ABSTRACT

ERYURUK, GUNCE. Three Essays on Dynamic Panel Data Estimation. (Under the direction of Mehmet Caner).

This dissertation consists of three essays, first two of which consider a new estimation method of dynamic panel data models and the last one considers an application of these models. The first essay (Chapter 1) offers empirical likelihood (EL) estimation of dynamic panel data models, which provide great flexibility to empirical researchers. EL estimation method is shown to have great advantages in usual settings, however little is known on the relative merits of these estimators in panel data models. With this essay, we try to fill that gap by establishing the asymptotic properties of the EL estimator for a dynamic panel model with individual effects when both the time and the cross-section dimensions tend to infinity. We give the conditions under which this estimator is consistent and asymptotically normal. In the second essay (Chapter 2), via a Monte Carlo study, we assess the relative finite sample performances of EL, generalized method of moments, and limited information maximum likelihood estimators for an autoregressive panel data model when there are many moment conditions. We also extend our results to the many weak moments settings. Our results suggest that when the overall performances are concerned, in terms of median, interquartile range and median absolute error of the estimators, in both strong and weak moments settings, EL is more reliable. In the final essay (Chapter 3) we consider an application of dynamic panel data models to examine the determinants of the allocation of state highway funds using panel data for North Carolina's 100 counties for the years 1990 to 2005. We make two main contributions with this essay. First, although there have been numerous studies of highway funding at the state level, to our knowledge, there is no analysis at the sub-state or county levels. Second, by using dynamic panel data models and sophisticated methods to estimate them, we account for any potential persistence in the process of adjustment toward an equilibrium, besides, unlike most of the previous studies, we control for the unobserved county heterogeneity and time effects that explain spatial differences, which may cause omitted variable problem if ignored.

Three Essays on Dynamic Panel Data Estimation

by
Gunce Eryuruk

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fullfillment of the
requirements for the Degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina

2009

APPROVED BY:

Mehmet Caner Chair of Advisory Committee	Atsushi Inoue
A. Ronald Gallant	Xiaoyong Zheng

DEDICATION

Babama,
sınırlarımı zorlamayı öĝrettiĝi için

(To my father
who taught me to push my boundaries)

BIOGRAPHY

Gunce Eryuruk was born in 1979 in Eskişehir, a city in west-central Turkey. Immediately after her birth the family moved to İzmir, a city on the west cost of Turkey, where she lived until she finished highschool. She attended college in İstanbul at Boĝaziçi University where she received a B.A. in Economics in 2002.

In August of 2002, she started the Ph.D program in Economics at North Carolina State University and received her doctorate degree in the summer of 2009. Gunce will be an assistant professor at Instituto Tecnológico Autónomo de México (ITAM) in Mexico City starting in August of 2009.

ACKNOWLEDGMENTS

I owe my deepest gratitude to Dr. Mehmet Caner, my advisor, for his invaluable guidance, support, and encouragement when I needed the most. My deepest gratitude extends equally to Dr. Ron Gallant for granting me the privilege to benefit from his endless knowledge and helping me in any way possible, to Dr. Atsushi Inoue for his open door policy whenever I sought help, and to Dr. Xiaoyong Zheng for his helpful comments. I am also grateful to Dr. Alastair Hall for helping me with the initial idea about my topic. I extend my sincerest gratitude to Dr. Michael Walden for providing me the insight and the tools to help me finish my dissertation.

My special thanks go to my best friends Jared Walton and Sevgin Yılmaz for supporting me to the end.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

The Time Series and Cross-Section Asymptotics of Empirical Likelihood Estimator in Dynamic Panel Data Models

1.1 Introduction

Dynamic panel data models offer great flexibility to empirical researchers. Many economic phenomena are dynamic in nature. Examples include household consumption, firms' factor demands, and countries' economic growth. Dynamic panel data models allow researchers to control for unobserved heterogeneity in adjustment dynamics between different individual units and thereby provide improved insights in such models.

When dynamic models are estimated using panel data, the "usual" least squares methods lead to inconsistent estimates for the parameters of the models when the time dimension (T) is short regardless of the cross sectional dimension (N). This inconsistency stems from the fact that the disturbance terms are correlated with the lagged endogenous variable[1]. Moreover, under large N fixed T asymptotics it is well known from Nickell (1981) that the standard maximum likelihood estimator suffers from an incidental parameter

[1]The nature and the magnitude of this inconsistency are well defined by Nickell (1981) and Sevestre & Trognon (1985) for fixed and random effects specifications, respectively.

problem leading to inconsistency. In order to avoid this problem, the literature has focused on generalized method of moments (GMM) estimator applied to first differences. Examples include Anderson & Hsiao (1982), Holtz-Eakin, Newey, & Rosen (1988), and Arellano & Bond (1991).

However, the standard GMM estimator obtained after first differencing has been found to suffer from substantial finite sample bias, especially when the instruments are weak[2] and the number of moments is large relative to the cross section sample size. See Alonso-Borrego & Arellano (1999).

This low precision of GMM is also evident in more general contexts. To improve the small sample properties of GMM estimators, a number of alternative estimators have been suggested, including, among others, EL, continuous updating (CU), and exponential tilting (ET) estimators. Newey & Smith (2004) show that these estimators are members of a class of generalized empirical likelihood (GEL) estimators. They use this structure to compare their higher order asymptotic properties with those of GMM. They find that EL has two theoretical advantages. First, its asymptotic bias does not grow with the number of moment restrictions, while the bias of GMM often does. This, as a result, suggests that in estimation of models with many moment conditions, the bias of EL will be less than the bias of GMM. Consequently, EL can be an important alternative to GMM in such applications. Furthermore, they show that under a symmetry condition, which may be satisfied in some instrumental variable settings, all GEL estimators inherit the small bias property of EL. The second theoretical advantage of EL estimator is that after it is bias-corrected using probabilities obtained from EL, it is higher order efficient relative to other bias-corrected estimators.

The purpose of this article is to provide further insight into the asymptotic properties of the EL estimator in dynamic panel data framework by allowing both N and T tend to infinity and to study its behavior for alternative relative rates of increase for N and T. This asymptotics is motivated by the increased availability of panel data sets covering different individuals, regions, and countries over a relatively long time period. Among the important examples of these data sets are the PSID household panel in the US, Penn World table and the balance sheet-based company panels. For panels in which T is not negligible

[2]This occurs when the series are highly autoregressive, i.e., the autoregressive parameter is close to one, and the relative variance of fixed effects to the variance of idiosyncratic shocks is large.

relative to N, the analysis of the asymptotic behavior of the estimators as both T and N tend to infinity may provide better approximations to the finite sample behavior of the estimators and hence may be useful in assessing alternative methods. Previously, this type of asymptotics is used by Alvarez & Arellano (2003). They derive the asymptotic properties of within groups (WG), GMM, and limited information maximum likelihood (LIML) estimators.

It is also the case that as T grows, for each period, the number of available lags which can be used as instruments for the equations in the first differences grows at the rate of $T(T-1)/2$. This corresponds to what is known in the literature as the "many moment conditions" situation. It is well known in the linear instrumental variable regression models that using many moments causes the usual Gaussian asymptotic approximation to be poor. In more general contexts this problem was addressed by, among others, Han & Phillips (2006) and Newey & Windmeijer (2007). They point out that the two-step GMM estimator can be very biased. On the other hand, GEL estimator has smaller bias but the usual standard errors are found to be too small. In their study, Newey & Windmeijer (2007), consider alternative asymptotics that addressed this problem, i.e., asymptotic properties of GEL and GMM under "many weak moment conditions". They find that the two-step GMM estimator is asymptotically biased under this scheme, whereas the GEL estimator is not. In addition, they find that, under the alternative asymptotics, GEL has a Gaussian limit distribution with asymptotic variance larger than the usual one and for this estimator they propose an appropriate variance estimator that is consistent under standard and alternative asymptotics.

It is natural to expect that the GEL estimators possess similar advantages in dynamic panel data models under the many moment conditions setting. Unfortunately, little is known on the relative merits of these estimators in dynamic panel data models under the double asymptotics, i.e., asymptotics taken as both T and N going to infinity. In a Monte Carlo study, Oĝuzoĝlu (2006) compares performance of a number of estimators including GMM, EL, transformed ML, minimum distance and bias corrected LSDV estimators in an autoregressive panel model for various parameter combinations. The results show that the biases of all estimators considered tend to increase as the autoregressive parameter gets larger. The increase in bias is the highest for LSDV, whereas EL is the least sensitive to changes in this parameter. Moreover, the bias of GMM does not decrease much as T

increases. When the overall performances are concerned, i.e., in terms of comparisons based on biases, standard deviations, and root mean square errors, EL performs the best.

In the same framework, Gonzalez (2007) considers the finite-sample size properties of the overidentification tests for a hybrid of EL and bootstrap estimators. Previously, a similar study was carried out by Brown & Newey (2001) and Bowsher (2002) for the GMM estimator. Gonzalez (2007) investigates whether the limitations encountered within GMM estimation are extended to EL-bootstrap estimator. Her results show that EL-bootstrap may be a good alternative to GMM estimator within this setting. She also applies this estimator using the cash-flow series data for 174 US firms.

Although a few studies considered the finite sample performance of the EL estimator in dynamic panel data models, none of them, to our knowledge, analyze its asymptotic performance explicitly under the aforementioned settings. This article tries to fill this gap. Specifically, we establish the asymptotic properties of the EL estimator for a first-order autoregressive model with individual effects when both N and T tend to infinity. We show that this estimator is consistent and asymptotically normal. We also compare the asymptotic properties of EL estimator with those of the GMM and LIML estimators, which are popular in empirical research.

The chapter is organized as follows. Section 2 presents the model and the estimators. In section 3 we establish the asymptotic properties of the EL estimator. For comparison purposes we give those for the LIML and GMM estimators, as well. Section 4 concludes. Proofs are relegated to the Appendix A.

1.2 The Model and The Estimators

1.2.1 The Model

We consider a first order univariate autoregressive panel data model given by

$$y_{it} = \alpha_0 y_{i,t-1} + \eta_i + v_{it}, \quad \text{for} \quad t = 1, \ldots, T; \;\; i = 1, \ldots, N \tag{1.1}$$

where y_{it} is the observable variable whose dynamics are of interest; for example, local government expenditure variable, $|\alpha_0| < 1$, η_i is the fixed effect representing the unobserved heterogeneity among individuals, and v_{it} is the idiosyncratic variable with zero mean and variance σ^2 given $\eta_i, y_{i0}, \ldots, y_{i,t-1}$ and has no autocorrelation. We assume that y_{i0} is observed. Define $x_{it} \equiv y_{i,t-1}$.

The parameter of interest is α_0. Our goal is to analyze the asymptotic properties of EL estimator of this parameter. For comparison purposes we are going to consider that of GMM and LIML estimators. Next we shall define these estimators.

1.2.2 The Estimators

The GMM Estimator. The GMM estimator considered here is a version developed by Arellano & Bover (1995), which simplifies characterization of the "weight matrix" in GMM estimation. Arellano & Bover (1995) eliminate the fixed effect η_i in (1.1) by applying Helmert's transformation. For example, the t-th element of transformed v_{it} can be written as:

$$v_{it}^* = c_t[v_{it} - \frac{1}{T-t}(v_{i,t+1} + \ldots + v_{iT})] \qquad t = 1, \ldots, T-1$$

where $c_t^2 = (T-1)/(T-t+1)$. That is, to each of the first $(T-1)$ observations the mean of the remaining future observations available in the sample is subtracted. The weighting c_t is introduced to equalize the variances. This transformation can be applied by using the forward orthogonal deviations operator, A, where

$$A = diag[\frac{T-1}{T}, \ldots, \frac{1}{2}]^{1/2} \begin{bmatrix} 1 & -\frac{1}{T-1} & -\frac{1}{T-1} & \ldots & -\frac{1}{T-1} & -\frac{1}{T-1} & -\frac{1}{T-1} \\ 0 & 1 & -\frac{1}{T-2} & \ldots & -\frac{1}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -1 \end{bmatrix}.$$

Equation (1.1) with the variables stacked over t can be written as

$$y_i = \alpha_0 x_i + \eta_i \imath_T + v_i$$

where $y_i = [y_{i1}, \ldots, y_{iT}]'$, $x_i = [x_{i1}, \ldots, x_{iT}]'$, $v_i = [v_{i1}, \ldots, v_{iT}]'$, and $\imath$ dimension T vector of ones. Operating A on this equation produces the transformed model:

$$y_i^* = \alpha_0 x_i^* + v_i^* \tag{1.2}$$

where $y_i^* = Ay_i$, $x_i^* = Ax_i$, $v_i^* = Av_i$. Note that the fixed effect are eliminated because $A\imath = 0$. Also, $A'A = Q_T \equiv I_T - \imath_T \imath_T'/T$ (Q_T is known as WG operator) and $AA' = I_{T-1}$. Thus, if $Var(v_i) = \sigma^2 I_T$, the vector of errors in orthogonal deviations also has $Var(v_i^*) = \sigma^2 I_{T-1}$.

Let $z_{it} = [x_{i1}, \ldots, x_{it}]'$. The model (1.2) and the stated conditions imply the following moment conditions

$$E[z_{it} v_{it}^*] = 0 \qquad t = 1, \ldots, T-1. \tag{1.3}$$

There are $m \equiv T(T-1)/2$ orthogonality conditions. These moment conditions can be written, more compactly, as

$$E[Z_i' v_i^*] = 0,$$

where

$$\underset{(T-1)\times m}{Z_i} = \begin{bmatrix} \underset{1\times 1}{z_{i1}}' & 0 & \ldots & 0 \\ 0 & \underset{1\times 2}{z_{i2}}' & 0 \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \underset{1\times (T-2)}{z_{iT-1}}' \end{bmatrix} = \begin{bmatrix} y_{i0} & 0 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & y_{i0} & y_{i1} & & 0 & & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \ldots & y_{i0} & \ldots & y_{i(T-2)} \end{bmatrix}.$$

The constant variance of v_{it} given $\eta_i, y_{i0}, \ldots, y_{i,t-1}$ implies that

$$E(Z_i' v_i^* v_i^{*\prime} Z_i) = \sigma^2 E(Z_i' Z_i). \tag{1.4}$$

Therefore, letting $x^* = (x_1^{*\prime}, \ldots x_N^{*\prime})'$ and $y^* = (y_1^{*\prime}, \ldots y_N^{*\prime})'$, an asymptotically efficient GMM estimator of α_0 based on the moment conditions in (1.3) is given by

$$\widehat{\alpha}_{GMM} = \frac{x^{*\prime} Z(Z'Z)^{-1} Z y^*}{x^{*\prime} Z(Z'Z)^{-1} Z x^*}$$

where $Z = (Z_1', \ldots, Z_N')'$.

The LIML Estimator. A non-robust analog of the LIML estimator of the simultaneous equations literature solves the following problem:

$$\widehat{\alpha}_{LIML} = \arg\min_{\alpha} \frac{(y^* - \alpha x^*)' Z (Z'Z)^{-1} Z' (y^* - \alpha x^*)}{(y^* - \alpha x^*)'(y^* - \alpha x^*)}.$$

The robust LIML analog, or "continuously updated" GMM estimator in the terminology of Hansen, Heaton, & Yaron (1996), can be written as

$$\widehat{\alpha}_{CU} = \arg\min_{\alpha} (y^* - \alpha x^*)' Z \left(\sum_{i=1}^{N} Z_i' (y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)' Z_i \right)^{-1} Z' (y^* - \alpha x^*).$$

In the robust version, instead of keeping σ^2 fixed in the weighting matrix of the GMM criterion, it is continuously updated by making it a function of the arguments in the estimating criterion.

The EL Estimator. Empirical Likelihood estimation (Qin & Lawless (1994) and Imbens (1997)) is a one-step method that achieves the same first-order asymptotic efficiency as robust GMM.

The empirical likelihood estimator maximizes a multinomial pseudo likelihood (or empirical likelihood) function subject to the orthogonality conditions. Letting p_i be the probability of observation i, the multinomial log likelihood of the data is given by the empirical likelihood estimator:

$$L = \sum_{i=1}^{N} \ln p_i.$$

The EL estimator maximizes this function subject to the restrictions

$$p_i \geq 0, \sum_{i=1}^{N} p_i = 1 \text{ and } \sum_{i=1}^{N} p_i Z_i' (y_i^* - \alpha x_i^*) = 0.$$

The Lagrangian is given by

$$\mathcal{L} = \sum_{i=1}^{N} p_i + \phi(1 - \sum_{i=1}^{N} p_i) - N\lambda' \sum_{i=1}^{N} p_i Z_i' (y_i^* - \alpha x_i^*),$$

where λ and ϕ are Lagrange multipliers. Taking the derivative of $\mathcal{L}$ with respect to p_i we obtain the following first-order conditions

$$\frac{1}{p_i} - \phi - N\lambda' Z_i'(y_i^* - \alpha x_i^*) = 0.$$

Multiplying by p_i and adding equations we get $\phi = N$. Hence,

$$p_i = \frac{1}{N}\left(\frac{1}{1+\lambda' Z_i'(y_i^* - \alpha x_i^*)}\right).$$

The multipliers of the moment conditions can be determined as implicit functions $\lambda(\alpha)$ solving (for a given value of α):

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{1+\lambda' Z_i'(y_i^* - \alpha x_i^*)}\right) Z_i'(y_i^* - \alpha x_i^*) = 0$$

such that $1 + \lambda' Z_i'(y_i^* - \alpha x_i^*) \geq 1/N$.

The concentrated likelihood function for α:

$$\mathcal{L}_c(\alpha) = \prod_{i=1}^{N}\frac{1}{N}\left(\frac{1}{1+\lambda' Z_i'(y_i^* - \alpha x_i^*)}\right).$$

Therefore, the EL estimator is given by

$$\widehat{\alpha}_{EL} = \arg\min_{\alpha}\sum_{i=1}^{N}\ln[1+\lambda(\alpha)' Z_i'(y_i^* - \alpha x_i^*)].$$

A computationally useful alternative expression for α_{EL} is

$$\widehat{\alpha}_{EL} = \arg\min_{\alpha}\widehat{Q}(\alpha), \text{ where } \widehat{Q}(\alpha) = \max_{\lambda}\frac{1}{NT}\sum_{i=1}^{N}\ln[1+\lambda' Z_i'(y_i^* - \alpha x_i^*)].$$

1.3 The Asymptotic Properties Of The Estimators

In this section we derive the asymptotic properties of the previous estimators when both N and T tend to infinity. Following Alvarez & Arellano (2003), we make the following assumptions:

Assumption 1. $\{v_{it}\}$ $(t = 1, \ldots T; i = 1, \ldots N)$ *are i.i.d across time and individuals and* *independent of* η_i *and* y_{i0} *with* $E[v_{it}] = 0$, $Var[v_{it}] = \sigma^2$, *and finite moments up to fourth order.*

Assumption 2. *The initial observations satisfy*

$$y_{i0} = \frac{\eta_i}{1-\alpha_0} + \omega_{i0}$$

where ω_{i0} *is independent of* η_i *and i.i.d. with the steady state distribution of the homogenous process, so that* $\omega_{i0} = \sum_{j=0}^{\infty} \alpha_0^j v_{i,(-j)}$.

Assumption 3. η_i *are i.i.d. across individuals with* $E[\eta_i] = 0$, $Var[\eta_i] = \sigma_\eta^2$ *and finite fourth order moment.*

Note that under these assumptions, the moment conditions given in (1.3) do not represent all the available moment conditions available. Ahn & Schmidt (1995) present additional moment conditions and argue that they are important in improving the GMM estimation in highly persistent samples. However, we focus only on the moment conditions in (1.3) as they remain valid under much weaker assumptions.

1.3.1 The GMM and the LIML Estimators

Alvarez & Arellano (2003) show that under the stated assumptions as both N and T tend to infinity, provided $(\log T)/N \to 0$, $\widehat{\alpha}_{GMM}$ is consistent for α_0:

$$\widehat{\alpha}_{GMM} \xrightarrow{p} \alpha_0$$

Moreover, provided $T/N \to c$, $0 \leq c \leq \infty$

$$\sqrt{NT}\left[\widehat{\alpha}_{GMM} - \left(\alpha_0 - \frac{1}{N}(1+\alpha_0)\right)\right] \xrightarrow{d} \mathcal{N}\left(0, 1-\alpha_0^2\right).$$

Although, the number of moment conditions, m tend to infinity (at the rate T^2), their result show that $\widehat{\alpha}_{GMM}$ remains consistent. However, in the structural equation setting, when both the number of instruments and the sample size tend to infinity, while their ratio tends to a positive constant, the two-stage least squares estimator is shown to be inconsistent (Kunitomo (1980), Morimune (1983), and Bekker (1994)). The intuition for this consistency of $\widehat{\alpha}_{GMM}$ is defined by Alvarez & Arellano (2003) as in structural equation setting too many instruments produces over fitting and undesirable closeness to the OLS coefficients. Here a large number of instruments is associated with larger values of T and in such a case closeness to OLS, which is the WG estimator, becomes increasingly desirable because "endogeneity bias" tends to zero as $T \to \infty$.

For, LIML estimator, they show that, under the stated assumptions, as both N and T tend to infinity, provided $T/N \to c$, $0 \leq c \leq 2$, $\widehat{\alpha}_{LIML}$ is consistent for α_0:

$$\widehat{\alpha}_{LIML} \xrightarrow{p} \alpha_0$$

Moreover,

$$\sqrt{NT}\left[\widehat{\alpha}_{LIML} - \left(\alpha_0 - \frac{1}{2N-T}(1+\alpha_0)\right)\right] \xrightarrow{d} \mathcal{N}\left(0, 1-\alpha_0^2\right).$$

Note that GMM and LIML estimators are both asymptotically normal with the same asymptotic variance, however, unless $T/N \to 0$, they exhibit a bias term in their asymptotic distributions differing in its order of magnitude: $((1+\alpha)/N$ for GMM and $((1+\alpha)/(2N-T)$ for LIML. Provided $T < N$, the LIML has a smaller asymptotic bias.

1.3.2 The EL Estimator

For consistency and asymptotic normality of EL estimator some additional assumptions are needed. Let $\lambda_{min}(S)$ and $\lambda_{max}(S)$ denote the smallest and the largest eigenvalues of a symmetric matrix S, respectively.

Assumption 4. *(i) There is $C > 0$ such that $1/C \leq \lambda_{min}\big(E[Z_i'(y_i^* - \alpha x_i^*)$ $(y_i^* - \alpha x_i^*)'Z_i]\big), \lambda_{max}\left(E[Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)'Z_i]\right) \leq C$, and $\lambda_{max}\left(E[Z_i'x_i^*x_i^{*\prime}Z_i]\right) \leq C$; (ii) $\sup_{\alpha}\|\frac{1}{N}\sum_{i=1}^{N} Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)'Z_i - E[Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)'Z_i]\| \xrightarrow{p} 0$.*

Assumption 5. *For a constant C and $\gamma > 2$, $E|v_{it}|^{2\gamma} < C$ and $E|\eta_i|^{2\gamma} < C$.*

Assumption 4 puts restrictions on the rate at which T and hence the number of moment conditions, m, can grow relative to N. Assumption 5 puts further moment restriction on the error and the fixed effect terms. Note that Assumption 5 along with Assumption 2 imply that $E|y_{i0}|^{2\gamma} < C$.

The following is a consistency result for the EL estimator.

Theorem 1. *Let Assumptions 1-5 hold. Then as both N and T tend to infinity, provided $N^{1/\gamma}T^{3-2/\gamma}\sqrt{T^2/N} \to 0$, $\gamma > 2$, $\widehat{\alpha}_{EL}$ is consistent for α_0:*

$$\widehat{\alpha}_{EL} \xrightarrow{p} \alpha_0.$$

For the consistency of the EL estimator a further restriction is need on the relative rates at which T and N can grow. This is also the case in Newey & Windmeijer (2007) for a general cross sectional model.

Next we give the asymptotic normality result. This result is parallel to Theorem 3 of Newey & Windmeijer (2007).

Theorem 2. *Let Assumptions 1-5 hold. Then as both N and T tend to infinity, provided $N^{1/\gamma}T^{3-2/\gamma}\sqrt{T^2/N} \to 0$, $\gamma > 2$, and $T^{11}/N \to 0$,*

$$\sqrt{NT}(\widehat{\alpha}_{EL} - \alpha_0) \xrightarrow{d} N(0, 1 - \alpha_0^2).$$

Note that Newey & Windmeijer (2007) give an asymptotic variance of GEL estimator as a summation of two terms. The first term corresponds to the conventional asymptotic variance term of GMM. The additional term can be considered as a "higher order" variance term in asymptotic theory with fixed number of moment conditions. They note that this term can be important even when the sample size is large under certain conditions, which includes weak moments. Under the restrictions on the relative rates on N and T given by Theorem 2, the additional terms, in our case, tend to zero[3]. Hence, the gradient of the EL objective function at the true parameter, α_0, scaled by $\sqrt{NT}$ takes the following form:

$$\sqrt{NT}\frac{\partial\widehat{Q}(\alpha_0)}{\partial\alpha} = \frac{1}{\sqrt{NT}}x^{*\prime}Z\left(Z'v^*v^{*\prime}Z\right)^{-1}Z'v^* + o_p(1).$$

The asymptotic variance of $\sqrt{NT}\frac{\partial\widehat{Q}(\alpha_0)}{\partial\alpha}$ converges in probability to

$$\frac{1}{\sqrt{NT}}x^{*\prime}Z\left(Z'v^*v^{*\prime}Z\right)^{-1}Z'x^* \xrightarrow{p} \frac{1}{1-\alpha_0^2}.$$

Then, as shown in Appendix A, theorem 2 follows from theorem 3 of Newey & Windmeijer (2007).

[3]The proof is available upon request from the author.

Theorem 2 requires that for asymptotic normality T has to grow much slower than N does. More specifically, it is required that $\lim(T/N) \to c = 0$. This condition is much more strict than $0 \leq c \leq 2$ requirement which is needed for asymptotic normality of LIML. Moreover, for LIML, when $c = 0$, the asymptotic bias disappears and the two estimators become asymptotically equal.

In the literature, performances of these estimators are compared through Monte Carlo studies. Although, to our knowledge none of these studies have compared all these three estimators in the same model under the same specifications, they have been compared pairwise on not so different settings. In their Monte Carlo study, Newey & Windmeijer (2007) consider GMM and GEL estimators in a panel data model with heterogenous idiosyncratic errors and predetermined variables which are introduced via inclusion of a lagged independent variable as an explanatory variable. They compare their interquartile range and median bias. They show that when the instruments are strong[4] biases are negligible for these estimators with comparable interquartile ranges. When the instruments are weaker[5], the GMM estimators are downward biased, whereas the CU estimator is median unbiased but exhibits a larger interquartile range than the GMM estimators. Moreover, when they included lags of dependent variable as sequential instruments additional to the sequential lags of the independent variable the GMM estimators become more downward biased whereas the CU estimator is still median unbiased with an interquartile range decreased by 15%.

Oĝuzoĝlu (2006) compares the performances of EL and GMM estimators, among the others, based on their bias, standard deviation and root mean square, in a model identical to the one we consider here. He shows that in the case of strong instruments the EL estimator performs better in general. In the case of weak instruments, EL estimator, again, exhibits the best performance except when T is small. In this case, its performance suffers mostly from its high variance.

Although, for EL estimator, the condition on the rate at which T can grow is much more strict than that of LIML and GMM for asymptotic normality, Monte Carlo studies suggest that EL usually outperforms LIML or GMM in various settings.

[4]The case in which the coefficient of the lagged independent variable is equal to 0.40.
[5]the case in which the coefficient of the lagged independent variable is equal to 0.85.

1.4 Conclusion

We show that in autoregressive panel data models, the EL estimator that uses all the moment conditions based on all the available lags at each period are consistent and asymptotically normal when both N and T tend to infinity. When showing normality, we applied Newey & Windmeijer (2007) method that they use for a general cross sectional model. For the EL estimator, for normality, the required condition on the relative rates at which N and T can grow turns out to be much more strict than that of the LIML and GMM estimators. Under this restriction, the LIML and GMM asymptotic biases disappear. Therefore, all three estimators that we consider have the same asymptotic distribution.

Chapter 2

A Comparative Study of Common Dynamic Panel Data Model Estimators

2.1 Introduction

Dynamic panel data models are of great interest in both macroeconomic and microeconomic applications. In estimation of these models generalized method moments (GMM) are widely used.

Unfortunately, the standard GMM estimator obtained after first differencing has been found to suffer from substantial finite sample bias, especially when the instruments are weak and large in number compared to the cross section sample size. See Alonso-Borrego & Arellano (1999).

Motivated by this problem, many alternative estimators are suggested in the literature. Among others, Hahn, Hausman & Kuersteiner (2001) proposed several moment-based estimators with bias-reducing properties. One of their estimators extend the Nagar (1959) bias-corrected estimator (including the limited information maximum likelihood estimator) to dynamic panel data models. Blundell & Bond (1998) motivated by Arellano & Bover (1995) offer an extended GMM estimator in which lagged first differences of the series are also used as instruments for the levels equations. This extended version outperforms the standard GMM estimator especially in cases where lagged levels of the series used as instruments are only weakly correlated with subsequent first differences.

In this article we analyze another estimator that has many important advantages over GMM, namely, empirical likelihood (EL) estimator (Owen, 1990; Qin & Lawless, 1994; Imbens, 1997). Advantages of EL estimator can be stated as follows. First of all, EL is a one-step estimator, hence, a source of bias that arises from the choice of preliminary estimator is eliminated with EL estimator. Second, EL wights are sensitive to the validity of moment conditions, which is a great theoretical advantage when the moment restrictions are weakly defined. Third, it has been shown that GMM that uses the optimal weighting matrix attends the semi-parametric efficiency bound (Chamberlain 1987). EL shares the same first order asymptotic property (Smith 1997). Moreover, Newey & Smith (2004) reports that EL has certain advantages relating to higher order asymptotics. Namely, EL does not have some components of the second order bias, characteristic of GMM estimators, and its asymptotic bias, unlike GMM, does not increase with the number of ever-identifying restrictions.

It is natural to expect that EL estimator possesses similar advantages in dynamic panel data models as well. Unfortunately, little is known on the relative merits of this estimator in these models. With this article we are aiming to provide some insight into the finite sample behavior of the EL estimator for these models. We are particulary interested in the case where there are many moment conditions and the extension to the many and weak moment conditions case. These are the settings that are very common in dynamic panel models. We also compare the performance of EL estimator with other well known estimators of dynamic panel data models, namely GMM and LIML, for alternative relative values for the cross sectional dimension (N) and the time series dimension (T).

2.2 The Model and The Estimators

2.2.1 The Model

We consider a first order autoregressive model with unobserved individual-specific effects

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it}, \quad \text{for} \quad t = 1, \ldots, T; \;\; i = 1, \ldots, N \tag{2.1}$$

where $|\alpha| < 1$ and v_{it} is the idiosyncratic variable with zero mean and variance σ^2 given $\eta_i, y_{i0}, \ldots, y_{i,t-1}$ and has no autocorrelation. We assume that y_{i0} is observed.

It is convenient to write the model in (2.1) in the form:

$$y_i = \alpha y_{i(-1)} + \eta_i \imath_T + v_i \tag{2.2}$$

where $y_i = [y_{i1}, \ldots, y_{iT}]'$, $y_{i(-1)} = [y_{i0}, \ldots, y_{i,T-1}]'$, $v_i = [v_{i1}, \ldots, v_{iT}]'$, and $\imath$ dimension T vector of ones.

Our goal is to compare the finite sample behavior of GMM, LIML, and EL estimators of α. Next we shall define these estimators.

2.2.2 The Estimators

The GMM Estimator. The GMM estimator considered here is a version developed by Arellano & Bover (1995), which simplifies characterization of the "weight matrix" by using the orthogonal deviations operator A which is the $(T-1) \times T$ upper triangular matrix such that $A'A = Q_T \equiv I_T - \imath_T \imath_T'/T$ (Q_T is known as WG operator) and $AA' = I_{T-1}$. Thus, if $Var(v_i) = \sigma^2 I_T$, the vector of errors in orthogonal deviations also has $Var(v_i^*) = \sigma^2 I_{T-1}$. Note that the fixed effect are eliminated because $A\imath = 0$. Operating A on (2.2) produces the transformed model:

$$y_i^* = \alpha y_{i(-1)}^* + v_i^* \tag{2.3}$$

where $y_i^* = Ay_i$, $y_{i(-1)}^* = Ay_{i(-1)}$, $v_i^* = Av_i$.

Let $z_{it} = [y_{i0}, \ldots, y_{i,t-1}]'$. The model (2.3) and the stated conditions imply the following moment conditions

$$E[z_{it} v_{it}^*] = 0 \qquad t = 1, \ldots, T-1. \tag{2.4}$$

There are $m \equiv T(T-1)/2$ orthogonality conditions. These moment conditions can be written, more compactly, as

$$E[Z_i' v_i^*] = 0,$$

where

$$\underset{(T-1)\times m}{Z_i} = \begin{bmatrix} \underset{1\times 1}{z_{i1}}' & 0 & \dots & 0 \\ 0 & \underset{1\times 2}{z_{i2}}' & 0\dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \underset{1\times (T-2)}{z_{iT-1}}' \end{bmatrix} = \begin{bmatrix} y_{i0} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i0} & y_{i1} & & 0 & & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \dots & y_{i0} & \dots & y_{i(T-2)} \end{bmatrix}.$$

The constant variance of v_{it} given $\eta_i, y_{i0}, \dots, y_{i,t-1}$ implies that

$$E(Z_i' v_i^* v_i^{*\prime} Z_i) = \sigma^2 E(Z_i' Z_i). \tag{2.5}$$

Therefore, letting $y^*_{(-1)} = (y^{*\prime}_{1(-1)}, \dots y^{*\prime}_{N(-1)})'$ and $y^* = (y^{*\prime}_1, \dots y^{*\prime}_N)'$, an asymptotically efficient GMM estimator of α based on the moment conditions in (2.4) is given by

$$\widehat{\alpha}_{GMM} = \frac{y^{*'}_{(-1)} Z(Z'Z)^{-1} Z y^*}{y^{*'}_{(-1)} Z(Z'Z)^{-1} Z y^*_{(-1)}}$$

where $Z = (Z_1', \dots, Z_N')'$.

The LIML Estimator. The LIML analog estimator solves the following problem:

$$\widehat{\alpha}_{LIML} = \arg\min_{\alpha} \frac{(y^* - \alpha y^*_{(-1)})' Z(Z'Z)^{-1} Z'(y^* - \alpha y^*_{(-1)})}{(y^* - \alpha y^*_{(-1)})'(y^* - \alpha y^*_{(-1)})}. \tag{2.6}$$

In our Monte Carlo study, following Alvarez & Arellano (2003), we use a simple expression for $\widehat{\alpha}_{LIML}$. It can be noted that the minimized criterion in (2.6) is the minimum generalized characteristic root $\hat{\ell}$ of the polynomial equation:

$$\det[\ell(W^{*\prime}W^*) - W^{*\prime}Z(Z'Z)^{-1}Z'W^*] = 0 \tag{2.7}$$

where $W^* = (y^* : y^*_{(-1)})$. The first order conditions for (2.6) are

$$(1, -a)[W^{*\prime}Z(Z'Z)^{-1}Z'W^* - \hat{\ell}(W^{*\prime}W^*)] \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0 \tag{2.8}$$

from which we obtain

$$\widehat{\alpha}_{LIML} = \frac{y^{*\prime}_{(-1)} Z(Z'Z)^{-1} Z' y^* - \hat{\ell}(y^{*\prime}_{(-1)} y^*)}{y^{*\prime}_{(-1)} Z(Z'Z)^{-1} Z' y^*_{(-1)} - \hat{\ell}(y^{*\prime}_{(-1)} y^*_{(-1)})}. \tag{2.9}$$

The EL Estimator. Empirical Likelihood estimation (Qin & Lawless (1994) and Imbens (1997)) is a one-step method that achieves the same first-order asymptotic efficiency as robust GMM.

The empirical likelihood estimator maximizes a multinomial pseudo likelihood (or empirical likelihood) function subject to the orthogonality conditions. Letting p_i be the probability of observation i, the multinomial log likelihood of the data is given by the empirical likelihood estimator:

$$L = \sum_{i=1}^{N} \ln p_i.$$

The EL estimator maximizes this function subject to the restrictions

$$p_i \geq 0, \sum_{i=1}^{N} p_i = 1 \text{ and } \sum_{i=1}^{N} p_i Z_i'(y_i^* - \alpha y^*_{i(-1)}) = 0.$$

The Lagrangian is given by

$$\mathcal{L} = \sum_{i=1}^{N} p_i + \phi(1 - \sum_{i=1}^{N} p_i) - N\lambda' \sum_{i=1}^{N} p_i Z_i'(y_i^* - \alpha y^*_{i(-1)}),$$

where λ and ϕ are Lagrange multipliers. Taking the derivative of $\mathcal{L}$ with respect to p_i we obtain the following first-order conditions

$$\frac{1}{p_i} - \phi - N\lambda' Z_i'(y_i^* - \alpha y^*_{i(-1)}) = 0.$$

Multiplying by p_i and adding equations we get $\phi = N$. Hence,

$$p_i = \frac{1}{N}\left(\frac{1}{1 + \lambda' Z_i'(y_i^* - \alpha y^*_{i(-1)})}\right).$$

The multipliers of the moment conditions can be determined as implicit functions $\lambda(\alpha)$

solving (for a given value of α):

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{1+\lambda' Z_i'(y_i^* - \alpha y_{i(-1)}^*)}\right) Z_i'(y_i^* - \alpha y_{i(-1)}^*) = 0$$

such that $1 + \lambda' Z_i'(y_i^* - \alpha y_{i(-1)}^*) \geq 1/N$.

The concentrated likelihood function for α:

$$\mathcal{L}_c(\alpha) = \prod_{i=1}^{N} \frac{1}{N}\left(\frac{1}{1+\lambda' Z_i'(y_i^* - \alpha y_{i(-1)}^*)}\right).$$

Therefore, the EL estimator is given by

$$\widehat{\alpha}_{EL} = \arg\min_{\alpha} \sum_{i=1}^{N} \ln[1 + \lambda(\alpha)' Z_i'(y_i^* - \alpha y_{i(-1)}^*)].$$

A computationally useful alternative expression for α_{EL} is

$$\widehat{\alpha}_{EL} = \arg\min_{\alpha} \widehat{Q}(\alpha), \text{ where } \widehat{Q}(\alpha) = \max_{\lambda} \frac{1}{NT} \sum_{i=1}^{N} \ln[1 + \lambda' Z_i'(y_i^* - \alpha y_{i(-1)}^*)].$$

2.3 Monte Carlo Study and Results

In this section we report some Monte Carlo simulations of EL, GMM, and LIML estimators for various combinations of T and N values. Our focus is specifically on large T and moderated N values. As T gets large the number of orthogonality conditions also gets large (it grows at the rate of $T(T-1)/2$). Hence, the purpose of these experiments is to compare the biases of these estimators for different values of T and N when many moment conditions exist. As an extension we consider the case where the moment conditions are also weak besides being many. Weak moment conditions case occurs when the lagged levels of the series are only weakly correlated with subsequent orthogonal deviations, i.e., the series $[y_{i0}, \ldots, y_{i,t-1}]'$ is weakly correlated with $y_{i,t-1}^*$ for $t = 1, \ldots, T-1$. In our model, instruments available for the transformed equations become weak either as the autoregressive parameter (α) approaches unity or as the variance of the individual effects (η_i) increases relative to the variance of the transient shocks (v_{it}) (Bond, Hoeffler & Temple, 2001).

For all cases we conducted 1000 replications from the model specified in section 2 under normality, i.e. each sample consists of N independent observation of $(y_{i0}, y_{i1}, \ldots, y_{iT})$ generated from the process $y_{i0} = (1-\alpha)^{-1}\eta_i + (1-\alpha^2)^{-1/2}v_{i0}$, $y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it}$ for $t = 1, \ldots, T$ with $v_i = (v_{i0}, v_{i1}, \ldots, v_{iT})' \sim N(0, \sigma^2 I)$ and $\eta_i \sim N(0, \sigma_\eta^2)$ independent of v_i.

For the first set of results (the many strong instruments case), σ^2 and σ_η^2 are 1 and 0, respectively. We let α to take the values of 0.2, 0.5 and 0.8. This design follows Alvarez & Arellano (2003) closely. For the second set of results, we increase σ_η^2 from 0 to 1 and 4 while holding σ^2 at 1 and let α to take low (0.2) and high (0.9) values.

The first set of results are summarized in tables 2.1 and 2.2. In Table 2.1 we report median, interquartile range (iqr), and median absolute error (mae) of the EL, GMM, and LIML estimators for $N = 100$ with $T^0 = 10$, 25, and 50, where $T^0 = T + 1$ (the actual number of time series observations in the data). Table 2.2 reports the similar results for $N = 50$.

Table 2.1 and 2.2 reveal that in all cases the median bias of GMM estimator is always larger than the median biases of the EL and LIML estimators. The EL and LIML biases are both very small, however, the ranking between the two is not obvious. When T is small relative to N and α is small EL bias is smaller than the LIML bias. The difference between them gets smaller for a square panel with $T^0 = 50$ and $N = 50$.

When it comes to dispersion, GMM always has a smaller interquartile range than the other two estimators. Again, the ranking between EL and LIML is not clear. When $T^0 = 25$ LIML interquartile range is always smaller, however, for the other cases there is not an obvious order.

Finally, for median absolute errors, Table 2.1 and 2.2 show that except one case (the case when $T^0 = 50$, $N = 100$, and $\alpha = 0.2$) the GMM median absolute error is always the smallest. The ranking is less obvious between EL and LIML for this comparison criterion as well. In Table 2.1 for $T^0 = 10$ and 50 and $\alpha = 0.2$ and 0.8 EL median absolute error is smaller than LIML. In Table 2.2 it is smaller for $T^0 = 10$ and $\alpha = 0.2$, however as T gets larger and closer to N, LIML median absolute error becomes smaller, although the difference between them is very small especially when α is large.

For the weak instruments setup, we first let the autoregressive parameter α take a value close to unity while keeping the relative variance of individual effects (to the variance of idiosyncratic errors) low as in the previous cases, i.e., $\sigma^2 = 1$ and $\sigma_\eta^2 = 0$. The results for this case are reported in Table 2.3. To see the effect of increasing α on the performances of the estimators more conveniently we reproduce part of Table 2.1 here as well. Second, we let the relative variance of individual effects to increase, specifically $\sigma^2 = \sigma_\eta^2 = 1$, and α take low (0.2) and high (0.9) values. The results for this case are reported in Table 2.4. Third, we let the variance of individual effects increase to 4 while holding the variance of idiosyncratic errors still at 1 and α take low and high values. Table 2.5 reports the results for this case. For all these cases we let $N = 100$ and $T = 10$, 25, and 50[1].

Relative performances of EL, GMM and LIML estimators do not change when α is close to unity either. GMM is outperformed by EL and LIML estimators in terms of median, but outperforms in terms of interquartile range and median absolute error measures. Although, LIML performs slightly better than EL, the difference in terms of all three measures between them is very small.

Table 2.4 and 2.5 reveal interesting changes in the relative performances of the estimators. In Table 2.4, we give the results for the case when the variance of individual effects takes a positive value while the relative variance is not high. In terms of median, for $\alpha = 0.2$, LIML still performs better than the other two estimators. However, when the persistence in the series increases and the number of instruments gets large EL performs better than the rest. EL performs better than LIML also in terms of interquartile range and median absolute error when α approaches to unity or the number of instruments gets larger no matter if α is small or large.

The deterioration in the performance of LIML estimator due to high persistency becomes more obvious when the relative variance of individual effects gets larger (Table 2.5). When $\alpha = 0.9$ for all values of T EL dominates GMM and LIML estimators in terms of median. It is interesting to note that in this setting, when the number of instruments along with persistency increase the deterioration in the performance of LIML estimator in terms of all three measures becomes severe. On the other hand, the performance of EL estimator stays robust to the existence of weak instruments even in the extreme case.

Although EL and LIML both perform well in settings where instruments are strong, LIML performs slightly better than EL. However, when the over all performances are concerned in both strong and weak instrument cases, we conclude that EL is more reliable. Especially, when the instruments are many and weak LIML performs a lot worse as com-

[1]We run the same simulations for $N = 50$ as well, however, those results are not included here since the main conclusions did not change.

pared to EL and GMM. Panel data sets with these features are common (see Blundell & Bond (1998) for an example). Hence, these results suggest that when LIML is used in estimation of panel data models one should be cautious. If the data is long and persistent it would be better to use EL instead of LIML.

Table 2.1: Properties of *EL*, *GMM* & *LIML* Estimators: Many Instruments Case ($N = 100$)

	$\alpha = 0.2$			$\alpha = 0.5$			$\alpha = 0.8$		
	EL	GMM	LIML	EL	GMM	LIML	EL	GMM	LIML
				$T^0 = 10$					
median	0.1949	0.1830	0.1946	0.4908	0.4750	0.4912	0.7955	0.7801	0.8002
iqr	0.0589	0.0577	0.0595	0.0623	0.0575	0.0574	0.0571	0.0533	0.0566
mae	0.0294	0.0292	0.0298	0.0313	0.0285	0.0288	0.0279	0.0258	0.0281
				$T^0 = 25$					
median	0.1870	0.1825	0.1914	0.4841	0.4789	0.4902	0.7824	0.7756	0.7907
iqr	0.0334	0.0306	0.0309	0.0317	0.0295	0.0303	0.0272	0.0241	0.0250
mae	0.0166	0.0153	0.0155	0.0158	0.0149	0.0152	0.0136	0.0120	0.0126
				$T^0 = 50$					
median	0.1852	0.1824	0.1890	0.4830	0.4797	0.4881	0.7804	0.7771	0.7873
iqr	0.0197	0.0197	0.0208	0.0205	0.0193	0.0203	0.0147	0.0135	0.0149
mae	0.0097	0.0098	0.0104	0.0102	0.0098	0.0100	0.0074	0.0068	0.0075

iqr is the 75th-25th interquartile range; *mae* denotes the median absolute error; $\sigma^2 = 1$ and $\sigma_\eta^2 = 0$.

Table 2.2: Properties of *EL*, *GMM* & *LIML* Estimators: Many Instruments Case ($N = 50$)

	$\alpha = 0.2$			$\alpha = 0.5$			$\alpha = 0.8$		
	EL	GMM	LIML	EL	GMM	LIML	EL	GMM	LIML
					$T^0 = 10$				
median	0.1834	0.1625	0.1834	0.4790	0.4549	0.4843	0.7823	0.7510	0.7894
iqr	0.0837	0.0798	0.0858	0.0945	0.0817	0.0899	0.0855	0.0742	0.0793
mae	0.0418	0.0400	0.0429	0.0470	0.0413	0.0446	0.0423	0.0376	0.0399
					$T^0 = 25$				
median	0.1766	0.1666	0.1814	0.4714	0.4587	0.4778	0.7702	0.7578	0.7824
iqr	0.0492	0.0397	0.0428	0.0487	0.0401	0.0430	0.0390	0.0341	0.0370
mae	0.0245	0.0199	0.0215	0.0246	0.0201	0.0217	0.0195	0.0170	0.0186
					$T^0 = 50$				
Median	0.1689	0.1663	0.1690	0.4643	0.4603	0.4634	0.7658	0.7613	0.7667
iqr	0.0327	0.0276	0.0287	0.0319	0.0263	0.0284	0.0239	0.0194	0.0242
mae	0.0164	0.0138	0.0143	0.0160	0.0131	0.0142	0.0120	0.0097	0.0120

iqr is the 75th-25th interquartile range; *mae* denotes the median absolute error; $\sigma^2 = 1$ and $\sigma_\eta^2 = 0$.

Table 2.3: Properties of *EL*, *GMM* & *LIML* Estimators: Many Weak Instruments Case ($N = 100$; $\sigma^2 = 1$; $\sigma_\eta^2 = 0$)

	$\alpha = 0.2$			$\alpha = 0.9$		
	EL	GMM	LIML	EL	GMM	LIML
			$T^0 = 10$			
median	0.1949	0.1830	0.1946	0.8936	0.8787	0.8982
iqr	0.0589	0.0577	0.0595	0.0457	0.0433	0.0462
mae	0.0294	0.0292	0.0298	0.0232	0.0216	0.0229
			$T^0 = 25$			
median	0.1870	0.1825	0.1914	0.8892	0.8830	0.8961
iqr	0.0334	0.0306	0.0309	0.0191	0.0175	0.0185
mae	0.0166	0.0153	0.0155	0.0097	0.0088	0.0093
			$T^0 = 50$			
median	0.1852	0.1824	0.1890	0.8848	0.8825	0.8924
iqr	0.0197	0.0197	0.0208	0.0118	0.0105	0.0113
mae	0.0097	0.0098	0.0104	0.0058	0.0053	0.0056

iqr is the 75th-25th interquartile range; *mae* denotes the median absolute error.

Table 2.4: Properties of *EL*, *GMM* & *LIML* Estimators: Many Weak Instruments Case ($N = 100$; $\sigma^2 = 1$; $\sigma_\eta^2 = 1$)

	$\alpha = 0.2$			$\alpha = 0.9$		
	EL	GMM	LIML	EL	GMM	LIML
			$T^0 = 10$			
median	0.1924	0.1782	0.1929	0.8808	0.8197	0.8849
iqr	0.0759	0.0703	0.0728	0.1160	0.0961	0.1247
mae	0.0380	0.0351	0.0361	0.0587	0.0483	0.0624
			$T^0 = 25$			
median	0.1862	0.1799	0.1898	0.8601	0.8422	0.8662
iqr	0.0384	0.0349	0.0362	0.0533	0.0388	0.1020
mae	0.0192	0.0174	0.0181	0.0265	0.0188	0.0444
			$T^0 = 50$			
median	0.1838	0.1807	0.1877	0.8628	0.8574	0.8249
iqr	0.0217	0.0201	0.0221	0.0228	0.0214	0.1401
mae	0.0110	0.0101	0.0111	0.0114	0.0108	0.0595

iqr is the 75th-25th interquartile range; *mae* denotes the median absolute error.

Table 2.5: Properties of *EL*, *GMM* & *LIML* Estimators: Many Weak Instruments Case ($N = 100$; $\sigma^2 = 1$; $\sigma_\eta^2 = 4$)

	$\alpha = 0.2$			$\alpha = 0.9$		
	EL	GMM	LIML	EL	GMM	LIML
			$T^0 = 10$			
median	0.1916	0.1744	0.1923	0.8722	0.7692	0.8570
iqr	0.0805	0.0736	0.0782	0.1893	0.1413	0.3061
mae	0.0402	0.0368	0.0394	0.0931	0.0687	0.1439
			$T^0 = 25$			
median	0.1858	0.1794	0.1905	0.8433	0.8210	0.7606
iqr	0.0349	0.0344	0.0369	0.0728	0.0498	0.7120
mae	0.0174	0.0172	0.0183	0.0362	0.0248	0.1835
			$T^0 = 50$			
median	0.1825	0.1797	0.1862	0.8564	0.8501	0.5444
iqr	0.0222	0.0210	0.0222	0.0245	0.0198	0.8230
mae	0.0112	0.0105	0.0112	0.0120	0.0099	0.3104

iqr is the 75th-25th interquartile range; *mae* denotes the median absolute error.

Chapter 3

Economic and Political Determinants of Local Highway Spending in North Carolina: A Dynamic Panel Regression Analysis

Public spending, because it is influenced by elected officials, can be swayed by political considerations as well as socioeconomic factors. Previous studies have confirmed the importance of political measures in the allocation of general public spending as well in spending for infrastructure and highway projects. This study examines the determinants of the allocation of state highway funds in North Carolina to counties during the period 1990-2005. Determinants were divided between socioeconomic measures capturing population, job market, property, and road conditions, and political factors related to the party affiliation and county support of the governor and a power ranking of local legislators. Only socioeconomic measures were found to be related to the allocation of new road construction funds. However, both socioeconomic and political factors were statistically linked to the distribution of highway maintenance funds. Still, state decision-makers appear to primarily use highway funds to narrow income disparities between counties.

3.1 Introduction

Public choice economists and political scientists have long recognized that factors other than the common interest of the public influence the allocation of public funds. Important among these can be political considerations of elected officials who vote on aspects of public spending. This raises the possibility that public spending may not always be directed to the most deserving (by some objective measure) households or regions, but may be spent in such a way that enhances the objectives of politicians.

Indeed, there is a significant literature examining the possible impacts of political factors on public spending projects. One set of studies demonstrates the role of political ideology on the level of total government spending. Politicians and parties classified as "right of center" tend to be associated with lower spending levels, while those termed "left of center" can work toward higher spending amounts (Bilek, 2005; Cadot, Roller, & Stephan, 2006; Cruz, 2004). Other studies have found public spending decisions can be influenced by the power of interest groups lobbying elected officials (Gradstein, 2003). Another set of research focuses on explaining differences in federal public spending among U.S. states. This work has concluded that the political position of the state and of the state's congressional delegation can have independent effects on the level of federal spending. Characteristics like the tenure of the state's senior representative in Congress, the state's population per senator, the margin by which the sitting president won the state, and whether the majority of the state's congressional delegation is of the same party as the president have been shown to be associated with higher levels of federal spending in the state (Hoover & Pecorino, 2005; Boyle & Matheson, 2008).

Public spending on infrastructure (including highways), because of its size and dominance of public over private funding, has specifically been the subject of several studies analyzing political influence, and notable findings have been made. The political ideology of the state can matter in highway spending, but opposite of the findings for general government spending, here the conclusion is that"left-leaning" states spend less on highways and "right-leaning" states spend more (Witko & Newmark, 2003). Likewise, interest groups, such as the relative number of truckers in a state, have been associated with higher levels of state road spending (Congleton & Bennett, 1995). States without term-limited governors and with more stable party control of the state legislature have been linked to a lower stock of infrastructure (Crain & Oakley, 1995). In studies specifically looking at the distribution of federal road funds to the states during the New Deal, the political power of the state in Congress - in terms of the state's electoral votes and members of the state's congressional delegation being in leadership positions - has been shown to be positively related to the level of federal road spending (Anderson & Tollison, 1991). Similar results have been found for the allocation of federal highway demonstration grants (Gamkhar & Ali, 2008).

The general conclusion from these studies is that political factors can have an impact on the distribution of public spending - including public highway spending - but the impacts do not necessarily dominate other considerations. All the cited studies included non-political explanatory variables in their analyses which were also significantly related to the variation in public spending. Clearly, however, in seeking explanations for the variation in highway spending over both geography and time, it seems important to test for both socioeconomic and political factors together.

This study focuses on the allocation of public highway funds to counties in North Carolina. We focus specifically on North Carolina to study highway funding because the vast majority of revenue for roads flows to the state level, where the monies are then distributed by state level decision makers to local (county) areas. This system has led to long-time concerns about the degree of political influence in the spending allocation decisions. Our analysis differs from previous work in two respects. First, although there have been numerous studies of highway funding at the state level, to our knowledge, there is no analysis at the sub-state or county levels.

Second, we use panel data to analyze the effect of socioeconomic and political factors in the allocation of public highway funds. Most studies at this level of aggregation are based on cross-sectional data, which may potentially suffer from omitted variables. Apart from greater efficiency of parameter estimates, unobserved county heterogeneity and time effects that explain spatial differences in county highway spending can be explicitly taken into account. The use of panel data also provides flexibility in modeling the nature and the timing of socioeconomic and political variables impact on highway spending. In this analysis, therefore, we consider dynamic models, one of which also includes lagged value of highway spending as an explanatory variable. To estimate dynamic highway spending models, we use generalized method of moments (GMM) techniques that allow us to control for the endogeneity of lagged highway spending as well as that of socioeconomic variables, which could bias the estimated coefficients.

We first describe the econometric models and briefly discuss the estimation techniques applied to estimate the dynamic panel models in Section 3.2. The county-level data used for the analysis are outlined in Section 3.3, while estimation results from dynamic panel data models of highway spending are presented in Section 3.4. Finally, we conclude the paper in Section 3.5.

3.2 Empirical Models and Estimation Methods

Models

Our purpose in this research is to measure and compare the influences of non-political - that is, objective socioeconomic - factors and political considerations on highway spending among counties in North Carolina. Hence, the general form of the model to be estimated is

$$HS = f(SES, POL), \tag{3.1}$$

where HS is highway spending, SES is a vector of socioeconomic measures, and POL is a vector of political measures.

We consider two types of highway spending: Spending for new construction projects (HSN) and maintenance (HSM) projects, so equation (3.1) is estimated individually for each of these spending types to allow for potentially different impacts of the independent variables.

It is likely there a lag between the decisions about the allocation of highway funds among counties and the actual spending of those funds. Also, decision makers would consider movements in a county's socioeconomic characteristics as well as level values in funding allocations. Accordingly, for empirical implementations we let the general specification of model (3.1) to the highway spending on maintenance and the highway spending on construction models take the following form

$$HS_{it} = \alpha + \delta_1' \overset{\triangle}{SES}_{i,t-p} + \delta_2' SES_{i,t-p} + \beta_0' POL_{it} + \ldots + \beta_p' POL_{i,t-p} + \mu_i + \tau_t + \varepsilon_{it}, \tag{3.2}$$

where $\overset{\triangle}{SES}_{i,t-p} = SES_{it} - SES_{i,t-p}$, i.e., the change in economic variables over the period p, ε is an i.i.d. error term, and the first and the second components of the subscripts index counties and years respectively. The time-invariant county-specific component (μ) is included to account for unobserved or omitted heterogeneity across counties that does not very over time (e.g., climate, topography, and geographical location). The county-invariant time-specific component (τ) is used to capture any economic shocks that affect highway spending that are common to all counties but vary across time (e.g., fiscal policy of federal and state governments). HS is either HSM or HSN and when it is HSM p takes the value 2 and when it is HSN it takes the value 4. This specification of lag structure and lengths is explained in Section 3.4.

Besides model (3.2) we also consider an extension of this model:

$$HS_{it} = \alpha + \delta HS_{i,t-1} + \delta_1' \overset{\triangle}{SES}_{i,t-p} + \delta_2' SES_{i,t-p} + \beta_0' POL_{it} + \ldots + \beta_p' POL_{i,t-p} + \mu_i + \tau_t + \varepsilon_{it}. \tag{3.3}$$

Model (3.3) accounts for any potential persistence in the process of adjustment toward an equilibrium by including the lagged dependent variable among the explanatory variables of model (3.2). The parameter δ reflects this adjustment dynamics.

Dynamic Panel Model Estimation

The basic problem in estimation of model (3.3) is that the ordinary least squares (OLS) and other conventional panel regression techniques such as fixed effects (FE) and random effects (RE) estimators are generally biased and inconsistent. More specifically, due to the correlation between the lagged dependent variable and the error term OLS estimator becomes biased and inconsistent even if the ε are not serially correlated. For the FE estimator, the Within transformation wipes out the individual effects, but the transformed lagged dependent variable will still be correlated with the transformed errors even if the error term is serially uncorrelated. Although, the magnitude of this bias depends on the time dimension (see Nickell (1981) and Alvarez & Arellano (2003) for the nature of this bias), in finite samples this bias is shown to be severe. the same problem occurs with the random effects generalized least squares (GLS) estimator, which uses quasi-demeaning transformation (see Baltagi (1995) Chapter 2) to eliminate the individual effects. As a solution to this problem generalized method of moments (GMM) techniques are widely used in estimation of panel data model.

The standard GMM estimator, which is known as the "first difference" GMM estimator of Arellano & Bond (1991), takes the first difference of model (3.3) to eliminate the individual effects (μ_i) and uses all possible lags of HS_{it} in time period $t-2$ and earlier, i.e., $(HS_{i,t-2}, HS_{i,t-3}, \ldots, HS_{i1})$, as instruments for $(HS_{i,t-1} - HS_{i,t-2})$ that is correlated with the $MA(1)$ residuals $(\varepsilon_{it} - \varepsilon_{i,t-1})$ from the differencing. The same strategy is applied to form instruments for other explanatory variables that are allowed to be endogenous in the sense that they are correlated with contemporaneous and earlier shocks. This feature enables us to avoid simultaneity bias due to the endogeneity of economic variables. Since the current and past levels of highway spending could influence the current levels of economic variables, such as employment, population, property values, paved and unpaved highway milage, ... etc., $E(SES_{it}\varepsilon_{is}) \neq 0$ for $t \geq s$, the economic variables $\overset{\triangle}{SES}_{i,t-p}$ can be treated as endogenous as they include the current values. In this case the moment conditions $E(HS_{i,t-s}\Delta\varepsilon_{it}) = 0$ for $s \geq 2$ and $E(\overset{\triangle}{SES}_{i,t-p-s}\Delta\varepsilon_{it}) = 0$ for $s \geq 2$[1] can be used.

The first difference GMM estimator, however, may suffer from a weak instruments problem when the time series are highly persistent or the variance of the individual specific effects (μ_i) is relatively large compared to that of transitory shocks (ε_{it}) (Blundell & Bond, 1998; 2000). In such a case only weak correlation may exist between the lagged levels of the series and their subsequent first differences, implying that the available instruments used in the GMM estimator in first differences are less informative.

The "system GMM" estimator subsequently developed by Arellano & Bover (1995) and Blundell & Bond (1998) is preferred to the first-difference GMM estimator when estimating dynamic models with persistent panel data. Providing that $E(\Delta HS_{it}\mu_i) = 0$, the lagged first difference of highway spending $(\Delta HS_{i,t-1})$ is valid as an instrument for $HS_{i,t-1}$ in the levels equation (3.3), in addition to the instruments that are available after first-differencing. The analogous strategy can be applied to any explanatory variable treated as endogenous. Assuming that the economic variables SES_{it} are endogenous we have $\overset{\triangle}{SES}_{i,t-p}$ endogenous hence, we can use their lagged first differences $(\Delta\overset{\triangle}{SES}_{i,t-p-1})$ as instruments for $\overset{\triangle}{SES}_{i,t-p}$ in the levels equations, providing that $E(\Delta SES_{it}\mu_i) = 0$. This exploits additional moment conditions for equations in levels, $E(\Delta HS_{i,t-1}(\mu_i + \varepsilon_{it})) = 0$ and $E(\Delta\overset{\triangle}{SES}_{i,t-p-1}(\mu_i + \varepsilon_{it})) = 0$ for $t \geq 3$, that are combined with the moment conditions for equations in first differences exploited in the first-difference GMM estimator. The sys-

[1]We could also use $(SES_{i,t-2}, SES_{i,t-3}, \ldots, SES_{i1})$ as instruments for $\overset{\triangle}{SES}_{i,t-p}$ and have the moment conditions $E(SES_{i,t-s}\varepsilon_{it}) = 0$ for $s \geq 2$, however using both of the sets of moment conditions makes one of them redundant.

tem GMM thus estimates regressions in differences and regressions in levels simultaneously. Blundell & Bond (1998; 2000) show that the use of the system GMM estimator provides dramatic gains in efficiency and reduces the biases of the GMM estimator in first differences if the series is relatively short in time and highly autoregressive.

To assess the validity of instruments employed in GMM estimations, we consider three specification tests. First, we perform the Arellano-Bond test for serial correlation. The assumption of no serial correlation in the ε_{it} is essential for the validity of instruments. If the disturbances are not serially correlated, there should be evidence of significant negative first order serial correlation in the first-differenced residuals $(\hat{\varepsilon}_{it} - \hat{\varepsilon}_{i,t-1})$, and no evidence of second order serial correlation in the differenced residuals. Hence, Arellano-Bond test checks for these serial correlations in differenced residuals. Secondly, we examine the orthogonality conditions of the set of instruments using the Sargan test of overidentifying restrictions. Finally, the Difference Sargan test is employed to assess whether additional moment conditions in level equations are valid and can be used in the system GMM framework.

In addition to these specification tests, to evaluate the performance of GMM estimators we also report OLS and FE estimates and compare them with GMM estimates. Due to the presence of the lagged-dependent variable in a dynamic panel model, the OLS estimator tends to have an upward bias in the coefficient of the lagged-dependent variable, while the FE estimator produces a downward-biased estimate of the autoregressive coefficient (e.g., Arellano & Bond, 1991). Therefore, one may expect the OLS and FE estimates to form an approximate upper and lower bound, respectively, for the true parameter of the lagged-dependent variable.

In implementing both GMM estimators for dynamic panel models, to avoid "many instruments" problem we typically employ three lagged levels in time period $t-2, \ldots, t-7$ as instruments for GMM estimations in differenced equations, and one period lagged first differences of variables (e.g., $HS_{i,t-1} - HS_{i,t-2}$) for GMM in levels equations.

3.3 Data

Data for HS are annual spending amounts from the North Carolina Department of Transportation for each of North Carolina's 100 counties for the years 1990 to 2005[2].[3] The spending data were available separately for new construction projects (HSN) and maintenance (HSM) projects. A collection of socioeconomic variables are included in SES, among them population, employment, income, wealth (property value), road, and vehicle measures. The definitions of all variables as well as their means and standard deviations are presented in Appendix B (Table B.1). All dollar values are expressed in 2005 dollars. These measures are used to capture the range of "objective" factors public decision-makers might reasonably use in allocating highway funds.

Finding reasonable measures for POL is one of the most challenging parts of the analysis. Common measures of political power used in previous studies include tenure (years of service) of elected officials, membership of an elected official on an appropriations committee, and position of an elected official at a legislative leadership level.

This study uses four measures for POL, two relating to the governor and two relating to members of the General Assembly. One gubernatorial measure is simply the party affiliation of the governor (categorical variable GOVDEM, taking the value 1 for a Democrat governor and 0 for a Republican governor). This variable can test if there is systematic difference in highway spending across the state between governors of the two parties. The second gubernatorial measure is an interaction term (DEM*GOV) between the party affiliation of the governor and the party registration in the county. Specifically, the term multiplies the categorical variable for the governor (GOVDEM) by the ratio of registered Democrats to registered Republicans in the county (DEMS). A statistically significant positive parameter estimate on the term can be interpreted as road spending favoring counties with greater relative percentages of voters having the same party affiliation as the governor.

The political measures related to a county's representatives in the state General Assembly are based on a power ranking of those members. During each biennial legislative session the North Carolina Center for Public Policy Research polls legislators, lobbyists based in North Carolina, prominent legislative liaisons, and news correspondents covering

[2]North Carolina Department of Transportation, "Maintenance and Construction Spending by County," Raleigh, North Carolina, 2007.

[3]We thank Michael Walden for his suggestions on the choice of variables as well as the models.

the General Assembly on the effectiveness of each General Assembly member. Those polled base their ratings on several criteria, including the ability of the legislator to sway the opinion of colleagues, skill at guiding legislation through the chamber, respect commanded from peers, and knowledge and expertise in legislative matters. The power ratings are averaged for each member, and the average is converted to a "power ranking", with "1" being the highest in each of the two General Assembly chambers.

For each county for each year, two legislative political measures were created from the rankings. The first, HIGH, represents the highest political power ranking received by any member of the county's legislative delegation. Since there are different total members in the General Assembly's two chambers (50 in the Senate and 120 in the House), the power rankings in each chamber were first divided by the total membership, with the result subtracted from 1, so that higher readings on the remaining final value (a number between 0 and 1) corresponded to a higher power ranking.

HIGH then took the highest final value for members of the county's legislative delegation. The second measure was constructed by first averaging the final values (those measured between 0 and 1) separately for the county's Senate delegation and House delegation. Then, those two measures were averaged. The result, AVER, is the average power ranking for the county's legislative delegation, where political power between the two chambers is weighed equally. Because each legislative session is two years in length, HIGH and AVER each have identical values in two year periods.

Given this choice of explanatory variables, we are also concerned with potential endogeneity of socioeconomic variables that are presented in recent changes, which could influence estimation results. Any shock that affects the highway spending is likely to affect the present and future values of the socioeconomic variables as well. Accordingly, in addition to the dynamic highway spending model (3.3) assuming that all variables except the lagged highway spending are exogenous, we also consider an alternative specification in which the socioeconomic variables presented in recent changes are treated as endogenous using their lagged values as instruments in the GMM estimation.

3.4 Results

We start by presenting empirical results for model (3.2), i.e., the model that does not include any lagged dependent variable as an explanatory variable, for maintenance spending and construction spending in Table (3.1) and Table (3.2) respectively.

In determination of the lag structure, construction spending and maintenance spending were separately regressed on the explanatory socioeconomic and political variables using alternative lag lengths from one year to five years[4] The explanatory power (R^2) of the regressions with alternative lag lengths was compared and the lag length with the greatest explanatory power for each spending type was chosen to be used in the analysis. For construction, this was a lag length of four years, and for maintenance it was a lag length of two years. These lag lengths make sense, in that the pre-construction activities associated with a new road (design, required impact studies, land assembly) necessarily take longer than the planning involved with road maintenance projects. Because of the presence of serial correlation among the socioeconomic variables, only the level value of the variable corresponding to the lag length was used along with the change in the variable's value between the current year and the lag year. For the political variables, no serial correlation was present, so the level values for all the years in the lag period were included.

The first three columns report coefficients estimated by pooled OLS, fixed effects, and random effects estimators. The common result of all three methods is that economic variables have more significant effects than political variables in explaining both types of highway spending. This can be seen from the significance levels of the coefficients individually[5].

Among the socioeconomic variables, both change in population and recent changes in paved road milage are positively associated with maintenance spending. Although, population density and paved highway milage seem to have a positive impact on maintenance spending, the fixed effects coefficients for these variables turn out to be not significant. On the other hand, maintenance spending is lower with higher recent changes in population density, with employment, with recent changes in median income, and with property value.

[4] FE was used as the estimation technique for these regressions.

[5] Moreover, we also tested the significance of political variables jointly, the null hypothesis of no joint significance is always rejected. Whereas, the same hypothesis is always rejected for the socioeconomic variables.

Among the political variables, only the contemporaneous measure on HIGH, the highest political power ranking among the county's legislative delegation, has a positive and statistically significant effect on county road maintenance spending. The findings suggest that the difference between the county with the highest ranked power legislator and the county with the lowest ranked power legislator is around \$1 million of additional highway maintenance spending in the county[6].

The determinants of new road construction in counties are different. Employment and recent changes in number of registered vehicles are positively associated with new road construction, however, population density and recent changes in employment are negatively associated with new road construction. Although, population density and paved highway milage are important determinants of new road construction, there is not an agreement on the direction of their effects between these three estimators.

Nevertheless, these estimated results have relied on very restrictive assumptions that the disturbance in equation (3.2) has no serial correlation and all independent variables are strictly exogenous. Therefore, we relax these assumptions by considering two alternative fixed effects specifications of equation (3.2).

We first estimate the fixed effects model in which the error term exhibits first-order autocorrelation:

$$\varepsilon_{it} = \rho \varepsilon_{i,t-1} + \nu_{it}, \tag{3.4}$$

where ρ is the first-order coefficient of autocorrelation and μ_{it} is the stochastic disturbance term. Note that this approach introduces error dynamics and hence makes model (3.2) a special case of model (3.3) by partially accounting for lagged responses of highway spending.

Second, we address the potential endogeneity of change in socioeconomic variables using a two-stage least squares (2SLS) estimator with instrumental variables. An important concern when using this method is the choice of appropriate instruments. Based on diagnostics for instrument exogeneity and relevance, we find that lagged levels of change in socioeconomic variables in time period $t-2, \ldots, t-7$ are valid instruments for the current level of these variables. The J-test of overidentifying restrictions does not reject the null hypothesis that these instruments are exogenous. As for the relevance, following Baum, Heriot, & Stillman (2003), we tested this by examining the fit of the first stage regressions.

[6]Recall that HIGH takes a value of between 0 (lowest ranking) and 1 (highest ranking, so a one unit difference approximates the difference between the highest and lowest ranked legislators.

To be more precise, consider a regression equation with a set of endogenous variables $X1$, say, and a set of instruments $Z = Z1 + Z2$, where, $Z2$ is the subsets of instruments that are included and $Z1$ is the subsets of instruments that are not included in the regression equation. The first stage regressions are reduced form regressions of the endogenous variables $X1$ on the full set of instruments Z; the relevant test statistics here relate to the explanatory power of the excluded instruments $Z1$ in these regressions. A statistic commonly used, as recommended e.g., by Bound, Jaeger, & Baker (1995), is the R^2 of the first-stage regression with the included instruments "partialled-out". Alternatively, this may be expressed as the F-test of the joint significance of the $Z1$ instruments in the first-stage regression. However, for models with multiple endogenous variables, as in our case, these indicators may not be sufficiently informative. In these situations, Smith (1997) proposed a statistic: a "partial R^2"[7] measure that takes the inter-correlations among the instruments into account. For a model containing a single endogenous regressor, the two R^2 measures are equivalent. The distribution of Shea's partial R^2 statistic has not been derived, but it may be interpreted like any R^2. As a rule of thumb, if an estimated equation yields a large value of the standard (Bound, Jaeger, & Baker (1995)) partial R^2 and a small value of the Shea measure, one may conclude that the instruments lack sufficient relevance to explain all the endogenous regressors, and the model may be essentially unidentified. Following this advise, we compared the partial R^2 and the Shea measures obtained from the first-stage regression. For none of the endogenous socioeconomic variables these measures were considerable different from each other.

The empirical findings from these specifications are reported in columns 4 and 5. Overall, the results when taking into account potential autocorrelation in the disturbance and potential endogeneity are consistent with those obtained under stricter assumptions and consistent with each other. Also, when the potential autocorrelation is taken into account, the low degrees of the first-order autoregression, 0.265 for maintenance spending and 0.614 for construction spending, suggest that model (3.2) may not be seriously misspecified. According to both of the specifications, for maintenance spending, the lagged level of population density has a positive effect. The political variable HIGH still has a positive

[7]The Shea partial R^2 statistic may be easily computed according to the simplification presented in Godfrey (1999), who demonstrates that Shea's statistic for endogenous regressor i may be expressed as $R_p^2 = \frac{\nu_{i,i}^{OLS}}{\nu_{i,i}^{IV}} \left[\frac{(1-R_{IV}^2)}{(1-R_{OLS}^2)}\right]$ where $\nu_{i,i}$ is the estimated asymptotic variance of the coefficient.

and statistically significant effect on county road maintenance spending. The magnitude of its effect is even bigger when it is estimated in these settings. On the other hand, lagged employment, lagged property value, and recent changes in property value has a still negative but statistically more significant effect on maintenance spending. On the other hand, spending on new road construction is lower with larger recent changes in employment and higher with higher lagged paved highway milage.

Table 3.3 reports results from various specification of model (3.3), the model with lagged dependent variable included as an explanatory variable to account for dynamic responses of county maintenance spending to changes in socioeconomic and political variables. Table 3.4 reports the same for road construction spending. In columns 1 and 2, we report, respectively, OLS and FE estimates of the highway spending models in levels. Columns 3 and 5 represent coefficient estimates using the first-difference GMM and system GMM estimators, respectively, in which only the lagged highway spending variable is instrumented. The corresponding GMM estimates that also account for the potential endogeneity of socioeconomic variables, as discussed in Section 3.3, are reported in columns 4 and 6.

We first consider the GMM estimation results from columns 3 to 6 in Table 3.3. The coefficient for the lagged maintenance spending is positive and statistically significant at the 1 percent level in all specifications considered, indicating the presence of adjustment dynamics. This may indicate that model (3.3) may be a better specification in explaining maintenance spending. If that is so, the parameter estimates of model (3.2) are biased and inconsistent as they may partially pick up the effects of lagged maintenance spending. The coefficient for lagged maintenance spending obtained from first-difference GMM estimator is much smaller than the lower bound provided by the FE estimator. Despite controlling for the potential endogeneity of socioeconomic variables, the coefficient for the lagged maintenance spending is still lower than the corresponding FE estimate. These results raise concerns regarding the problem of weak instruments when using the standard GMM estimator in first differences with persistent series. Although Column 1 does not show much persistency in maintenance spending, estimating first-order autoregressive model for each of the other variables considered in this study using OLS and FE, we find that many, such as population, population density, employment, median income, and number of registered vehicles, are highly persistent. As such, the strong persistence of the panel data used in this analysis is likely to result in poor performance of the first-difference GMM estimator.

The system GMM estimator, which exploits additional moment conditions in the levels equations, appears more appropriate. We find evidence supports the work by Blundell & Bond (1998 and 2000), which suggests that due to weak instruments one can expect first-difference GMM estimates to be biased in the direction of the FE. Either under the strict exogeneity assumption on all explanatory variables (System GMM1) or considering the potential endogeneity of recent changes of socioeconomic variables (System GMM2), the system GMM estimates of the first-order autoregressive coefficient as shown in columns 5 and 6 are considerably higher than the difference GMM results (0.124 and 0.171) and lies in the interval between the OLS and FE estimates, which forms a lower and upper bound from 0.199 to 0.397. Moreover, tests for first-order and second-order serial correlation are comfortably passed, and the Difference Sargan tests do not reject the validity of instruments that are employed in the system GMM framework.

Table 3.4 reveals that similar results are also valid for the corresponding model for road construction. Evidence in favor of the system GMM estimation method is even more obvious in this model. Tests of first-order and second-order serial correlation and Sargan test is more comfortably passed.

Taking account the above results, it is advisable to use model (3.3) along with system GMM estimation in explaining highway spending. We first consider these results for maintenance spending, i.e., columns 5 and 6 of Table 3.3. The system GMM1 and the system GMM2 reveal parallel results. Among the socioeconomic variables recent changes in population, recent changes in unemployment rate, lagged number of registered vehicles, lagged paved and unpaved highway milage, and recent changes in paved highway milage have positive and significant effect on maintenance spending. On the other hand, maintenance spending is negatively affected by lagged population, recent changes in median income, lagged property value. Although, the political variables are jointly not significance, the second lag of variable DEM, registered Democrats to registered Republicans, and the first lag of $DEM * GOV$, Interaction of Democrat governor and ratio of registered Democrats to registered Republicans, seem to have a positive and significant effect on maintenance spending.

Next, we consider the system GMM1 and system GMM2 estimates of model (3.3) for road construction spending (columns 5 and 6 of Table 3.4). Higher levels of new road construction in the county are also associated with higher levels of lagged employment, with greater recent changes in median income, and with greater recent changes in property valuation. In contrast, spending on new road construction is lower with larger lagged population density, with greater recent changes in employment, and with larger levels of lagged paved highway milage. Political variables have an insignificant effect on new road construction spending as well. Only political variable HIGH has a significant effect with its current (negative) and first lag (positive) values.

3.5 Conclusion

The debate over the influence of political considerations on public spending is a long one. With this paper we contribute to the literature on the issue by estimating the influence of socioeconomic and political factors on the disbursement of state highway funds to county projects in North Carolina over the period 1990 and 2005. Based on the results obtained from several alternative modeling frameworks consistently suggest that political variables have almost no impact on the spending for either new road construction spending or maintenance spending.

Among the economic variables, the most consistent result for new road construction spending is the negative effect of change in employment. That is, new road construction spending is more likely to go to counties with slow growth in employment. This suggests that road construction is at least partially used as an economic development tool. On the other hand, the most consistent results for maintenance spending can be stated as follows: Higher levels of maintenance spending in the county are also associated with larger lagged value for the number of registered vehicles, with larger lagged values of paved and unpaved road mileage, and with greater recent changes in paved highway mileage, whereas, the higher levels of maintenance are associated with lower lagged values of property value and with lower recent changes in property value. The interpretation is that maintenance spending goes to counties with more road mileage and more vehicles. Also, road maintenance goes to low wealth counties. Again, the reason might be to provide jobs or to stimulate economic development.

Table 3.1: Estimation Results for Model (2) (Maintenance)

	(1)	(2)	(3)	(4)	(5)
Variables	Pooled OLS	Fixed Effects	Random Effects	Fixed Effects/AR(1)	Fixed Effects/2SLS
DEN_{t-2}	7.035***	12.724	9.240***	27.418**	40.567***
$DEN_t - DEN_{t-2}$	-87.833***	-144.836**	-136.573***	-121.524	88.892
POP_{t-2}	-0.046***	0.095***	-0.029**	0.041	-0.036
$POP_t - POP_{t-2}$	0.260***	0.505***	0.429***	0.464***	0.015
EMP_{t-2}	-0.021***	-0.071***	-0.009	-0.076***	-0.086***
$EMP_t - EMP_{t-2}$	0.050**	-0.021	0.030	-0.072**	-0.009
$UNEMP_{t-2}$	51.580	-49.382	-18.210	-67.479	-3.077
$UNEMP_t - UNEMP_{t-2}$	19.459	43.132	54.680	19.393	6.826
INC_{t-2}	-0.034***	-0.026	-0.041***	0.015	0.012
$INC_t - INC_{t-2}$	-0.048**	-0.055**	-0.067***	-0.040	-0.019
$PROP_{t-2}$	-0.00018***	-0.0004***	-0.00031***	-0.00041***	-0.00045**
$PROP_t - PROP_{t-2}$	-0.0001	-0.0002**	-0.00011*	-0.0002***	-0.00031***
VHC_{T-2}	0.111***	-0.036	0.082***	-0.013	0.109
$VHC_t - VHC_{t-2}$	-0.054*	-0.094***	-0.045	-45.809	-0.153***
PAV_{t-2}	5.926***	6.248	6.006***	12.848***	4.649
$PAV_t - PAV_{t-2}$	17.724***	10.197**	10.135**	11.323***	7.469
$UNPAV_{t-2}$	4.724***	4.895	6.169***	12.707**	6.895
$UNPAV_t - UNPAV_{t-2}$	2.710	0.027	0.433	10.648	-17.713

Continued

Table 3.1 – Continued

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Random Effects	(4) Fixed Effects/AR(1)	(5) Fixed Effects/2SLS
$HIGH_t$	1338.652**	1115.803*	1074.336*	1212.713**	169.7504**
$HIGH_{t-1}$	-807.138	-528.251	-604.666	-349.935	-1092.425
$HIGH_{t-2}$	-558.654	204.293	187.580	213.755	212.217
$AVER_t$	-15.926	-209.133	-177.884	-170.438	-624.938
$AVER_{t-1}$	576.413	364.611	429.912	333.820	731.462
$AVER_{t-2}$	99.256	4.769	-73.277	-19.756	-132.057
$DEMS_t$	134.605	-38.083	83.716	-166.219	-69.054
$DEMS_{t-1}$	387.087	254.561	253.741	399.828	283.088
$DEMS_{t-2}$	-481.148**	-168.408	-270.255	-187.920	-165.802
$DEM_t * GOV_t$	19.161	-89.291	-93.370	-52.372	-62.900
$DEM_{t-1} * GOV_{t-1}$	-65.242	-59.265	-67.002	-71.070	-68.537
$DEM_{t-2} * GOV_{t-2}$	18.746	6.941	38.352	14.870	42.822
Constant	1593.975**	939.158	1937.512**	-5961.166***	-2169.446
R^2	0.786	0.205	0.188	0.198	0.1859
Observations	1510	1510	1510	1410	1315
AR(1) coefficient	-	-	-	0.265	-

Notes: Coefficients for county-specific effects, year-specific effects as well as estimated t-statistics are omitted for brevity, but they are available upon request. *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Table 3.2: Estimation Results for Model (2) (Construction)

	(1)	(2)	(3)	(4)	(5)
Variables	Pooled OLS	Fixed Effects	Random Effects	Fixed Effects/AR(1)	Fixed Effects/2SLS
DEN_{t-4}	-57.413***	280.748***	-30.425*	-133.508	139.346
$DEN_t - DEN_{t-4}$	108.998	-486.867*	-275.979	-969.202**	177.028
POP_{t-4}	0.117**	-1.089***	-0.058	0.072	-0.770*
$POP_t - POP_{t-4}$	-0.345	0.046	0.060	1.364*	-1.928
EMP_{t-4}	0.374***	0.961***	0.391***	0.270	0.195
$EMP_t - EMP_{t-4}$	-0.736***	-0.630***	-0.728***	-1.051***	-1.030***
$UNEMP_{t-4}$	-57.534	-197.518	-297.226	-351.341	3957.404**
$UNEMP_t - UNEMP_{t-4}$	-224.981	-175.776	-291.965	-294.380	2676.556**
INC_{t-4}	0.164**	-0.285**	0.066	-0.277	1.283**
$INC_t - INC_{t-4}$	0.376***	-0.068	0.180*	-0.028	2.073**
$PROP_{t-4}$	0.002***	-0.00043	0.00163***	-0.00043	0.0059*
$PROP_t - PROP_{t-2}$	-0.0002	-0.0018***	-0.00039	-0.00003	-0.0015
VHC_{t-4}	-0.054	0.691***	0.119	0.378*	-0.449
$VHC_t - VHC_{t-4}$	0.299*	0.715***	0.395**	0.321*	0.949
PAV_{t-4}	-9.372***	215.513***	-6.849**	45.397**	297.314***
$PAV_t - PAV_{t-4}$	50.826***	28.882*	6.711	0.711	47.401
$UNPAV_{t-4}$	9.187	210.085***	-11.043	7.939	400.583***
$UNPAV_t - UNPAV_{t-4}$	3.774	-41.377	-71.469**	3.999	-48.394

Continued

Table 3.2 – Continued

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Random Effects	(4) Fixed Effects/AR(1)	(5) Fixed Effects/2SLS
$HIGH_t$	-4764.780	-1551.601	-3479.534	-3014.728	5242.896
$HIGH_{t-1}$	2012.167	2219.818	3022.763	3496.417	10500*
$HIGH_{t-2}$	-2064.270	1325.462	-1113.155	1794.203	-2483.020
$HIGH_{t-3}$	3825.279	2201.976	2334.662	1394.391	315.039
$HIGH_{t-4}$	-96.668	1858.621	1439.711	3869.524	-1350.143
$AVER_t$	3945.967	808.898	1897.689	3294.900	-2877.398
$AVER_{t-1}$	-4026.255	-4427.130	-4273.745	-1742.159	-12100*
$AVER_{t-2}$	6086.651	2582.483	3490.784	2182.088	5182.144
$AVER_{t-3}$	-1918.250	-1763.841	-692.465	1750.967	2230.994
$AVER_{t-4}$	-442.801	-1142.840	-1752.292	-3962.090	-3957.883
$DEMS_t$	3476.150*	2827.996	2495.808	1591.157	2955.979
$DEMS_{t-1}$	-1531.976	-902.352	-1176.601	-386.198	-2764.577
$DEMS_{t-2}$	-1906.339	-1053.413	-1556.408	-1852.366	-1115.401
$DEMS_{t-3}$	383.198	-255.289	191.198	161.582	2505.260
$DEMS_{t-4}$	-211.748	356.867	181.120	873.614	41.268
$DEM_t * GOV_t$	89.484	14.502	37.483	-97.500	-
$DEM_{t-1} * GOV_{t-1}$	-39.999	61.775	6.806	-18.049	283.450
$DEM_{t-2} * GOV_{t-2}$	449.489	371.433	378.353	292.789	743.928

Continued

Table 3.2 – Continued

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Random Effects	(4) Fixed Effects/AR(1)	(5) Fixed Effects/2SLS
$DEM_{t-3} * GOV_{t-3}$	143.041	270.405	163.712	192.035	-222.337
$DEM_{t-4} * GOV_{t-4}$	-276.173	-53.176	-249.822	-87.429	276.830
Constant	-3171.025	-168000***	3336.647	-15800***	-289000
R^2	0.802	0.412	0.365	0.163	0.205
Observations	1476	1476	1476	1376	1188
AR(1) coefficient	-	-	-	0.614	-

Notes: Coefficients for county-specific effects and estimated t-statistics are omitted for brevity, but they are available upon request. *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively. For the construction model the significance of time-specific effects is rejected at the 1 percent level, hence they are excluded in the regressions.

Table 3.3: Estimation Results for Model (3) (Maintenance)

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Diff GMM1	(4) Diff GMM2	(5) System GMM1	(6) System GMM2
$MAINT_{t-1}$	0.396***	0.199***	0.124***	0.171***	0.302***	0.367***
DEN_{t-2}	4.394**	14.819	23.806	14.548	4.892	2.641
$DEN_t - DEN_{t-2}$	-59.843*	-138.619*	-234.693*	-148.947	-60.827	-13.724
POP_{t-2}	-0.0317***	0.0655*	0.0482	0.0693	-0.0294**	-0.0322***
$POP_t - POP_{t-2}$	0.234***	0.464***	0.879***	0.550**	0.293***	0.166*
EMP_{t-2}	-0.0066	-0.0632***	-0.0736**	-0.0621**	-0.0054	-0.0035
$EMP_t - EMP_{t-2}$	0.00198	-0.0612**	-0.0909**	-0.0834**	-0.0204	-0.0235
$UNEMP_{t-2}$	-23.233	-57.963	237.145***	185.575***	63.947*	49.107
$UNEMP_t - UNEMP_{t-2}$	34.662	29.658	88.391	108.753**	84.254**	85.173**
INC_{t-2}	-0.03189***	-0.00132	0.23746***	0.06813*	-0.01399	-0.01794
$INC_t - INC_{t-2}$	-0.0460*	-0.0409	0.0973***	0.0117	-0.0413**	-0.0448**
$PROP_{t-2}$	-0.00016***	-0.00378***	-0.00043**	-0.00048***	-0.00022**	-0.0002**
$PROP_t - PROP_{t-2}$	-0.0001*	-0.00019***	-0.00028**	-0.00025**	-0.0002*	-0.00017
VHC_{T-2}	0.0716***	-0.0221	-0.0052	-0.0101	0.0751***	0.0807***
$VHC_t - VHC_{t-2}$	-0.0477*	-0.0663*	-0.0251	-0.0502	-0.0394	-0.0229
PAV_{t-2}	3.498***	5.833	17.919	6.816	3.683***	3.324***
$PAV_t - PAV_{t-2}$	4.262	9.324**	20.628***	18.599***	16.135***	14.937***
$UNPAV_{t-2}$	3.414***	4.718	45.168*	9.871	3.388**	3.529***

Continued

Table 3.3 – Continued

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Diff GMM1	(4) Diff GMM2	(5) System GMM1	(6) System GMM2
$UNPAV_t - UNPAV_{t-2}$	0.587	2.101	6.950	3.420	-0.120	6.695
$HIGH_t$	947.909*	1129.932*	601.810	1343.028*	1026.338*	1075.100*
$HIGH_{t-1}$	-804.693	-577.441	-440.370	-626.636	-551.172	-689.418
$HIGH_{t-2}$	410.300	320.697	-1640.995**	-665.839	-893.420*	-484.332
$AVER_t$	92.384	-150.270	485.946	338.601	289.972	450.307
$AVER_{t-1}$	404.453	321.659	331.297	272.520	222.041	99.196
$AVER_{t-2}$	-494.756	-179.991	972.645	406.686	344.455	49.583
$DEMS_t$	9.096	-150.892	479.040***	51.819	17.340	117.229
$DEMS_{t-1}$	388.698	363.852	8.841	320.730*	341.937*	289.055*
$DEMS_{t-2}$	-368.559*	-192.825	-321.603**	-299.099**	-361.953***	-396.466***
$DEM_t * GOV_t$	-33.532	-41.971	5.430	23.475	46.807	48.821*
$DEM_{t-1} * GOV_{t-1}$	-67.549	-57.452	-57.944*	-44.613	-54.530*	-83.957**
$DEM_{t-2} * GOV_{t-2}$	64.142	17.744	180.265***	35.883	15.478	47.737
Sargan	-	-	0.268	1.000	0.847	1.000
Diff Sargan	-	-	-	-	1.000	1.000
AR(1)	-	-	0.000	0.000	0.000	0.000
AR(2)	-	-	0.812	0.991	0.595	0.404
Observations	1414	1414	1298	1298	1414	1414

Continued

Table 3.3 – Continued

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Pooled OLS	Fixed Effects	Diff GMM1	Diff GMM2	System GMM1	System GMM2

Notes: Coefficients for county-specific effects, year-specific effects as well as estimated t-statistics are omitted for brevity, but they are available upon request. *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively. In columns 3 to 6, the results are from one-step GMM robust estimator, which are consistent with possible heterogeneity. We consider the restricted instrument set as described in Section 3.2. P-values are reported for AR(1), AR(2), Sargan, Diff Sargan. AR(1) and AR(2) are the Arellano and Bond test for first-order and second-order serial correlation. Sargan is a test of the overidentifying restrictions for the GMM estimations, Diff Sargan is a test of the additional moment conditions used in the system GMM estimations.

Table 3.4: Estimation Results for Model (3) (Construction)

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Diff GMM1	(4) Diff GMM2	(5) System GMM1	(6) System GMM2
$CONST_{t-1}$	0.702***	0.557***	0.211*	0.531	0.588***	0.693***
DEN_{t-4}	-31.526***	108.521**	-47.107	121.238	-41.837*	-32.220*
$DEN_t - DEN_{t-4}$	136.852	-361.798*	-604.326	-375.912***	151.403	144.614
POP_{t-4}	0.0598*	-0.4049**	0.6903	-0.4157***	0.0887*	0.0659
$POP_t - POP_{t-4}$	-0.1127	0.2921	1.7072**	0.3236	-0.0507	-0.1276
EMP_{t-4}	0.193***	0.459***	0.124	0.400	0.239***	0.194***
$EMP_t - EMP_{t-4}$	-0.517***	-0.630***	-1.120***	-0.663***	-0.575***	-0.527***
$UNEMP_{t-2}$	23.349	-222.208	-	-	-21.865	112.586
$UNEMP_t - UNEMP_{t-4}$	-107.682	-207.773	-	-	-	-
INC_{t-4}	0.06614	-0.05747	-0.48334*	-0.22945***	0.1028	0.0825
$INC_t - INC_{t-4}$	0.234***	0.120	-0.331*	-0.025***	0.171*	0.246**
$PROP_{t-4}$	-0.00042*	-0.00105*	-0.00127	-0.0013***	-0.0004	-0.00041
$PROP_t - PROP_{t-2}$	0.00061***	-0.00037	-0.00046	-0.00045***	0.00068*	0.00062*
VHC_{t-4}	0.0229	0.245	-0.347	0.367	0.0178	0.0196
$VHC_t - VHC_{t-4}$	0.102	0.366**	-0.123	0.430	-0.0561	0.110
PAV_{t-4}	-5.939***	126.605***	148.564	106.080	-7.389***	-6.575***
$PAV_t - PAV_{t-4}$	17.410	8.141	11.438	0.676**	20.288	21.085
$UNPAV_{t-4}$	12.943**	131.644***	151.461	91.196	15.697***	13.269***

Continued

Table 3.4 – Continued

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Diff GMM1	(4) Diff GMM2	(5) System GMM1	(6) System GMM2
$UNPAV_t - UNPAV_{t-4}$	35.172	-8.322	17.072	-51.935***	39.493	29.134
$HIGH_t$	-5413.511*	-2594.415	-1116.227	-2609.459***	-5849.860**	-6073.584*
$HIGH_{t-1}$	6185.701*	6024.805*	4237.930*	4540.506	6739.353*	7069.165*
$HIGH_{t-2}$	-3925.352	-2088.789	2124.720	-356.113***	-3776.644	-4412.261
$HIGH_{t-3}$	3609.903	3157.358	2683.148*	1862.166	3965.291	4604.541
$HIGH_{t-4}$	-715.774	436.669	2120.480	-1507.739***	-1929.326	-2276.778
$AVER_t$	4908.623*	2975.577	3267.238	2285.626	6091.017*	5810.708
$AVER_{t-1}$	-3996.638	-4278.976	-1272.390	-1165.967***	-4923.177	-4968.733
$AVER_{t-2}$	6110.857	4504.602	-	-	6159.579	6816.880
$AVER_{t-3}$	-2548.078	-2089.902	-	-	-3375.412	-4139.166
$AVER_{t-4}$	-1892.830	-1589.687	-	-	-	-
$DEMS_t$	2857.483*	2202.084	-413.264	66.532	2155.171***	-
$DEMS_{t-1}$	-2360.078	-1786.577	-	-	-1796.929*	2256.583***
$DEMS_{t-2}$	-1391.751	-1589.615	-	-	-333.370	-2077.020**
$DEMS_{t-3}$	1043.134	1117.522	-	-	-	-70.041
$DEMS_{t-4}$	-25.571	317.326	-	-	-	-40.192
$DEM_t * GOV_t$	-110.227	-112.989	-	-	14.157	53.571
$DEM_{t-1} * GOV_{t-1}$	-12.518	34.433	-	-	29.998	145.377

Continued

Table 3.4 – Continued

Variables	(1) Pooled OLS	(2) Fixed Effects	(3) Diff GMM1	(4) Diff GMM2	(5) System GMM1	(6) System GMM2
$DEM_{t-2} * GOV_{t-2}$	384.635	385.540	-	-	212.845*	-
$DEM_{t-3} * GOV_{t-3}$	12.945	86.941	-	-	-	-
$DEM_{t-4} * GOV_{t-4}$	-208.503	-219.798	-	-	-	-
Sargan	-	-	0.343	1	0.907	1
Diff Sargan	-	-	-	-	1	1
AR(1)	-	-	0.014	0.000	0.001	0.001
AR(2)	-	-	0.005	0.018	0.050	0.053
Observations	1380	1380	1270	1270	1380	1380

Notes: Coefficients for county-specific effects and estimated t-statistics are omitted for brevity, but they are available upon request. *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively. For the construction model the significance of time-specific effects is rejected at the 1 percent level, hence they are excluded in the regressions. In columns 3 to 6, the results are from one-step GMM robust estimator, which are consistent with possible heterogeneity. We consider the restricted instrument set as described in Section 3.2. P-values are reported for AR(1), AR(2), Sargan, Diff Sargan. AR(1) and AR(2) are the Arellano and Bond test for first-order and second-order serial correlation. Sargan is a test of the overidentifying restrictions for the GMM estimations, Diff Sargan is a test of the additional moment conditions used in the system GMM estimations.

Bibliography

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

Alonso-Borrego, C. & Arellano, M. (1999). Systematically normalized instrumental-variable estimation using panel data. *Journal of Business & Econometric Statistics*, *17*, 36–49.

Alvarez, J. & Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, *71*, 1121–1159.

Anderson, G & Tollison, R. (1991). Congressional influence and patterns of new deal spending. *Journal of Law and Economics*, *34*, 161–175.

Anderson, T. W. & Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, *18*, 47–82.

Arellano, M. (2003). *Panel Data Econometrics.* Oxford University Press, New York, NY, U.S.A.

Arellano, M., & Bond, S. R. (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., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*, 29–51.

Baltagi, B. H. (1995). Econometric analysis of panel data. John Wiley, Chichester.

Baum, C. F., Heriot, M. E., & Stillman, S. (2003). Instrumental variables and GMM: estimation and testing. Working paper No. 545, Boston College, USA.

Bekker, P. A. (1994). Alternative approximations to the distributions of instrumental variable estimators. *Econometrica*, *62*, 657–681.

Bilek, A. (2005). Local public choice of school spending: some original results from the French data. Working Paper, The University of Paris, Paris.

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

Blundell, R. & Bond, S. (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews*, *19*, 321–340.

Bond, S. R. (2002). Dynamic panel data models: a guide to micro data method and practice. *Portuguese Economic Journal* , *1*, 141–162.

Bond, S., Hoeffler, A., & Temple, J. (2001). GMM estimation of empirical growth models. CEPR Discussion Paper No. 3048.

Bound, J., Jaeger, D. A., & Baker, R. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association*, *90*, 443–450.

Bowsher, C. G. (2002). On testing overidentifying restrictions in dynamic panel data models. *Economics Letters*, *77*, 211–220.

Boyle, M. & Matheson, M. (2008). Determinants of the distribution of congressional earmarks across states. Faculty Research Series Paper No. 08-06, Dept. of Economics, College of the Holy Cross, Worcester, MA, USA.

Brown, B. & Newey, W. (2001). GMM, efficient bootstrapping, and improved inference. *Journal of Business & Econometric Statistics*, *20*, 507–517.

Cadot, O., Roller, L. H., & Stephan, A. (2006). Contribution to productivity or pork barrel: the two faces of infrastructure. *Journal of Public Economics*, *90*, 1133–1153.

Chamberlain, G. (1987). Asymptotic efficiecy in estimation with conditional moment restrictions. *Journal of Econometrics*, *34*, 305–334.

Congleton, R. & Bennett, R. (1995). On the political economy of state highway expenditures: some evidence of the relative performance of alternative public choice models. *Public Choice*, *84*, 1–24.

Crain, W.M. & Oakley, L. (1995). The politics of infrastructure. *Journal of Law and Economics*, *38*, 1–17.

Cruz, J. (2004). An empirical analysis of political influences in the public choice of portuguese municipalities. Paper presented at the annual meeting of the European Public Choice Society, Berlin, Germany.

Gamkhar, S. & Ali, H. (2008). The political economy of grant allocations: the case of federal highway demonstration grants. *The Journal of Federalism*, *38*, 1–21.

Godfrey, L. G. (1999). Instrument relevance in multivariate linear models. *Review of Economics & Statistics*, *81(3)*, 550–552.

Gonzalez, A. (2007). Empirical likelihood: improved inference within dynamic panel data models. ESE Discussion Papers 154, Edinburgh School of Economics, University of Edinburgh.

Gradstein, M. (2003). The political economy of public spending on education, inequality, and growth. Research Working Paper 3162, Washington: The World Bank.

Guggenberger, P. & Smith, R. J. (2005). Generalized empirical likelihood estimators and test under partial, weak, and strong identification. *Econometric Theory*, *21*, 667–709.

Hahn, J., Hausman, J., & Kuersteiner, G. (2001). Bias corrected instrumental variables estimation for dynamic panel models with fixed effects. Unpublished manuscript.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton, NJ, U.S.A.

Han, C. & Phillips, P C. B. (2006). GMM with many moment conditions. *Econometrica*, *74*, 147–192.

Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite sample properties of some alternative GMM estimators obtained from financial market data. *Journal of Business & Econometric Statistics*, *14*, 262–280.

Holtz-Eakin, D., Newey, W., & Rosen, H. (1988). Estimating vector autoregressions with panel data. *Econometrica*, *56*, 1371–1395.

Hoover, G. & Pecorino, P. (2005). The political determinants of federal expenditures at the state level. *Public Choice*, *123*, 95–113.

Imbens, G. W. (1997). One step estimators for over-identified generalized method of moments models. *Review of Economics Studies*, *64*, 359–383.

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

Kunitomo, N. (1980). Asymptotic expansions of the distributions of estimators in a linear functional relationship and simultaneous equations. *Journal of the American Statistical Association*, *75*, 693–700.

Morimune, K. (1983). Approximate distribution of k-class estimators when the degree of overidentifiability is large compared with the sample size. *Econometrica*, *51*, 821–841.

Nagar, A. L. (1959). The bias and moment matrix of the general k-class estimators of the parameters in simulation equations. *Econometrica*, *27*, 575–595.

Newey, W. & Smith, R. J. (2004). Higher-order properties of GMM and generalized empirical likelihood estiamtors. *Econometrica*, *72*, 219–255.

Newey, W. & Windmeijer, F. (2007). GMM with many weak moment conditions. Working paper, MIT.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, *49*, 1417–1426.

Oĝuzoĝlu, U. (2006). Empirical likelihood estimation of dynamic panel data models with fixed effects. Unpublished manuscript.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *Annals of Statistics*, *19*, 1725–1747.

Qin, J. & Lawless, J. (1994). Empirical likelihood and generalized estimating equations. *Annals of Statistics*, *22*, 300–325.

Sevestre, P. & Trognon, A. (1985). A note on autoregressive error components models. *Journal of Econometrics*, *28*, 231–245.

Shea, J. (1997). Instrument relevance in multivariate linear models: A simple measure. *Review of Economics & Statistics*, *79(2)*, 348–352.

Smith, R. J. (1997). Alternative semi-parametric likelihood approaches to generalized method of moments estimation *Economic Journal*, *107*, 503–519.

Witko, C. & Newmark, A. (2003). Economic competition and development spending in the States, 1964-2000. Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, USA.

Zilliak, J. P. (1997). Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators. *Journal of Business & Econometric Statistics*, *15*, 419–431.

Appendices

Appendix A

Through out the Appendix, we let C to denote a generic positive constant that may be different in different uses. Also, let $w.p.a.1$ stand for "with probability approaching one". To show the order of each term, we frequently employ results in Alvarez & Arellano (2003). AA refers to the paper of Alvarez & Arellano (2003); also AA (.) reads "the formula (.) in Alvarez & Arellano (2003)".

Lemma 1. $\|\frac{1}{N}\sum_{i=1}^{N} Z_i' v_i^*\| = O_p\left(\frac{T}{\sqrt{N}}\right)$.

Proof. By independence across individuals we have

$$\begin{aligned} E\left(\|\frac{1}{N}\sum_{i=1}^{N} Z_i' v_i^*\|^2\right) &= E\left(\frac{1}{N^2}\sum_{i=1}^{N} v_i^{*\prime} Z_i \sum_{i=1}^{N} Z_i' v_i^*\right) \\ &= \frac{1}{N^2}\sum_{i=1}^{N} E\left(v_i^{*\prime} Z_i Z_i' v_i^*\right) + \frac{2}{N^2}\sum_{i=1}^{N}\sum_{j>i} E\left(v_i^{*\prime} Z_i\right) E\left(Z_j' v_j^*\right) \\ &= \frac{\sigma^2}{N} tr\left\{E\left(Z_i' Z_i\right)\right\} \\ &= \frac{\sigma^2}{N}\sum_{t=1}^{T-1}\sum_{s=0}^{t-1} E\left(y_{is}^2\right). \end{aligned} \tag{A.1}$$

Recall

$$\begin{aligned} E\left(Z_i' Z_i\right) &= E\left\{\underbrace{\begin{bmatrix} z_{i1} & 0 & \dots & 0 \\ 0 & z_{i2} & 0\dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & z_{i,T-1} \end{bmatrix}}_{q\times(T-1)} \underbrace{\begin{bmatrix} z_{i1}' & 0 & \dots & 0 \\ 0 & z_{i2}' & 0\dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & z_{i,T-1}' \end{bmatrix}}_{(T-1)\times q}\right\} \\ &= E\begin{bmatrix} z_{i1}z_{i1}' & 0 & \dots & 0 \\ 0 & z_{i2}z_{i2}' & 0\dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & z_{i,T-1}z_{i,T-1}' \end{bmatrix}, \end{aligned}$$

where

$$z_{it}z_{it}' = \begin{bmatrix} y_{i0}^2 & \cdots & y_{i0}y_{i,T-1} \\ y_{i1}y_{i0} & \cdots & y_{i1}y_{i,T-1} \\ \vdots & \ddots & \vdots \\ y_{i,T-1}y_{i0} & \cdots & y_{i,T-1}^2 \end{bmatrix}.$$

Let $\omega_{it} = y_{it} - \frac{\eta_i}{(1-\alpha_0)}$. Then we have

$$\sum_{s=0}^{t-1} y_{is}^2 = \sum_{s=1}^{t} \omega_{i,s-1}^2 - 2t\bar{\bar{\omega}}_{i(-1)} \frac{\eta_i}{(1-\alpha_0)} + t\frac{\eta_i^2}{(1-\alpha_0)^2},$$

where $\bar{\bar{\omega}}_{i(-1)} = \frac{1}{t}\sum_{s=1}^{t} \omega_{i,s-1}$.

Under Assumptions 1–3, we have

$$E\left(\omega_{i,s-1}^2\right) = \frac{\sigma^2}{1-\alpha_0^2},$$

$$E\left(\bar{\bar{\omega}}_{i(-1)} \frac{\eta_i}{(1-\alpha_0)}\right) = 0,$$

$$E\left(\frac{\eta_i^2}{(1-\alpha_0)^2}\right) = \frac{\sigma_\eta^2}{(1-\alpha_0)^2}.$$

Therefore,

$$\sum_{s=0}^{t-1} E\left(y_{is}^2\right) = t\left(\frac{\sigma^2}{1-\alpha_0^2} + \frac{\sigma_\eta^2}{(1-\alpha_0)^2}\right),$$

and for $C = \sigma^2\left(\frac{\sigma^2}{1-\alpha_0^2} + \frac{\sigma_\eta^2}{(1-\alpha_0)^2}\right)$, (A.1) becomes

$$\frac{1}{N}\sum_{t=1}^{T-1} Ct = C\frac{(T-1)T}{2N}.$$

Hence the conclusion follows by Markov's Inequality. ■

Lemma 2.

$$E\left[\sup_{\alpha}\|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma\right] = O(T^{3\gamma-2}).$$

Proof. Consider

$$
\begin{aligned}
\|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma &= \left[\sum_{t=1}^{T-1}(y_{it}^* - \alpha x_{it}^*)^2 \sum_{s=1}^{t} y_{i,s-1}^2\right]^{\frac{\gamma}{2}} \\
&\leq (T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1}\left|y_{it}^* - \alpha x_{it}^*\right|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}} \\
&= (T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1}\left|(\alpha_0 - \alpha)x_{it}^* - v_{it}^*\right|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}.
\end{aligned}
$$

The inequality in the second line is obtained using Loève's c_r inequality. From AA (A43), we can write

$$
\begin{aligned}
x_{it}^* &= \psi_t\Big(y_{i,t-1} - \frac{\eta_i}{1-\alpha}\Big) - c_t\widetilde{v}_{itT}, \text{ where} \\
c_t &= \sqrt{\frac{T-t}{T-t+1}}, && \text{(A.2)} \\
\psi_t &= c_t(1 - \frac{\alpha\phi_{T-t}}{T-t}), && \text{(A.3)} \\
\phi_j &= \frac{1-\alpha^j}{1-\alpha}, \text{ and} && \text{(A.4)} \\
\widetilde{v}_{itT} &= \frac{1}{T-t}(\phi_{T-t}v_{it} + \ldots + \phi_1 v_{i,T-1}). && \text{(A.5)}
\end{aligned}
$$

Using this expression for x_{it}^*, we have, for $\beta = \alpha - \alpha_0$,

$$
\begin{aligned}
\|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma &= (T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1}\left|v_{it}^* + \beta c_t\widetilde{v}_{itT} - \beta\psi_t\big(y_{i,t-1} - \frac{\eta_i}{1-\alpha}\big)\right|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}} \\
&\leq C(T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1}\Big[\left|v_{it}^*\right|^\gamma + \left|\beta c_t\widetilde{v}_{itT}\right|^\gamma + \left|\beta\psi_t y_{i,t-1}\right|^\gamma \\
&+ \left|\beta\psi_t\frac{\eta_i}{1-\alpha}\right|^\gamma\Big]\Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}},
\end{aligned}
$$

where we used the fact that $\left|\sum_{i=1}^m a_i\right|^\gamma \leq c_\gamma \sum_{i=1}^m |a_i|^\gamma$ with c_γ being $m^{\gamma-1}$.

Taking the expectation of $\sup_\alpha \|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma$ we get

$$E\Big[\sup_\alpha \|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma\Big] \leq C(T-1)^{\frac{\gamma}{2}-1} \sum_{t=1}^{T-1} \Big\{ E\Big[\big|v_{it}^*\big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big] \tag{A.6}$$
$$+ \sup_\alpha \big|\beta c_t\big|^\gamma E\Big[\big|\widetilde{v}_{itT}\big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big]$$
$$+ \sup_\alpha \big|\beta \psi_t\big|^\gamma E\Big[\big|y_{i,t-1}\big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big]$$
$$+ \sup_\alpha \big|\beta \psi_t\big|^\gamma E\Big[\Big|\frac{\eta_i}{1-\alpha}\Big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big]\Big\}.$$

Now, we consider the first, second, and fourth elements on the right hand side. Note that v_{it}^* and $\widetilde{v}_{itT}$ includes the present and future legs of the error terms, i.e., $v_{it}, v_{i,t+1}, \ldots, v_{iT}$, hence by Assumption 1 we have

$$E\Big[\big|v_{it}^*\big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big] = E\big|v_{it}^*\big|^\gamma E\Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}},$$

$$E\Big[\big|\widetilde{v}_{itT}\big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big] = E\big|\widetilde{v}_{itT}\big|^\gamma E\Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}},$$

$$E\Big[\Big|\frac{\eta_i}{1-\alpha}\Big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big] = E\Big|\frac{\eta_i}{1-\alpha}\Big|^\gamma E\Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}.$$

Under Assumption 5 $E\big|v_{it}^*\big|^\gamma$, $E\big|\widetilde{v}_{itT}\big|^\gamma$, and $E\big|\eta_i\big|^\gamma$ are constant. Therefore, the order of magnitude of $E\Big[\sup_\alpha \|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma\Big]$ is determined by the third term in (A.6).

We have

$$E\Big[\big|y_{i(t-1)}\big|^\gamma \Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big] \leq E\Big[\big|y_{i,t-1}\big|^\gamma t^{\frac{\gamma}{2}-1} \sum_{s=1}^{t} \big|y_{i,s-1}\big|^\gamma\Big]$$
$$= t^{\frac{\gamma}{2}-1} \sum_{s=1}^{t} E\Big[\big|y_{i,t-1} y_{i,s-1}\big|^\gamma\Big], \tag{A.7}$$

where we have used Loève's c_r inequality.

Solving (1.1) recursively we obtain

$$y_{i,t-1} = \alpha^{t-1} y_{i0} + \frac{1-\alpha^{t-1}}{1-\alpha}\eta_i + \sum_{k=0}^{t-2}\alpha^k v_{i,t-1-k} \tag{A.8}$$

and using (A.8) along with Loève's c_r inequality in (A.7), and under Assumption 2 and Assumption 5 we get

$$\begin{aligned}
&t^{\frac{\gamma}{2}-1}\sum_{s=1}^{t} E\Big|\alpha^{s+t-2}y_{i0}^2 + \frac{1-\alpha^{t-1}}{1-\alpha}\alpha^{s-1}\eta_i y_{i0} + \sum_{k=0}^{t-2}\alpha^k\alpha^{s-1}v_{i,t-1-k}y_{i0} + \frac{1-\alpha^{s-1}}{1-\alpha}\alpha^{t-1}y_{i0}\eta_i \qquad \text{(A.9)}\\
&+\frac{1-\alpha^{s-1}}{1-\alpha}\frac{1-\alpha^{t-1}}{1-\alpha}\eta_i^2 + \sum_{k=0}^{t-2}\alpha^k\frac{1-\alpha^{s-1}}{1-\alpha}v_{i,t-1-k}\eta_i + \sum_{k=0}^{s-2}\alpha^k\alpha^{t-1}y_{i0}v_{i,s-1-k}\\
&+\sum_{k=0}^{s-2}\alpha^k\frac{1-\alpha^{t-1}}{1-\alpha}v_{i,s-1-k}\eta_i + \sum_{j=0}^{t-2}\sum_{k=0}^{s-2}\alpha^{k+j}v_{i,t-1-j}\eta_i v_{i,s-1-k}\Big|^{\gamma}\\
&\le Ct^{\frac{\gamma}{2}-1}\sum_{s=1}^{t}\alpha^{\gamma(s+t-2)}E\big|y_{i0}^2\big|^{\gamma} + \Big|\frac{1-\alpha^{t-1}}{1-\alpha}\alpha^{s-1}\Big|^{\gamma}E\big|\eta_i\big|^{\gamma}E\big|y_{i0}\big|^{\gamma}\\
&+\big|\alpha^{s-1}\big|^{\gamma}E\big|y_{i0}\big|^{\gamma}E\Big|\sum_{k=0}^{t-2}\alpha^k\alpha^{s-1}v_{i,t-1-k}\Big|^{\gamma} + \Big|\frac{1-\alpha^{s-1}}{1-\alpha}\alpha^{t-1}\Big|^{\gamma}E\big|y_{i0}\big|^{\gamma}E\big|\eta_i\big|^{\gamma}\\
&+\Big|\frac{1-\alpha^{s-1}}{1-\alpha}\frac{1-\alpha^{t-1}}{1-\alpha}\Big|^{\gamma}E\big|\eta_i^2\big|^{\gamma} + \Big|\frac{1-\alpha^{s-1}}{1-\alpha}\Big|^{\gamma}E\big|\eta_i\big|^{\gamma}E\Big|\sum_{k=0}^{t-2}\alpha^k v_{i,t-1-k}\Big|^{\gamma}\\
&+\big|\alpha^{t-1}\big|^{\gamma}E\big|y_{i0}\big|^{\gamma}E\Big|\sum_{k=0}^{s-2}\alpha^k v_{i,s-1-k}\Big|^{\gamma} + \Big|\frac{1-\alpha^{t-1}}{1-\alpha}\Big|^{\gamma}E\big|\eta_i\big|^{\gamma}E\Big|\sum_{k=0}^{s-2}\alpha^k v_{i,s-1-k}\Big|^{\gamma}\\
&+E\Big|\sum_{j=0}^{t-2}\sum_{k=0}^{s-2}\alpha^{k+j}v_{i(t-1-j)}\eta_i v_{i,s-1-k}\Big|^{\gamma}\\
&\le Ct^{\frac{\gamma}{2}-1}\Big\{\frac{1-|\alpha^{\gamma}|^t}{1-|\alpha^{\gamma}|} + \left|\frac{1-\alpha^{t-1}}{1-\alpha}\right|\frac{\left(1-|\alpha^{\gamma}|^t\right)}{\left(1-|\alpha^{\gamma}|\right)} + (t-1)^{\gamma-1}\frac{\left(1-|\alpha^{\gamma}|^t\right)\left(1-|\alpha^{\gamma}|^{t-1}\right)}{\left(1-|\alpha^{\gamma}|\right)^2}\\
&+|\alpha^{\gamma}|^{t-1}\frac{1}{|1-\alpha|^{\gamma}}\left(t+\frac{1-|\alpha^{\gamma}|^t}{1-|\alpha^{\gamma}|}\right) + \left|\frac{1-\alpha^{t-1}}{1-\alpha}\right|^{\gamma}\left|\frac{1}{1-\alpha}\right|^{\gamma}\left(t+\frac{1-|\alpha^{\gamma}|^t}{1-|\alpha^{\gamma}|}\right)\\
&+\frac{1}{|1-\alpha|^{\gamma}}\left(t+\frac{1-|\alpha^{\gamma}|^t}{1-|\alpha^{\gamma}|}\right)\frac{\left(1-|\alpha^{\gamma}|^t\right)}{\left(1-|\alpha^{\gamma}|\right)}(t-1)^{\gamma-1} + |\alpha|^{t-1}\frac{1}{\left(1-|\alpha^{\gamma}|\right)}\sum_{s=1}^{t}(s-1)^{\gamma-1}\\
&+\left|\frac{1-\alpha^{t-1}}{1-\alpha}\right|\frac{1}{\left(1-|\alpha^{\gamma}|\right)}\sum_{s=1}^{t}(s-1)^{\gamma-1} + (t-1)^{\gamma-1}\frac{1}{\left(1-|\alpha^{\gamma}|\right)^2}\sum_{s=1}^{t-1}s^{\gamma-1}.
\end{aligned}$$

Hence, to determine the order of magnitude of $E\Big[\sup_{\alpha}\|Z_i'(y_i^*-\alpha x_i^*)\|^\gamma\Big]$ we consider

$$
\begin{aligned}
&C(T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1}\sup_{\alpha}\left|\beta\psi_t\right|^\gamma E\Big[\left|\omega_{i,t-1}\right|^\gamma\Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big]\\
&\le C(T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1} E\Big[\left|\omega_{i,t-1}\right|^\gamma\Big(\sum_{s=1}^{t} y_{i,s-1}^2\Big)^{\frac{\gamma}{2}}\Big]\\
&\le C(T-1)^{\frac{\gamma}{2}-1}\sum_{t=1}^{T-1} Ct^{\frac{\gamma}{2}-1}\Big\{\frac{1}{1-|\alpha^\gamma|}+\frac{1}{|1-\alpha|\,(1-|\alpha^\gamma|)}+(t-1)^{\gamma-1}\frac{1}{(1-|\alpha^\gamma|)^2}\\
&+\frac{1}{|1-\alpha|^\gamma}\left(t+\frac{1}{1-|\alpha^\gamma|}\right)+\frac{1}{|1-\alpha|^{2\gamma}}\left(t+\frac{1}{1-|\alpha^\gamma|}\right)\\
&+\frac{1}{|1-\alpha|^\gamma}\left(t+\frac{1}{1-|\alpha^\gamma|}\right)\frac{1}{(1-|\alpha^\gamma|)}(t-1)^{\gamma-1}+\frac{1}{(1-|\alpha^\gamma|)}\sum_{s=1}^{t}(s-1)^{\gamma-1}\\
&+\frac{1}{|1-\alpha|\,(1-|\alpha^\gamma|)}\sum_{s=1}^{t}(s-1)^{\gamma-1}+(t-1)^{\gamma-1}\frac{1}{(1-|\alpha^\gamma|)^2}\sum_{s=1}^{t-1}s^{\gamma-1}\Big\}\\
&=O(T^{3\gamma-2}),
\end{aligned}
$$

where the first inequality uses the fact that $\left|\beta\psi_t\right|^\gamma$ is bounded in t and α and the second inequality uses the fact that $1-|\alpha^\gamma|^t$, $1-\alpha^{t-1}$ and $|\alpha|^{t-1}$ are bounded in t. For the last result we used the fact that $T^{-(v+1)}\sum_{t=1}^{T}t^v\to 1/(v+1)$ (cf. Hamilton (1994), Proposition 17.4 (h)). ■

Proof of Theorem 1. Combining the result in Lemma 2 with the first result in Appendix of Guggenberger & Smith (2005), we have

$$\max_{i\le N}\sup_{\alpha}\|Z_i'(y_i^*-\alpha x_i^*)\| = O_p(N^{1/\gamma}(E[\sup_{\alpha}\|Z_i'(y_i^*-\alpha x_i^*)\|^\gamma])^{1/\gamma}) = O_p(N^{1/\gamma}T^{3-2/\gamma}). \quad \text{(A.10)}$$

By the hypothesis of the theorem, namely $N^{1/\gamma}T^{3-2/\gamma}\sqrt{T^2/N}\to 0$ for $\gamma>2$, there exists τ_N such that $\frac{T}{\sqrt{N}}=o(\tau_N)$ and
$\tau_N=o(N^{-1/\gamma}T^{-3+2/\gamma})$. Let $L_N=\{\lambda:\|\lambda\|\le\tau_N\}$. Note that

$$\sup_{\lambda\in L_N,|\alpha|<1,i\le N}\left|\lambda' Z_i'(y_i^*-\alpha x_i^*)\right|\le\tau_N\max_{i\le N}\sup_{|\alpha|<1}\left\|Z_i'(y_i^*-\alpha x_i^*)\right\| = O_p(\tau_N N^{1/\gamma}T^{3-2/\gamma})\to 0.$$

Note that the multipliers of the moment conditions, λ, have to satisfy the condition
$\lambda' Z_i'(y_i^*-\alpha x_i^*)>-1$, for all $i=1,\ldots,N$. Let $\widehat{L}(\alpha)$ be the set of λs that satisfies this

condition, i.e. $\widehat{L}(\alpha) = \{\lambda : \lambda' Z_i'(y_i^* - \alpha x_i^*) > -1, i = 1, \ldots, N\}$. Therefore, there exists a C such that $w.p.a.1$, for all $|\alpha| \leq 1$, $\lambda \in L_N$, $i \leq N$

$$L_N \subset \widehat{L}(\alpha), \; -C \leq \frac{-1}{[1 + \lambda' Z_i'(y_i^* - \alpha x_i^*)]^2} \leq -C^{-1}, \; \left| \frac{1}{[1 + \lambda' Z_i'(y_i^* - \alpha x_i^*)]^3} \right| \leq C. \quad \text{(A.11)}$$

Let $\widehat{P}(\alpha, \lambda) = \frac{1}{NT} \sum_{i=1}^N \ln\left[1 + \lambda' Z_i'(y_i^* - \alpha x_i^*)\right]$. By a Taylor expansion around $\lambda = 0$ with Lagrange remainder for all $\lambda \in L_N$

$$\widehat{P}(\alpha, \lambda) = \lambda' \frac{1}{NT} \sum_{i=1}^N Z_i'(y_i^* - \alpha x_i^*) - \lambda' \frac{1}{T} \left[\frac{1}{N} \sum_{i=1}^N \frac{Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)' Z_i}{\left[1 + \overline{\lambda}' Z_i'(y_i^* - \alpha x_i^*)\right]^2} \right] \lambda$$

where $\bar{\lambda}$ lies between $\widehat{\lambda}$ and 0. By Assumption 4 (i), *LemmaA*0 of Newey & Windmeijer (2007), we have, $w.p.a.1$ $\lambda_{\min}\left(\frac{1}{N} \sum_{i=1}^N Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)' Z_i\right) \geq C^{-1}$ and $\lambda_{\max}\left(\frac{1}{N} \sum_{i=1}^N Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)' Z_i\right) \geq C$. Then, $w.p.a.1$ for all $|\alpha| < 1$ and $\lambda \in L_N$,

$$\lambda' \frac{1}{NT} \sum_{i=1}^N Z_i'(y_i^* - \alpha x_i^*) - C\frac{1}{T}\|\lambda\|^2 \leq \widehat{P}(\alpha, \lambda) \leq \lambda' \frac{1}{NT} \sum_{i=1}^N Z_i'(y_i^* - \alpha x_i^*) - C^{-1}\frac{1}{T}\|\lambda\|^2 \quad \text{(A.12)}$$

$$\leq \|\lambda\| \| \frac{1}{NT} \sum_{i=1}^N Z_i'(y_i^* - \alpha x_i^*)\| - C^{-1}\frac{1}{T}\|\lambda\|^2$$

Let $\widetilde{\lambda} = \arg\max_{\lambda \in L_N} \widehat{P}(\alpha_0, \lambda)$. By the right hand inequality in eq. (A.12),

$$0 = \widehat{P}(\alpha_0, 0) \leq \widehat{P}(\alpha_0, \widetilde{\lambda}) \leq \|\widetilde{\lambda}\| \| \frac{1}{NT} \sum_{i=1}^N Z_i' v_i^* \| - C^{-1}\frac{1}{T}\|\widetilde{\lambda}\|^2.$$

Subtracting $-C^{-1}\frac{1}{T}\|\widetilde{\lambda}\|^2$ from both sides and dividing through by $-C^{-1}\frac{1}{T}\|\widetilde{\lambda}\|$ and using the result of Lemma 1 gives

$$\|\widetilde{\lambda}\| \leq C \| \frac{1}{N} \sum_{i=1}^N Z_i' v_i^* \| = O_p(\frac{T}{\sqrt{N}}).$$

Now, following the same arguments in the proof of Lemma A3 of Newey & Windmeijer

(2007), it can be shown that $\|\widehat{\lambda}\| = O_p(\frac{T}{\sqrt{N}})$.

Now, expanding around $\lambda = 0$ (note that we let $\widehat{\alpha} \equiv \widehat{\alpha}_{EL}$ for simplification) to obtain

$$
\begin{aligned}
\widehat{Q}(\widehat{\alpha}) = \widehat{P}(\widehat{\alpha}, \widehat{\lambda}) &= \frac{1}{NT}\sum_{i=1}^{N} \ln 1 + \frac{1}{NT}\sum_{i=1}^{N} \frac{(y_i^* - \widehat{\alpha} x_i^*)' Z_i}{[1 + \lambda' Z_i'(y_i^* - \widehat{\alpha} x_i^*)]}|_{\lambda=0}\widehat{\lambda} \\
&- \frac{1}{2}\frac{1}{NT}\widehat{\lambda}' \left[\sum_{i=1}^{N} \frac{Z_i'(y_i^* - \widehat{\alpha} x_i^*)(y_i^* - \widehat{\alpha} x_i^*)' Z_i}{[1 + \lambda' Z_i'(y_i^* - \widehat{\alpha} x_i^*)]^2}|_{\lambda=0}\right]\widehat{\lambda} \\
&+ \frac{1}{3}\frac{1}{NT}\sum_{i=1}^{N} \frac{1}{\left[1 + \bar{\lambda}' Z_i'(y_i^* - \widehat{\alpha} x_i^*)\right]^3}\left[(y_i^* - \widehat{\alpha} x_i^*)' Z_i \hat{\lambda}\right]^3 \\
&= \frac{1}{NT}\sum_{i=1}^{N}(y_i^* - \widehat{\alpha} x_i^*)' Z_i \widehat{\lambda} - \frac{1}{2}\frac{1}{NT}\widehat{\lambda}' \sum_{i=1}^{N} Z_i'(y_i^* - \widehat{\alpha} x_i^*)(y_i^* - \widehat{\alpha} x_i^*)' Z_i \widehat{\lambda} \\
&+ \hat{r}
\end{aligned}
$$

where

$$
\hat{r} = \frac{1}{3}\frac{1}{NT}\sum_{i=1}^{N} \frac{1}{\left[1 + \bar{\lambda}' Z_i'(y_i^* - \widehat{\alpha} x_i^*)\right]^3}\left[(y_i^* - \widehat{\alpha} x_i^*)' Z_i \hat{\lambda}\right]^3
$$

and $\|\bar{\lambda}\| \leq \|\widehat{\lambda}\|$. We have, $w.p.a.1$

$$
\begin{aligned}
|\widehat{r}| &\leq \frac{1}{3}\|\widehat{\lambda}\| \max_{1\leq i\leq N} \sup_{\alpha} \|Z_i'(y_i^* - \widehat{\alpha} x_i^*)\| \frac{1}{T}\widehat{\lambda}' \left[\frac{1}{N}\sum_{i=1}^{N} \frac{Z_i'(y_i^* - \widehat{\alpha} x_i^*)(y_i^* - \widehat{\alpha} x_i^*)' Z_i}{\left[1 + \bar{\lambda}' Z_i'(y_i^* - \widehat{\alpha} x_i^*)\right]^3}\right]\widehat{\lambda} \\
&\leq C\|\widehat{\lambda}\| \max_{1\leq i\leq N} \sup_{\alpha} \|Z_i'(y_i^* - \widehat{\alpha} x_i^*)\| \frac{1}{T}\widehat{\lambda}' \left[\frac{1}{N}\sum_{i=1}^{N} Z_i'(y_i^* - \widehat{\alpha} x_i^*)(y_i^* - \widehat{\alpha} x_i^*)' Z_i\right]\widehat{\lambda} \\
&\leq O_p(\sqrt{T^2/N} N^{1/\gamma} T^{3-2/\gamma}) C \frac{1}{T}\|\widehat{\lambda}\|^2 = o_p(\frac{T}{N}).
\end{aligned}
$$

Also, $\widehat{\lambda}$ solves the equations:

$$
\frac{1}{NT}\sum_{i=1}^{N} \frac{1}{[1 + \lambda' Z_i'(y_i^* - \widehat{\alpha} x_i^*)]} Z_i'(y_i^* - \widehat{\alpha} x_i^*) = 0.
$$

Expanding around $\lambda = 0$:

$$
0 = \frac{1}{NT}\sum_{i=1}^{N} \frac{1}{[1 + \lambda' Z_i'(y_i^* - \widehat{\alpha} x_i^*)]} Z_i'(y_i^* - \widehat{\alpha} x_i^*) \,|_{\lambda=0}
$$

$$- \quad \frac{1}{NT}\sum_{i=1}^{N}\frac{1}{[1+\lambda' Z_i'(y_i^*-\widehat{\alpha}x_i^*)]^2}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\mid_{\lambda=0}\widehat{\lambda}$$
$$+ \quad \frac{1}{2}\frac{1}{NT}\sum_{i=1}^{N}\frac{1}{\left[1+\bar{\lambda}' Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]^3}\left[(y_i^*-\widehat{\alpha}x_i^*)'Z_i\hat{\lambda}\right]^2 Z_i'(y_i^*-\widehat{\alpha}x_i^*)$$

$$\Rightarrow 0 = \frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*) - \frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\widehat{\lambda}+\widehat{R}$$

where

$$\widehat{R} = \frac{1}{2}\frac{1}{NT}\sum_{i=1}^{N}\frac{1}{\left[1+\bar{\lambda}' Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]^3}\left[(y_i^*-\widehat{\alpha}x_i^*)'Z_i\hat{\lambda}\right]^2 Z_i'(y_i^*-\widehat{\alpha}x_i^*)$$

and $\bar{\lambda}$ lies in between 0 and $\widehat{\lambda}$. Note that $\max_{i\leq N}|\bar{\lambda}' Z_i'(y_i^*-\widehat{\alpha}x_i^*)| \leq \max_{i\leq N}|\widehat{\lambda}' Z_i'(y_i^*-\widehat{\alpha}x_i^*)| \leq \|\widehat{\lambda}\| \max_{1\leq i\leq N}\sup_{\alpha}\|Z_i'(y_i^*-\alpha x_i^*)\| \to 0$.

Then we have

$$\begin{aligned}
\|\widehat{R}\| &\leq C\max_{1\leq i\leq N}\Big(\sup_{\alpha}\|Z_i'(y_i^*-\alpha x_i^*)\|\frac{1}{\left|1+\bar{\lambda}' Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right|^3}\Big) \\
&\times \frac{1}{T}\widehat{\lambda}'\Big(\frac{1}{N}\sum_{i=1}^{N}Z_i'(y_i^*-\alpha x_i^*)(y_i^*-\alpha x_i^*)'Z_i\Big)\widehat{\lambda} \\
&\leq C\max_{1\leq i\leq N}\sup_{\alpha}\|Z_i'(y_i^*-\alpha x_i^*)\|\frac{1}{T}\|\widehat{\lambda}\|^2 = O_p(N^{1/\gamma}T^{3-2/\gamma}T/N) = o_p(\sqrt{T/N}).
\end{aligned}$$

Solving for $\widehat{\lambda}$:

$$\widehat{\lambda} = \left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1}\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)+\widehat{R}\right].$$

Plugging into $\widehat{Q}(\widehat{\alpha})$:

$$\begin{aligned}
\widehat{Q}(\widehat{\alpha}) =& \left[\frac{1}{NT}\sum_{i=1}^{N}(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1} \qquad (A.13)\\
&\times\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)+\widehat{R}\right] - \frac{1}{2}\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)+\widehat{R}\right]' \\
&\times\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1}\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)+\widehat{R}\right] + o_p(\frac{T}{N})
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{2}\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]'\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1}\\
&\times\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]-\frac{1}{2}\widehat{R}'\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1}\widehat{R}\\
&+o_p(\frac{T}{N})\\
\leq&\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]'\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1}\\
&\times\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]-\|\widehat{R}\|^2C+o_p(\frac{T}{N})\\
=&\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]'\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)(y_i^*-\widehat{\alpha}x_i^*)'Z_i\right]^{-1}\\
&\times\left[\frac{1}{NT}\sum_{i=1}^{N}Z_i'(y_i^*-\widehat{\alpha}x_i^*)\right]+o_p(\frac{T}{N}).
\end{aligned}
$$

The first term on the right side of (A.13) is the objective function of continuously updating GMM estimator (CUE). Hence the last conclusion shows that the difference of the CUE and EL objective functions converges uniformly to zero in α. The remainder of the proof then follows from the proof for of Theorem 3 (the consistency of the LIML estimator) of AA upon noting that in a linear model under homoskedasticity CUE is the LIML estimator as mentioned in Section 1.2.2. ■

Lemma 3. *If Assumption 4 (i) holds, Assumption 8 (ii) of Newey & Windmeijer (2007) also holds.*

Proof. In Newey & Windmeijer (2007) notation

$$
\begin{aligned}
&|a\big[\Omega^k(\widetilde{\alpha})-\Omega^k(\alpha)\big]b| = |a\big[E[Z_i'(y_i^*-\widetilde{\alpha}x_i^*)(-x_i^*)'Z_i]-E[Z_i'(y_i^*-\alpha x_i^*)(-x_i^*)'Z_i]\big]b|\\
&=|a\big[E[Z_i'W_i^*\begin{pmatrix}1\\-\widetilde{\alpha}\end{pmatrix}\begin{pmatrix}0 & -1\end{pmatrix}W_i^{*\prime}Z_i]-E[Z_i'W_i^*\begin{pmatrix}1\\-\alpha\end{pmatrix}\begin{pmatrix}0 & -1\end{pmatrix}W_i^{*\prime}Z_i]\big]b|\\
&=|a\big[E[Z_i'W_i^*\begin{pmatrix}0 & 1\\0 & \widetilde{\alpha}\end{pmatrix}-\begin{pmatrix}0 & 1\\0 & \alpha\end{pmatrix}W_i^{*\prime}Z_i]\big]b|
\end{aligned}
$$

$$= |tr\left\{\begin{pmatrix} 0 & 0 \\ 0 & \widetilde{\alpha}-\alpha \end{pmatrix} E\left[W_i^{*\prime} Z_i b a' Z_i' W_i^*\right]\right\}|$$
$$= |\widetilde{\alpha}-\alpha||E[(x_i^{*\prime} Z_i b a' Z_i' x_i^*)]|$$
$$= |\widetilde{\alpha}-\alpha||E[(a' Z_i' x_i^*)(b' Z_i' x_i^*)]|$$
$$\leq |\widetilde{\alpha}-\alpha| E[(a' Z_i' x_i^*)^2]^{1/2} E[(b' Z_i' x_i^*)^2]^{1/2}$$
$$\leq C|\widetilde{\alpha}-\alpha|\|a\|\|b\|,$$

where for the second to last inequality, we used Cauchy-Schwartz inequality and the last inequality is obtained by noting that

$$E[(a' Z_i' x_i^*)^2] = a' E[Z_i' x_i^* x_i^{*\prime} Z_i] a \leq \|a\|^2 \lambda_{\max}\{E[Z_i' x_i^* x_i^{*\prime} Z_i]\} \leq C\|a\|^2,$$

by Rayleigh quotient and Assumption 4.

The other parts of 8 (ii) follow by $\Omega^{k,l}(\alpha)$ and $\Omega^{kl}(\alpha)$ not depending on α. ■

Proof of Theorem 2. We proceed by verifying all the hypotheses of Theorem 3 of Newey & Windmeijer (2007). Note that $Z_i'(y_i^* - \alpha x_i^*)$ is twice continuously differentiable and that its first derivative does not depend on α, so Assumption 7 is satisfied. Also, for our case, Assumption 9 (i) of Newey & Windmeijer (2007) holds from the hypothesis of the theorem, namely $N^{1/\gamma} T^{3-2/\gamma} \sqrt{T^2/N} \to 0$ for $\gamma > 2$.

Now, we show that their Assumption 6 and Assumption 9 (ii) also hold. In Newey & Windmeijer (2007) notation, we have

$$\left(E[\|g_i\|^4] + E[\|G_i\|^4]\right)\frac{m}{n} = \left(E[\|Z_i' v_i^*\|^4] + E[\|Z_i' x_i^*\|^4]\right)\frac{T}{N}$$
$$\leq \left(CE[\|Z_i\|^4] + E[\|x_i^*\|^4 \|Z_i\|^4]\right)\frac{T}{N}.$$

The order of magnitude of the second term in the summation dominates that of the first term. Therefore, it is sufficient to show that $\left(E[\|Z_i' x_i^*\|^4]\right)\frac{T}{N} \to 0$.

Following the similar steps as in Lemma 2, we note that $E[\|Z_i' x_i^*\|^\gamma] = E[\|Z_i'(y_i^* - \alpha x_i^*)\|^\gamma] = O(T^{3\gamma-2})$. Hence, for $\gamma = 4$, we have

$$E[\|Z_i' x_i^*\|^4] = O(T^{10})$$

Hence, the first part of Assumption 6 and Assumption 9 (ii) of Newey & Windmeijer (2007) hold since $T^{11}/N \to 0$.

The second part of their Assumption 6 follows from Assumption 4 (i), and the rest holds by the model being linear in α.

The parts of Assumption 8 (i) of Newey & Windmeijer (2007) follow similarly upon noting that, for $W_i^* = (y_i^* : x_i^*)$ by Assumption 4 (i) we have

$$\sup_{\alpha} \| \frac{1}{N} \sum_{i=1}^{N} Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)' Z_i - E[Z_i'(y_i^* - \alpha x_i^*)(y_i^* - \alpha x_i^*)' Z_i] \|$$

$$= \sup_{\alpha} \| \begin{pmatrix} 1 & -\alpha \end{pmatrix} \left(\frac{1}{N} \sum_{i=1}^{N} W_i^{*\prime} Z_i Z_i' W_i^* - E[W_i^{*\prime} Z_i Z_i' W_i^*] \right) \begin{pmatrix} 1 \\ -\alpha \end{pmatrix} \| \to 0.$$

Hence, using Newey & Windmeijer (2007) notation, we have

$$\widehat{\Omega}^k(\alpha) - \Omega^k(\alpha) = \frac{1}{N} \sum_{i=1}^{N} Z_i'(y_i^* - \alpha x_i^*)(-x_i^*)' Z_i - E[Z_i'(y_i^* - \alpha x_i^*)(-x_i^*)' Z_i]$$

$$= \begin{pmatrix} 1 & -\alpha \end{pmatrix} \left(\frac{1}{N} \sum_{i=1}^{N} W_i^{*\prime} Z_i Z_i' W_i^* - E[W_i^{*\prime} Z_i Z_i' W_i^*] \right) \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

$$\widehat{\Omega}^{k,l}(\alpha) - \Omega^{k,l}(\alpha) = \frac{1}{N} \sum_{i=1}^{N} Z_i'(x_i^*)(x_i^*)' Z_i - E[Z_i'(x_i^*)(x_i^*)' Z_i]$$

$$= \begin{pmatrix} 0 & -1 \end{pmatrix} \left(\frac{1}{N} \sum_{i=1}^{N} W_i^{*\prime} Z_i Z_i' W_i^* - E[W_i^{*\prime} Z_i Z_i' W_i^*] \right) \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \text{and}$$

$$\widehat{\Omega}^{kl}(\alpha) = \Omega^{kl}(\alpha) = 0.$$

Finally, assumption 8 (ii) is satisfied by Lemma 3. ■

Appendix B

Table B.1: Description of Variables and Descriptive Statistics

			Standard Deviations		
Variables	Definition	Mean	Overall	Within	Between
MAINT	Annual highway spending on new construction in county, in 2005 (thousand) dollars	19200	31200	27300	15400
CONST	Annual highway spending on maintenance in county, in 2005 (thousand) dollars	7760	5079	4606	2186
POP	Population in county, in persons	76872	102387	101721	15264
DEN	Population density in county, in persons per square mile	150	182	25	181
EMP	Employment in county, in persons	34934	64056	63713	9053
UNEMP	Unemployment rate in county, in %	5.751	2.464	1.929	1.545
INC	Median family income in county, in 2005 dollars	47719	9785	8824	4316
PROP	Property valuation in county, in 2005 (million) dollars	5.390	9.502	9.177	2.619
VHC	Number of vehicles registered in county	58683	76174	75567	12070
PAV	Paved highway mileage in county	687.859	371.687	369.361	54.801
UNPAV	Unpaved highway mileage in county	94.967	71.570	60.459	38.746
GOVDEM	Categorical variable indicating governor is a Democrat	0.813	0.390	0	0.390
DEMS	Registered Democrats to registered Republicans ratio	3.082	3.221	2.960	1.302
DEM*GOV	Interaction of Democrat governor and ratio of registered Democrats to registered Republicans	2.282	2.708	2.124	1.691
HIGH	Highest political power ranking received by any member of the county's legislative delegation	0.731	0.245	0.170	0.177

Continued

Table B.1 – Continued

Variables	Definition	Mean	Standard Deviations		
			Overall	Within	Between
AVER	Average political power ranking of county's legislative delegation	0.522	0.224	0.163	0.154

Sources: North Carolina Department of Transportation, North Carolina State Data Center, North Carolina Center for Public Policy Research.