Appendix A

A1. Estimation of total factor productivity (TFP)

To analyse productivity effects of FDI spillovers we consider linearized Cobb Douglas production function for firm i in time period t estimated for each industry and country separately that accounts for heterogeneity of different production technologies, quality and intensity of inputs. In logs, the production function to be estimated takes the following form:

$$ln(Y_{it}) = \beta_k ln(K_{it}) + \beta_l ln(L_{it}) + \beta_m ln(M_{it}) + \omega_{it} + v_{it}$$
(A1)

where Y_{it} , K_{it} , M_{it} denote natural logarithms of deflated sales, tangible fixed assets and material costs, respectively. L_{it} is the log of the number of employees for firm *i* at time *t*. Composite error term ($\omega_{it} + v_{it}$) includes firm specific productivity denoted by (ω_{it}) and (v_{it}) which represents measurement error or unanticipated productivity shock (Eberhardt and Helmers, 2010). Hence, for the construction of TFP sample we need information on firms' sales, tangible fixed assets, number of employees and expenditure on materials. Firms with missing, negative or zero values on either of the variables of interest are dropped from the sample. We have also eliminated observations for which accounting rules are violated.

If the unobserved productivity shocks are correlated with the firm's input choices, standard estimation techniques will yield biased input coefficient estimates that affect TFP estimates as well. Recently, Gandhi et al. (2016) argued that proxy variable methods suffer from identification issues. When it contains flexible inputs, the production function cannot be identified since there is not enough exogenous variation other than the inputs included in the function. Estimation of value-added production function may not solve the identification problem causing the dispersion and heterogeneity of TFP to be overstated and may end up with a distorted image of productivity distribution. Gandhi et al. (2016) propose to use gross output production function with at least one flexible input. They establish identification by exploiting information in the first order condition with respect to the flexible input from the firm's static profit maximisation problem. Apart from controlling for simultaneity and value added bias, the additional advantage of this estimator is that it can be used to estimate a wide arrange of functional forms; it does not rely on the strong assumption of scalar unobservability to invert the proxy demand function; nor it requires that the demand for intermediate inputs is strictly increasing in productivity. To account for simultaneity bias in Cobb Douglas production function estimation, we rely on Gandhi et al. (2016) approach.

Table A1: Model diagnostics						
Specification	(1)	(2)	(3)	(4)		
Spatial matrix	1 st neighbours only	1 st 1 st and 2 nd neighbours neighbours only equal weights		Inverse distance normalized matrix		
Baseline model						
Wald test H _o : (wX's=0)	12.57***	37.45***	35.86***	162.62***		
Wald test H_0 : ($\theta = -\beta \rho$)	22.61***	46.07***	45.37***	174.14***		
Manufacturing sector						
Wald test H _o : (wX's=0)	14.17***	46.37***	41.33***	51.82***		
Wald test H_0 : ($\theta = -\beta \rho$)	20.87***	50.97***	46.42***	59.08***		
Service sector						
Wald test H _o : (wX's=0)	13.02***	34.16***	33.61***	148.39***		
Wald test H_0 : ($\theta = -\beta \rho$)	22.95***	42.45***	43.02***	156.84***		
Number of regions	217	217	217	217		
Number of observations	1085	1085	1085	1085		

A2. Summary statistics and model diagnostics

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level of significance respectively. Source: Authors' calculations

Table A2: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
lnTFPall	1,085	2.420304	1.382995	-2.426913	6.411024
InTFPservice	1,085	0.9400552	1.248842	-4.60517	4.07375
InTFPmanuf.	1,085	2.13258	1.444654	-2.426913	6.320672
Horizontal	1,085	12.49287	10.07815	0	68.81055
Backward	1,085	11.90092	9.262812	0	58.73926
Forward	1,085	10.25198	8.626129	0	59.2662
Urbanization	1,085	3.686636	6.131696	0.0136045	59.80432
Localization	1,085	8.581083	4.053897	4.154872	55.55556
Hcapital	1,085	68.63952	38.30403	14.29933	188.517
Intangibles	1,085	2.61135	3.431242	0.0354925	24.03264

A3. Data representativeness

	Total economy		Services		Manufacturing	
	2006	2011	2006	2011	2006	2011
BG	0.94***	0.93***	0.99***	0.99***	0.89***	0.90***
CZ	0.96***	0.97***	0.98***	0.98***	0.77***	0.78***
HR	0.99***	0.99***	0.99***	0.99***	0.97***	0.97***
HU	0.98***	1.00***	0.96***	0.99***	0.92***	0.97***
PL	0.94***	0.94***	0.93***	0.84***	0.88***	0.83***
RO	0.90***	0.90***	0.99***	0.99***	0.87***	0.84***
SI	0.99***	1.00***	0.99***	0.99***	0.89***	0.94***
SK	0.97***	0.97***	0.97***	0.98***	0.69*	0.81***
Entire sample	0,96***	0.96***	0.90***	0.97***	0.94***	0.95***

 Table A3: Pairwise correlation coefficients of NUTS3 regions share in total employment of analysed countries in 2006-2011 according to the Eurostat SBS and Amadeus TFP sample

***,** and * denote statistical significance at 1%, 5% and 10% levels respectively

Table A4: Pairwise correlation coefficients of NUTS3 regions share in total value added of analysed countries in 2006-2011 according to the Eurostat SBS and Amadeus TFP sample

	Total economy		Services		Manufacturing	
	2006	2011	2006	2011	2006	2011
BG	0.76***	0.91***	0.99***	0.99***	0.73***	0.82***
CZ	0.99***	0.98***	0.99***	0.99***	0.94***	0.91***
HR	0.99***	0.99***	0.99***	0.99***	0.99***	0.99***
HU	0.93***	1.00***	0.90***	0.99***	0.89***	0.97***
PL	0.97***	0.95***	0.98***	0.95***	0.58***	0.45***
RO	0.99***	0.99***	0.99***	0.99***	0.92***	0.85***
SI	0.99***	0.99***	0.99***	0.99***	0.96***	0.95***
SK	0.98***	0.99***	0.99***	0.99***	0.84***	0.90***
Entire sample	0.92***	0.95***	0.92***	0.97***	0.88***	0.91***

***,** and * denote statistical significance at 1%, 5% and 10% levels respectively



Figure A1: NUTS3 regions share in **total employment** of analysed countries according to the Eurostat SBS and Amadeus TFP sample – total economy, manufacturing and services

Figure A2: NUTS3 regions share in **total value aded** of analysed countries according to the Eurostat SBS and Amadeus TFP sample – total economy, manufacturing and services

