0.0154	0.0118	0.0580	0.0335	0.0257	0.0124
			0.5 FD	0.0274	0.1505	0.0803	0.0548	0.0715	0.0162	0.0870	0.0533	0.0488	0.0193	0.0117	0.0595	0.0323	0.0247	0.0093
			0.5 SYS	0.0207	0.1010	0.0505	0.0305	0.0209	0.0142	0.0705	0.0405	0.0345	0.0163	0.0107	0.0535	0.0300	0.0218	0.0116
			1 FD	0.0241	0.1315	0.0673	0.0498	0.0583	0.0120	0.0665	0.0335	0.0248	0.0199	0.0086	0.0470	0.0248	0.0177	0.0085
			1 SYS	0.0166	0.0840	0.0415	0.0248	0.0164	0.0111	0.0580	0.0283	0.0193	0.0139	0.0081	0.0435	0.0230	0.0172	0.0098
			2 FD	0.0167	0.0965	0.0500	0.0412	0.0389	0.0073	0.0355	0.0173	0.0117	0.0115	0.0051	0.0285	0.0140	0.0092	0.0066
			2 SYS	0.0120	0.0590	0.0295	0.0178	0.0129	0.0070	0.0345	0.0163	0.0098	0.0083	0.0051	0.0275	0.0140	0.0093	0.0070
			5 FD	0.0071	0.0425	0.0240	0.0187	0.0158	0.0030	0.0145	0.0068	0.0042	0.0038	0.0023	0.0115	0.0058	0.0035	0.0028
			5 SYS	0.0058	0.0310	0.0155	0.0110	0.0073	0.0030	0.0145	0.0068	0.0040	0.0031	0.0023	0.0115	0.0058	0.0035	0.0030
			10 FD	0.0033	0.0200	0.0113	0.0083	0.0065	0.0015	0.0070	0.0033	0.0020	0.0016	0.0012	0.0060	0.0028	0.0017	0.0014
			10 SYS	0.0030	0.0170	0.0090	0.0062	0.0048	0.0015	0.0075	0.0033	0.0020	0.0016	0.0012	0.0060	0.0028	0.0017	0.0014

Table IV-13: Relative standard deviation of $\hat{\gamma}$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	50	sw 0.1	FD	0.1023	0.8920	0.6763	0.4803	0.2384	0.0614	0.9285	0.8118	0.5820	0.1199	0.0493	1.0600	0.9760	0.8012	0.1753
			SYS	0.0839	0.8175	0.6055	0.3933	0.1939	0.0547	0.9285	0.8013	0.5528	0.1666	0.0447	1.0030	0.9205	0.7578	0.2248
		0.2	FD	0.0987	0.8670	0.6450	0.4390	0.2370	0.0609	0.9340	0.8078	0.6007	0.1291	0.0483	1.0210	0.9535	0.8023	0.1820
			SYS	0.0808	0.7525	0.5485	0.3477	0.1756	0.0546	0.8965	0.7635	0.5433	0.1894	0.0431	0.9150	0.8620	0.7163	0.2405
		0.5	FD	0.0937	0.7885	0.5553	0.3498	0.2221	0.0529	0.7145	0.6398	0.5612	0.1601	0.0397	0.8265	0.7915	0.7268	0.2028
			SYS	0.0739	0.5720	0.3950	0.2767	0.1641	0.0478	0.6675	0.5948	0.4943	0.2315	0.0356	0.7300	0.7043	0.6328	0.2544
		1	FD	0.0813	0.5990	0.3860	0.2322	0.1920	0.0411	0.4030	0.3270	0.2845	0.1499	0.0279	0.4810	0.4780	0.4842	0.2089
			SYS	0.0610	0.3950	0.2570	0.1938	0.1373	0.0370	0.3615	0.3090	0.2962	0.2198	0.0269	0.4260	0.4310	0.4433	0.2383
		2	FD	0.0539	0.3485	0.2010	0.1363	0.1434	0.0238	0.1690	0.1038	0.0810	0.0545	0.0161	0.1965	0.1855	0.1855	0.1400
			SYS	0.0388	0.2435	0.1535	0.1062	0.0904	0.0229	0.1530	0.0918	0.0855	0.1001	0.0160	0.1720	0.1600	0.1748	0.1573
		5	FD	0.0219	0.1305	0.0750	0.0665	0.0606	0.0102	0.0505	0.0248	0.0162	0.0120	0.0066	0.0500	0.0375	0.0338	0.0311
			SYS	0.0190	0.1065	0.0633	0.0453	0.0345	0.0096	0.0480	0.0225	0.0148	0.0168	0.0065	0.0445	0.0313	0.0313	0.0363
		10	FD	0.0107	0.0615	0.0360	0.0308	0.0263	0.0050	0.0235	0.0108	0.0067	0.0051	0.0034	0.0190	0.0120	0.0098	0.0081
			SYS	0.0099	0.0560	0.0295	0.0217	0.0168	0.0048	0.0225	0.0103	0.0065	0.0053	0.0034	0.0180	0.0108	0.0093	0.0085
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	sw 0.1	FD	0.0410	0.5925	0.4760	0.3118	0.1015	0.0286	0.8180	0.7903	0.5472	0.0568	0.0179	0.9295	0.9040	0.7447	0.1540
			SYS	0.0325	0.6205	0.5018	0.3220	0.1419	0.0249	0.8240	0.7613	0.5143	0.1166	0.0169	0.9185	0.8900	0.7502	0.2371
		0.2	FD	0.0414	0.5880	0.4598	0.2898	0.1040	0.0280	0.7965	0.7690	0.5820	0.0688	0.0182	0.9035	0.8883	0.7478	0.1609
			SYS	0.0325	0.5925	0.4570	0.2915	0.1404	0.0241	0.7685	0.6958	0.5102	0.1486	0.0167	0.8770	0.8570	0.7397	0.2570
		0.5	FD	0.0388	0.5190	0.3835	0.2167	0.1034	0.0248	0.6150	0.5875	0.5228	0.0994	0.0156	0.7395	0.7503	0.6837	0.1775
			SYS	0.0300	0.4685	0.3460	0.2348	0.1393	0.0214	0.5830	0.5260	0.4408	0.1971	0.0145	0.7060	0.7185	0.6650	0.2714
		1	FD	0.0355	0.3865	0.2600	0.1300	0.1179	0.0177	0.3165	0.2635	0.2253	0.0956	0.0123	0.4500	0.4618	0.4583	0.1776
			SYS	0.0251	0.3255	0.2330	0.1707	0.1196	0.0162	0.3255	0.2678	0.2407	0.1763	0.0116	0.4250	0.4433	0.4562	0.2424
		2	FD	0.0238	0.1860	0.1063	0.0610	0.1043	0.0103	0.1070	0.0745	0.0518	0.0273	0.0072	0.1760	0.1773	0.1772	0.1206
			SYS	0.0166	0.1765	0.1295	0.0932	0.0781	0.0100	0.1205	0.0828	0.0678	0.0731	0.0071	0.1610	0.1625	0.1783	0.1551
		5	FD	0.0102	0.0595	0.0358	0.0400	0.0518	0.0043	0.0250	0.0135	0.0082	0.0049	0.0032	0.0350	0.0330	0.0310	0.0264
			SYS	0.0082	0.0630	0.0453	0.0330	0.0294	0.0043	0.0280	0.0153	0.0102	0.0144	0.0031	0.0305	0.0280	0.0312	0.0346
		10	FD	0.0050	0.0285	0.0200	0.0198	0.0208	0.0022	0.0110	0.0055	0.0028	0.0024	0.0016	0.0115	0.0093	0.0082	0.0066
			SYS	0.0046	0.0275	0.0180	0.0130	0.0108	0.0022	0.0120	0.0058	0.0032	0.0039	0.0015	0.0105	0.0083	0.0085	0.0078
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw 0.1	FD	0.0287	0.5710	0.4773	0.3033	0.0818	0.0189	0.7975	0.7688	0.5222	0.0480	0.0129	0.9230	0.9163	0.7785	0.1534
			SYS	0.0221	0.6170	0.5033	0.3200	0.1385	0.0160	0.8050	0.7425	0.4982	0.1094	0.0116	0.8975	0.8870	0.7778	0.2444
		0.2	FD	0.0295	0.5570	0.4535	0.2800	0.0873	0.0178	0.7660	0.7553	0.5598	0.0596	0.0132	0.8910	0.8985	0.7752	0.1606
			SYS	0.0229	0.5905	0.4538	0.2882	0.1386	0.0150	0.7455	0.6855	0.5022	0.1425	0.0119	0.8470	0.8528	0.7622	0.2641
		0.5	FD	0.0274	0.4925	0.3743	0.2103	0.0988	0.0162	0.5965	0.5745	0.5088	0.0881	0.0119	0.7335	0.7513	0.6930	0.1763
			SYS	0.0208	0.4560	0.3410	0.2330	0.1406	0.0142	0.5790	0.5218	0.4360	0.1914	0.0107	0.6965	0.7118	0.6718	0.2769
		1	FD	0.0242	0.3640	0.2518	0.1237	0.1124	0.0120	0.3105	0.2583	0.2185	0.0876	0.0086	0.4385	0.4495	0.4438	0.1733
			SYS	0.0166	0.3200	0.2315	0.1702	0.1191	0.0111	0.3355	0.2693	0.2355	0.1716	0.0081	0.4270	0.4355	0.4443	0.2400
		2	FD	0.0167	0.1595	0.0918	0.0413	0.1045	0.0073	0.0985	0.0723	0.0510	0.0240	0.0051	0.1685	0.1643	0.1642	0.1119
			SYS	0.0120	0.1720	0.1263	0.0913	0.0755	0.0070	0.1210	0.0858	0.0690	0.0723	0.0051	0.1665	0.1575	0.1695	0.1479
		5	FD	0.0071	0.0425	0.0288	0.0372	0.0511	0.0030	0.0205	0.0120	0.0073	0.0038	0.0023	0.0325	0.0305	0.0285	0.0236
			SYS	0.0058	0.0545	0.0408	0.0315	0.0286	0.0030	0.0255	0.0150	0.0103	0.0144	0.0023	0.0310	0.0278	0.0293	0.0326
		10	FD	0.0033	0.0215	0.0168	0.0183	0.0205	0.0015	0.0080	0.0040	0.0023	0.0018	0.0012	0.0100	0.0083	0.0073	0.0059
			SYS	0.0030	0.0215	0.0148	0.0115	0.0096	0.0015	0.0090	0.0048	0.0028	0.0035	0.0012	0.0095	0.0075	0.0077	0.0071

Table IV-14: Relative RMSE of $\hat{\gamma}$ over N, T, sw and γ

5.2.2 Relative Bias and Standard Deviation and RMSE of β (Short-Run Effects)

					T													
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	50	0.1	FD	0.0033	0.1540	-0.0839	-0.7186	-0.8561	-0.0073	0.0397	-0.1595	-0.7335	-0.8148	0.0018	-0.0704	-0.2174	-0.4485	-0.2613
			SYS	0.1011	0.2005	0.1588	-0.0005	-0.2512	0.0841	0.1993	0.1816	-0.1328	-0.7571	0.1461	0.2481	0.3084	0.2471	-0.4863
		0.2	FD	0.0083	0.0438	-0.0771	-0.3803	-0.5004	0.0072	0.0030	-0.1159	-0.4210	-0.4101	-0.0068	-0.0907	-0.2031	-0.3366	-0.1840
			SYS	0.0686	0.1364	0.0884	-0.0701	-0.2810	0.0552	0.1149	0.0757	-0.1735	-0.6085	0.0711	0.0766	0.0739	0.0296	-0.3429
		0.5	FD	0.0048	0.0192	-0.0583	-0.1741	-0.1360	-0.0016	-0.0063	-0.0680	-0.2047	-0.2261	-0.0074	-0.0718	-0.1476	-0.2360	-0.1273
			SYS	0.0211	0.0340	-0.0258	-0.1335	-0.2195	0.0105	0.0121	-0.0380	-0.1735	-0.3694	0.0099	-0.0366	-0.0811	-0.1350	-0.1448
		sw 1	FD	0.0013	0.0011	-0.0414	-0.0914	-0.0337	0.0026	-0.0058	-0.0273	-0.0761	-0.1203	-0.0054	-0.0410	-0.0871	-0.1438	-0.1062
			SYS	0.0059	-0.0054	-0.0430	-0.0968	-0.1461	0.0032	-0.0127	-0.0424	-0.0977	-0.2083	-0.0017	-0.0324	-0.0684	-0.1123	-0.0990
		2	FD	0.0006	-0.0050	-0.0257	-0.0338	-0.0022	0.0014	-0.0015	-0.0094	-0.0203	-0.0339	-0.0032	-0.0162	-0.0330	-0.0514	-0.0616
			SYS	0.0020	-0.0075	-0.0331	-0.0552	-0.0780	0.0011	-0.0072	-0.0213	-0.0369	-0.0699	-0.0020	-0.0131	-0.0269	-0.0432	-0.0583
		5	FD	0.0002	-0.0029	-0.0063	-0.0046	0.0054	-0.0004	-0.0016	-0.0030	-0.0057	-0.0072	-0.0006	-0.0030	-0.0064	-0.0095	-0.0130
			SYS	0.0006	-0.0024	-0.0094	-0.0178	-0.0288	-0.0003	-0.0025	-0.0051	-0.0093	-0.0121	-0.0003	-0.0021	-0.0048	-0.0079	-0.0129
		10	FD	0.0004	-0.0012	-0.0015	-0.0008	0.0027	0.0003	-0.0004	-0.0015	-0.0023	-0.0028	-0.0002	-0.0008	-0.0020	-0.0028	-0.0036
			SYS	0.0006	-0.0009	-0.0024	-0.0056	-0.0096	0.0003	-0.0007	-0.0019	-0.0030	-0.0040	-0.0001	-0.0002	-0.0015	-0.0022	-0.0034
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	0.1	FD	0.0169	0.0732	-0.1158	-0.6475	-0.8554	-0.0041	0.0306	-0.1708	-0.7324	-0.8063	-0.0065	-0.0750	-0.2196	-0.3933	-0.2471
			SYS	0.0023	0.1021	0.1161	-0.0224	-0.2220	0.0335	0.1656	0.1581	-0.2842	-0.8570	0.0421	0.0984	0.0687	-0.1356	-0.8811
		0.2	FD	0.0125	0.0480	-0.0946	-0.3646	-0.3753	-0.0053	-0.0031	-0.1120	-0.4127	-0.4389	-0.0085	-0.0790	-0.1921	-0.3108	-0.1645
			SYS	0.0041	0.0918	0.0621	-0.0757	-0.2313	0.0178	0.0956	0.0347	-0.2816	-0.6957	0.0109	0.0042	-0.0544	-0.1895	-0.5697
		0.5	FD	0.0016	0.0079	-0.0571	-0.1644	-0.1383	-0.0018	-0.0101	-0.0622	-0.2009	-0.2171	-0.0037	-0.0690	-0.1498	-0.2317	-0.1209
			SYS	0.0006	0.0229	-0.0244	-0.1140	-0.1992	0.0033	0.0051	-0.0472	-0.1985	-0.3851	0.0006	-0.0494	-0.1126	-0.1920	-0.2059
N		sw 1	FD	0.0004	-0.0030	-0.0394	-0.0829	-0.0307	-0.0015	-0.0061	-0.0242	-0.0700	-0.1101	-0.0014	-0.0398	-0.0862	-0.1383	-0.0929
			SYS	0.0006	-0.0082	-0.0408	-0.0824	-0.1277	-0.0005	-0.0089	-0.0379	-0.1007	-0.2010	-0.0006	-0.0345	-0.0742	-0.1225	-0.1078
		2	FD	-0.0001	-0.0057	-0.0222	-0.0302	0.0046	-0.0002	-0.0029	-0.0077	-0.0180	-0.0315	-0.0007	-0.0150	-0.0307	-0.0511	-0.0555
			SYS	-0.0002	-0.0086	-0.0287	-0.0502	-0.0687	-0.0002	-0.0048	-0.0151	-0.0318	-0.0629	-0.0005	-0.0129	-0.0264	-0.0457	-0.0596
		5	FD	0.0002	-0.0034	-0.0064	-0.0044	0.0074	-0.0001	-0.0015	-0.0033	-0.0057	-0.0083	0.0000	-0.0030	-0.0061	-0.0092	-0.0122
			SYS	0.0001	-0.0029	-0.0092	-0.0185	-0.0276	-0.0001	-0.0012	-0.0034	-0.0069	-0.0102	0.0000	-0.0022	-0.0048	-0.0081	-0.0133
		10	FD	-0.0001	-0.0015	-0.0021	-0.0006	0.0027	0.0000	-0.0007	-0.0014	-0.0024	-0.0033	0.0000	-0.0009	-0.0017	-0.0026	-0.0033
			SYS	-0.0001	-0.0010	-0.0031	-0.0059	-0.0107	0.0000	-0.0004	-0.0009	-0.0020	-0.0030	0.0000	-0.0005	-0.0012	-0.0022	-0.0034
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	0.1	FD	0.0126	0.1095	-0.1295	-0.6861	-0.7957	0.0029	0.0314	-0.1782	-0.7234	-0.7809	-0.0077	-0.0822	-0.2198	-0.3984	-0.2308
			SYS	0.0197	0.1268	0.1346	-0.0146	-0.2032	0.0036	0.1509	0.1141	-0.3730	-0.8900	0.0299	0.1115	0.0831	-0.1627	-0.9449
		0.2	FD	0.0011	0.0437	-0.0974	-0.3829	-0.3780	-0.0008	0.0048	-0.1179	-0.4011	-0.4334	-0.0068	-0.0838	-0.1998	-0.3147	-0.1698
			SYS	0.0130	0.1076	0.0718	-0.0762	-0.2207	-0.0009	0.0839	0.0021	-0.3266	-0.7186	0.0125	0.0102	-0.0569	-0.2163	-0.6023
		0.5	FD	0.0007	0.0089	-0.0605	-0.1684	-0.1216	-0.0014	-0.0082	-0.0683	-0.1965	-0.2170	-0.0033	-0.0699	-0.1522	-0.2354	-0.1198
			SYS	0.0058	0.0274	-0.0258	-0.1180	-0.2023	-0.0013	0.0032	-0.0588	-0.2038	-0.3953	0.0006	-0.0480	-0.1125	-0.2001	-0.2138
		sw 1	FD	0.0002	-0.0039	-0.0422	-0.0850	-0.0295	-0.0001	-0.0050	-0.0237	-0.0685	-0.1053	-0.0014	-0.0382	-0.0829	-0.1359	-0.0923
			SYS	0.0008	-0.0070	-0.0425	-0.0846	-0.1323	-0.0002	-0.0083	-0.0381	-0.1016	-0.1988	-0.0002	-0.0329	-0.0716	-0.1229	-0.1119
		2	FD	0.0002	-0.0064	-0.0224	-0.0323	0.0063	-0.0003	-0.0027	-0.0073	-0.0188	-0.0315	0.0000	-0.0139	-0.0278	-0.0468	-0.0519
			SYS	0.0003	-0.0089	-0.0283	-0.0509	-0.0682	-0.0005	-0.0045	-0.0147	-0.0325	-0.0631	0.0002	-0.0127	-0.0246	-0.0427	-0.0575
		5	FD	0.0001	-0.0033	-0.0064	-0.0042	0.0078	0.0000	-0.0013	-0.0032	-0.0059	-0.0082	-0.0001	-0.0026	-0.0054	-0.0086	-0.0109
			SYS	0.0000	-0.0027	-0.0091	-0.0179	-0.0285	0.0000	-0.0010	-0.0031	-0.0068	-0.0101	0.0000	-0.0021	-0.0044	-0.0076	-0.0122
		10	FD	-0.0001	-0.0013	-0.0020	-0.0006	0.0035	0.0000	-0.0006	-0.0015	-0.0024	-0.0033	0.0000	-0.0008	-0.0016	-0.0024	-0.0032
			SYS	0.0000	-0.0009	-0.0029	-0.0058	-0.0102	0.0000	-0.0003	-0.0009	-0.0020	-0.0030	0.0000	-0.0005	-0.0012	-0.0020	-0.0033

Table IV-15: Relative bias $\hat{\beta} - \beta$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	50	0.1	FD	1.0124	1.0277	1.0709	1.1754	1.5371	0.5415	0.5704	0.5757	0.5640	0.7515	0.3525	0.3405	0.3208	0.3691	0.4805
			SYS	0.3602	0.3269	0.3643	0.3743	0.3650	0.2451	0.2552	0.2710	0.2779	0.2267	0.2127	0.2192	0.2161	0.2483	0.2340
		0.2	FD	0.5120	0.5071	0.5679	0.6355	0.8208	0.2706	0.2714	0.2790	0.2974	0.3584	0.1581	0.1674	0.1634	0.1677	0.2240
			SYS	0.2915	0.2964	0.3332	0.3270	0.3306	0.1984	0.1947	0.2066	0.2171	0.1782	0.1343	0.1432	0.1447	0.1630	0.1748
		0.5	FD	0.2012	0.2158	0.2159	0.2492	0.3415	0.1121	0.1103	0.1145	0.1207	0.1520	0.0663	0.0665	0.0705	0.0751	0.0918
			SYS	0.1751	0.1781	0.1891	0.2059	0.2066	0.1047	0.0987	0.1043	0.1063	0.1082	0.0659	0.0642	0.0719	0.0752	0.0921
		1	FD	0.0952	0.1024	0.1084	0.1296	0.1784	0.0539	0.0552	0.0551	0.0604	0.0766	0.0336	0.0366	0.0368	0.0412	0.0455
			SYS	0.0894	0.0939	0.1046	0.1141	0.1278	0.0516	0.0514	0.0520	0.0571	0.0616	0.0351	0.0373	0.0378	0.0421	0.0466
		2	FD	0.0488	0.0472	0.0522	0.0671	0.0935	0.0260	0.0276	0.0283	0.0290	0.0338	0.0173	0.0179	0.0190	0.0206	0.0230
			SYS	0.0455	0.0455	0.0527	0.0609	0.0685	0.0242	0.0261	0.0270	0.0277	0.0312	0.0175	0.0182	0.0191	0.0205	0.0226
		5	FD	0.0178	0.0196	0.0225	0.0274	0.0356	0.0102	0.0108	0.0105	0.0112	0.0119	0.0068	0.0069	0.0075	0.0080	0.0095
			SYS	0.0175	0.0195	0.0226	0.0258	0.0312	0.0097	0.0099	0.0098	0.0105	0.0113	0.0069	0.0069	0.0075	0.0079	0.0093
		10	FD	0.0088	0.0095	0.0108	0.0132	0.0164	0.0051	0.0050	0.0052	0.0052	0.0057	0.0033	0.0035	0.0037	0.0040	0.0047
			SYS	0.0087	0.0095	0.0111	0.0132	0.0161	0.0049	0.0047	0.0049	0.0050	0.0056	0.0034	0.0034	0.0037	0.0040	0.0046
		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	0.1	FD	0.4225	0.4493	0.4636	0.5368	0.6800	0.2500	0.2537	0.2569	0.2740	0.3593	0.1490	0.1489	0.1491	0.1622	0.2071
			SYS	0.1265	0.1198	0.1447	0.1513	0.1561	0.0986	0.1082	0.1559	0.1510	0.1139	0.0978	0.1006	0.1138	0.1186	0.0992
		0.2	FD	0.2112	0.2125	0.2406	0.2676	0.3541	0.1261	0.1218	0.1256	0.1399	0.1793	0.0725	0.0769	0.0763	0.0790	0.1062
			SYS	0.1085	0.1092	0.1354	0.1438	0.1397	0.0836	0.0888	0.1087	0.1049	0.0884	0.0609	0.0672	0.0710	0.0804	0.0772
		0.5	FD	0.0875	0.0883	0.0945	0.1106	0.1432	0.0464	0.0491	0.0516	0.0581	0.0766	0.0314	0.0320	0.0310	0.0345	0.0415
			SYS	0.0709	0.0733	0.0817	0.0914	0.0904	0.0431	0.0469	0.0477	0.0478	0.0509	0.0300	0.0311	0.0307	0.0352	0.0410
		1	FD	0.0425	0.0431	0.0468	0.0552	0.0766	0.0239	0.0249	0.0254	0.0280	0.0351	0.0152	0.0164	0.0175	0.0184	0.0202
			SYS	0.0393	0.0408	0.0448	0.0519	0.0578	0.0232	0.0235	0.0236	0.0251	0.0271	0.0157	0.0162	0.0174	0.0185	0.0206
		2	FD	0.0205	0.0220	0.0233	0.0286	0.0390	0.0115	0.0122	0.0126	0.0132	0.0154	0.0077	0.0083	0.0083	0.0090	0.0105
			SYS	0.0191	0.0207	0.0235	0.0266	0.0316	0.0110	0.0115	0.0116	0.0119	0.0137	0.0076	0.0081	0.0082	0.0091	0.0105
		5	FD	0.0077	0.0080	0.0091	0.0122	0.0149	0.0045	0.0048	0.0049	0.0052	0.0053	0.0032	0.0031	0.0033	0.0037	0.0041
			SYS	0.0073	0.0078	0.0094	0.0118	0.0131	0.0044	0.0047	0.0046	0.0047	0.0048	0.0032	0.0030	0.0033	0.0036	0.0040
		10	FD	0.0037	0.0039	0.0046	0.0059	0.0071	0.0022	0.0022	0.0022	0.0023	0.0025	0.0015	0.0016	0.0016	0.0018	0.0021
			SYS	0.0036	0.0039	0.0049	0.0060	0.0070	0.0022	0.0022	0.0021	0.0022	0.0024	0.0015	0.0015	0.0016	0.0018	0.0020
		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	0.1	FD	0.3048	0.3039	0.3218	0.3705	0.4605	0.1632	0.1682	0.1809	0.2027	0.2510	0.1128	0.1097	0.1131	0.1205	0.1511
			SYS	0.0864	0.0859	0.1058	0.1096	0.1122	0.0629	0.0745	0.1157	0.1054	0.0753	0.0633	0.0677	0.0877	0.0908	0.0748
		0.2	FD	0.1509	0.1536	0.1665	0.1856	0.2582	0.0810	0.0845	0.0907	0.1036	0.1284	0.0547	0.0559	0.0559	0.0595	0.0743
			SYS	0.0779	0.0828	0.0962	0.1060	0.0983	0.0543	0.0602	0.0791	0.0734	0.0639	0.0442	0.0482	0.0516	0.0567	0.0570
		0.5	FD	0.0595	0.0604	0.0649	0.0777	0.1045	0.0329	0.0346	0.0353	0.0393	0.0511	0.0218	0.0223	0.0243	0.0251	0.0293
			SYS	0.0492	0.0525	0.0607	0.0662	0.0672	0.0296	0.0323	0.0324	0.0336	0.0354	0.0209	0.0223	0.0237	0.0245	0.0294
		1	FD	0.0298	0.0297	0.0326	0.0390	0.0544	0.0162	0.0170	0.0185	0.0203	0.0259	0.0108	0.0114	0.0132	0.0133	0.0144
			SYS	0.0269	0.0278	0.0316	0.0375	0.0397	0.0155	0.0161	0.0169	0.0177	0.0198	0.0106	0.0113	0.0130	0.0131	0.0148
		2	FD	0.0147	0.0148	0.0170	0.0206	0.0272	0.0076	0.0085	0.0090	0.0092	0.0104	0.0057	0.0060	0.0063	0.0064	0.0074
			SYS	0.0132	0.0144	0.0168	0.0200	0.0211	0.0074	0.0080	0.0082	0.0081	0.0096	0.0057	0.0059	0.0062	0.0063	0.0072
		5	FD	0.0054	0.0059	0.0067	0.0083	0.0105	0.0032	0.0033	0.0033	0.0034	0.0035	0.0023	0.0023	0.0025	0.0026	0.0030
			SYS	0.0051	0.0058	0.0068	0.0082	0.0092	0.0031	0.0032	0.0031	0.0031	0.0032	0.0022	0.0023	0.0024	0.0025	0.0029
		10	FD	0.0027	0.0028	0.0032	0.0041	0.0050	0.0015	0.0016	0.0016	0.0017	0.0017	0.0012	0.0011	0.0012	0.0013	0.0014
			SYS	0.0026	0.0029	0.0034	0.0042	0.0051	0.0015	0.0016	0.0015	0.0016	0.0016	0.0012	0.0011	0.0011	0.0013	0.0014

Table IV-16: Relative standard deviation of $\hat{\beta}$ over N, T, sw and γ

				T														
				5					10					20				
				γ					γ					γ				
N		sw		0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
50	sw	0.1	FD	1.0119	1.0387	1.0736	1.3772	1.7587	0.5412	0.5715	0.5971	0.9251	1.1082	0.3523	0.3475	0.3873	0.5807	0.5468
			SYS	0.3739	0.3834	0.3972	0.3742	0.4430	0.2590	0.3237	0.3261	0.3079	0.7903	0.2580	0.3310	0.3765	0.3502	0.5396
		0.2	FD	0.5118	0.5087	0.5728	0.7403	0.9610	0.2705	0.2712	0.3020	0.5153	0.5445	0.1581	0.1903	0.2606	0.3760	0.2898
			SYS	0.2993	0.3262	0.3446	0.3343	0.4338	0.2059	0.2259	0.2200	0.2778	0.6340	0.1519	0.1624	0.1625	0.1656	0.3849
		0.5	FD	0.2012	0.2166	0.2235	0.3039	0.3674	0.1121	0.1105	0.1331	0.2376	0.2725	0.0667	0.0979	0.1635	0.2477	0.1569
			SYS	0.1763	0.1813	0.1908	0.2453	0.3014	0.1051	0.0994	0.1110	0.2035	0.3849	0.0666	0.0739	0.1084	0.1545	0.1716
		1	FD	0.0952	0.1024	0.1160	0.1585	0.1814	0.0540	0.0554	0.0615	0.0972	0.1426	0.0340	0.0550	0.0946	0.1496	0.1155
			SYS	0.0895	0.0940	0.1131	0.1496	0.1940	0.0517	0.0530	0.0671	0.1131	0.2172	0.0351	0.0494	0.0782	0.1199	0.1094
		2	FD	0.0488	0.0474	0.0582	0.0751	0.0935	0.0260	0.0276	0.0298	0.0354	0.0478	0.0175	0.0241	0.0381	0.0553	0.0658
			SYS	0.0455	0.0461	0.0622	0.0822	0.1037	0.0243	0.0270	0.0344	0.0461	0.0765	0.0176	0.0224	0.0330	0.0478	0.0625
		5	FD	0.0178	0.0198	0.0234	0.0277	0.0359	0.0102	0.0109	0.0109	0.0126	0.0139	0.0068	0.0075	0.0099	0.0124	0.0161
			SYS	0.0175	0.0196	0.0245	0.0313	0.0424	0.0097	0.0102	0.0111	0.0140	0.0165	0.0069	0.0072	0.0090	0.0112	0.0159
		10	FD	0.0088	0.0095	0.0109	0.0133	0.0166	0.0051	0.0050	0.0054	0.0057	0.0063	0.0033	0.0036	0.0042	0.0049	0.0059
			SYS	0.0087	0.0095	0.0113	0.0144	0.0188	0.0049	0.0047	0.0053	0.0058	0.0069	0.0034	0.0034	0.0040	0.0046	0.0057
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
250	sw	0.1	FD	0.4226	0.4550	0.4777	0.8409	1.0925	0.2499	0.2554	0.3084	0.7820	0.8827	0.1491	0.1667	0.2654	0.4254	0.3224
			SYS	0.1265	0.1574	0.1854	0.1528	0.2714	0.1041	0.1978	0.2220	0.3218	0.8645	0.1064	0.1407	0.1329	0.1801	0.8866
		0.2	FD	0.2115	0.2178	0.2584	0.4522	0.5159	0.1261	0.1218	0.1682	0.4357	0.4740	0.0730	0.1103	0.2066	0.3206	0.1958
			SYS	0.1085	0.1426	0.1489	0.1624	0.2702	0.0854	0.1304	0.1140	0.3005	0.7013	0.0618	0.0673	0.0894	0.2058	0.5749
		0.5	FD	0.0875	0.0886	0.1103	0.1981	0.1990	0.0464	0.0501	0.0808	0.2091	0.2302	0.0316	0.0760	0.1529	0.2343	0.1279
			SYS	0.0709	0.0767	0.0852	0.1461	0.2188	0.0432	0.0472	0.0671	0.2041	0.3884	0.0300	0.0584	0.1167	0.1952	0.2100
		1	FD	0.0425	0.0432	0.0611	0.0996	0.0825	0.0240	0.0256	0.0351	0.0754	0.1156	0.0152	0.0430	0.0879	0.1395	0.0951
			SYS	0.0393	0.0416	0.0605	0.0974	0.1401	0.0231	0.0251	0.0446	0.1038	0.2028	0.0157	0.0381	0.0763	0.1239	0.1097
		2	FD	0.0205	0.0227	0.0322	0.0416	0.0392	0.0115	0.0125	0.0148	0.0223	0.0351	0.0077	0.0171	0.0318	0.0519	0.0565
			SYS	0.0191	0.0224	0.0371	0.0568	0.0756	0.0110	0.0125	0.0191	0.0340	0.0644	0.0077	0.0152	0.0276	0.0466	0.0605
		5	FD	0.0077	0.0087	0.0111	0.0130	0.0166	0.0045	0.0051	0.0059	0.0077	0.0098	0.0032	0.0043	0.0070	0.0099	0.0129
			SYS	0.0073	0.0083	0.0132	0.0220	0.0306	0.0044	0.0048	0.0057	0.0083	0.0113	0.0032	0.0037	0.0058	0.0089	0.0139
		10	FD	0.0037	0.0042	0.0051	0.0059	0.0076	0.0022	0.0023	0.0026	0.0034	0.0041	0.0015	0.0018	0.0024	0.0032	0.0039
			SYS	0.0036	0.0040	0.0058	0.0084	0.0127	0.0022	0.0022	0.0023	0.0030	0.0039	0.0015	0.0016	0.0020	0.0029	0.0039
				0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
500	sw	0.1	FD	0.3049	0.3229	0.3467	0.7796	0.9193	0.1632	0.1711	0.2539	0.7513	0.8202	0.1130	0.1370	0.2471	0.4162	0.2758
			SYS	0.0885	0.1532	0.1712	0.1106	0.2321	0.0629	0.1682	0.1624	0.3876	0.8932	0.0700	0.1304	0.1208	0.1863	0.9479
		0.2	FD	0.1508	0.1597	0.1928	0.4255	0.4577	0.0809	0.0846	0.1487	0.4142	0.4520	0.0551	0.1007	0.2074	0.3202	0.1853
			SYS	0.0789	0.1358	0.1200	0.1305	0.2416	0.0543	0.1032	0.0791	0.3348	0.7214	0.0459	0.0492	0.0768	0.2236	0.6050
		0.5	FD	0.0594	0.0610	0.0887	0.1854	0.1603	0.0329	0.0356	0.0769	0.2004	0.2230	0.0221	0.0733	0.1542	0.2368	0.1233
			SYS	0.0495	0.0592	0.0660	0.1353	0.2132	0.0296	0.0324	0.0672	0.2065	0.3969	0.0209	0.0529	0.1150	0.2016	0.2158
		1	FD	0.0298	0.0299	0.0533	0.0935	0.0619	0.0162	0.0178	0.0300	0.0714	0.1084	0.0108	0.0399	0.0840	0.1365	0.0934
			SYS	0.0269	0.0286	0.0529	0.0925	0.1381	0.0155	0.0180	0.0417	0.1032	0.1997	0.0106	0.0348	0.0728	0.1236	0.1129
		2	FD	0.0147	0.0161	0.0281	0.0383	0.0280	0.0076	0.0089	0.0116	0.0209	0.0332	0.0057	0.0152	0.0285	0.0472	0.0524
			SYS	0.0132	0.0169	0.0329	0.0547	0.0714	0.0074	0.0092	0.0168	0.0334	0.0638	0.0057	0.0140	0.0253	0.0432	0.0579
		5	FD	0.0054	0.0068	0.0093	0.0093	0.0131	0.0032	0.0036	0.0045	0.0068	0.0090	0.0023	0.0035	0.0059	0.0090	0.0113
			SYS	0.0051	0.0064	0.0114	0.0197	0.0300	0.0031	0.0033	0.0044	0.0074	0.0106	0.0022	0.0031	0.0050	0.0080	0.0126
		10	FD	0.0027	0.0031	0.0038	0.0042	0.0061	0.0015	0.0018	0.0021	0.0030	0.0037	0.0012	0.0013	0.0020	0.0028	0.0035
			SYS	0.0026	0.0030	0.0045	0.0072	0.0114	0.0015	0.0016	0.0018	0.0025	0.0034	0.0012	0.0012	0.0017	0.0024	0.0035

Table IV-17: Relative RMSE of $\hat{\beta}$ over N, T, sw and γ

5.2.3 Relative Bias, Standard Deviation and RMSE of β* (Long-Run Effects)

N		sw			T = 5, γ = 0	0.2	0.4	0.6	0.8	T = 10, γ = 0	0.2	0.4	0.6	0.8	T = 20, γ = 0	0.2	0.4	0.6	0.8
N	50	sw	0.1	FD	-0.0333	-0.0333	-0.3458	-0.8230	-1.3319	-0.0257	-0.1467	-0.4478	-0.8534	-0.8371	-0.0225	-0.2596	-0.5222	-0.7464	-0.5601
				SYS	0.0904	0.0199	-0.1598	-0.3572	-0.5731	0.0736	-0.0189	-0.2258	-0.5234	-0.8551	0.1316	0.0040	-0.1862	-0.4140	-0.7266
			0.2	FD	-0.0172	-0.1073	-0.3157	-0.5492	1.0099	-0.0081	-0.1738	-0.4167	-0.6866	-0.5770	-0.0265	-0.2699	-0.5095	-0.6957	-0.5202
				SYS	0.0605	-0.0187	-0.1890	-0.3716	-0.5449	0.0484	-0.0790	-0.2815	-0.5404	-0.7738	0.0638	-0.1179	-0.3147	-0.5005	-0.6610
			0.5	FD	-0.0013	-0.1038	-0.2657	-0.3916	-0.8044	-0.0111	-0.1456	-0.3374	-0.5557	-0.4944	-0.0241	-0.2255	-0.4387	-0.6316	-0.5115
				SYS	0.0279	-0.0662	-0.2054	-0.3656	-0.5011	0.0103	-0.1225	-0.3046	-0.5176	-0.6662	0.0063	-0.1795	-0.3711	-0.5531	-0.5702
			1	FD	-0.0006	-0.0790	-0.1888	-0.5313	0.3534	-0.0057	-0.0848	-0.1917	-0.3414	-0.4129	-0.0142	-0.1385	-0.3041	-0.5009	-0.5070
				SYS	0.0156	-0.0658	-0.1612	-0.2772	-0.4215	0.0004	-0.0831	-0.1972	-0.3671	-0.5687	-0.0045	-0.1188	-0.2719	-0.4632	-0.5331
			2	FD	-0.0012	-0.0409	-0.0784	0.0150	-2.5029	-0.0012	-0.0303	-0.0644	-0.1147	-0.1628	-0.0066	-0.0579	-0.1362	-0.2545	-0.3933
				SYS	0.0065	-0.0400	-0.0998	-0.1580	-0.2934	0.0004	-0.0316	-0.0682	-0.1364	-0.3203	-0.0032	-0.0482	-0.1166	-0.2382	-0.4171
			5	FD	-0.0007	-0.0040	-0.0017	0.0454	0.2801	-0.0010	-0.0066	-0.0120	-0.0190	-0.0165	-0.0017	-0.0115	-0.0280	-0.0544	-0.1170
				SYS	0.0011	-0.0105	-0.0305	-0.0526	-0.1069	-0.0007	-0.0068	-0.0120	-0.0197	-0.0571	-0.0011	-0.0086	-0.0218	-0.0489	-0.1328
			10	FD	0.0005	-0.0003	0.0064	0.0244	0.0824	0.0000	-0.0018	-0.0038	-0.0050	-0.0022	-0.0005	-0.0025	-0.0075	-0.0148	-0.0303
				SYS	0.0015	-0.0034	-0.0076	-0.0150	-0.0292	0.0001	-0.0018	-0.0035	-0.0047	-0.0073	-0.0003	-0.0012	-0.0058	-0.0132	-0.0312
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	sw	0.1	FD	0.0087	-0.0579	-0.3227	-0.7542	-10.4843	-0.0091	-0.1418	-0.4554	-0.8518	-0.8271	-0.0110	-0.2483	-0.5125	-0.7129	-0.5331
				SYS	0.0008	-0.0410	-0.1606	-0.3372	-0.4914	0.0317	-0.0319	-0.2311	-0.5956	-0.9026	0.0404	-0.1058	-0.3290	-0.5932	-0.9391
			0.2	FD	0.0085	-0.0773	-0.2999	-0.5447	-0.0205	-0.0102	-0.1653	-0.4104	-0.6840	-0.5455	-0.0118	-0.2477	-0.4919	-0.6746	-0.4906
				SYS	0.0042	-0.0447	-0.1828	-0.3536	-0.4985	0.0162	-0.0789	-0.2919	-0.5919	-0.8082	0.0111	-0.1754	-0.3977	-0.6153	-0.7876
			0.5	FD	-0.0002	-0.0972	-0.2389	-0.3532	0.8128	-0.0051	-0.1389	-0.3236	-0.5492	-0.4243	-0.0073	-0.2132	-0.4326	-0.6200	-0.4849
				SYS	0.0024	-0.0791	-0.2030	-0.3409	-0.4791	0.0044	-0.1206	-0.2928	-0.5159	-0.6539	-0.0001	-0.1910	-0.3994	-0.5950	-0.6185
			1	FD	-0.0025	-0.0791	-0.1711	-0.2036	-0.5173	-0.0015	-0.0757	-0.1672	-0.3010	-0.3415	-0.0033	-0.1358	-0.3005	-0.4888	-0.4687
				SYS	0.0010	-0.0771	-0.1654	-0.2648	-0.4036	0.0016	-0.0810	-0.1817	-0.3372	-0.5286	-0.0011	-0.1261	-0.2847	-0.4784	-0.5462
			2	FD	-0.0008	-0.0355	-0.0685	-0.0224	0.8183	-0.0005	-0.0258	-0.0520	-0.0850	-0.1057	-0.0012	-0.0555	-0.1324	-0.2497	-0.3619
				SYS	0.0004	-0.0453	-0.1015	-0.1614	-0.2830	0.0004	-0.0314	-0.0646	-0.1191	-0.2719	-0.0005	-0.0499	-0.1208	-0.2464	-0.4188
			5	FD	0.0005	-0.0029	0.0031	0.0445	0.2580	-0.0001	-0.0051	-0.0095	-0.0141	-0.0088	-0.0002	-0.0106	-0.0268	-0.0527	-0.1053
				SYS	0.0006	-0.0140	-0.0343	-0.0594	-0.1227	0.0000	-0.0060	-0.0113	-0.0193	-0.0615	0.0000	-0.0086	-0.0224	-0.0516	-0.1324
			10	FD	-0.0001	0.0009	0.0059	0.0246	0.0849	0.0002	-0.0017	-0.0029	-0.0039	-0.0011	0.0000	-0.0030	-0.0073	-0.0142	-0.0281
				SYS	0.0000	-0.0042	-0.0113	-0.0200	-0.0426	0.0002	-0.0017	-0.0030	-0.0044	-0.0152	0.0001	-0.0022	-0.0059	-0.0141	-0.0323
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw	0.1	FD	0.0082	-0.0250	-0.3363	-0.7811	-0.2657	0.0001	-0.1387	-0.4561	-0.8442	-0.8083	-0.0090	-0.2537	-0.5153	-0.7222	-0.5227
				SYS	0.0186	-0.0214	-0.1491	-0.3324	-0.4811	0.0028	-0.0411	-0.2546	-0.6410	-0.9235	0.0304	-0.0918	-0.3192	-0.6135	-0.9723
			0.2	FD	0.0003	-0.0785	-0.3026	-0.5598	0.0768	-0.0023	-0.1552	-0.4120	-0.6732	-0.5348	-0.0084	-0.2502	-0.4992	-0.6828	-0.4940
				SYS	0.0135	-0.0324	-0.1753	-0.3532	-0.4942	-0.0009	-0.0855	-0.3115	-0.6154	-0.8201	0.0137	-0.1658	-0.3984	-0.6341	-0.8064
			0.5	FD	0.0004	-0.0962	-0.2435	-0.3595	1.0937	-0.0025	-0.1354	-0.3250	-0.5424	-0.4142	-0.0050	-0.2134	-0.4348	-0.6248	-0.4832
				SYS	0.0077	-0.0752	-0.2041	-0.3442	-0.4862	-0.0009	-0.1227	-0.3009	-0.5175	-0.6564	0.0012	-0.1887	-0.3978	-0.6014	-0.6266
			1	FD	-0.0014	-0.0810	-0.1743	-0.2138	0.9602	-0.0012	-0.0750	-0.1657	-0.2966	-0.3298	-0.0022	-0.1326	-0.2939	-0.4809	-0.4633
				SYS	0.0022	-0.0778	-0.1682	-0.2682	-0.4091	-0.0004	-0.0838	-0.1836	-0.3352	-0.5237	0.0002	-0.1257	-0.2802	-0.4734	-0.5465
			2	FD	-0.0004	-0.0365	-0.0691	-0.0317	0.8178	-0.0010	-0.0251	-0.0515	-0.0866	-0.1051	-0.0004	-0.0532	-0.1234	-0.2348	-0.3443
				SYS	0.0009	-0.0472	-0.1014	-0.1627	-0.2806	-0.0008	-0.0325	-0.0669	-0.1223	-0.2716	0.0003	-0.0517	-0.1168	-0.2364	-0.4074
			5	FD	0.0002	-0.0030	0.0046	0.0473	0.2598	-0.0001	-0.0050	-0.0097	-0.0147	-0.0094	-0.0001	-0.0101	-0.0248	-0.0490	-0.0957
				SYS	0.0005	-0.0136	-0.0334	-0.0593	-0.1247	-0.0001	-0.0061	-0.0120	-0.0208	-0.0626	0.0001	-0.0092	-0.0221	-0.0490	-0.1257
			10	FD	0.0000	0.0007	0.0064	0.0249	0.0889	0.0000	-0.0016	-0.0031	-0.0041	-0.0013	0.0000	-0.0027	-0.0067	-0.0131	-0.0253
				SYS	0.0001	-0.0041	-0.0107	-0.0200	-0.0420	0.0000	-0.0017	-0.0033	-0.0049	-0.0155	0.0000	-0.0023	-0.0059	-0.0131	-0.0301

Table IV-18: Relative bias $\hat{\beta}^* - \beta^*$ over N, T, sw and γ

N	sw		T = 5					T = 10					T = 20				
			γ					γ					γ				
			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
50	0.1	FD	1.0041	0.8781	0.7943	0.8858	5.2302	0.5310	0.4678	0.3817	0.3142	0.5573	0.3478	0.2759	0.1998	0.1726	0.2890
		SYS	0.3373	0.2455	0.2276	0.2221	0.8246	0.2348	0.1947	0.1628	0.1411	0.1354	0.2095	0.1772	0.1318	0.1132	0.1252
	0.2	FD	0.4863	0.4265	0.4377	0.5488	44.5403	0.2659	0.2283	0.1903	0.1684	0.2859	0.1623	0.1420	0.1058	0.0823	0.1409
		SYS	0.2656	0.2297	0.2255	0.2069	0.2851	0.1902	0.1544	0.1312	0.1144	0.1011	0.1384	0.1240	0.0950	0.0833	0.0938
	0.5	FD	0.1967	0.2000	0.1973	0.9056	9.0806	0.1179	0.1054	0.0964	0.0942	0.1632	0.0796	0.0698	0.0584	0.0474	0.0683
		SYS	0.1676	0.1607	0.1676	0.1636	0.1439	0.1056	0.0918	0.0836	0.0758	0.0650	0.0788	0.0686	0.0616	0.0502	0.0611
	1	FD	0.1088	0.1230	0.1447	10.2501	11.4399	0.0606	0.0639	0.0659	0.0740	0.1337	0.0469	0.0494	0.0455	0.0410	0.0504
		SYS	0.1008	0.1057	0.1172	0.1293	0.1279	0.0562	0.0593	0.0611	0.0646	0.0627	0.0490	0.0506	0.0485	0.0433	0.0473
	2	FD	0.0620	0.0786	0.1276	0.4971	111.5346	0.0333	0.0367	0.0390	0.0487	0.1148	0.0271	0.0284	0.0309	0.0351	0.0429
		SYS	0.0592	0.0666	0.0845	0.1032	0.1164	0.0320	0.0354	0.0376	0.0465	0.0728	0.0273	0.0298	0.0323	0.0359	0.0404
	5	FD	0.0266	0.0386	0.0610	0.1250	0.4301	0.0133	0.0146	0.0155	0.0202	0.0475	0.0109	0.0125	0.0140	0.0181	0.0322
		SYS	0.0263	0.0349	0.0479	0.0676	0.1034	0.0127	0.0138	0.0151	0.0202	0.0431	0.0110	0.0124	0.0141	0.0191	0.0326
	10	FD	0.0134	0.0194	0.0293	0.0522	0.1110	0.0067	0.0072	0.0076	0.0101	0.0214	0.0055	0.0062	0.0070	0.0094	0.0175
		SYS	0.0136	0.0188	0.0257	0.0399	0.0721	0.0066	0.0070	0.0076	0.0102	0.0216	0.0057	0.0062	0.0070	0.0096	0.0180
			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
250	0.1	FD	0.4111	0.3894	0.3547	0.3801	278.8005	0.2454	0.2110	0.1688	0.1520	0.3092	0.1495	0.1226	0.0938	0.0776	0.1297
		SYS	0.1196	0.0938	0.0948	0.0879	0.0708	0.0942	0.0822	0.0942	0.0785	0.0777	0.0980	0.0809	0.0697	0.0540	0.0508
	0.2	FD	0.2028	0.1829	0.1849	0.2036	4.5739	0.1253	0.1050	0.0865	0.0787	0.1645	0.0750	0.0656	0.0508	0.0393	0.0666
		SYS	0.1030	0.0881	0.0942	0.0915	0.0723	0.0795	0.0726	0.0720	0.0578	0.0555	0.0616	0.0573	0.0469	0.0394	0.0377
	0.5	FD	0.0828	0.0829	0.0872	0.1127	18.2929	0.0480	0.0482	0.0464	0.0449	0.0977	0.0373	0.0324	0.0260	0.0226	0.0301
		SYS	0.0686	0.0682	0.0707	0.0728	0.0643	0.0439	0.0435	0.0407	0.0351	0.0366	0.0354	0.0314	0.0257	0.0220	0.0233
	1	FD	0.0447	0.0505	0.0601	0.1042	45.7394	0.0273	0.0284	0.0319	0.0404	0.0755	0.0217	0.0223	0.0206	0.0179	0.0209
		SYS	0.0420	0.0457	0.0499	0.0572	0.0601	0.0259	0.0258	0.0270	0.0306	0.0318	0.0218	0.0221	0.0208	0.0184	0.0181
	2	FD	0.0269	0.0358	0.0508	0.1118	2.0144	0.0145	0.0159	0.0178	0.0238	0.0615	0.0121	0.0130	0.0139	0.0150	0.0194
		SYS	0.0246	0.0294	0.0360	0.0436	0.0560	0.0141	0.0151	0.0165	0.0204	0.0308	0.0120	0.0128	0.0142	0.0157	0.0176
	5	FD	0.0118	0.0167	0.0259	0.0522	0.1503	0.0058	0.0066	0.0070	0.0089	0.0204	0.0054	0.0056	0.0064	0.0082	0.0143
		SYS	0.0111	0.0146	0.0199	0.0282	0.0412	0.0058	0.0065	0.0068	0.0087	0.0163	0.0053	0.0055	0.0063	0.0084	0.0142
	10	FD	0.0059	0.0083	0.0134	0.0221	0.0492	0.0029	0.0031	0.0035	0.0043	0.0095	0.0026	0.0028	0.0032	0.0041	0.0080
		SYS	0.0059	0.0080	0.0119	0.0170	0.0293	0.0029	0.0032	0.0034	0.0042	0.0091	0.0025	0.0028	0.0032	0.0043	0.0082
			0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
500	0.1	FD	0.2960	0.2611	0.2434	0.2592	18.2195	0.1610	0.1398	0.1188	0.1148	0.2192	0.1134	0.0904	0.0709	0.0559	0.0942
		SYS	0.0810	0.0672	0.0680	0.0648	0.0519	0.0612	0.0575	0.0711	0.0569	0.0523	0.0629	0.0546	0.0545	0.0409	0.0378
	0.2	FD	0.1448	0.1324	0.1282	0.1388	6.5709	0.0801	0.0718	0.0633	0.0583	0.1203	0.0573	0.0478	0.0365	0.0288	0.0470
		SYS	0.0738	0.0663	0.0676	0.0688	0.0520	0.0532	0.0483	0.0533	0.0404	0.0412	0.0457	0.0415	0.0344	0.0272	0.0280
	0.5	FD	0.0566	0.0573	0.0582	0.0770	9.4820	0.0342	0.0334	0.0322	0.0333	0.0635	0.0269	0.0233	0.0202	0.0157	0.0215
		SYS	0.0482	0.0479	0.0534	0.0540	0.0492	0.0300	0.0294	0.0278	0.0263	0.0256	0.0246	0.0227	0.0199	0.0151	0.0171
	1	FD	0.0320	0.0354	0.0422	0.0702	8.1905	0.0190	0.0207	0.0224	0.0278	0.0541	0.0155	0.0156	0.0158	0.0135	0.0151
		SYS	0.0295	0.0314	0.0370	0.0424	0.0422	0.0176	0.0186	0.0188	0.0210	0.0231	0.0150	0.0149	0.0154	0.0132	0.0131
	2	FD	0.0187	0.0255	0.0355	0.0724	0.8894	0.0097	0.0108	0.0128	0.0165	0.0422	0.0090	0.0099	0.0105	0.0110	0.0144
		SYS	0.0176	0.0210	0.0262	0.0317	0.0393	0.0094	0.0102	0.0116	0.0136	0.0224	0.0089	0.0096	0.0105	0.0109	0.0125
	5	FD	0.0082	0.0124	0.0194	0.0365	0.1111	0.0041	0.0045	0.0050	0.0064	0.0149	0.0038	0.0041	0.0048	0.0057	0.0104
		SYS	0.0079	0.0106	0.0142	0.0203	0.0288	0.0041	0.0045	0.0048	0.0062	0.0118	0.0038	0.0040	0.0047	0.0057	0.0101
	10	FD	0.0042	0.0061	0.0092	0.0159	0.0348	0.0021	0.0022	0.0024	0.0030	0.0068	0.0020	0.0020	0.0023	0.0029	0.0055
		SYS	0.0041	0.0058	0.0081	0.0121	0.0214	0.0021	0.0022	0.0024	0.0030	0.0064	0.0020	0.0020	0.0023	0.0029	0.0057

Table IV-19: Relative standard deviation of $\hat{\beta}^*$ over N, T, sw and γ

					T														
					5					10					20				
					γ					γ					γ				
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
N	50	sw	0.1	FD	1.0042	0.8782	0.8659	1.2088	5.3946	0.5314	0.4901	0.5882	0.9093	1.0054	0.3483	0.3788	0.5591	0.7661	0.6302
				SYS	0.3490	0.2462	0.2780	0.4206	1.0038	0.2459	0.1955	0.2783	0.5421	0.8657	0.2473	0.1772	0.2281	0.4292	0.7373
			0.2	FD	0.4864	0.4396	0.5395	0.7762	44.5295	0.2659	0.2869	0.4580	0.7069	0.6439	0.1644	0.3050	0.5204	0.7006	0.5390
				SYS	0.2723	0.2303	0.2941	0.4252	0.6149	0.1962	0.1734	0.3106	0.5524	0.7804	0.1523	0.1711	0.3287	0.5074	0.6676
			0.5	FD	0.1966	0.2252	0.3309	0.9862	9.1117	0.1183	0.1797	0.3509	0.5636	0.5206	0.0831	0.2360	0.4426	0.6333	0.5160
				SYS	0.1699	0.1738	0.2651	0.4005	0.5214	0.1060	0.1530	0.3159	0.5232	0.6693	0.0790	0.1922	0.3762	0.5554	0.5735
			1	FD	0.1087	0.1462	0.2378	10.2588	11.4397	0.0609	0.1062	0.2027	0.3493	0.4340	0.0489	0.1470	0.3074	0.5025	0.5095
				SYS	0.1019	0.1245	0.1993	0.3058	0.4404	0.0562	0.1021	0.2064	0.3727	0.5721	0.0492	0.1292	0.2762	0.4652	0.5352
			2	FD	0.0620	0.0885	0.1497	0.4971	111.5069	0.0333	0.0476	0.0752	0.1246	0.1992	0.0279	0.0645	0.1397	0.2569	0.3956
				SYS	0.0596	0.0777	0.1307	0.1887	0.3156	0.0320	0.0475	0.0778	0.1442	0.3285	0.0274	0.0566	0.1210	0.2409	0.4190
			5	FD	0.0266	0.0388	0.0610	0.1329	0.5131	0.0133	0.0159	0.0196	0.0278	0.0502	0.0111	0.0170	0.0313	0.0574	0.1214
				SYS	0.0263	0.0364	0.0568	0.0856	0.1487	0.0128	0.0154	0.0192	0.0282	0.0715	0.0111	0.0151	0.0260	0.0525	0.1368
			10	FD	0.0134	0.0194	0.0300	0.0576	0.1383	0.0067	0.0074	0.0085	0.0113	0.0215	0.0055	0.0067	0.0103	0.0175	0.0349
				SYS	0.0136	0.0190	0.0267	0.0426	0.0777	0.0066	0.0071	0.0083	0.0112	0.0228	0.0057	0.0063	0.0091	0.0163	0.0360
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	250	sw	0.1	FD	0.4110	0.3935	0.4794	0.8445	278.8582	0.2454	0.2542	0.4856	0.8652	0.8829	0.1498	0.2769	0.5210	0.7171	0.5486
				SYS	0.1196	0.1024	0.1865	0.3484	0.4964	0.0994	0.0882	0.2495	0.6008	0.9060	0.1060	0.1332	0.3362	0.5956	0.9404
			0.2	FD	0.2028	0.1985	0.3523	0.5815	4.5716	0.1256	0.1958	0.4194	0.6886	0.5697	0.0759	0.2562	0.4945	0.6757	0.4951
				SYS	0.1031	0.0987	0.2056	0.3652	0.5037	0.0811	0.1071	0.3006	0.5947	0.8101	0.0625	0.1845	0.4004	0.6166	0.7885
			0.5	FD	0.0827	0.1278	0.2543	0.3707	18.3018	0.0483	0.1470	0.3269	0.5510	0.4354	0.0380	0.2157	0.4334	0.6204	0.4858
				SYS	0.0686	0.1044	0.2150	0.3486	0.4834	0.0441	0.1282	0.2956	0.5170	0.6549	0.0353	0.1936	0.4003	0.5954	0.6190
			1	FD	0.0448	0.0938	0.1814	0.2287	45.7194	0.0274	0.0808	0.1702	0.3036	0.3497	0.0220	0.1375	0.3013	0.4891	0.4692
				SYS	0.0420	0.0896	0.1727	0.2709	0.4080	0.0259	0.0850	0.1837	0.3386	0.5296	0.0218	0.1280	0.2854	0.4787	0.5465
			2	FD	0.0269	0.0504	0.0852	0.1139	2.1733	0.0145	0.0302	0.0550	0.0883	0.1223	0.0122	0.0570	0.1331	0.2502	0.3624
				SYS	0.0246	0.0540	0.1077	0.1672	0.2885	0.0141	0.0349	0.0666	0.1208	0.2736	0.0120	0.0515	0.1216	0.2469	0.4192
			5	FD	0.0118	0.0170	0.0261	0.0686	0.2986	0.0058	0.0083	0.0118	0.0167	0.0222	0.0054	0.0120	0.0276	0.0533	0.1063
				SYS	0.0111	0.0202	0.0396	0.0658	0.1294	0.0058	0.0088	0.0132	0.0212	0.0636	0.0053	0.0102	0.0233	0.0523	0.1331
			10	FD	0.0059	0.0083	0.0147	0.0330	0.0982	0.0029	0.0035	0.0046	0.0058	0.0096	0.0026	0.0041	0.0080	0.0148	0.0292
				SYS	0.0059	0.0090	0.0164	0.0262	0.0517	0.0029	0.0036	0.0046	0.0061	0.0177	0.0025	0.0035	0.0068	0.0147	0.0333
					0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8	0	0.2	0.4	0.6	0.8
	500	sw	0.1	FD	0.2959	0.2622	0.4151	0.8230	18.2123	0.1609	0.1969	0.4713	0.8520	0.8375	0.1137	0.2693	0.5202	0.7244	0.5311
				SYS	0.0831	0.0705	0.1639	0.3386	0.4839	0.0612	0.0706	0.2643	0.6435	0.9250	0.0698	0.1068	0.3238	0.6149	0.9730
			0.2	FD	0.1447	0.1538	0.3286	0.5767	6.5681	0.0801	0.1710	0.4168	0.6757	0.5481	0.0578	0.2547	0.5005	0.6834	0.4962
				SYS	0.0750	0.0738	0.1879	0.3598	0.4969	0.0532	0.0982	0.3161	0.6167	0.8212	0.0476	0.1710	0.3999	0.6347	0.8069
			0.5	FD	0.0566	0.1119	0.2503	0.3677	9.5402	0.0343	0.1395	0.3266	0.5434	0.4191	0.0273	0.2147	0.4352	0.6250	0.4837
				SYS	0.0488	0.0891	0.2110	0.3484	0.4886	0.0300	0.1262	0.3022	0.5181	0.6569	0.0246	0.1901	0.3983	0.6015	0.6268
			1	FD	0.0321	0.0884	0.1793	0.2250	8.2425	0.0190	0.0778	0.1672	0.2979	0.3342	0.0156	0.1335	0.2944	0.4811	0.4635
				SYS	0.0295	0.0840	0.1721	0.2715	0.4112	0.0176	0.0858	0.1846	0.3358	0.5242	0.0150	0.1266	0.2806	0.4735	0.5467
			2	FD	0.0187	0.0446	0.0777	0.0790	1.2079	0.0098	0.0274	0.0531	0.0882	0.1133	0.0090	0.0542	0.1238	0.2350	0.3446
				SYS	0.0176	0.0516	0.1048	0.1658	0.2833	0.0094	0.0340	0.0679	0.1230	0.2726	0.0089	0.0526	0.1172	0.2367	0.4076
			5	FD	0.0082	0.0127	0.0199	0.0597	0.2826	0.0041	0.0066	0.0109	0.0160	0.0177	0.0038	0.0109	0.0253	0.0493	0.0963
				SYS	0.0079	0.0173	0.0362	0.0627	0.1280	0.0041	0.0075	0.0129	0.0217	0.0637	0.0038	0.0101	0.0226	0.0494	0.1261
			10	FD	0.0042	0.0062	0.0113	0.0296	0.0954	0.0021	0.0028	0.0040	0.0051	0.0069	0.0020	0.0034	0.0071	0.0134	0.0259
				SYS	0.0041	0.0070	0.0134	0.0234	0.0472	0.0021	0.0028	0.0041	0.0058	0.0168	0.0020	0.0030	0.0064	0.0134	0.0306

Table IV-20: Relative RMSE of $\hat{\beta}^*$ over N, T, sw and γ

				Researcher's Focus / Interest								
				γ			β			β*=β/(1−γ)		
				Bias	Variance	RMSE	Bias	Variance	RMSE	Bias	Variance	RMSE
sw low	N low	T low	γ low	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
			γ large	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
		T large	γ low	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
			γ large	**FD-GMM**	**SYS-GMM**	**FD-GMM**	**FD-GMM**	**SYS-GMM**	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
	N large	T low	γ low	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
			γ large	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**FD-GMM**	**SYS-GMM**	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
		T large	γ low	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
			γ large	**FD-GMM**	**FD-GMM**	**FD-GMM**	**FD-GMM**	**SYS-GMM**	**FD-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
sw large	N low	T low	γ low	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM	FD-GMM	SYS-GMM	SYS-GMM
		T large	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM
	N large	T low	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM
		T large	γ low	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
			γ large	FD-GMM	SYS-GMM	FD-GMM	SYS-GMM	SYS-GMM	SYS-GMM	FD-GMM	SYS-GMM	FD-GMM

Table IV-21: Decision rules according to the Monte Carlo experiment (full set of instruments used)

Bibliography

Ahn, S.C. & Schmidt, P., 1995. Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, 68(1), pp.5–27.

Alberini, A. & Filippini, M., 2011. Response of residential electricity demand to price: The effect of measurement error. *Energy Economics*.

Anderson, T.W. & Hsiao, Cheng, 1981. Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*, 76(375), pp.598–606.

Anselin, L., 2010. Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1), pp.3–25.

Anselin, L., Gallo, J.L. & Jayet, H., 2008. Spatial panel econometrics. *The econometrics of panel data*, pp.625–660.

Arellano, M. & Bond, S., 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2), pp.277–297.

Asplund, M., Friberg, R. & Wilander, F., 2007. Demand and distance: evidence on cross-border shopping. *Journal of public Economics*, 91(1-2), pp.141–157.

Balestra, P. & Nerlove, M., 1966. Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas. *Econometrica*, 34(3), pp.585–612.

Baltagi, B. H & Li, D., 2006. Prediction in the Panel Data Model with Spatial Correlation: The Case of Liquor. *Center for Policy Research. Paper*, 81. Available at: http://surface.syr.edu/cpr/81.

Baltagi, B. H. et al., 2007. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics*, 140(1), pp.5–51.

Baltagi, B.H., 2005. *Econometric Analysis of Panel Data*, Wiley. Available at: http://books.google.ch/books?id=yTVSqmufge8C.

Baltagi, B.H., Egger, P. & Pfaffermayr, M., 2009. A generalized spatial panel data model with random effects. *Center for Policy Research Working Paper No*.

Baltagi, B.H. & Griffin, J.M., 1984. Short and long run effects in pooled models. *International Economic Review*, 25(3), pp.631–645.

Baltagi, B.H. & Long, L., 2008. Testing for random effects and spatial lag dependence in panel data models. *Center for Policy Research Working Paper No 102*.

Baltagi, B.H., Song, S.H. & Koh, W., 2003. Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117(1), pp.123–150.

Baltagi, Badi H. et al., 2003. Homogeneous, heterogeneous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption. *Empirical Economics*, 28, pp.795–811.

Banfi, S., Filippini, M. & Hunt, L.C., 2005. Fuel tourism in border regions: The case of Switzerland. *Energy economics*, 27(5), pp.689–707.

Banfi, S., Horehajava, A. & Filippini, M., 2006. Hedonic price functions for Zurich and Lugano with special focus on electrosmog. In *3rd World Congress of Environmental and Resource Economists*.

Baranzini, A. & Weber, S., (last), 2012. *Élasticité-prix de la demande d'essence en Suisse*, Haute école de gestion de Genève. Available at: http://doc.rero.ch/lm.php?url=1000,44,9,20120117094514-WC/CR_1_12_Baranzini.pdf.

Blázquez, L., Boogen, N. & Filippini, Massimo, 2013. Residential electricity demand in Spain: New empirical evidence using aggregate data. *Energy Economics*, 36(0), pp.648–657.

Blonigen, B.A. et al., 2007. FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review*, 51(5), pp.1303–1325.

Blundell, R. & Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), pp.115–143.

Brons, M.R.E. et al., 2006. *A meta-analysis of the price elasticity of gasoline demand: a system of equations approach*, Tinbergen Institute. Available at: http://books.google.com/books?id=1F_LPAAACAAJ.

Clark, T.S. & Linzer, D.A., 2012. *Should I Use Fixed or Random Effects?*, Working paper, available at http://polmeth. wustl. edu/mediaDetail. php. Available at: http://userwww.service.emory.edu/~dlinzer/ClarkLinzer-REFE-Mar2012.pdf [Accessed December 15, 2012].

Coats, R.M., 1995. A note on estimating cross-border effects of state cigarette taxes. *National Tax Journal*, 48, pp.573–584.

A. Colin Cameron and Pravin K. Trivedi, *Microeconometrics Using STATA, Revised Edtion*, Stata Press.

Coughlin, C.C. & Segev, E., 2000. Foreign Direct Investment in China: A Spatial Econometric Study. *World Economy*, 23(1), pp.1–23.

Dahl, C.A., 2012. Measuring global gasoline and diesel price and income elasticities. *Energy Policy*, 41(0), pp.2–13.

Deaton, A. & Muellbauer, J., 1980. *Economics and Consumer Behavior*, Cambridge University Press. Available at: http://books.google.ch/books?id=B81RYQsx2l0C.

Debarsy, N. & Ertur, C., 2010. Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, 40(6), pp.453–470.

Drukker, D. M, Prucha, I. R & Raciborski, R., 2011. *A command for estimating spatial-autoregressive models with spatial-autoregressive disturbances and additional endogenous variables*, Technical report, Stata.

Drukker, David M., 2008. Econometric analysis of dynamic panel-data models using Stata. In *Summer North American STATA Users Group Meeting, July 2008*. Available at: http://stata.mobi/meeting/snasug08/drukker_xtdpd.pdf [Accessed March 18, 2013].

Egger, P. & Larch, M., 2008. Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of International Economics*, 76(2), pp.384–399.

Egger, P., Pfaffermayr, M. & Winner, H., 2005a. An unbalanced spatial panel data approach to US state tax competition. *Economics Letters*, 88(3), pp.329–335.

Egger, P., Pfaffermayr, M. & Winner, H., 2005b. Commodity taxation in a [] linear'world: a spatial panel data approach. *Regional Science and Urban Economics*, 35(5), pp.527–541.

Elhorst, J.P., 2010. Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, 5, pp.9–28.

Farsi, M. & Filippini, M., 2004. Regulation and measuring cost-efficiency with panel data models: application to electricity distribution utilities. *Review of Industrial Organization*, 25(1), pp.1–19.

Greene, W.H., 2003. *Econometric analysis*, Prentice Hall. Available at: http://books.google.ch/books?id=JJkWAQAAMAAJ.

H. Kelejian, H. & Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*, 104(2), pp.219–257.

Hsiao, C., 2007. Panel data analysis—advantages and challenges. *Test*, 16(1), pp.1–22.

Kapoor, M., Kelejian, Harry H. & Prucha, Ingmar R., 2007. Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1), pp.97–130.

Kelejian, H.H. & Prucha, I.R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1), pp.99–121.

Kiviet, J.F., 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68(1), pp.53–78.

Kmenta, J., 1997. *Elements of econometrics*, University of Michigan Press. Available at: http://books.google.ch/books?id=w9TrAAAAMAAJ.

Leal, A., López-Laborda, J. & Rodrigo, F., 2009. Prices, taxes and automotive fuel cross-border shopping. *Energy Economics*, 31(2), pp.225–234.

Lee, L. & Yu, J., 2010. Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40(5), pp.255–271.

Lee, L.F., 2004. Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica*, 72(6), pp.1899–1925.

LeSage, J.P. & Pace, R.K., 2009. *Introduction to spatial econometrics*, CRC Press. Available at: http://books.google.com/books?id=EKiKXcgL-D4C.

Di Matteo, L. & Di Matteo, R., 1996. An analysis of Canadian cross-border travel. *Annals of Tourism Research*, 23(1), pp.103–122.

Michaelis, P., 2004. Tanktourismus–eine szenario-analyse. *Zeitschrift für Verkehrswissenschaft*, 75(2), pp.110–125.

Pennerstorfer, D., 2008. Strategische Interaktion und r\äumlicher Preiswettbewerb im Treibstoffeinzelhandel. *Eine r\äumlich-\ökonometrische Analyse, WU Vienna University of Economics and Business*.

Pfaffermayr, Michael, 2011. The Hausman test in a Cliff. and Ord. panel model. *The Econometrics Journal*, 14(1), pp.48–76.

Pirotte, A. & Madre, J.L., 2011. Car Traffic Elasticities A Spatial Panel Data Analysis of French Regions. *Journal of Transport Economics and Policy (JTEP)*, 45(3), pp.341–365.

Pock, M., 2010. Gasoline demand in Europe: New insights. *Energy Economics*, 32(1), pp.54–62.

Rietveld, P., Bruinsma, F. & Van Vuuren, D., 2001. Spatial graduation of fuel taxes; consequences for cross-border and domestic fuelling. *Transportation Research Part A: Policy and Practice*, 35(5), pp.433–457.

Roodman, D., 2007. How to do xtabond2: An introduction to difference and system GMM in Stata.

Schleiniger, R. & Universität Zürich. Institut für Empirische Wirtschaftsforschung, 1995. *The demand for gasoline in Switzerland: in the short and in the long run*, Institute for Empirical Research in Economics, University of Zurich. Available at: http://books.google.com/books?id=_94zGwAACAAJ.

Sen, M. & Bera, A.K., 2011. Specification Testing for Panel Spatial Models.

Stakhovych, S. & Bijmolt, T.H.A., 2009. Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, 88(2), pp.389–408.

Wasserfallen, W. & G\üntensperger, H., 1988. Gasoline consumption and the stock of motor vehicles:: An empirical analysis for the Swiss economy. *Energy economics*, 10(4), pp.276–282.

Ziliak, J.P., 1997. Efficient Estimation With Panel Data When Instruments Are Predetermined: An Empirical Comparison of Moment-Condition Estimators. *Journal of Business & Economic Statistics*, 15(4), pp.419–431.

Curriculum Vitae

I was born on 8 September 1981 as the second twin of Marianne and Roland Heimsch.

EDUCATION

2009 - 2012	Doctoral studies at the Centre for Energy Policy and Economics (CEPE), ETH Zurich
2005 - 2007	Master of Science ETH in Management, Technology and Economics, ETH Zurich
2002 - 2005	Bachelor of Science in Mechanical Engineering, ETH Zurich
1995 - 2001	Grammar school at Kantonsschule am Burggraben, St. Gallen
1988 - 1994	Basic education

WORK EXPERIENCE

2012 - present	Lecturer in statistics at Kalaidos Fachhochschule Schweiz, Zurich
2006 - 2012	Lecturer in physics, chemistry and mathematics at Wirtschaftsinformatikschule Schweiz (WISS) and IFA, Zurich
2004 - 2006	Assistant at the chairs of fluid dynamics and mechanics
2006	Internship in Consulting, b&m management, Zurich
2005	Internship Bühler AG, Uzwil and Pöyry Group, Zurich