A spatial nonparametric analysis of local multipliers

Margherita Gerolimetto and Stefano Magrini
Department of Economics
Ca' Foscari University Venice - Italy
e-mail: margherita.gerolimetto@unive.it s.magrini@unive.it

Preliminary draft – Please do not quote without authors' permission

Abstract

In this paper we present a spatial nonparametric analysis of local multipliers. In particular, we first present a spatial equilibrium model that clarifies the conditions under which the local nontradable multiplier exists and, at the same time, suggests that the multiplier of a city might depend on the size of the shock occurring to the local tradable sector in a nonlinear fashion as well as on shocks in other cities of the system. We then estimate the effect of an exogenous shock to tradable employment on the employment in the nontradable sector and, in order to achieve this, we first develop a two-step nonparametric estimator that allows for spatial effects. Our analysis shows that spatial effects must be allowed for and, when this is done, the local nontradable multiplier depends nonlinearly on the size of the shock to the tradable sector.

Keywords: Local Multipliers, Spatial dependence

1 Introduction

State and local governments employ large amounts of public funds in economic development policies aimed at attracting and fostering new economic activities or at retaining existing ones. The outcomes of these efforts could be new or more stable jobs, higher income and wealth and an improved tax base. Indeed, it appears that job creation has actually represented the primary goal sought after by policy makers (Eberts, 2005) possibly because this may also lead to fiscal benefits in the form of an increase of tax revenue net of public expenditure. In fact, due to multiplier effects, the total increase in local jobs can be greater than the increase in jobs in assisted businesses when these businesses produce tradable goods. For this reason, most of local development interventions are actually targeted to the tradable sector.

In the US, a precise account of the total amount of resources involved in these activities is almost an impossible task given the large number of agencies involved and the even larger number of policies being implemented. However, a survey of state-funded programs conducted in 1998 (Poole *et al.*, 1999), calculated that the states allocated approximately $4.6 billion in tax incentive programs and $6.3 billion in non-tax incentives. This figure excludes tax and other financial incentives, as well as job training and infrastructure incentives provided by local (substate) governments; it also excludes local development efforts carried out under the leadership of non-governmental organizations. Eberts (2005) reports an overall estimate of $30 billion a year devoted to local development initiatives through direct and indirect funds which implies more than $2,000 per targeted job.

In a recent paper, Moretti (2010) proposes a simple methodology to assess the effectiveness of local development interventions in creating new jobs. The theoretical framework that underlies the empirical analysis builds on the traditional general equilibrium setting by Rosen (1979) and Roback (1982). Differently from the original Rosen-Roback set-up, in Moretti's framework local shocks to local labor markets are not necessarily fully capitalized in the price of land as the local supply of labor is not necessarily infinitely elastic and the local housing supply in not necessarily perfectly inelastic (see also Moretti, 2011). Inside a city, a positive shock to the labor demand of an industry within the tradable sector, possibly induced by a local development initiative, has positive indirect effects on employment in the nontradable sector as well as in other industries within the tradable sector. There are however offsetting general equilibrium effects that pass through local labor and housing markets. In particular, if the elasticity of the local supply of labor is not elastic (which, in turn, depends on the degree of geographical mobility of workers and on the elasticity of housing supply), the initial increase in employment crowds out employment in other industries due to a local increase in real wages. Local multipliers are thus the net effect of indirect and general equilibrium effects.

In the present paper, we develop a spatial equilibrium model that shares many features with Moretti's analysis while following in its set-up the models by Merrifield (1987 and 1990) and McGregor *et al.* (2000). In particular, our model shows that a positive exogenous productivity shock in the tradable sector of a city may trigger a nontradable multiplier in the same city. This same model, however, suggests the existence of other interesting effects: on the one hand, a local nontradable multiplier might as well be triggered by shocks occurring to other cities of the system; on the other hand, the size of the local nontradable multiplier might nonlinearly depend on the size of the shock.

The effects suggested by the theoretical model might have very important implications for the empirical analysis. To establish the magnitude of the local nontradable multiplier, Moretti estimates a very simple model

$$y = \alpha + \beta x + \epsilon \tag{1}$$

in which y and x are the change in the log number of jobs in the nontradable and tradable sectors respectively. Estimates of β can thus be interpreted as elasticities and from these it also possible to derive, for each new job in the tradable sector, the number of additional jobs created in the nontradable sector. These estimates are obtained either via OLS or, in order to deal with endogeneity concerns, via instrumental variables estimation. According the latter estimates, approximately 1.6 additional jobs are created in the nontradable sector of a city for each new job in the tradable sector of the same city.[1]

[1] De Blasio and Menon (2011) apply this same framework to data on Italian local labour markets but find no significant effects on nontradable industries.

Despite its simplicity, this approach appears to be able to provide rather meaningful estimates. One particularly appealing feature is that the exogenous variation is directly attributed to the tradable sector which is the one that attracts most of the local development initiatives (De Blasio and Menon, 2011). More generally, as emphasized by Moretti, this approach represents a valid alternative to the traditional methodology, local Input-Output analysis, that tends to overlook the employment effect for nontradables as well as the offsetting general equilibrium effects.

According to our model, however, in addition to endogeneity, two other crucial issues must be addressed in the empirical analysis in order to avoid misleading results. Firstly, the underlying relation between exogenous and induced variations in local employment is not necessarily linear. Secondly, within our framework space clearly plays a key role through labor flows. Focussing on this second issue, it must be recognized that, from an econometric point of view, neglecting space might lead to omit spatially structured covariates thus running the risk of obtaining biased estimates. Under these circumstances, identifying valid instrumental variables could be extremely difficult as the omitted factors are often unobservable. In addition, from a more interpretative point of view, ignoring space in the empirical model means overlooking that the exogenous variation that benefits one city does not necessarily come exclusively from the tradable sector of the same city. Consider, for example, a city in which there has been no local exogenous variation in the tradable sector. Implicit in model (1), no changes in nontradable sector employment should be observed. However, such changes could be induced by exogenous variations in the tradable sectors of other cities, with an intensity that is possibly negatively correlated with relative distance. To the extent in which this happens, the local multiplier estimated through model (1) might be biased. In particular, a positive (negative) effect on nontradable sector employment from additional tradable jobs in neighboring cities leads to overestimate (underestimate) the local multiplier.

To estimate model (1) we opt for a nonparametric approach. Operatively, to do so we develop a two-step nonparametric regression estimator that, moving from the standard Local Linear regression Estimator (LLE), allows for spatial effects. More in details, we draw on the work by Martins-Filho and Yao (2009) who establish a set of sufficient conditions for the asymptotic normality of the LLE estimator and propose a two step procedure for nonparametric regression with spatially dependent data that does not require *a priori* parametric assumptions on spatial dependence. Information on its structure is in fact drawn from a nonparametric estimate of the errors spatial covariance matrix. The finite sample performance of this estimator is then assessed via an extensive Monte Carlo experiment. Finally, to address endogeneity concerns, we include this new estimator in both steps of the nonparametric Instrumental Variable estimator proposed by Newey *et al.* (1999).

The structure of the paper is as follows. In Section 2 we present our economic model. In Section 3 we introduce our new nonparametric estimator as well as the Monte Carlo experiment. In Section 4 we report the empirical analysis and Section 5 concludes.

2 The economic model

2.1 General set up

Consider two cities (A and B) within a system of many. Each city features two sectors: one produces consumption goods which are exported to the rest of the system and the other produces goods which are consumed locally. Residents in each city maximize utility obtained by spending all their income either on goods imported from the rest of the system or on locally produced ones. There are two factors of production, capital and labor; the supply of capital is assumed to be infinitely elastic at an exogenously determined rental rate, labor supply is instead an increasing function of real consumption wage due to migrations between the two cities. Mobility across cities is not perfect as a consequence of idiosyncratic preferences for location. Similar to most models of this type, city income is total wage income as capital is assumed to be owned and produced outside the cities. Production takes place in both sectors according to a linearly homogeneous production function with constant elasticity of substitution. In addition, both sectors are in perfect competition thus implying profit maximization and zero profits.

Similar to Merrifield (1987, 1990) and McGregor *et al.* (2000) we first present the fundamental relationships in levels and then solve the model in proportionate-change terms.[2] In the exposition we shall focus on the relationships characterizing city Abut an analogous set of equations holds for city B with the only exception that in the latter no exogenous efficiency shock occurs in the tradable sector. Demands for tradable and nontradable goods produced in city A are

$$q_1^A = q_1(p_1^A, p_1^x)$$

$$q_2^A = q_2(p_2^A, n^A \cdot w^A)$$

where p_1^A is the price of the tradable good produced in city A, p_1^X is the price of the tradable good in the external market, p_2^A is the price of the nontradable good in city A, n^A is the population and w^A is the wage rate in A. The price of the tradable good is

$$p_1^A = p_1(w^A, e^A)$$

where e^A is the efficiency level in the tradable sector. The price of the nontradable good is given by

$$p_2^A = p_2(w^A)$$

that enters the consumer price index of city A

$$c^A = c(p_2^A)$$

Note that the price index depends only on the price of the nontradable good; this is because the tradable good produced in A is fully exported and the (constant) price of the imported tradable good, p_1^X, is taken as the numeraire.

[2] The notation adopted here borrows heavily from the set up in McGregor *et al.*, 2000.

Given cost-minimization behavior of firms, labor demands in the two sectors are

$$n_1^{dA} = n_1^d \left(q_1^A, e^A, \frac{w^A}{p_1^A} \right)$$

$$n_2^{dA} = n_2^d \left(q_2^A, \frac{w^A}{p_2^A} \right)$$

while aggregate labor demand in city $A$is

$$n^{dA} = n_1^{dA} + n_2^{dA}$$

Labor supply in city A is represented by

$$n^{sA} = n(V^A, V^B, L^{AB}, \bar{n}) \quad (2)$$

where V^A and V^B are indirect utility levels in the two cities, L^{AB} is a parameter describing the relevance of idiosyncratic preferences for location in mobility decisions[3] and $\bar{n} = n^A + n^B$ is the total number of individuals residing in the two cities assumed, as in Moretti (2011), to be constant. In order to understand the logic behind equation (2), in line with Moretti (2010) (see also Maier and Fischer, 1985) we consider a random utility framework in which the indirect utility of individual i in city A is

$$V^{iA} = V \left(V_d(w^A, c^A), exp(\epsilon^{iA}) \right)$$

where V_d is a deterministic component (which might also include local amenities as in the Rosen-Roback framework; Rosen, 1982, Roback, 1979) while $exp(\epsilon^{iA})$ is the random component representing individual's i idiosyncratic preferences for city A.

Assuming a Cobb-Douglas specification for the utility function, letting ϕ denote the (constant, both in time and across cities) share of expenditure on nontradables and taking logs we obtain

$$log\, V^{iA} = log\, w^A - \phi log\, c^A + \epsilon^{iA} \quad (3)$$

[3] This parameter can be supposed to depend also on the physical distance between alternative locations

Under the utility maximization rule, individual i will choose $A$with probability

$$P^A = P(log\, V^{iA} - log\, V^{iB}) = P(\epsilon^{iB} - \epsilon^{iA} < log\, V^{iA} - log\, V^{iB})$$

For *iid* Gumbell distributed random terms ϵ^i and under the Independence of Irrelevant Alternatives) hypothesis, the difference $\epsilon^{iB} - \epsilon^{iA}$ follows a logistic distribution with scale parameter (L^{AB}) and variance $(L^{AB})^2\pi^2/3$. The probability of choosing city A is

$$P^A = \frac{exp\left(\frac{log\, V^{iA}}{L^{AB}}\right)}{exp\left(\frac{log\, V^{iA}}{L^{AB}}\right) + exp\left(\frac{log\, V^{iB}}{L^{AB}}\right)}$$

or, using equation (3) and the corresponding one for city B,

$$P^A = \frac{\left(\frac{w^A}{c^{A\phi}}\right)^{\frac{1}{L^{AB}}}}{\left(\frac{w^A}{c^{A\phi}}\right)^{\frac{1}{L^{AB}}} + \left(\frac{w^B}{c^{B\phi}}\right)^{\frac{1}{L^{AB}}}}$$

and

$$P^B = \frac{\left(\frac{w^B}{c^{B\phi}}\right)^{\frac{1}{L^{AB}}}}{\left(\frac{w^A}{c^{A\phi}}\right)^{\frac{1}{L^{AB}}} + \left(\frac{w^B}{c^{B\phi}}\right)^{\frac{1}{L^{AB}}}}$$

In equilibrium, the expected number of individuals moving from A to B and *vice-versa* must be equal thus leading to the following spatial equilibrium condition

$$P^A \cdot n^B = P^B \cdot n^A$$

From this, substituting the just derived probabilities P^A and P^B, we obtain

$$n^A = n^B \left[\frac{w^A}{w^B}\left(\frac{c^A}{c^B}\right)^{-\phi}\right]^{\frac{1}{L^{AB}}} \tag{4}$$

which underlies the labor supply equation (2).

Finally, the labor market clearing condition in city A

$$n^{dA} = n^{sA} = n^A \tag{5}$$

closes the model.

As anticipated, we solve the model in proportionate change terms as this allows linearization and the interpretation of the parameters of the model as elasticities or shares.[4] The growth rate of the demand for tradable goods produced in city A is then represented by

$$\dot{q}_1^A = -\eta^A(\dot{p}_1^A) \tag{6}$$

where η^A is the price elasticity of demand for city's A exports. In spatial equilibrium models within the Rosen-Roback framework, it is often assumed that the city or region is a small supplier within a perfectly competitive interregional or international market for tradable goods and, consequently, a single price is imposed to this market. In line with McGregor *et al.* (2000), we instead consider a more general setting in which the price of tradable goods realized in city A can differ from the price of tradable goods realized elsewhere; clearly, the single price assumption represents a special case of this setting that applies when $\eta \to \infty$. However, here we explicitly consider the situation in which the interregional demand for tradable goods is highly elastic and consequently we assume that η takes on a large (although not infinite) value.

Given our assumption about the utility functional form, the growth rate of the demand for nontradable goods in city A is

$$\dot{q}_2^A = -\dot{p}_2^A + \dot{n}^A + \dot{w}^A \tag{7}$$

Assuming linearly homogeneous CES production functions, the growth rates of labor demand in the two sectors are

$$\dot{n}_1^{dA} = \dot{q}_1^A - \dot{e}^A - \sigma_1^A(\dot{w}^A - \dot{e}^A - \dot{p}_1^A) \tag{8}$$

$$\dot{n}_2^{dA} = \dot{q}_2^A - \sigma_2^A(\dot{w}^A - \dot{p}_2^A) \tag{9}$$

[4] The dot notation symbolizes variables in proportionate change terms.

where σ_1^A and σ_2^A are the factor elasticities of substitution in the production of, respectively, tradable and nontradable goods and $\dot{e}^A$ is the labor-augmenting efficiency increase in the production of tradable goods in city A arising from local economic development policies.

Given firms' cost minimization behavior, the growth rates of prices are

$$\dot{p}_1^A = \alpha_1^A(\dot{w}^A - \dot{e}^A) \tag{10}$$

$$\dot{p}_2^A = \alpha_2^A(\dot{w}^A) \tag{11}$$

in which α_1^A and α_2^A are, respectively, the share of labor in the output of tradable and nontradable sectors in city A.

Denoting with β^j, $j = A, B$ the initial proportion of labor force employed in the tradable sector of city j, the proportional change in total labor demand in city A is then given by

$$\dot{n}^{dA} = \beta^A \dot{n}_1^{dA} + (1 - \beta^A)\dot{n}_2^{dA} \tag{12}$$

Since the share of labor in the production of nontradable goods is

$$\alpha_2 = \frac{n_2 w}{p_2 q_2}$$

we can show that the growth rate of the consumer price index in city A is

$$\dot{c}^A = \left[\frac{1 - \beta^A}{\alpha_2^A}\right]\dot{p}_2^A \tag{13}$$

Now, letting λ be the initial proportion of city A's population in the overall (given) number of individuals residing in the two cities, i.e.

$$\lambda = \frac{n^A}{\bar{n}}$$

from equation (4) it is possible to derive the proportionate change in labor supply in city A

$$\dot{n}^{sA} = \frac{1 - \lambda}{L^{AB}}(\dot{w}_A - \dot{w}_B) - \phi\frac{1 - \lambda}{L^{AB}}(\dot{c}_A - \dot{c}_B) \tag{14}$$

Finally, city A's labor market clearing condition in proportionate change terms is

$$\dot{n}^{dA} = \dot{n}^{sA} = \dot{n}^{A} \tag{15}$$

2.2 Model's equilibrium

At the outset, we must stress that the model's solution reached in the present Section applies to the case of small disturbances whereby λ, β^A and β^B can be assumed to remain constant. We shall see later how the equilibrium multiplier effect is modified when this assumption is removed.

We start solving the model by concentrating on the labor market equilibrium of city A. In particular, by substituting equations (6) and (10) into (8) and equations (7) and (11) into (9) we obtain

$$\dot{n}_1^{dA} = -\dot{w}^A \Lambda_1^A + \dot{e}^A(\Lambda_1^A - 1) \tag{16}$$

where

$$\Lambda_1^A = \left[(1-\alpha_1^A)\sigma_1^A + \alpha_1^A \eta^A\right]$$

and

$$\dot{n}_2^{dA} = \dot{w}^A \Lambda_2^A + \dot{e}^A(\Lambda_1^A - 1) \tag{17}$$

where

$$\Lambda_2^A = \frac{1-\alpha_2^A - \sigma_2^A(1-\alpha_2^A)}{\beta^A} - \Lambda_1^A$$

These, in turn, lead to the following growth rate for the general equilibrium total labor demand in city A

$$\dot{n}^{dA} = \left[(1-\beta^A)\Lambda_2^A - \beta^A \Lambda_1^A\right] \dot{w}^A + \dot{e}^A(\Lambda_1^A - 1) \tag{18}$$

Note that Λ_1 and Λ_2 are the general equilibrium wage elasticities of labor demand in, respectively, the tradable and nontradable sectors. While Λ_1 is always positive, the sign of Λ_2 is potentially ambiguous. Here, however, since the price elasticity of the tradable goods η is assumed to be large, Λ_1 is also large and Λ_2 is negatively signed and large in absolute terms.

Moving to the labor supply side, through equations (11) and (13) and the corresponding ones for city B, equation (14) yields

$$\dot{n}^{sA} = \frac{1-\lambda}{L^{AB}}\left(\Gamma^A\dot{w}^A - \Gamma^B\dot{w}^B\right) \tag{19}$$

where $\Gamma^j = 1-\phi(1-\beta^j)$ with $j = A, B$.

Then, from the labor market clearing condition (15) we get the equilibrium growth rates of wages in the two cities

$$\dot{w}^A = \frac{(\Lambda_1^A - 1)\Omega^B}{\Omega^A\Omega^B - \frac{(1-\lambda)\lambda}{(L^{AB})^2}\Gamma^A\Gamma^B}\dot{e}^A \tag{20}$$

$$\dot{w}^B = \frac{(\Lambda_1^A - 1)\Gamma^A\frac{\lambda}{L^{AB}}}{\Omega^A\Omega^B - \frac{(1-\lambda)\lambda}{(L^{AB})^2}\Gamma^A\Gamma^B}\dot{e}^A \tag{21}$$

where

$$\Omega^A = \beta^A\Lambda_1^A - (1-\beta^A)\Lambda_2^A + \frac{1-\lambda}{L^{AB}}\Gamma^A$$

and

$$\Omega^B = \beta^B\Lambda_1^B - (1-\beta^B)\Lambda_2^B + \frac{\lambda}{L^{AB}}\Gamma^B$$

which are both positively signed sincer $\Lambda_1^j > 0$ and $\Lambda_2^j < 0$, with $j = A, B$.

In order to ensure the stability of the system, in each city, as wage increases, labor supply must increase more rapidly than labor demand. Given that $\Omega_A > 0$ and $\Omega_B > 0$, this requirement can be shown to imply

$$\Omega^A\Omega^B > \frac{(1-\lambda)\lambda}{(L^{AB})^2}\Gamma^A\Gamma^B$$

which, in turn, entails $\dot{w}^A > 0$ and $\dot{w}^B > 0$.

By substituting the growth rates of wages in the two cities in (19) it is possible to show that

$$\Gamma^A\dot{w}^A - \Gamma^B\dot{w}^B > 0$$

so that

$$\dot{n}^{sA} > 0 \qquad \dot{n}^{sB} < 0$$

In other words, a positive shock to the efficiency of production in the tradable sector of city A induces a shift of population from B to A. Having established this, we can now concentrate on the conditions under which a nontradable employment multiplier exists.

2.3 Local nontradable multiplier

Within this theoretical framework, the local employment multiplier in city A is

$$M^A = \frac{dn_1^A + dn_2^A}{dn_1^A} = 1 + \frac{(1-\beta^A)}{\beta^A}\frac{\dot{n}_2^A}{\dot{n}_1^A} \tag{22}$$

and therefore includes the elasticity of nontradable with respect to tradable employment

$$\dot{n}_2^A/\dot{n}_1^A$$

The first step is to assess the existence of M^A which, in our framework, requires both growth rates in the elasticity to be positive whenever a positive shock ($\dot{e}^A > 0$) occurs. Moving from equation (16), it is possible to show that

$$\dot{n}_1^A = \Delta^B\left[-(1-\beta^A)(\Lambda_1^A + \Lambda_2^A) + \frac{1-\lambda_A}{L_{AB}}\Gamma_A\right] - \frac{\lambda_A}{L_{AB}}(1-\beta^A)(\Lambda_1^A + \Lambda_2^A) \tag{23}$$

where $\Delta^B = \beta^B\Lambda_1^B - (1-\beta^B)\Lambda_2^A$ and $\Lambda_1^A + \Lambda_2^A = \frac{1-\alpha_2^A - \sigma_2^A(1-\alpha_2^A)}{\beta^A}$. Since in our setting the elasticity η is assumed to be large, we have that Δ_B is also large and positive; hence, what determines the sign of (23) is $\Lambda_1^A + \Lambda_2^A$ which is finite and whose sign depends on σ_2^A: $\Lambda_1^A + \Lambda_2^A \leq 0$ when $\sigma_2^A \geq 1$; $\Lambda_1^A + \Lambda_2^A > 0$ when $\sigma_2^A < 1$. More specifically

- if $\Lambda_1^A + \Lambda_2^A \leq 0$, then $\dot{n}_1^A$ is always positive
- if $\Lambda_1^A + \Lambda_2^A > 0$, then $\begin{array}{l} \dot{n}_1^A \geq 0 \quad if \quad \Lambda_1^A + \Lambda_2^A \leq \frac{(1-\lambda_A)\Gamma_A}{L_{AB}(1-\beta^A)} \\ \dot{n}_1^A < 0 \quad if \quad \Lambda_1^A + \Lambda_2^A > \frac{(1-\lambda_A)\Gamma_A}{L_{AB}(1-\beta^A)} \end{array}$

The just derived conditions, therefore, ensure that a positive shock hitting the tradable sector ($\dot{e}^A > 0$) actually lead to the creation of new jobs in the same sector ($\dot{n}_1^A > 0$). Having established this, we can now concentrate on the conditions for the existence of the local multplier, *i.e.* the conditions under which $\dot{n}_2^A > 0$. Moving from equation (17), it is possible to show that

$$\dot{n}_2^A = \Delta^B \left[\beta^A (\Lambda_1^A + \Lambda_2^A) + \frac{1-\lambda_A}{L_{AB}} \Gamma_A \right] - \frac{\lambda_A}{L_{AB}} \beta^A (\Lambda_1^A + \Lambda_2^A) \tag{24}$$

Also in this case, what determines the sign of (24) is $\Lambda_1^A + \Lambda_2^A$. In particular,

- if $\Lambda_1^A + \Lambda_2^A \geq 0$, then $\dot{n}_2^A$ is always positive
- if $\Lambda_1^A + \Lambda_2^A < 0$, then $\begin{array}{l} \dot{n}_2^A > 0 \quad if \quad \Lambda_1^A + \Lambda_2^A > -\frac{(1-\lambda_A)\Gamma_A}{L_{AB}\beta^A} \\ \dot{n}_2^A < 0 \quad if \quad \Lambda_1^A + \Lambda_2^A < -\frac{(1-\lambda_A)\Gamma_A}{L_{AB}\beta^A} \end{array}$

To sum up, therefore, the condition under which a positive shock effectively induces the creation of new jobs in the tradable sector and, at the same time, ensures the existence of the nontradable multiplier is

$$-\frac{(1-\lambda_A)\Gamma_A}{L_{AB}\beta^A} < \Lambda_1^A + \Lambda_2^A \leq \frac{(1-\lambda_A)\Gamma_A}{L_{AB}(1-\beta^A)} \tag{25}$$

Moving towards the empirical analysis, two crucial aspects must be taken under consideration. First, it is clear from equation (21) establishing a link between $\dot{e}^A$ and the growth rate of wages in B, that a shock occurring in city A spreads its effects to city B as well. More specifically, as emphasized above, a positive shock to the efficiency of production in the tradable sector of city A induces a shift of population from B to A. Equation (19) then tells us that the size of this population shift (and, therefore, of the effect induced to city B as well) is, other things being equal, inversely related to the distance between the two cities through L_{AB}. This clearly reveals a potential spatial dependence issue in the empirical analysis of the multiplier.

The second critical aspect to be addressed relates to the potential nonlinearities in the relationship between tradable and nontradable employment growth. As previously emphasised, the employment effects of a productivity shock are derived under the assumptions that shocks are small. However, when trying to empirically assess existence and size of the nontradable multiplier it must be recognized that, given the rather long time span required in empirical analyses in order to obtain interpretable results, it is quite possible that the small disturbance assumption gets violated. When this occurs, λ and βs cannot be regarded as constants any longer. As recalled above, a positive productivity shock in the tradable sector of city A induces a migration flow from B to A, hence determining an increase in λ; in addition, equations (23) and (24) tell us that the employment growth rates in the two sectors can be different thus altering the sector composition of the local labor markets (β).

To consider how these changes might affect the elasticity, we focus on the difference between the employment growth rates

$$\dot{n}_2^{dA} - \dot{n}_1^{dA} = (\Lambda_1^A + \Lambda_2^A)\dot{w}^A \tag{26}$$

as this greatly simplifies the presentation without altering the implications. In addition, we emphasize we assumed that the price elasticity of demand for city's A exports η^A is large and, consequently, that Λ_1^A is positive and large while Λ_2^A is negative and large in absolute terms. For simmetry, we make an analogous assumption with respect to these elasticities for city B.

Let us start from the effect of an increase in λ. Through partial differentiation it is easy to show that

$$\frac{\partial(\dot{n}_2^{dA} - \dot{n}_1^{dA})}{\partial\lambda} > 0 \qquad iff \qquad \Lambda_1^A + \Lambda_2^A > 0$$

but also that the effect is, in absolute terms, quite small. Hence, whenever the difference in general equilibrium elasticities is positive, the size of the multiplier in city A gets, to a small extent, larger (smaller) as the relative size of its population increases.

We then move to the effects of the changes in the sectoral compositions of employment. First, note from equation (26) that, whenever $\Lambda_1^A + \Lambda_2^A$ is positive (negative), nontradable employment grows more (less) them tradable and, consequently, β^A increases (decreases) and, if city B shares the same structural parameters with A, β^A decreases (increases).

Starting from the partial differentiation of (26) with respect to β^A, we get

$$\frac{\partial(\dot{n}_2^{dA} - \dot{n}_1^{dA})}{\partial\beta^A} > 0 \qquad iff \qquad \Lambda_1^A + \Lambda_2^A < 0$$

implying that, whenever $\Lambda_1^A + \Lambda_2^A$ is positive (negative), the size of the elasticity in city A gets smaller (larger) as β^A increases.

Finally, it possible to show that

$$\frac{\partial(\dot{n}_2^{dA} - \dot{n}_1^{dA})}{\partial\beta^B} > 0 \qquad iff \qquad \Lambda_1^A + \Lambda_2^A > 0$$

but also, as in the case of a variation in λ, that the effect is, in absolute terms, small. In other words, whenever $\Lambda_1^A + \Lambda_2^A$ is positive (negative), an increase in the relative size of the nontradable sector of city B leads to a small decrease (increase) in the nontradable employment elasticity of city A.

To sum up, it follows clearly from the above discussion that the tools employed in the empirical investigation must allow for the possibility that the size of the local nontradable multiplier in one city depends, on the one hand, on city size in a nonlinear fashion and, on the other, on shock occurring elsewhere in the system. In the following section, we therefore present a procedure to estimate model (1) that, being nonparametric, is flexible with respect to the functional form and, at the same time, is specifically designed to allow for spatial effects.

3 The spatial nonparametric regression estimator

3.1 Modeling spatial dependence

In the analysis of cross-section data quite often researchers have to face problems of misspecifications arising from dependence across spatially organised observational units.

Indeed, it is unlikely that the explanatory variables in a regression model can capture spatial unobservable factors and neglecting them is configured as a typical omitted variable scenario, difficult to treat with usual methods, like instrumental variables, because spatial dependence is often due to latent, although relevant, influences. To deal with spatial dependence the spatial econometrics literature offers variety of models (among others, Le Sage and Pace, 2009), a general example of which is the Spatial Durbin Model

$$\mathbf{Y} = \rho W \mathbf{Y} + \beta \mathbf{X} - \lambda W \mathbf{X} \beta + \epsilon \tag{27}$$

where $\mathbf{Y}$ is an $n \times 1$ vector, $\mathbf{X}$ is an $n \times p$ matrix[5], $\epsilon \sim N(0, \sigma^2 I_n)$ is a $n \times 1$ vector of innovations, $-1 < \rho < 1$, $-1 < \lambda < 1$, β is a $p \times 1$ vector of parameters and W is a $n \times n$ spatial weights matrix whose w_{ij} elements are non negative when $i \neq j$ and zero otherwise.

Analyses conducted resorting to models such as (27) have recently attracted a strong critical attention (Corrado and Fingleton, 2012; Gibbons and Overman, 2012; McMillen, 2012; Pinske and Slade, 2010; Fingleton and Lopez–Bazo, 2006). In particular, concerns concentrate on the set of very restrictive assumptions underlying commonly adopted models obtainable from (27) with appropriate constraints (*e.g.*, the Spatial Autoregressive and the Spatial Error models), on the difficulty in achieving identification and on the fact that these models are often estimated in a rather mechanical way with the aim of producing estimates with satisfactory statistical properties. Despite the fact that these alternative specifications may appear equally plausible in terms of statistical properties, they usually have widely different economic meanings and often lead to contraddictory policy implications.

In particular, Pinske and Slade (2010) and McMillen (2012) emphasize the dangers arising from reling heavily on parametric representations of spatial relations when, in fact, the functional form of these relations is, in general, unknown. As a promising

[5] For simplicity's sake we assume $p = 1$. All that follows can be generalised to the multivariate case $p \geq 2$.

way to address these concerns, they then indicate the employment of semiparametric or nonparametric models. Here, we follow this indication and generalize it so to include not just the functional form of spatial relations but, more generally, the overall relationship between regressand and regressor(s).

3.2 Nonparametric regression with dependent errors

Nonparametric regression has now become quite a standard statistical tool when the functional form is possibly of an unknown type. Indeed, given a model such as

$$y = m(x) + \epsilon$$

where ϵ is the i.i.d. error term and $m(x)$ is a smooth function, linearity of $m(x)$ can not be always safely assumed. Under these circumstances, the parametric literature typically offers Non Linear Least Squares Estimates, but this requires to conjecture a specific functional form with respect to which the minimization problem have to be solved. When making assumptions on the functional form of m is not possible or not recommended, nonparametric methods represent a valuable solution.

In general, the estimate of a nonparametric regression can be obtained by means of some smoothing methods. One of the most commonly adopted estimation technique is the Nadaraya-Watson estimator (Nadaraya, 1964; Watson, 1964):

$$\hat{m}(x) = \frac{\sum_{j=1}^{n} K(\frac{x-X_j}{h})Y_j}{\sum_{j=1}^{n} K(\frac{x-X_j}{h})} \tag{28}$$

where h is the bandwidth, the parameter that controls the degree of smoothness. The Nadaraya-Watson estimator is in fact a special case of local polynomial regression that applies when the degree of the smoothing polinomial is 0 and for this reason the Nadaraya-Watson is also know as the Local Constant Estimator (LCE). When the degree of the smoothing polynomial is 1 instead of 0, the smoother becomes the local linear estimator

(LLE):

$$\hat{m}(x) = \frac{\sum_{j=1}^{n} K(\frac{x-X_j}{h})Y_j}{\sum_{j=1}^{n} K(\frac{x-X_j}{h})} + (x - \bar{X}_w)\frac{\sum_{j=1}^{n} K(\frac{x-X_j}{h})(X_j - \bar{X}_w)Y_j}{\sum_{j=1}^{n} K(\frac{x-X_j}{h})(X_j - \bar{X}_w)^2} \tag{29}$$

where

$$\bar{X}_w = \frac{\sum_{j=1}^{n} K(\frac{x-X_j}{h})X_j}{\sum_{j=1}^{n} K(\frac{x-X_j}{h})}$$

Similarly to the parametric regression environment, nonparametric regression estimators generally assume i.i.d. error terms. In case of lack of independence, Robinson (2008, 2009) derives consistency and asymptotic distribution theory for the local constant regression estimator in relation to various kinds of spatial data. Other authors (for example, Xiao *et al.*, 2003; Lin and Carroll, 2000; Ruckstuhl *et al.*, 2000; Wang, 2003) study possible extensions of the nonparametric regression to a non i.i.d. errors setting, where errors can be correlated and heteroschedastic. In all cases, however, a parametric structure for the dependence must be assumed beforehand and this might represent a serious limitation since, as highlighted by Martins-Filho and Yao (2009), most asymptotic results for the LCE in case of dependent errors are unfortunately contingent on the assumptions made on the covariance structure and it is not possible to generalize their application to different parametric structures. Stimulated by this lack of generality, attention within the nonparametric literature has focussed on estimators that, by incorporating the information contained in the error covariance structure, outperform, both asymptotically and in finite samples, traditional nonparametric ones. In particular, Martins-Filho and Yao (2009) develop a two-step procedure whose asymptotic validity is proved under rather general covariance structures.

More formally, Martins-Filho and Yao consider the following nonparametric regression:

$$\mathbf{Y} = m(\mathbf{X}) + \mathbf{u} \tag{30}$$

where the error term $\mathbf{u}$ is such that $E(u_i) = 0$, $\forall i = 1, ..., n$, and $E(u_i, u_j) = \omega_{ij}(\theta_0)$, $\theta_0 \in R^p$, $p < \infty$.

In their Theorem 2 (Martins-Filho and Yao, 2009; page 312), the authors demonstrate the asymptotic normality and convergence rate of the traditional LLE of model (30). In addition, they observe that the traditional LLE of (30), $\hat{m}$, does not exploit the information contained in the error term correlation structure. Therefore, to improve its performance, they suggest a two-step procedure that incorporates this information in order to yield spherical error terms. More in detail, let $\Omega(\theta_0)$ denote the $n \times n$ matrix with elements ω_{ij} and $P(\theta_0)$ be a $n \times n$ matrix such that $\Omega(\theta_0) = P(\theta_0)P(\theta_0)'$. Now, by defining the new regressand as $\mathbf{Z} = P(\theta_0)^{-1}\mathbf{Y} + (I_n - P(\theta_0)^{-1})m(\mathbf{X})$, Martins-Filho and Yao replace the original regression with the following

$$\mathbf{Z} = m(\mathbf{X}) + \epsilon \tag{31}$$

where the error terms $\epsilon = P(\theta_0)^{-1}\mathbf{u}$ are now spherical by construction. The new estimator, $\check{m}(\mathbf{X})$, is simply the LLE of (31). With an additional assumption constraining the nature of the stochastic process $\mathbf{u}$ to be a linear transformation of i.i.d. processes, the authors show $\check{m}(\mathbf{X})$ to represent an improvement over $\hat{m}(\mathbf{X})$ in terms of efficiency (Theorem 3). To guarantee the bias from the first stage estimator to be smaller than the leading bias coming from the second stage, as is usual in the literature on two-stage nonparametric regression, undersmoothing in the first stage is required.

Since $\mathbf{Z}$ is not observed (it depends on the unknown $m(\mathbf{X})$ and $P(\theta_0)$), Martins-Filho and Yao propose a feasible version of the $\check{m}(\mathbf{X})$ estimator. This estimator, $\dot{m}(\mathbf{X})$, is based on an observed regressand

$$\dot{\mathbf{Z}} = P(\hat{\theta})^{-1}\mathbf{Y} + (I_n - P(\hat{\theta}^{-1})\hat{m}(\mathbf{X})$$

where a pilot local linear estimate $\hat{m}(\mathbf{X})$ is used in place of $m(\mathbf{X})$ and $P(\hat{\theta})$ in place of $P(\theta_0)$. The authors also provide an asymptotic result (Theorem 4) that guarantees that, as long as a consistent estimate $\hat{\theta}$ is plugged in the expression $P(\theta_0)$, the feasible estimator $\dot{m}(\mathbf{X})$ is asymptotically equivalent to $\check{m}(\mathbf{X})$.

3.3 A new nonparametric regression estimator for spatially dependent data

Building on the theoretical background of the two-step nonparametric regression estimator by Martins-Filho and Yao, we can now introduce our proposal of a nonparametric regression estimator for spatially dependent data (SNP).

Since spatial dependence has not been included among the forms of dependence considered by Martins-Filho and Yao, here further assumptions need to be made on the form of the spatial covariance matrix of the error term. In particular, apart from adopting assumptions A1-A6 characterising the original framework developed by these authors (Martins-Filho and Yao, 2009; page 311 and 313), we: *i.* suppose the error term to possess a spatial covariance matrix such that the weighted average of the main diagonal elements converge as $n \to \infty$ and *ii.* impose spatial mixing conditions, as in Jenish and Prucha (2009).[6]

A peculiar feature of our procedure is that the spatial covariance matrix is estimated nonparametrically starting from a direct representation of spatial dependence. The logic underlying the estimate of the covariance matrix of the error term is that, unless the form of dependence is of interest itself, it is better not to parametrize it. From this viewpoint, Robinson (1987) estimates the residuals variance, conceived as an unknown function of the explanatory variables, by a nearest neighbours nonparametric regression and proves that asymptotic properties of the estimated residuals variance (as well as the other parameters involved in the multiple regression) are guaranteed if k the number of nearest neighbours, increases slowly with the sample size. In developing our estimator, we follow Robinson's logic relatively to a multivariate regression with non spherical errors and estimate consistently the spatial covariance matrix through a nonparametric methodology, called spline correlogram, whose details will be presented in the next subsections.

[6] Note that Cliff-Ord models trivially meet these assumptions.

3.3.1 Nonparametric estimation of the spatial covariance matrix

A commonly adopted approach to express the elements of a generic spatial covariance matrix Ω is through a direct representation of the dependence as some function of the distance separating sites s_i and s_j. In such an instance, the spatial autocovariance function is defined by

$$\gamma(s_i, s_j) = \sigma^2 f(d_{ij}, \phi) \tag{32}$$

and the spatial autocorrelation function by

$$\rho(s_i, s_j) = f(d_{ij}, \phi) \tag{33}$$

where d_{ij} is the distance between sites i,j and $f(.)$ is a decaying function such that $\frac{\partial f}{\partial d_{ij}} < 0$, $|f(d_{ij}, \phi)| \leq 1$ with ϕ being an appropriate vector of parameters. Within this framework, the spatial covariance matrix Ω is positive definite and composed by elements ω_{ij}, obtained through function $\gamma(s_i, s_j)$ in observed distances across sites. These features of matrix Ω follow directly from the stationarity and isotropy assumptions underpinning the existence of a spatial covariance function.[7]

Bjørnstad and Falck (2001) propose a nonparametric estimate of the spatial covariance matrix moving from a continuous nonparametric positive semidefinite estimator of $f(d_{ij}, \phi)$ in (33), called spline correlogram. In particular, they build on the seminal work of Hall and Patil (1994) who, in turn, develop a kernel estimator of the spatial autocorrelation function $\rho(s_i, s_j)$:

$$\tilde{\rho}(s_i, s_j) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} K(d_{ij}/a)(\hat{\rho}_{ij})}{\sum_{i=1}^{n} \sum_{j=1}^{n} K(d_{ij}/a)} \tag{34}$$

where K is a kernel function, a is a bandwidth and $\hat{\rho}_{ij}$ is the sample correlation

$$\hat{\rho}_{ij} = \frac{(z_i - \bar{z})(z_j - \bar{z})}{1/n \sum_{l=1}^{n} (z_l - \bar{z})^2} \tag{35}$$

[7] These assuptions are certainly met when Ω represents the spatial covariance matrix of homoskedastic errors of a Cliff-Ord type model.

in which $\bar{z} = 1/n \sum_{l=1}^{n} z_l$ is the sample mean. Hall and Patil (1994) demonstrate that the estimator in (34) can be tuned (by tuning a) so that $\tilde{\rho}(d) \rightarrow 0$ for any smooth functional form of $\rho(d)$.

Starting from the estimator in (34), Bjørnstad and Falck (2001) opt for a cubic B-spline as a smoother[8]. Given N pairs (x_i, y_i), $i = 1, ..., N$, the smoothing spline solves the fitting problem by selecting the function f that minimizes the penalized residual sum of squares (RSS)

$$RSS(f(x), \lambda) = \sum_{i=1}^{N} \{y_i - f(x_i)\}^2 + \lambda \int \left\{f''(t)\right\}^2 dt \tag{36}$$

where λ is a fixed smoothing parameter (Hastie *et al.*, 2009). The first term measures closeness to the data, the second term penalizes curvature in the function; λ represents a trade-off between the two, varying from very rough fits ($\lambda = 0$) to very smooth fits($\lambda = \infty$). The asymptotic kernel, equivalent to a cubic B-spline is:

$$K(d/a) = \frac{1}{2} exp\left(-\frac{|d/a|}{\sqrt{2}}\right) sin\left(-\frac{|d/a|}{\sqrt{2}} + \frac{\pi}{4}\right) \tag{37}$$

where, once more, d denotes a generic measure of distance. The advantage in using the B-spline is in that this smoother adapts better to irregularly spaced data and produces a consistent estimate of the covariance function (Hyndman and Wand, 1997). A standardized way to express the degree of smoothing when a spline smoother is used is by the equivalent degrees of freedom; in other words, λ can be specified by fixing the degrees of freedom. It has been shown that fixing the degree of smoothing using the cross validation (see Green and Silverman, 1994, and Hastie *et al.*, 2009, for more details) and assuming a true covariance function $\rho(s_i, s_j)$ that is C2-differentiable (i.e. with continuous 1st and 2nd derivatives) guarantees results with asymptotic properties.

In addition, since the estimator $\tilde{\rho}(s_i, s_j)$ must be not only consistent but also positive semidefinite, and this is not necessarily guaranteed by the estimator in equation (34),

[8] Silverman (1984) pointed out that the smoothing spline is essentially a local kernel average with a variable bandwidth.

Bjørnstad and Falck resort to a Fourier-filter method (Hall *et al.*, 1994). The latter works as follows: firstly the Fourier transform of $\tilde{\rho}(s_i, s_j)$ is calculated, then all negative excursions of the transformed function are set to zero and, finally, a nonparametric positive semidefinite estimate is obtained of the spatial correlation function by backtransformation.

3.3.2 The SNP estimator in practice

Consider the following general nonparametric model:

$$\begin{aligned} \mathbf{Y} &= M(\mathbf{X}) + \mathbf{u} \\ \mathbf{u} &= \theta W \mathbf{u} + \epsilon \end{aligned} \tag{38}$$

Operatively, the SNP estimator of (38) can be described as follows:

0. *Pilot fit*: estimate $m(\mathbf{X})$ with a local polynomial smoother, where the bandwidth h is chosen following an optimal rule. As for the degree of the polynomial, we consider $p = 1$. The output is $\hat{\mathbf{u}} = \mathbf{Y} - \hat{m}(\mathbf{X})$.

1. *Nonparametric covariance matrix estimation*: using the spline correlogram, obtain $\hat{V}$, the estimated spatial covariance matrix of $\hat{\mathbf{u}}$.

2. *Final fit*: feed the procedure with the information obtained from the estimate of the spatial covariance matrix $\hat{V}$ by running a modified regression where $\mathbf{Y}$ is replaced by $\mathbf{Z} = \hat{m}(\mathbf{X}) + L^{-1}\hat{\mathbf{u}}$ and L is obtained by taking the Cholevsky decomposition of $\hat{V}$.

We emphasize that SNP allows to estimate models like (27) and all those specifications obtainable from it with appropriate constraints (*i.e.*, the traditional Spatial Autoregressive and Spatial Error models). To see this, start from (38) and note that

- $M(\mathbf{X}) = \mathbf{X}\beta$ and $\theta = \lambda$ in case of Spatial Error Model

- $M(\mathbf{X}) = (I - \rho W)^{-1}\mathbf{X}\beta$ and $\theta = \rho$ in case of Spatial Lag Model

- $M(\mathbf{X}) = (I - \rho W)^{-1}(I - \lambda W)\mathbf{X}\beta$ and $\theta = \rho$ in case of Spatial Durbin Model

The function $M(X)$ in (38) consists of two components: one is $m(x)$, the function that characterizes the models in their original form[9], the second is a spatial factor that depends on the typology of spatial model that can possibly be the data generating process. We stress that estimating function M via SNP does not require assumptions on the functional form of M. This aspect is crucial since it means that no hypotheses are made neither on m nor the spatial factor that, multiplied by $m(X)$, generates $M(X)$.

3.4 Monte Carlo study

This Section describes a Monte Carlo experiment[10] that aims at showing, via simulations, the finite sample performance of our procedure in comparison with a traditional non-parametric method that does not take the presence of spatial dependence into account. The purpose therefore is to investigate the effective improvement in regression estimation results when spatial dependence is not neglected.

We carry out the Monte Carlo experiment considering several nonlinear specifications for the Spatial Error Model

$$\begin{aligned} \mathbf{Y} &= m(\mathbf{X}) + \mathbf{u} \\ \mathbf{u} &= \lambda W \mathbf{u} + \epsilon \end{aligned} \tag{39}$$

In particular, we consider the following specifications:

- A $\quad m(x) = sin(5\pi x)$
- B $\quad m(x) = 2 + sin(7.1(x - 3.2))$
- C $\quad m(x) = 1 - 48x + 218x^2 - 315x^3 + 145x^4$
- D $\quad m(x) = 10exp(-10x)$

[9] $m(X)$ is not necessarily required to be linear since we are estimating the regression with nonparametric tools

[10] The code has been written in Matlab (Matlab 7.7.0, R2008b).

- E $\quad m(x) = (-1 + 2x) + 0.95exp(-40(-1 + 2x)^2)$
- F $\quad m(x) = 1/(1 + exp(-6 + 12x))$
- G $\quad m(x) = \begin{cases} exp(x - 0.33) & if \quad x < 0.33 \\ exp(-2(x - 0.33)) & if \quad x \geq 0.33 \end{cases}$

the shapes of which are depicted in Figure 1.

The simulated data set length is $N = 50, 100, 200$ and the number of Monte Carlo replications per experiment is 1000. The regressor is drawn from a uniform distribution, $X \sim U(0, 1)$, while the disturbance term is generated as a vector of normally distributed random variables, $\epsilon \sim N(0, \sigma^2)$, where σ is set to obtain three alternative levels for the pseudo-R^2 $(0.2, 0.5, 0.8)$. In addition, units are assumed to belong to a circular world and, similar to Kelejian and Prucha (1999 and 2007) and Kapoor *et al.* (2007), the spatial weights matrix W is such that each observation is directly related to the six units immediately surrounding it (three on each side) it in the ordering. Specifically, the matrix is such that all nonzero elements are equal and, following common practice, the matrix is row-normalized. Finally, λ takes on three alternative values $(0.3, 0.5, 0.8)$ corresponding to low, intermediate and strong spatial dependence, giving us a total of 189 $(7 \times 3 \times 3 \times 3)$ experiments.

We employ two estimation methods: the traditional local linear estimator (NP) and the procedure proposed in the previous Section (SNP), implemented with a local linear estimator. For all simulations we use the gaussian kernel with bandwidths that minimize the cross-validation criterion.[11]

An estimator's performance is measured by calculating the median across replications of the Mean Integrated Squared Error (MISE) obtained in each replication. A direct comparison of the relative performance of the two estimators is then carried out through the ratio between the median MISE of SNP with respect to the median MISE on NP. The results of the complete set of experiments are reported in Table 1.

[11] To guarantee the required degree of undersmoothing, the bandwith in the pilot estimate of the SNP estimator is $h = N^{(-1/10)}g$ where g is optimal bandwith obtained via the cross-validatory criterion.

Overall, results are quite good as median ratios are in almost all cases below 1, with no appreciable differences across the different functional forms. Median ratios are close to 1 for low levels of spatial dependence ($\lambda = 0.3$) and for the smallest sample size ($N = 50$) while they display significant reductions as the strength of spatial dependence and the size of the sample increase. In particular, SNP procedure visibly outperforms the traditional local linear estimator when λ reaches 0.8 and $N = 200$, obtaining median values of the MISE that are approximately 25 percent smaller in several cases.

4 Empirical analysis

We estimate the following regression model:

$$\Delta N^{NT} = m(\Delta N^{T}) + \epsilon \tag{40}$$

where ΔN^{NT} and ΔN^{T} are, respectively, the change in the log number of jobs in the nontradable and tradable sector.

The period covered in the analysis runs from 1980 to 2010. Employment data are estimated using the 5 percent samples from federal census data for 1980 and using the 1 percent samples from the American Community Survey for 2010 obtained from the Integrated Public Use Microdata Series (IPUMS). We paid particular attention to the distinction between tradable and non tradable sectors. As emphasied in Section 2.3, the existence of a multiplier rests on the hypothesis that the parameter η, *i.e.*, the price elasticity of demand for tradable goods, is large. From a pratical point of view, this means that the actual classification of tradable industries must be effectively able to identify goods and services that share this distinctive feature. Consequently, rather that adopting the traditional classification that identifies tradable industries with manufacturing, we use the two-digit NAICS code classification provided by Hufbauer and Vieiro (2013) following the approach developed by Jensen and Kletzer (2005). According to this approach, when production is concentrated at a distance from consumption within the US, as inferable from a locational Gini coefficient exceeding 0.1, the activity is classified as 'tradable'.[12]

To isolate exogenous shifts in the demand for labor in the tradable sector, we use the instrument suggested by Moretti (2010), *i.e.* a weighted average of nationwide employment growth by 84 industries (classified according to the 1990 Census Bureau industrial classification scheme) within the tradable sector, with weights reflecting their location-specific employment share in 1980.

The territorial unit of the analysis is the metropolitan area since this is considered the most appropriate in approximating the boundaries of local labor markets. Due to limitations in the availability of data on employment at the level of disaggregation required for the construction of the instrument, we are forced to restrict the analysis to a subset of 123 MSA.

We produce two sets of estimates. The first employs OLS and, among nonparametric methods, the local linear estimator (NP) and the corresponding spatial nonparametric estimator (SNP) described in Section 3. The second resorts to the Instrumental Variable (IV) counterparts of the above methods.

Before proceeding with the empirical analysis, we need to clarify how the SNP estimator has been included within the IV framework. In order to be consistent with the nonparametric nature of the SNP estimator, we resort to the literature on nonparametric IV estimators which offers various possibilities (see Horowitz (2011) for a good review of them). In particular, we opted for the control function model (Newey *et al.*, 1999). Given a dependent variable Y, an explanatory variable X that may be endogenous and an instrument W, the model is written as follows

$$Y = m(X) + u \tag{41}$$

[12] Since data in the Integrated Public Use Microdata Series are originally classified according to the 1990 Census Bureau industrial classification scheme, we mapped them into the two-digit NAICS code classification using industry code crosswalks provided by the US Census Bureau. As in Moretti (2010) we exclude agriculture, mining, government and the military.

and

$$X = h(W) + V \tag{42}$$

where g and h are unknown functions and

$$E(V|W) = 0$$

and

$$E(U|X, V)$$

and it follows that

$$E(Y|X, V) = g(X) + k(V) \tag{43}$$

If V were observed, g could be estimated using a variety of estimators for nonparametric additive models. However, (Newey *et al.*, 1999) prove that V can be consistently estimated by the residuals from the nonparametric regression of h in (42) and the estimated V can be used in place of the true one to estimate g in (43). In sum, the method proposed by Newey *et al.* (1999) consists of two steps. In the first step, the nonparametric regression of X on Z generates a residuals set $\hat{V}$. The second step is then the nonparametric regression of Y on X and $\hat{V}$.

Within this theoretical framework, we run the second step of the procedure by Newey *et al.* (1999) by resorting to generalized additive models (Hastie *et al.*, 1993). More specifically, since in our case we also need to treat spatial dependence, we estimate the generalized additive models (GAM, hereafter) by means of our SNP estimator.

The outcome of the first set of estimates is shown in Figure 2. In both panels the horizontal dotted line represents the level of elasticity returned from the OLS estimate, equal to 0.659 (and statistically highly significant) and higher than the value by Moretti (2010).[13]

[13] It must be clarified, however, that our analysis covers a longer time period and is based on a different spatial unit.

In the left panel we compare this estimate to the one obtained through a traditional local linear estimator. In addition, we report 5% confidence bands for the nonparametric estimate in order to understand whether there are significant differences with the OLS. It can be noted that confidence bands always include the OLS elasticity value, thus suggesting the absence of any significant difference between the two estimates and, in addition, that the linear specification might not be inappropriate.

It is also worth noting that the null hypothesis of spatial independence in the residuals is strongly rejected by the Moran's I test (based on an inverse of distance weight matrix and 9999 permutations) both for the OLS and local linear estimate (Table 2). A radically different picture then emerges when we move to the right panel in which we report the outcome of the estimate that allows for spatial dependence. Here, the 5% confidence intervals suggest that the elasticity obtained through the SNP estimator is significantly different from the OLS over a large part of the domain. More in detail, the estimated elasticity is approximately constant around a value of 0.55 (marginally different from the OLS from a statistical point of view) up to a value just above 1 in terms of tradable growth (corresponding approximately to a three-fold increase of this sector over a 30-year period) and then increases at a rather constant rate up to a level of 1.1. As shown in Table 2, as expected it is now possible to safely accept the null of spatial independence.

The outcome of the second set of estimates is reported in Figure 3. As explained above, these estimates try to address endogeneity concerns via an Instrumental Variable approach. In all estimates, we employed as instrument the growth rate in the tradable sector that the local economy would have experienced had local sector grown at the same rate of the national economy. Similar to what we have done before, the horizontal dotted line represents the level of OLS-IV elasticity, which is now equal to 0.424 (and statistically highly significant). Also in this case, the value we find is higher that the corresponding one reported by Moretti; however, due to the different classification of tradable and nontradable sectors, our elasticity yields an estimate of only 0.34 (rather than 1.59) jobs created in the nontradable sector of a given city for each additional job in the tradable sector of the same city.

Moving to the nonparametric analysis, we observe that IV estimates of the elasticities are, in general, lower that those found previously. Once again, the estimate obtained though a local linear estimator is, to a large extent, non statistically distinguishable from the OLS-IV elasticity (left panel of Figure 3). In addition, the estimated shape of the NP-IV elasticity might suggest the presence of a threshold corresponding to a value just above 1 in terms of tradable growth (in line to the SNP estimate of Figure 2). However, also in this case there is strong evidence against the null hypothesis of spatial independence in the residuals as reported in Table 3. The right panel then report the outcome of an estimate that explicitly allows for both endogeneity and spatial dependence concerns. The SNP-IV elasticity is in general lower than the SNP one, but shares with the latter the main features: the estimated elasticity is rather constant (but significantly different from the OLS-IV elasticity) up to a level just in excess of 1 in tradable growth, and then increases at an apparently constant rate, thus confirming the idea of a possible threshold. Once more, the Moran's test suggests that there should not spatial dependence in the residuals.

Finally, exploiting the local nature of the nonparametric estimate, we calculate for each metropolitan area in the dataset the jobs in the nontradable sector created by an additional job in the tradable one. As shown in Figure 4, there is a wide variety of responses to a shock in the tradable sector as the number of additional job in the nontradable sector ranges between 0.1 to almost 0.8. In addition, coherently with the main features of the SNP estimates, for shocks larger than a three-fold increase in the tradable sector over the 30-year horizon, the response from the nontradable sector appears to increase with the size of the shock.

5 Conclusions

In this paper we have presented a spatial equilibrium model that clarifies the conditions under which the local nontradable multiplier exists and, consequently, a local economic development policy implemented in the tradable sector of a city determines an employment expansion in the local nontradable sector.

The model, however, also suggests that this local productivity shock produces effects that depend nonlinearly on the size of the shock and, at the same time, may spread to other cities. These indications raise some concerns about the validity of recent results on the size of the multiplier. To address these concerns we have therefore suggested that the simple methodology indicated by Moretti (2010) should be implemented through a spatial nonparametric procedure, *i.e.*, a procedure that, being nonparametric, is flexible with respect to the functional form and, at the same time, is specifically designed to allow for spatial effects. We have therefore developed a spatial nonparametric (SNP) estimator and demostrated its consistence via Monte Carlo experiments.

To establish the size of the local nontradable multiplier have then applied the SNP procedure to data on employment by narrowly defined industries for 123 US metropolitan areas between 1980 and 2010, paying particular attention to the distinction between tradable and non tradable sectors. In addition, to check whether spatial dependence and nonlinearities might actually affect the results, we have employed the OLS and the traditional nonparametric local linear estimator, while adopting an Instrumental Variable approach to address endogeneity concerns.

Our results suggest that allowing for spatial dependence significantly modifies the picture arising from more conventional estimators. In particular, using SNP we find that the relationship between the change in tradable jobs and the change in nontradable jobs is constant up to a threshold size of the shock and then increase at an approximately constant rate. There results therofore appear to confirm suppositions arising from the theoretical model with respect to the nonlinearity of the the response from the nontradable sector and to the existence of effects that spill outside the boundaries of the local labor market.

References

Bjørnstad, O.N., Falck, W. (2001): Nonparametric Spatial Covariance Functions: Estimation and Testing. *Enviromental and Ecological Statistics*, **8**, 53–70.

Cliff, A.D. and Ord, J.K. (1973): Spatial Autocorrelation. London: Pion.

Corrado, L., Fingleton, B. (2012): Where Is the Economics in Spatial Econometrics? *Journal of Regional Science*, **52**, 210–239.

de Blasio, G., Menon, C. (2011): Local Effects of Manufacturing Employment Growth in Italy. *Giornale degli Economisti e Annali di Economia*, **70**, 101–112.

Eberts, R.W. (2005): Overview of State and Local Economic Development Policies and Practice in the United States. In Local Governance for Promoting Employment: Comparing the Performance of Japan and Seven Countries, Gigure, S., Higuchi, Y. and the Japan Institute for Labour Policy and Training, (eds.), Tokyo: The Japan Institute for Labour Policy and Training.

Fingleton, B., Lopez–Bazo, E. (2006): Empirical Growth Models with Spatial Effects *Papers in Regional Science*, **85**, 177–198.

Gibbons, S., Overman, H. (2012): Mostly Pointless Spatial Econometrics? *Journal of Regional Science*, **52**, 172–191.

Green, P.J., Silverman, B.W. (1994): Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach. London: Chapman & Hall

Hall, P., Patil, P. (1994): Properties of Nonparametric Estimators of Autocovariance for Stationary Random Fields. *Probability Theory and Related Fields*, **99**, 399–424.

Hall, P., Fisher, N.I., Hoffman, B. (1994): On the Nonparametric Estimation of Covariance Functions. *Annals of Statistics*, **22**, 2115–2134.

Hastie, T., Tibshirani R., Friedman J. (2009): *The Elements of Statistical Learning.* Berlin: Springer.

Horowitz, J.L. (2011): Applied Nonparametric Instrumental Variables Estimation Econometrica **79**, 347–394.

Hufbauer G.C. Vieiro, M. (2013): Corporate Taxation and US MNCs: Ensuring a Competitive Economy Policy Brief 13-9, Washington, DC: Peterson Institute for International Economics.

Hyndman, R.J., Wand, M.P. (1997): Nonparametric Autocovariance Function Estimation. *Australian Journal of Statistics*, **39**, 313–324.

Jenish, N., Prucha I. (2009): Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, **150**, 86–98.

Jensen, J.B., Kletzer, L. (2005): Tradable services: understanding the scope and impact of services outsourcing. Working paper series, paper n. 05-9. Washington, DC: Institute of International Economics.

Kapoor, M., Kelejan H.H., Prucha I. (2007): Panel data models with spatially correlated error components. *Journal of Econometrics*, **140**, 97-130.

Kelejian, H.H. and Prucha, I.R. (1998): A Generalized Spatial two-stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate and Economics*, **17**, 99–121.

Kelejian, H.H., Prucha, I.R. (1999): A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review*, **40**, 509–533.

Kelejian, H.H., Prucha, I.R. (2007): HAC Estimation In a Spatial Framework. *Journal of Econometrics*, **140**, 131–154.

LeSage J.P., Pace, R.K. (2009): Introduction to Spatial Econometrics. Boca Raton (FL): Chapman and Hall.

Lin, X., Carroll, R.J. (2000): Nonparametric Function Estimator for Clustered Data when the Predictor is Measured with/without Error. *Journal of the American Statistical Association*, **95**, 520–534.

Maier, G., Fischer, M.M. (1985): Random Utility Modelling and Labour Supply Mobility Analysis *Papers in Regional Science*, **58**, 21–33.

Martins-Filho, C., Yao, F. (2009): Nonparametric Regression Estimation with General Parametric Error Covariance. *Journal of Multivariate Analysis*, **100**, 309–333.

McGregor, P.G., McVittie, E.P., Swales, J.K., Yin, Y.P. (2000): The Neoclassical Economic Base Multiplier. *Journal of Regional Science*, **40**, 1–31.

McMillen, D.P. (2012): Perspectives on Spatial Econometrics: Linear Smoothing With Structured Models *Journal of Regional Science*, **52**, 192–209.

Merrifield, J. (1987): A Neoclassical Anatomy of the Economic Base Multiplier *Journal of Regional Science*, **27**, 283–294.

Merrifield, J. (1990): A Practical Note on the Neoclassical Economic Base Marginal Multiplier *Journal of Regional Science*, **30**, 123–127.

Moretti, E. (2010): Local Multipliers. *American Economic Review: Papers and Proceedings*, **100**, 1–7.

Moretti, E. (2011): Local Labor Markets. In Ashenfelter, O. and Card, D. (eds.) Handbook of Labor Economics **4B**, Amsterdam: North-Holland Elsevier.

Nadaraya, E.A. (1964): On Estimating Regression. *Theory of Probability and its Applications* **9**, 141–142.

Newey, W.K., Powell J.L., Vella F. (1999): Nonparametric Estimation of Simultaneous Equation Models. *Econometrica* **67**, 565–603.

Pinske, J., Slade, M.E. (2010): The Future of Spatial Econometrics. *Journal of Regional Science*, **50**, 103–117.

Poole, K.E., Erickcek, G.A., Iannone, D.T., McCrea, N., Salem, P. (1999): Evaluating Business Development Incentives. Washington, D.C.: National Association of State Development Agencies.

Roback, J. (1982): Wages, Rents and the Quality of Life. *Journal of Political Economy*, **90**, 1257–1278.

Robinson, P.M. (1987) Asymptotically Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form. *Econometrica*, **55**, 875–891.

Robinson, P.M. (2008): Developments in the Analysis of Spatial Data. *Journal of the Japanese Statistical Society*, **38**, 87–96.

Robinson, P.M. (2011): Asymptotic Theory for Nonparametric Regression with Spatial data. *Journal of Econometrics*, **165**, 5–19.

Rosen, S. (1979): Wage-based Indexes of Urban Quality of Life. In Miezkowski, P. and Straszheim, M.R. (eds.) Current Issues in Urban Economics. Baltimore: Johns Hopkins University Press.

Ruckstuhl, A.F., Welsh, A.H. and Carroll, R.J. (2000): Nonparametric Function Estimation of the Relationship between two Repeatedly Measured Variables. *Statistica Sinica*, **10**, 51–71.

Silverman B.W. (1984) Spline Smoothing: The Equivalent Variable Kernel Method. *Annals of Statists*, **12**, 898–916.

Wang, N. (2003): Marginal Nonparametric Kernel Regression Accounting for within Subject Correlation. *Biometrika*, **90**, 43–52.

Watson, G.S. (1964): Smooth Regression Analysis. *Sankhya, Series A*, **26**, 359–372.

Xiao, Z., Linton, O.B., Carroll, R.J. and Mammen, E. (2003): More Efficient Local Polynomial Estimation in Nonparametric Regression with Autocorrelated Errors. *Journal of the American Statistical Association*, **98**, 980–992.

Figures and Tables

Figure 1: Monte Carlo experiment: functional forms

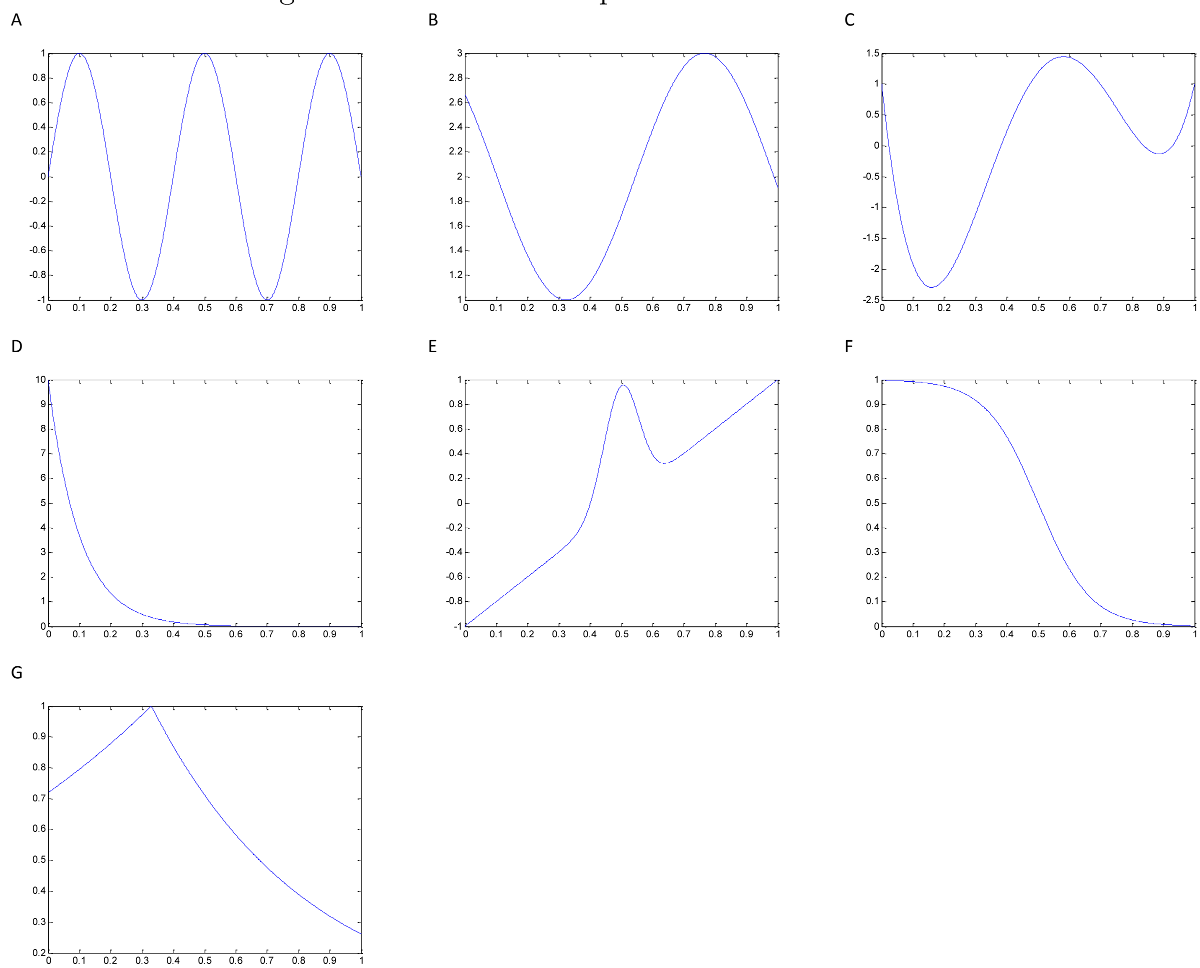
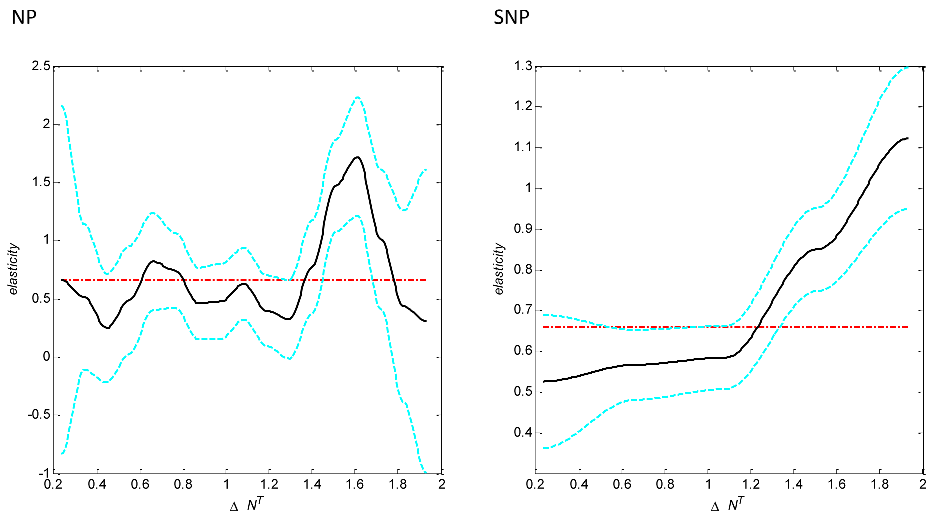
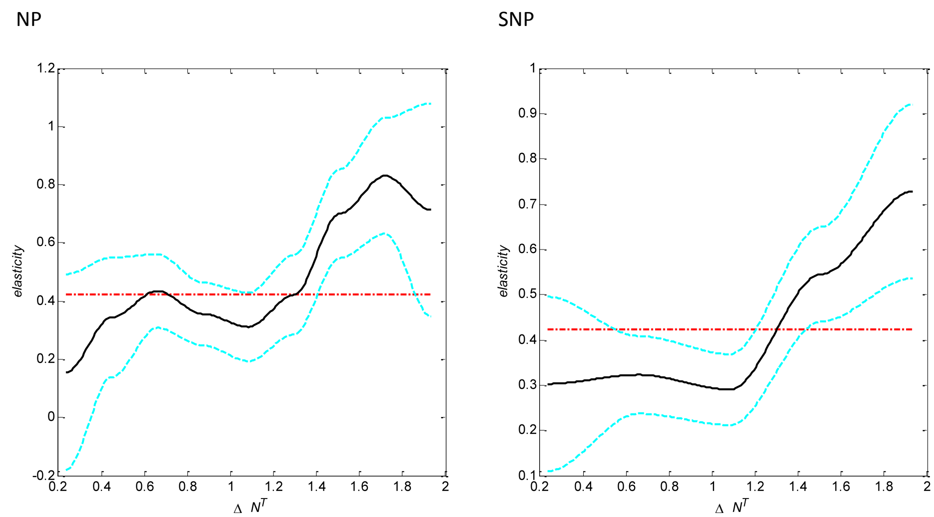


Figure 2: Elasticity estimates



NOTE: The NP elasticity is obtained using a local linear estimator; the SNP estimate is obtained using a spatial local linear estimator as described in the text. Bandwidths have been selected using the generalized cross validation method. The distance matrix for the SNP estimator contains Euclidean distances across metropolitan areas centroids.

Figure 3: Elasticity estimates (IV)



NOTE: The NP elasticity is estimated through an additive model with a local linear estimator for each element; the SNP estimate is obtained through an additive model with a spatial local linear estimator. Bandwidths have been selected using the generalized cross validation method. The distance matrix for the SNP estimator contains Euclidean distances across metropolitan areas centroids.

Figure 4: Additional jobs (IV-SNP)

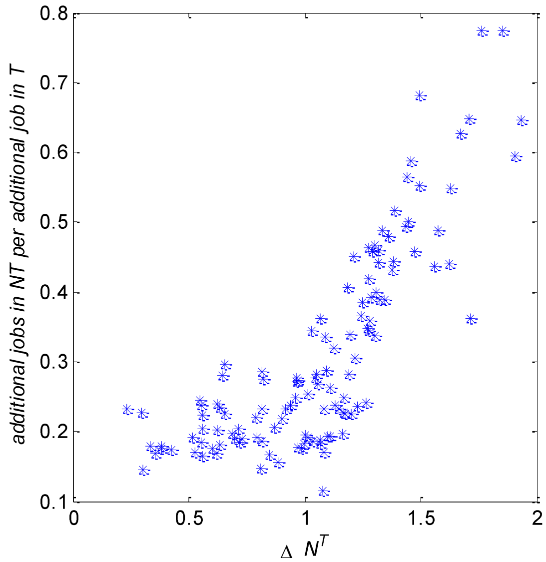


Table 1: Monte Carlo results

			A			B			C	
pseudo-R^2	**N**	**0.3**	**0.5**	**0.8**	**0.3**	**0.5**	**0.8**	**0.3**	**0.5**	**0.8**
0.2	**50**	0.98	0.97	0.95	0.98	0.99	0.94	1.02	1.00	0.95
0.2	**100**	0.98	0.94	0.89	1.02	0.98	0.88	0.98	0.94	0.87
0.2	**200**	0.97	0.93	0.79	0.98	0.93	0.79	0.97	0.90	0.83
0.5	**50**	1.00	0.98	0.94	1.00	0.98	0.89	1.02	1.01	0.92
0.5	**100**	0.99	0.93	0.81	1.01	0.95	0.81	0.99	0.96	0.82
0.5	**200**	0.98	0.93	0.75	0.98	0.92	0.73	0.97	0.91	0.79
0.8	**50**	1.00	0.98	0.89	1.00	0.97	0.90	1.02	0.99	0.89
0.8	**100**	1.00	0.95	0.81	0.99	0.96	0.78	1.01	0.96	0.82
0.8	**200**	0.98	0.92	0.73	0.98	0.93	0.77	0.98	0.93	0.76

			D			E			F	
pseudo-R^2	**N**	**0.3**	**0.5**	**0.8**	**0.3**	**0.5**	**0.8**	**0.3**	**0.5**	**0.8**
0.2	**50**	1.02	1.00	0.94	0.99	0.97	0.92	1.01	0.96	0.95
0.2	**100**	1.01	0.95	0.88	0.98	0.96	0.88	0.99	0.97	0.84
0.2	**200**	0.98	0.96	0.85	0.99	0.96	0.79	0.98	0.94	0.79
0.5	**50**	1.01	0.98	0.94	1.00	0.97	0.90	1.02	0.98	0.97
0.5	**100**	1.02	0.95	0.88	1.00	0.95	0.89	1.01	0.97	0.87
0.5	**200**	0.99	0.92	0.78	0.99	0.92	0.86	0.99	0.94	0.82
0.8	**50**	1.01	0.98	0.94	1.00	0.98	0.92	1.00	0.98	0.98
0.8	**100**	0.99	0.97	0.82	1.00	0.97	0.84	1.01	0.96	0.89
0.8	**200**	0.99	0.93	0.74	0.98	0.92	0.77	0.99	0.94	0.78

			G	
pseudo-R^2	**N**	**0.3**	**0.5**	**0.8**
0.2	**50**	0.99	0.98	0.89
0.2	**100**	0.99	0.96	0.87
0.2	**200**	0.98	0.96	0.82
0.5	**50**	1.00	0.98	0.92
0.5	**100**	0.98	0.94	0.88
0.5	**200**	0.99	0.92	0.80
0.8	**50**	1.01	0.97	0.91
0.8	**100**	1.00	0.95	0.87
0.8	**200**	0.98	0.94	0.77

NOTE: Ratio (SNP over NP) of the median across replications of the MISE

Table 2: Spatial dependence diagnostics

	p-level		
	OLS	**NP**	**SNP**
Moran's I	0.0000	0.0003	0.3426
LM error	0.0000		
Robust LM error	0.0002		
LM lag	0.0393		
Robust LM lag	0.8599		

NOTE: Spatial dependence tests employ an inverse of distance weight matrix

Table 3: Spatial dependence diagnostics (IV estimates)

	p-level		
	OLS	**NP**	**SNP**
Moran's I	0.0000	0.0050	0.6979
LM error	0.0000		
Robust LM error	0.0044		
LM lag	0.0013		
Robust LM lag	0.1586		

NOTE: Spatial dependence tests employ an inverse of distance weight matrix