

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Real Estate Market and Financial Stability in US Metropolitan Areas: a Dynamic Model with Spatial Effects

### This is the author's manuscript

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1659271> since 2018-01-28T11:54:18Z

*Published version:*

DOI:10.1016/j.regsciurbeco.2014.08.003

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Real Estate Market and Financial Stability in US Metropolitan Areas: a Dynamic Model with Spatial Effects\*

Francesco Moscone,  
Brunel University<sup>†</sup>

Elisa Tosetti,  
Brunel University and CEPE, ETH Zurich

Alessandra Canepa  
Brunel University

August 8, 2014

## Abstract

This paper investigates spatio-temporal variations in ex-post credit risk in the United States, as a function of real estate prices, loan purchases made by government sponsored enterprises, and a set of local characteristics during the recent housing boom and bust.

We model bank's non-performing loans as a first-order dynamic panel data regression model with group-specific effects and spatial autoregressive errors. To estimate this model, we develop an *ad-hoc* generalized method of moments procedure which consists of augmenting moments proposed by the panel literature to estimate short  $T$ , pure dynamic panels, with a set of quadratic conditions in the disturbances. Results on estimation of the empirical model point at the negative impact of real estate prices on non-performing loans. Further, our results show that a rise in the number of real estate mortgages backed by government-sponsored enterprises increase non-performing loans, thus deteriorating the quality of banks' loan portfolio.

**Keywords:** Non performing loans, house prices, dynamic panels, spatial dependence, GMM estimator.

**JEL Classification:** C10, C31, C33.

---

\*The authors acknowledge financial support from ESRC (Ref. no. RES-061-25-0317).

<sup>†</sup>Corresponding author: Francesco Moscone. email: moscone.francesco@brunel.ac.uk. Tel: +44(0)1895266833. Room: EJ057. Fax: +4401142685766. Brunel Business School, Uxbridge, Middlesex UB8 3PH, United Kingdom.

# 1 Introduction

Since the second half of 2007, the United States experienced a severe financial crisis that spread to the financial sector of European and Asian economies and triggered a deep, worldwide, recession. The US housing market and its interaction with the financial system has been pointed as the main cause of such crisis, through the build-up of a bubble in real estate markets that eventually collapsed.

Housing booms and busts are often associated with systemic financial stress (Herring and Wachter (1999), Reinhart and Rogoff (2008)). Among others, Reinhart and Rogoff (2008) show that six major historical episodes of banking crises in advanced economies, since mid-1970, were associated with housing-bust. The authors also report that this pattern was found in many emerging countries including the devastating Asian financial crisis of 1997-1998. A number of studies on the recent housing boom in the United States emphasise the link between a decrease in lending standards and a sharp expansion in loan delinquency in the prime and subprime mortgage market (see, for example Dell’Ariccia and Marquez (2006), and Mayer, Pence, and Sherlund (2009)). The rapid boom in house prices experienced over the period 1997-2005, accompanied by a reduction in lending standards, led to the point that many people were able to purchase properties they couldn’t afford otherwise. Over these years, an important component of the mortgage credit was in the form of subprime lending targeted to borrowers providing little or no down payment, with questionable and troubled credit histories, and minimal income requirement for loan origination. Mortgages with balloon payments, variable interest rates, and/or interest-only periods, were often sold on the presumption that individuals could refinance their mortgages at later stages. Further, individual mortgages were put into pools of assets out of which the so-called mortgage backed securities were created and sold both within the US and abroad. When house prices began to fall below the nominal value of loans, both speculative buyers and owner-occupiers that were unable to repay their mortgages could not roll them or sell their properties and, as a consequence, started to default. Public opinion has also pointed at the involvement of government-sponsored enterprises (GSEs) Fannie Mae and Freddie Mac<sup>1</sup> in the subprime mortgage market as bearing responsibility for the financial crisis. These agencies set affordable housing goals in order to support mortgages to low-income borrowers and other high-risk groups, in specific neighbourhoods and geographic areas, by purchasing and securitising mortgages in the form of mortgage-backed securities. For several years, these GSEs have provided safe and stable means of lending to buyers who did not have access to prime credit. However, in the more recent years, with the growth of private-label securitization, we assist to a deterioration in mortgage underwriting standards of GSEs and excessive risk taking, to compete with private-label securitization for market share. This has resulted in the oversupply of underpriced housing finance that led, in 2006, to an increasing number of borrowers, often with poor credit, who were unable to pay their mortgages, ultimately causing a rapid increase in home foreclosures (Bolotnyy (2012)).

In this paper we investigate spatio-temporal variations in ex-post credit risk as a function of real

---

<sup>1</sup>In the United States, the most common securitisation trusts are Fannie Mae and Freddie Mac, US government-sponsored enterprises. Ginnie Mae, a US government-sponsored enterprise backed by the full faith and credit of the US government, guarantees its investors receive timely payments, but buys limited numbers of mortgage notes. Other private institutions also securitize mortgages. These are known as "private-label" mortgage securities.

estate prices, GSEs loan purchases, and a set of local, socio-economic characteristics in the United States. We take non-performing loans (NPLs) as proxy for ex-post credit risk. As a proxy for real estate prices we focus on house prices of residential properties, using data from the Federal Housing Finance Agency on loan purchases made by the GSEs Fannie Mae and Freddie Mac. We explore the impact of house prices on NPLs across US metropolitan areas, both in the period of housing boom, in the years 2000 to 2005, and during the house-price bubble bursting, over the years 2006 to 2011. Dividing the sample period into two subsets is also justified by the structural break in house prices observed towards the end of the first sub-period.

Economic theory has formulated a number of hypotheses to explain the relationship between financial stability and real estate prices. Some authors suggest that increases in house prices reduce the risk of real estate financing perceived by banks, thus inducing excessive lending to risky real estate borrowers (Dell’Ariccia and Marquez (2006)). In addition, rising house prices may encourage the riskiest investors to bet on further price increases, leading to a rise in the demand of credit. These factors work in the same direction and tend to increase the bank exposure to risky assets, thus suggesting a positive relationship between NPLs and real estate prices, as increasing bank loans also increase ex-post credit risk. Other theories instead predict a negative relation. For example, the collateral value hypothesis asserts that, in a period of rising house prices, the value of the collateral increases thus improving borrowers’ financial position, which in turn reduces the associated risk of default (Koetter and Poghosyan (2010)). During the bursting of the bubble, theoretical models also suggest that, when house prices start to fall below the nominal value of loans, both speculative buyers and owner-occupiers that are unwilling or unable to repay their mortgages, find it difficult to roll over their loans or sell their properties. As a consequence, default rates increase and we expect a negative relationship between NPLs and real estate prices.

In our empirical application, we also wish to investigate how loan purchases made by the GSEs Fannie Mae and Freddie Mac affect NPLs, both before and during the house-price bubble bursting. These agencies had more than 40 percent of total US mortgage debt outstanding on their balance sheets at the height of the housing bubble, and experienced a financial collapse along with the rest of the market. Understanding the size of impact of GSEs’ loan purchases on NPLs is of great interest for institutional investors, and policy makers wishing to regulate the housing market.

Previous studies on the determinants of NPLs use data either at country- or at bank-level. However, data at country-level do not allow to capture the heterogeneity within an economy. For example, there is a wide range of variation in the structure and performance of the housing market across the US territory in terms of housing values exposure to subprime loans, foreclosure rates as well as demographic and economic factors. It is likely that these variations in housing market are reflected in the quality of bank loan portfolios. Hence, differently from previous works, in this paper we consider as statistical unit the Metropolitan Statistical Area (MSA). A MSA is a geographical region in the US with a relatively high population density at its core, and close economic ties throughout the area. Given that this unit of aggregation is the target of many policy interventions, exploring the impact of real estate prices and other local characteristics on credit risk at MSA level is of great interest.

Following existing literature, we adopt a dynamic specification for NPLs and focus on a first-order dynamic panel data regression model. We condition on a set of macroeconomic indicators,

such as personal income and unemployment, that are well known to influence borrowers' balance sheet and their debt servicing capacity. However, we observe that other socio-economic factors may also affect NPLs, such as the degree of urbanization, deprivation and crime, which are notoriously difficult to quantify and are well known to be geographically concentrated. Accordingly, in our empirical model we allow errors to be spatially correlated and assume that they follow a spatial autoregressive process. Ignoring spatial dependence, when this is present in the data, leads to inefficient estimates, which may cause wrong inferences. The availability of reliable models is very important for all market participants, including institutional investors, those who regulate housing, GSEs, mortgage lenders, and related financial institutions. In our regression specification, we also incorporate MSA-specific effects, and control for MSA-specific heteroskedasticity, to allow for heterogeneity in the characteristics of borrowers across different MSAs.

To estimate this model, we develop an *ad-hoc* generalized method of moments (GMM) procedure which consists of augmenting moments proposed by the panel literature to estimate pure dynamic panels, with a set of quadratic conditions in the disturbances. Recent years have witnessed an emerging interest towards the use of GMM for estimating regression models with spatially correlated disturbances. The proposed model is in line with the framework advanced by Mutl (2006). However, the work in Mutl (2006) relays on the restrictive assumption of homoskedastic group-specific effects and idiosyncratic errors, which does not apply to our empirical study. For example, it is likely that the conditional variance of ex-post credit risk varies with characteristics such as house prices, or income. From a computational point of view, our proposed approach is also simpler and more parsimonious as it requires a smaller set of parameters to be estimated. This method is a general procedure that can be used in many other areas of economics such as labour economics, health economics or macroeconomics to name few, where the variable of interest is characterized by both spatial and temporal patterns. We test the small sample properties of our GMM estimator by the means of a small Monte Carlo exercise, presented in the Appendix.

In the following, Section 2 briefly reviews the literature on the determinants of NPLs. Section 3 sets out the framework of a regression model with spatially correlated disturbances, while Section 4 introduces the GMM estimator. Sections 5 and 6 describe data and empirical results, respectively. Finally, Section 7 concludes.

## 2 Review of related literature

A number of studies examine the relationship between the real estate market and bank exposures. Hilbers, Lei, and Zacho (2001) use probit and logit models to estimate the likelihood of a financial crisis conditional on country characteristics and the real residential property price index, for 11 countries. The authors find that a downturn in residential property prices increases the probability of financial sector distress. Koetter and Poghosyan (2010) test two competing views of the relation between nominal estate prices and bank distress: the *collateral value* and the *deviation* hypotheses. The former suggests a negative relation because increasing house prices rises the market value of collateral on outstanding real estate loans, thus enhancing the financial positions of bank customers (Kiyotaki and Moore (1997)). The deviation theory conjectures that soaring house prices can lead to the accumulation of risks by banks due to moral hazard and adverse selection problems, inducing

banks to lend excessively to risky real estate borrowers at unreasonably low rates (Bernanke and Gertler (1995)). According to this view, departures of house prices from their fundamental value increase bank's probability of default. Using data on 78 regional real estate markets in Germany, Koetter and Poghosyan (2010) find evidence that larger departures of house prices from their fundamental value increase the bank's probability of default, as stated in the deviation hypothesis. Gimeno and Martinez-Carrascal (2010) use Spanish data and find evidence that house purchase loans depend positively on house prices. However, they also find evidence for causality from loans to prices when loans depart from their long-run levels. An alternative approach is taken by Mian and Sufi (2009), who investigate the reasons for the rapid expansion in the supply of mortgage credit and increase in house prices in the period 2001 to 2005, and the subsequent mortgage default crisis of 2007, at zip code level, in the US. The authors wish to explore whether the rapid growth in mortgage debt and house price are due to a greater willingness by lenders to assume risk that led to a reduction in the risk premium (*supply* explanation), or rather to increases in productivity or economic opportunities (*demand* explanation). They find that zip codes with high unfulfilled demand (at the beginning of the sample period) experienced a sharp relative decrease in denial rates and a relative increases in mortgage credit and house prices over time, despite the fact that they also experienced negative relative income and employment growth. Results are strongly consistent with the supply hypothesis, also pointing at the important role of securitization in credit expansion.

Endogenous developments in the financial market can greatly amplify the effect of small income shocks, through the so-called *financial accelerator* mechanism (Bernanke, Gertler, and Gilchrist (1996)). In particular, positive shocks to household income translate into wider house price increases in geographical areas where people can borrow against a larger fraction of their housing value (thus having a high loan-to-value) such as in the US and UK, and smaller in countries where such leverage ratios are lower (e.g. Italy). Empirical evidence on such financial accelerator for a set of countries can be found in Almeida, Campello, and Liu (2006).

Empirical research also suggests that banks bad loans are closely related to the economic and business cycle. Modelling mortgage arrears in the book building societies in United Kingdom, Brookes, Dicks, and Pradhan (1994) find that increases in unanticipated inflation rate heavily affect mortgage defaults. Rinaldi and Sanchis-Arellano (2006), focusing on household NPLs for a panel of euro area countries over the years 1989 to 2004, show that disposable income, unemployment and monetary conditions strongly affect NPLs. Berge and Boye (2007), focusing on the Nordic banking system over the period from 1993 to 2005, show that problem loans are highly sensitive to real interest rates and unemployment. Jappelli, Pagano, and Maggio (2008) investigate how households arrears are influenced by household indebtedness, using data on 11 European countries and the US, and find that insolvencies tends to be associated to greater households' indebtedness, and that institutional arrangements play an important role in determining the size and fragility of household credit markets. Beck, Jakubik, and Piliou (2013) study the macroeconomic determinants of NPLs across 75 countries during the past decade, and find that real GDP growth, share prices, the exchange rate, and the lending interest rate significantly affect asset quality.

A further strand of literature emphasises the effect of bank-specific characteristics on the quality of loans. Factors such as bad management of banks with poor skills in credit scoring, banks' risk attitude and diversification opportunities, or banks' size have been pointed as important in

determining loan quality and NPLs. We refer to Louzis, Vouldis, and Metaxas (2012) for a review of this literature.

In this paper, to study variations in NPLs over time and across territory, we extend the GMM approach for pure dynamic panels, to allow for spatial dependence in disturbances. Hence, it is of interest to briefly introduce the reader to the literature on GMM estimation of panels in the presence of spatial dependence. Kelejian and Prucha (1999) first proposed GMM estimation of regression models with spatial autoregressive (SAR) disturbances, in a single cross sectional setting. They suggested the use of three moment conditions that exploit the properties of disturbances entailed by a standard set of assumptions. In the last few years, a sizeable literature has been developed to extend this procedure. Liu, Lee, and Bollinger (2012) and Liu, Lee, and Bollinger (2010) suggested a set of linear and quadratic conditions in the error term, where the matrices appearing in the linear and quadratic forms have bounded row and column norms (see also Lee (2007)). These moments can be robustified against unknown heteroskedasticity by assuming that the diagonal elements of the inner matrices are zero (see Lin and Lee (2010)). Lee and Liu (2010) have extended this framework to estimate SAR models with higher-order spatial lags. Kelejian and Prucha (2009) have generalized their work to incorporate spatial lags in the dependent variable as well as unknown heteroskedasticity. This setting has been further extended to estimate a spatial panel regression model with group-specific coefficients, both under the random effects and fixed effects assumptions (Kapoor, Kelejian, and Prucha (2007), Mutl and Pfaffermayr (2011), and Moscone and Tosetti (2011)). Druska and Horrace (2004) have introduced the Kelejian and Prucha GMM within the framework of a panel with SAR disturbances, time dummies and time-varying spatial weights, while Fingleton (2008a) and Fingleton (2008b) have extended it to the case of a regression model with spatial moving average disturbances.

While GMM estimation of spatial panel data models with strictly exogenous regressors has been widely investigated, little work has been undertaken so far on the estimation of panel data models that include both spatial and temporal dynamics. Lee and Yu (2010a) considered GMM estimation of a panel with fixed effects, a time lagged dependent variable, and a spatially lagged dependent variable among the regressors. The authors suggest to eliminate individual effects by applying an orthonormal transformation and then use linear and quadratic conditions to estimate the unknown parameters. Korniotis (2010) have proposed a bias-corrected least squares dummy variable estimator of a time-space recursive model, and compare its small sample properties with those of an instrumental variables-type estimator, showing a worse performance for the latter. Maximum likelihood estimation of a dynamic, stationary panel with fixed effects and spatial lags in the dependent variable and error is considered in Yu, de Jong, and Lee (2007), Yu, de Jong, and Lee (2008), and Lee and Yu (2010b).

### 3 The empirical model

As mortgage arrears are likely to be persistent over time, in our empirical study we adopt a dynamic specification for NPLs. In particular, let  $y_{it}$  be the NPL on the  $i$ th MSA at time  $t$ , and suppose that it is generated by the following first-order dynamic panel data model, for  $i = 1, 2, \dots, N, t = 1, 2, \dots, T$ ,

with  $T > 2$ ,

$$y_{it} = \alpha_i + \lambda y_{i,t-1} + \beta' \mathbf{x}_{it} + u_{it}, \quad (1)$$

where  $\alpha_i$  are the group-specific effects,  $\mathbf{x}_{it}$  is a  $k$ -dimensional vector of determinants of NPLs, possibly correlated with  $\alpha_i$ ,  $\lambda$  is a scalar parameter of the lagged dependent variable,  $\beta$  is a  $k$ -dimensional parameter vector. We assume that  $\mathbf{x}_{it}$  is made up of the following sets of variables:

$$\mathbf{x}_{it} = (HP_{it} | GSE_{i,t-1} | MACRO_{it} | BANK_{it}).$$

In the above,  $HP_{it}$  and  $GSE_{i,t-1}$  are our key variables explaining NPLs. More specifically,  $HP_{it}$  is a measure of real house prices in the  $i$ th MSA at time  $t$ ,<sup>2</sup> while  $GSE_{i,t-1}$  is the number of real estate mortgages backed by government-sponsored enterprises Fannie Mae and Freddie Mac at time  $t - 1$ . There exists a large literature suggesting that the sharp increase in subprime lending played an important role in the creation of recent US housing bubble (Mian and Sufi (2009); Bolotnyy (2012)). Although most of this increase was due to non-agency serviced private label mortgage-backed securities, during the height of the housing bubble, almost 40 per cent of newly issued private-label subprime securities were purchased by Fannie Mae and Freddie Mac. It is plausible to think that it is the GSE purchase in the year of origination of the loan rather than the GSE purchase in the year of default that affect credit risk. As a proxy of this temporal lag in the effect of the GSE purchases, we have included GSE purchased loans lagged by one period in the model.

In our regression, we also control for a set variables reflecting the state of the local economy and social conditions,  $MACRO_{it}$ . These are real per-capita personal income ( $INCOME_{it}$ ), unemployment rate ( $UNEMP_{it}$ ), mortgage interest rate ( $IRATE_{it}$ ), and population density ( $POPDENS_{it}$ ). A growing economy is likely to be associated with rising incomes and reduced financial distress. Accordingly, we expect growth in real per-capita personal income and decline in unemployment rate to diminish NPLs. A rise in mortgage interest rate weakens borrowers' debt servicing capacity, particularly if loan rates are variable. Therefore, in general, we would expect NPLs to be positively related with mortgage interest rate. We include population density as a proxy of regional deprivation which may positively affect NPLs.

Finally, we incorporate in our regression a vector of variables related to the financial sector ( $BANK_{it}$ ), computed at aggregate level for each MSA. We have selected these variables with the aim to control for differences across MSAs in the concentration of banks, and riskiness of banks' loan portfolios. In particular,  $BANK_{it}$  includes the concentration of assets within the MSA ( $HHI_{it}$ ), the equity-to-assets ratio ( $EQASS_{it}$ ), bank size ( $SIZE_{it}$ ) and the number of branches within the MSA ( $BRANCHES_{it}$ ). The variable  $HHI_{it}$  is measured as an Herfindahl index of assets, which is equal to the sum of the squared bank shares of assets, calculated each year and for each MSAs in the sample. Some papers associate a higher concentration of assets to an increase in banking system fragility (concentration-fragility hypothesis), given that large banks may be more difficult to monitor than small banks. On the contrary, other studies support the view that banking system concentration enhance stability (concentration-stability hypothesis), as it signals less competition and hence greater market power and profits, thus reducing incentives for bankers to take excessive

---

<sup>2</sup>Some studies (see, for example, Koetter and Poghosyan (2010)), rather than using real house prices, focus on nominal house prices, i.e., without accounting for the variation in underlying macroeconomic fundamentals. We have also tried estimating our model using nominal house prices and obtained very similar results.



risk (see Beck (2007) for a review of the literature). The variable equity-to-asset ratio represents a key measure of the level of banks' capitalisation, indicating a bank's ability to cover unexpected losses. Under the current Basel I and II regulations banks are required to meet a minimum 8 per cent capital buffer. Most banks hold a capital buffer above the legally required value, but this is subject to individual banks. The equity-to-asset ratio reflects the "moral hazard" hypothesis, first discussed by Keeton and Morris (1987). Banks with relatively low capital, measured by a small equity-to-assets ratio, tend to make riskier loans, which in turn results in higher non-performing loans on average. Bank size is a proxy for diversification opportunities and in this paper is measured as the average total assets held by banks within the MSA. Empirical evidence supports a negative relationship between NPLs and bank size, as larger banks have better risk management practices which eventually lead to lower credit risk and NPLs (Salas and Saurina (2002)). Finally, coming to the variable  $BRANCHES_{it}$ , the branching deregulation in the US has significantly expanded the supply of mortgage credit; increase in mortgage credit is expected to have a significant, positive, impact on loan losses (Favara and Imbs (2011)).

We recognise that, in addition to these variables, many other bank-specific factors can be related to NPLs. However, in order to keep the model tractable and parsimonious, we have only included for few key variables, and rely in our econometric approach to deal with potential omitted-variable problems. It is also important to observe that some of the above regressors may be endogenously determined. For instance, markets with increasing defaults may experience depreciation in house prices and land value due to lower expected returns on investment properties. In our empirical work we will deal with the problem of potential endogeneity of house price and the financial variables by following the suggestion by Blundell, Bond, and Windmeijer (2000). Specifically, as explained in Section 5, we exclude from the set of instruments that are adopted for estimation, those that use contemporaneous and one-period lagged values of the endogenous regressors.

Most empirical works investigating the determinants of NPLs do not take into consideration that disturbance are likely to be spatially correlated. Socio-economic factors may also affect NPLs, such as degree of urbanization, deprivation and crime, which are notoriously difficult to quantify and are well known to be spatially correlated. We refer to Triki and Maktouf (2012) for further discussion. Spatial correlation among NPLs in neighbouring MSAs may also occur as the consequence of measurements errors. Given the above discussion, we assume that the error term,  $u_{it}$ , is generated by the following SAR process

$$u_{it} = \delta \sum_{j=1}^N w_{ij} u_{jt} + \varepsilon_{it}, \quad (2)$$

where  $w_{ij}$  are elements of a  $N \times N$  spatial weights matrix,  $\mathbf{W}$ .

In the next section, we introduce a GMM estimator for the unknown parameters  $\lambda$ ,  $\beta$  and  $\delta$  in equations (1)-(2).

## 4 Methods

In the following, we focus on consistent estimation of  $\lambda$ ,  $\beta$  and  $\delta$  via GMM. In order to distinguish the true parameters from other possible values in the parameter space, we denote by  $\lambda_0$ ,  $\beta_0$ ,  $\delta_0$ ,

and  $\sigma_{0i}^2$  the true parameters, which generated an observed sample.

## 4.1 Moment conditions

Consider model (1)-(2) and assume that the following assumptions hold:

**Assumption 1** :  $\varepsilon_{it}$  are independently distributed random variables with zero mean, variance  $0 < E(\varepsilon_{it}^2) = \sigma_{0i}^2 \leq \sigma_{\max}^2 < \infty$ , and such that  $E|\varepsilon_{it}|^{4+\eta} \leq K < \infty$  for some  $\eta > 0$  and for  $i = 1, 2, \dots, N; t = 1, 2, \dots, T$ .

**Assumption 2** The group-specific effects,  $\alpha_i$ , and the errors,  $\varepsilon_{it}$ , satisfy:

$$E(\alpha_i) = 0, E(\varepsilon_{it}) = 0, i = 1, 2, \dots, N; t = 1, 2, \dots, T, \quad (3)$$

$$E(\varepsilon_{is}\varepsilon_{it}) = 0, i = 1, 2, \dots, N; s \neq t = 1, 2, \dots, T, \quad (4)$$

$$E(\alpha_i\varepsilon_{it}) = 0, E(\mathbf{x}_{it}\varepsilon_{it}) = \mathbf{0}, i = 1, 2, \dots, N; t = 1, 2, \dots, T. \quad (5)$$

**Assumption 3** The main diagonal elements of  $\mathbf{W}$  are zero. The row and column norms of the matrices  $\mathbf{W}$  and  $(\mathbf{I}_N - \delta\mathbf{W})^{-1}$  are bounded.

**Assumption 4**  $\delta_0 \in [c_l, c_u]$ , with  $-\infty < c_l, c_u < \infty$ , and  $(\mathbf{I}_N - \delta\mathbf{W})^{-1}$  is non-singular for all  $\delta \in [c_l, c_u]$ .

The existence of moments of order higher than four stated in Assumption 1 is needed for applicability of the central limit theorem for triangular arrays by Kelejian and Prucha (2001). In Assumption 2, conditions (4) require serially uncorrelated errors, while (5) exclude the  $\mathbf{x}_{it}$  process to be endogenously determined. We observe that this assumption allows the group-specific effects to be correlated to the included regressors,  $\mathbf{x}_{it}$ . The following assumptions concerning the initial conditions are also taken

$$E(y_{i0}\varepsilon_{it}) = 0, i = 1, 2, \dots, N, t = 1, 2, \dots, T. \quad (6)$$

We observe that (3)-(5) and (6) are standard in the literature on GMM estimation of dynamic panels, and refer to Arellano and Bond (1991) and Blundell, Bond, and Windmeijer (2000) for further discussion. Under (3)-(6), and focusing on the equation expressed in first difference, Arellano and Bond (1991) suggest the following  $T(T-1)/2$  population moment conditions:

$$E[y_{is}(\Delta y_{it} - \lambda_0 \Delta y_{i,t-1} - \beta'_0 \Delta \mathbf{x}_{it})] = 0, s = 0, 1, \dots, t-2, t = 2, 3, \dots, T. \quad (7)$$

If, in addition to condition (5),  $\mathbf{x}_{it}$  are strictly exogenous, i.e. if  $E(\mathbf{x}_{is}\varepsilon_{it}) = 0$ , for all  $s$  and  $t$ , then the following  $T(T-1)$  additional moments can be used

$$E[\mathbf{x}_{is}(\Delta y_{it} - \lambda_0 \Delta y_{i,t-1} - \beta'_0 \Delta \mathbf{x}_{it})] = 0, \text{ for } s = 1, 2, \dots, T; t = 2, \dots, T, \quad (8)$$

while in the case  $\mathbf{x}_{it}$  are weakly exogenous, namely if  $E(\mathbf{x}_{is}\varepsilon_{it}) = 0$ , for  $s = 1, 2, \dots, t$  and for all  $t$ , then there are only  $T(T-1)/2$  additional moments available:

$$E[\mathbf{x}_{is}(\Delta y_{it} - \lambda_0 \Delta y_{i,t-1} - \beta'_0 \Delta \mathbf{x}_{it})] = 0, \text{ for } s = 1, 2, \dots, t-1; t = 2, \dots, T. \quad (9)$$

Under no spatial error dependence, and in absence of extra information about the dynamic process, a GMM estimator based on conditions (7) and (8) (or (9)) is asymptotically normal and efficient in the class of estimators based on linear moment conditions (Hansen (1982), Chamberlain (1987)). However, as shown by Blundell and Bond (1998), its performance deteriorates as the variance of the group-effects is large relative to the variance of the idiosyncratic error, or when the parameter attached to the lagged dependent variable is close to one. Indeed, in these cases it is possible to show that the instruments are only weakly related with the endogenous differences (see also Binder, Hsiao, and Pesaran (2005) and Kiviet (2007)). To deal with this problem, Arellano and Bover (1995), Blundell and Bond (1998) and Blundell, Bond, and Windmeijer (2000) proposed to focus on the dynamic equation expressed in levels rather than in first-differences, and suggested a set of instruments valid under certain assumptions on the initial conditions of the dynamic process. In particular, suppose that, in addition to (3)-(6), the conditions

$$E(\Delta y_{i1} \alpha_i) = 0, \quad (10)$$

hold. Then the following  $(T-1)^2/2$  moment conditions are available for the equation in levels, (1):

$$E[\Delta y_{is}(y_{it} - \lambda_0 y_{i,t-1} - \beta'_0 \mathbf{x}_{it})] = 0, \text{ for } s = 1, \dots, t-1; t = 2, 3, \dots, T. \quad (11)$$

Further, if regressors,  $\mathbf{x}_{it}$ , satisfy

$$E(\Delta \mathbf{x}_{i1} \alpha_i) = 0, \quad (12)$$

then, under strict exogeneity, the  $T^2$  conditions

$$E[\Delta \mathbf{x}_{is}(y_{it} - \lambda_0 y_{i,t-1} - \beta'_0 \mathbf{x}_{it})] = 0, \text{ } s = 1, 2, \dots, T; t = 1, 2, \dots, T, \quad (13)$$

can also be used, while under weak exogeneity, we have the  $T^2/2$  moments

$$E[\Delta \mathbf{x}_{is}(y_{it} - \lambda_0 y_{i,t-1} - \beta'_0 \mathbf{x}_{it})] = 0, \text{ for } s = 1, 2, \dots, t-1; t = 1, 2, \dots, T. \quad (14)$$

We observe that, if (7)-(8) (or (9)) and (11), (13) (or (14)) are used jointly, then some of the conditions in (11)-(14) are redundant. In this case, in addition to (7)-(8), only the  $(T-1)$  conditions

$$E[\Delta y_{i,t-1}(y_{it} - \lambda_0 y_{i,t-1} - \beta'_0 \mathbf{x}_{it})] = 0, \text{ for } t = 2, 3, \dots, T, \quad (15)$$

and, under either strictly or weakly exogenous regressors,

$$E[\Delta \mathbf{x}_{it}(y_{it} - \lambda_0 y_{i,t-1} - \beta'_0 \mathbf{x}_{it})] = 0, \text{ for } t = 1, 2, \dots, T, \quad (16)$$

can be used. Conditions (7)-(8) and (15)-(16) yield the so-called *system* GMM, first proposed by Blundell and Bond (1998) in the context of a pure autoregressive panel data model. It is convenient to rewrite moments (7)-(16) in the compact form:

$$E[\mathbf{Z}'(\mathbf{q} - \mathbf{G}\gamma_0)] = 0, \quad (17)$$

where  $\gamma_0 = (\lambda_0, \beta'_0)'$ ,  $\mathbf{q} = (\mathbf{q}'_1, \mathbf{q}'_2, \dots, \mathbf{q}'_N)'$ ,  $\mathbf{Z} = (\mathbf{Z}'_1, \mathbf{Z}'_2, \dots, \mathbf{Z}'_N)'$ ,  $\mathbf{G} = (\mathbf{G}'_1, \mathbf{G}'_2, \dots, \mathbf{G}'_N)'$ . The vectors  $\mathbf{q}_i$  and the matrices  $\mathbf{Z}_i$ ,  $\mathbf{G}_i$ ,  $i = 1, 2, \dots, N$ , depending on the three possible sets of conditions (and under the further assumption of strictly exogenous regressors), are given by:

(i) Under the *difference moment conditions* (7) and (8):

$$\begin{aligned} \mathbf{Z}_{i.} &= \mathbf{Z}_{i.}^d \\ &\quad (T-1) \times (1+2k)T(T-1)/2 \\ &= \begin{pmatrix} y_{i0}, \mathbf{x}'_{i1}, \dots, \mathbf{x}'_{iT} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & y_{i0}, y_{i1}, \mathbf{x}'_{i1}, \dots, \mathbf{x}'_{iT} & \dots & \mathbf{0} \\ & & \ddots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & y_{i0}, \dots, y_{i,T-2}, \mathbf{x}'_{i1}, \dots, \mathbf{x}'_{iT} \end{pmatrix}, \end{aligned} \quad (18)$$

$$\mathbf{q}_{i.} = \mathbf{q}_{i.}^d = \begin{pmatrix} \Delta y_{i2} \\ \dots \\ \Delta y_{iT} \end{pmatrix}, \mathbf{G}_{i.} = \mathbf{G}_{i.}^d = \begin{pmatrix} \Delta y_{i1} & \Delta \mathbf{x}'_{i2} \\ \vdots & \vdots \\ \Delta y_{i,T-1} & \Delta \mathbf{x}'_{iT} \end{pmatrix}. \quad (19)$$

(ii) Under the *level moment conditions* (11) and (13):

$$\begin{aligned} \mathbf{Z}_{i.} &= \mathbf{Z}_{i.}^\ell \\ &\quad (T-1) \times [2kT + (T-1)](T-1)/2 \\ &= \begin{pmatrix} \Delta y_{i1}, \Delta \mathbf{x}'_{i1}, \dots, \Delta \mathbf{x}'_{iT} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Delta y_{i1}, \Delta y_{i2}, \Delta \mathbf{x}'_{i1}, \dots, \Delta \mathbf{x}'_{iT} & \dots & \mathbf{0} \\ & & \ddots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \Delta y_{i1}, \dots, \Delta y_{i,T-1}, \Delta \mathbf{x}'_{i1}, \dots, \Delta \mathbf{x}'_{iT} \end{pmatrix} \end{aligned} \quad (20)$$

$$\mathbf{q}_{i.} = \mathbf{q}_{i.}^\ell = \begin{pmatrix} y_{i2} \\ \dots \\ y_{iT} \end{pmatrix}, \mathbf{G}_{i.} = \mathbf{G}_{i.}^\ell = \begin{pmatrix} y_{i1} & \mathbf{x}'_{i2} \\ \dots & \mathbf{x}'_{iT} \end{pmatrix}. \quad (21)$$

(iii) Under both *difference and level moment conditions*:

$$\mathbf{Z}_{i.} = \mathbf{Z}_{i.}^{sys} = \begin{pmatrix} \mathbf{Z}_{d,i} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Delta y_{i1}, \Delta \mathbf{x}'_{i2} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & \Delta y_{i2}, \Delta \mathbf{x}'_{i3} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \Delta y_{i,T-1}, \Delta \mathbf{x}'_{iT} \end{pmatrix} \quad (22)$$

$$\mathbf{q}_{i.} = \mathbf{q}_{i.}^{sys} = \begin{pmatrix} \mathbf{q}_{i.}^\ell \\ \mathbf{q}_{i.}^d \end{pmatrix}, \mathbf{G}_{i.} = \mathbf{G}_{i.}^{sys} = \begin{pmatrix} \mathbf{G}_{i.}^d \\ \mathbf{G}_{i.}^\ell \end{pmatrix}. \quad (23)$$

In addition to moments (17), following Kelejian and Prucha (1999), Liu, Lee, and Bollinger (2010), and others, we propose the following  $r$  quadratic conditions in the error term for estimation of the spatial parameter:

$$\frac{1}{2N(T-1)} E [\Delta \boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \boldsymbol{\varepsilon}] = 0, \quad \ell = 1, 2, \dots, r, \quad (24)$$

where

$$\Delta \boldsymbol{\varepsilon} = [(\mathbf{I}_N - \delta_0 \mathbf{W}) \otimes \mathbf{I}_{T-1}] \Delta \mathbf{u} = [(\mathbf{I}_N - \delta_0 \mathbf{W}) \otimes \mathbf{I}_{T-1}] (\Delta \mathbf{y} - \Delta \mathbf{G}^d \boldsymbol{\gamma}_0),$$

$\Delta \mathbf{y} = (\Delta \mathbf{y}'_1, \Delta \mathbf{y}'_2, \dots, \Delta \mathbf{y}'_N)'$  with  $\Delta \mathbf{y}_i = (\Delta y_{i2}, \Delta y_{i3}, \dots, \Delta y_{iT})'$ ,  $\mathbf{G}^d = (\mathbf{G}_1^d, \mathbf{G}_2^d, \dots, \mathbf{G}_N^d)'$  with  $\mathbf{G}_i^d$  provided in (19), and  $\mathbf{A}_\ell$  are  $N \times N$  non-stochastic matrices with generic elements  $a_{ij,\ell}$ , and having bounded row and column norms. Following the work by Liu, Lee, and Bollinger (2010), to render estimation robust against unknown heteroskedasticity we assume that the matrices inside the quadratic form have zero diagonal elements, i.e.,  $a_{ii,\ell} = 0$ , for  $i = 1, 2, \dots, N$  and  $\ell = 1, 2, \dots, r$ . We refer to Moscone and Tosetti (2011) for a discussion on various possible choices for the inner matrices in (24). In the Appendix, we describe the GMM estimator of the SAR coefficient based on moments (24). We next introduce a two-step GMM estimator of the slope parameters in (1), that accounts for spatial dependence.

## 4.2 Two-step estimation of the slope parameters

Estimation of  $\gamma_0$  can proceed adopting the following two-step procedure. First, compute the conventional one-step GMM estimator, which ignores spatial error dependence, and is given by:

$$\hat{\gamma} = (\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})^{-1} \mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{Zq}, \quad (25)$$

where

$$\mathbf{S}_{ZW} = \mathbf{Z}' \mathbf{G}, \mathbf{S}_{Zq} = \mathbf{Z}' \mathbf{q}, \quad (26)$$

$$\mathbf{D}_u = \mathbf{Z}' (\mathbf{I}_N \otimes \mathbf{P}) \mathbf{Z}, \quad (27)$$

with  $\mathbf{P}$  being

$$\mathbf{P} = \underset{(T-1) \times (T-1)}{\mathbf{P}^d} = \begin{pmatrix} 2 & -1 & \dots & 0 \\ -1 & 2 & & \vdots \\ \vdots & & \ddots & -1 \\ 0 & \dots & -1 & 2 \end{pmatrix}, \quad (28)$$

$$\mathbf{P} = \underset{(T-1) \times (T-1)}{\mathbf{P}^l} = \mathbf{I}_{T-1}, \quad (29)$$

$$\mathbf{P} = \underset{2(T-1) \times 2(T-1)}{\mathbf{P}^{sys}} = \begin{pmatrix} \mathbf{P}^d & 0 \\ 0 & \mathbf{P}^l \end{pmatrix}, \quad (30)$$

for the three sets of moments, respectively. Hence, calculate the residuals:

$$\Delta \hat{u}_{it} = \Delta y_{it} - \hat{\gamma}' \Delta \mathbf{g}_{it}^d, \quad (31)$$

which can be used in the minimization problem (A.10), set out in the Appendix, to obtain  $\hat{\delta}$ . Finally, apply to the variables a Cochrane-Orcutt-type transformation to get, in the case of difference or level moment conditions,<sup>3</sup>  $\tilde{\mathbf{Z}} = [(\mathbf{I}_N - \hat{\delta} \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{Z}$ ,  $\tilde{\mathbf{q}} = [(\mathbf{I}_N - \hat{\delta} \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{q}$ , and  $\tilde{\mathbf{G}} = [(\mathbf{I}_N - \hat{\delta} \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{G}$  and compute a two-step estimator that accounts for spatial correlation:

$$\hat{\gamma}^{II} = \left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1} \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{Zq}, \quad (32)$$

---

<sup>3</sup>In the case both difference and level conditions are taken the appropriate transformations are  $\tilde{\mathbf{Z}}^{sys} = [(\mathbf{I}_N - \hat{\delta} \mathbf{S}) \otimes \mathbf{I}_{2(T-1)}] \mathbf{Z}^{sys}$ ,  $\tilde{\mathbf{q}}^{sys} = [(\mathbf{I}_N - \hat{\delta} \mathbf{S}) \otimes \mathbf{I}_{2(T-1)}] \mathbf{q}^{sys}$ , and  $\tilde{\mathbf{W}}^{sys} = [(\mathbf{I}_N - \hat{\delta} \mathbf{S}) \otimes \mathbf{I}_{2(T-1)}] \mathbf{W}^{sys}$ .

where now

$$\tilde{\mathbf{S}}_{ZW} = \tilde{\mathbf{Z}}' \tilde{\mathbf{G}}, \tilde{\mathbf{S}}_{Zq} = \tilde{\mathbf{Z}}' \tilde{\mathbf{q}}, \quad (33)$$

$$\tilde{\mathbf{D}}_u = \tilde{\mathbf{Z}}' (\mathbf{I}_N \otimes \mathbf{P}) \tilde{\mathbf{Z}}, \quad (34)$$

and  $\mathbf{P}$  being one of the three matrices (28)-(30). In the following, we provide a theorem for consistency and asymptotic normality of GMM estimators (25) and (32) based on difference conditions (7) and (8). Suppose that the following assumption holds:

**Assumption 5** *The matrix  $\frac{1}{N(T-1)} (\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})$  has finite elements and is non-singular; the matrix  $\frac{1}{N(T-1)} \mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{Z}^d$  has finite elements and is full rank.*

**Theorem 1** *Suppose the Assumptions 1-5 are satisfied. Then the one-step estimator  $\hat{\gamma}$ , given by (25) and based on conditions (7) and (8), is consistent and asymptotically normal for  $N \rightarrow \infty$  and fixed  $T$  with*

$$\hat{\gamma} - \gamma_0 \stackrel{a}{\sim} N \left( \mathbf{0}, (\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})^{-1} \mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{Z}^d [\mathbf{R}_0 \Sigma \mathbf{R}_0' \otimes \mathbf{P}^d] \mathbf{Z}^d \mathbf{D}_u^{-1} \mathbf{S}_{ZW} (\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})^{-1} \right). \quad (35)$$

with  $\mathbf{R}_0 = (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1}$ . Further, let  $\hat{\delta}$  be the solution of (A.10) based on residuals (31). Then the two-step estimator  $\hat{\gamma}^{II}$ , given by (32) and based on conditions (7) and (8), is consistent for  $\gamma_0$ , and asymptotically normal for  $N \rightarrow \infty$  and fixed  $T$ , with

$$\hat{\gamma}^{II} - \gamma_0 \stackrel{a}{\sim} N \left( \mathbf{0}, (\mathbf{S}^{*'}_{ZW} \mathbf{D}_u^{*-1} \mathbf{S}^*_{ZW})^{-1} \mathbf{S}^{*'}_{ZW} \mathbf{D}_u^{*-1} \mathbf{Z}^{*'} (\Sigma \otimes \mathbf{P}^d) \mathbf{Z}^* \mathbf{D}_u^{*-1} \mathbf{S}^*_{ZW} (\mathbf{S}^{*'}_{ZW} \mathbf{D}_u^{*-1} \mathbf{S}^*_{ZW})^{-1} \right), \quad (36)$$

where  $\mathbf{S}^*_{ZW} = [(\mathbf{I}_N - \delta_0 \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{S}_{ZW}$ ,  $\mathbf{D}_u^* = [(\mathbf{I}_N - \delta_0 \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{D}_u$ , and  $\mathbf{Z}^* = [(\mathbf{I}_N - \delta_0 \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{Z}$ .

**Proof.** Consider

$$\hat{\gamma} - \gamma_0 = (\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})^{-1} \mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{Z}' \Delta \mathbf{u}, \quad (37)$$

$$\hat{\gamma}^{II} - \gamma_0 = \left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1} \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{Z}}' \Delta \tilde{\mathbf{u}}. \quad (38)$$

Note that (37) can be written as

$$\hat{\gamma} - \gamma_0 = (\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})^{-1} \mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{Z}' [(\mathbf{I}_N - \delta_0 \mathbf{S})^{-1} \otimes \mathbf{D}] \boldsymbol{\varepsilon}.$$

Result (35) follows by applying the central limit theorem provided in Kelejian and Prucha (1998) (see page 112), since under Assumption 5 it is easily seen that the matrix  $(\mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{S}_{ZW})^{-1} \mathbf{S}'_{ZW} \mathbf{D}_u^{-1} \mathbf{Z}' [(\mathbf{I}_N - \delta_0 \mathbf{S})^{-1} \otimes \mathbf{D}] \boldsymbol{\varepsilon}$  has finite elements and bounded row and column norms. Now consider (38), and note that, given the  $\sqrt{N}$ -consistency of  $\delta$ ,

$$\left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1} - (\mathbf{S}^{*'}_{ZW} \mathbf{D}_u^{*-1} \mathbf{S}^*_{ZW})^{-1} = O \left( \frac{1}{\sqrt{N}} \right), \quad (39)$$

$$\tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{Z}}' \Delta \tilde{\mathbf{u}} - \mathbf{S}^{*'}_{ZW} \mathbf{D}_u^{*-1} \mathbf{Z}^{*'} \Delta \boldsymbol{\varepsilon} = O \left( \frac{1}{\sqrt{N}} \right) \quad (40)$$

It follows that (36) holds. ■

Similar results can be obtained using level moments under analogous conditions. An estimator of the asymptotic variance of  $\hat{\gamma}^{II}$ , robust to heteroskedasticity in errors is

$$Var(\hat{\gamma}^{II}) = \left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1} \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{Z}}' (\mathbf{I}_N \otimes \mathbf{P}_e) \tilde{\mathbf{Z}} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{ZW} \left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_u^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1}, \quad (41)$$

with

$$\mathbf{P}_e = \frac{1}{N} \sum_{i=1}^N \left( \tilde{\mathbf{q}}_i - \tilde{\mathbf{G}}_i \hat{\gamma}^{II} \right) \left( \tilde{\mathbf{q}}_i - \tilde{\mathbf{G}}_i \hat{\gamma}^{II} \right)'. \quad (42)$$

The two-step GMM estimator  $\hat{\gamma}^{II}$  can be computed using optimal weights (see Arellano and Bover (1995)) as follows:

$$\hat{\gamma}_e^{II} = \left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_e^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1} \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_e^{-1} \tilde{\mathbf{S}}_{Zq}, \quad (43)$$

where  $\tilde{\mathbf{S}}_{ZW}$  and  $\tilde{\mathbf{S}}_{Zq}$  are given by (33), and  $\tilde{\mathbf{D}}_e = \tilde{\mathbf{Z}}' (\mathbf{I}_N \otimes \mathbf{P}_e) \tilde{\mathbf{Z}}$ , with  $\mathbf{P}_e$  given by (42). An estimator of the variance of  $\hat{\gamma}_e^{II}$  robust to heteroskedasticity in errors is

$$Var(\hat{\gamma}_e^{II}) = \left( \tilde{\mathbf{S}}'_{ZW} \tilde{\mathbf{D}}_e^{-1} \tilde{\mathbf{S}}_{ZW} \right)^{-1}. \quad (44)$$

In the appendix we provide a small Monte Carlo exercise to investigate the properties of the above estimators.

## 5 Data and descriptive statistics

Data on NPLs and the other bank-related variables are collected from the database Statistics on Depository Institutions maintained by the Federal Deposit Insurance Corporation (FDIC). This database contains detailed information on all FDIC-insured commercial banks and saving institutions, which represents the great majority in the US. In this paper, bank-related variables for each MSA are computed as sums of quantities associated to branches located within the MSA. By doing this, we are implicitly assuming that all mortgage borrowers live in the MSA of the branch office where they take the loan. However, it is important to observe that a branch located in a MSA could also lend to customers residing in other MSAs for example through brokers. According to Xu and Zhang (2012), evidence from Home Mortgage Disclosure Act data suggests that less than half of mortgage loans made during 2005-2008 were originated by banks chartered in the borrower's state with a branch office in the same MSA where the borrowers reside, the remaining were either originated by a bank chartered outside of the borrower's state, or by a bank without a branch office in the borrower's MSA. To mitigate this problem of misallocation of loans to MSAs, we have decided to drop from our sample all inter-States banks, operating across several US States. By doing this, the amount of loans for which the borrower and lender are located in the same MSA rises to over 72 per cent of total loans (see Xu and Zhang (2012) for details). After this cleaning procedure, our sample includes between 6,700 and 9,600 institutions located in 366 MSAs, over the period 2000 to 2011.<sup>4</sup>

---

<sup>4</sup>We note that the number of MSAs for which we have financial information reduces to 357 towards the end of the sample period.

Bank NPLs ( $NPL_{it}$ ) are defined as the sum of loans that are in arrears, i.e. borrowed money upon which the debtor has not made her scheduled payments, for at least 90 days. Technically, we measure arrears as the sum of 90 days or more past due loans for 1-to-4 family residential properties plus their loans that are past due in nonaccrual status. We then divide the sum of all arrears held by any bank belonging to the MSA, by total gross loans held by banks belonging to the MSA. Given that the ratio of non-performing loans to total loans takes values in the range  $[0, 1]$ , in the regression we use as dependent variable its logit transformation  $y_{it} = \ln [NPL_{it} / (1 - NPL_{it})]$ .

The variable  $EQASS_{it}$  is calculated as the ratio of equities divided by total assets owned by any banks within the MSA, while  $HHI_{it}$  is the Herfindahl index of total assets in the  $i$ th MSA at time  $t$ .

Data on NPLs and other bank-related variables are then matched with data at MSA level for the same period on house prices, GSE loan purchases, and local socio-economic conditions. House prices and data on GSE purchases are collected from the Federal Housing Finance Agency (FHFA). Specifically,  $HP_{it}$  refers to the average house price of single-family properties within the  $i$ th MSA whose mortgages have been purchased or securitized by Fannie Mae or Freddie Mac.<sup>5</sup> Similarly,  $GSE_{it}$  is the number of single-family mortgages purchased by Fannie Mae or Freddie Mac within the  $i$ th MSA. In our regression analysis, we have divided this variable by total population in the MSA. Data on per capita-personal income, unemployment and mortgage rate are gathered from Bureau of Economic Analysis, Bureau of Labor Statistics and FHFA, respectively. Finally, population density is taken from the Census. We observe that data on mortgage interest rate are at State level.

As discussed in Section 1, to investigate the effect of housing price and GSE loan purchases on NPLs, the sample is split into two sub-samples, with the first sub-sample covering the years 2000-2005, preceding the bust of the real estate bubble, and the second sub-period covering the house-price bubble bursting, over the years 2006 to 2011. In addition, to take into account for possible endogeneity of house price and the financial variables, in estimation we follow Blundell, Bond, and Windmeijer (2000) and for these variables we excluded from the set of instruments in (9) and (14) those for  $s = t - 1$ .

Table 6 reports the descriptive statistics for the variables under study in the two sub-periods. The statistics indicate a deterioration in the financial and economic conditions in the second period, characterised, on average, by larger NPLs and unemployment rates, and a real income only slightly growing. Mortgage interest rates decrease in the second period. Such decrease can be explained by the monetary policy response to the financial crisis, which has reduced considerably interest rates with the aim to inject liquidity in the system. On the other side, GSE loan purchases decrease dramatically in the second period. Such reduction is probably due to the US government policy response to the financial meltdown after 2005. GSEs were placed under conservatorship in 2008 in order to prop up their mortgage activities and the number of government backed mortgages decreased. While descriptive statistics indicate a rise in average real house prices between the two sub-periods, the dynamic of this variable can be better appreciated by looking at Figure 1, which shows the temporal pattern of average house prices, total GSE loan purchases and average NPLs

---

<sup>5</sup>We observe that, since  $HP_{it}$  only includes houses with conforming, conventional loans, it does not include information on house prices purchased via jumbo mortgages, i.e., in an amount above conventional conforming loan limits.



in the US. This figure shows that the rapid rise in house prices experienced in the first half of the sample period is followed by a decrease in the second half. At the same time, NPLs show an initial stable pattern below 1 per cent, while blowing up to above 3 per cent towards the end of the sample period. As for GSE loan purchases, these grow consistently until 2003, year in which they fall precipitously, remaining low for the following years.

Figure 2 and 3 show the quantile distribution of house prices and NPLs, respectively, in the two sub-periods and across metro areas. The two figures indicate a marked heterogeneity in these variables across the US. Also, it is interesting to observe that, when passing from boom to bust sub-period, both variables show a sharp increase in their geographical concentration. In particular, during the housing bust, NPLs and house prices tend to distribute in clusters, with the East and West coasts showing the highest values, and the Midwest region characterised by clusters of values belonging to the lowest quantiles.

## 6 Estimation results

Tables 7 and 8 present the estimated parameters for the two sub-periods. In both tables, the upper panel presents the estimated parameters for the model using conventional GMM estimation with no spatial errors, whereas the estimation results using the proposed GMM approach are reported in the bottom panel. Following our Monte Carlo experiments, we report three alternative GMM estimators, the GMM estimator using difference conditions (7)-(8) (column I, GMM-DIF), using level conditions (11), (13) (column II, GMM-LEV) and the GMM estimator based on a combination of these two sets (column III, GMM-SYS). Given the high degree of heterogeneity across MSAs, we report standard errors that are robust to heteroskedasticity in errors.

Looking at the results, from the top and the bottom panels in Table 7, as expected NPLs are characterised by significant temporal dynamics, with  $\hat{\lambda}$  ranging between 0.23 and 0.42, depending on the set of moments considered for estimation. Results also show that house prices negatively affect NPLs for all estimators. This result seem to be in line with the collateral hypothesis outlined in Koetter and Poghosyan (2010), according to which real estate price appreciation prevents (sub-prime) mortgage borrowers from defaulting (see also Daglish (2009)). The estimated coefficient for  $GSE_{i,t-1}$  is not significant for all estimators. As for the covariates on the economic and social conditions, Table 7 shows that, after controlling for spatial dependence, growth in personal income reduces NPLs. Unemployment rate, by negatively impacting on the cash flow streams of households and increasing the debt burden, has a positive effect on NPLs. These results confirm that a rising income and decreasing unemployment rate is associated with reduced financial distress, and lower probability rates of default (Lawrence (1995)). In general, our results corroborate evidence in Rinaldi and Sanchis-Arellano (2006) on the impact of current income and unemployment rate on the probability of default. The coefficient attached to interest rate is positive for most estimators, but, after controlling for spatial dependence, statistically insignificant. We observe that our data on the interest rate is at State level, and thus our estimates for this coefficient may not capture well the relationship between NPLs and mortgage interest rate. Finally, as expected, population density has a positive impact on NPLs, although the effect is significant only when controlling for spatial dependence. Moving to the bank-specific variables, the estimated coefficient of the equity ratio,

although showing the correct sign, is in general not significant. Focusing on HHI, it is interesting to observe that the estimated coefficient attached to this variable is positive and significant, supporting the concentration-fragility view in the period of housing boom. Coming to the variable  $SIZE_{it}$ , from Table 7 it appears that small banks suffer from higher NPL levels during boom period. This is probably due to the fact that smaller banks have less market power, less economies of scale, and less diversification opportunities among their customers and products (Salas and Saurina (2002)). The variable  $BRANCHES_{it}$  has a negative sign although not significant, no matter the model under consideration.

Results on the estimation on the second sub-period are reported in Table 8, and confirm a temporal dynamic in NPLs similar to the first sub-period, and a negative effect of real estate prices on NPLs. Such negative coefficient relative to a period characterised by house prices depreciating rapidly, may be explained by the fact that speculative buyers and owner-occupiers that are unwilling or unable to repay their mortgages, find it difficult to roll over their loans or sell their properties, and hence start to default. Contrary to the first sub-period, the estimated coefficient for  $GSE_{i,t-1}$  is now positive and significant once controlled for spatial dependence, indicating that a rise in the number of real estate mortgages backed by government-sponsored enterprises grows NPLs, thus deteriorating the quality of banks' loan portfolio. Such result has interesting policy implications. The activity of GSEs, directed to enhance home-ownership opportunities for the population, should not have such effect on financial stability, which is known to have a negative impact on real economy thus offsetting the primary goals of GSEs.

As for the covariates on the economic and social conditions, results are very similar to those in the first sub-period, for disposable income unemployment rate and population density. The coefficient attached to interest rate is in most cases positive, but, after controlling for spatial dependence, statistically insignificant. This is probably due to the aggressive monetary policy response to the financial crisis, which adopted several measures to inject liquidity in the system, including exceptionally low interest rates. Finally, as expected, population density has a positive impact on NPLs, although the effect is significant only when controlling for spatial dependence. Coming to the bank-specific variables, the estimated coefficient of the equity ratio turns to be significant, indicating that during housing bust, MSAs with high proportion of thinly capitalized banks are exposed to greater financial instability. Indeed, banks with low capitalization may invest more on high-risk assets, resulting in an impaired loan quality. In this respect, our result is in line with Davis and Zhu (2005) who find that capital ratio reduce NPLs, while positively influencing bank margin, the loan provision and return on assets. The authors find that banks with high NPLs typically have low capital ratios. Focusing on HHI, it is interesting to observe that the estimated coefficient attached to this variable, while positive and significant in the first sub-period, is now negative and significant. Hence, our results seem to support the concentration-stability hypothesis in the years during the bubble bust. Empirical evidence in favour of the concentration-stability view can be found in several studies (see Beck (2007) for a review), although these are mostly carried at country level. Finally, the variable  $SIZE_{it}$  has the expected negative sign, while  $BRANCHES_{it}$  has no significant effect on NPLs.

Table 7 and 8 also show strong and significant spatial effects in NPLs both in the period of housing boom and bust, indicating that some unobservable affecting NPLs are geographically concentrated. It is interesting to observe that the spatial coefficient rises between the first and the

second sub-periods from around 0.4-0.5 in the first sub-period to over 0.7. Such growth in the spatial effects confirms the increased geographical concentration in NPLs observed in Figure 3, and may be explained by the deterioration in the economic and social conditions experienced in the second sub-period, which is also likely to be accompanied by a higher geographical concentration of unobservables, such as poverty and criminality.

It is important to observe that the three estimators DIF, LEVEL and SYS yield sometimes different results in terms of signs and size of estimated coefficients. Given the high degree of heterogeneity that we expect in this study, and the temporal persistence in the variable of our model, in general we believe that the GMM-SYS is more appropriate in this application. Our specification tests indicate the validity of the instruments used for estimation in all cases, and the absence of second order error serial correlation.

## 7 Concluding remarks

This paper has investigated spatio-temporal variations in ex-post credit risk as a function of real estate prices, loan purchases made by GSEs, and a set of local characteristics, using data on US metropolitan statistical area over the period 2000 to 2011. We have given a number of contributions to existing empirical literature on nonperforming loans. Differently from previous work, we have used data at metro level, to properly capture the effect of local social, economic and financial conditions on financial stability. Our results point to a significant negative impact of real estate prices on ex-post risk, both during and before the bust of the bubble. In a period of house prices rising fast, this result corroborates the hypothesis that wealth can play the role of a buffer in case of unexpected shocks or that housing wealth can be used as collateral to ease access to credit. During the bursting of the bubble, when house prices start falling below the nominal value of loans, the negative impact of real estate prices on NPLs is explained by an increase in default rates due to speculative buyers and owner-occupiers that are unwilling or unable to repay their mortgages and have difficulties in selling their properties. Our results also indicate a significant positive impact of GSE loan purchases on ex-post risk, only in the period during the bust of the bubble. Hence, in a period of crisis, the activity of GSEs seems to contribute to enhancing financial fragility, rather than working as a economic cushion to mortgage markets. We also found a marked spatial concentration of unobservables, that rises consistently during the bubble bust. Such result may be explained by the worsening of social and economic conditions, which in turn may have accentuated the spatial clustering of poverty and deprivation across the territory in this period.

Another major contribution of this paper has been to extend existing econometric methods adopted to study the determinants of NPLs, to account for possible spatial dependence present in the data. To this end, we have developed an *ad-hoc* GMM procedure to estimate a first-order dynamic panel data regression model with group-specific effects and spatial autoregressive errors. This procedure may be adopted to investigate a large number of economic problems characterised by both spatial and temporal patterns. For instance, they may be useful for estimating cross-country growth regressions as in Caselli, Esquivel, and Lefort (1996), studying spatio-temporal patterns in consumption behaviour (see, for example, Browning and Collado (2007)), or exploring the dynamics in the production of firms as in Blundell and Bond (2000).

## References

- Almeida, H., M. Campello, and C. Liu (2006). The financial accelerator: Evidence from international housing markets. *Review of Finance* 10, 1–32.
- Alvarez, J. and M. Arellano (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica* 71, 1121–1159.
- Arellano, M. and S. R. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58, 277–297.
- Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68, 29–52.
- Beck, R., P. Jakubik, and A. Piloju (2013). Non-performing loans: What matters in addition to the economic cycle? ECB Working Paper No. 1515.
- Beck, T. (2007). Bank concentration and fragility. impact and mechanics. In M. Carey and R. M. Stulz (Eds.), *The Risks of Financial Institutions*. National Bureau of Economic Research.
- Berge, T. O. and K. G. Boye (2007). An analysis of banks’ problem loans. *Economic Bulletin* 2, 65–76.
- Bernanke, B. and M. Gertler (1995). Inside of a black box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives* 9, 27–48.
- Bernanke, B., M. Gertler, and S. Gilchrist (1996). The financial accelerator and the flight to quality. *The Review of Economics and Statistics* 78, 1–15.
- Binder, M., C. Hsiao, and M. H. Pesaran (2005). Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory* 21, 795–837.
- Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115–143.
- Blundell, R. and S. Bond (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews* 19, 321–340.
- Blundell, R., S. Bond, and F. Windmeijer (2000). Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator. IFS Working Papers W00/12, Institute for Fiscal Studies.
- Bolotnyy, V. (2012). The government-sponsored enterprises and the mortgage crisis: The role of the Affordable Housing Goals. Finance and Economics Discussion Series n. 2012-25, Federal Reserve Board.
- Brookes, M., M. Dicks, and M. Pradhan (1994). An empirical model of mortgage arrears and repossessions. *Economic Modelling* 11, 134–144.
- Browning, M. and M. D. Collado (2007). Habits and heterogeneity in demands: a panel data analysis. *Journal of Applied Econometrics* 22, 625–640.
- Caselli, F., G. Esquivel, and F. Lefort (1996). Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth* 1, 363–389.

- Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics* 34, 305–334.
- Daglish, T. (2009). What motivates a subprime borrower to default? *Journal of Banking & Finance* 33, 681–693.
- Davis, E. P. and H. Zhu (2005). Commercial property prices and bank performance. BIS Working Papers n. 175.
- Dell’Ariccia, G. and R. Marquez (2006). Lending booms and lending standards. *Journal of Banking and Finance* 61, 2511–2546.
- Druska, V. and W. C. Horrace (2004). Generalized moments estimation for spatial panels: indonesian rice farming. *American Journal of Agricultural Economics* 86, 185–198.
- Favara, G. and J. Imbs (2011). Credit supply and the price of housing. 2011 Meeting Papers 1342, Society for Economic Dynamics.
- Fingleton, B. (2008a). A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. *Empirical Economics* 34, 35–57.
- Fingleton, B. (2008b). A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. *Spatial Economic Analysis* 3, 27–44.
- Gimeno, R. and C. Martinez-Carrascal (2010). The relationship between house prices and house purchase loans: The spanish case. *Journal of Banking & Finance* 34, 1849–1855.
- Greene, W. (2002). *Econometric Analysis* (Fifth ed.). Prentice Hall.
- Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* 50, 1029–1054.
- Herring, R. J. and S. Wachter (1999). Real estate booms and banking busts: An international perspective. Center for Financial Institutions Working Papers 99-27, Wharton School Center for Financial Institutions, University of Pennsylvania.
- Hilbers, P., Q. Lei, and L. Zacho (2001). Real estate market developments and financial sector soundness. IMF Working Paper n. 129.
- Hsiao, C., M. H. Pesaran, and A. K. Tahmiscioglu (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of Econometrics* 109, 107–150.
- Jappelli, T., M. Pagano, and M. D. Maggio (2008). Households’ indebtedness and financial fragility. CSEF Working Paper 208 (Naples: Center for Studies of Economics and Finance).
- Kapoor, M., H. H. Kelejian, and I. Prucha (2007). Panel data models with spatially correlated error components. *Journal of Econometrics* 140, 97–130.
- Keeton, W. and C. Morris (1987). Why do banks’ loan losses differ? *Economic review*.
- Kelejian, H. H. and I. Prucha (1998). A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17, 99–121.

- Kelejian, H. H. and I. Prucha (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* 40, 509–533.
- Kelejian, H. H. and I. R. Prucha (2001). On the asymptotic distribution of the moran i test with applications. *Journal of Econometrics* 104, 219–257.
- Kelejian, H. H. and I. R. Prucha (2009). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics* 157, 53–67.
- Kiviet, J. F. (2007). Judging contending estimators by simulation: tournaments in dynamic panel data models. In G. D. A. Phillips and E. Tzavalis (Eds.), *The Refinement of Econometric Estimation and Test Procedures*. Cambridge University Press.
- Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy* 105, 211–248.
- Koetter, M. and T. Poghosyan (2010). Real estate markets and bank stability. *Journal of Banking & Finance* 34, 1129–1138.
- Korniotis, G. M. (2010). Estimating panel models with internal and external habit formation. *Journal of Business & Economic Statistics* 28, 145–158.
- Lawrence, E. C. (1995). Consumer default and the life cycle model. *Journal of Money Credit and Banking* 27, 939–954.
- Lee, L. F. (2007). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics* 137, 489–514.
- Lee, L. F. and X. Liu (2010). Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. *Econometric Theory* 26, 187–230.
- Lee, L. F. and J. Yu (2010a). Efficient GMM estimation of spatial dynamic panel data models with fixed effects. Ohio State University.
- Lee, L.-F. and J. Yu (2010b). Some recent developments in spatial panel data model. *Regional Science and Urban Economics* 40, 255–271.
- Lin, X. and L. F. Lee (2010). GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics* 157, 34–52. Mimeo, Ohio State University.
- Liu, X., L. F. Lee, and C. R. Bollinger (2010). An efficient GMM estimator of spatial autoregressive models. *Journal of Econometrics* 159, 303–319.
- Liu, X., L. F. Lee, and C. R. Bollinger (2012). Improved efficient quasi maximum likelihood estimator of spatial autoregressive models. *Journal of Econometrics Forthcoming*. Mimeo.
- Louzis, D. P., A. T. Vouldis, and V. L. Metaxas (2012). Macroeconomic and bank-specific determinants of non-performing loans in greece: A comparative study of mortgage, business and consumer loan portfolios. *Journal of Banking & Finance* 36, 1012–1027. Bank of Greece Working Paper n. 118.
- Mayer, C., K. Pence, and S. M. Sherlund (2009). The rise in mortgage defaults. *Journal of Economic Perspectives* 23, 27–50.

- Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis. *Quarterly Journal of Economics* 124, 1449–1496.
- Moscone, F. and E. Tosetti (2011). GMM estimation of spatial panels with fixed effects and unknown heteroskedasticity. *Regional Science and Urban Economics*, Forthcoming.
- Mutl, J. (2006). Dynamic panel data models with spatially correlated disturbances. PhD Thesis, University of Maryland.
- Mutl, J. and M. Pfaffermayr (2011). The Hausman test in a Cliff and Ord panel model. *The Econometrics Journal* 14, 48–76.
- Reinhart, C. M. and K. S. Rogoff (2008). Banking crises: An equal opportunity menace. NBER Working Paper No. 14587.
- Rinaldi, L. and A. Sanchis-Arellano (2006). Household debt sustainability: what explains household non-performing loans? an empirical analysis. ECB Working Paper n. 570.
- Salas, V. and J. Saurina (2002). Credit risk in two institutional regimes: Spanish commercial and savings banks. *Journal of Financial Services Research* 22, 203–224.
- Triki, M. B. and S. Maktouf (2012). Cross-country empirical studies of banking crisis: A spatial durbin model. *European Journal of Business and Management* 4, 181–192.
- Xu, Y. and J. Zhang (2012). Nonlocal mortgage lending and the second market involvement. *Journal of Real Estate Literature* 20, 307–322.
- Yu, J., R. de Jong, and L. F. Lee (2007). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both N and T are large: A nonstationary case. Mimeo, Ohio State University.
- Yu, J., R. de Jong, and L. F. Lee (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both N and T are large. *Journal of Econometrics* 146, 118–137.

# Appendices

In these appendices we first introduce the GMM estimator of the SAR coefficient, and prove its consistency and asymptotic normality. We then provide results for a small Monte Carlo exercise. For our statistical derivations, it is useful to introduce the following lemma.

**Lemma 1** *Let  $\Delta\boldsymbol{\varepsilon} = (\Delta\boldsymbol{\varepsilon}'_1, \Delta\boldsymbol{\varepsilon}'_2, \dots, \Delta\boldsymbol{\varepsilon}'_N)'$ ,  $\Delta\boldsymbol{\varepsilon}_{i.} = (\Delta\varepsilon_{i2}, \dots, \Delta\varepsilon_{iT})'$ , be a  $N(T-1)$ -dimensional vector with  $\varepsilon_{it}$  satisfying Assumption 1, and let  $\mathbf{A}_\ell$ , for  $\ell = 1, 2, \dots, r$ , be non-stochastic matrices with zero diagonal elements. We have, for  $\ell = 1, 2, \dots, r$ ,*

$$\frac{1}{2N(T-1)} E [\Delta\boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta\boldsymbol{\varepsilon}] = 0, \quad (.45)$$

$$Var \left[ \frac{1}{2N(T-1)} \Delta\boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{I}_T) \Delta\boldsymbol{\varepsilon} \right] = \frac{6(T-2) + 4}{4N^2(T-1)^2} Tr \left[ (\boldsymbol{\Sigma} \mathbf{A}_\ell)^2 + \boldsymbol{\Sigma} \mathbf{A}_\ell \mathbf{A}'_\ell \boldsymbol{\Sigma} \right] \quad (.46)$$

$$Cov \left[ \frac{1}{2N(T-1)} \Delta\boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{I}_T) \Delta\boldsymbol{\varepsilon}, \Delta\boldsymbol{\varepsilon}' (\mathbf{A}_h \otimes \mathbf{I}_T) \Delta\boldsymbol{\varepsilon} \right] = \frac{6(T-2) + 4}{4N^2(T-1)^2} Tr (\boldsymbol{\Sigma} \mathbf{A}_\ell \boldsymbol{\Sigma} \mathbf{A}_h + \boldsymbol{\Sigma} \mathbf{A}_\ell \mathbf{A}'_h \boldsymbol{\Sigma}), \quad (.47)$$

$$E \left[ \frac{1}{2N(T-1)} \Delta\boldsymbol{\varepsilon} \Delta\boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{I}_T) \Delta\boldsymbol{\varepsilon} \right] = \mathbf{0}. \quad (.48)$$

**Proof.** The above can be proved using results on moments of quadratic forms (see Moscone and Tosetti (2011))<sup>6</sup> and noting that we can rewrite  $\Delta\boldsymbol{\varepsilon}$  as follows:

$$\Delta\boldsymbol{\varepsilon}_{N(T-1) \times 1} = (\mathbf{I}_N \otimes \mathbf{D}) \boldsymbol{\varepsilon}_{NT \times 1},$$

where  $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_1, \boldsymbol{\varepsilon}'_2, \dots, \boldsymbol{\varepsilon}'_N)'$ , with  $\boldsymbol{\varepsilon}_{i.} = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$ , and  $\mathbf{D}$  is a rectangular matrix given by:

$$\mathbf{D}_{(T-1) \times T} = \begin{pmatrix} -1 & 1 & 0 & & 0 & 0 \\ 0 & -1 & 1 & & & \\ 0 & & -1 & \dots & & \\ \dots & & & \dots & 1 & 0 \\ 0 & & & \dots & -1 & 1 \end{pmatrix}, \quad (.49)$$

so that we can rewrite the quadratic form as follows:

$$\frac{1}{2N(T-1)} \Delta\boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta\boldsymbol{\varepsilon} = \frac{1}{2N(T-1)} \boldsymbol{\varepsilon}' (\mathbf{A}_\ell \otimes \mathbf{D}' \mathbf{D}) \boldsymbol{\varepsilon},$$

where the elements of the  $NT$ -dimensional vector,  $\boldsymbol{\varepsilon}$ , are independently distributed, and

$$\mathbf{D}' \mathbf{D}_{T \times T} = \begin{pmatrix} 1 & -1 & 0 & & 0 & 0 \\ -1 & 2 & -1 & & & \\ 0 & -1 & 2 & \dots & & \\ \dots & & & \dots & & \\ 0 & & & \dots & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}.$$

■

<sup>6</sup>A detailed proof is available upon request.



## A Estimation of the SAR coefficient

Suppose that we can find an estimator,  $\hat{\gamma}$ , consistent for the unknown parameters,  $\gamma_0$ . This can be obtained by computing the GMM estimator based solely on the set of moments (17). In a non-spatial setting (i.e., under  $\delta_0 = 0$ ), this estimator is unbiased and consistent for  $N$  tending to infinity, and  $T$  fixed (see Arellano and Bond (1991) and Alvarez and Arellano (2003)). As it will be shown below (see, in particular, Theorem 1), this is true also when errors are spatially correlated. Let  $\mathcal{M}(\delta) = [\mathcal{M}_1(\delta), \dots, \mathcal{M}_r(\delta)]'$  be a vector containing the  $r$  conditions (24), and consider their empirical counterpart given by  $\mathbf{M}_{NT}(\delta) = [M_{NT,1}(\delta), \dots, M_{NT,r}(\delta)]'$ , where:

$$M_{NT,\ell}(\delta) = \frac{1}{2N(T-1)} \Delta \hat{\varepsilon}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \hat{\varepsilon}(\delta),$$

with

$$\Delta \hat{\varepsilon}(\delta) = [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] \Delta \hat{\mathbf{u}} = [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] (\Delta \mathbf{y} - \mathbf{G}^d \hat{\gamma}),$$

The following proposition holds.

**Proposition 1** *Under Assumptions 1-4 we have, for all  $\delta \in [c_l, c_u]$ ,*

$$\frac{1}{2N(T-1)} (\Delta \hat{\varepsilon}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \hat{\varepsilon}(\delta) - \Delta \varepsilon(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \varepsilon(\delta)) = O_p\left(\frac{1}{N}\right), \quad (\text{A.1})$$

$$\frac{1}{2N(T-1)} [\Delta \varepsilon(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \varepsilon(\delta) - E(\Delta \varepsilon(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \varepsilon(\delta))] = O_p\left(\frac{1}{\sqrt{N}}\right). \quad (\text{A.2})$$

**Proof.** We now sketch the proof, and refer to Liu, Lee, and Bollinger (2010), Lee (2007), Kapoor, Kelejian, and Prucha (2007), Kelejian and Prucha (2009), and Moscone and Tosetti (2011) for further details on the convergence of quadratic forms. First, consider

$$\begin{aligned} \Delta \hat{\varepsilon}(\delta) &= [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] \Delta \hat{\mathbf{u}} = [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] (\Delta \mathbf{y} - \mathbf{G}^d \hat{\gamma}) \\ &= [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] [\mathbf{G}^d (\hat{\gamma} - \gamma_0) + (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \Delta \varepsilon] \\ &= [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{G}^d (\hat{\gamma} - \gamma_0) + [(\mathbf{I}_N - \delta \mathbf{W}) (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{I}_{T-1}] \Delta \varepsilon, \end{aligned}$$

Noting that  $(\mathbf{I}_N - \delta \mathbf{W}) (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1}$  can be also written as

$$\begin{aligned} (\mathbf{I}_N - \delta \mathbf{W}) (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} &= (\mathbf{I}_N - \delta_0 \mathbf{W} + \delta_0 \mathbf{W} - \delta \mathbf{W}) (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \\ &= \mathbf{I}_N + (\delta_0 - \delta) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} = \mathbf{P}(\delta), \end{aligned} \quad (\text{A.3})$$

we can rewrite  $\Delta \varepsilon(\delta)$  and  $\Delta \hat{\varepsilon}(\delta)$  as follows

$$\begin{aligned} \Delta \varepsilon(\delta) &= [(\mathbf{I}_N - \delta \mathbf{W}) (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{I}_{T-1}] \Delta \varepsilon = \left[ (\mathbf{I}_N + (\delta_0 - \delta) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{D}) \right] \varepsilon \\ &= [\mathbf{P}(\delta) \otimes \mathbf{I}_{T-1}] \Delta \varepsilon \end{aligned} \quad (\text{A.4})$$

$$\Delta \hat{\varepsilon}(\delta) = [(\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{G}^d (\hat{\gamma} - \gamma_0) + [\mathbf{P}(\delta) \otimes \mathbf{I}_{T-1}] \Delta \varepsilon. \quad (\text{A.5})$$

To prove (A.1), note that

$$\begin{aligned}
& \frac{1}{2N(T-1)} \Delta \hat{\boldsymbol{\varepsilon}}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \hat{\boldsymbol{\varepsilon}}(\delta) \\
= & \frac{1}{2N(T-1)} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} [(\mathbf{I}_N - \delta \mathbf{W})' \mathbf{A}_\ell (\mathbf{I}_N - \delta \mathbf{W}) \otimes \mathbf{I}_{T-1}] \mathbf{G}^d (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) \\
& + \frac{2}{2N(T-1)} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} [(\mathbf{I}_N - \delta \mathbf{W})' \mathbf{A}_\ell \mathbf{P}(\delta) \otimes \mathbf{I}_{T-1}] \Delta \boldsymbol{\varepsilon} \\
& + \Delta \boldsymbol{\varepsilon}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \boldsymbol{\varepsilon}(\delta) \\
= & \frac{1}{2N(T-1)} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} (\mathbf{B}_\ell \otimes \mathbf{I}_{T-1}) \mathbf{G}^d (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) \\
& + \frac{2}{NT} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} (\mathbf{C}_\ell \otimes \mathbf{I}_{T-1}) \Delta \boldsymbol{\varepsilon} + \Delta \boldsymbol{\varepsilon}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \boldsymbol{\varepsilon}(\delta).
\end{aligned}$$

where  $\mathbf{B}_\ell = (\mathbf{I}_N - \delta \mathbf{W})' \mathbf{A}_\ell (\mathbf{I}_N - \delta \mathbf{W})$ ,  $\mathbf{C}_\ell = (\mathbf{I}_N - \delta \mathbf{W})' \mathbf{A}_\ell \mathbf{P}(\delta)$ . Under Assumptions 3-4  $\mathbf{B}_\ell$  and  $\mathbf{C}_\ell$ , have row and column norms that are uniformly bounded. Given the  $\sqrt{N}$ -consistency of  $\hat{\boldsymbol{\gamma}}$ , it is easily seen that

$$\begin{aligned}
\frac{1}{2N(T-1)} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} (\mathbf{B}_\ell \otimes \mathbf{I}_{T-1}) \mathbf{G}^d (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) & \leq K \frac{1}{2N(T-1)} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} \mathbf{G}^d (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) \\
& = O_p\left(\frac{1}{N}\right), \\
\frac{2}{2N(T-1)} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} (\mathbf{C}_\ell \otimes \mathbf{D}) \boldsymbol{\varepsilon} & = \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{j=1}^N c_{\ell,ij} (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)' \mathbf{G}^{d'} \mathbf{D} \boldsymbol{\varepsilon}_j. \\
& = O_p\left(\frac{1}{N}\right).
\end{aligned}$$

It follows that

$$\frac{1}{2N(T-1)} \Delta \hat{\boldsymbol{\varepsilon}}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \hat{\boldsymbol{\varepsilon}}(\delta) = \frac{1}{2N(T-1)} \Delta \boldsymbol{\varepsilon}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \boldsymbol{\varepsilon}(\delta) + O_p\left(\frac{1}{N}\right),$$

which proves (A.1). As for (A.2), using (A.4) we have

$$\frac{1}{2N(T-1)} \Delta \boldsymbol{\varepsilon}(\delta)' (\mathbf{A}_\ell \otimes \mathbf{I}_{T-1}) \Delta \boldsymbol{\varepsilon}(\delta) = \frac{1}{NT} \boldsymbol{\varepsilon}' [\mathbf{P}(\delta)' \mathbf{A}_\ell \mathbf{P}(\delta) \otimes \mathbf{D}] \boldsymbol{\varepsilon} \quad (\text{A.6})$$

where  $\mathbf{P}(\delta)$  is given by (A.3), and has uniformly bounded row and column norms. Using Lemma 1, the mean of (A.6) satisfies

$$\begin{aligned}
E \left\{ \frac{1}{2N(T-1)} \boldsymbol{\varepsilon}' [\mathbf{P}(\delta)' \mathbf{A}_\ell \mathbf{P}(\delta) \otimes \mathbf{D}] \boldsymbol{\varepsilon} \right\} & = \frac{1}{2N(T-1)} \text{Tr}(\boldsymbol{\Sigma} \mathbf{P}(\delta)' \mathbf{A}_\ell \mathbf{P}(\delta)) \text{Tr}(\mathbf{D}' \mathbf{D}) \\
& = \frac{1}{N} \text{Tr}(\boldsymbol{\Sigma} \mathbf{P}(\delta)' \mathbf{A}_\ell \mathbf{P}(\delta)) = O(1).
\end{aligned}$$

Let  $\mathbf{B}_\ell = \mathbf{P}(\delta)' \mathbf{A}_\ell \mathbf{P}(\delta)$  with elements  $b_{ij,\ell}$ , and note that the diagonal elements of  $(\mathbf{B}_\ell \otimes \mathbf{D})$  are  $2b_{ii,\ell}$ , for  $i = 1, 2, \dots, N, t = 1, 2, \dots, T$ . Then the variance of (A.6) satisfies

$$\begin{aligned} \text{Var} \left[ \frac{1}{2N(T-1)} \boldsymbol{\varepsilon}' (\mathbf{B}_\ell \otimes \mathbf{D}'\mathbf{D}) \boldsymbol{\varepsilon} \right] &= \frac{1}{4N^2(T-1)^2} \sum_{i=1}^N 4b_{ii,\ell}^2 [E(\varepsilon_{it}^4) - 3\sigma_{0i}^4] \\ &\quad + \frac{1}{2N^2(T-1)^2} \text{Tr}[(\mathbf{D}'\mathbf{D})^2] \text{Tr}[(\boldsymbol{\Sigma}\mathbf{B}_\ell)^2 + \boldsymbol{\Sigma}\mathbf{B}_\ell\mathbf{B}_\ell'\boldsymbol{\Sigma}] \\ &= O\left(\frac{1}{N}\right), \end{aligned} \quad (\text{A.7})$$

which proves (A.2). ■

Let

$$\mathbf{V}(\boldsymbol{\theta}_0) = \lim_{N \rightarrow \infty} E[\mathbf{NM}_{NT}(\delta_0) \mathbf{M}_{NT}(\delta_0)'], \quad (\text{A.8})$$

where  $\boldsymbol{\theta}_0 = (\delta_0, \sigma_{01}^2, \dots, \sigma_{0N}^2)'$ . Given Lemma 1 (see in particular result (.46)), the above matrix has generic  $(\ell, h)^{th}$  element,  $v_{\ell h}$ , given by

$$v_{\ell h} = \lim_{N \rightarrow \infty} \frac{6(T-2) + 4}{4N(T-1)^2} \text{Tr}[\boldsymbol{\Sigma}\mathbf{A}_\ell \boldsymbol{\Sigma}\mathbf{A}_h + \boldsymbol{\Sigma}\mathbf{A}_\ell \mathbf{A}_h' \boldsymbol{\Sigma}]. \quad (\text{A.9})$$

We observe that the factor  $\frac{6(T-2)+4}{4N(T-1)^2}$  appears in the above expression because  $\text{Tr}[(\mathbf{D}'\mathbf{D})^2] = 6(T-2) + 4$  enters in the expectation (where  $\mathbf{D}$  is given in (.49)). Under the assumption of bounded row and column norms of the matrices  $\mathbf{A}_\ell$  and  $\mathbf{A}_h$ , it is easily seen that  $v_{\ell h} = O(1)$ . We take up the following assumptions needed for identifiability of parameters (see also Moscone and Tosetti (2011)):

**Assumption 6** *The matrix  $\mathbf{V}(\boldsymbol{\theta}_0)$  is non-singular, i.e. we assume  $\lambda_r(\mathbf{V}(\boldsymbol{\theta}_0)) \geq K > 0$ .*

**Assumption 7** *There exists at least one moment condition, the  $\ell^{th}$ , for which we have either  $\text{Tr}[\boldsymbol{\Sigma}\mathbf{A}_\ell \mathbf{W}(\mathbf{I}_N - \delta_0 \mathbf{W})^{-1}] \neq 0$ , or  $\text{Tr}[\boldsymbol{\Sigma}((\mathbf{I}_N - \delta_0 \mathbf{W}')^{-1} \mathbf{W}' \mathbf{A}_\ell \mathbf{W}(\mathbf{I}_N - \delta_0 \mathbf{W})^{-1})] \neq 0$ .*

The GMM estimator  $\hat{\delta}$  of  $\delta_0$  is the solution to the following optimization problem

$$\hat{\delta} = \arg \min_{\delta \in [c_l, c_u]} \{\mathbf{M}_{NT}(\delta)' \mathbf{Q}_{NT} \mathbf{M}_{NT}(\delta)\}, \quad (\text{A.10})$$

where  $[c_l, c_u]$  is the parameter space (see Assumption 4), and  $\mathbf{Q}_{NT}$  is a  $r \times r$ , positive definite, weighting matrix, such that

$$\mathbf{Q}_{NT} \xrightarrow{p} \mathbf{Q}.$$

The following theorem states that  $\hat{\delta}$  is consistent for  $\delta_0$  and establishes its asymptotic distribution.

**Theorem 2** *Under Assumptions 1-7,  $\hat{\delta}$  in (A.10) is consistent for  $\delta$  as  $N \rightarrow \infty$  and for  $T$  fixed. Further, we have*

$$\sqrt{N}(\hat{\delta} - \delta_0) \overset{a}{\sim} N\left(0, (\mathbf{d}'\mathbf{Q}\mathbf{d})^{-1} \mathbf{d}'\mathbf{Q}\mathbf{V}\mathbf{Q}\mathbf{d} (\mathbf{d}'\mathbf{Q}\mathbf{d})^{-1}\right), \quad (\text{A.11})$$

where  $\mathbf{d} = \mathbf{d}(\boldsymbol{\theta}_0) = -\lim_{N \rightarrow \infty} E\left[\frac{\partial}{\partial \delta} \mathbf{M}_{NT}(\delta) \Big|_{\delta=\delta_0}\right]$ .

**Proof.** Consistency and asymptotic normality of  $\hat{\delta}$  can be proved using results from Proposition and following the same lines of reasoning as in Moscone and Tosetti (2011). See also Kelejian and Prucha (1999), Liu, Lee, and Bollinger (2010), Lee (2007), and Kelejian and Prucha (2009) for further details on consistency of GMM estimators of spatial models. ■

The efficient GMM estimator can be obtained by imposing, in (A.10), the optimal weights given by  $\mathbf{Q} = \mathbf{Q}^* = \mathbf{V}^{-1}$  (see Greene (2002) on this). Notice that the  $\ell$ th element of  $\mathbf{d}$  is (see Appendix A.1)

$$d_\ell = \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr} \left( \boldsymbol{\Sigma} (\mathbf{A}_\ell + \mathbf{A}'_\ell) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \right). \quad (\text{A.12})$$

Since  $\mathbf{Q}^*$  and  $\mathbf{d}$  depend on  $\boldsymbol{\theta}_0$ , they can be proxied by  $\mathbf{Q}^* = \mathbf{Q}^*(\hat{\boldsymbol{\theta}})$ , and  $\mathbf{d} = \mathbf{d}(\hat{\boldsymbol{\theta}})$ , where  $\hat{\boldsymbol{\theta}} = \left( \hat{\delta}, \frac{1}{2(T-1)} \sum_{t=1}^T (\Delta \hat{\varepsilon}_{1t})^2, \dots, \frac{1}{2(T-1)} \sum_{t=1}^T (\Delta \hat{\varepsilon}_{Nt})^2 \right)'$ .

## A.1 The elements of $\mathbf{d}$

We now show that  $\left[ \frac{\partial}{\partial \delta} \mathbf{M}_{NT}(\delta) \right]_{\delta=\delta_0} \xrightarrow{p} \lim_{N \rightarrow \infty} E \left[ \frac{\partial}{\partial \delta} \mathbf{M}_{NT}(\delta) \right]_{\delta=\delta_0}$ , and derive the elements of the vector  $\mathbf{d}$ , introduced in Theorem 2. First note that

$$\begin{aligned} \frac{\partial}{\partial \delta} \Delta \boldsymbol{\varepsilon}(\delta) &= \left[ \frac{\partial}{\partial \delta} \left( \mathbf{I}_N + (\delta_0 - \delta) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{I}_{T-1} \right) \right] \Delta \boldsymbol{\varepsilon} \\ &= - \left[ \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{I}_{T-1} \right] \Delta \boldsymbol{\varepsilon}. \end{aligned}$$

Hence, following similar lines of reasoning as in Moscone and Tosetti (2011), we obtain for the  $\ell$ th empirical moment

$$\begin{aligned} \frac{\partial}{\partial \delta} M_{NT,\ell}(\delta) &= - \frac{1}{2N(T-1)} \Delta \boldsymbol{\varepsilon}(\delta)' \left[ \mathbf{A}_\ell \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{I}_{T-1} \right] \Delta \boldsymbol{\varepsilon} \\ &\quad - \frac{1}{NT} \Delta \boldsymbol{\varepsilon}' \left[ (\mathbf{I}_N - \delta_0 \mathbf{W}')^{-1} \mathbf{W}' \mathbf{A}_\ell \otimes \mathbf{I}_{T-1} \right] \Delta \boldsymbol{\varepsilon}(\delta) \\ &= - \frac{1}{2N(T-1)} \Delta \boldsymbol{\varepsilon}' \left[ \left( \mathbf{I}_N + (\delta_0 - \delta) (\mathbf{I}_N - \delta_0 \mathbf{W}')^{-1} \mathbf{W}' \right) \mathbf{A}_\ell \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \otimes \mathbf{I}_{T-1} \right] \Delta \boldsymbol{\varepsilon} \\ &\quad - \frac{1}{2N(T-1)} \Delta \boldsymbol{\varepsilon}' \left[ (\mathbf{I}_N - \delta_0 \mathbf{W}')^{-1} \mathbf{W}' \mathbf{A}_\ell \left( \mathbf{I}_N + (\delta_0 - \delta) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \right) \otimes \mathbf{I}_{T-1} \right] \Delta \boldsymbol{\varepsilon}. \end{aligned}$$

Thus, at  $\delta_0$ ,

$$\frac{\partial}{\partial \delta} M_{NT,\ell}(\delta_0) = - \frac{1}{2N(T-1)} \Delta \boldsymbol{\varepsilon}' \left\{ \left[ (\mathbf{A}_\ell + \mathbf{A}'_\ell) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \right] \otimes \mathbf{I}_{T-1} \right\} \Delta \boldsymbol{\varepsilon},$$

The mean of  $-\frac{\partial}{\partial \delta} M_{NT,\ell}(\delta_0)$  is

$$\begin{aligned} d_\ell &= -E \left[ \frac{\partial}{\partial \delta} M_{NT,\ell}(\delta_0) \right] = \frac{1}{2N(T-1)} E \left\{ \boldsymbol{\varepsilon}' \left[ \left( (\mathbf{A}_\ell + \mathbf{A}'_\ell) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \right) \otimes \mathbf{D}' \mathbf{D} \right] \boldsymbol{\varepsilon} \right\} \\ &= \frac{1}{N} \text{Tr} \left[ \boldsymbol{\Sigma} (\mathbf{A}_\ell + \mathbf{A}'_\ell) \mathbf{W} (\mathbf{I}_N - \delta_0 \mathbf{W})^{-1} \right]. \end{aligned}$$

Further, following similar lines of reasoning as in (A.7), it is possible to show that the variance of  $\text{Var} \left[ -\frac{\partial}{\partial \delta} M_{NT,\ell}(\delta_0) \right]_{\delta=\delta_0} = O\left(\frac{1}{N}\right)$ . It follows that  $-\frac{\partial}{\partial \delta} \mathbf{M}_{NT}(\delta) \Big|_{\delta=\delta_0} \xrightarrow{p} \mathbf{d}$ , where  $\mathbf{d} = (d_1, d_2, \dots, d_r)'$ .

## B Monte Carlo evidence

We consider the following data generating process

$$y_{it} = \alpha_i(1 - \lambda) + \lambda y_{i,t-1} + \beta x_{it} + u_{it}, \quad t = -m + 1, -m + 2, \dots, 0, 1, \dots, T, \quad (\text{B.1})$$

$$y_{i,-m} = \alpha_i + \beta x_{i,-m} + u_{i,-m}, \quad (\text{B.2})$$

with

$$u_{it} = \delta \sum_{j=1}^N w_{ij} u_{jt} + \varepsilon_{it}, \quad t = -m, -m + 1, -m + 2, \dots, 0, 1, \dots, T, \quad (\text{B.3})$$

and  $\varepsilon_{it} \sim IIDN(0, \sigma_i^2)$ ,  $\sigma_i^2 \sim IIDU(0.05, 0.95)$ ,  $t = -m, -m + 1, \dots, 0, 1, \dots, T$ . We assume the spatial weights matrix  $\mathbf{W}$  is a row standardised regular lattice of 1st order, with elements  $w_{ij} = 1$  if units  $i$  and  $j$  are contiguous and  $w_{ij} = 0$  otherwise. The spatial weight matrix is defined in a circular fashion, whereby the first cross section unit is placed adjacent to the last unit. We discard the first  $m$  observations, using the observations  $t = 0$  through  $T$  for estimation. We assume that the regressor,  $x_{it}$ , is generated by

$$x_{it} = \alpha_i + \zeta_{it}, \quad t = -m, -m + 1, \dots, 0, 1, \dots, T. \quad (\text{B.4})$$

where

$$\zeta_{it} = \phi \zeta_{i,t-1} + \epsilon_{it}, \quad t = -49 - m, -48 - m, \dots, 0, 1, \dots, T \quad (\text{B.5})$$

$$\epsilon_{it} \sim IIDN(0, \sigma_{i,\epsilon}^2), \quad \epsilon_{i,-m-50} = 0, \quad (\text{B.6})$$

$\sigma_{\epsilon,i}^2$  is fixed such that  $R_{\Delta y_i}^2 = 0.4$  under no spatial error dependence (i.e.,  $\delta = 0$ ) (see, in particular, formula (8.5) in Hsiao, Pesaran, and Tahmiscioglu (2002)). We discard the first 50 observations of  $\zeta_{it}$  and use the remaining  $(T + 1 + m)$  observations for generating  $x_{it}$  and  $y_{it}$ . The individual-specific coefficients are fixed across experiments and set to:

$$\alpha_i = \frac{q_i - 1}{\sqrt{2}}, \quad q_i \sim \chi_1^2, \quad (\text{B.7})$$

In the simulations, we set  $\beta = 1$ , and try  $\lambda = 0.3, 0.7$ ,  $\delta = 0.0, 0.3, 0.7$ . We consider  $N = 300, 500$ , keeping  $T = 5$ , fixed, and running 1,000 replications for each experiment.

We provide results for the conventional GMM estimators with optimal weights for  $\lambda$  and  $\beta$  ignoring spatial dependence (see, for example, equation (3.2) in Arellano and Bover (1995)), and for the two-step estimator (43)-(44), using either difference conditions (7)-(8) (GMM-DIF), or level conditions (11), (13) (GMM-LEV), or a combination of these two sets (GMM-SYS). Since it is known that the conventional GMM estimators with optimal weights has size distortions in small samples, we also report results for the conventional one-step GMM-DIF estimators for  $\lambda$  and  $\beta$  ignoring spatial dependence (see equation (25)), and for the estimator given in equations (32), (41). Finally, we provide the small sample properties of the estimator for  $\delta$  obtained from (A.10) and using optimal weights, with  $\mathbf{Q}^*(\hat{\theta})$  and  $\mathbf{d}(\hat{\theta})$  as proxies for  $\mathbf{Q}^*$  and  $\mathbf{d}$ . In the computation of  $\hat{\delta}$ , we adopt for moments (24) the inner matrices suggested by Kelejian and Prucha (2009), and set  $r = 2$  with  $\mathbf{A}_1 = \mathbf{W}$ ,  $\mathbf{A}_2 = \mathbf{W}'\mathbf{W} - \text{diag}(\mathbf{W}'\mathbf{W})$ . This choice is made merely for computational convenience, since  $\mathbf{A}_1$  and  $\mathbf{A}_2$  do not depend on unknown parameters and minimization of (A.10) in the first step does not require a preliminary estimation of the unknown parameters.

We assess the performance of estimators by computing their bias, RMSE, size and power. In computing size and power, we adopt a significance level of 5 per cent; the power of the estimator of a parameter,  $\pi_0$ , is calculated under the alternative hypothesis to  $H_1 : \pi = \pi_0 - 0.1$ .

## B.1 Results

Table 1 shows results for the conventional, one-step GMM-DIF estimator of  $\lambda$  and  $\beta$ , for the GMM-DIF estimator corrected for spatial correlation using formulas (32), (41), and for the corresponding estimated SAR parameter. The bias and RMSE of conventional GMM-DIF are small, and decrease as  $N$  gets large, for all values of  $\delta$ , corroborating the theoretical results provided in the first part of Theorem 1. When  $\delta = 0$ , the conventional GMM-DIF for  $\lambda$  and  $\beta$  is correctly sized for all choices of  $N$ , while it is subject to size distortions when  $\delta > 0$ . The over-rejection tendency is due to the use of inappropriate standard errors, and appears to be substantial in the case where the true value of spatial parameter is relatively large ( $\delta = 0.7$ ). In contrast, the GMM-DIF estimator corrected for spatial dependence is correctly sized, reflecting the fact that the estimated variance is a consistent estimator of the true variance.

Tables 2, 3 and 4 provide results for the conventional GMM-DIF, GMM-LEV and GMM-SYS estimators using optimal weights, for the GMM-DIF, GMM-LEV and GMM-SYS estimators corrected for spatial correlation using formulas (43)-(44) (i.e.,  $\hat{\gamma}_e^{II}$ ), and for the corresponding estimated SAR parameters. The first panel of these tables shows that, when  $\delta = 0$ , the conventional GMM estimators with optimal weights for  $\lambda$  and  $\beta$  are correctly sized for large  $N$ . However, they show some size distortions when  $N = 300$ . This result is in line with existing findings in the literature, indicating that the estimated asymptotic standard errors of the conventional two-step GMM estimator are downward biased in small samples. The second and third panels in Tables 2-4 show that, when  $\delta > 0$ , the conventional GMM estimators, ignoring spatial dependence, are severely oversized even when  $N$  is large. In contrast, the empirical sizes of the GMM estimators corrected for spatial dependence are very close to the nominal size, for all values of the spatial parameters, for large  $N$ . Tables 1-4 also show that the GMM estimators for  $\delta$  are always correctly sized, for any sets of moments taken to compute the slope parameters, and for all choices of  $N$ .

Table 5 provides results for a set of experiments where data have been generated as in our previous experiments, with the only difference that we now adopt as spatial weights matrix that we have used for our empirical application, based on the inverse of the geographical distance between MSAs ( $N = 366$  and  $T = 6$ ). To save space, we only report the output for  $\lambda, \delta = 0.3, 0.7$ . Results seem to be robust to the choice of the spatial weights matrix, whether this is irregular, or based on contiguity or geographical distance. To conclude, our results indicate that, for the combination of  $N$  and  $T$  in our empirical study, and using a spatial weights matrix based on geographical distance, the proposed GMM estimators performs quite well.

Table 1: Monte Carlo results for the conventional GMM-DIF estimator and the two-step GMM-DIF estimator corrected for spatial correlation

$N$	Par.	$\lambda_0 = 0.3$				$\lambda_0 = 0.7$			
		Bias	RMSE	Size	Power	Bias	RMSE	Size	Power
		$\delta_0 = 0.0$				$\delta_0 = 0.0$			
300	$\hat{\lambda}$	-0.026	0.063	0.053	0.583	-0.040	0.081	0.053	0.530
500		-0.013	0.054	0.057	0.650	-0.022	0.063	0.053	0.600
300	$\hat{\lambda}^{II}$	-0.025	0.063	0.050	0.580	-0.039	0.081	0.057	0.533
500		-0.012	0.054	0.057	0.643	-0.022	0.063	0.053	0.590
300	$\hat{\beta}$	0.034	0.184	0.050	0.573	-0.025	0.186	0.050	0.457
500		0.028	0.158	0.050	0.593	0.004	0.139	0.050	0.450
300	$\hat{\beta}^{II}$	0.033	0.184	0.050	0.573	-0.024	0.186	0.053	0.463
500		0.028	0.158	0.053	0.593	0.004	0.139	0.057	0.450
300	$\hat{\delta}$	-0.002	0.047	0.033	0.593	-0.001	0.047	0.043	0.557
500		0.002	0.038	0.056	0.747	0.003	0.038	0.050	0.750
		$\delta_0 = 0.3$				$\delta_0 = 0.3$			
300	$\hat{\lambda}$	-0.027	0.065	0.097	0.597	-0.043	0.086	0.127	0.537
500		-0.014	0.055	0.093	0.697	-0.024	0.066	0.113	0.587
300	$\hat{\lambda}^{II}$	-0.024	0.062	0.053	0.590	-0.038	0.080	0.050	0.523
500		-0.012	0.053	0.050	0.643	-0.022	0.062	0.050	0.597
300	$\hat{\beta}$	0.037	0.196	0.080	0.580	-0.025	0.198	0.083	0.647
500		0.031	0.165	0.077	0.593	0.002	0.147	0.077	0.640
300	$\hat{\beta}^{II}$	0.029	0.180	0.053	0.560	-0.026	0.185	0.050	0.663
500		0.027	0.155	0.053	0.580	0.004	0.137	0.053	0.650
300	$\hat{\delta}$	-0.007	0.041	0.047	0.750	-0.006	0.040	0.043	0.733
500		-0.002	0.032	0.057	0.917	0.000	0.032	0.060	0.903
		$\delta_0 = 0.7$				$\delta_0 = 0.7$			
300	$\hat{\lambda}$	-0.034	0.086	0.223	0.693	-0.065	0.122	0.290	0.650
500		-0.019	0.064	0.223	0.727	-0.037	0.086	0.230	0.630
300	$\hat{\lambda}^{II}$	-0.024	0.062	0.053	0.603	-0.037	0.079	0.057	0.510
500		-0.012	0.052	0.050	0.643	-0.022	0.060	0.050	0.613
300	$\hat{\beta}$	0.055	0.300	0.130	0.517	-0.040	0.320	0.120	0.763
500		0.049	0.236	0.110	0.513	-0.008	0.233	0.100	0.727
300	$\hat{\beta}^{II}$	0.025	0.169	0.050	0.580	-0.026	0.176	0.053	0.767
500		0.025	0.146	0.055	0.597	0.002	0.129	0.053	0.760
300	$\hat{\delta}$	-0.007	0.025	0.053	1.000	-0.006	0.024	0.057	1.000
500		-0.003	0.019	0.080	1.000	-0.001	0.019	0.053	1.000

We compute  $\hat{\gamma} = (\hat{\lambda}, \beta)'$  using equation (25), and  $\hat{\gamma}^{II} = (\hat{\lambda}^{II}, \beta^{II})'$  using formula (32), and (41) for its variance. We compute  $\hat{\delta}$  using residuals  $\Delta \hat{u}_{it} = \Delta y_{it} - \hat{\gamma}' \Delta \mathbf{w}_{it}^d$ .

Table 2: Monte Carlo results for the conventional GMM-DIF estimator using optimal weights and the two-step GMM-DIF estimator corrected for spatial correlation

$N$	Par.	$\lambda_0 = 0.3$				$\lambda_0 = 0.7$			
		Bias	RMSE	Size	Power	Bias	RMSE	Size	Power
		$\delta_0 = 0.0$				$\delta_0 = 0.0$			
300	$\hat{\lambda}_e$	-0.025	0.067	0.080	0.567	-0.037	0.086	0.110	0.507
500		-0.012	0.055	0.050	0.630	-0.020	0.064	0.073	0.580
300	$\hat{\lambda}_e^{II}$	-0.024	0.068	0.057	0.557	-0.036	0.088	0.060	0.503
500		-0.012	0.055	0.053	0.623	-0.019	0.064	0.063	0.580
300	$\hat{\beta}_e$	0.034	0.186	0.060	0.593	-0.019	0.187	0.070	0.640
500		0.028	0.158	0.063	0.580	0.009	0.138	0.057	0.607
300	$\hat{\beta}_e^{II}$	0.034	0.188	0.060	0.590	-0.017	0.189	0.063	0.650
500		0.028	0.159	0.050	0.580	0.010	0.138	0.053	0.603
300	$\hat{\delta}$	-0.002	0.047	0.053	0.593	-0.001	0.047	0.043	0.557
500		0.002	0.038	0.057	0.747	0.003	0.038	0.050	0.750
		$\delta_0 = 0.3$				$\delta_0 = 0.3$			
300	$\hat{\lambda}_e$	-0.027	0.070	0.103	0.577	-0.040	0.092	0.143	0.517
500		-0.012	0.056	0.107	0.660	-0.021	0.066	0.097	0.563
300	$\hat{\lambda}_e^{II}$	-0.024	0.067	0.060	0.550	-0.036	0.088	0.080	0.507
500		-0.011	0.054	0.050	0.627	-0.019	0.063	0.057	0.567
300	$\hat{\beta}_e$	0.038	0.197	0.073	0.593	-0.018	0.198	0.073	0.650
500		0.028	0.165	0.073	0.593	0.006	0.146	0.047	0.617
300	$\hat{\beta}_e^{II}$	0.031	0.185	0.057	0.583	-0.019	0.187	0.063	0.630
500		0.028	0.155	0.053	0.580	0.009	0.136	0.057	0.613
300	$\hat{\delta}$	-0.007	0.041	0.047	0.750	-0.006	0.040	0.043	0.733
500		-0.002	0.042	0.057	0.917	0.000	0.032	0.050	0.903
		$\delta_0 = 0.7$				$\delta_0 = 0.7$			
300	$\hat{\lambda}_e$	-0.035	0.089	0.257	0.663	-0.064	0.131	0.300	0.643
500		-0.016	0.065	0.213	0.700	-0.033	0.088	0.207	0.607
300	$\hat{\lambda}_e^{II}$	-0.024	0.066	0.063	0.540	-0.039	0.088	0.063	0.490
500		-0.012	0.053	0.053	0.630	-0.021	0.062	0.057	0.567
300	$\hat{\beta}_e$	0.062	0.306	0.143	0.753	-0.024	0.326	0.140	0.650
500		0.041	0.240	0.107	0.707	-0.004	0.237	0.083	0.613
300	$\hat{\beta}_e^{II}$	0.032	0.177	0.057	0.770	-0.022	0.181	0.057	0.630
500		0.027	0.147	0.054	0.793	0.006	0.130	0.053	0.610
300	$\hat{\delta}$	-0.007	0.025	0.053	1.000	-0.006	0.024	0.060	1.000
500		-0.003	0.019	0.050	1.000	-0.001	0.019	0.053	1.000

We compute  $\hat{\gamma}_e = (\hat{\lambda}_e, \beta_e)'$  using equation (3.2) in Arellano and Bover (1995),

and  $\hat{\gamma}_e^{II} = (\hat{\lambda}_e^{II}, \beta_e^{II})'$  using equation (43), and (44) for its variance.

We compute  $\hat{\delta}$  using residuals  $\Delta \hat{u}_{it} = \Delta y_{it} - \hat{\gamma}_e' \Delta \mathbf{w}_{it}^d$ .



Table 3: Monte Carlo results for the conventional GMM-LEV estimator using optimal weights and the two-step GMM-LEV estimator corrected for spatial correlation

$N$	Par.	$\lambda_0 = 0.3$				$\lambda_0 = 0.7$			
		Bias	RMSE	Size	Power	Bias	RMSE	Size	Power
		$\delta_0 = 0.0$				$\delta_0 = 0.0$			
300	$\hat{\lambda}_e$	0.000	0.065	0.100	0.467	0.002	0.045	0.063	0.713
500		0.012	0.047	0.057	0.517	0.005	0.031	0.057	0.860
300	$\hat{\lambda}_e^{II}$	-0.002	0.066	0.090	0.490	0.001	0.046	0.067	0.717
500		0.011	0.047	0.053	0.513	0.005	0.032	0.050	0.853
300	$\hat{\beta}_e$	-0.001	0.173	0.060	0.647	-0.003	0.138	0.060	0.783
500		-0.017	0.133	0.050	0.663	0.000	0.105	0.047	0.703
300	$\hat{\beta}_e^{II}$	0.004	0.175	0.060	0.637	-0.002	0.140	0.060	0.780
500		-0.015	0.133	0.053	0.663	0.001	0.105	0.047	0.710
300	$\hat{\delta}$	-0.001	0.047	0.053	0.580	-0.001	0.047	0.040	0.567
500		0.003	0.038	0.057	0.747	0.003	0.038	0.057	0.750
		$\delta_0 = 0.3$				$\delta_0 = 0.3$			
300	$\hat{\lambda}_e$	-0.002	0.065	0.097	0.487	0.001	0.046	0.090	0.723
500		0.011	0.047	0.090	0.567	0.005	0.032	0.077	0.857
300	$\hat{\lambda}_e^{II}$	-0.002	0.065	0.070	0.490	0.001	0.046	0.067	0.710
500		0.011	0.047	0.063	0.523	0.005	0.032	0.047	0.873
300	$\hat{\beta}_e$	0.003	0.180	0.083	0.640	-0.002	0.145	0.080	0.780
500		-0.014	0.141	0.067	0.667	0.002	0.111	0.063	0.700
300	$\hat{\beta}_e^{II}$	0.004	0.170	0.063	0.630	0.001	0.139	0.080	0.780
500		-0.016	0.131	0.053	0.670	0.000	0.104	0.057	0.710
300	$\hat{\delta}$	-0.007	0.041	0.057	0.740	-0.004	0.040	0.057	0.733
500		-0.001	0.033	0.057	0.903	0.001	0.032	0.057	0.900
		$\delta_0 = 0.7$				$\delta_0 = 0.7$			
300	$\hat{\lambda}_e$	-0.009	0.080	0.230	0.583	-0.004	0.058	0.180	0.690
500		0.005	0.055	0.177	0.667	0.004	0.044	0.160	0.803
300	$\hat{\lambda}_e^{II}$	-0.001	0.062	0.060	0.467	0.002	0.045	0.060	0.693
500		0.012	0.047	0.050	0.500	0.006	0.032	0.050	0.880
300	$\hat{\beta}_e$	0.022	0.249	0.143	0.777	0.011	0.214	0.113	0.747
500		0.000	0.186	0.113	0.773	0.007	0.161	0.093	0.757
300	$\hat{\beta}_e^{II}$	0.000	0.162	0.050	0.730	0.002	0.134	0.057	0.783
500		-0.016	0.128	0.057	0.773	-0.001	0.100	0.043	0.720
300	$\hat{\delta}$	-0.007	0.025	0.050	1.000	-0.004	0.023	0.053	1.000
500		-0.002	0.019	0.053	1.000	0.000	0.019	0.057	1.000

We compute  $\hat{\gamma}_e = (\hat{\lambda}_e, \beta_e)'$  using equation (3.2) in Arellano and Bover (1995),

and  $\hat{\gamma}_e^{II} = (\hat{\lambda}_e^{II}, \beta_e^{II})'$  using equation (43), and (44) for its variance.

We compute  $\hat{\delta}$  using residuals  $\Delta \hat{u}_{it} = \Delta y_{it} - \hat{\gamma}_e' \Delta \mathbf{w}_{it}^d$ .

Table 4: Monte Carlo results for the conventional GMM-SYS estimator using optimal weights and the two-step GMM-SYS estimator

$N$	Par.	$\lambda_0 = 0.3$				$\lambda_0 = 0.7$			
		Bias	RMSE	Size	Power	Bias	RMSE	Size	Power
		$\delta_0 = 0.0$				$\delta_0 = 0.0$			
300	$\hat{\lambda}_e$	-0.002	0.034	0.120	0.920	-0.002	0.031	0.130	0.950
500		0.005	0.026	0.073	0.973	0.002	0.024	0.090	0.993
300	$\hat{\lambda}_e^{II}$	-0.001	0.034	0.120	0.907	-0.001	0.032	0.140	0.940
500		0.005	0.026	0.073	0.967	0.002	0.024	0.083	0.993
300	$\hat{\beta}_e$	0.000	0.058	0.083	0.867	0.000	0.055	0.073	0.933
500		-0.001	0.046	0.073	0.860	0.000	0.042	0.057	0.917
300	$\hat{\beta}_e^{II}$	0.000	0.059	0.083	0.873	0.001	0.056	0.083	0.927
500		-0.001	0.047	0.077	0.850	0.000	0.042	0.067	0.913
300	$\hat{\delta}$	-0.001	0.048	0.040	0.560	-0.001	0.048	0.040	0.557
500		0.003	0.039	0.057	0.733	0.003	0.039	0.063	0.737
		$\delta_0 = 0.3$				$\delta_0 = 0.3$			
300	$\hat{\lambda}_e$	-0.003	0.035	0.063	0.910	-0.004	0.032	0.060	0.950
500		0.005	0.026	0.063	0.970	0.002	0.024	0.063	0.990
300	$\hat{\lambda}_e^{II}$	-0.001	0.034	0.063	0.923	-0.001	0.032	0.057	0.940
500		0.005	0.026	0.050	0.977	0.002	0.023	0.060	0.993
300	$\hat{\beta}_e$	-0.001	0.061	0.063	0.843	0.000	0.059	0.060	0.997
500		-0.001	0.048	0.050	0.823	0.000	0.044	0.053	0.997
300	$\hat{\beta}_e^{II}$	-0.001	0.059	0.057	0.860	0.001	0.055	0.057	0.920
500		-0.001	0.046	0.053	0.867	0.000	0.041	0.050	0.940
300	$\hat{\delta}$	-0.001	0.040	0.057	0.707	-0.001	0.040	0.057	0.700
500		0.002	0.032	0.053	0.890	0.002	0.032	0.053	0.883
		$\delta_0 = 0.7$				$\delta_0 = 0.7$			
300	$\hat{\lambda}_e$	-0.008	0.051	0.240	0.850	-0.011	0.049	0.230	0.893
500		0.002	0.035	0.167	0.920	-0.001	0.035	0.170	0.950
300	$\hat{\lambda}_e^{II}$	0.000	0.033	0.060	0.913	-0.001	0.030	0.067	0.940
500		0.005	0.025	0.057	0.980	0.002	0.023	0.053	0.993
300	$\hat{\beta}_e$	-0.001	0.087	0.117	0.837	-0.003	0.098	0.167	0.970
500		-0.002	0.064	0.077	0.840	-0.001	0.069	0.113	0.967
300	$\hat{\beta}_e^{II}$	-0.002	0.059	0.063	0.800	-0.001	0.055	0.063	0.847
500		0.000	0.044	0.050	0.873	0.001	0.040	0.053	0.980
300	$\hat{\delta}$	-0.001	0.023	0.045	0.997	-0.001	0.023	0.049	0.997
500		0.001	0.018	0.053	1.000	0.001	0.018	0.057	1.000

We compute  $\hat{\gamma}_e = (\hat{\lambda}_e, \beta_e)'$  using equation (3.2) in Arellano and Bover (1995),

and  $\hat{\gamma}_e^{II} = (\hat{\lambda}_e^{II}, \beta_e^{II})'$  using equation (43), and (44) for its variance.

We compute  $\hat{\delta}$  using residuals  $\Delta \hat{u}_{it} = \Delta y_{it} - \hat{\gamma}_e' \Delta \mathbf{w}_{it}^d$ .

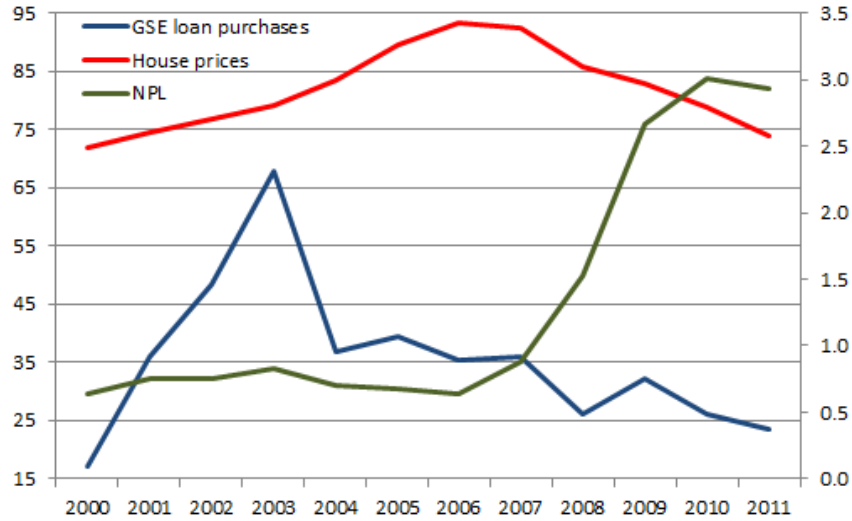
Table 5: Monte Carlo results for GMM-DIF, GMM-LEV and GMM-SYS estimators using a real world, distance based, spatial weights matrix

	Par.	Bias	RMSE	Size	Power	Bias	RMSE	Size	Power
		$\lambda_0 = 0.3, \delta_0 = 0.3$				$\lambda_0 = 0.7, \delta_0 = 0.3$			
GMM-DIFF	$\hat{\lambda}_e$	-0.028	0.091	0.076	0.268	-0.030	0.096	0.086	0.270
	$\hat{\lambda}_e^{II}$	-0.022	0.096	0.054	0.274	-0.023	0.082	0.044	0.288
	$\hat{\beta}_e$	0.009	0.183	0.068	0.290	-0.027	0.127	0.068	0.280
	$\hat{\beta}_e^{II}$	0.014	0.169	0.044	0.284	-0.014	0.110	0.044	0.282
	$\hat{\delta}$	0.016	0.099	0.048	0.402	0.019	0.085	0.042	0.396
GMM-LEV	$\hat{\lambda}_e$	0.023	0.052	0.078	0.792	0.011	0.031	0.126	0.944
	$\hat{\lambda}_e^{II}$	0.011	0.043	0.056	0.778	0.005	0.029	0.056	0.922
	$\hat{\beta}_e$	0.012	0.063	0.136	0.628	0.005	0.041	0.070	0.800
	$\hat{\beta}_e^{II}$	0.007	0.052	0.056	0.620	0.003	0.039	0.058	0.814
	$\hat{\delta}$	0.014	0.087	0.045	0.398	0.013	0.085	0.044	0.398
GMM-SYS	$\hat{\lambda}_e$	-0.002	0.032	0.082	0.912	-0.003	0.029	0.076	0.936
	$\hat{\lambda}_e^{II}$	-0.003	0.029	0.074	0.940	-0.003	0.027	0.052	0.958
	$\hat{\beta}_e$	0.001	0.057	0.104	0.588	-0.001	0.045	0.078	0.728
	$\hat{\beta}_e^{II}$	0.001	0.049	0.074	0.610	0.000	0.041	0.048	0.736
	$\hat{\delta}$	0.013	0.094	0.049	0.408	0.012	0.095	0.048	0.370
		$\lambda_0 = 0.3, \delta_0 = 0.7$				$\lambda_0 = 0.7, \delta_0 = 0.7$			
GMM-DIFF	$\hat{\lambda}_e$	-0.032	0.078	0.246	0.272	-0.032	0.070	0.274	0.266
	$\hat{\lambda}_e^{II}$	-0.021	0.091	0.042	0.374	-0.021	0.071	0.052	0.284
	$\hat{\beta}_e$	0.029	0.188	0.140	0.376	-0.040	0.128	0.186	0.294
	$\hat{\beta}_e^{II}$	0.020	0.166	0.046	0.272	-0.008	0.132	0.042	0.282
	$\hat{\delta}$	-0.040	0.093	0.048	0.488	-0.033	0.087	0.046	0.486
GMM-LEV	$\hat{\lambda}_e$	0.031	0.066	0.328	0.862	0.019	0.036	0.224	0.964
	$\hat{\lambda}_e^{II}$	0.006	0.039	0.054	0.802	0.002	0.028	0.058	0.932
	$\hat{\beta}_e$	0.020	0.064	0.154	0.676	0.001	0.037	0.092	0.814
	$\hat{\beta}_e^{II}$	0.004	0.046	0.046	0.644	0.009	0.043	0.054	0.820
	$\hat{\delta}$	-0.024	0.098	0.052	0.530	-0.023	0.097	0.044	0.532
GMM-SYS	$\hat{\lambda}_e$	-0.004	0.041	0.160	0.850	-0.007	0.039	0.146	0.874
	$\hat{\lambda}_e^{II}$	-0.003	0.028	0.056	0.950	-0.003	0.026	0.052	0.962
	$\hat{\beta}_e$	0.001	0.067	0.144	0.594	-0.002	0.057	0.126	0.708
	$\hat{\beta}_e^{II}$	0.000	0.046	0.058	0.634	-0.001	0.040	0.058	0.754
	$\hat{\delta}$	-0.024	0.076	0.048	0.528	-0.024	0.076	0.044	0.526

Table 6: Descriptive statistics

	2000-2005		2006-2011	
	Mean	Std.err.	Mean	Std.err.
<i>NPL</i> (%)	0.71	0.80	1.78	2.13
<i>HP</i>	145.42	28.43	180.13	34.23
<i>GSE</i> (1,000s)	203.73	439.97	125.99	286.01
<i>INCOME</i> (\$)	15,640	2,510	16,460	2,660
<i>UNEMP</i> (%)	5.195	1.74	7.06	3.11
<i>IRATE</i>	6.51	0.83	5.67	0.78
<i>POPDENS</i> (people per km <sup>2</sup> )	226.04	200.64	240.64	210.76
<i>EQASS</i>	0.10	0.03	0.10	0.03
<i>HHI</i>	0.42	0.26	0.44	0.27
<i>SIZE</i> (1,000\$)	343,455	416,350	440,257	685,448
<i>BRANCHES</i> (n.)	10.59	15.19	9.20	13.45

Figure 1: Temporal pattern of average real house price, total GSE loan purchases (in 100,000s) (left axis), and average nonperforming loans (right axis)



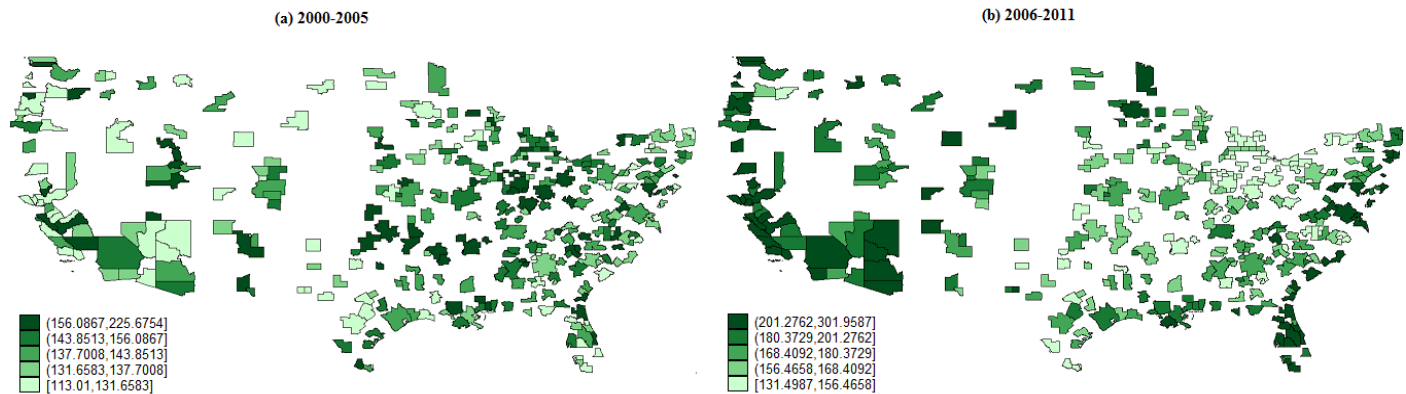


Figure 2: Quantile distribution of real house prices in US MSAs, in the years 2000 to 2005 (left) and 2006 to 2011 (right)

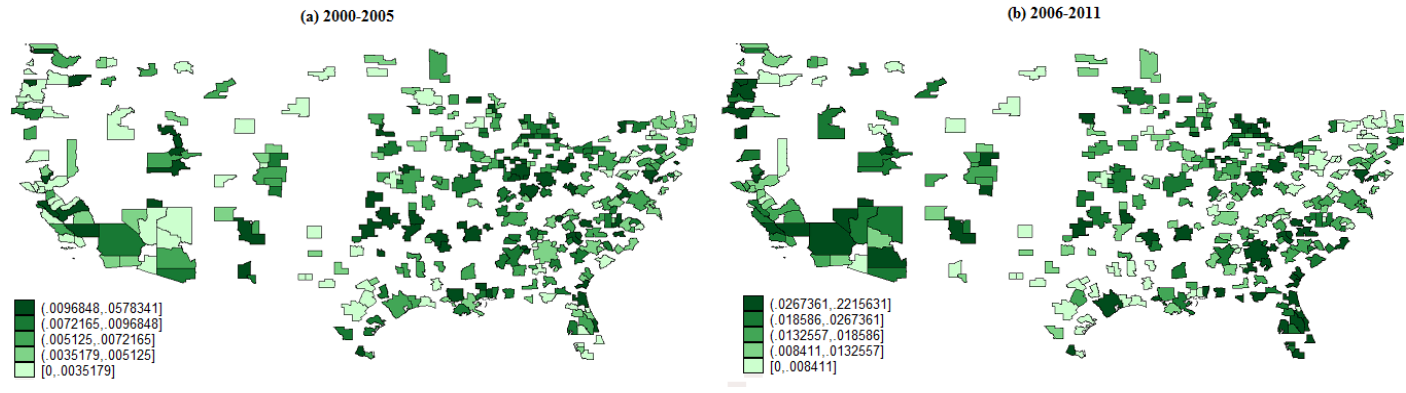


Figure 3: Quantile distribution of non performing loans in US MSAs, in the years 2000 to 2005 (left) and 2006 to 2011 (right)

Table 7: Determinants of non-performing loans in the period pre-bubble (2000 to 2005)

	(I): GMM-DIF		(II): GMM-LEV		(III): GMM-SYS	
	CONVENTIONAL GMM					
	Par.	S.E.	Par.	S.E.	Par.	S.E.
$NPL_{i,t-1}$	0.286*	0.071	0.424*	0.642	0.327*	0.061
$HP_{it}$	-1.618*	0.656	-1.483*	0.359	-1.476*	0.344
$GSE_{i,t-1}$	0.216	0.171	0.017	0.032	0.014	0.030
$INCOME_{it}$	-4.118*	2.493	-0.702*	0.245	-1.482*	0.532
$URATE_{it}$	0.699*	0.313	0.241	0.225	0.366*	0.124
$POP_{it}$	2.776	2.841	0.175	0.190	0.129	0.150
$IRATE_{it}$	0.937*	0.405	0.320	0.330	0.337	0.316
$EQASS_{it}$	-0.977	0.474	-0.147	0.309	-0.299*	0.092
$HHI_{it}$	1.357*	0.997	0.881	0.505	0.595*	0.390
$SIZE_{it}$	-0.366*	0.108	-0.035*	0.016	-0.043*	0.015
$BRANCHES_{it}$	-0.529	0.541	-0.069	0.106	-0.056	0.094
	SPATIAL GMM					
$NPL_{i,t-1}$	0.233*	0.086	0.425*	0.064	0.256*	0.051
$HP_{it}$	-0.456*	0.744	-1.543*	0.382	-1.562*	0.609
$GSE_{i,t-1}$	-0.198	0.279	0.021	0.032	0.184	0.115
$INCOME_{it}$	-3.437*	3.068	0.398	0.651	-1.027*	0.340
$URATE_{it}$	0.649*	0.284	0.992*	0.237	1.406*	0.287
$POP_{it}$	1.940*	0.932	0.185*	0.095	0.208*	0.098
$IRATE_{it}$	1.183	0.623	0.347	0.342	0.276	0.231
$EQASS_{it}$	0.797	0.585	-0.121	0.314	-0.516	0.304
$HHI_{it}$	1.366*	0.465	0.784*	0.224	0.678*	0.197
$SIZE_{it}$	-0.307*	0.274	-0.201*	0.088	-0.166*	0.031
$BRANCHES_{it}$	-0.719	0.803	-0.085	0.118	-0.379	0.177
$\hat{\delta}$	0.262*	0.247	0.358*	0.231	0.318*	0.187
AR(1)	-3.15*	[0.00]	-4.04*	[0.00]	-4.13*	[0.00]
AR(2)	-0.98	[0.33]	-0.40	[0.69]	-0.36	[0.72]
Hansen	124.66	[0.34]	156.16	[0.11]	189.70	[0.12]

Notes: (\*) denote 5 per cent significance level respectively.

Standard errors are reported in in round brackets, while  $p$ -value are shown in square brackets

Table 8: Determinants of non-performing loans in the period of the bubble bursting (2006 to 2011)

	(I): GMM-DIF		(II): GMM-LEV		(III): GMM-SYS	
	CONVENTIONAL GMM					
	Par.	S.E.	Par.	S.E.	Par.	S.E.
$NPL_{i,t-1}$	0.290*	0.037	0.441*	0.055	0.350*	0.036
$HP_{it}$	-1.526*	0.354	-0.739*	0.271	-0.872*	0.268
$GSE_{i,t-1}$	0.241*	0.109	0.527*	0.128	0.434*	0.085
$INCOME_{it}$	-2.820*	1.105	-1.176*	0.321	-1.000*	0.311
$URATE_{it}$	1.339*	0.194	1.184*	0.176	1.033*	0.141
$POP_{it}$	2.718	1.681	0.110	0.091	0.502	0.462
$IRATE_{it}$	0.454*	0.151	0.452*	0.207	0.407*	0.127
$EQASS_{it}$	-0.621*	0.231	-0.129	0.332	-0.523*	0.174
$HHI_{it}$	-0.430	0.584	-0.526	0.905	-0.157	0.439
$SIZE_{it}$	-0.195*	0.064	-0.109*	0.028	-0.078*	0.019
$BRANCHES_{it}$	-0.267	0.249	-0.302	0.171	-0.193*	0.090
	SPATIAL GMM					
$NPL_{i,t-1}$	0.240*	0.055	0.269*	0.050	0.260*	0.051
$HP_{it}$	-0.840*	0.986	-0.713*	0.479	-0.750*	0.264
$GSE_{i,t-1}$	0.231*	0.091	0.365*	0.108	0.380*	0.114
$INCOME_{it}$	-2.367*	2.451	-1.029*	0.337	-1.031*	0.339
$URATE_{it}$	1.268*	0.304	1.214*	0.187	1.337*	0.249
$POP_{it}$	2.441*	0.935	0.196*	0.095	0.197*	0.087
$IRATE_{it}$	0.350	0.236	0.341	0.228	0.418	0.227
$EQASS_{it}$	-0.701*	0.051	-0.503*	0.194	-0.514*	0.101
$HHI_{it}$	-0.736	1.230	0.673*	0.076	0.684*	0.091
$SIZE_{it}$	-0.469*	0.219	-0.149*	0.024	-0.161*	0.078
$BRANCHES_{it}$	-0.193	0.872	-0.320*	0.155	-0.362	0.272
$\hat{\delta}$	0.406*	0.1917	0.558*	0.231	0.551*	0.227
AR(1)	-3.84*	[0.00]	-4.23*	[0.00]	-4.74*	[0.00]
AR(2)	-0.34	[0.73]	-0.29	[0.77]	0.07	[0.94]
Hansen	294.55	[1.00]	270.10	[1.00]	308.40	[1.00]

Notes: (\*) denote 5 per cent significance level respectively.

Standard errors are reported in in round brackets, while  $p$ -value are shown in square brackets