Foreign Direct Investment and Growth Symbiosis: A Semiparametric System of Simultaneous Equations Analysis

Nadine McCloud*
Department of Economics
University of the West Indies at Mona

Michael S. Delgado†
Department of Agricultural Economics
Purdue University

Subal C. Kumbhakar‡
Department of Economics
Binghamton University

February 19, 2017

Abstract

We characterize the types of interactions between foreign direct investment (FDI) and economic growth, and analyze the effect of institutional quality on such interactions. To do this analysis, we develop a class of instrument-based semiparametric system of simultaneous equations estimators for panel data and prove that our estimators are consistent and asymptotically normal. Our new methodological tool suggests that across developed and developing economies, causal, heterogeneous symbiosis and commensalism are the most dominant types of interactions between FDI and economic growth. Higher institutional quality facilitates, impedes or has no effect on the interactions between FDI and economic growth.

Keywords: Foreign direct investment; economic growth; institutional quality; parameter heterogeneity; semiparametric system of equations model; nonparametric method of moments; instrumental variables.

---

*Nadine McCloud, Department of Economics, University of the West Indies at Mona, Kingston 7, Jamaica. Phone: 876-970-6559, Fax: 876-977-1483, Email: nadine.mccloud02@uwimona.edu.jm.

†Corresponding author: Michael S. Delgado, Department of Agricultural Economics, Purdue University, West Lafayette, IN 47906-2056. Phone: 765-494-4211, Fax: 765-494-9176, Email: delgado2@purdue.edu.

‡Subal C. Kumbhakar, Department of Economics, Binghamton University, PO Box 6000, Binghamton, NY 13902-6000. Phone: 607-777-4762. Fax: 607-777-2681. Email: kkar@binghamton.edu.

Elizabeth Byrd and Shamar Stewart provided excellent research assistance. We thank participants at the 2013 China Meeting of the Econometric Society, the 2013 Asian Meeting of the Econometric Society, and the Department of Economics Research Seminar at the University of the West Indies at Mona for helpful comments. McCloud thanks the support of the Mona Research Fellowship from the Office of the Principal at University of the West Indies at Mona. This research was supported in part by computational resources provided by Information Technology at Purdue – Rosen Center for Advanced Computing, Purdue University, West Lafayette, Indiana. This paper has been presented under its previous title "Are There Feedbacks Between Foreign Direct Investment And Economic Growth? A Semiparametric System Of Simultaneous Equations Analysis With Instrumental Variables".

1 Introduction

In theory, FDI can provide an important source of capital, technology, and other productivity elements that are important for economic growth, but are otherwise lacking because of insufficient domestic investment (Borensztein, De Gregorio & Lee 1998). In theory, also, FDI can reduce or have no effect on economic growth. Consequently, a large strand of literature has been devoted to analyzing the nature of the effect of FDI on economic growth across countries.[1] However, knowledge of only the effect of FDI on growth may be insufficient to derive useful policy prescriptions on the relationship between these two variables. Useful policy prescriptions may be better derived from analyses on the (joint) interaction between FDI and growth; such general analyses lend themselves to answering questions such as whether countries experience (i) growth-FDI symbiosis – a positive effect of FDI on growth and a positive effect of growth on FDI, or (ii) FDI-commensalism – a positive effect of FDI on growth but no effect of growth on FDI. More important, the type of interaction between FDI and economic growth can shed light on the existence of direct multiplier benefits stemming from increases in either FDI or economic growth.

In this paper, we characterize the types of interactions between FDI and economic growth, and analyze the effect of institutional quality on such interactions. To do this analysis, we propose a semiparametric panel model of a system of simultaneous equations that allows FDI and economic growth to be modeled simultaneously and uses instrumental variables to identify causal effects. Our model is semiparametric – that is, we represent the coefficients on all regressors in all equations as unknown smooth functions of a measure of institutional quality, and unobserved country- and year-specific factors (fixed effects) – for a few reasons. One, we adopt the view that substantial linear and nonlinear forms of parameter heterogeneity that stem from, among other sources, cross-country differences in institutional quality, exist in empirical growth models (see, for e.g., Durlauf 2001, Minier 2007, Durlauf, Kourtellos & Tan 2008, Huynh & Jacho-Chávez 2009*a*, Huynh & Jacho-Chávez 2009*b*), and in particular in the effect of FDI on economic growth (see, for e.g., Borensztein et al. 1998, Alfaro et al. 2004, Durham 2004, Papaioannou 2009, Kottaridi & Stengos 2010, Delgado et al. 2014, McCloud & Kumbhakar 2012). Thus, we do not assume a priori that all countries use identical technologies to produce goods and services. In a recent review of the FDI literature, Alfaro & Johnson (2013) emphasize the importance of incorporating measures of institutional quality into empirical models of FDI and growth. Moreover, and although highly possible, the existence of parameter heterogeneity in empirical FDI models has not been considered in the literature. Two, unlike standard panel models, we abstract from the use of neutral ("proper" or additively separable) fixed effects and incorporate *non-neutral* fixed effects to reflect the presence of unobserved parameter heterogeneities that may influence the FDI and growth equations in many ways beyond a simple translation of each equation. For example, changes in FDI inflows within a firm may lead to changes in input composition of the production process and organizational structure, which are likely to be associated with changes in economic growth.[2]

[1] See, for example, Balasubramanyam, Salisu & Sapsford (1996), Borensztein et al. (1998), Alfaro, Chanda, Kalemli-Ozcan & Sayek (2004), Durham (2004), Carkovic & Levine (2005), Kottaridi & Stengos (2010), Delgado, McCloud & Kumbhakar (2014) and McCloud & Kumbhakar (2012), and the relevant references cited therein. See, also, Alfaro & Johnson (2013) for an excellent review.

[2] In essence, our empirical specification can be viewed as a mixture of the standard parametric system of equations model of FDI and growth used by Li & Liu (2005), and the semiparametric smooth varying coefficient growth model used by Delgado et al. (2014). Note that the system model of Li & Liu (2005) was fully parametric, and, more important, (i) assumed that, by virtue of parameter homogeneity, the interaction between economic growth is the same across countries and (ii) did not control for country- and time-specific effects. The model used by Delgado et al. (2014) was restricted to

Semiparametric system of equations estimation is still in its infancy (Welsh & Yee 2006, Henderson, Kumbhakar, Li & Parmeter 2015), and has not been used in any empirical economic studies of which we are aware. In the absence of a conditional distributional assumption on the response vector given the set of covariates, we opt to use the generalized method of moments (GMM) approach by Hansen (1982) to estimate our system of equations. The general unspecified form of our coefficient functions precludes estimation of our system with parametric GMM estimation. The estimators of the unknown coefficient functions, however, can be obtained using nonparametric GMM methods. In comparison to the literature on GMM and parametric systems of simultaneous equations, relatively little is known about using the GMM approach to estimate semiparametric systems of simultaneous equations and the asymptotic properties of the resultant estimators. To fill this gap in the literature, we therefore derive a broad class of local-linear GMM estimators – by coupling the GMM approach with local-linear estimation (see, e.g., Fan & Gijbels 1996) and the nonparametric system of equations symmetrical kernel-weighting approach in Welsh & Yee (2006) – for in-depth theoretical analysis. We establish the consistency and asymptotic normality of our class of system estimators. We propose a standard two-step estimation procedure that potentially yields more efficient estimates than a one-step systems estimator in the case that the errors across equations are indeed correlated. Our use of non-neutral fixed effects – in lieu of their neutral counterparts – circumvents the need to remove the fixed effects via some type of weighting or first difference transformation prior to estimation, to avoid biased and inconsistent estimates of, in particular, the marginal effects. Further, our use of generalized product kernels (Racine & Li 2004) allows us to avoid the incidental parameters problem associated with many parametric panel models that include dummy variables to account for unobserved effects.

Our theoretical framework can be seen as a generalization of both the *single* equation models proposed by Li, Huang, Li & Fu (2002), Das (2005), Cai, Das, Xiong & Wu (2006), Cai & Li (2008), Tran & Tsionas (2009), and Cai & Xiong (2012), as well as the *multivariate* response models of Welsh & Yee (2006) and Henderson et al. (2015). One important difference between our class of system estimators and those of Welsh & Yee (2006) and Henderson et al. (2015) is that we allow for correlation between any of the conditioning variables and the error term; hence, these other papers consider a system of reduced form equations, whereas we consider a system of structural equations. Such an important distinction renders our theoretical analysis a nontrivial extension of these aforementioned studies; we, however, show the numerical and asymptotic links between our estimators and some of these studies.[3] Our model and estimators can be used to empirically analyze a wide range of economic and non-economic phenomena. Moreover, the theoretical contributions of this paper are of independent interest and complement the relevant existing theoretical works.

Implementation of our new methodological tool and its application to a panel of 114 developed and developing countries over the period 1984 to 2010 yield that across developed and developing economies, causal, heterogeneous symbiosis and FDI-commensalism are the most dominant types of interactions between FDI and economic growth. The latter interaction suggests that in some countries there is no direct multiplier benefit between FDI and economic growth. Estimates of the smoothing parameters for our measure of institutional quality, and unobserved country- and year-specific factors

a single equation specification of the effects of FDI on growth rates. The model considered here is therefore substantially more general than either of those empirical specifications.

[3]We focus on local-linear estimators and derive a class of semiparametric system GMM estimators that includes instrumental variables estimators, seemingly unrelated regressions, and both semiparametric and nonparametric estimators. In contrast, Henderson et al. (2015) focus on local-constant estimators of a varying coefficient seemingly unrelated regression model, deriving both unrestricted and restricted estimators; in addition, they develop a consistent model specification test to accompany their estimators.

substantiate our claim that these three factors induce nonlinear forms of parameter heterogeneity in our simultaneous equation model, and are important in the growth-FDI nexus. In particular, we find empirical support for the use of non-neutral – rather than neutral – fixed effects.

We begin in Section 2 with a formal setup of our semiparametric system of simultaneous equations model – through which we will examine the FDI-growth nexus – and then derive our proposed class of semiparametric systems estimators. We present the large sample theory for the estimators in Section 3. We provide our empirical model and a description of our data, including the instrumental variables, in Section 4. We present our empirical results in Section 5, and Section 6 provides our conclusions. We provide detailed proofs for our large sample theory in the technical appendix to this paper.

2 Semiparametric System of Simultaneous Equations Estimation

To unveil empirical evidence on the types of interactions between economic growth and FDI, and the effect of institutional quality on such interactions, we put forward a very general semiparametric simultaneous system of equations model. We develop a novel class of semiparametric estimators suited for obtaining consistent estimates from different formulations of our semiparametric simultaneous system of equations model. From henceforth, we use the term *vector* to mean a *column* vector, unless otherwise stated.

To begin, consider in general form a J-variate semiparametric system of simultaneous equations

$$\begin{array}{rcl} y_{1,it} & = & Y'_{-1,it}\lambda_1(Z_{1,it}) + X'_{1,it}\gamma_1(Z_{1,it}) + \epsilon_{1,it} \\ \vdots & & \vdots \\ y_{J,it} & = & Y'_{-J,it}\lambda_J(Z_{J,it}) + X'_{J,it}\gamma_J(Z_{J,it}) + \epsilon_{J,it} \end{array} \tag{2.1}$$

where the j-th equation is

$$y_{j,it} = Y'_{-j,it}\lambda_j(Z_{j,it}) + X'_{j,it}\gamma_j(Z_{j,it}) + \epsilon_{j,it}, \tag{2.2}$$

for $j = 1, \ldots, J$, $i = 1, \ldots, N$, and $t = 1, \ldots, T$. In equation j for cross-sectional unit i in time period t, $y_{j,it}$ is a scalar response variable, $Y_{-j,it}$ is a p_j-dimensional vector of endogenous variables that includes at least one $y_{j_1,it}$ with $j_1 \neq j$; hence, the presence of $Y_{j,it}$ in each equation renders the system non-triangular. In addition, $X_{j,it}$ is a k_j-dimensional vector of exogenous variables in which the first entry is equal to 1, $Z_{j,it} \in \mathbb{R}^{d_j}$ is a vector of exogenous variables, $\lambda_j(\cdot)$ and $\gamma_j(\cdot)$ are unknown Borel measurable functions of conformable dimensions, and $\epsilon_{j,it}$ is the idiosyncratic error term.[4] Notice that, for the general derivation, we assume that the elements of $Z_{j,it}$ are continuously distributed; in practice, this assumption is easily relaxed to accommodate mixed categorical and continuous data using the important tools developed by Racine & Li (2004).

Our main interest is in the set of unknown coefficient functions $\{\lambda_j(\cdot)\}$, which clearly captures the types of interactions between the pairs y_j and y_{j_1} with $j_1 \neq j$. To characterize all interactions between any pair y_j and y_{j_1} with $j_1 \neq j$, we adopt the following taxonomy from the biological literature:

Definition 2.1. Let $l_j \in \{1, \ldots, J\}$ and $\lambda_j(\cdot) = \{\lambda_{j,l_j}(\cdot) : \mathbb{R}^{d_j} \to \mathbb{R},\ l_j \neq j\}$. Assume that for cross-sectional unit i in time period t the effect of y_j and y_{j_1} with $j_1 \neq j$ can be *positive*, *negative* or

[4] It is common in panel data models to define a one- or two-way error component specification for the idiosyncratic noise. We, however, follow the panel nonparametric GMM models of Cai & Li (2008) and Tran & Tsionas (2009) in our theoretical specification.

zero, and *vice versa*. Between the pair of variables $(y_{j,it}, y_{j_1,it})$ we say there exists:

(a) *symbiosis* if $\lambda_{j,j_1}(\cdot), \lambda_{j_1,j}(\cdot) > 0$;

(b) $y_{j_1,it}$-*commensalism* if $\lambda_{j,j_1}(\cdot) > 0$ and $\lambda_{j_1,j}(\cdot) = 0$;

(c) *synnercrosis* if $\lambda_{j,j_1}(\cdot), \lambda_{j_1,j}(\cdot) < 0$;

(d) $y_{j_1,it}$-*antagonistic symbiosis* if $\lambda_{j,j_1}(\cdot) > 0$ and $\lambda_{j_1,j}(\cdot) < 0$;

(e) $y_{j,it}$-*ammensalism* if $\lambda_{j,j_1}(\cdot) < 0$ and $\lambda_{j_1,j}(\cdot) = 0$;

(f) *non-symbiosis* if $\lambda_{j,j_1}(\cdot) = \lambda_{j_1,j}(\cdot) = 0$.

Remark 2.2. As mentioned in the preamble of this paper, plausible theoretical predictions are that the effect of FDI on economic growth *and* the effect of economic growth on FDI can be *positive*, *negative* or *zero*. Moreover, within a country there can be *symbiosis* between FDI and growth in one time period, but growth-commensalism in another time period as a result of, say, certain country-specific policies. Thus, this general taxonomy seems quite fitting for characterizing all possible theoretical interactions between FDI and growth. In other empirical applications, however, only a subset of this taxonomy may be applicable due to theoretical constraints of the signs of several $\{\lambda_j(\cdot)\}$ coefficient functions.

To proceed with estimation, we reformulate (2.2) as

$$y_{j,it} = \widetilde{X}'_{j,it} g_j(Z_{j,it}) + \epsilon_{j,it}, \tag{2.3}$$

where $\widetilde{X}'_{j,it} := (Y'_{-j,it}, X'_{j,it})$ and $g_j(Z_{j,it}) := (\lambda'_j(Z_{j,it}), \gamma'_j(Z_{j,it}))'$ and $m_j := p_j + k_j$.[5] Our estimators are motivated by the following conditional moment condition that we assume to hold throughout our various theoretical settings:

$$E[\epsilon_{it}|Z_{it}] = 0, \tag{2.4}$$

where $\epsilon_{it} = (\epsilon_{1,it}, \ldots, \epsilon_{J,it})'$, and $Z_{it} = (Z'_{1,it}, \ldots, Z'_{J,it})'$. In the absence of a distributional assumption on ϵ_{it}, we opt to use the generalized method of moments (GMM) approach to estimate our system of equations. The general unspecified form of our coefficient functions $g_j(\cdot)$ in (2.3) precludes estimation of our system with parametric GMM estimation. The estimators of $g_j(\cdot)$, however, can be obtained using nonparametric GMM methods. For our economic analysis, we are interested in estimating these unknown coefficient functions and their derivatives in all equations. Consequently, we first linearize the $g_j(\cdot)$ in each equation using local-linear approximation (Fan & Gijbels 1996); we apply this method to our system of equations as follows. We assume each g_j is sufficiently smooth and consider a first-order Taylor series expansion of $g_j(Z_{j,it})$ around a fixed point z_j in a neighborhood of $\{Z_{j,it}\}$, so that the s^{th} component of this expansion is

$$g^s_j(Z_{j,it}) \approx a^s_j + (b^s_j)'(Z_{j,it} - z_j), \; s = 1, \ldots, m_j, \tag{2.5}$$

where $b^s_j := \partial g^s_j(z_j)/\partial z_j$, a $d_j \times 1$ vector of first-order derivatives. Note that for the j-th equation, the remainder term of the second-order Taylor series expansion of the s-th component of $g_j(Z_{j,it})$,

[5] In general, $Z_{j,it}$ is required to be the same across $g_j(\cdot)$ for any j because of substantial econometric difficulties that arise in estimation of a semiparametric varying coefficient model in which the coefficient variables differ across coefficients.

$g_j^s(Z_{j,it})$, is

$$R_j^s(Z_{j,it}, z_j) = g_j^s(Z_{j,it}) - a_j^s - (b_j^s)'(Z_{j,it} - z_j) - \frac{1}{2}(Z_{j,it} - z_j)'\nabla^2 g_j^s(z_j)(Z_{j,it} - z_j), \tag{2.6}$$

and $R_j(Z_{j,it}, z_j) = (R_j^1(Z_{j,it}, z_j), R_j^2(Z_{j,it}, z_j), \ldots, R_j^{m_j}(Z_{j,it}, z_j))'$ is a m_j-dimensional vector. Define $\bar{R}_j^s(Z_{j,it}, z_j) := \frac{1}{2}(Z_{j,it} - z_j)'\nabla^2 g_j^s(z_j)(Z_{j,it} - z_j)$ to be the second order term in the expansion, and $\bar{R}_j(Z_{j,it}, z_j) = (\bar{R}_j^1(Z_{j,it}, z_j), \bar{R}_j^2(Z_{j,it}, z_j), \ldots, \bar{R}_j^{m_j}(Z_{j,it}, z_j))'$.

Combining (2.3) and the first-order approximation in (2.5) we obtain

$$y_{j,it} \approx U_{j,it}'\alpha_j + \epsilon_{j,it}, \tag{2.7}$$

where $U_{j,it} := \begin{pmatrix} \widetilde{X}_{j,it} \\ \widetilde{X}_{j,it} \otimes (Z_{j,it} - z_j) \end{pmatrix}$ is a vector of dimension $m_j(d_j + 1)$, $\otimes$ is the Kronecker product operator, and the corresponding coefficient vector is $\alpha_j := (a_j^1, \ldots, a_j^{m_j}, (b_j^1)', \ldots, (b_j^{m_j})')'$. Now stacking observations by T, then by N, and then by J gives the compact system formulation

$$y \approx U\alpha + \epsilon, \tag{2.8}$$

where $y = (y_1', \ldots, y_J')'$, $U = block\ diag(U_1, \ldots, U_J)$ so that for each j, U_j is a matrix of $NT \times m_j(d_j+1)$ observations on all right-hand side variables, $\alpha = (\alpha_1', \ldots, \alpha_J')'$, and $\epsilon = (\epsilon_1', \ldots, \epsilon_J')'$ with $\epsilon_j = (\epsilon_{j,11}, \ldots, \epsilon_{j,1T}, \ldots, \epsilon_{j,21}, \ldots, \epsilon_{j,2T}, \ldots, \epsilon_{j,N1}, \ldots, \epsilon_{j,NT})'$.

We now assume the existence of additional information in the form of instruments, W, to ensure the identification of the α parameter in the system in (2.8). For the population moment conditions, let $V_{j,it} := (W_{j,it}', Z_{j,it}')'$ and $V_{it} = (V_{1,it}', \ldots, V_{J,it}')'$. Thus,

$$E(\epsilon_{it}|V_{it}) = 0. \tag{2.9}$$

In light of our moment equality in (2.9), for any measurable function $Q(V_{it})$,

$$E(\epsilon_{it}|V_{it}) = 0 \iff E(Q(V_{it})\epsilon_{it}|V_{it}) = 0. \tag{2.10}$$

In essence, a plethora of conditional and unconditional moment equations can be generated from (2.10) using different specifications of $Q(V_{it})$. In the spirit of Cai & Li (2008), we choose for each equation j, $Q_{j,it} := Q(V_{j,it}) = \begin{pmatrix} W_{j,it} \\ W_{j,it} \otimes (Z_{j,it} - z_j)/h_j \end{pmatrix}$, which is a low-order polynomial vector of dimension $l_j(d_j+1)$ in $W_{j,it}$ and $Z_{j,it}$, l_j is the dimension of $W_{j,it}$, and $l_j \geq m_j$ for identification. In addition, the first entry of the vector $W_{j,it}$ is equal to one. Clearly, this simple form of $Q_{j,it}$ may not be the optimal form of the instruments for our model of interest. Newey (1990), for example, provides a mechanism for obtaining optimal instruments. However, deriving the functional form of optimal instruments for our model is beyond the scope of this paper.

We consider two types of local-linear GMM estimators for theoretical analysis. We draw on insights from the work of Welsh & Yee (2006) to guide us in constructing such estimators. Using a nonparametric exogenous system of seemingly unrelated regressions (SUR) – in which the regressor differs across equations – Welsh & Yee (2006) document that (i) consistency of the local-linear estimator via weighted least squares hinges on the position of the kernel weights in the unconditional moment equations, and (ii) under a homoscedasticity assumption, there is no gain in asymptotic efficiency from accounting for the correlation across errors – that is, there is no gain in smoothing jointly over smoothing marginally.

To yield a local-linear GMM estimator that is consistent, the nonparametric and multivariate nature of our model in (2.8) and our use of kernel smoothing therefore suggest careful consideration of our functional form for the system-based unconditional moment equation that is implied by (2.10). For ease of exposition, we let $Q = block\ diag(Q_1, \ldots, Q_J)$ so that for each j, Q_j is a matrix of $NT \times l_j(d_j+1)$ observations on the variables in $Q_{j,it}$. Also, let the system kernel matrix $K = block\ diag(K_1, \ldots, K_J)$ where $K_j = diag(K_{h_j}(Z_{j,11} - z_j), \ldots, K_{h_j}(Z_{j,NT} - z_j))$ with $K_{h_j}(\cdot) := h_j^{-d_j} K_j(\cdot/h_j)$, a kernel function in $\mathbb{R}^{d_j}$ for equation j. Define $\tilde{m}_j := m_j(d_j + 1)$, $\tilde{m} := \sum_{j=1}^{J} \tilde{m}_j$, and similarly, $\tilde{l}_j := l_j(d_j + 1)$, $\tilde{l} := \sum_{j=1}^{J} \tilde{l}_j$.

For our *first* local-linear GMM estimator, we assume that $A := Var(\epsilon\epsilon'|V)$ is a known $NT \times NT$ positive definite weighting matrix, and seek the nonparametric GMM system estimator $\widehat{\alpha}$ such that the following unconditional moment requirement is satisfied

$$Q'K^{1/2}A^{-1}K^{1/2}(y - U\alpha) = 0. \tag{2.11}$$

This moment condition corresponds to a local-linear GMM generalized least squares (GLS) estimator. In (2.11), we adopt the system of equations kernel weighting structure of Welsh & Yee (2006) that guarantees consistency of $\widehat{\alpha}$. Intuitively, the use and position of $K^{1/2}$ in (2.11) ensures that the cross-product of residuals, which are in the off-diagonal entries, are weighed symmetrically. This moment condition, however, represents an inconsistent system of $\widetilde{l}$ equations in $\widetilde{m}$ unknowns, which will not yield a unique estimator of α. We can premultiply (2.11) by a suitable scaling matrix to ensure we have a consistent system of equations for uniquely identifying the local-linear GMM-GLS estimator $\widehat{\alpha}$. In the spirit of Cai & Li (2008), we choose $U'K^{1/2}A^{-1}K^{1/2}Q$ as the $\widetilde{m} \times \widetilde{l}$ scaling matrix so that (2.11) becomes

$$U'K^{1/2}A^{-1}K^{1/2}Q \cdot Q'K^{1/2}A^{-1}K^{1/2}(y - U\alpha) = 0. \tag{2.12}$$

Then solving (2.12) gives

$$\widehat{\alpha} = [U'(K^{1/2}A^{-1}K^{1/2})QQ'(K^{1/2}A^{-1}K^{1/2})U]^{-1}[U'(K^{1/2}A^{-1}K^{1/2})QQ'(K^{1/2}A^{-1}K^{1/2})y]. \tag{2.13}$$

Remark 2.3. Note that $\widehat{\alpha}$, as defined by (2.13), does not take into account the variance-covariance moment matrix $Var(Q'K^{1/2}A^{-1}K^{1/2}\epsilon)$, and is therefore not the fully-efficient GMM estimator of α. Deriving a formula for the fully-efficient GMM estimator of α that is predicated on (2.11) will yield a very long expression that will provide no additional insights beyond what can be extracted from the asymptotic theory of $\widehat{\alpha}$.

Using a nonparametric exogenous vector measurement error model (a nonparametric system of SUR that has identical covariates across equations), Welsh & Yee (2006) also document that (i) the position of the kernel weights in the unconditional moment equations is immaterial for consistency of the local linear estimator via weighted least squares, and (ii) in some instances, and even under the homoscedasticity assumption, ignoring the correlations in errors across equations can result in a large loss in efficiency – that is, there can be gains in smoothing jointly over smoothing marginally. A semiparametric system of simultaneous equations in which the coefficient covariates are identical across equations is the model we use in our empirical application; moreover, and in light of the findings in Welsh & Yee (2006), a local-linear GMM estimator for such a model has different asymptotic properties from those of $\widehat{\alpha}$.

Thus, for our second local-linear GMM estimator, we consider another GMM-based local-linear system estimator but for the system model with $Z_{1,it} = Z_{2,it} = \cdots = Z_{J,it} = Z_{it}$. For this model, we also assume $h_1 = h_2 = \cdots = h_J = h$, and $K_1 = K_2 = \cdots = K_J = K$ to carry out local-linear estimation. Also, we assume that Γ^{-1} is a known $\tilde{l} \times \tilde{l}$ positive definite weighting matrix and seek the local-linear GMM system estimator $\widehat{\alpha}_{GMM}$ such that

$$\widehat{\alpha}_{GMM} = \arg\min_{\alpha} \, (y - U\alpha)' \widetilde{K} Q \Gamma^{-1} Q' \widetilde{K} (y - U\alpha), \tag{2.14}$$

where $\widetilde{K} = K \otimes I_J$, and Q and U are as previously defined but with Z_{it} in lieu of $Z_{j,it}, \forall\ j$. Then

$$\widehat{\alpha}_{GMM} = \Big[U' \widetilde{K} Q \Gamma^{-1} Q' \widetilde{K} U\Big]^{-1} \Big[U' \widetilde{K} Q \Gamma^{-1} Q' \widetilde{K} y\Big]. \tag{2.15}$$

3 Asymptotic Properties

To establish the asymptotic properties of our class of estimators, $\widehat{\alpha}$ and $\widehat{\alpha}_{GMM}$, we adopt the scaling approach in Cai & Li (2008) by defining $H := \mathit{block\ diag}(H_1, \ldots, H_J)$ so that for each j, $H_j := \mathit{diag}(I_{m_j}, h_j I_{d_j m_j})$ where I_{m_j} represents an identity matrix of size m_j. We develop and discuss the asymptotic properties of $\widehat{\alpha}$, and then state the corresponding properties for $\widehat{\alpha}_{GMM}$, in that order.

For ease of exposition, we define $\widehat{\alpha} := [S_n' S_n]^{-1} S_n' T_n$, where $n := NT$, and

$$S_n = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) U, \quad \text{and} \quad T_n = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) y.$$

We define $\widetilde{U}_{j,it} := H_j^{-1} U_{j,it} = \begin{pmatrix} \widetilde{X}_{j,it} \\ \widetilde{X}_{j,it} \otimes (Z_{j,it} - z_j)/h_j \end{pmatrix}$, so that

$$H\widehat{\alpha} = [\widetilde{S}_n' \widetilde{S}_n]^{-1} \widetilde{S}_n' T_n,$$

where

$$\widetilde{S}_n = S_n H^{-1} = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) U H^{-1} = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) \widetilde{U},$$

with $\widetilde{U} := UH^{-1}$. Then, we decompose T_n as follows:

$$T_n = \widetilde{S}_n H \alpha + T_n^* + B_n + R_n,$$

in which $T_n^* = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) \epsilon$, $B_n = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) \widetilde{X} \bar{R}$, $R_n = \frac{1}{n} Q' (K^{1/2} A^{-1} K^{1/2}) \widetilde{X} R$, and $\widetilde{X} := \mathit{block\ diag}\ (\widetilde{X}_1, \ldots, \widetilde{X}_J)$, $\widetilde{X}_j$ is a $n \times m_j$ matrix of observations on $\widetilde{X}_{j,it}$, and $\bar{R} := (\bar{R}_1', \ldots, \bar{R}_J')'$ and $R := (R_1', \ldots, R_J')'$ are vectors of dimension $\sum_{j=1}^J m_j$. We seek to establish the asymptotic properties of a properly normalized variant of $H(\widehat{\alpha} - \alpha)$, which we express as

$$H(\widehat{\alpha} - \alpha) - (\widetilde{S}_n' \widetilde{S}_n)^{-1} (\widetilde{S}_n' B_n) - (\widetilde{S}_n' \widetilde{S}_n)^{-1} (\widetilde{S}_n' R_n) = (\widetilde{S}_n' \widetilde{S}_n)^{-1} (\widetilde{S}_n' T_n^*). \tag{3.1}$$

On the left-hand side of (3.1), the second term will determine the asymptotic bias, whereas the third term will be shown to be asymptotically negligible. The term on the right-hand side of (3.1) will be shown to be asymptotically normal.

Without loss of generality, and in light of our empirical analysis of interest, from henceforth we restrict our theoretical developments to a *bivariate* semiparametric system of simultaneous equations. Some additional notations are in order. We define

$$\mu_{j,2}(K_j) := \int u_j u_j' K_j(u_j) du_j, \text{ and } \nu_{j,0} := \int K_j^2(u_j) du_j,$$
$$\Omega_j = \Omega_j(z_j) = E[W_{j,it}\widetilde{X}_{j,it}' | Z_{j,it} = z_j],$$
$$\Omega_j^* = \Omega_j^*(z_j) = Var[W_{j,it}\epsilon_{j,it} | Z_{j,it} = z_j],$$
$$\Omega_{jk}^{lm}(z_1, z_2) := E\{W_{j,it}W_{k,it}'\epsilon_{l,it}\epsilon_{m,it} | Z_{1,it} = z_1, Z_{2,it} = z_2\}, \text{ for } j,k,l,m = \{1,2\},$$
$$\Omega_{jk} = \Omega_{jk}(z_1, z_2) = E[W_{j,it}\widetilde{X}_{k,it}' | Z_{1,it} = z_1, Z_{2,it} = z_2], \text{ for } j \neq k = \{1,2\},$$
$$S_j = S_j(z_j) := \begin{pmatrix} \Omega_j & \mathbf{0} \\ \mathbf{0}' & \Omega_j \otimes \mu_{j,2}(K_j) \end{pmatrix}, \text{ and } S = block\ diag(S_1, S_2),$$
$$S_j^* = S_j^*(z_j) := \begin{pmatrix} \Omega_j^* \nu_{j,0} & \mathbf{0} \\ \mathbf{0}' & \Omega_j^* \otimes \mu_{j,2}(K_j^2) \end{pmatrix}, \text{ and } S^* = block\ diag(S_1^*, S_2^*),$$
$$B_j(z_j) = \int \begin{pmatrix} \Omega_j A_j(u_j, z_j) \\ \{\Omega_j A_j(u_j, z_j)\} \otimes u_j \end{pmatrix} K_j(u_j) du_j, \text{ and } A_j(u_j, z_j) = \begin{pmatrix} u_j' \nabla^2 g_j^1(z_j) u_j \\ \vdots \\ u_j' \nabla^2 g_j^{m_j}(z_j) u_j \end{pmatrix},$$

with $\nabla^2 g_j^s(z_j) := \partial g_j^s(z_j)/\partial z_j \partial z_j'$, and $B(z) = \big(B_1(z_1)', B_2(z_2)'\big)'$, and the dimension of $\mathbf{0}$, the *zero* matrix, differs according to context in which it is used. In addition, define

$$G_{1t}^{(jk,lm)}(z_1, z_2) := E\{W_{j,i1}W_{k,it}'\epsilon_{l,i1}\epsilon_{m,it} | Z_{i1} = z_1, Z_{it} = z_2\}.$$

The following assumptions are needed to establish the asymptotic properties of $\widehat{\alpha}$, our local-linear GMM-GLS estimator, in the case of large N and small T. Note that we will use the vector notation ϵ_{it} to mean $(\epsilon_{1,it}, \epsilon_{2,it})'$, and similar notations for W_{it}, X_{it}, Y_{it}, Z_{it}, etc.

Assumption A.1. (i) $\{(W_{it}, X_{it}, Y_{it}, Z_{it}, \epsilon_{it})\}$ are *i.i.d.* across the i index for each fixed t and strictly stationary over t for each fixed i, $E|\epsilon_{j,it}|^2 < \infty$, and $E\|W_{j,it}\tilde{X}_{k,it}'\|^2 < \infty$, $E\|W_{j,it}W_{k,it}'\|^2 < \infty$, where $\|A\|$ is the Frobenius norm for a finite-dimensional matrix A; this norm reduces to the usual Euclidean norm if A is a column vector.
(ii) The conditional variance of ϵ_{it} is Σ a bivariate positive definite matrix defined as

$$\Sigma(v) := Var(\epsilon_{it} | V_{it} = v) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

Assumption A.2. For each $t \geq 1$, $G_{1t}^{(jk,lm)}(z_1, z_2)$ and $f_{1t}(z_1, z_2)$, the joint density of Z_{i1} and Z_{it}, are continuous at (z_1, z_2). Also, for each z_j, $\Omega_j(z_j)$, $\Omega_{jk}(z_1, z_2)$, and $f_j(z_j)$ are bounded away from zero, where $f_j(z_j)$ is the marginal density function of $Z_{j,it}$. Further, $\sup_{t\geq 1} |G_{1t}^{(jk,lm)}(z_1, z_2) f_{1t}(z_1, z_2)| \leq M(z_1, z_2) < \infty$ for some arbitrary function $M(z_1, z_2)$. In addition, $g_j(z_j)$ and $f_j(z_j)$ are both twice continuously differentiable at $z_j \in \mathbb{R}^{d_j}$. The joint density of $Z_{it} = (Z_{1,it}', Z_{2,it}')'$ is $f(z) = f(z_1, z_2)$, and the partial derivatives $f^{(j)}(z) = \partial f(z)/\partial z_j$ and $f^{(j,k)}(z) = \partial^2 f(z)/\partial z_j \partial z_k'$ exist and are continuous.

Assumption A.3. The kernel functions $K_j(\cdot)$ are even, nonnegative, bounded density functions with compact support.

Assumption A.4. The instrumental variable W_{it} satisfies the conditions that $E(\epsilon_{it} | W_{it}, Z_{it}) = 0$ and $E[\pi(V_{it})\pi(V_{it})' | Z_{it} = z]$ is of full rank for all z, where $\pi(V_{it}) = E(\widetilde{X}_{it} | V_{it})$.

Assumption A.5. For $j = 1, 2$, (i) $h_j \to 0$, (ii) $Nh_j^{d_j} \to \infty$, and (iii) $Nh_j^{d_j} h_k^{d_k} \to \infty$ as $N \to \infty$, and $h_j^{d_j}/h_k^{d_k} \to 1$ for $j \neq k$.

Assumption A.6. There exists some arbitrary $\delta > 0$ such that $E\{|\epsilon_{j,it} W_{k,it}|^{(2+\delta)}|Z_j = u_j, Z_k = u_k\}$ is continuous at $u_j = z_j$ and $u_k = z_k$.

Remark 3.1. These assumptions contain some standard regularity conditions in the GMM and nonparametric literatures for panel data models with large N and small T. Moreover, these conditions represent generalizations to those in Welsh & Yee (2006) and Cai & Li (2008). Assumption A.1 extends the orthodox assumptions in the single-equation panel data models to a two-equation case. Note that $E|\epsilon_{j,it}\epsilon_{k,it}| < \infty$, an assumption that we omit, follows from $E|\epsilon_{j,it}|^2 < \infty$ and an application of the Cauchy-Schwarz inequality. Assumption A.2 provides bounds and smoothness conditions on the functionals in the proofs. The twice differentiability condition on the marginal distributions and functions is slightly stronger than warranted because it is possible to impose a Lipschitz condition on the first derivative of these marginals in lieu of the assumption of existence and continuity of the second derivative. The use of such general substitution, however, would lead to more cumbersome notations. Assumption A.3 renders $K_j(\cdot)$ a member of the class of second-order kernels. The nonnegativity and boundedness of $K_j(\cdot)$ are used several times in the proofs. Assumption A.4 is the identification condition. Assumption A.5 states that the each bandwidth is a null sequence of positive integers, and provides minimal conditions on the bandwidths to ensure consistency of the corresponding kernel estimators. Also, Assumption A.5 requires that the two bandwidths have the same order of magnitude. Assumption A.6 provides a Liapounov's condition, which we use in establishing asymptotic normality.

Define $\theta := \{\sigma_1^2(1-\rho^2)\}^{-1}$, $\beta := -\rho\{\sigma_1\sigma_2(1-\rho^2)\}^{-1}$, and $\gamma := \{\sigma_2^2(1-\rho^2)\}^{-1}$ with $|\rho| < 1$. By virtue of Assumption A.1, we can express the matrices $\Upsilon := \Sigma^{-1}$ and A^{-1} as follows:

$$\Upsilon = \begin{pmatrix} \theta & \beta \\ \beta & \gamma \end{pmatrix}, \quad A^{-1} = \Upsilon \otimes I_{NT} := \begin{pmatrix} D_\theta & \vdots & D_\beta \\ \dots & \dots & \dots \\ D_\beta & \vdots & D_\gamma \end{pmatrix}. \tag{3.2}$$

Finally, we define $D_{\theta\gamma} := block\ diag(\theta I_{\tilde{l}_1}, \gamma I_{\tilde{l}_2})$, $\tilde{D}_j := h_j^{d_j} I_{\tilde{m}_j}$, $\tilde{D} := block\ diag(\tilde{D}_1, \tilde{D}_2)$, and $\iota_{\tilde{m}_j}$ is an $\tilde{m}_j$-dimensional unit vector so that

$$\tilde{h}_{\tilde{m}}^2 := \big(h_1^2 \iota_{\tilde{m}_1}', h_2^2 \iota_{\tilde{m}_2}'\big)', \ \tilde{h}_{\tilde{l}}^2 := \big(h_1^2 \iota_{\tilde{l}_1}', h_2^2 \iota_{\tilde{l}_2}'\big)',$$
$$\tilde{I}_{\tilde{m}}^2 := block\ diag\big(h_1^2 I_{\tilde{m}_1}, h_2^2 I_{\tilde{m}_2}\big), \ \tilde{I}_{\tilde{l}}^2 := block\ diag\big(h_1^2 I_{\tilde{l}_1}, h_2^2 I_{\tilde{l}_2}\big)$$
$$\tilde{f}_{\tilde{m}}(z) := block\ diag\big(f_1(z_1) I_{\tilde{m}_1}, f_2(z_2) I_{\tilde{m}_2}\big), \ \tilde{f}_{\tilde{l}}(z) := block\ diag\big(f_1(z_1) I_{\tilde{l}_1}, f_2(z_2) I_{\tilde{l}_2}\big).$$

The following results establish consistency and asymptotic normality of $\widehat{\alpha}$.

Proposition 3.2. *If Assumptions A.1 to A.5 hold, then*

(i) $\widetilde{S}_n = \tilde{f}_{\tilde{l}}(z) D_{\theta\gamma} S\{1 + o_{\mathbb{P}}(1)\}$,

(ii) $B_n = \frac{1}{2}\tilde{I}_{\tilde{l}}^2 \tilde{f}_{\tilde{l}}(z) D_{\theta\gamma} B(z) + o_{\mathbb{P}}(\tilde{h}_{\tilde{l}}^2)$, *and*

(iii) $R_n = o_{\mathbb{P}}(\tilde{h}_{\tilde{l}}^2)$.

Proposition 3.3. *If Assumptions A.1 to A.5 hold, then*

$$n\tilde{\mathcal{D}}Var(T_n^*) = \tilde{f}_{\tilde{l}}(z)D_{\theta\gamma}^2 S^*, \tag{3.3}$$

where $\tilde{\mathcal{D}} :=$ *block diag*$(\tilde{\mathcal{D}}_1, \tilde{\mathcal{D}}_2)$ *with* $\tilde{\mathcal{D}}_j := h_j^{d_j} I_{\tilde{l}_j}$*, for* $j = 1, 2$*.*

Theorem 3.4. *(i) If Assumptions A.1 to A.5 hold, then*

$$H(\widehat{\alpha} - \alpha) - \frac{\tilde{I}_{\tilde{m}}^2}{2}B^*(z) = o_{\mathbb{P}}(\tilde{h}_{\tilde{m}}^2) + O_{\mathbb{P}}\big(n^{-1/2}\tilde{D}^{-1/2}\iota_{\tilde{m}}\big), \tag{3.4}$$

where $B^*(z) := \big(S'S\big)^{-1}\big(S'B(z)\big)$*.*

(ii) If Assumptions A.1 to A.6 hold, then

$$\big(n^{1/2}\tilde{D}^{1/2}\big)\Bigg[H(\widehat{\alpha} - \alpha) - \frac{\tilde{I}_{\tilde{m}}^2}{2}B^*(z) + o_{\mathbb{P}}(\tilde{h}_{\tilde{m}}^2)\Bigg] \xrightarrow{d} N\big(\mathbf{0}, \tilde{f}_{\tilde{m}}^{-1}(z)\Delta\big), \tag{3.5}$$

where $\Delta := (S'S)^{-1}S'S^*S(S'S)^{-1}$*.*

Remark 3.5. The block-diagonal nature of the matrices S and S^* implies that the estimators of the coefficient functions and their derivatives are asymptotically uncorrelated across equations. This result is equivalent to assuming the errors across equations are conditionally uncorrelated, that is $\rho = 0$. Thus, under specific conditions, $\widehat{\alpha}$ – the local-linear GMM-GLS estimator – does not yield any gain in smoothing jointly over smoothing marginally. Indeed, decomposing $B^*(z)$ and Δ reveals that

$$B^*(z) = \Big(B_{1,g}(z_2)'|\mathbf{0}'|B_{2,g}(z_2)'|\mathbf{0}'\Big)',$$

with

$$B_{j,g}(z_j) = \int A_j(u_j, z_j)K_j(u_j)du_j = \Big[tr\big(\nabla^2 g_j^s(z_j)\mu_{j,2}(K_j)\big)\Big]_{m_j\times 1},$$

and $\Delta :=$ *block diag*(Δ_1, Δ_2), and $\Delta_j =$ *diag*$\Big\{\nu_{j,0}\Omega_{j,g}, \Omega_{j,g}\otimes[\mu_{j,2}^{-1}(K_j)\mu_{j,2}(K_j^2)\mu_{j,2}^{-1}(K_j)]\Big\}$ with $\Omega_{j,g} = (\Omega_j'\Omega_j)^{-1}\Omega_j'\Omega_j^*\Omega_j(\Omega_j'\Omega_j)^{-1}$, which are identical to the asymptotic bias and variance terms from a bivariate variant of the main results of Cai & Li (2008). Note that these results also demonstrate the parallels between certain properties of GMM estimators for parametric and nonparametric system models.

As previously mentioned, the results in Theorem 3.4 are quite general and therefore nest the asymptotic properties of several other estimators including, for example, a few in Welsh & Yee (2006). We now demonstrate the links between our $\widehat{\alpha}$ and some estimator in the existing literature. The comparable estimators in Welsh & Yee (2006) are derived from a set of nonparametric SUR models. We first consider a semiparametric estimator of the Welsh & Yee (2006) econometric modeling framework that is predicated on the assumption of $E[\epsilon_{j,it}|\widetilde{X}_{j,it}'] = 0$. To begin, we suppose $W_{j,it} = \widetilde{X}_{j,it},\ \forall\ i, j, t$, and define $\widetilde{Q}_{j,it} := \begin{pmatrix} \widetilde{X}_{j,it} \\ \widetilde{X}_{j,it}\otimes(Z_{j,it} - z_j)/h_j \end{pmatrix}$. We define the system estimator in this case as $\tilde{\alpha}$, where

$$\tilde{\alpha} = [\widetilde{Q}'(K^{1/2}A^{-1}K^{1/2})U]^{-1}[\widetilde{Q}'(K^{1/2}A^{-1}K^{1/2})y]. \tag{3.6}$$

To derive the asymptotic properties, we simply put $W_{j,it}$ in lieu of $\widetilde{X}_{j,it}$ in all relevant expressions on page 9 and relabel, for example, in the following way:

$$\widetilde{\Omega}_j := E[\widetilde{X}_{j,it}\widetilde{X}'_{j,it}|Z_{j,it} = z_j],$$
$$\widetilde{\Omega}^*_j := Var[\widetilde{X}_{j,it}\epsilon_{j,it}|Z_{j,it} = z_j],$$
$$\widetilde{\Omega}^{lm}_{jk}(z_1, z_2) := E\{\widetilde{X}_{j,it}\widetilde{X}'_{k,it}\epsilon_{l,it}\epsilon_{m,it}|Z_{1,it} = z_1, Z_{2,it} = z_2\}, \text{ for } j, k, l, m = \{1, 2\},$$
$$\widetilde{\Omega}_{jk} := \Omega_{jk}(z_1, z_2) = E[\widetilde{X}_{j,it}\widetilde{X}'_{k,it}|Z_{1,it} = z_1, Z_{2,it} = z_2], \text{ for } j \neq k = \{1, 2\}.$$

We relabel $\widetilde{S}_j$, $\widetilde{B}_j(z_j)$, $\widetilde{B}(z)$, $\widetilde{S}$, $\widetilde{S}^*$, and $\widetilde{G}^{(jk,lm)}_{1t}$ in a similar manner. To this end, the results of Propositions 3.2 and 3.3 continue to hold but with $\widetilde{\widetilde{S}}_n$, $\widetilde{S}$, $\widetilde{B}_n$, $\widetilde{R}_n$, $\widetilde{T}^*_n$ and $\widetilde{S}^*$ respectively in lieu of $\widetilde{S}_n$, S, B_n, R_n, T^*_n and S^*. The following theorem establishes consistency and asymptotic normality of the GMM estimator $\tilde{\alpha}$ from the semiparametric system of SUR model.

Theorem 3.6. *Suppose* $E[\epsilon_{j,it}|\widetilde{X}'_{j,it}] = 0$.

(i) If Assumptions A.1 to A.5 hold, then

$$H(\tilde{\alpha} - \alpha) - \frac{\tilde{I}^2_{\tilde{m}}}{2}\tilde{B}^*(z) = o_{\mathbb{P}}(\tilde{h}^2_{\tilde{m}}) + O_{\mathbb{P}}\big(n^{-1/2}\tilde{D}^{-1/2}\iota_{\tilde{m}}\big), \tag{3.7}$$

where $\tilde{B}^*(z) := \tilde{S}^{-1}\tilde{B}(z)$.

(ii) If Assumptions A.1 to A.6 hold, then

$$\big(n^{1/2}\tilde{D}^{1/2}\big)\left[H(\tilde{\alpha} - \alpha) - \frac{\tilde{I}^2_{\tilde{m}}}{2}\tilde{B}^*(z) + o_{\mathbb{P}}(\tilde{h}^2_{\tilde{m}})\right] \xrightarrow{d} N\big(\mathbf{0}, \tilde{f}^{-1}_{\tilde{m}}(z)\tilde{\Delta}\big), \tag{3.8}$$

where $\tilde{\Delta} := \tilde{S}^{-1}\tilde{S}^*\tilde{S}^{-1}$.

We can use our preceding results to obtain an estimator for a purely nonparametric SUR model characterized by

$$y_{j,it} = g_j(Z_{j,it}) + \epsilon_{j,it}, \; g_j(\cdot) : \mathbb{R}^{d_j} \to \mathbb{R}, \text{ for } j = 1, 2. \tag{3.9}$$

To begin, we set $W_{j,it} = \widetilde{X}_{j,it} = 1$, $\breve{Q}_{j,it} := \begin{pmatrix} 1 \\ (Z_{j,it} - z_j)/h_j \end{pmatrix}$, and $\breve{U}_{j,it} := \begin{pmatrix} 1 \\ (Z_{j,it} - z_j) \end{pmatrix}$. We define the system estimator in this case as $\breve{\alpha}$, where

$$\breve{\alpha} = [\breve{Q}'(K^{1/2}A^{-1}K^{1/2})\breve{U}]^{-1}[\breve{Q}'(K^{1/2}A^{-1}K^{1/2})y]. \tag{3.10}$$

The results of Propositions 3.2 and 3.3 continue to hold but with $\breve{\widetilde{S}}_n$, $\breve{S}$, $\breve{B}_n$, $\breve{R}_n$, $\breve{T}^*_n$ and $\breve{S}^*$ respectively in lieu of $\widetilde{S}_n$, S, B_n, R_n, T^*_n and S^*. The following corollary establishes consistency and asymptotic normality of $\breve{\alpha}$.

Corollary 3.7. *Suppose in 2.3* $\widetilde{X}_{j,it} = 1, \; \forall \; i, j, t$. *That is, suppose the purely nonparametric SUR in (3.9) is the model of interest.*

(i) If Assumptions A.1 to A.5 hold, then

$$H(\breve{\alpha} - \alpha) - \frac{\tilde{I}^2_{\tilde{m}}}{2}\breve{B}^*(z) = o_{\mathbb{P}}(\tilde{h}^2_{\tilde{m}}) + O_{\mathbb{P}}\big(n^{-1/2}\tilde{D}^{-1/2}\iota_{\tilde{m}}\big), \tag{3.11}$$

where $\breve{B}^*(z) := \breve{S}^{-1}\breve{B}(z)$.

(ii) If Assumptions A.1 to A.6 hold, then

$$\left(n^{1/2}\tilde{D}^{1/2}\right)\left[H(\breve{\alpha}-\alpha)-\frac{\tilde{I}_{\tilde{m}}^2}{2}\breve{B}^*(z)+o_{\mathbb{P}}(\tilde{h}_{\tilde{m}}^2)\right]\xrightarrow{d}N\big(\mathbf{0},\tilde{f}_{\tilde{m}}^{-1}(z)\breve{\Delta}\big),\tag{3.12}$$

where $\breve{\Delta} := \breve{S}^{-1}\breve{S}^*\breve{S}^{-1}$.

Remark 3.8. Consider the specific case of Corollary 3.7 in which $Z_{j,it}$ is a scalar – as in Welsh & Yee (2006). Then it is straightforward to show that $\breve{B}^*(z)$ simplifies to the 4-dimensional vector $\big(\mu_{1,2}(K_1)\nabla^2 g_1(z), 0, \mu_{2,2}(K_2)\nabla^2 g_2(z), 0\big)'$, whereas $\breve{\Delta}$ becomes the 4×4 matrix *block diag*$\big(\breve{\Delta}_1,\breve{\Delta}_2\big)$ with

$$\breve{\Delta}_j=\sigma_j^2(z)\left(\begin{array}{cc}\nu_{j,0} & 0\\ 0 & \mu_{j,2}(K_j^2)(\mu_{j,2}(K_j))^{-2}\end{array}\right).$$

This specific case captures the asymptotic properties of a scaled variant of the nonparametric SUR estimator in Welsh & Yee (2006), for panel data and under the assumption that the errors $\{\epsilon_{it}\}$ are *i.i.d.* across i for each fixed t. Note that, unlike the unscaled $\breve{\alpha}$ in Welsh & Yee (2006), scaling of our $\breve{\alpha}$ by the H matrix renders (i) the biases of the derivative estimators asymptotically zero and (ii) the estimators of $g_j(z_j)$ and $\nabla g_j(z_j)$ asymptotically uncorrelated. However, similar to Welsh & Yee (2006), the estimates of $\big(g_1(z_1),\nabla g_1(z_1)\big)'$ and $\big(g_2(z_2),\nabla g_2(z_2)\big)'$ are asymptotically uncorrelated. These observations also apply to $\widehat{\alpha}$ and $\tilde{\alpha}$.

For $\widehat{\alpha}_{GMM}$, our second estimator, we seek to establish the asymptotic properties of a properly normalized variant of $H(\widehat{\alpha}_{GMM}-\alpha)$, which we express as

$$H(\widehat{\alpha}_{GMM}-\alpha)-(\widetilde{S}_n'\Gamma^{-1}\widetilde{S}_n)^{-1}(\widetilde{S}_n'\Gamma^{-1}B_n)-(\widetilde{S}_n'\Gamma^{-1}\widetilde{S}_n)^{-1}(\widetilde{S}_n'\Gamma^{-1}R_n)=(\widetilde{S}_n'\Gamma^{-1}\widetilde{S}_n)^{-1}(\widetilde{S}_n'\Gamma^{-1}T_n^*),\tag{3.13}$$

where we redefine $\widetilde{S}_n, B_n, R_n$ and T_n^* respectively as

$$\widetilde{S}_n=\frac{1}{n}Q'\widetilde{K}\widetilde{U},\ B_n=\frac{1}{n}Q'\widetilde{K}\widetilde{X}\bar{R},\ R_n=\frac{1}{n}Q'\widetilde{K}\widetilde{X}R,\ T_n^*=\frac{1}{n}Q'\widetilde{K}\epsilon.$$

The following assumptions are needed to establish the asymptotic properties of $\widehat{\alpha}_{GMM}$ in the case of large N and small T.

Assumption B.1. For each $t\geq 1$, $G_{1t}(z_1,z_2)$ and $f_{1t}(z_1,z_2)$, the joint density of Z_{i1} and Z_{it}, are continuous at $(z_1=z, z_2=z)$. Also, for each z, $\Omega(z)>0$ and $f(z)>0$, where $f(z)$ is the marginal density function of Z_{it}. Further, $\sup_{t\geq 1}|G_{1t}(z,z)f_{1t}(z,z)|\leq M(z)<\infty$ for some function $M(z)$. Finally, $g(z)$ and $f(z)$ are both twice continuously differentiable at $z\in\mathbb{R}^d$.

Assumption B.2. The kernel function $K(\cdot)$ is an even, nonnegative, and bounded density function with compact support.

Assumption B.3. $h\to 0$ and $Nh^d\to\infty$ as $N\to\infty$.

Assumption B.4. There exists some $\delta>0$ such that $E\{|\epsilon_{it}W_{it}|^{(2+\delta)}|Z=u\}$ is continuous at $u=z$.

Clearly, Assumptions B.1 to B.4 are simplifications of Assumptions A.2, A.3, A.5, A.6, respectively. We now state the asymptotic properties of $\widehat{\alpha}_{GMM}$.

Proposition 3.9. *Suppose $Z_{1,it} = Z_{2,it} = Z_{it}$, $h_1 = h_2 = h$, and $K_1 = K_2 = K$. If Assumptions A.1, A.4 and B.1 to B.3 hold, then*

(i) $\widetilde{S}_n = f(z)S\{1 + o_{\mathbb{P}}(1)\}$,

(ii) $B_n = \frac{1}{2}h^2 f(z)B(z) + o_{\mathbb{P}}(h^2)$,

(iii) $R_n = o_{\mathbb{P}}(h^2)$,

(iv) $nh^d Var(T_n^*) = f(z)S^{**}$, *where*

$$S^{**} := \begin{pmatrix} S_1^* & S_{12}^* \\ S_{21,}^* & S_2^* \end{pmatrix}, \ S_{12}^* := \begin{pmatrix} \Omega_{12}^{12}\nu_0 & \mathbf{0} \\ \mathbf{0}' & \Omega_{12}^{12} \otimes \mu_2(K^2) \end{pmatrix}, \ S_{21}^* = (S_{12}^*)'.$$

Remark 3.10. The result in Proposition 3.9 (iv) suggests that the diagonal and off-diagonal block terms in $Var(T_n^*)$ are of the same order of magnitude: an efficacy of the assumption of common Z's across equations; consequently, if we impose this assumption on our estimator $\widehat{\alpha}$, its asymptotic variance will not be a block diagonal matrix.

Theorem 3.11. *Suppose $Z_{1,it} = Z_{2,it} = Z_{it}$, $h_1 = h_2 = h$, and $K_1 = K_2 = K$.*

(i) If Assumptions A.1, A.4 and B.1 to B.3 hold, then

$$H(\widehat{\alpha}_{GMM} - \alpha) - \frac{h^2}{2}B^*(z) = o_{\mathbb{P}}(h^2) + O_{\mathbb{P}}\big(n^{1/2}h^{d/2}\big), \tag{3.14}$$

where $B^*(z) := \big(S'\Gamma^{-1}S\big)^{-1}\big(S'\Gamma^{-1}B(z)\big)$.

(ii) If Assumptions A.1, A.4 and B.1 to B.4 hold then

$$n^{1/2}h^{d/2}\left[H(\widehat{\alpha}_{GMM} - \alpha) - \frac{h^2}{2}B^*(z) + o_{\mathbb{P}}(h^2)\right] \xrightarrow{d} N\big(\mathbf{0}, f^{-1}(z)\Delta_{GMM}\big), \tag{3.15}$$

where $\Delta_{GMM} := (S'\Gamma^{-1}S)^{-1}S'\Gamma^{-1}S^{**}\Gamma^{-1}S(S'\Gamma^{-1}S)^{-1}$.

Remark 3.12. To estimate our empirical bivariate simultaneous model in the ensuing section, we use the $\widehat{\alpha}_{GMM}$ estimator and assume that the Z variables and the kernel density function are the same but allow for the bandwidths to differ across equations. The salient asymptotic properties of $\widehat{\alpha}_{GMM}$ are not eliminated by the assumption of different bandwidths across equations. In addition, we use a wild-bootstrap to conduct inference; it is well-known that in practice researchers often wish to avoid relying on the asymptotic variance for conducting inference because of the relatively slow rate of convergence. For details on using the bootstrap to conduct inference in nonparametric regression models, we refer the reader to Henderson & Parmeter (2015).

4 Empirical Application

4.1 An Empirical Simultaneous Model of Growth and FDI

Our empirical bivariate semiparametric system of equations model allows for the economic growth rate and FDI to be modeled simultaneously. We measure economic growth rate, GRO_{it}, as the growth rate of real per capita GDP, and FDI_{it} as the share of FDI inflows to GDP. We let $i = 1, 2, \ldots, N$ denote country index, and $t = 1, 2, \ldots, T$ denote the time period. Our bivariate model takes the form

$$\begin{aligned} GRO_{it} &= FDI_{it}\lambda_1(Z_{it}) + X'_{1,it}\gamma_1(Z_{it}) + \epsilon_{1,it} \\ FDI_{it} &= GRO_{it}\lambda_2(Z_{it}) + X'_{2,it}\gamma_2(Z_{it}) + \epsilon_{2,it}. \end{aligned} \tag{4.1}$$

In our empirical model of (4.1), $X_{j,it}$ is a k_j-dimensioned vector of control variables for equations $j = 1, 2$, such that the first entry in the vector is equal to one; $X_{j,it}$ may share common elements across j. $\gamma_j(\cdot)$ and $\lambda_j(\cdot)$ are unknown smooth coefficient functions of conformable dimensions. We presume that Z_{it} is a d-dimensioned vector of environmental variables, which may include a mix of continuous and discrete regressors (Racine & Li 2004, Li & Racine 2010). We assume that Z_{it} is constant across both equations and across each of the m_j coefficient functions. That is, we maintain the hypothesis of the same sources of parameter heterogeneities in the growth and FDI equations. As in our general model in Section 3, the errors $\epsilon_{j,it}$ are assumed to be mean zero disturbances that are correlated across equations, and all other model assumptions are assumed to be satisfied.

4.2 Data Overview

Our data are primarily derived from the 2012 World Development Indicators database published by the World Bank, unless otherwise specified. Our sample contains an unbalanced panel of 114 developed and developing countries, spanning the period 1984-2010. We average our data into 9 non-overlapping 3-year panels to reduce the influence of serial correlation on our results. Also, using time-averaged data can partially mitigate the effect of purely random measurement error and provide more reliable estimates for our variables of interest. Due to time averaging and a dearth of data on some variables for some countries, our effective sample contains 463 total observations.

4.3 Environmental Variables

We include a mix of continuous and discrete environmental variables in our specification of Z_{it}. Specifically, we include an index of corruption in Z_{it}. The index of corruption comes from the International Country Risk Guide published by the Political Risk Services Group. This measure of institutional quality is defined as "actual or potential corruption in the form of excessive patronage, nepotism, job reservations, 'favor-for-favors', secret party funding, and suspiciously close ties between politics and business."[6] The corruption index ranges from 0 to 6, with 0 representing high levels of corruption and 6 representing low levels of corruption.

We further allow for unobserved heterogeneity in all the coefficient functions across both countries and time, through an unordered country variable and ordered year categorical variable. The advantage of including country and year indicators in each of our coefficients is that we can control for country and time invariant effects – i.e., fixed effects – in a *non-neutral* fashion. That is, the country and year indicators capture *any* country and time invariant factors that induce heterogeneity in the intercept *and* slope coefficients across countries and time, which are likely to be present in our empirical model.

[6]This measure of corruption is quite popular in empirical works and, particularly has been used to study the effects of corruption on economic growth (Mauro 1995), investment (Mauro 1998), and the intersection between economic growth and FDI (McCloud & Kumbhakar 2012, Delgado et al. 2014), to name only a few.

4.4 Explanatory Variables

The control variables we choose for the growth equation pertain to both the benchmark neoclassical growth specification and the macroeconomic policy ideology. The control variables we choose for the FDI equation are empirically and (or) theoretically related to FDI.

4.4.1 The Growth Equation

We specify $X_{1,it}$ to contain initial income, the growth rate of the population, the rate of investment in physical capital, the inflation rate, government consumption, and openness. Initial income, population growth, and the rate of physical capital investment are the traditional neoclassical 'Solow' growth variables and are defined as follows. Initial income is the log of GDP per capita at the beginning of each 3-year panel period; the growth rate of the population is the annual percentage change in the total population; and investment in physical capital is defined to be gross capital formation as a percentage of GDP. The inflation rate is the annual percentage change in the consumer price index. Government consumption is current period government expenditure on goods and services, excluding military spending on government capital formation. Openness is the sum of exports and imports as a percentage of GDP, and comes from the Penn World Table version 7.1 of Heston, Summers & Aten (2012).

Identification. Under our maintained assumption that economic growth and FDI are determined simultaneously, there is concern that FDI is endogenous in our growth equation. In the growth equation, a valid instrumental variable for FDI is one that is correlated with FDI, but uncorrelated with economic growth, conditional on both X_1 (i.e., 'Solow' and macroeconomic policy variables) and Z (i.e., institutional quality and unobservable country and year effects).

In general, the literature that has investigated single equation growth models with FDI as the key variable has failed to unearth a generally suitable instrumental variable for FDI.[7] Borensztein et al. (1998), Durham (2004) and Delgado et al. (2014) find evidence that lagged values of FDI perform well and mitigate at least part of the endogeneity of contemporaneously measured FDI. However, after considering which instrumental variables are available for panel data (i.e., with country and year variation) and not internal to FDI (i.e., lagged measures), we are left with one other potential source of exogenous variation in FDI: the total area of the country to measure country size. The general intuition for using the total land area as an instrument is that, all else equal, FDI is attracted to larger countries. Yet, there is no reason to believe that, given our set of control variables, economic growth is directly correlated with the size of the country. In addition, Borensztein et al. (1998) find empirical evidence to support the validity and strength of this instrumental variable. Hence, we consider the log of the total land area in square kilometers of a country as an instrumental variable for FDI in our growth equation. We assert that our use of an instrumental variable, in conjunction with our robust array of conditioning variables and generalized interactive fixed effects, are able to alleviate any concerns that endogeneity of FDI is driving any of our results.

[7] Notable contributions proposing a variety of instrumental variables for FDI – such as lagged values of FDI, some time-invariant measures of institutional quality, and total area of the country – include panel and cross-sectional growth studies by Borensztein et al. (1998), Alfaro et al. (2004), Durham (2004), and Delgado et al. (2014).

4.4.2 The FDI Equation

In the FDI equation, we specify $X_{2,it}$ to contain schooling, trade openness, the inflation rate, the foreign exchange rate, and the log of total GDP. The level of schooling in the host economy may have explanatory power in the FDI equation. Often, foreign firms bring advanced technology into the host economy that requires a relatively more skilled labor force relative to that required by pre-existing technology. Hence, all else equal, a more highly educated labor force may be more inviting for FDI.We define schooling to be the net enrollment rate in secondary school; and trade openness and inflation are the same as those measures in the growth equation. The foreign exchange rate comes from the Penn World Table version 7.1 of Heston et al. (2012), and is defined as the exchange rate to United States dollars. Each of these control variables are important correlates of FDI, measuring the relative attractiveness, macroeconomic conditions and policies, and ease of entry into the host country (foreign exchange, trade openness, and inflation), the size of the country (log of total GDP), and the degree of absorptive capacity (schooling).

Identification. Endogeneity of the growth rate likely arises because foreign investors view strong economic growth as a favorable metric of financial returns, and as a general measure of economic and institutional stability. Appropriate instrumental variables for the growth rate in the FDI equation must be factors that are correlated with economic growth, but uncorrelated with FDI, conditional on both X_2 (i.e., macroeconomic stability and schooling variables) and Z (i.e., institutional quality and country and year effects). One important set of growth correlates that are unlikely to be correlated with FDI are demographic growth correlates – specifically, the fertility rate and life expectancy. Henderson, Papageorgiou & Parmeter (2012) find robust econometric evidence that the demographic growth variables have a nontrivial relationship with economic growth, using robust nonparametric estimators; hence, there is ample evidence that these demographic variables are correlated with growth. We argue that these variables are uncorrelated with FDI decisions, as FDI decisions are typically related to investment risk and private return. Of course, fertility and life expectancy may be correlated with economic or institutional factors that may also determine economic risk and return factors that influence FDI. However, conditional on the set of economic and institutional factors in X_2 and Z, we maintain that all potential indirect linkages between demographic variables and FDI have been accounted for. Hence, demographic growth variables are suitable instruments for economic growth in the FDI equation. The fertility rate is defined as the average number of births per woman, and life expectancy is defined as the life expectancy at birth measured in years.

4.5 Practical Implementation of our Estimator and Goodness of Fit Measures

With regards to the practical implementation of our proposed estimator in the context of the economic growth and FDI model in (4.1), a few clarifications are appropriate. Since Z_{it} is assumed to contain a mix of continuous variables and unordered and ordered discrete categorical variables, we adopt the generalized product kernel technique of Racine & Li (2004) and Li & Racine (2010). We define the product kernel function $K_{h_j}(\cdot) := h_j^{-d_j} K_j(\cdot/h_j)$ to be

$$K_j(\cdot) = \prod_{c=1}^{d_j^c} k^c \left(\frac{Z_{j,it}^c - z_j^c}{h_j^c} \right) \prod_{u=1}^{d_j^u} k^u(Z_{j,it}^u - z_j^u; h_j^u) \prod_{o=1}^{d_j^o} k^o(Z_{j,it}^o - z_j^o; h_j^o) \tag{4.2}$$

in which

$$k^c\left(\frac{Z^c_{j,it}-z^c_j}{h^c_j}\right)=\frac{1}{\sqrt{2\pi}}\exp\left[\frac{1}{2}\left(\frac{Z^c_{j,it}-z^c_j}{h^c_j}\right)^2\right] \tag{4.3}$$

is a univariate Gaussian kernel function used for each of the d^c_j continuous variables in $Z_{j,it}$,

$$k^u(Z^u_{j,it}-z^u_j;h^u_j)=\begin{cases}1 & \text{if } Z^u_{j,it}-z^u_j=0\\ h^u_j & \text{if } Z^u_{j,it}-z^u_j\neq 0\end{cases} \tag{4.4}$$

is a univariate discrete kernel function used for each of the d^u_j unordered discrete variables in $Z_{j,it}$, and

$$k^o(Z^o_{j,it}-z^o_j;h^o_j)=\begin{cases}1 & \text{if } Z^o_{j,it}-z^o_j=0\\ {h^o_j}^{|Z^o_{j,it}-z^o_j|} & \text{if } Z^o_{j,it}-z^o_j\neq 0\end{cases} \tag{4.5}$$

is a univariate discrete kernel function used for each of the d^o_j ordered discrete variables in $Z_{j,it}$ (Li & Racine 2007). In the above product kernel setup, h^c_j is a d^c_j-dimensioned vector of bandwidths for the continuous variables, and h^u_j and h^o_j are d^u_j- and d^o_j-dimensioned vectors of unordered and ordered discrete variable bandwidths.[8]

We select the optimal smoothing parameters, $\{h^c_j,h^u_j,h^o_j\}$ using the method of least squares cross validation. The method of least squares cross validation selects $\{h^c_j,h^u_j,h^o_j\}$ by minimizing the following criterion function

$$\underset{\{h^c_j,h^u_j,h^o_j\}}{\arg\min}\sum_{j=1}^{J}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[y_{j,it}-\widetilde{X}'_{-j,-it}\widehat{a}_j(Z_{-j,-it})\right]^2 \tag{4.6}$$

in which $\widetilde{X}'_{-j,-it}\widehat{a}_j(Z_{-j,-it})$ is the leave-one-out nonparametric generalized method of moments estimate of $\widetilde{X}'_{j,it}\widehat{a}_j(Z_{j,it})$. Although we have constructed our empirical model such that $Z_{j,it}=Z_{it}$ for each j, our cross validation procedure selects a set fixed bandwidths for each variable in Z_{it} in each equation; note that we do not restrict the bandwidths to be fixed across equations. Given our small number of cross-sectional units, we opt to not use our asymptotic distribution in computing the standard error of our estimate. We obtain standard error for our estimate of α from a wild bootstrap procedure based on 399 replications, which corrects for heteroscedasticity of unknown form.

We provide three separate measures of the goodness of fit for each equation in our model. The first measure is the in-sample R^2 calculated as $R^2_j=\text{cor}\left[y_{j,it},\widetilde{X}'_{j,it}\widehat{a}_j(Z_{j,it})\right]^2$, the square of the correlation between the observed dependent variable in equation j and its estimated counterpart. The second measure is the out-of-sample R^2, and the third measure is the out-of-sample Average Squared Prediction Error ($ASPE$), calculated as $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[y_{j,it}-\widetilde{X}'_{j,it}\widehat{a}_j(Z_{j,it})\right]^2$ for each equation. The advantage of using out-of-sample measures of fit is that these measures are typically robust to over-fitting, which can sometimes inflate in-sample measures of fit.[9]

[8]In our empirical model, $d^c_j=d^u_j=d^o_j=1$. Note that the presence of discrete components in z_j has changed the interpretation of some of the regularity conditions in Sections 2 and 3. In particular, (i) $z_j\in\mathbb{R}^{d_j}$ should be interpreted as $(z^c_j,z^u_j,z^o_j)\in\mathbb{R}\times A^u\times A^o$, the product space where A^u and A^o denote the finite support of z^u_j and z^o_j, respectively, and (ii) the derivative with respect to z_j should be interpreted as the derivative with respect to z^c_j.

[9]To implement our out-of-sample goodness of fit measures, we sample 80 percent of our data without replacement and fit our model. We then use our estimates to predict on the 20 percent hold out sample and calculate both the R^2 and $ASPE$. We repeat this procedure 1,000 times, and report the mean R^2 and $ASPE$, in order to avoid any potential biases in our measures of fit induced by our choice of sample splits. We refer the reader to Racine & Parmeter (2013) for additional details on model evaluation, including adjustments to the optimal bandwidth parameter to account for different sample sizes arising from the out-of-sample splits.

5 Empirical Results

We now turn to the results from applying our semiparametric instrumental variables systems estimator to our novel empirical model of economic growth and FDI in (4.1). We consider estimates of (2.15) that use first stage estimates of A as our weighting matrix. Since our semiparametric systems estimator provides observation-specific estimates and standard errors, we summarize these estimates using 45 degree gradient plots that are depicted by Figures 1 to 2 (for details on the plots, see Henderson, Kumbhakar & Parmeter 2012). The 45 degree gradient plots found in the lower panels of the figures show the observation specific function estimates plotted on the 45 degree line, with 95 percent observation specific confidence intervals plotted above and below each point estimate. If the horizontal dotted line at zero lies outside of each observation specific confidence interval, then that point estimate is statistically significant.

5.1 Characterizing the Types of Interactions between Economic Growth and FDI

The corresponding 45 degree gradient plot in Figure 1 shows that the majority of these estimates are statistically significant at the 5 percent level. For the FDI equation, Figure 1 reveals that the distribution of growth coefficient estimates is generally positive, and the corresponding 45 degree plot reveals that many of these positive estimates are statistically significant. There is a clear absence of statistical parity among the estimated coefficients for FDI inflows and growth.

Our results in Figure 1 yield an important observation – empirically, FDI has positive, negative or no effect on economic growth, and vice versa. This observation therefore parallels the theoretical predictions of the effect on FDI on growth and the effect of growth on FDI. Hence, our semiparametric system of simultaneous equations, coupled with the taxonomy in Definition 2.1, seems quite appropriate for our analysis of the types of interactions between growth and FDI.

To characterize the types of interactions between growth and FDI on the basis of the taxonomy in Definition 2.1, we use the following criterion: a country is placed in the category, for example, symbiosis if at least 50 percent of its estimated effects of FDI on growth and its estimated effects of growth on FDI is positive and statistically significant. This criterion, therefore, is not all-inclusive; for some countries, the estimated effect of, say, FDI on growth changes sign and (or) statistical significance across time periods. That is, on the basis of this criterion a country can belong to two or no categories in Definition 2.1.

Table 1 contains the number of countries that can be characterized according to our Definition 2.1 and criterion. In this table, we use the term *distinct* to refer to countries that fall in only one category. We find that the most dominant relationships between FDI and growth are symbiosis and FDI-commensalism; that is, for a large number of countries either (i) FDI has a positive effect on growth and growth has a positive effect on FDI, or (ii) FDI has a positive effect on growth but growth has no effect on FDI. The number of countries that are in the symbiosis category is 63, of which 45 are distinct. To get an understanding of the economic significance of the nature of the growth-FDI interaction, consider two countries Australia and Turkey that fall into only the symbiosis category. For the period 1993 to 1995, for example, our estimates reveal that in Australia a 1 percent increase in FDI leads to a 0.987 percent increase in economic growth and a 1 percent increase in economic growth leads to a 0.097 percent increase in FDI; for this same period, in Turkey a 1 percent increase in FDI leads to a 0.228 percent increase in economic growth and a 1 percent increase in economic growth leads to a 0.062 percent increase in FDI. These symbiosis estimates imply that appreciable direct multiplier effects exist between FDI and growth in Australia and Turkey. These multiplier effects suggest that (at least some) resources these two countries may devote to strengthening the effect of, say, FDI on economic growth can be reallocated to other correlates of economic growth.

The dominance of the FDI-commensalism category means that many more countries experience FDI-commensalism than those that experience growth-commensalism – FDI has no effect on growth but growth has a positive effect on FDI. Only 2 countries experience non-symbiosis – FDI has no effect on growth and growth has no effect on FDI – which suggests that at least a one-way relationship between growth and FDI exists in almost all countries. In addition, and in fact, only developing countries (very few) fall into this latter category, as well as the categories of FDI-antagonistic symbiosis, synnercrosis and growth-commensalism. In addition, there is no country that experiences FDI-ammensalism – FDI has no effect on growth and growth has a negative effect on FDI.

5.2 Effect of Institutional Improvement

Recall that for this paper, an improvement in institutional quality means a reduction in the level of corruption. Figure 2 provides a set of the results of an improvement in institutional quality on the coefficients in the instrumental variables model. The 45 degree plots show that many of these effects are significant in the FDI coefficient case, but that many of the growth coefficient partials are statistically insignificant. However, it is clear that there are subsets of growth coefficient partials that are negative and positive and significant. Overall, an improvement in institutional quality weakens, strengthens or has no impact on the interactions between FDI and growth.

5.3 Cross-Validated Bandwidths and Model Specification

One important way we glean additional insight from our model about the nature of parameter heterogeneity is to examine the cross-validated bandwidths used for regression estimation. It is becoming increasingly well known that if a cross-validated bandwidth on a continuous regressor lies below 2 or 3 times the standard deviation of the regressor in a local linear regression, then that variable is chosen by the cross-validation procedure to enter nonlinearly into the regression model (for details, see Li & Racine 2007). For discrete variables, a bandwidth that is less than unity implies nonlinear, nontrivial interactions in the regression. While examining the cross-validated bandwidths does not amount to a formal model specification test, the bandwidths generate insight into nature of the model that best fits the data: the cross-validated bandwidths in our model shows that for each environmental variable – corruption, country effect, and year effect – the best fit of our model to the data is one that incorporates nonlinear interactive effects. That is, we find that our bandwidths are less than their upper bounds, which is a signal that any ad hoc parametric linear restriction is not justified by the data.[10] Further, existence of nonlinear interactions does not provide insight into correct parametric specification; hence our bandwidth analysis signals that parametric restrictions on the functional form of heterogeneity within our model should be carefully considered and supported by appropriate model specification tests. We finally note that since the degree of smoothing varies across equations for each regressor, we conclude that the nature of these nontrivial interactions differs across equations as well.

[10]See, also, Henderson, Papageorgiou & Parmeter (2012) for a detailed discussion of the implications of cross-validated bandwidths for model specification in an empirical growth context.

5.4 Developed versus Developing Countries

Our foregoing empirical results strongly suggest that institutional quality can impinge on the growth-FDI interactions across countries. In addition, it is well known that developed countries, on average, have better institutional quality than their developing counterparts. On the basis of such difference in institutional quality, developed countries that experience *symbiosis* and *FDI-commensalism* may have a natural comparative advantage over their developing counterparts. On the surface, two empirical regularities – which are not necessarily mutually exclusive – may lend credence to such comparative advantage. *One*, there are two main types of FDI that flow to host countries: vertical FDI – investment that allows for different components of a final good to be produced in different countries with different factor intensities; and horizontal FDI – investment that allows for the entire production process of a final good to be replicated in a foreign country that is within close proximity to major foreign markets. On average, developed host countries receive mostly horizontal FDI, whereas developing host countries receive mostly vertical FDI. *Two*, differences in institutional quality is associated with, among other things, differences in investment climate, factor endowments and thus absorptive capacity, direct and indirect transaction costs, and organizational structure of firms and industries.

To investigate this concern, we examine whether the conditional densities of the growth and FDI effects differ between OECD and non-OECD countries using kernel densities and boxplots. In Figure 3, the top graph shows superimposed kernel density plots of FDI effects (from the growth equation) for both OECD and non-OECD countries, whereas the bottom graph shows superimposed kernel density plots of growth effects (from the FDI equation) for both OECD and non-OECD countries. In Figure 4, the top graph shows boxplots across quartiles of FDI effects for both OECD and non-OECD countries, whereas the bottom graph shows boxplots across quartiles of growth effects for both OECD and non-OECD countries. A cursory glance at these graphs suggests a discernible difference in both sets of growth and FDI effects between OECD and non-OECD countries. However, a formal test of the difference in densities is warranted prior to drawing inferences from these graphs; we use the nonparametric kernel-based test for equality of distributions by Li, Maasoumi & Racine (2009) to test for statistical differences between the OECD and non-OECD distributions of FDI and growth effects. The application of this nonparametric kernel-based test to our estimated FDI and growth effects yields p-values of 0.0000 under the null hypothesis of equality of the OECD and non-OECD for both FDI and growth densities. Thus, we reject the null hypothesis of equality of OECD and non-OECD distributions for both sets of growth and FDI effects. Therefore, these formal statistical tests confirm our intuition that OECD and non-OECD countries have statistically different interactions between FDI and economic growth.

Looking specifically at Figure 3, it is clear that for both FDI and growth effects the distribution for OECD countries is generally centered at zero, but has a fat right-tail that indicates a subset of non-zero effects. Non-OECD countries, on the other hand, do not have a large mass at zero for either FDI or growth effects, and are distributed generally over positive, non-zero values. These differences suggest that the *symbiosis* and *FDI-commensalism* between growth and FDI are not substantial in many OECD countries. This finding is consistent with our earlier results that indicate an important interactive relationship between growth, FDI, and institutional quality: countries that have better institutions, on average, have smaller *symbiosis* and *FDI-commensalism* interactions between growth and FDI. Non-OECD countries, on the other hand, have sizeable *symbiosis* and *FDI-commensalism* interactions; for *symbiosis*, this indicates that in the absence of high quality institutions, FDI is an important component for economic growth *and* economic growth rates are important for attracting

FDI.

These results have important implications, particularly for developing countries. First, it is clear that in many developing countries, FDI is a key factor for growth; yet, since growth is crucial for attracting FDI, it is clear that countries looking to improve growth rates through FDI may not be successful given that they are not relatively attractive to FDI investors. Improvements in institutional quality may be one way to circumvent this cycle. Our results also indicate that countries with high levels of institutional quality (e.g., OECD countries) may not gain much from improving growth rates by pursuing policies aimed at attracting FDI.

We turn to Figure 4 to focus on the distribution of FDI and growth effects across OECD and non-OECD countries within each quartile. These boxplots provide an alternative view into the differences in our estimates across developed and developing countries. The top panel in the figure shows the distribution of FDI effects in the growth equation across OECD and non-OECD countries; the bottom panel shows the distributions for the growth effects in the FDI equation. It is clear from the top panel that the distribution of estimates within the first and third percentile are generally wider for non-OECD countries, with a wider interquartile range and higher mean for non-OECD countries in each group. We do not see much difference in FDI effects at the second percentile across OECD and non-OECD countries, and we see a slightly wider interquartile range for OECD countries at the highest quartile. Interestingly, we find that the interquartile range is higher, with higher mean estimate, within each quartile of growth effects (FDI equation) for non-OECD countries. Therefore, although developed countries may have a natural comparative advantage because of their higher level of institutional quality, the magnitudes of their *symbiosis* and *FDI-commensalism* interactions between FDI and economic growth are smaller than those of their developing counterparts.

5.5 An Alternative Instrumental Variables Specification

We also estimate a semiparametric system of simultaneous equations model that uses the fertility rate, which supplants the life expectancy rate, as an instrument for growth in our FDI equation. These results parallel the preceding reported results that are predicated on the life expectancy rate. That is, across developed and developing economies, causal, heterogeneous *symbiosis* and *FDI-commensalism* are the most dominant types of interactions between FDI and economic growth. Higher institutional quality facilitates, impedes, or has no effect on the interactions between FDI and economic growth. In addition, our out-of-sample goodness of fit measures for this alternative model are lower than their counterparts in the preceding model; this observation lends credence to the preceding model.

6 Conclusion

In theory, FDI inflows can have positive, negative or no effect on economic growth, and vice versa. If within a country FDI has a positive effect on growth and growth has a positive effect on FDI – our concept of symbiosis – then FDI-promoting strategies for fostering and sustaining economic growth have added and direct multiplier benefits. In such a country, policymakers can therefore reallocate scarce resources to, for example, other correlates of economic development. To date, however, no study has analyzed empirically the types of interactions between economic growth and FDI that may exist within and across countries and the effect of institutional quality on such interactions.

In this paper, we characterize the types of interactions between FDI and economic growth, and analyze the effect of institutional quality on such interactions. To do so, we propose a novel semiparametric system of simultaneous equations model with the economic growth rate and FDI as a bivariate response. Our model unifies several important aspects of the empirical growth and FDI literatures, including (i) the joint determination of economic growth and FDI, (ii) nonlinear and nontrivial interactions of institutional quality with each of the conditioning variables, (iii) an instrumental variables approach for identification, (iv) unobserved heterogeneity (country- and time-specific effects) of unknown and non-neutral form, (v) and correlations in errors across equations. Only a few existing papers have explored a subset of these important model structures.

To estimate the coefficient functions and their derivatives for the proposed bivariate response growth-FDI model, we derive and establish the large sample properties of a class of semiparametric system of simultaneous equations estimators. We show using rigorous proof that our class of systems estimators is both consistent and asymptotically normal. We emphasize that our econometric model is fully generalizable to J separate equations, and is in no way restricted to the empirical growth-FDI context in which our model is framed. Our proposed class of systems estimators is a generalization of several important econometric models, including the fully parametric systems generalized method of moments estimator, the single-equation nonparametric generalized method of moments estimator, and the nonparametric system of equations (i.e., without endogeneity) estimators. Our proposed class of estimators is relatively straightforward to implement and, more important, has a wide range of applicability to economic and non-economic data.

Our proposed semiparametric system of equations model, and associated specification tools, suggests that across developed and developing economies, causal, heterogeneous (i) symbiosis and (ii) FDI-commensalism are the most dominant types of interactions between FDI and economic growth; that is, for a large number of countries either (i) FDI has a positive effect on growth and growth has a positive effect on FDI, or (ii) FDI has a positive effect on growth but growth has no effect on FDI, respectively. We further find that higher institutional quality facilitates, impedes or has no effect on the interactions between FDI and economic growth.

These findings are strong evidence in support of research advocating a more tailored, country-specific set of macroeconomic policies for the relationship between economic growth and FDI. Additionally, we uncover substantial heterogeneity in terms of interactions between our conditioning variables in each equation and institutional quality and country- and time-specific effects. Is it well-known that neglected heterogeneity can lead to misleading inferences on the parameters of interest. Thus, our findings underscore the importance of accounting for different sources of heterogeneities in a flexible – rather than the traditionally ad hoc parametric – manner to obtain consistent and generally reliable results. In essence, our new-fangled semiparametric system of simultaneous equations model coupled with its instrument-based estimator seems appropriate for assessing empirically the types of interactions between growth and FDI.

Consistent with the parametric GMM toolkit, several theoretical extensions of our framework are possible. In our current work, we maintain the standard assumptions of (relatively) few instruments and that the instruments are strong. It would certainly be interesting to develop nonparametric GMM estimators for system of simultaneous equations with many IVs. A Hansen J-Test for overidentification in the context of nonparametric GMM models would also be a useful tool.

Technical Appendix

In this appendix, we assume $C \in (0, \infty)$ is an arbitrary bounded constant. Recall that $n \equiv NT$; we use these terms interchangeably. The integral symbol represents a multiple integral of varying dimensions depending on the context in which it is used. We provide the proofs for only Propositions 3.2, 3.3 and 3.9 and Theorem 3.4 because the proofs for Corollary 3.7 and Theorems 3.6 and 3.11 are less involved. Many of the ensuing proofs use convergence in mean square.

Proof of Proposition 3.2: (i) Note that

$$\widetilde{S}_n = \frac{1}{n}\begin{pmatrix} Q_1' K_1^{1/2} D_\theta K_1^{1/2} U_1' H_1^{-1} & Q_1' K_1^{1/2} D_\beta K_2^{1/2} U_2' H_2^{-1} \\ Q_2' K_2^{1/2} D_\beta K_1^{1/2} U_1' H_1^{-1} & Q_2' K_2^{1/2} D_\gamma K_2^{1/2} U_2' H_2^{-1} \end{pmatrix} := \begin{pmatrix} \widetilde{S}_{n,11} & \widetilde{S}_{n,12} \\ \widetilde{S}_{n,21} & \widetilde{S}_{n,22} \end{pmatrix}. \tag{A.1}$$

The proofs for $\widetilde{S}_{n,11}$ and $\widetilde{S}_{n,22}$ follow directly from Cai and Li (2008) [Proof of Proposition (i)], which yields $\widetilde{S}_{n,11} = \widetilde{f}_1(z_1) D_\theta S_1\{1 + o_{\mathbb{P}}(1)\}$ and $\widetilde{S}_{n,22} = \widetilde{f}_2(z_2) D_\gamma S_2\{1 + o_{\mathbb{P}}(1)\}$. To complete the proof of Proposition 3.2, it therefore remains to show that (ia) $\widetilde{S}_{n,12} = o_{\mathbb{P}}(1)$ and (ib) $\widetilde{S}_{n,21} = o_{\mathbb{P}}(1)$.

We now prove part (ib); the proof of part (ia) can be easily established using the approach below.

$$\begin{aligned}
E[\widetilde{S}_{n,21}] &= E\left\{\frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta Q_{2,it}\widetilde{U}_{1,it}' K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)\right\} \\
&= \beta E\left\{Q_{2,it}\widetilde{U}_{1,it}' K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)\right\} \\
&= \beta E\begin{pmatrix} W_2\widetilde{X}_1' & W_2\widetilde{X}_1' \otimes (Z_1-z_1)'/h_1 \\ \widetilde{X}_1 W_2' \otimes (Z_1-z_1)/h_1 & W_2\widetilde{X}_1' \otimes (Z_1-z_1)(Z_1-z_1)'/h_1^2 \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(Z_1-z_1)K_{h_2}^{1/2}(Z_2-z_2) \\
&= \beta E\begin{pmatrix} \Omega_{21}(Z_1,Z_2) & \Omega_{21}(Z_1,Z_2) \otimes (Z_1-z_1)'/h_1 \\ \Omega_{21}(Z_1,Z_2)' \otimes (Z_1-z_1)/h_1 & \Omega_{21}(Z_1,Z_2) \otimes (Z_1-z_1)(Z_1-z_1)'/h_1^2 \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(Z_1-z_1)K_{h_2}^{1/2}(Z_2-z_2) \\
&= \beta \int \begin{pmatrix} \Omega_{21}(u_1,u_2) & \Omega_{21}(u_1,u_2) \otimes (u_1-z_1)'/h_1 \\ \Omega_{21}(u_1,u_2)' \otimes (u_1-z_1)/h_1 & \Omega_{21}(u_1,u_2) \otimes (u_1-z_1)(u_1-z_1)'/h_1^2 \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(u_1-z_1)K_{h_2}^{1/2}(u_2-z_2)f(u_1,u_2)du_1du_2 \\
&= \beta \int \begin{pmatrix} \Omega_{21}(z_1+h_1c_1, z_2+h_2c_2) & \Omega_{21}(z_1+h_1c_1, z_2+h_2c_2) \otimes c_1' \\ \Omega_{21}(z_1+h_1c_1, z_2+h_2c_2)' \otimes c_1 & \Omega_{21}(z_1+h_1c_1, z_2+h_2c_2) \otimes c_1c_1' \end{pmatrix} \\
&\qquad \times h_1^{-d_1/2}h_2^{-d_2/2}h_1^{d_1}h_2^{d_2}K_1^{1/2}(c_1)K_2^{1/2}(c_2)f(z_1+h_1c_1, z_2+h_2c_2)dc_1dc_2 \\
&= O\big(h_1^{d_1/2}h_2^{d_2/2}\big) = o(1),
\end{aligned}$$

where the second equality is by virtue of Assumption A.1, the fourth equality follows from law of iterative expectations (LIE), the sixth equality uses a change of variable, and the remaining equalities are consequences of changes in the implied canonical differential form, Lebesgue Dominated Convergence Theorem, and Assumptions A.2, A.3, and A.5.

We now show that $Var[\widetilde{s}_{n,21}] \to 0$ as $Nh_j^{d_j}h_k^{d_k} \to \infty$. We define

$$\widetilde{s}_{n,21}^{rs} := \frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta W_{2,itr}\widetilde{X}_{1,its}K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)$$

where $W_{2,itr}$ is the r-th element of $W_{2,it}$ and $\widetilde{X}_{1,its}$ is the s-th element of $\widetilde{X}_{1,it}$. Then, by Assumption A.1, we obtain

$$
\begin{aligned}
Var[\widetilde{s}_{n,21}^{rs}] &= \beta^2 \frac{1}{NT^2} Var\left\{ \sum_{t=1}^{T} W_{2,itr}\widetilde{X}_{1,its} K_{h_1}^{1/2}(Z_{1,it}-z_1) K_{h_2}^{1/2}(Z_{2,it}-z_2) \right\} \\
&= V_1 + V_2, \\
\text{where } \; V_1 &= \frac{\beta^2}{NT} Var\left\{ W_{2,i1r}\widetilde{X}_{1,i1s} K_{h_1}^{1/2}(Z_{1,i1}-z_1) K_{h_2}^{1/2}(Z_{2,i1}-z_2) \right\}, \\
V_2 &= \beta^2 \frac{2}{NT} \sum_{t=1}^{T-1} (T-t) Cov\Big(V_{21}, V_{2(t+1)}\Big), \\
V_{21} &= W_{2,i1r}\widetilde{X}_{1,i1s} K_{h_1}^{1/2}(Z_{1,i1}-z_1) K_{h_2}^{1/2}(Z_{2,i1}-z_2), \\
V_{2(t+1)} &= W_{2,i(t+1)r}\widetilde{X}_{1,i(t+1)s} K_{h_1}^{1/2}(Z_{1,i(t+1)}-z_1) K_{h_2}^{1/2}(Z_{2,i(t+1)}-z_2).
\end{aligned}
$$

Now, by Assumptions A.1 and A.2, and for a fixed T, it is straightforward to show that $V_1 \leq \frac{C}{NTh_1^{d_1}h_2^{d_2}}$. By similar arguments and using the Cauchy-Schwarz result that $|Cov(X,Y)| \leq Var(X)Var(Y)$ yield $|V_2| \leq \frac{CT}{Nh_1^{d_1}h_2^{d_2}}$. Hence, we obtain $Var[\widetilde{s}_{n,21}^{rs}] = O\Big(\big(Nh_1^{d_1}h_2^{d_2}\big)^{-1}\Big) = o(1)$ as required. Therefore, we have

$$
\frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T} \beta W_{2,it}\widetilde{X}_{1,it}' K_{h_1}^{1/2}(Z_{1,it}-z_1) K_{h_2}^{1/2}(Z_{2,it}-z_2) = o_{\mathbb{P}}(1).
$$

By invoking similar steps to those above, we deduce that

$$
\frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T} \beta W_{2,it}\widetilde{X}_{1,it}' \otimes (Z_{1,it}-z_1)'/h_1 \cdot K_{h_1}^{1/2}(Z_{1,it}-z_1) K_{h_2}^{1/2}(Z_{2,it}-z_2) = o_{\mathbb{P}}(1),
$$

$$
\frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T} \beta W_{2,it}\widetilde{X}_{1,it}' \otimes (Z_{1,it}-z_1)(Z_{1,it}-z_1)'/h_1^2 \cdot K_{h_1}^{1/2}(Z_{1,it}-z_1) K_{h_2}^{1/2}(Z_{2,it}-z_2) = o_{\mathbb{P}}(1).
$$

Thus, the proof of part (i) is complete.

(ii) Note that

$$
B_n = \frac{1}{n}\begin{pmatrix} Q_1' K_1^{1/2} D_\theta K_1^{1/2} \widetilde{X}_1 \overline{R}_1 + Q_1' K_1^{1/2} D_\beta K_2^{1/2} \widetilde{X}_2 \overline{R}_2 \\ Q_2' K_2^{1/2} D_\beta K_1^{1/2} \widetilde{X}_1 \overline{R}_1 + Q_2' K_2^{1/2} D_\gamma K_2^{1/2} \widetilde{X}_2 \overline{R}_2 \end{pmatrix} := \begin{pmatrix} B_{n,11} + B_{n,12} \\ B_{n,21} + B_{n,22} \end{pmatrix}. \tag{A.2}
$$

The proofs for $B_{n,11}$ and $B_{n,22}$ follow directly from Cai and Li (2008) [Proof of Proposition (ii)], which yields $B_{n,11} = (h_1^2/2) f_1(z_1) B_1(z_1) + o_{\mathbb{P}}(h_1^2)$ and $B_{n,22} = (h_2^2/2) f_2(z_2) B_2(z_2) + o_{\mathbb{P}}(h_2^2)$. To complete the proof, it remains to show that (iia) $B_{n,12} = o_{\mathbb{P}}(1)$ and (iib) $B_{n,21} = o_{\mathbb{P}}(1)$.

For (iia),

$$
E[B_{n,12}] = E\left\{ \frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T} \beta Q_{1,it}\widetilde{X}_{2,it}' \overline{R}_2(Z_{2,it}-z_2) K_{h_1}^{1/2}(Z_{1,it}-z_1) K_{h_2}^{1/2}(Z_{2,it}-z_2) \right\}
$$

$$
\begin{aligned}
h_2^{-2}E[B_{n,12}] &= E\beta\begin{pmatrix} W_{1,it}\widetilde{X}'_{2,it}A_2\big((Z_{2,it}-z_2)/h_2\big) \\ W_{1,it}\widetilde{X}'_{2,it}A_2\big((Z_{2,it}-z_2)/h_2\big)\otimes(Z_{1,it}-z_1)/h_1 \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2) \\
&= \beta\int\begin{pmatrix} \Omega_{12}(u_1,u_2)A_2\big((u_2-z_2)/h_2\big) \\ \Omega_{12}(u_1,u_2)A_2\big((u_2-z_2)/h_2\big)\otimes(u_1-z_1)/h_1 \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(u_1-z_1)K_{h_2}^{1/2}(u_2-z_2)f(u_1,u_2)du_1du_2 \\
&= h_1^{d_1/2}h_2^{d_2/2}\beta\int\begin{pmatrix} \Omega_{12}(z_1+c_1h_1,z_2+c_2h_2)A_2(c_2) \\ \Omega_{12}(z_1+c_1h_1,z_2+c_2h_2)A_2(c_2)\otimes c_1 \end{pmatrix} \\
&\qquad \times K_1^{1/2}(c_1)K_2^{1/2}(c_2)f(z_1+c_1h_1,z_2+c_2h_2)dc_1dc_2 \\
&= O\big(h_1^{d_1/2}h_2^{d_2/2}\big).
\end{aligned}
$$

Thus, $E[B_{n,12}] = O\big(h_1^{d_1/2}h_2^{(d_2+4)/2}\big)$. In addition, any (r,s)-entry of the $Var(B_{n,12})$ converges to zero.

Similarly, for (iib), we can show that $E[B_{n,21}] = O\big(h_1^{(d_1+4)/2}h_2^{d_2/2}\big)$, and any (r,s)-entry of the $Var(B_{n,21})$ converges to zero. Therefore, $B_{n,12} = o_{\mathbb{P}}(1)$ and $B_{n,21} = o_{\mathbb{P}}(1)$, which respectively *do not statistically dominate* $B_{n,11}$ and $B_{n,22}$. Hence, the proof of part (ii) is complete.

(iii) Note that

$$
R_n = \frac{1}{n}\begin{pmatrix} Q_1'K_1^{1/2}D_\theta K_1^{1/2}\widetilde{X}_1R_1 + Q_1'K_1^{1/2}D_\beta K_2^{1/2}\widetilde{X}_2R_2 \\ Q_2'K_2^{1/2}D_\beta K_1^{1/2}\widetilde{X}_1R_1 + Q_2'K_2^{1/2}D_\gamma K_2^{1/2}\widetilde{X}_2R_2 \end{pmatrix} := \begin{pmatrix} R_{n,11}+R_{n,12} \\ R_{n,21}+R_{n,22} \end{pmatrix}. \tag{A.3}
$$

The proofs for $R_{n,11}$ and $R_{n,22}$ follow directly from Cai and Li (2008) [Proof of Proposition (iii)], which yield that $R_{n,11} = o_{\mathbb{P}}(h_1^2)$ and $R_{n,22} = o_{\mathbb{P}}(h_2^2)$. To complete the proof, we now show that (iiia) $R_{n,12} = o_{\mathbb{P}}\big(h_1^{d_1/2}h_2^{(d_2+4)/2}\big)$ and (iiib) $R_{n,21} = o_{\mathbb{P}}\big(h_1^{(d_1+4)/2}h_2^{d_2/2}\big)$.

We prove part (iiib); by symmetry, the proof of part (iiia) easily follows.

$$
\begin{aligned}
E[R_{n,21}] &= E\Big\{\frac{1}{n}\sum_{i=1}^N\sum_{t=1}^T \beta Q_{2,it}\widetilde{X}'_{1,it}R_1(Z_{1,it}-z_1)K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)\Big\} \\
h_1^{-2}E[R_{n,21}] &= \beta E\begin{pmatrix} W_{2,it}\widetilde{X}'_{1,it}\ h_1^{-2}R_1(Z_{1,it},z_1) \\ W_{2,it}\widetilde{X}'_{1,it}\ h_1^{-2}R_1(Z_{1,it},z_1)\otimes(Z_{1,it}-z_1)/h_1\big) \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2) \\
&= \beta\int\begin{pmatrix} \Omega_{21}(u_1,u_2)\ h_1^{-2}R_1(u_1,z_1) \\ \Omega_{21}(u_1,u_2)\ h_1^{-2}R_1(u_1,z_1)\otimes(u_1-z_1)/h_1\big) \end{pmatrix} \\
&\qquad \times K_{h_1}^{1/2}(u_1-z_1)K_{h_2}^{1/2}(u_2-z_2)f(u_1,u_2)du_1du_2,
\end{aligned}
$$

where the last equality is a consequence of LIE. Applying a change of variables, the result that $h_1^{-2}R_1(z_1+c_1h_1,z_1) = o(1)$, Lebesgue Dominated Convergence Theorem, and Assumptions A.1 to A.3, we obtain $R_{n,21} = o_{\mathbb{P}}\big(h_1^{(d_1+4)/2}h_2^{d_2/2}\big)$ as required. By symmetry, it is straightforward to derive the result that $R_{n,12} = o_{\mathbb{P}}\big(h_1^{d_1/2}h_2^{(d_2+4)/2}\big)$. Furthermore, any (r,s)-entry of both the $Var(R_{n,21})$ and $Var(R_{n,12})$ converges to zero. In essence, the terms $R_{n,11}$ and $R_{n,22}$ stochastically dominate their counterparts. Therefore, we obtain the desired result. $\square$

Proof of Proposition 3.3: Since $E[T_n^*] = \mathbf{0}$, we write $n\tilde{\mathcal{D}}Var(T_n^*) = n\tilde{\mathcal{D}}E[T_n^* T_n^{*\prime}]$. Now

$$
\begin{aligned}
E[T_n^* T_n^{*\prime}] &= E\begin{pmatrix} T_{n1}^* T_{n1}^{*\prime} & T_{n1}^* T_{n2}^{*\prime} \\ T_{n2}^* T_{n1}^{*\prime} & T_{n2}^* T_{n2}^{*\prime} \end{pmatrix}, \qquad (A.4)\\
\text{where } \; T_{n1}^* &= \frac{1}{n} Q_1' K_1^{1/2} D_\theta K_1^{1/2} \epsilon_1 + \frac{1}{n} Q_1' K_1^{1/2} D_\beta K_2^{1/2} \epsilon_2 = T_{n,11}^* + T_{n,12}^*,\\
T_{n2}^* &= \frac{1}{n} Q_2' K_2^{1/2} D_\beta K_1^{1/2} \epsilon_1 + \frac{1}{n} Q_2' K_2^{1/2} D_\gamma K_2^{1/2} \epsilon_2 = T_{n,21}^* + T_{n,22}^*.
\end{aligned}
$$

To prove that

$$
n\tilde{\mathcal{D}}Var(T_n^*) = \tilde{f}_{\tilde{l}}(z) D_{\theta\gamma}^2 S^*, \qquad (A.5)
$$

we will show that the off-diagonal block terms for $E[T_n^* T_n^{*\prime}]$ in A.4 are of smaller order than its $(1,1)$ and $(2,2)$ main-diagonal block terms, which are of orders $O\{(nh_1^{d_1})^{-1}\}$ and $O\{(nh_2^{d_2})^{-1}\}$ respectively.

(i) To compute $E[T_{n1}^* T_{n1}^{*\prime}]$, note that

$$
T_{n1}^* T_{n1}^{*\prime} = T_{n,11}^* T_{n,11}^{*\prime} + T_{n,11}^* T_{n,12}^{*\prime} + T_{n,12}^* T_{n,11}^{*\prime} + T_{n,12}^* T_{n,12}^*. \qquad (A.6)
$$

For the first term in (A.6), we have,

$$
E[T_{n,11}^* T_{n,11}^{*\prime}] = Var(T_{n,11}^*) = Var\left\{ \frac{1}{n} \sum_{i=1}^{N} \sum_{t=1}^{T} \theta [Q_{1,it} \epsilon_{1,it} K_{h_1}(Z_{1,it} - z_1)] \right\} = V_{11,1} + V_{11,2}, \qquad (A.7)
$$

$$
\begin{aligned}
\text{where } \; V_{11,1} &= \frac{\theta^2}{n} Var\left\{ Q_{1,i1} \epsilon_{1,i1} K_{h_1}(Z_{1,i1} - z_1) \right\},\\
V_{11,2} &= 2\frac{\theta^2}{nT} \sum_{t=1}^{T-1} (T-t) Cov\Big(Q_{1,i1} \epsilon_{1,i1} K_{h_1}(Z_{1,i1} - z_1), Q_{1,i(t+1)} \epsilon_{1,i(t+1)} K_{h_1}(Z_{1,i(t+1)} - z_1) \Big).
\end{aligned}
$$

By Assumptions A.1 and A.2, and invoking similar steps to Cai and Li (2008) [Proof of Proposition 2], we obtain $nh_1^{d_1} V_{11,1} \to \theta^2 \tilde{f}_1(z_1) S_1^*$ and

$$
\begin{aligned}
&Cov\Big(Q_{1,i1} \epsilon_{1,i1} K_{h_1}(Z_{1,i1} - z_1), Q_{1,i(t+1)} \epsilon_{1,i(t+1)} K_{h_1}(Z_{1,i(t+1)} - z_1) \Big)\\
&\quad = E\Big\{ Q_{1,i1} Q_{1,i(t+1)}' \epsilon_{1,i1} \epsilon_{1,i(t+1)} K_{h_1}(Z_{1,i1} - z_1) K_{h_1}(Z_{1,i(t+1)} - z_1) \Big\}\\
&\quad \to f_{1,1(t+1)}(z_1, z_1) \begin{pmatrix} G_{1,t+1}^{(11,11)}(z_1, z_1) & \mathbf{0} \\ \mathbf{0}' & G_{1,t+1}^{(11,11)}(z_1, z_1) \otimes \mu_{1,2}(K_1^2) \end{pmatrix}.
\end{aligned}
$$

Hence, $V_{11,2} = O(n^{-1})$ and therefore by virtue of Assumption A.2 we obtain

$$
nh_1^{d_1} Var[T_{n,11}^*] \to \theta^2 \tilde{f}_1(z_1) S_1^*.
$$

For the second term in (A.6), and using Assumption A.1 we have,

$$
\begin{aligned}
E[T^*_{n,11}{T^*_{n,12}}'] &= \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}\theta\beta[Q_{1,it}Q'_{1,it}\epsilon_{1,it}\epsilon_{2,it}K^{3/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t\neq s=1}^{T}\theta\beta[Q_{1,it}Q'_{1,is}\epsilon_{1,it}\epsilon_{2,is}K_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_1}(Z_{1,is}-z_1)K^{1/2}_{h_2}(Z_{2,is}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t=1}^{T}\theta\beta[Q_{1,it}Q'_{1,lt}\epsilon_{1,it}\epsilon_{2,lt}K_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_1}(Z_{1,lt}-z_1)K^{1/2}_{h_2}(Z_{2,lt}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t\neq s=1}^{T}\theta\beta[Q_{1,it}Q'_{1,ls}\epsilon_{1,it}\epsilon_{2,ls}K_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_1}(Z_{1,ls}-z_1)K^{1/2}_{h_2}(Z_{2,ls}-z_2)] \\
&= O(N^{-1}) = o(1).
\end{aligned}
$$

To see this observe the following. The third and fourth summands in $E[T^*_{n,11}{T^*_{n,12}}']$ are zero by Assumption A.1. For the first summand in $E[T^*_{n,11}{T^*_{n,12}}']$, note that

$$
\begin{aligned}
&\frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}\theta\beta[Q_{1,it}Q'_{1,it}\epsilon_{1,it}\epsilon_{2,it}K^{3/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)] \\
&= \frac{\theta\beta}{n}E\begin{pmatrix} W_{1,it}W'_{1,it} & W_{1,it}W'_{1,it}\otimes(Z_{1,it}-z_1)'/h_1 \\ W'_{1,it}W_{1,it}\otimes(Z_{1,it}-z_1)/h_1 & W_{1,it}W'_{1,it}\otimes(Z_{1,it}-z_1)(Z_{1,it}-z_1)'/h_1^2 \end{pmatrix} \\
&\quad\times\epsilon_{1,it}\epsilon_{2,it}K^{3/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2) \\
&= \frac{\theta\beta}{n}\int\begin{pmatrix} \Omega^{12}_{11}(u_1,u_2) & \Omega^{12}_{11}(u_1,u_2)\otimes(u_1-z_1)'/h_1 \\ \Omega^{12}_{11}(u_1,u_2)'\otimes(u_1-z_1)/h_1 & \Omega^{12}_{11}(u_1,u_2)\otimes(u_1-z_1)(u_1-z_1)'/h_1^2 \end{pmatrix} \\
&\quad\times K^{3/2}_{h_1}(u_1-z_1)K^{1/2}_{h_2}(u_2-z_2)f(u_1,u_2)du_1du_2 \\
&= h_1^{-d_1/2}h_2^{d_2/2}\frac{\theta\beta}{n}\int\begin{pmatrix} \Omega^{12}_{11}(z_1+h_1c_1,z_2+h_2c_2) & \Omega^{12}_{11}(z_1+h_1c_1,z_2+h_2c_2)\otimes c'_1 \\ \Omega^{12}_{11}(z_1+h_1c_1,z_2+h_2c_2)'\otimes c_1 & \Omega^{12}_{11}(z_1+h_1c_1,z_2+h_2c_2)\otimes c_1c'_1/h_1^2 \end{pmatrix} \\
&\quad\times K^{3/2}_1(c_1)K^{1/2}_2(c_2)f(z_1+h_1c_1,z_2+h_2c_2)dc_1dc_2.
\end{aligned}
$$

For a fixed T and by invoking Assumptions A.2, A.3 and A.5, this first summand is $O(N^{-1})$. Similarly, the second summand in $E[T^*_{n,11}{T^*_{n,12}}']$ is $o(1)$.

For the fourth term in (A.6),

$$
\begin{aligned}
E[T^*_{n,12}{T^*_{n,12}}'] = Var(T^*_{n,12}) &= Var\left\{\frac{1}{n}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta[Q_{1,it}\epsilon_{2,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)]\right\} \\
&= \frac{\beta^2}{n}V_{12,1}+V_{12,2}, \\
\text{where}\quad V_{12,1} &= Var\left(Q_{1,i1}\epsilon_{2,i1}K^{1/2}_{h_1}(Z_{1,i1}-z_1)K^{1/2}_{h_2}(Z_{2,i1}-z_2)\right), \\
V_{12,2} &= 2\frac{\beta^2}{nT}\sum_{t=1}^{T-1}(T-t)Cov\Big(Q_{1,i1}\epsilon_{2,i1}K^{1/2}_{h_1}(Z_{1,i1}-z_1)K^{1/2}_{h_2}(Z_{2,i1}-z_2), \\
&\quad Q_{1,i(t+1)}\epsilon_{2,i(t+1)}K^{1/2}_{h_1}(Z_{1,i(t+1)}-z_1)K^{1/2}_{h_2}(Z_{2,i(t+1)}-z_2)\Big).
\end{aligned}
$$

$$
\begin{aligned}
V_{12,1} &= E\left(Q_{1,i1}Q_{1,i1}'\epsilon_{2,i1}^2 K_{h_1}(Z_{1,i1}-z_1)K_{h_2}(Z_{2,i1}-z_2)\right)\\
&= E\begin{pmatrix} W_{1,i1}W_{1,i1}' & W_{1,i1}W_{1,i1}'\otimes(Z_{1,i1}-z_1)'/h_1 \\ W_{1,i1}'W_{1,i1}\otimes(Z_{1,i1}-z_1)/h_1 & W_{1,i1}W_{1,i1}'\otimes(Z_{1,i1}-z_1)(Z_{1,i1}-z_1)'/h_1^2 \end{pmatrix}\\
&\qquad\qquad\qquad\qquad \times\epsilon_{2,it}^2 K_{h_1}(Z_{1,i1}-z_1)K_{h_2}(Z_{2,i1}-z_2)\\
&= \int\begin{pmatrix} \Omega_{11}^{22}(u_1,u_2) & \Omega_{11}^{22}(u_1,u_2)\otimes(u_1-z_1)'/h_1 \\ \Omega_{11}^{22}(u_1,u_2)'\otimes(u_1-z_1)/h_1 & \Omega_{11}^{22}(u_1,u_2)\otimes(u_1-z_1)(u_1-z_1)'/h_1^2 \end{pmatrix}\\
&\qquad\qquad\qquad\qquad \times K_{h_1}(u_1-z_1)K_{h_2}(u_2-z_2)f(u_1,u_2)du_1du_2\\
&\to f(z_1,z_2)\begin{pmatrix} \Omega_{11}^{22}(z_1,z_2) & \mathbf{0} \\ \mathbf{0}' & \Omega_{11}^{22}(z_1,z_2)\otimes\mu_{1,2}(K_1) \end{pmatrix}. \qquad \text{(A.8)}
\end{aligned}
$$

Then, for a fixed T, $V_{12,1} = O(N^{-1}) = o(1)$. In a similar manner, we obtain $V_{12,2} = o(1)$.

(ii) Note that by symmetry, $E[T_{n1}^* {T_{n2}^*}'] = E[T_{n2}^* {T_{n1}^*}']$. To compute $E[T_{n1}^* {T_{n2}^*}']$, we use the decomposition

$$
T_{n1}^* {T_{n2}^*}' = T_{n,11}^* {T_{n,21}^*}' + T_{n,11}^* {T_{n,22}^*}' + T_{n,12}^* {T_{n,21}^*}' + T_{n,12}^* T_{n,22}^*. \qquad \text{(A.9)}
$$

For the first term in (A.9), and by Assumptions A.1 and A.2,

$$
\begin{aligned}
E[T_{n,11}^* {T_{n,21}^*}'] &= \frac{1}{n^2}E\sum_{i=1}^N\sum_{t=1}^T \theta\beta[Q_{1,it}Q_{2,it}'\epsilon_{1,it}^2 K_{h_1}^{3/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)]\\
&+ \frac{1}{n^2}E\sum_{i=1}^N\sum_{t\neq s=1}^T \theta\beta[Q_{1,it}Q_{2,is}'\epsilon_{1,it}\epsilon_{1,is}K_{h_1}(Z_{1,it}-z_1)K_{h_1}^{1/2}(Z_{1,is}-z_1)K_{h_2}^{1/2}(Z_{2,is}-z_2)]\\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^N\sum_{t=1}^T \theta\beta[Q_{1,it}Q_{2,lt}'\epsilon_{1,it}\epsilon_{1,lt}K_{h_1}(Z_{1,it}-z_1)K_{h_1}^{1/2}(Z_{1,lt}-z_1)K_{h_2}^{1/2}(Z_{2,lt}-z_2)]\\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^N\sum_{t\neq s=1}^T \theta\beta[Q_{1,it}Q_{2,ls}'\epsilon_{1,it}\epsilon_{1,ls}K_{h_1}(Z_{1,it}-z_1)K_{h_1}^{1/2}(Z_{1,ls}-z_1)K_{h_2}^{1/2}(Z_{2,ls}-z_2)]\\
&= O(N^{-1}) = o(1).
\end{aligned}
$$

Similarly, for the second term in (A.9),

$$
\begin{aligned}
E[T_{n,11}^* {T_{n,22}^*}'] &= \frac{1}{n^2}E\sum_{i=1}^N\sum_{t=1}^T \theta\gamma[Q_{1,it}Q_{2,it}'\epsilon_{1,it}\epsilon_{2,it}K_{h_1}(Z_{1,it}-z_1)K_{h_2}(Z_{2,it}-z_2)]\\
&+ \frac{1}{n^2}E\sum_{i=1}^N\sum_{t\neq s=1}^T \theta\gamma[Q_{1,it}Q_{2,is}'\epsilon_{1,it}\epsilon_{2,is}K_{h_1}(Z_{1,it}-z_1)K_{h_2}(Z_{2,is}-z_2)]\\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^N\sum_{t=1}^T \theta\gamma[Q_{1,it}Q_{2,lt}'\epsilon_{1,it}\epsilon_{2,lt}K_{h_1}(Z_{1,it}-z_1)K_{h_2}(Z_{2,lt}-z_2)]\\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^N\sum_{t\neq s=1}^T \theta\gamma[Q_{1,it}Q_{2,ls}'\epsilon_{1,it}\epsilon_{2,ls}K_{h_1}(Z_{1,it}-z_1)K_{h_2}(Z_{2,ls}-z_2)]\\
&= o(1).
\end{aligned}
$$

For the third term in (A.9),

$$\begin{aligned}
E[T^*_{n,12}T^*_{n,21}{}'] &= \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}\beta^2[Q_{1,it}Q'_{2,it}\epsilon_{1,it}\epsilon_{2,it}K_{h_1}(Z_{1,it}-z_1)K_{h_2}(Z_{2,it}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t\neq s=1}^{T}\beta^2[Q_{1,it}Q'_{2,is}\epsilon_{1,is}\epsilon_{2,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2) \\
&\qquad \times K^{1/2}_{h_1}(Z_{1,is}-z_1)K^{1/2}_{h_2}(Z_{2,is}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t=1}^{T}\beta^2[Q_{1,it}Q'_{2,lt}\epsilon_{1,lt}\epsilon_{2,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2) \\
&\qquad \times K^{1/2}_{h_1}(Z_{1,lt}-z_1)K^{1/2}_{h_2}(Z_{2,lt}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t\neq s=1}^{T}\beta^2[Q_{1,it}Q'_{2,ls}\epsilon_{1,ls}\epsilon_{2,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2) \\
&\qquad \times K^{1/2}_{h_1}(Z_{1,ls}-z_1)K^{1/2}_{h_2}(Z_{2,ls}-z_2)] \\
&= o(1).
\end{aligned}$$

For the fourth term in (A.9),

$$\begin{aligned}
E[T^*_{n,12}T^*_{n,22}{}'] &= \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}\beta\gamma[Q_{1,it}Q'_{2,it}\epsilon^2_{2,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{3/2}_{h_2}(Z_{2,it}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t\neq s=1}^{T}\beta\gamma[Q_{1,it}Q'_{2,is}\epsilon_{2,it}\epsilon_{2,is}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)K^{1/2}_{h_2}(Z_{2,is}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t=1}^{T}\beta\gamma[Q_{1,it}Q'_{2,lt}\epsilon_{2,it}\epsilon_{2,lt}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)K^{1/2}_{h_2}(Z_{2,lt}-z_2)] \\
&+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t\neq s=1}^{T}\beta\gamma[Q_{1,it}Q'_{2,ls}\epsilon_{2,it}\epsilon_{2,ls}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)K^{1/2}_{h_2}(Z_{2,ls}-z_2)] \\
&= o(1).
\end{aligned}$$

(iii) To compute $E[T^*_{n2}T^*_{n2}{}']$, note that

$$T^*_{n2}T^*_{n2}{}' = T^*_{n,21}T^*_{n,21}{}' + T^*_{n,21}T^*_{n,22}{}' + T^*_{n,22}T^*_{n,21}{}' + T^*_{n,22}T^*_{n,22}. \tag{A.10}$$

For the first term in (A.10), we proceed as follows,

$$\begin{aligned}
E[T^*_{n,21}T^*_{n,21}{}'] = Var(T^*_{n,21}) &= \frac{1}{n^2}Var\left\{\sum_{i=1}^{N}\sum_{t=1}^{T}\beta Q_{2,it}\epsilon_{1,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)\right\} \\
&= \frac{\beta^2}{NT^2}Var\left\{\sum_{t=1}^{T}Q_{2,it}\epsilon_{1,it}K^{1/2}_{h_1}(Z_{1,it}-z_1)K^{1/2}_{h_2}(Z_{2,it}-z_2)\right\} \\
&= \frac{\beta^2}{n}V_{21,1} + V_{21,2},
\end{aligned}$$

$$\text{where } V_{21,1} = Var\left\{ Q_{2,it}\epsilon_{1,it}K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2) \right\},$$

$$V_{21,2} = \frac{\beta^2}{nT}\sum_{t=1}^{T-1}(T-t)Cov\Big(Q_{2,i1}\epsilon_{1,i1}K_{h_1}^{1/2}(Z_{1,i1}-z_1)K_{h_2}^{1/2}(Z_{2,i1}-z_2),$$
$$Q_{2,i(t+1)}\epsilon_{1,i(t+1)}K_{h_1}^{1/2}(Z_{1,i(t+1)}-z_1)K_{h_2}^{1/2}(Z_{2,i(t+1)}-z_2)\Big).$$

Using the steps in (A.8), we can show that

$$V_{21,1} \to f(z_1,z_2)\begin{pmatrix} \Omega_{22}^{11}(z_1,z_2) & \mathbf{0} \\ \mathbf{0}' & \Omega_{22}^{11}(z_1,z_2)\otimes\mu_{2,2}(K_2) \end{pmatrix}.$$

Hence, for a fixed T, $V_{21,1} = O(N^{-1})$. Similarly, we obtain $V_{21,2} = o(1)$.

For the second term in (A.10),

$$\begin{aligned} E[T^*_{n,21}T^{*\prime}_{n,22}] &= \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}\beta\gamma[Q_{2,it}Q'_{2,it}\epsilon_{1,it}\epsilon_{2,it}K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{3/2}(Z_{2,it}-z_2)] \\ &+ \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t\neq s=1}^{T}\beta\gamma[Q_{2,it}Q'_{2,is}\epsilon_{1,it}\epsilon_{2,is}K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)K_{h_2}(Z_{2,is}-z_2) \\ &+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t=1}^{T}\beta\gamma[Q_{2,it}Q'_{2,lt}\epsilon_{1,it}\epsilon_{2,lt}K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)K_{h_2}(Z_{2,lt}-z_2)] \\ &+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t\neq s=1}^{T}\beta\gamma[Q_{2,it}Q'_{2,ls}\epsilon_{1,it}\epsilon_{2,ls}K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2)K_{h_2}(Z_{2,ls}-z_2)] \\ &= o(1). \end{aligned}$$

For the fourth term in (A.10),

$$E[T^*_{n,22}T^{*\prime}_{n,22}] = Var(T^*_{n,22}) = \frac{1}{n^2}Var\left\{\sum_{i=1}^{N}\sum_{t=1}^{T}\gamma[Q_{2,it}\epsilon_{2,it}K_{h_2}(Z_{2,it}-z_2)]\right\}.$$

Similar to the above proof of $E[T^*_{n,11}T^{*\prime}_{n,11}]$, we can easily show that

$$nh_2^{d_2}Var(T^*_{n,22}) \to \gamma^2\tilde{f}_2(z_2)S_2^*.$$

In essence, the off-diagonal block terms for $E[T_n^*T_n^{*\prime}]$ in A.4 are of smaller order than its $(1,1)$ and $(2,2)$ main-diagonal block terms, which are of orders $O\{(nh_1^{d_1})^{-1}\}$ and $O\{(nh_2^{d_2})^{-1}\}$ respectively. Therefore, this completes the proof of Proposition 3.3. □

Proof of Theorem 3.4: We apply the Cramér-Wold device to assist in establishing asymptotic normality, given the multivariate nature of our semiparametric system estimator. We introduce some additional notations for ease of exposition. We define

$$
\begin{aligned}
\widetilde{A}_{it}^{-1} \;&:=\; \left(\begin{array}{cc} K_{h_1}^{1/2}(Z_{1,it}-z_1) & 0 \\ 0 & K_{h_2}^{1/2}(Z_{2,it}-z_2) \end{array}\right)\left(\begin{array}{cc} \theta & \beta \\ \beta & \gamma \end{array}\right)\left(\begin{array}{cc} K_{h_1}^{1/2}(Z_{1,it}-z_1) & 0 \\ 0 & K_{h_2}^{1/2}(Z_{2,it}-z_2) \end{array}\right) \\
&= \left(\begin{array}{cc} \theta K_{h_1}(Z_{1,it}-z_1) & \beta K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2) \\ \beta K_{h_1}^{1/2}(Z_{1,it}-z_1)K_{h_2}^{1/2}(Z_{2,it}-z_2) & \gamma K_{h_2}(Z_{2,it}-z_2) \end{array}\right)
\end{aligned}
$$

For any $\lambda \in \mathbb{R}^{\widetilde{l}}$ such that $\|\lambda\| = 1$, we set $\eta_{it} = \lambda'\widetilde{\mathcal{D}}^{1/2}Q_{it}\widetilde{A}_{it}^{-1}\epsilon_{it}$, where $Q_{it} = block\ diag(Q_{1,it}, Q_{2,it})$ and $\epsilon_{it} = (\epsilon_{1,it}, \epsilon_{2,it})'$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Thus, we have

$$n^{1/2}\lambda'\widetilde{\mathcal{D}}^{1/2}T_n^* = \frac{1}{\sqrt{n}}\sum_{i=1}^{N}\sum_{t=1}^{T}\eta_{it}.$$

By Assumption A.2 and Proposition 3.3, and for any $i = 1, \ldots, N$ and $t = 1, \ldots, T$, we obtain

$$Var(\eta_{it}) = \eta^2(z)\big(1 + o(1)\big), \text{ and } \sum_{t=2}^{T}\big|Cov(\eta_{i1}, \eta_{it})\big| = o(1),$$

where $\eta^2(z) := \lambda'\tilde{f}(z)D_{\theta\gamma}^2 S^*\lambda$. Thus, $Var(n^{1/2}\lambda'\widetilde{\mathcal{D}}^{1/2}T_n^*) = \eta^2(z)\big(1 + o(1)\big)$.

Continuing in this way, it remains to show that the Lyapounov condition holds. This is easily achieved by invoking the stipulated assumptions, Minkowski's inequality and similar steps to Proof of Theorem 2 in Cai & Li (2008). $\square$

Proof of Proposition 3.9: Note that

$$\widetilde{S}_n = \frac{1}{n}\left(\begin{array}{cc} Q_1'K\widetilde{U_1} & \mathbf{0} \\ \mathbf{0}' & Q_2'K\widetilde{U_2} \end{array}\right) := \left(\begin{array}{cc} \widetilde{S}_{n,11} & \mathbf{0} \\ \mathbf{0}' & \widetilde{S}_{n,22} \end{array}\right), \tag{A.11}$$

$$B_n = \frac{1}{n}\left(\begin{array}{c} Q_1'K\widetilde{X}_1\bar{R}_1 \\ Q_2'K\widetilde{X}_2\bar{R}_2 \end{array}\right) := \left(\begin{array}{c} B_{n,11} \\ B_{n,22} \end{array}\right), \tag{A.12}$$

$$R_n = \frac{1}{n}\left(\begin{array}{c} Q_1'K\widetilde{X}_1R_1 \\ Q_2'K\widetilde{X}_2R_2 \end{array}\right) := \left(\begin{array}{c} R_{n,11} \\ R_{n,22} \end{array}\right), \tag{A.13}$$

$$T_n^* = \frac{1}{n}\left(\begin{array}{c} Q_1'K\epsilon_1 \\ Q_2'K\epsilon_2 \end{array}\right) := \left(\begin{array}{c} T_{n,1}^* \\ T_{n,2}^* \end{array}\right). \tag{A.14}$$

Then $E[\widetilde{S}_n], E[B_n]$, and $E[R_n]$ follow directly from the results in Cai & Li (2008), and we have the desired result for (i), (ii) and (iii) of Proposition 3.9. For Proposition 3.9 (iv), note that

$$Var(T_n^*) = E\left(\begin{array}{cc} T_{n,1}^*{T_{n,1}^*}' & T_{n,1}^*{T_{n,2}^*}' \\ T_{n,2}^*{T_{n,1}^*}' & T_{n,2}^*{T_{n,2}^*}' \end{array}\right), \tag{A.15}$$

and $T_{n,1}^*{T_{n,1}^*}' = \frac{1}{n^2}Q_1'K\epsilon_1\epsilon_1'KQ_1$, $T_{n,1}^*{T_{n,2}^*}' = \frac{1}{n^2}Q_1'K\epsilon_1\epsilon_2'KQ_2$, and $T_{n,2}^*{T_{n,2}^*}' = \frac{1}{n^2}Q_2'K\epsilon_2\epsilon_2'KQ_2$. Using the results in Cai & Li (2008), it is easy to show that

$$nh^dE[T_{n,1}^*{T_{n,1}^*}'] = f(z)S_1^* \text{ and } nh^dE[T_{n,2}^*{T_{n,2}^*}'] = f(z)S_2^*.$$

Thus, it suffices to show that the off-diagonal block terms in (A.15) are also of the order of magnitude $n^{-1}h^{-d}$. We only consider the (1,2) block-entry in (A.15), as the result for the (2,1) block-entry will follow by virtue of symmetry. To begin, we express $E[T^*_{n,1}T^{*\prime}_{n,2}]$ as

$$\begin{aligned} E[T^*_{n,1}T^{*\prime}_{n,2}] &= \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}[Q_{1,it}Q'_{2,it}\epsilon_{1,it}\epsilon_{2,it}K_h^2(Z_{it}-z_1)] \\ &+ \frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t\neq s=1}^{T}[Q_{1,it}Q'_{2,is}\epsilon_{1,it}\epsilon_{2,is}K_h(Z_{it}-z_1)K_h(Z_{is}-z_2)] \\ &+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t=1}^{T}[Q_{1,it}Q'_{2,lt}\epsilon_{1,it}\epsilon_{2,lt}K_h(Z_{it}-z_1)K_h(Z_{lt}-z_2)] \\ &+ \frac{1}{n^2}E\sum_{i\neq l=1}^{N}\sum_{t\neq s=1}^{T}[Q_{1,it}Q'_{2,ls}\epsilon_{1,it}\epsilon_{2,ls}K_h(Z_{it}-z_1)K_h(Z_{ls}-z_2)]. \end{aligned} \tag{A.16}$$

The third and fourth terms in (A.16) are zero by Assumption A.1. For the first term in (A.16), we have

$$\begin{aligned} &\frac{1}{n^2}E\sum_{i=1}^{N}\sum_{t=1}^{T}[Q_{1,it}Q'_{2,it}\epsilon_{1,it}\epsilon_{2,it}K_h^2(Z_{it}-z_1)] \\ &= \frac{1}{n}E[Q_{1,it}Q'_{2,it}\epsilon_{1,it}\epsilon_{2,it}K_h^2(Z_{it}-z_1)] \\ &= \frac{1}{n}E\begin{pmatrix} W_1W_2' & W_1W_2'\otimes(Z-z)'/h \\ W_2W_1'\otimes(Z-z)/h & W_1W_2'\otimes(Z-z)(Z-z)'/h^2 \end{pmatrix}\epsilon_1\epsilon_2K_h^2(Z-z) \\ &= \frac{1}{n}E\begin{pmatrix} \Omega_{12}^{12}(Z) & \Omega_{12}^{12}(Z)\otimes(Z-z)'/h \\ \Omega_{12}^{12}(Z)'\otimes(Z-z)/h & \Omega_{12}^{12}(Z)\otimes(Z-z)(Z-z)'/h^2 \end{pmatrix}K_h^2(Z-z) \\ &= \frac{1}{n}\int\begin{pmatrix} \Omega_{12}^{12}(u) & \Omega_{12}^{12}(u)\otimes(u-z)'/h \\ \Omega_{12}^{12}(u)'\otimes(u-z)/h & \Omega_{12}^{12}(u)\otimes(u-z)(u-z)'/h^2 \end{pmatrix}K_h^2(u-z)f(u)du \\ &= \to \frac{1}{nh^d}f(z)\begin{pmatrix} \Omega_{12}^{12}(z)\nu_0 & \mathbf{0} \\ \mathbf{0}' & \Omega_{12}^{12}(z)\otimes\mu_2(K^2) \end{pmatrix} := \frac{1}{nh^d}S^*_{12} \end{aligned}$$

Hence, this completes the proof of Proposition 3.9. □

Proof of Theorem 3.11: This is straightforward given the above results in the proofs of Theorem 3.4 and Proposition 3.9, and the results in Cai & Li (2008). □

References

Alfaro, L., Chanda, A., Kalemli-Ozcan, S. & Sayek, S. (2004), 'FDI and economic growth: The role of local financial markets', *Journal of International Economics* **64**, 89–112.

Alfaro, L. & Johnson, M. S. (2013), 'Foreign direct investment and growth', *The Evidence and Impact of Financial Globalization* pp. 299–309.

Balasubramanyam, V. N., Salisu, M. & Sapsford, D. (1996), 'Foreign direct investment and growth in EP and IS countries', *Economic Journal* **106**, 92–105.

Borensztein, E., De Gregorio, J. & Lee, J. W. (1998), 'How does foreign direct investment affect economic growth?', *Journal of International Economics* **45**, 115–135.

Cai, Z., Das, M., Xiong, H. & Wu, X. (2006), 'Functional coefficient instrumental variables models', *Journal of Econometrics* **133**, 207–241.

Cai, Z. & Li, Q. (2008), 'Nonparametric estimation of varying coefficient dynamic panel data models', *Econometric Theory* **24**, 1321–1342.

Cai, Z. & Xiong, H. (2012), 'Partially varying coefficient instrumental variables models', *Statistica Neerlandica* **66**, 85–110.

Carkovic, M. & Levine, R. (2005), 'Does foreign direct investment accelerate economic growth?', *Does Foreign Direct Investment Promote Economic Development* .

Das, M. (2005), 'Instrumental variables estimators for nonparametric models with discrete endogenous variables', *Journal of Econometrics* **124**, 335–361.

Delgado, M. S., McCloud, N. & Kumbhakar, S. C. (2014), 'A generalized empirical model of corruption foreign direct investment and growth', *Journal of Macroeconomics* **42**, 298–316.

Durham, J. B. (2004), 'Absorptive capacity and the effects of foreign direct investment and equity foreign portfolio investment on economic growth', *European Economic Review* **48**, 285–306.

Durlauf, S. N. (2001), 'Manifesto for a growth econometrics', *Journal of Econometrics* **100**, 65–69.

Durlauf, S. N., Kourtellos, A. & Tan, C. M. (2008), 'Are any growth theories robust?', *Economic Journal* **118**, 329–346.

Fan, J. & Gijbels, I. (1996), *Local Polynomial Modeling and its Applications*, New York: Chapman and Hall.

Hansen, L. P. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica* **50**, 1029–1054.

Henderson, D. J., Kumbhakar, S. C., Li, Q. & Parmeter, C. F. (2015), 'Smooth coefficient estimation of a seemingly unrelated regression', *Journal of Econometrics* **189**, 148–162.

Henderson, D. J., Kumbhakar, S. C. & Parmeter, C. F. (2012), 'A simple method to visualize results in nonlinear regression models', *Economics Letters* **117**, 578–581.

Henderson, D. J., Papageorgiou, C. & Parmeter, C. F. (2012), 'Growth empirics without parameters', *Economic Journal* **122**, 125–154.

Henderson, D. J. & Parmeter, C. F. (2015), *Applied Nonparametric Econometrics*, Cambridge University Press.

Heston, A., Summers, R. & Aten, B. (2012), 'Penn World Table Version 7.1', Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Huynh, K. P. & Jacho-Chávez, D. T. (2009*a*), 'Growth and governance: A nonparametric analysis', *Journal of Comparative Economics* **37**, 121–143.

Huynh, K. P. & Jacho-Chávez, D. T. (2009*b*), 'A nonparametric quantile analysis of growth and governance', *Advances in Econometrics* **25**, 193–221.

Kottaridi, C. & Stengos, T. (2010), 'Foreign direct investment, human capital and nonlinearities in economic growth', *Journal of Macroeconomics* **32**, 858–871.

Li, Q., Huang, C. J., Li, D. & Fu, T. (2002), 'Semiparametric smooth coefficient models', *Journal of Business and Economic Statistics* **20**, 412–422.

Li, Q., Maasoumi, E. & Racine, J. S. (2009), 'A nonparametric tests for equality of distributions with mixed categorical and continuous data', *Journal of Econometrics* **148**, 186–200.

Li, Q. & Racine, J. S. (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.

Li, Q. & Racine, J. S. (2010), 'Smooth varying coefficient estimation and inference for qualitative and quantitative data', *Econometric Theory* **26**, 1607–1637.

Li, X. & Liu, X. (2005), 'Foreign direct investment and economic growth: An increasingly endogenous relationship', *World Development* **33**, 393–407.

Mauro, P. (1995), 'Corruption and growth', *Quarterly Journal of Economics* **110**, 681–712.

Mauro, P. (1998), 'Corruption and the composition of government expenditure', *Journal of Public Economics* **69**, 263–279.

McCloud, N. & Kumbhakar, S. C. (2012), 'Institutions, foreign direct investment, and growth: A hierarchical Bayesian approach', *Journal of the Royal Statistical Society: Series A* **175**, 83–105.

Minier, J. (2007), 'Institutions and parameter heterogeneity', *Journal of Macroeconomics* **29**, 595–611.

Newey, W. K. (1990), 'Efficient instrumental variables estimation of nonlinear models', *Econometrica* **58**, 809–837.

Papaioannou, E. (2009), 'What drives international financial flows? Politics, institutions and other determinants', *Journal of Development Economics* **88**, 269–281.

Political Risk Services Group (2010), 'International Country Risk Guide', `http://www.prsgroup.com/icrg.aspx`.

Racine, J. S. & Li, Q. (2004), 'Nonparametric estimation of regression functions with both categorical and continuous data', *Journal of Econometrics* **119**, 99–130.

Racine, J. S. & Parmeter, C. F. (2013), 'Data driven model evaluation: A test for revealed performance', *Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics* pp. 308–345.

Tran, K. C. & Tsionas, E. G. (2009), 'Local GMM estimation of semiparametric panel data with smooth coefficient models', *Econometric Reviews* **29**, 39–61.

Welsh, A. H. & Yee, T. W. (2006), 'Local regression for vector responses', *Journal of Statistical Planning and Inference* **136**, 3007–3031.

World Bank (2012), 'World Development Indicators [CD-ROM]', Washington, D.C.

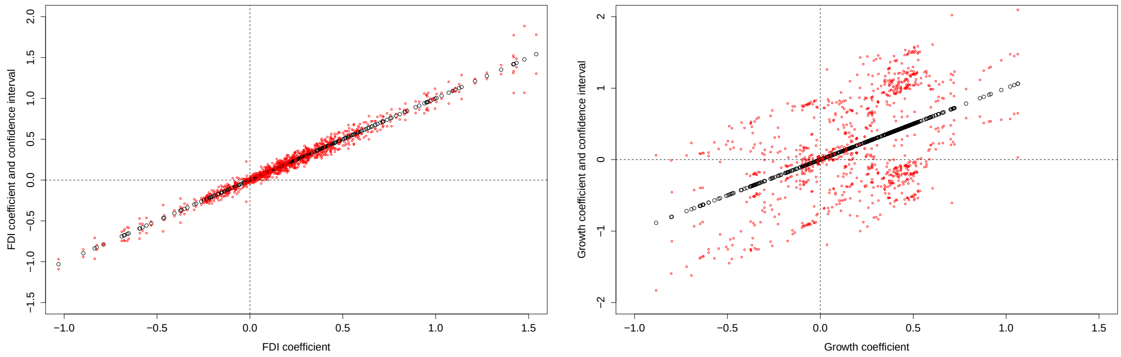


Figure 1: Significance plots for FDI inflows and growth coefficient functions for the model with instrumental variables.

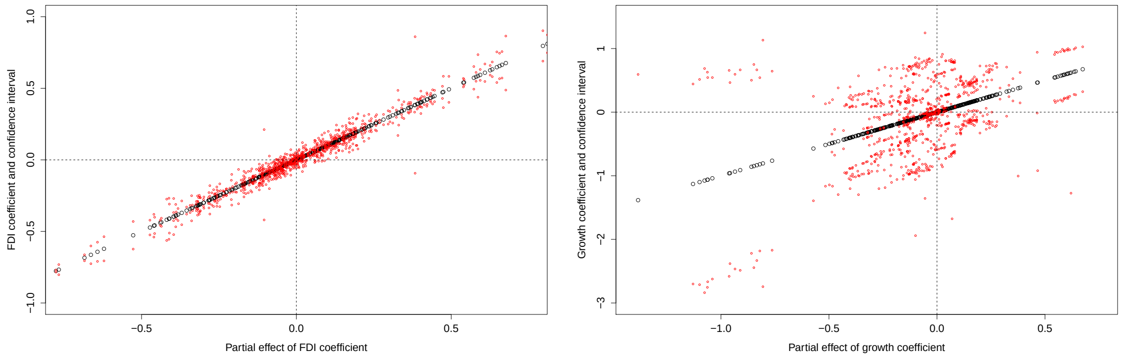
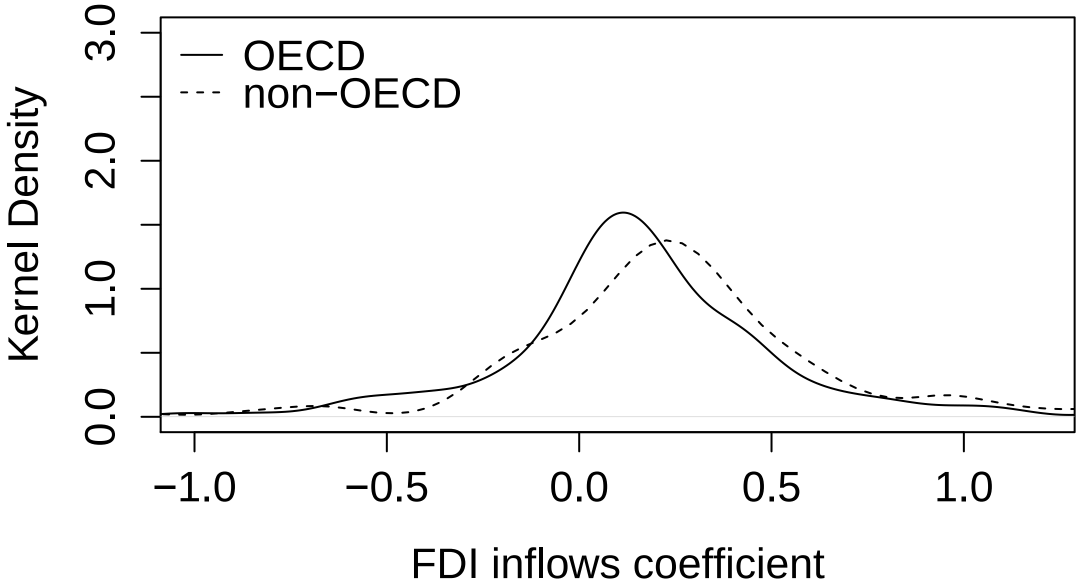


Figure 2: Significance plots for the partial effect of FDI inflows and growth coefficient functions with respect to corruption for the model with instrumental variables.



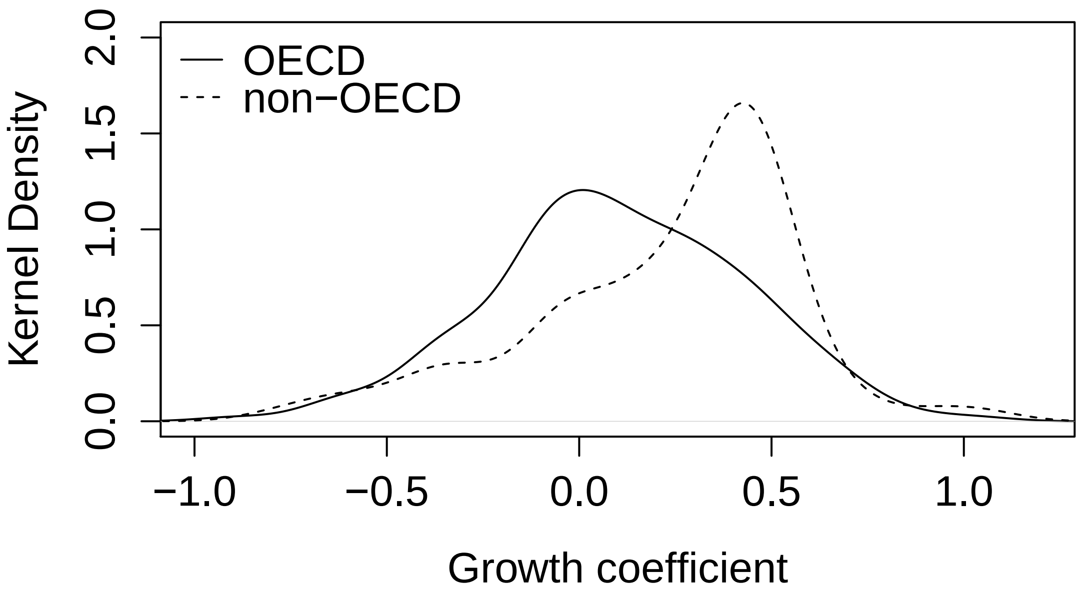


Figure 3: Kernel density plots of FDI inflows and growth coefficient function estimates for OECD and non-OECD countries based on the model with instrumental variables.

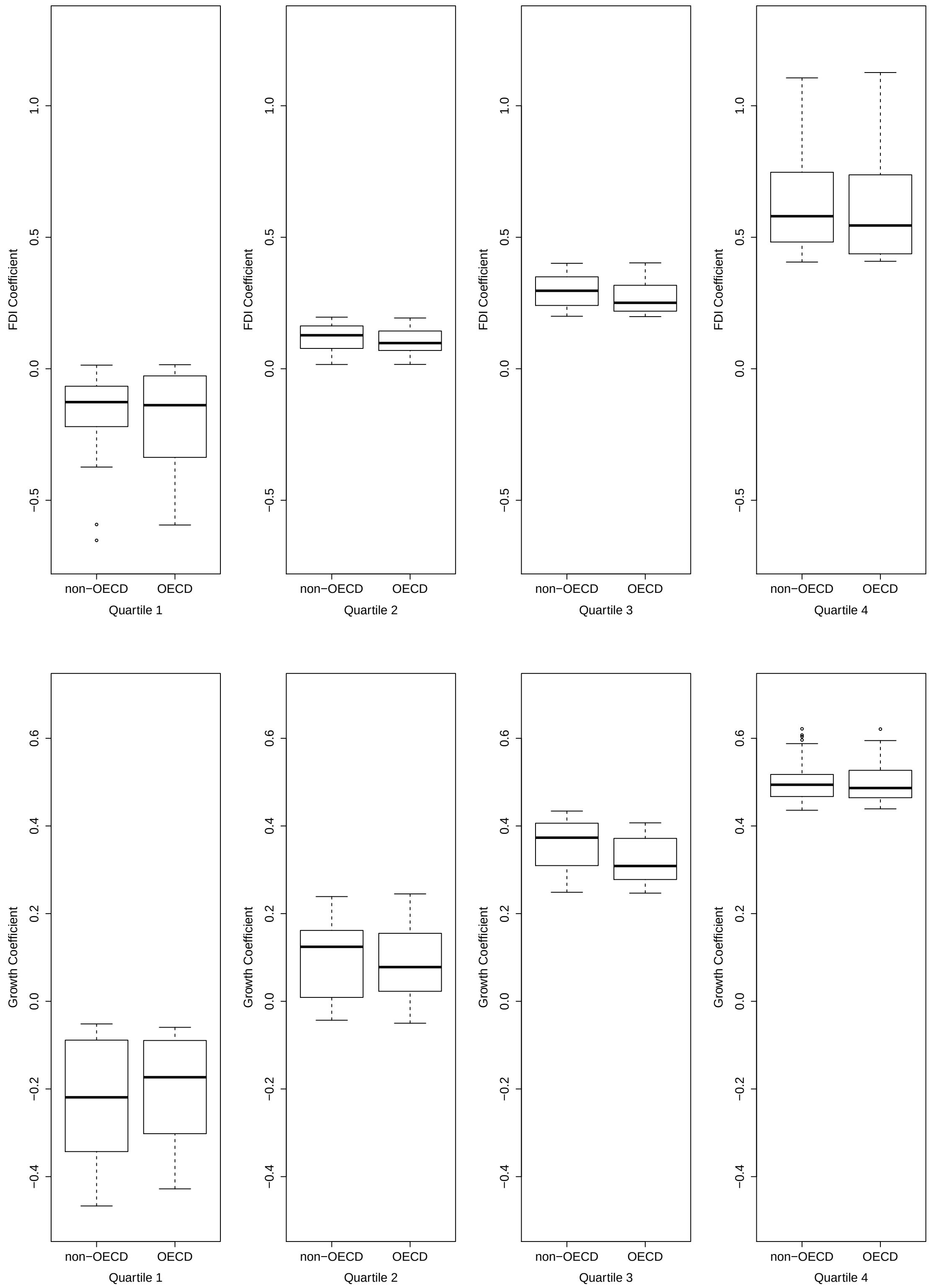


Figure 4: Box plots across quartiles of FDI inflows and growth coefficient function estimates for OECD and non-OECD countries based on the model with instrumental variables.

Table 1: Characterizing the Types of Interactions between Economic Growth and FDI

		Growth Effect		
		Positive	Negative	Zero
FDI Effect	Positive	**Symbiosis**: 63 countries (45 distinct)	**FDI-Antagonistic Symbiosis**: 3 *countries*^{d}	**FDI-Commensalism**: 37 countries (21 distinct)
	Negative	**Growth-Antagonistic Symbiosis**: 11 countries (5 distinct)	**Synnercrosis**: 3 *countries*^{d} (1 distinct)	**Growth-Ammensalism**: 4 countries (3 distinct)
	Zero	**Growth-Commensalism**: 1 *country*^{d} (not distinct)	**FDI-Ammensalism**: No Countries	**non-Symbiosis**: 2 *Countries*^{d}

1. To characterize the types of interactions between growth and FDI on the basis of the taxonomy in Definition 2.1, we use the following criterion: a country is placed in the category, for example, *symbiosis* if at least 50 percent of its estimated effects of FDI on growth *and* its estimated effects of growth on FDI is *positive* and *statistically significant* at least at the 5% level.
2. *d* are cells with *only* developing countries.
3. *distinct* refers to countries that exhibit *only* one type of interaction on the basis of our criterion.