

Geographically Weighted Panel Regression

Fernando Bruna (f.bruna@udc.es)^a
Danlin Yu (yud@mail.montclair.edu)^b

^a University of A Coruña, Economics and Business Department, Campus de Elviña s/n, 15071 A Coruña, Spain. Phone number: +34 981-167000. Fax number: +34 981-167070.

^b Department of Earth & Environmental Studies, Montclair State University. Montclair, NJ, 07043 USA. Phone number: +1 973-6554313. Fax number: +1 973-6554072.

Abstract

This paper discusses Yu's (2010) method of Geographically Weighted Panel Regression (GWPR) and estimates a New Economic Geography (NEG) 'wage-type-of equation' derived by Bruna (2013) using panel data for 206 European regions. Unlike cross-sectional or pooled models using data in levels, panel data models with fixed effects need to be interpreted in terms of changes of variables because of the time-demeaning transformation. GWPR is based on local panel estimates using weighted data for subsamples of nearest locations. Contrary to other extensions of GWR, this method finds local estimates considering the information of the whole sample period.

Surprisingly only Paredes and Iturra (2012) have applied cross-sectional GWR to a wage equation. However, an explanatory variable of Market Potential in this equation is particularly attractive to study local variations. Particularly, a Harris's (1954) variable of Market Potential is an inverse distance weighted sum of the income of the other regions in the sample. Using this weighting scheme the estimates for demeaned Market Potential mainly collect local spillovers from the variations of income in the nearest neighbors. Therefore, the GWPR estimates with fixed effects show how these local spillovers change across Europe. The estimates of Market Potential are found to be substantially higher for Portugal, Spain, South of France and North of Italy.

Keywords:

Local models, GWPR, fixed effects panel data, NEG, wage equation, Market Potential, European regions

JEL codes: C23, C51, R12

Draft 1.4 - October, 2013

1. Introduction

Geographically Weighed Regression (GWR) is the most frequently used method to study spatial nonstationary relationships due to intrinsically different relationships across space or due to misspecification of a global model. GWR is a useful tool for regional analysis and policy making (Ali et al., 2007; McMillen and Redfearn, 2010). The method was pioneered by Brunsdon, Fotheringham, and Charlton (1996) and McMillen (1996). It is basically a repeated estimation of a local regression at each point in space with a subsample of cross-sectional data properly weighted according to their proximity to each regression point. It has been applied many times in cross-sectional settings.

The possible different weighting functions to be used are determined by the concept of spatial closeness selected by the researcher. Crespo et al. (2007), Huang et al. (2010), Wrenn and Sam (2012), Yu (2013) and Wu et al. (2013) extended the concept of ‘closeness’ in GWR: data points close in both space and time dimensions can have a greater influence in the estimations of local parameters for an observation. The name ‘geographically and temporally weighted regression’ (GTWR) is appropriate to describe the procedure used by these authors. However, their approach is basically an extension of the cross-sectional GWR weighting function. The methodology followed here is different because it obtains local estimates with econometric techniques which are by their very nature based on repeated data points for every spatial observation: pooled and panel data estimations.

For the type of data used here the estimates obtained when pooling data in levels for cross-sections of different periods tend to be similar to those obtained by the separate estimation of each cross-section. Therefore, the paper focuses on the GWR extension to panel data estimation, particularly to panel data with fixed individual effects. The relevance of this contribution is given by the fact that the estimation with fixed unobserved effects is done with time-demeaned data. Therefore, the resulting estimates have a different interpretation from the cross-sectional (or pooled) estimates with variables in levels. The ‘within’ transformation makes the regression model to be interpreted in terms of the effects of changes of the explanatory variables on the changes of the dependent variable (‘changes’ with respect to the regional means). Contrary to cross-sectional economic data in levels for countries or regions, time-demeaned data can be highly volatile. Therefore, when pooling demeaned data the estimates might be very sensitive to the inclusion of time effects in the specification and can be very different from those obtained by pooling variables in levels. Indeed fixed effects panel estimates are likely to have more similar magnitude to the estimates obtained when the data is pooled after taking first differences or after taking differences of more periods. Therefore, cross-sectional or pooled estimates using data in levels are not comparable with pooled estimates using demeaned or differenced data. In summary, during 15 years the GWR literature has developed a set of arguments about the convenience of studying spatial nonstationarity in cross-sectional regressions, omitting panel data models. This omission is very relevant because panel data models are not particular cases of cross-sectional regressions but produce estimates with a very different meaning. Though both approaches introduce the time dimension, GWTR estimates are not comparable with the estimates that can be obtained under a local approach to panel data estimation.

There are only two basic references paying attention to the use of GWR with econometric techniques involving several periods, which have been published out of the main econometric circles. On one hand, Yu (2010) has proposed the Geographically Weighted Panel Regression (GWPR), applying standard panel data techniques to locally weighted subsets of the data based

on the spatial dimension. Cai, Yu and Oppenheimer (2012) make another application of Yu's method. On the other hand, Lin (2011) derives maximum likelihood estimators of spatial panel data GWR models (SPDGWR), considering both spatial autocorrelation and heterogeneity. At the moment there is no published empirical implementation of this last approach. This paper follows Yu's (2010) GWPR method of subsetting the data for each observation before estimating with panel data methods. The idea is simple but powerful because it allows obtaining local panel data estimates and it can be extended to local spatial panel data models in future work.

This paper presents GWPR in a particular context, the very much studied wage equation of the New Economic Geography (NEG). A variable of Market Potential or Market Access in this equation collects locational information, generally using geographical distances. However, as far as the authors know, Paredes and Iturra (2012) provide the only (cross-sectional) GWR estimation of a wage-type of equation in the large NEG empirical literature. Any measure of Market Potential with a distance exponent close to -1 mainly gathers information about the nearest neighbors. Therefore, it can be said that the GWR estimates of Market Potential are a kind of 'local estimates of local (neighboring) effects'. GWPR estimates of Market Potential capture local differences of regional spillovers from the variations of the GVA of the nearest neighbors. Given the novelty of this idea part of the paper is devoted to explain it.

Additionally, the paper emphasizes the differences between regressions with data in levels and demeaned data. This emphasis has the double intention of highlighting the contribution of GWPR and showing some aspects that must be considered for future work. At this early stage of the development of the GWPR approach, some issues, such as the role of time effects in local fixed effects estimates, are not addressed here. Too the paper stresses the software approach through different packages of R¹ which are related with the work presented here. The empirical part of the paper shows a GWPR exploratory analysis of an NEG wage-type-of equation for the European regions, which reveals the spatial heterogeneity of the local estimates.

The rest of the paper is organized as follows. The next section presents the NEG theoretical framework and the econometrics of panel data. The following section shows why a variable of Market Potential can be considered as an indicator of neighboring effects. The fourth section discusses the GWPR method in the context of the nonparametric literature and the subsequent section presents the empirical analysis. A final section concludes and an Appendix describes the sample and data.

2. Theoretical and econometric framework: NEG's wage equation and panel data

The so called "wage equation" of the NEG predicts that regional wages are a function of the size of the markets available to each region. This equation is considered to be very successful in the empirical literature (Redding, 2011). The wage equation in Fujita et al. (1999, chap.4) has been extended by Head and Mayer (2006) to control for human capital, building on Redding and Venables's (2004) version of the model. Bruna (2013) follows a similar approach to include capital stock per worker in the equation. The theoretical equation is derived for the agglomerating sector of region i , though the empirical literature frequently uses data for the total regional economies, as it will be done here. A version of the cross-sectional "wage-type of equation" in logarithmic form for region $i = 1, \dots, n$ can be:

$$\ln w_i = \alpha + \beta_1 \ln k_i + \beta_2 \ln h_i + \beta_3 \ln RMP_i \quad (1)$$

¹ <http://www.r-project.org/>

where w_i are wages, k_i is per capita capital stock, h_i is per capita human capital stock and the RMP_i term is called Market Access by Redding and Venables (2004) and Real Market Potential by Head and Mayer (2006). The Real Market Potential of region i is an indicator of its accessibility to the markets and includes trade cost and a measure of the degree of competition (“supply index”) in those markets. Equation (1) has an intercept (α) derived from the parameters of the model that are assumed to be common in all regions in the basic setting, especially total factor productivity.

The control variables can be considered as proxies for exogenous time-varying regional productivity differences. Alternatively, a regional variable of total factor productivity can be added to the equation to justify unobserved time-invariant fixed effects in an empirical estimation, as in equation (3) below. As discussed by Breinlich (2006), human and physical capital can be considered endogenous under a NEG setting (Redding and Schott, 2003). However, the extension of the wage equation to control for capital stock allows reducing possible biases in the estimates of Market Potential (Fingleton, 2006) and obtaining more cautious estimates of the effect of Market Potential. Additionally, capital stock can collect the exogenous European regional and transport policies and its inclusion in the equation allows the comparison of a wage-type-of equation with an expanded production function. The observational equivalence (Head and Mayer, 2004) of the NEG’s wage equation means that there is other theories that are consistent with the data, apart from the explanation offered by NEG. The dependent variable is frequently proxied by income per capita or per worker and, even when data on wages is used, wages proxy productivity too (Feldstein, 2008). Therefore, an empirical cross-sectional wage type-of equation is similar to a development accounting exercise with a production function including neighboring effects (Bruna, 2013; Bruna et al., 2013). These latter effects are collected by the empirical variable representing Real Market Potential, as it will be discussed in the next section. Theoretically Real Market Potential (RMP_i) collects the demand of other regions to region i , which is endogenous under the general equilibrium setting of the NEG. Instrumental variables estimation is not used here in order to focus on the exploration of the local variations of the model.

Generalizing the notation, an estimable extension of the cross-sectional equation (1) to pooled data of T periods can be represented as:

$$y_{it} = \alpha + \beta' x_{it} + u_t + u_{it} \quad (2)$$

where $t = 1, \dots, T$ and u_t are $T - 1$ possible common shocks to all regions in each period. The term u_{it} collects the effects of omitted variables and departures from the assumptions of the theoretical model.

As discussed by Wooldridge’s (2010) and other panel data authors, using the same notation, a panel data extension of equation (2) including unobserved time-invariant regional individual effects, u_i , is:

$$y_{it} = \alpha + \beta' x_{it} + u_i + u_t + u_{it} \quad (3)$$

where u_i collects omitted regional variables which are assumed to have an approximately constant role to explain the temporal levels of y_{it} in different regions. In a panel model with fixed effects, the unobserved u_i are removed through a within transformation of the data. Averaging equation (3) over $t = 1, \dots, T$ produces the following cross-sectional equation:

$$\bar{y}_i = \alpha + \beta' \bar{x}_i + u_i + \bar{u}_t + \bar{u}_i \quad (4)$$

and subtracting equation (4) from equation (3) produces the estimable fixed effects panel model, with the variables in deviations to the regional means:

$$y_{it} - \bar{y}_i = \beta'(x_{it} - \bar{x}_i) + (u_t - \bar{u}_t) + (u_{it} - \bar{u}_i) \quad (5)$$

where $(u_t - \bar{u}_t)$ is equivalent to $T - 1$ period dummies. This model can be estimated by standard OLS by pooling the demeaned data and the results allow estimating the unobserved fixed effects ($\hat{u}_i$). This advantage has the cost that all observable time-invariant variables are removed from the estimation too through the within transformation ($x_{it} - \bar{x}_i = 0$).

Keeping the same notation for the parameters, an alternative way of removing the unobserved u_i from equation (3) is the first differences transformation:

$$y_{it} - y_{it-1} = \alpha + \beta'(x_{it} - x_{it-1}) + (u_t - u_{t-1}) + (u_{it} - u_{it-1}) \quad (6)$$

where an intercept is usually added and $(u_t - u_{t-1})$ is equivalent to $T - 2$ time dummies. The same than equations (5), equation (6) can be estimated with OLS by pooling the transformed data.

Though the same β notation is used here for simplicity, the interpretation of the estimated coefficients ($\hat{\beta}$) is completely different in model (2), using variables in levels (x_{it}), when compared to models (5), using demeaned data ($x_{it} - \bar{x}_i$), or to model (6), using first differenced data ($x_{it} - x_{it-1}$). The different meaning of the results is even clearer starting from an equation like (1), where the x_i “levels” of the variables are already transformed through logarithms: $y_{it} = \ln Y_{it}$ and $x_{it} = \ln X_{it}$. The first difference of the logarithm of a variable is the instantaneous growth rate of the variable. Therefore, keeping the notation for the coefficients of the explanatory variables, in discrete time equation (6) is similar to a pooled model of one-period growth rates:

$$g_{y_{it}} = \alpha + \beta' g_{x_{it}} + v_{it} \quad (7)$$

Equations (6) and (7) are not identical to the fixed effects model in equation (5) but they are based in changes of variables (in logarithms) too. The ‘within’ transformation makes the fixed effects panel data model to be interpreted in terms of the effects of changes of the explanatory variables on the changes of the dependent variable (‘changes’ with respect to the regional means).

This creates at least four related issues which are relevant when comparing the interpretation of cross-sectional/pooled models and fixed effects panel models, at least with the type of data used in this research. First, the significance of the variables can change dramatically when pooling data in levels or demeaned data. Pritchett (2001) and Boulhold et al. (2008) comment about negative or non-significant estimates of human capital with country data and panel data with fixed effects. The results below (Table 1) confirm the non-significant role of human capital in the panel estimation.

Second, time-demeaned data, as well as growth rates, can be highly volatile. On the contrary, the pooled model (2) compares the relative levels of the variables in logarithms, which vary smoothly in time. Indeed, fixed effects panel estimates are likely to have more similar magnitude to the estimates obtained when the data is pooled after taking first differences or after taking differences of more periods. They tend to be very different from the estimates obtained by pooling variables in levels.

Third, because of the previous reason, the estimates obtained when demeaned data is pooled might be very sensitive to the inclusion of time effects in the specification (Table 1). However, cross-sectional estimates with variables in levels tend to be similar for different time periods and the estimates of a pooled model in level tend to be pretty insensitive to the inclusion of time effects.

Fourth, the cross-sectional dispersion of the variables in cross-sectional or pooled models in levels is related with the regional relative wealth and its spatial distribution in different samples. For instance, the levels of regional income tend to vary smoothly over the European space there-

fore they are highly spatially autocorrelated. But the regional income variations are less correlated with the GVA variations of the neighbors (see Table 2 below). This is a key distinction when one variable in the model collects neighboring effects, as in spatial econometrics models or, in a different form, as in an equation including a variable of Market Potential.

This discussion is especially relevant when the set of explanatory variables includes the logarithm of stock variables. For instance, the growth rate of per capita capital stock is similar to the ratio of investment to capital stock divided by the growth rate of population. However, the purpose of the paper is to illustrate the use of GWPR in a wage-type of equation. No attempt is done in this paper to adapt the panel estimations to growth theory (Boulhol et al., 2008).

In summary, the estimates obtained with panel data models (with fixed effects) are not comparable with those obtained with cross-sectional and pooled model. Therefore, the lack of an analysis of spatial nonstationary relationships in panel data models is a shortcoming of the previous (cross-sectional) GWR literature. In order to redress this situation, this paper presents the GWPR approach. Before that, the following section emphasizes again how different are the results of a pooled and a fixed effects panel model as a way of motivating the need of GWPR.

3. The data. Market Potential as an indicator of neighboring effects

The Appendix provides details about the sample and the variables. Human capital is proxied by the share of the population who has successfully completed education in Science and Technology (S&T) at the third level and is employed in a S&T occupation. Missing data in this variable were imputed with a polynomial of degree 2 on the regional time trend of each region. In a similar way to some other NEG's empirical research, wages are proxied by per capita income, measured as per capita gross value added (GVA). Real Market Potential (RMP_{it}) is proxied by a Harris's (1954) measure of Market Potential, built with GVA too. Norway and Switzerland are excluded from the sample because of lack of capital stock data though their regions are included in the measure of Market Potential.

Harris (1954) defines an indicator of the Market Potential of region i as an inverse distance weighted sum of the market size (GVA) of all the other regions in the sample. The same than in NEG derived measures of RMP_{it} , trade costs are proxied by physical distances², though a proxy for the non-observable competition index is omitted in Harris's Market Potential. This measure is regarded as a rough proxy of the NEG concept of RMP_{it} or Market Access (Combes et al., 2008, p.305), but both Breinlich (2006) and Head and Mayer (2006) find similar results for the European regions with a Harris's definition of Market Potential than with more sophisticated proxies calculated using gravity equations as Redding and Venables (2004). A full measure of Market Potential should include a proxy for the internal market size of each region. Omitting the internal markets introduces measurement error by reducing the access measure of some economically larger locations (Breinlich, 2006; Head and Mayer, 2006). This is especially relevant for the regions of Stockholm, Brussels, Berlin, Hamburg, Madrid, Paris, Vienna, Athens and (Inner) London (Bruna et al., 2013). However, the measurement of the internal market size is problematic (Frost and Spence, 1995; Kordi et al., 2012) and its inclusion severely aggravates the endogeneity problems of a wage-type-of equation (GVA in both sides of the equation). Moreover, the inclu-

² Actually, the interpretation of empirical results is more general because physical distances proxy "relative" trade costs (Yotov, 2012) and capture non-trade-related barriers (Linders et al., 2008) and interactions (Rodríguez-Pose, 2011) too. The estimation of a wage equation using any measure based on distances is sensitive to these factors.

sion of internal markets makes more difficult to interpret the estimates of Market Potential in terms of location. Given that the focus of this paper is the geographical distribution of local parameters, only the external markets are considered when measuring Market Potential, as Brakman et al. (2009) and other authors do.

The inverse distance weighting scheme in Harris's measure of (External) Market Potential can be justified by the robust finding in the gravity equations literature of a trade elasticity to distance close to -1 (Head and Mayer, 2013), though this number is based on trade statistics at the country level. Market Potential is a nonstandardized inverse distance weighted spatial lag of income using all the observations in the sample, instead of being built with the typical row-standardized weights for a few neighboring regions used in Spatial Econometrics³. But the inverse distance weighting scheme implies strong distance decay, especially for peripheral regions. Any measure of Market Potential or Market Access based on distance exponents close to -1 overweighs the nearest neighbors. Bruna (2013) ran 220 regressions of the cross-sectional equation (1) by building the variable of Market Potential only with the information of the first nearest neighbor, then with the two first nearest neighbors and so on until the 219 neighbors in the data set. The conclusion is that Market Potential is significant because it captures the locational information of the nearest neighbor. All the other neighbors do not add relevant information to the regression. In the same direction, Bruna et al. (2013) conclude that when working with time-demeaned data or first differences, the variations of Market Potential are an indicator of the income or production spillovers from the nearest neighbors. Given that a Harris's measure of Market Potential collects neighboring effects, a wage-type-of equation is especially attractive to be analyzed with GWR techniques, as Paredes and Iturra (2012) do. In the context of panel data, the GWPR estimates of Market Potential capture local differences of regional spillovers from the variations of the GVA of the nearest neighbors.

4. Baseline global models: pooled and fixed effects panel estimation

Before getting into GWPR, Table 1 presents some global models to emphasize the differences between cross-sectional/pooled estimations and panel data with fixed effects. Columns (4)-(6) show the results of estimating equation (3) without time effects, including time dummies and replacing them with a time trend, respectively. For comparative purposes columns (1)-(3) show the pooled estimations for the analogous specifications with the variables in levels. These pooled estimations are similar to cross-sectional estimations for particular years and gather information about the relative levels of variables. In a sample period of 14 years it can be expected that the regions that were relatively poor at the beginning of the period continue to be relatively poor at the end of the period, even if there is absolute convergence as it happens to be the case in this sample (Bruna et al., 2013). On the contrary, the panel estimations with fixed effects use time demeaned data, i.e., variations of variables, changes of variables with respect to the regional means in the sample period. The results of the panel estimations could change if each cross-sectional panel is built with data of several years. Here, year by year time demeaned data is used, as it is typical in panel estimations with fixed effects and more comparable with the pooled esti-

³ The standardization of inverse distance (summing to one) makes the weights to lose the economic interpretation of a distance decay (Anselin, 1988, pp.23–24).

mations. This means that the fixed effects panel estimations shown here collect average short-run (one year) effects.

Several conclusions can be drawn from Table 1. First, the estimation with time demeaned data in columns (4)-(6) alters the significance of the variables when comparing with the pooled estimation with data in levels. As mentioned above, human capital is not significant anymore (omitted from the table), probably due to the smooth changes of this variable. Second, the magnitude of the estimated coefficients changes dramatically when comparing both estimations methods. The estimates of per capita capital stock decrease around 70% when pooling demeaned data instead of levels. Capital stock collects information about the accumulated investment in each region. Therefore, in cross-sectional or pooled regressions the dispersion of the levels of capital stock gathers long run information about the relative wealth of each region. This relative wealth is what is captured by a cross-sectional regression, so the pooled estimate of per capita capital stock is around 0.7. On the contrary, demeaned capital stock is closely related to investment in each particular period. Using the analogy of demeaned logarithms with data in first differences of logarithms, regional economic growth on a particular period depends on regional investment on that period but it is affected by a number of other variables and shocks. Therefore the (short run) panel estimates using regional individual effects are around 0.2.

Table 1. Pooled and panel estimations 1995-2008 for 206 European regions

	<i>Pooled estimation (levels)</i>			<i>Panel with regional fixed effects</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	1.734*** (0.118)	1.475*** (0.116)	1.489*** (0.116)			
Per capita capital stock	0.646*** (0.010)	0.679*** (0.010)	0.678*** (0.010)	0.171*** (0.017)	0.188*** (0.017)	0.178*** (0.017)
Human capital	0.149*** (0.008)	0.165*** (0.008)	0.166*** (0.008)			
External Market Potential	0.139*** (0.007)	0.139*** (0.006)	0.139*** (0.006)	0.610*** (0.025)	0.984*** (0.083)	0.854*** (0.058)
Trend			-0.010*** (0.001)			-0.006*** (0.001)
Year dummies?	No	Yes	No	No	Yes	No
R-squared	0.792	0.806	0.805	0.793	0.796	0.795
Adj. R-squared	0.791	0.801	0.804	0.736	0.735	0.737
F	3662	743	2978	5134	692	3457
Sum sq. errors	68.43	64.00	64.13	4.37	4.32	4.34

Note: Table displays coefficients: * significant at 10% level; ** at 5% level; *** at 1% level. Standard errors are in brackets. All the variables are in logarithmic form. The dependent variable is per capita GVA. The total number of observations is 2884.

On the contrary, the estimates of External Market Potential increase between 300 and 600% when estimating with unobserved regional individual effects. The estimate goes from 0.1 when pooling data in levels to 0.6-1.0 when pooling demeaned data⁴. The inverse distance weighting scheme of the Market Potential variable makes the changes of this variable to collect growth spillovers. The high impact of the variation of GVA of a region on the variation of the per capita GVA of its close neighbors can be due to the (short run) diffusion of economic growth in space through trade, foreign direct investment, expectations, policies and knowledge spillovers. How-

⁴ The qualitative result is the same when Table 1 is repeated but omitting the variables of physical and human capital. The estimate of Market Potential is 0.4 for data in levels and 0.8-1.2 for demeaned data.

ever, the correlation of cross-sectional relative levels of Market Potential with the relative levels of per capita GVA might be due to NEG related channels of access to the markets or to common institutional and historical characteristics among close neighbors. This spatial similarity among the levels of GVA of neighboring regions is shown to have lower effects when explaining the cross-sectional dispersion of the levels of the per capita GVA.

As it was mentioned above, for each region the variations of variables are more volatile in time than the levels of variables. A third conclusion from Table 1 is that the inclusion of time effects is more relevant when using demeaned data than when using data in levels. Common shocks in each period (the European economic cycle) tend to extract from the dependent variable more information when that variable is measured in variations instead of levels. Similarly, replacing the year dummies by a time trend, as in columns (3) and (6), does not alter the estimates with pooled data but have some consequences on the magnitude of the estimates with demeaned data. When the data is pooled in levels, a time trend in the regression of a production function controlling for inputs can be considered to collect a common trend in the level of total factor productivity (TFP) during the sample period. However, a time trend in panel data with fixed effects must be interpreted as a constant variation of TFP (rate of technological progress). This is a strong assumption for short panels. With a sample period 1995-2008, as in this case, the variations of variables tend to present communalities along subperiods of the economic cycle. However, the regression diagnostics are similar in columns (5) and (6). For simplicity the specification with a time trend shown in column (6) of Table 1 is the baseline global model for the later GWPR.

Related to this, a final issue mentioned before is the dispersion and spatial distribution of the variables in levels and in deviations to the regional means of the sample period. Table 2 shows some statistics about this using the data of just one year. Cross-sectional data in deviations to the regional means of the whole period has higher dispersion and lower spatial autocorrelation than data in levels. For instance, the quartile coefficient of dispersion⁵ of the demeaned dependent variable is 20 times as greater as that of the variable in levels. Though the dispersion of demeaned Market Potential is higher too, there is no big difference from the dispersion of the levels of Market Potential because of the omission of the tails of the distribution and the smoothing role of the sum when building this variable. Additionally, the table shows the Moran's tests calculated using R spdep package (Bivand, 2013) for the variables in levels and after the within transformation. The variables are spatially autocorrelated in both cases, i.e., they present spatial clusters of high and low values⁶. However, this spatial correlation is lower for the variations of the variables because the short run changes of variables are more independent among close neighbors than the levels of the variables: high income regions tend to be located close to high income regions but economic growth in a particular year can differ more among neighbors. This is relevant because of the previous result of a high impact of the variations of Market Potential on the variations of per capita GVA in spite of the variations of both variables are less spatially autocorrelated than their levels.

⁵ The quartile coefficient of dispersion is a scale-free measure of dispersion omitting the tails of the distribution. It can be used when there are some negative values in a variable, contrary to the coefficient of variation.

⁶ Indeed the residuals of the models in Table 1 are spatially autocorrelated too, violating the OLS assumption of independence and calling for the estimation of spatial models. However, at this stage of development of GWPR, spatial local fixed effects panel models are still not studied here. Without the estimation of spatial model Cho et al. (2010) propose to minimize the spatial error Lagrange Multiplier test statistic for bandwidth selection of cross-sectional GWR models. Lin (2011) derives maximum likelihood estimators of spatial panel data GWR models.

Table 2. Dispersion and spatial autocorrelation of the variables for the cross-section of the year 2008

Variables	Data in levels			Data in deviations to the 1995-2008 means		
	Quartile coefficient of dispersion	Moran's test		Quartile coefficient of dispersion	Moran's test	
		I statistic	p-value		I statistic	p-value
Per capita GVA	0.018	0.618	0.000	0.351	0.463	0.000
Per capita capital stock	0.015	0.518	0.000	0.249	0.323	0.000
Human capital	-0.091	0.528	0.000	0.440	0.476	0.000
External Market Potential	0.037	0.919	0.000	0.050	0.773	0.000

Note: All the variables are in logarithmic form. Moran's tests use the randomisation assumption, which introduces a correction for departures from normality. The alternative hypothesis for the p-values is that Moran's I is greater than expected under the null hypothesis of absence of spatial autocorrelation. Zero p-values indicate the rejection of the null hypothesis and the presence of positive spatial autocorrelation. The weights matrix for this test is a row-standardized binary matrix to the 5 nearest neighbors.

The analysis in this section shows that the omission of panel data models is an important gap in the literature of geographically weighted regression. Additionally, it shows some aspects that are especially relevant in panel data, such as the role of time effects and spatial dependence. They will have to be carefully studied in later stages of the development of GWPR. The rest of the paper is devoted to present this method.

5. Methodology: geographically weighted panel regression

The standard econometric regression is based on the assumption that a “true” model exists that can be estimated conditional to a particular sample of data. On the contrary, the GWR model focuses on subsampling and weighting the data for each particular cross-sectional observation. The procedure allows studying how the estimates vary across space and comparing the results with those of a global model. The following summary about GWR is mainly based on the work by Charlton, Fotheringham and Brunsdon (2006), Charlton and Fotheringham (2009), Chasco, Vicéns and García (2008) and McMillen and Redfearn (2010).

As it was mentioned before, GWR is a technique developed for cross-sectional data. In order to obtain local estimates for the target location i , the data around location i will have to be properly weighted after defining the number of locations that are going to be subsampled to obtain local estimates for i (bandwidth). The literature of GTWR has extended this approach to consider time by assuming that data points close in both space and time dimensions can have a greater influence in the estimations of local parameters for i . The extension presented here is focused on econometric techniques which are by their very nature based on repeated observations for each location, with special attention to panel data with fixed effects. GWPR differs from GTWR in that once a bandwidth is chosen, all the time observations of each subsampled location has to receive the same weight in order to locally reproduce what a global model does when pooling data for different periods.

Before getting into these specifics it is useful to adopt a broad perspective about GWR. GWR is often considered a nonparametric procedure that fits individual regressions targeted to specific points, with more weight placed on observations that are closer to the target. The common practice is to use each observation, in turn, as the target point. McMillen (2010) and, specially, McMillen and Redfearn (2010) discuss the following most commonly employed nonparametric models. Only the two first methods can be considered totally non parametric because they use a local linear function to approximate a function $f(\cdot)$ that is constrained only to be smooth and continuous. Here the notation is simplified and two types of independent variables are considered, x and z :

- 1) The Locally Weighted Regression (LWR) model uses a weighting function (kernel function) to approximate $f(\cdot)$. At each target location i , the prediction of i is done weighting the differences of the normalized data of the independent variables with respect to the normalized data for the target point, $x_j - x_i$ and $z_j - z_i$. The kernel function (see below) determines the weight that observation j receives in estimating the value of y at target point i . Avoiding details, this can be generically represented as:

$$y = f(x, z) + u \quad (8)$$

- 2) The Kernel Regression model uses the kernel function based on $x_j - x_i$ and $z_j - z_i$ to weight the observations. However, only the values of y_j are used to predict y_i . This can be represented as:

$$y = f(y) + u \quad (9)$$

- 3) The Conditional Parametric Regression (CPAR) model is a special case of the LWR model which assumes that the set of explanatory variables (x, z) can be divided into portions that are fully nonparametric (x) and conditionally parametric (z) ⁷. For fixed values of z the model is a linear equation in which the coefficients vary with z :

$$y = \alpha(z) + \beta(z)x + u \quad (10)$$

The CPAR model becomes spatial when the conditionally parametric variables are the geographic coordinates of each point, latitude and longitude:

$$y = \alpha(la, lo) + \beta(la, lo)x + u \quad (11)$$

- 4) The Geographically Weighted Regression (GWR) model is a special case of the CPAR (LWR) model where the coordinates are replaced by straight-line distances (d) among observations and are typically omitted from the list of explanatory variables.

$$y = \alpha(d) + \beta(d)x + u \quad (12)$$

The GWR model appears to have first been used by McMillen (1996) and Brunsdon, Fotheringham, and Charlton (1996). McMillen (1996) uses the term LWR instead of GWR to emphasize that it is an application of the procedure developed originally by Cleveland and Devlin (1988). Fotheringham, Charlton and Brunsdon (1998) and McMillen and Redfearn (2010) consider GWR as a spatial extension of Casetti's (1972) expansion model, in which each parameter is allowed to deterministically depend on other variables. However, McMillen and Redfearn prefer to consider GWR as a part of the CPAR (LWR) family of models. From the point of view of Economics, the restrictive meaning of the word "geographically" in GWR has an advantage. In Economics, geographical coordinates have not meaning but distances can represent trade or informational costs, networking capacity, probability of common history or institutions. GWR is an established name and it is kept in this paper. But recognizing its status as a special case makes it easier to consider useful generalizations while providing a link to other literatures (McMillen and Redfearn, 2010)⁸.

In this context of multiple possibilities, the initial setup of GWPR appears to be a natural extension of the standard cross-sectional GWR model. Once a number of nearest neighbors is chosen to subsample the spatial observations, a weights matrix is built for each location i using the kernel function. Then, those weights are applied to all periods of the data in levels of the subsample of nearest neighbors of region i . After the weighting, all the panel data models available in

⁷ The semi-parametric model $y = \alpha(z) + \beta x + u$ is a constrained CPAR model where β does not vary with z .

⁸ For instance, the literature of nonparametric estimation of panel data models, reviewed by Gao and Li (2013) for the case of panel estimation with fixed effects.

R's plm package (Croissant and Millo, 2008)⁹ can be estimated to obtain local panel estimates for i . Therefore, in the case of panel data with fixed effects, the pooled estimates of weighted time-demeaned data consider the information of the whole sample period, as in a global panel data model, but using a local subsample for each regression point. Repeating the process for any location, a whole set of local GWPR estimates is obtained. In order to execute this process, some R's functions were developed based initially on R's package spgwr (Bivand and Yu, 2013), though in a later stage they will be extended to the framework of the recent package GWmodel (Gollini et al., 2013).

In GWR the choice of the kernel function used to select and weight each local sample has little effects on the results because the common functions share the property of declining weights with distance (Fotheringham et al., 2002; Yu, 2006). However, the bandwidth or distance to the target point is much more important because it determines how much an observation will be weighted and how such weights decline with distance. An adaptive bandwidth, also called 'window size', selects a different bandwidth h_i for each location, so the same number of nearest neighbors is considered for all the regression points. The adaptive bandwidth approach is often preferred due to its advantage that each regression point will have identical amount of local data points for local coefficient estimates¹⁰, albeit with different weighting schemes. The kernel function used here to weight the data of each local sample is an adaptive bisquare weighting function. The weight of the temporal observations of each location j in the estimation of the target point i is given by:

$$w_{ij} = \begin{cases} [1 - (d_{ij}/h_i)]^2 & \text{if } d_{ij} < h_i \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

A key issue when using GWR is bandwidth selection or, in this case, the selection of the number of nearest neighbors to be subsampled for each local estimation. This procedure is currently under development in GWPR. The cross-validation methods studied by Farber and Páez (2007) can be considered. An out of sample cross-validation procedure is being examined. Additionally, applying the AIC method to GWPR must consider the computation burden of dealing with T data points for each location.

One of the problems of GWR analyses and interpretation is that they are largely dependent on GWR maps. Apart from the general problems of choropleth maps when dealing with units of heterogeneous size, as it is the case in European regions, maps of the size of local parameters have a relative value if they do not inform about their statistical significance (Mennis, 2006; Wheeler, 2010; Matthews and Yang, 2012). In a similar way to what Mennis recommends, this visualization problem is reduced in the next section by excluding from the map the local coefficients with significance less than 90%.

Finally, Wheeler and Tiefelsdorf (2005) and Páez et al. (2011) has raised concerns about the potential correlations among local regression coefficients in cross-sectional GWR. R's packages gwrr (Wheeler, 2011) and GWmodel (Gollini et al., 2013) have developed techniques to measure and reduce problems of correlation. Those techniques can be extended to GWPR in future work.

⁹ A future extension of GWPR to spatial panel data models might use R's splm package (Millo and Piras, 2012) too.

¹⁰ The density of the observational units changes in space in most spatial datasets. The disaggregation level used in this paper for the European data is NUTS 2, which mainly follows national administrative divisions and combines countries with many observations with others where the data are sparse. NUT 2 territories have a range of population between 800.000 and 3 million people. Their average geographical area by country is very different too.

6. A GWPR exploratory analysis of a wage equation for the European regions

Except for obtaining an “optimal” bandwidth with either out-of-sample cross-validation or AIC measures, geographically weighted approaches can be used as exploratory tools too¹¹. A series of bandwidths can be selected and the resulting parameter surfaces examined at different levels of smoothing. In this sense, GW approaches are similar to a ‘spatial microscope’ (Fotheringham et al., 2002, chap. 6). Complex ‘movies’ can be constructed from GW methods by using a series of different bandwidths. At the current stage of research, the paper shows fixed effects panel data estimates of a wage-type-of equation for three different adaptive bandwidths of 15, 70 and 140 nearest neighbors. The choice is made of fairly exploratory nature to cover a relatively small, a somewhat average and a fairly large amount of nearest neighbors. The total number of possible nearest neighbors is 205. The estimated equation is the same than the one showed in column (6) of Table 1.

Figures 1 to 3 show quantile maps of the local fixed effects panel estimates of the logarithms of per capita capital stock and Market Potential with p-values lower than 0.1 for the three selected bandwidths. Figure 4 show the maps for local estimates of the trend for those bandwidths. Darker colors are associated with higher values of the variables. The first thing to note is that the lack of capital stock data for Norway and Switzerland could create edge effects when weighting the data of the nearest neighbors. The omission of the regions from these countries provokes visible consequences in the local estimates of per capita capital stock around Switzerland. This effect is not present in the global models of Table 1 because they are nonspatial models, assuming independence of the data for different locations. The exclusion of Norway and Switzerland from the sample does not affect the local estimates of Market Potential for two reasons. On one hand, the GVA of the regions of these two countries are considered when building the variable. On the other hand, even if those regions were omitted from the variable of Market Potential, the consequences would be limited given that the level of Market Potential is built as an inverse distance weighted sum of the GVA of all the other regions in the sample.

A first result which is relevant to validate the GWPR method is that the median local estimates for the three variables (calculating the median without excluding the insignificant estimates) are always very similar to the estimates of the global model. This means that the GW approach indeed localizes the global results, regardless of the bandwidth we choose.

However, the spatial distribution of the local estimates shows high heterogeneity. For the variable of per capita capital stock, there are significant negative estimates in the three first figures. For the sample period under consideration the panel data with fixed effects estimates for this variable for the regions of Spain and Portugal tend to be non-significant or negative.

On the contrary, with the exception of the regions of Greece and a few others, the local estimates of Market Potential tend to be significant and positive. A pattern seems to emerge in the local estimates for Portugal, Spain, South of France and North of Italy, especially in Figure 2. While the global estimate of Market Potential is 0.85 the local estimates range between 1.5 and 2.0 in these regions. Following the previous discussion about spillovers, this means that the variations of per capita GVA in regions of those areas are more sensitive to the variations of the GVA of their neighbors. In other words, the part of the variations of per capita income not explained by the variation of per capita capital stock in those regions is more dependent from their neighbors.

¹¹ The exploratory nature of GWR is more relevant when the goal of the researcher is to study causality. McMillen (2010) argues that the optimal bandwidth or window size is likely to be much larger when the objective is to estimate the marginal effect of x on y rather than to predict y directly. How much larger remains an open issue despite the voluminous literature on bandwidth selection.

Figure 1. Local GWPR estimates for 15 nearest neighbors (at least significant at 90%)

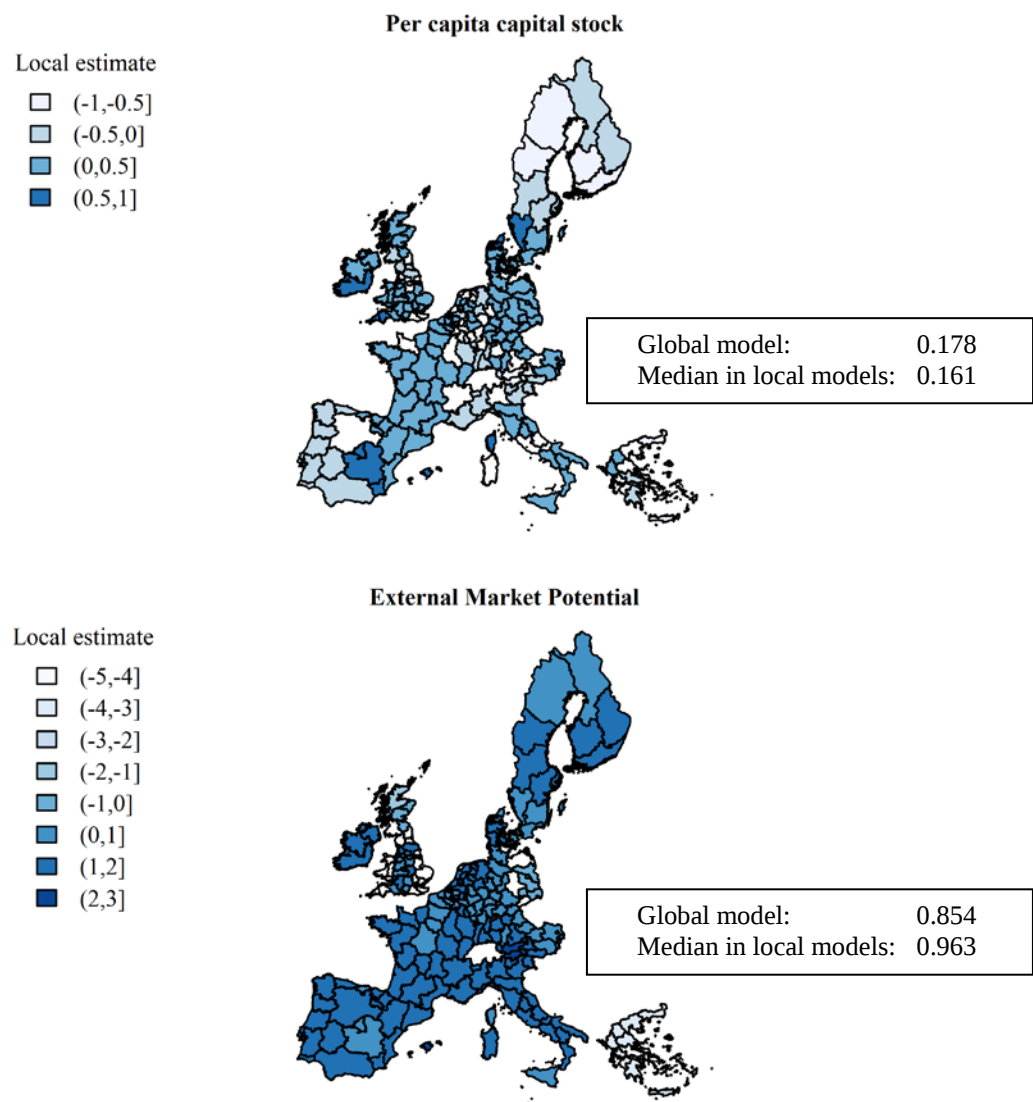


Figure 2. Local GWPR estimates for 70 nearest neighbors (at least significant at 90%)

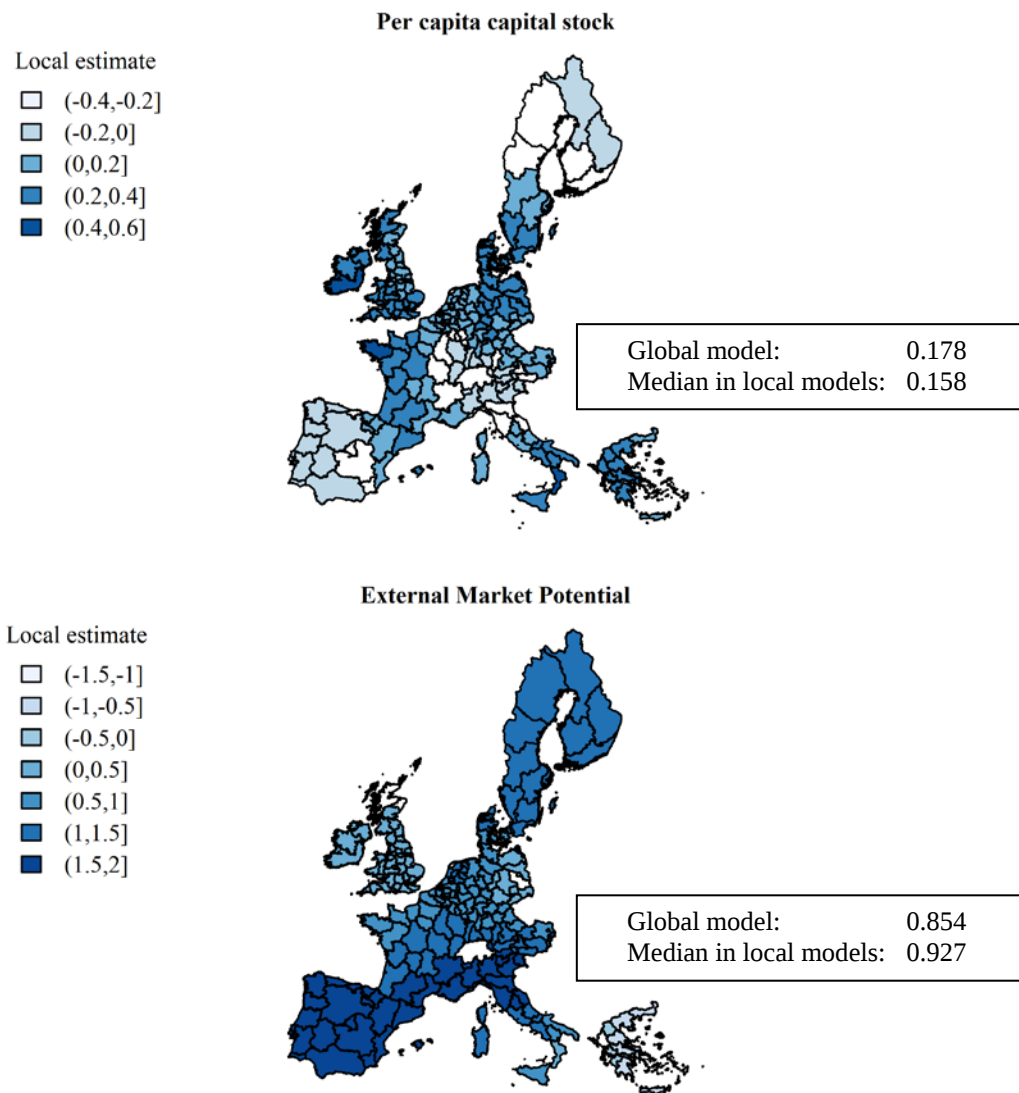


Figure 3. Local GWPR estimates for 140 nearest neighbors (at least significant at 90%)

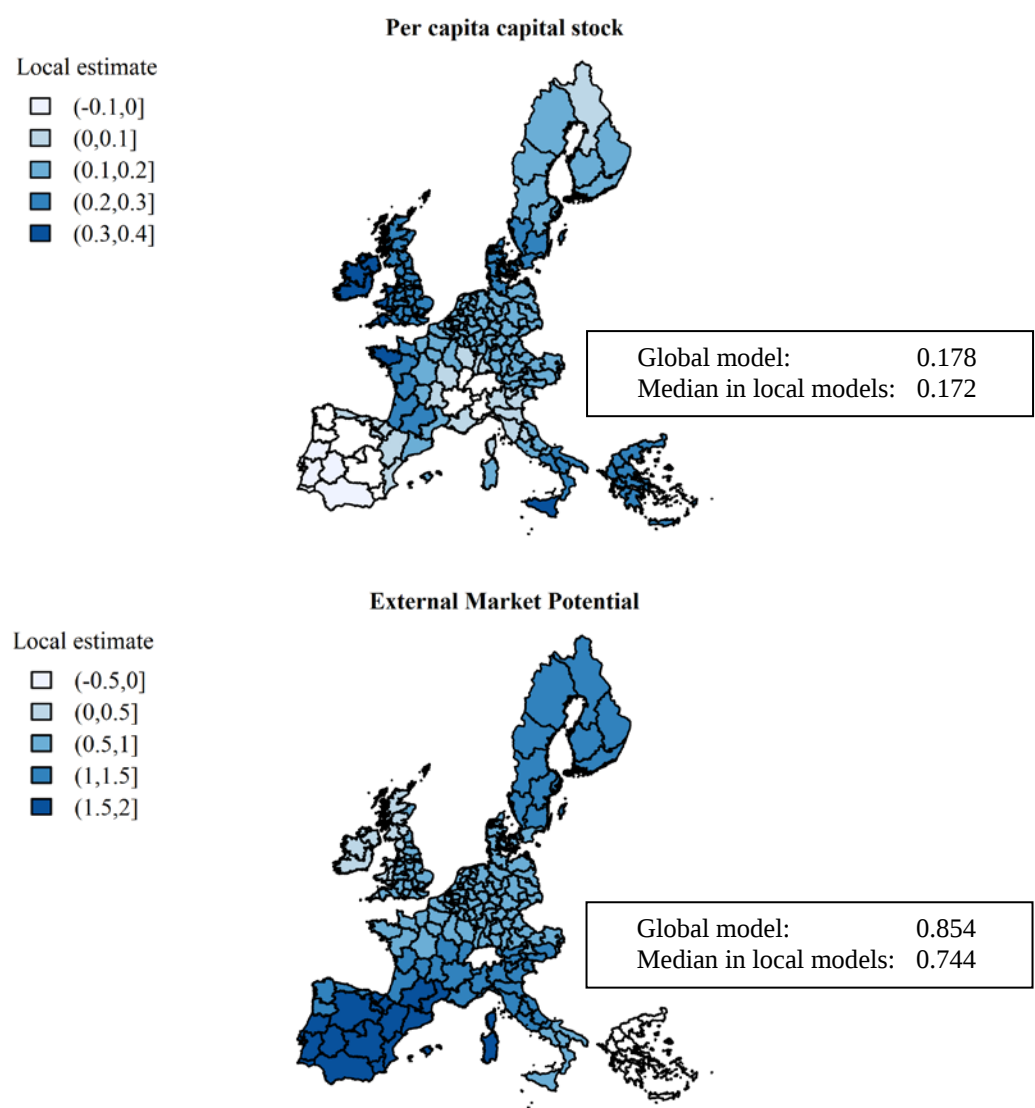


Figure 4 Local GWPR estimates for the variable “trend” (at least significant at 90%)

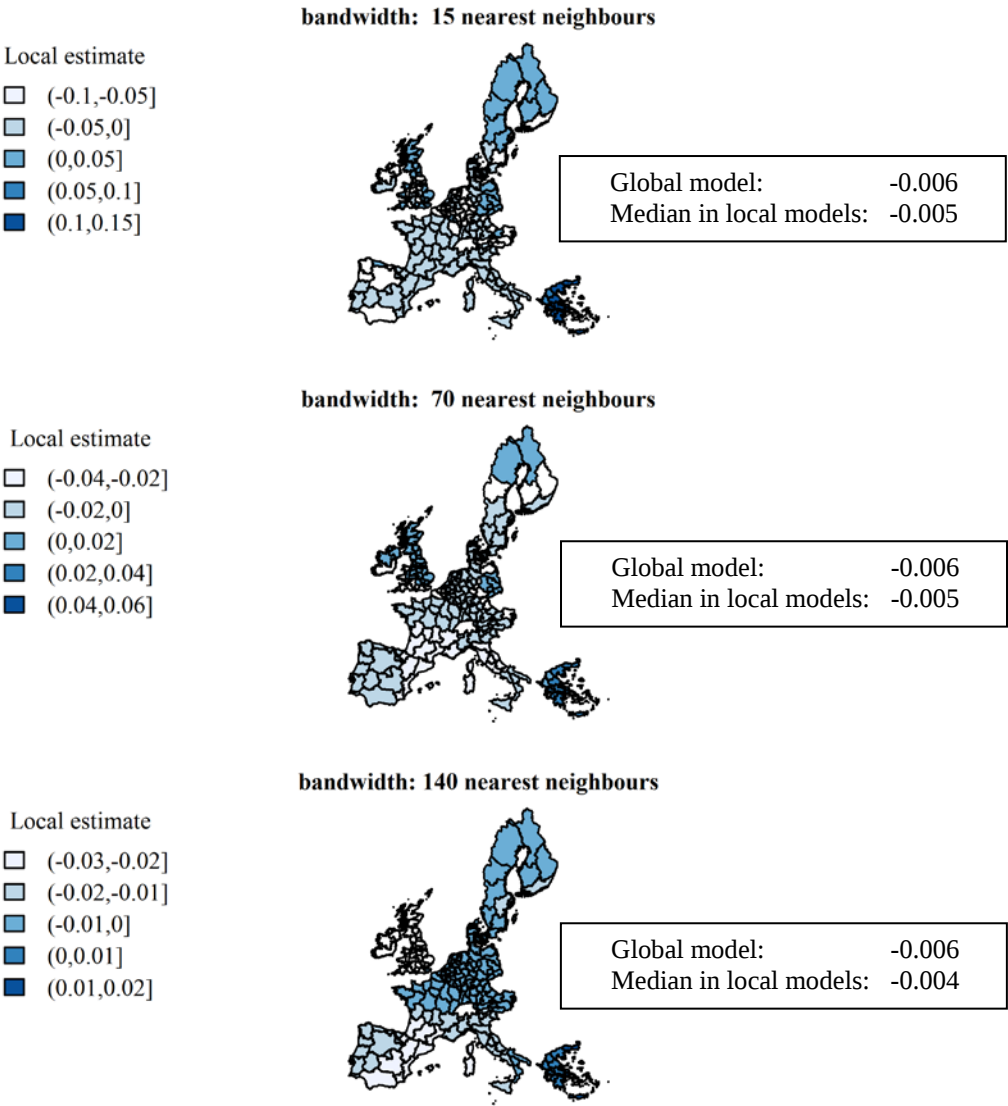
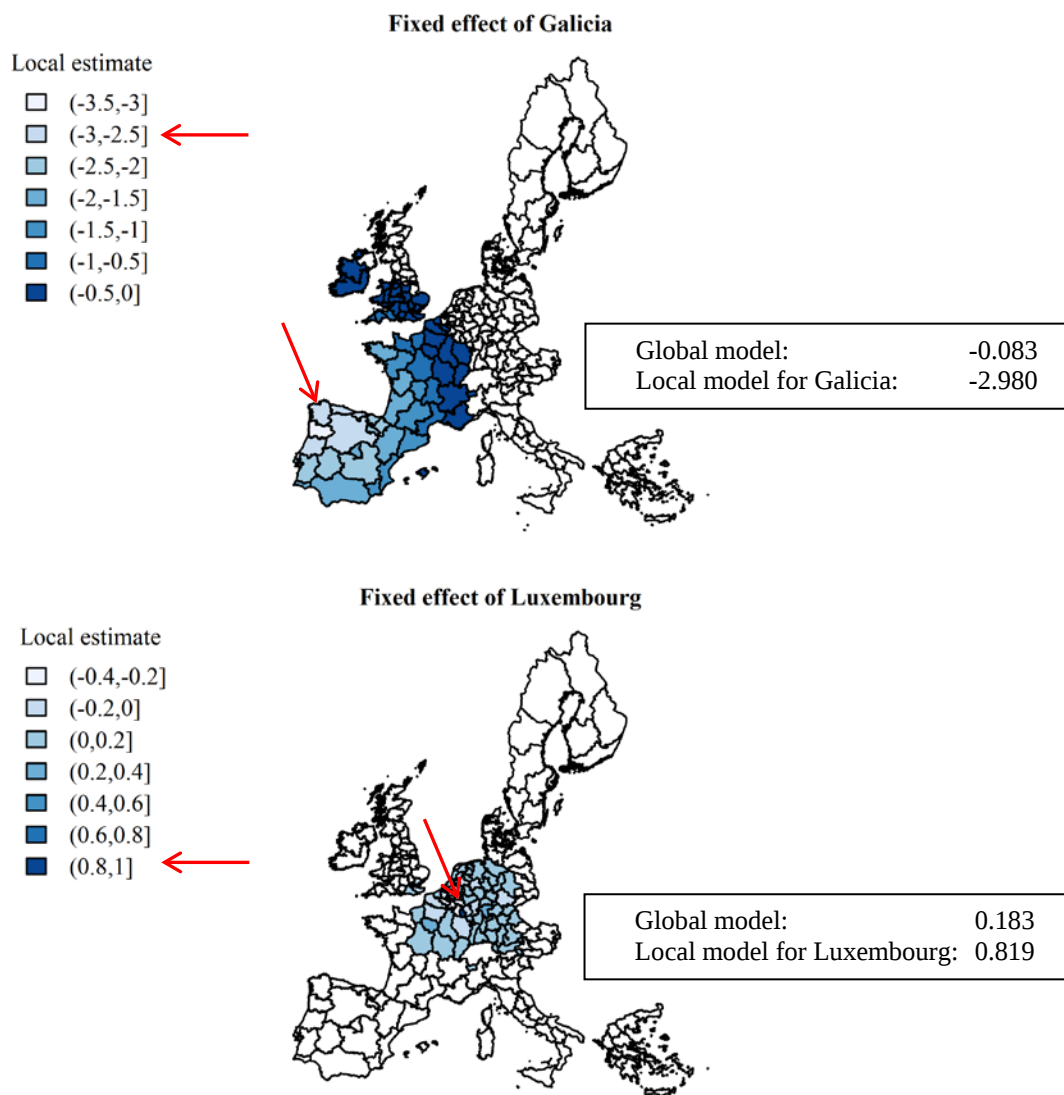


Figure 5. Fixed effect of two regions in all subsampled local estimations for 70 nearest neighbors



A final test of the results is show in Figure 5, just for the GWPR model estimated with 70 nearest neighbors. Given that in this case each region is subsampled and weighted in 71 panel data estimations, the maps show the level of the fixed effect for Galicia (Northwest of Spain) and Luxembourg in all the 71 estimations. This type of analysis is useful to study the sensitivity of the estimated local fixed effect to the bandwidth. However, the true estimated local fixed effect of Galicia and Luxembourg are those obtained in the local panel estimations for these two locations, which are market with an arrow in the maps. As it can be seen in the figure, the fixed effect of Galicia in the global model is -0.9 while it is -3.0 in the local estimation for Galicia including 70 nearest neighbors. With 140 nearest neighbors the fixed effect is similar, -2.7, but with 15 neighbors is very different, 2.0 (not show,). For Luxembourg, the fixed effect in the global model is 0.2, while it is 0.8 for a bandwidth of 70, 1.3 for 140 and 6.1 for 15 nearest neighbors. The estimated levels of the individual regional effects are very sensitive to sample selection.

7. Conclusions

Following Yu's (2010) approach this paper presents the GWPR method to fill the gap between the literature of GWR and the literature of panel data. The paper emphasizes the different interpretation of models estimated with cross-sectional or pooled data in levels and fixed effects panel data models, which use time-demeaned data. The marginal effects of demeaned data must be interpreted in terms of variations of variables. Therefore, GWPR allows studying spatial nonstationarity in models using time-demeaned data and, more generally, it can be considered a method extendible to any econometric technique based on repeated observations for each location.

This discussion is exemplified with a New Economic Geography wage-type-of equation. A novelty of this analysis is that GWPR estimates of a Harris's variable of Market Potential can be considered to capture local differences of regional spillovers from the variations of the GVA of the nearest neighbors. The exploratory analysis of the empirical part of the paper shows that when the data is subsampled and weighted, no matter what local sample sizes are arbitrarily chosen, the estimates from GWPR always fluctuate from the global estimates. This is to be expected for GW approaches. Although this research has not applied any bandwidth optimization technique as usually was done when applying cross-sectional GW analysis, the exploratory attempt does reveal some findings already. The local estimates of per capita capital stock tend to be negative or insignificant for Portugal and Spain. On the contrary, for 70 nearest neighbors, the local fixed effects estimates of Market Potential for Portugal, Spain, South of France and North of Italy have a magnitude which doubles the one of the global model.

GWPR is a method under development. The paper introduces some issues under current research and some future potential extensions, such as geographically weighted spatial panel data models. The future research foci of GWPR will be the development of optimization procedures, incorporating random effects and tests of whether or not fixed and/or random effects are significantly present.

Appendix: Data description

A.1 Sample of regions - The disaggregation level for the regional data is NUTS 2 (2006 version), which involves the basic regions for the application of regional policies. The sample includes 206 regions from 15 countries of the European Union: Austria, Belgium, Spain, Finland, France, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Sweden and United Kingdom. The following NUTS 2 regions are excluded: the Atlantic islands (the Spanish Canary Islands and the Portuguese Madeira and the Azores), the Spanish Ceuta and Melilla in the North African coast and the French Departments Guadeloupe, Guiana, Martinique and Reunion. Oil related regions are not excluded. Norway and Switzerland are omitted because of the lack of capital stock data but their 14 regions are included to compute Market Potential.

A.2 Variables - All the variables are in logarithmic form for the years 1995-2008. Cambridge Econometrics data is used for gross value added (GVA), capital stock and population. Per capita capital stock and per capita GVA are in 2000 year euros. Market Potential is built with GVA and it is in millions of 2000 euros.

Human capital stock (H_{it}) is proxied by the following Eurostat variable: share of the population who has successfully completed education in Science and Technology (S&T) at the third level and is employed in a S&T occupation. 9.7% of the observations 1995-2008 were missing

and imputed using R's Amelia II package (Honaker et al., 2011). The imputed data is the average of 5 multiple imputations with a small ridge prior predicting with a polynomial of degree 2 on the regional time trend of each region and including lags and leads: $H_{it} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 H_{it-1} + \beta_4 H_{it+1}$. This method allows imputing this control variable in mainly seven regions with high degree of missingness.

The External Market Potential of region $i = 1, \dots, n$ is defined as the inverse distance (d_{ij}) weighted sum of the GVA of all the other regions in the sample:

$$EMP_i = \sum_{j \neq i}^{n-1} \frac{GVA_j}{d_{ij}}$$

Geographical distances (d_{ij}) are measured as great circle distances among regional centroids calculated using GISCO's shape files (© EuroGeographics for the administrative boundaries).

References

- ALI, K., M. D. PARTRIDGE & M. R. OLFERT (2007), Can Geographically Weighted Regressions Improve Regional Analysis and Policy Making? *International Regional Science Review* 30, pp. 300–329.
- ANSELIN, L. (1988), *Spatial Econometrics: Methods and Models*. Springer.
- BIVAND, R. (2013), *spdep. Spatial dependence: weighting schemes, statistics and models. R package*.
- BIVAND, R. & D. YU (2013), *spgwr: Geographically weighted regression. R package*.
- BOULHOL, H., A. DE SERRES & M. MOLNÁR (2008), The contribution of economic geography to GDP per capita. *OECD Journal: Economic Studies* 2008, pp. 1–37.
- BRAKMAN, S., H. GARRETSEN & C. VAN MARREWIJK (2009), Economic Geography within and between European nations: The role of Market Potential and density across space and time. *Journal of Regional Science* 49, pp. 777–800.
- BREINLICH, H. (2006), The spatial income structure in the European Union—what role for Economic Geography? *Journal of Economic Geography* 6, pp. 593–617.
- BRUNA, F. (2013), *The observational equivalence of the NEG's wage-type-of equation* [“Economic Geography and Development in the European Space”, Chapter 3]. PhD. thesis. University of A Coruña, Spain.
- BRUNA, F., A. FAÍÑA & J. LOPEZ-RODRIGUEZ (2013), Market Potential and the curse of distance in the European regions. *Tijdschrift voor Economische en Sociale Geografie* (submitted).
- BRUNSDON, C., A. S. FOTHERINGHAM & M. E. CHARLTON (1996), Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis* 28, pp. 281–298.
- CAI, R., D. YU & M. OPPENHEIMER (2012), *Estimating the Effects of Weather Variations on Corn Yields using Geographically Weighted Panel Regression*, in 2012 Seattle, Washington: Agricultural and Applied Economics Association.
- CASETTI, E. (1972), Generating Models by the Expansion Method: Applications to Geographical Research. *Geographical Analysis* 4, pp. 81–91.
- CHARLTON, M. & A. S. FOTHERINGHAM (2009), *Geographically weighted regression. White paper*. National Centre for Geocomputation. National University of Ireland Maynooth.
- CHARLTON, M., S. FOTHERINGHAM & C. BRUNSDON (2006), *Geographically Weighted Regression*. NCRM Methods Review Papers 6. ESRC National Centre for Research Methods
- CHASCO YRIGOYEN, C., J. VICENS-OTERO & I. GARCÍA RODRÍGUEZ (2008), Modeling spatial variations in household disposable income with geographically weighted regression. *Estadística española* 50, pp. 321–360.
- CHO, S.-H., D. LAMBERT & Z. CHEN (2010), Geographically weighted regression bandwidth selection and spatial autocorrelation: an empirical example using Chinese agriculture data. *Applied Economics Letters* 17, pp. 767–772.

- CLEVELAND, W. S. & S. J. DEVLIN (1988), Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* 83, pp. 596–610.
- COMBES, P.-P., T. MAYER & J.-F. THISSE (2008), *Economic geography: the integration of regions and nations*. Princeton, N.J.: Princeton University Press.
- CRESPO, R., S. FOTHERINGHAM & M. E. CHARLTON (2007), Application of Geographically Weighted Regression to a 19-year set of house price data in London to calibrate local hedonic price models, in U. Demšar (ed.) *Proceedings of the 9th International Conference on Geocomputation*. Maynooth, Ireland: NCG, NUI.
- CROISSANT, Y. & G. MILLO (2008), Panel Data Econometrics in R: The plm Package. *Journal of Statistical Software* 27.
- FARBER, S. & A. PÁEZ (2007), A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. *Journal of Geographical Systems* 9, pp. 371–396.
- FELDSTEIN, M. (2008), Did wages reflect growth in productivity? *Journal of Policy Modeling* 30, pp. 591–594.
- FINGLETON, B. (2006), The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain. *Oxford Economic Papers* 58, pp. 501–530.
- FOTHERINGHAM, A. S., C. BRUNSDON & M. CHARLTON (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. 1st edition. Wiley.
- FOTHERINGHAM, A. S., M. E. CHARLTON & C. BRUNSDON (1998), Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A* 30, pp. 1905–1927.
- FROST, M. E. & N. A. SPENCE (1995), The rediscovery of accessibility and economic potential: the critical issue of self-potential. *Environment and Planning A* 27, pp. 1833–1848.
- FUJITA, M., P. KRUGMAN & A. J. VENABLES (1999), *The spatial economy: cities, regions and international trade*. Cambridge: The MIT Press.
- GAO, Y. & K. LI (2013), Nonparametric estimation of fixed effects panel data models. *Journal of Nonparametric Statistics* forthcoming, pp. 1–15.
- GOLLINI, I., B. LU, M. CHARLTON, C. BRUNSDON & P. HARRIS (2013), *GWmodel: an R Package for Exploring Spatial Heterogeneity using Geographically Weighted Models*. arXiv e-print 1306.0413.
- HARRIS, C. D. (1954), The Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers* 44, pp. 315–348.
- HEAD, K. & T. MAYER (2004), The empirics of agglomeration and trade, in J.V. Henderson & J.-F. Thisse (eds.) *Handbook of regional and urban economics*. North Holland. pp. 2609–2669.
- HEAD, K. & T. MAYER (2006), Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics* 36, pp. 573–594.
- HEAD, K. & T. MAYER (2013), Gravity Equations: Workhorse, Toolkit, and Cookbook. *CEPR Discussion Paper* 9322.
- HONAKER, J., G. KING & M. BLACKWELL (2011), Amelia II: A Program for Missing Data. *Journal of Statistical Software* 45.
- HUANG, B., B. WU & M. BARRY (2010), Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science* 24, pp. 383–401.
- KORDI, M., C. KAISER & A. S. FOTHERINGHAM (2012), A possible solution for the centroid-to-centroid and intra-zonal trip length problems, in J. Gensel, D. Josselin, & D. Vandembroucke (eds.) *Proceedings of the AGILE International Conference on Geographic Information Science*. AGILE. pp. 147–152.
- LIN, Z. (2011), ML Estimation of Spatial Panel Data Geographically Weighted Regression Model, in *2011 International Conference on Management and Service Science (MASS)*. pp. 1–4.

- LINDERS, G.-J. M., M. J. BURGER & F. G. VAN OORT (2008), A rather empty world: the many faces of distance and the persistent resistance to international trade. *Cambridge Journal of Regions, Economy and Society* 1, pp. 439–458.
- MATTHEWS, S. A. & T.-C. YANG (2012), Mapping the results of local statistics: Using geographically weighted regression. *Demographic Research* 26, pp. 151–166.
- McMILLEN, D. P. (1996), One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach. *Journal of Urban Economics* 40, pp. 100–124.
- McMILLEN, D. P. (2010), Issues In Spatial Data Analysis. *Journal of Regional Science* 50, pp. 119–141.
- McMILLEN, D. P. & C. L. REDFEARN (2010), Estimation And Hypothesis Testing For Nonparametric Hedonic House Price Functions. *Journal of Regional Science* 50, pp. 712–733.
- MENNIS, J. (2006), Mapping the Results of Geographically Weighted Regression. *The Cartographic Journal* 43, pp. 171–179.
- MILLO, G. & G. PIRAS (2012), splm: Spatial Panel Data Models in R. *Journal of Statistical Software* 47, pp. 1–38.
- PÁEZ, A., S. FARBER & D. WHEELER (2011), A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environment and Planning A* 43, pp. 2992–3010.
- PRITCHETT, L. (2001), Where Has All the Education Gone? *The World Bank Economic Review* 15, pp. 367–391.
- REDDING, S. J. (2011), Economic Geography: A Review of the Theoretical and Empirical Literature, in D. Bernhofen, R. Falvey, D. Greenaway, & U. Kreickemeier (eds.) *Palgrave Handbook of International Trade*. 1st edition Chapter 16: Palgrave Macmillan.
- REDDING, S. J. & P. K. SCHOTT (2003), Distance, skill deepening and development: will peripheral countries ever get rich? *Journal of Development Economics* 72, pp. 515–541.
- REDDING, S. J. & A. J. VENABLES (2004), Economic geography and international inequality. *Journal of International Economics* 62, pp. 53–82.
- RODRÍGUEZ-POSE, A. (2011), Economists as geographers and geographers as something else: on the changing conception of distance in geography and economics. *Journal of Economic Geography* 11, pp. 347–356.
- WHEELER, D. (2011), *gwrr: Geographically weighted regression with penalties and diagnostic tools*. R package.
- WHEELER, D. C. (2010), Visualizing and Diagnosing Coefficients from Geographically Weighted Regression Models, in B. Jiang & X. Yao (eds.) *Geospatial Analysis and Modelling of Urban Structure and Dynamics*. GeoJournal Library. Springer Netherlands. pp. 415–436.
- WHEELER, D. & M. TIEFELSDORF (2005), Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems* 7, pp. 161–187.
- WOOLDRIDGE, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. The MIT Press.
- WRENN, D. H. & A. G. SAM (2012), Geographically and temporally weighted likelihood regression: exploring the spatiotemporal determinants of land use change. *Regional Science and Urban Economics* (submitted).
- WU, K., B. LIU, B. HUANG & Z. LEI (2013), Incorporating the multi-cross-sectional temporal effect in Geographically Weighted Logit Regression, in L. Zhang & Y. Gu (eds.) *Information Systems and Computing Technology - Proceedings of the International Conference on Information Systems and Computing Technology, ISCT 2013*. London: CRC Press, Taylor & Francis Group. pp. 3–14.
- YOTOV, Y. V. (2012), A simple solution to the distance puzzle in international trade. *Economics Letters* 117, pp. 794–798.
- YU, D. (2006), Spatially varying development mechanisms in the Greater Beijing Area: a geographically weighted regression investigation. *The Annals of Regional Science* 40, pp. 173–190.

- YU, D. (2010), Exploring spatiotemporally varying regressed relationships: the geographically weighted panel regression analysis E. Guilbert, B. Lees, & Y. Leung (eds.). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38, pp. 134–139.
- YU, D. (2013), Understanding regional development mechanisms in Greater Beijing Area, China, 1995-2001, from a spatial-temporal perspective. *GeoJournal* (forthcoming).