Discussion paper (draft version)

Please don't quote.

Knowledge Does Not Fall Far from the Tree –

A Case Study on the Diffusion of Solar Cells in Germany

Johannes Rode[1], Alexander Weber[2]

March 1, 2011

Abstract

The purpose of this paper is to illuminate the geographical diffusion of photovoltaics installations in Germany quantitatively and to test if preexisting photovoltaic systems stimulate further installations nearby; thus we investigate to which extent knowledge flows depend on geographic proximity. We develop an econometric model, which is discrete in time and space, but the level of geographical agglomeration is adjustable in arbitrarily small steps. We find that the probability to install a photovoltaic system dependents on the geographic proximity to agents, who have previously installed a photovoltaic system. In conclusion, our results confirm that knowledge exchange attenuates with distance.

Key words: PV, photovoltaics, knowledge, technology, diffusion, spatial econometrics

JEL classification: O33, Q42, R12

[1]Contact: rode@stadtforschung.tu-darmstadt.de
[2]Contact: a.a.weber@gmx.de

1 Introduction

According to Marshall (1890), knowledge exchange decreases with distance. In other words, spatial proximity fosters knowledge diffusion between economic agents. Similarly, Audretsch and Feldman (1996) argue that transferring information may not be affected by distance, but exchanging knowledge is. Nevertheless, scholars of innovation do have problems to find evidence when it comes to defining and assessing the geographical dimension of knowledge externalities (Autant-Bernard, Mairesse and Massard, 2007). Krugman (1991a, p.53) even stated "knowledge flows are [...] invisible; they leave no paper trail by which they may be measured and tracked". However, the localization of knowledge externalities has been confirmed by using patent citations (Jaffe et. al., 1993) or spatially autocorrelated econometric models (Anselin et. al., 1997). Still, to our best knowledge there is a lack of studies confirming localized knowledge externalities by employing other data than patent citations.

In this paper, we study knowledge diffusion at the example of photovoltaics (PV) installations – i.e. solar cell systems to produce electric power – in Germany. We argue that – due to a very lucrative legislation, which assures a similar subsidy level all over the country – the diffusion of PV can be used as a proxy for the diffusion of knowledge about an attractive investment opportunity.

Fig. 1, in which each gray dot marks a PV system, shows that the spatial distribution of PV systems is inhomogeneous in Germany. Therefore, we study where PV systems are installed and which factors foster or hamper their diffusion.

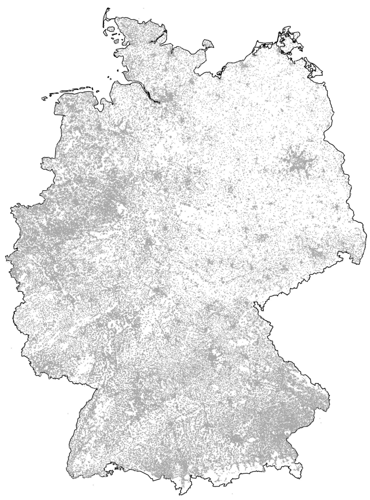

Figure 1: *Distribution of PV systems within Germany until 2009; each gray dot points a PV system*

The purpose of this paper is twofold: first, we aim at describing the heterogeneous geographical diffusion of photovoltaics installations in Germany quantitatively; second, we intend to test if preexisting photovoltaic systems stimulate further installations in their immediate vicinity, thus we investigate to which extent knowledge flows (about a profitable investment opportunity in a PV installation) depend on the geographic proximity of agents.

In order to do so, we base our analysis on a data set covering the PV installations in Germany until 2009 (some $550,000$). We geocoded the data set and employed it in an econometric model, which describes the geographical diffusion of PV installations. The model is discrete in time and space, but the level of geographical agglomeration employed is adjustable in arbitrarily small steps. We study if the knowledge and experience level regarding PV installations within a certain geographical unit, which we measure by previously installed PV systems, determine the amount of new installations. We control for – inter alia – the influence of spatial variations in global radiation, available installation space and the income level on investments in photovoltaics installations. In general, applying our modeling approach enables us to reach strongly detailed results in comparison to ordinary discrete spatial models since the usual boundary problems are avoided. Therefore, our model is applicable to study overlapping spatial units and data of very different levels of geographical agglomeration.

This paper is structured as follows. In chapter 2, we introduce the theoretical background. Thereafter, the method of the analysis, i.e. the model, the data and the level of geographical agglomeration studied are characterized in chapter 3. The following chapter 4 highlights and discusses the statistical results. Finally, chapter 5 summarizes the paper and gives an outlook for further research.

2 Theoretical Background

Although some authors argue that progress in information and communication technologies has downgraded the relevance of distance, Feldman (2002) and Redding and Venables (2004) find that physical distance is still of major importance.

2.1 Localization

Several scholars suggest that knowledge – especially the so called tacit part (Polanyi, 1967) – is localized since it builds on experience, which cannot be transferred easily as codification may not be possible (Breschi and Lissoni, 2001). In contrast, information or codifiable knowledge is thought to be transmittable over distance as it can easily be written down (Audretsch and Feldman, 1996). Similarly, Jaffe (1989) and Ponds et. al. (2007) state that face-to-face contacts are necessary to share tacit forms of knowledge. However, close interaction and daily contact are essential here. Geographical proximity only favors close interaction (Baptista, 1999), which in turn stimulates the exchange of tacit knowledge parts.

Several studies confirm this view by using patent data. Jaffe et. al. (1993) studied the state level and found a relevant effect of university research on corporate patents nearby. Eaton and Kortum (1996) examine the country level. Their analysis reveals that distance hampers the exchange of ideas; there seems to be a tendency for ideas to remain at home. Likewise, Audretsch and Feldman (1996) identify that industries where knowledge spillovers are of high importance are more likely to cluster than industries where knowledge externalities are uncommon, which are industries where research and development (R&D) efforts and highly skilled employees are less relevant. In addition, Keller (2002) approves that technological spillovers mainly happen on a local and not on a global scale since the benefits from spillovers would fade with distance.

Although it may even be possible to exchange the tacit part of knowledge over a long spatial distance, we expect that the likelihood of such an exchange increases with spatial proximity. In order to find out, we want to test if knowledge flows significantly fade with distance.

In summary, there is evidence that geographic or physical proximity is beneficial for transmitting knowledge, especially when tacit parts of knowledge should be transferred. It is just easier to share knowledge through a floor or corridor than between continents (Glaeser et. al., 1992) since distance at least enters by transportation costs (Krugman, 1991b). However, most studies confirming this theory rely upon patent data. That is why we believe there is a lack of research confirming the decreasing character of knowledge exchange with distance for different indicators of knowledge diffusion than patent citations.

We therefore employ a dataset of PV installations in Germany. But how to approach the analysis of spatial processes?

2.2 Approaches to spatial processes

Several possibilities exist when modeling spatial interactions. One option is employing usual regression models, which regard for spatial dependence by including a spatially lagged, thus autocorrelated variable. The lag arises from multiplying the variable with a spatial weights matrix (Anselin, 1988). However, this and similar approaches usually make sense for analyzing a data set where the level of geographical agglomeration has been predefined. Our data set consists of point data, which we can directly analyze or aggregate to a chosen level. This gives us the possibility to reach more detailed results.

A possibility to study point data is a spatial point pattern analysis (Bivand et. al., 2008; Ripley, 1977). However, these approaches are usually employed to find out whether a data set is homogeneously distributed in space or not, thus to draw conclusions to which extent the point data is clustered. This question is not in our focus since Fig. 1 directly shows the inhomogeneous distribution of PV systems in Germany. We therefore chose to develop our own approach, which is described in the following.

2.3 Knowledge diffusion function

Economic scholars use the concept of a production function to describe the output of a firm or even a region, which is reached by combining and processing several inputs. Traditionally, inputs as capital K, labor L and physical resources are assumed to lead to a physical output Y. Similarly, a knowledge production function can be assessed by assuming that human capital and efforts in R&D are combined to innovative output, e.g. measured by patents (Griliches, 1979; Jaffe, 1986).

In our analysis we lean on this concept and suggest that the amount of new PV installations is "produced" by the amount of previously installed PV systems, which we perceive as a proxy for localized knowledge exchange effects. Thus, we study if $\Delta PV_t\,(x, y)$ – reflecting the count of new PV installations within a certain radius around the spatial point (x, y) at time t – is positively affected by the amount of previously installed PV systems – $PV_{t-1}\,(x, y)$ – within the same geographical unit. Further, we have to control for characteristics of the local environment where – inter alia – the global radiation intensity may be influential.

As we take PV installations as a proxy for knowledge about a highly profitable investment possibility, we specify a knowledge diffusion function as follows:

$$\Delta PV_t\left(x,y\right) = f\left(PV_{t-1}\left(x,y\right), CONTROL_1\left(x,y\right), CONTROL_2\left(x,y\right)...\right). \tag{1}$$

3 Building the Model

To analyze the evolution of PV installations, we define an additive spatial aggregation function to measure the installation density for a given point (x_0, y_0). For simplicity, we chose to count all installations at a given time point t in the radius r:[3]

$$PV_t\left(x_0,y_0\right) = \#\text{ PV installations, which are within a distance of } d \leq r \text{ to } (x_0,y_0). \tag{2}$$

However, our hypothesis is that the likelihood of installing a PV system increases with the amount of previously installed PV systems. Thus, we study the annex of PV installations at time t at (x_0, y_0):

$$\Delta PV_t\left(x_0,y_0\right) = PV_t\left(x_0,y_0\right) - PV_{t-1}\left(x_0,y_0\right). \tag{3}$$

Fig. 2 illustrates schematically the relation between PV_{t-1} and ΔPV_t we want to test. In the figure gray dots are new events – PV systems – in the period specified above and black dots are old events. To clarify, we aim to explore if the probability to install a PV system increases with the density of previously installed PV systems nearby. According to our assumption that the installation of a PV system can be regarded as acquired knowledge about a profitable investment opportunity, we study to which extent knowledge flows depend on the geographic proximity of agents. Our hypothesis is as follows: The higher the density of installed PV systems at point in time $t-1$, the more PV systems are going to be installed at t since the costs of exchanging ideas tend to increase with spatial distance. Certainly, there may be situations where this is not true. As Autant-Bernard, Billand, Frachisse and Billand (2007) point out: if two persons are socially close, the exchange of ideas may be intense over long distances. However, we assume that the most interaction takes place at the local level and that tacit knowledge – including experience and trust – is mainly transmitted face-to-face, which implies interaction at the local level. Hence, we expect a spatial clustering of PV systems, which is already confirmed by Fig. 1, and a strong positive relationship between ΔPV_t and PV_{t-1}.

To explain different amounts of newly installed PV systems at different coordinates (x_0, y_0) and points in time t, we employ a pooled linear regression model:

$$\begin{aligned}\Delta PV_t(x,y) =&\beta_0 + \beta_1 PV_{t-1}(x,y) + \beta_2 GR(x,y) + \beta_3 INC_{t-1}(x,y) + \beta_4 PD_{t-1}(x,y)\\ &+ \beta_5 HOUS_{t-1}(x,y) + \beta_{6-21} DUMMY_{1994-2009} + \epsilon, \text{ with } \mathbb{E}(\epsilon) = 0, \mathbb{V}(\epsilon) > 0.\end{aligned} \tag{4}$$

GR, INC_t, PD_t, $HOUS_t$ are control variables for global solar radiation, household income, population density and the share of single and double family homes, respectively. They are all specific for a point (x_0, y_0)

[3] Incorporating the capacity (the nominal power in MW_p) of a PV system would allow us to regard for different relevance levels of PV systems, assuming that larger (more powerful) installations have a stronger impact on our aggregation function as they attract more attention. However, we suppose that differences between the installations, which can easily be about 10 to the power of 4, are not justifiable. Therefore, we included all PV installations without regarding for different capacities.

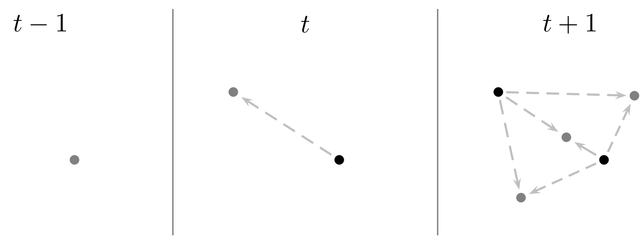

Figure 2: *Causal relationship of events (installations of PV systems) at different points in time, which should be tested; gray (black) dots are new (old) in the shown period*

and – except for global radiation – represent a specific point in time. All of these are calculated as weighted averages of the intersection of their spatial shape and the radius r around the point (x_0, y_0). This is illustrated in Fig. 3: The intersection area between the circle of the spatial aggregation function and the administrative districts (DEDnn) is taken to weight its value (e.g. INC_{t-1}).[4] We used the NUTS3 classification of 2006 where Germany was divided into 429 districts as most of the data was available for that classification.

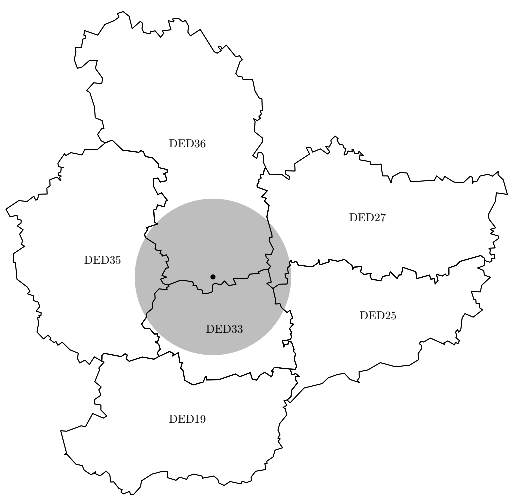

Figure 3: *Spatial aggregation of data defined for arbitrary areas*

The variables represent different factors possibly having an impact on the erection of photovoltaics installations: PV_{t-1} considers local knowledge spillover effects. GR directly influences earnings from the investment in a PV system. INC_{t-1} limits financing and risk-bearing possibilities. PD_{t-1} corrects for weaker effects of the other variables in sparsely populated areas and covers intensified knowledge exchange possibilities in densely populated areas. Finally, $HOUS_{t-1}$ is a proxy for available roofs for PV systems.[5] Regarding causality, we assume a time lag of one year. Thus, PV_{t-1}, INC_{t-1}, PD_{t-1} and $HOUS_{t-1}$

[4] In Fig. 3, DED19 stands for the rural district of Mittweida, DED25 for the rural district of Meißen, DED27 for the rural district of Riesa-Großenhain, DED33 for the rural district of Döbeln, DED35 for the rural district of Muldentalkreis and DED36 for the rural district of Torgau-Oschatz.

[5] More than 80% of the PV systems in Germany are installed on a roof since open-space systems receive a lower remuneration (Bundesverband Solarwirtschaft, 2011). We suppose that the probability to install a PV system may be higher in single and double family homes as the owner structure of these may encourage PV installations since only few parties have to agree upon the investment.

explain the new PV installations at time t. Our analysis includes the years from 1992 to 2009, allowing for 17 regression data points at each point (x, y). GR is the yearly average of 1981 until 2000 and is therefore constant over the whole period of study.

In Germany, a generous feed-in tariff makes investments in photovoltaic systems very profitable. Although the global radiation is low when compared with other countries in the south, the PV capacity installed per capita is by far the highest world wide in 2011. The feed-in tariff forces the electricity grid operators to accept the electricity produced by renewable energy to be fed into the grid and guarantees a fixed remuneration for electricity produced by PV for 20 years (Altrock et. al., 2006). Most importantly, the remuneration is financed through an apportionment by all consumers of electricity; thus the costs are born by all consumers.[6] The feed-in tariff for electricity produced by PV was changed yearly until 2009. Furthermore, PV installations became cheaper in this time period as production costs decreased due to learning effects. In order to cover these shifts in the incentive level to install a PV system, we include year dummies into our regression model. Further, the year dummies account for global effects of increasing PV installations in our panel data setting and prevent that a positive coefficient for PV_{t-1} is induced implicitly by our model specification. In other words, the time dummies enable us to study how the amount of previously installed PV systems affect new installations. They permit both the verification and the falsification of our hypothesis as positive and negative coefficients become possible for PV_{t-1}.[7]

See Table 1 for an overview of the variables including their description, our interpretation as a proxy, their source and their level of geographical agglomeration. Since the amendment of the Renewable Energy Sources Act on the 25th of October 2008, address data and the date of grid connection of all the PV installations in Germany are by law publicly available from the German transmission system operators (TSOs).[8] In order to conduct our analysis, we geocoded the data set of PV installations (some 550,000). Data upon the global radiation intensity was provided by the German Weather Service as 1-km raster data (DWD, 2010). Further data, $INC_{t-1}(x, y)$, $PD_{t-1}(x, y)$ and $HOUS_{t-1}(x, y)$, was taken from 2010's INKAR database (INKAR, 2010) and the German Statistical Office (DESTATIS, 2010a,b). This data was available at the district level.

Table 1: *Overview of the variables employed*

Variable	Description	Proxy	Source	Geogr. agglom. level
Explained variable				
$\Delta PV_t(x, y)$	New PV installations during t	–	TSOs	Geocoded point data
Explanatory variables				
$PV_{t-1}(x, y)$	Amount of PV installations until $t-1$	Knowledge and experience level regarding PV investments	TSOs	Geocoded point data
$INC_{t-1}(x, y)$	Household income at $t-1$	Economic wealth	INKAR (2010)	District level
$GR(x, y)$	Global solar radiation (average of $1981-2000$)	Earnings from PV	DWD (2010)	1-km raster data
$PD_{t-1}(x, y)$	Population density at $t-1$	Correct for sparsely populated areas	DESTATIS (2010a,b)	District level
$HOUS_{t-1}(x, y)$	Share of single and double family homes at $t-1$	Roof availability	INKAR (2010)	District level

[6] The more electricity one consumes, the more he or she has to subsidize renewable energy. However, the state is not directly involved in this subsidy, only by setting the legal framework.

[7] A positive coefficient would mean that a high PV density increases the amount of new PV installations whereas a negative coefficient describes a situation where areas with a sparse PV density catch up.

[8] The German TSOs are 50Hertz Transmission, Amprion, EnBW Transportnetze and Tennet.

To test our hypothesis, we created data sets by defining two different scanning rasters for spatial data aggregation with radius r. The scanning rasters provide points (x, y) to evaluate the respective spatial values of the regression model. The radius r defines the size of the area of study and was chosen to be sufficiently large that all points in-between several scanning points were covered. We generated:

Name	s [km]	r [km]	Resulting scanning points
Coarse raster	20	30	632
Fine raster	10	15	2974.

The $20\,km$ raster is displayed in Fig. 4. The pattern was generated by latitudinal lines each having a shortest distance of step width s to its adjacent line. On these lines, we determined points with a step width of s again. However, for the spatial density aggregation to be comparable with other regions, all points having a shortest distance less than the radius of r to the nearest border of Germany were deleted. This is also shown in Fig. 4 with $r = 30\,\text{km}$.

The spatial data was modeled, analysed and stored in a PostgreSQL database with the PostGIS extension. Interfacing to the spatial database was done via self-made C# programs, the rendering of maps with QuantumGIS and the final statistical analysis with R (R Development Core Team, 2010). All distance calculations were done on the WGS 84 ellipsoid.

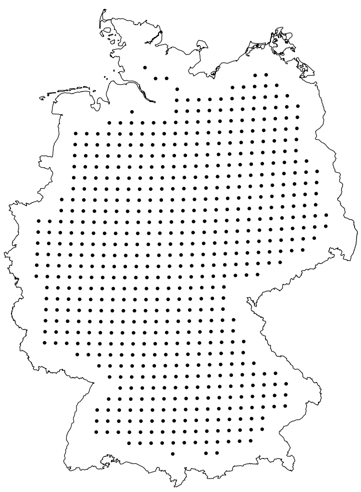

Figure 4: *Scanning raster with step width of* $20\,km$*; each black dot marks a scanning point*

4 Statistical Results and Discussion

We ran our regression model for the different raster setups, including modified versions were some variables were omitted. The descriptive statistics and the results for the $10\,km$ raster with a radius of $15\,km$ are given in Table 2 and in Table 3.

We used 4 methods to estimate our regression coefficients: ordinary least squares (OLS), maximum likelihood (ML) Poisson, quasi ML Poisson and ML negative binomial. Breusch-Pagan tests heavily suggested using heteroscedasticity- and autocorrelation-consistent standard errors for the OLS run

Table 2: *Descriptive statistics (step:* $10\,km$*, radius:* $15\,km$*)*

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Missing Values
ΔPV	0	0	8	64.92	72	1476	0
PV_{t-1}	0	2	14	180.4	155	4243	0
INC_{t-1}	634.4	1143	1259	1271	1414	2109	0
GR	940.5	988.6	1016	1030	1066	1164	0
PD_{t-1}	38.29	105.1	150	238.5	246.1	3438	0
$HOUS_{t-1}$	55.41	84.1	87.91	87.04	91.37	97.08	0

(Breusch and Pagan, 1979). Therefore, we state Newey-West standard errors, which are consistent to heteroscedasticity and autocorrelation, in the OLS regression tables (Newey and West, 1987). We also performed the rainbow test for the OLS regression, which indicates the probability of an specification error in the model (Utts, 1982). In general, the rainbow test confirms that a reasonable linear fit can be obtained for our model specification.

However, although the signs of the coefficients found do not differ in the different fitting models (except for some year dummies), we generally consider OLS as unsuitable since we model count data. Count data means: values cannot become negative and have no natural upper bound (Cameron and Trivedi, 1998). Count data is often modeled via a Poisson distribution. Under a Poisson distribution $\mathbb{E} = \mathbb{V}$, which does not hold for our data set, as our variance is higher than the expectation value. Therefore, we have an "overdispersed" data set, which leads to distorted coefficient errors under the ML Poisson model. To correct for this, we can use the quasi maximum likelihood method. Nevertheless, the coefficients are then still estimated under the Poisson assumption, which does obviously not hold, at least not in a strict sense. This problem can be resolved by choosing a negative binomial linear model, which represents an generalization of Poisson linear models. Comparing the ML Poisson and the ML negative binomial results by both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) suggest that the ML Negative Binomial fits the data better.[9]

When interpreting our regression results, the most striking observation is that all the explainable variables show – as expected – a positive coefficient for the scanning raster with step width $10\,km$. Further, the impact of all variables is found to be significant. According to our results, we approve that previously installed PV systems stimulate further installations nearby. Therefore, we conclude that knowledge flows about a highly profitable investment opportunity in PV installations depend on the geographic proximity of agents. Concerning the other variables, Table 5 in the Appendix shows that the income level has the highest explanatory power. After income level, global radiation contributes some explanation, which is consistent with our expectations. However, we find it striking that the effect of income is larger than the one of global radiation, which suggests that the feed-in tariff is high enough to make investments in PV even in less sunny parts of Germany profitable. We expected that sparsely populated areas also have less PV installations since the other factors cannot be utilized by as many individuals as in more densely populated areas. Further, we supposed that knowledge exchange would be facilitated in densely populated areas since interaction possibilities increase there. As we find a positive coefficient for PD_{t-1} we consider our hypothesis as confirmed. The "available roofs" information we intended to capture by $HOUS_{t-1}$ also has some explanatory power. We therefore see the assumption as confirmed that the

[9] Similarly, the AIC and the BIC values suggest that the full model with all the explanatory variables suites best as shown in Table 6, Table 7 and Table 8 of the Appendix.

Table 3: *Regression results (step: 10 km, radius: 15 km)*

	Model 1 OLS	Model 1 ML Poisson	Model 1 quasi ML Poisson	Model 1 ML negbin
Constant	-323.608^{***}	-11.102^{***}	-11.102^{***}	-12.163^{***}
	(9.910)	(0.027)	(0.151)	(0.143)
PV_{t-1}	0.233^{***}	0.000^{***}	0.000^{***}	0.001^{***}
	(0.005)	(0.000)	(0.000)	(0.000)
INC_{t-1}	0.013^{***}	0.001^{***}	0.001^{***}	0.003^{***}
	(0.003)	(0.000)	(0.000)	(0.000)
GR	0.155^{***}	0.004^{***}	0.004^{***}	0.004^{***}
	(0.006)	(0.000)	(0.000)	(0.000)
PD_{t-1}	0.025^{***}	0.001^{***}	0.001^{***}	0.002^{***}
	(0.002)	(0.000)	(0.000)	(0.000)
$HOUS_{t-1}$	1.675^{***}	0.059^{***}	0.059^{***}	0.045^{***}
	(0.077)	(0.000)	(0.001)	(0.001)
$Dummy_{1994}$	$-0.659^{\dagger}$	-0.263^{***}	-0.263	-0.272^{***}
	(0.377)	(0.030)	(0.168)	(0.040)
$Dummy_{1995}$	-1.109^{**}	-0.317^{***}	$-0.317^{\dagger}$	-0.585^{***}
	(0.393)	(0.030)	(0.169)	(0.042)
$Dummy_{1996}$	-0.351	0.673^{***}	0.673^{***}	0.241^{***}
	(0.431)	(0.024)	(0.135)	(0.037)
$Dummy_{1997}$	-0.151	1.041^{***}	1.041^{***}	0.505^{***}
	(0.351)	(0.023)	(0.128)	(0.036)
$Dummy_{1998}$	-0.398	1.213^{***}	1.213^{***}	0.625^{***}
	(0.405)	(0.022)	(0.125)	(0.036)
$Dummy_{1999}$	-0.581	1.424^{***}	1.424^{***}	0.885^{***}
	(0.550)	(0.022)	(0.123)	(0.035)
$Dummy_{2000}$	14.015^{***}	2.929^{***}	2.929^{***}	2.501^{***}
	(0.730)	(0.020)	(0.114)	(0.033)
$Dummy_{2001}$	35.733^{***}	3.707^{***}	3.707^{***}	3.204^{***}
	(1.004)	(0.020)	(0.113)	(0.033)
$Dummy_{2002}$	13.160^{***}	3.341^{***}	3.341^{***}	2.693^{***}
	(0.851)	(0.020)	(0.113)	(0.034)
$Dummy_{2003}$	7.561^{***}	3.386^{***}	3.386^{***}	2.691^{***}
	(0.962)	(0.020)	(0.113)	(0.034)
$Dummy_{2004}$	46.772^{***}	4.116^{***}	4.116^{***}	3.458^{***}
	(1.320)	(0.020)	(0.112)	(0.034)
$Dummy_{2005}$	62.177^{***}	4.362^{***}	4.362^{***}	3.709^{***}
	(1.801)	(0.020)	(0.112)	(0.034)
$Dummy_{2006}$	26.229^{***}	4.176^{***}	4.176^{***}	3.517^{***}
	(1.676)	(0.020)	(0.112)	(0.035)
$Dummy_{2007}$	26.001^{***}	4.251^{***}	4.251^{***}	3.577^{***}
	(2.084)	(0.020)	(0.112)	(0.035)
$Dummy_{2008}$	62.523^{***}	4.496^{***}	4.496^{***}	3.804^{***}
	(2.961)	(0.020)	(0.112)	(0.036)
$Dummy_{2009}$	40.279^{***}	4.413^{***}	4.413^{***}	3.769^{***}
	(2.456)	(0.020)	(0.113)	(0.038)
Dispersion parameter				1.169^{***}
				(0.009)
N	50558	50558	50558	50558
R^2	0.726			
adj. R^2	0.726			
Resid. sd	65.335			
Rainbow test (p-val.)	0.9784			
AIC		1687978.771		364094.527
BIC		1688755.888		364906.967
$\log L$		-843901.385		-181955.263

Standard errors in parentheses (Newey-West for OLS)
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

ownership structure implies energy efficiency considerations, which foster investment decisions towards PV installations. The year dummies were included to cover strong shifts in the feed-in tariff and were mainly found to be significant.

A test on variance inflation suggests not to discard any variable: Table 4 exemplarily shows the variance inflation factor (VIF) for our fine scanning raster and indicates that multicollinearity should be no problem in our analysis since all the values are close to 3 or well below (Marquardt, 1970; O'brien, 2007). Similarly low values for VIF were found for all the model specifications tested.

Table 4: *VIF (step: 10 km, radius: 15 km)*

	Model 1
PV_{t-1}	2.40
INC_{t-1}	3.03
GR	1.20
PD_{t-1}	2.50
$HOUS_{t-1}$	2.11
$Dummy_{1994}$	1.89
$Dummy_{1995}$	1.89
$Dummy_{1996}$	1.91
$Dummy_{1997}$	1.93
$Dummy_{1998}$	1.95
$Dummy_{1999}$	1.97
$Dummy_{2000}$	2.01
$Dummy_{2001}$	2.06
$Dummy_{2002}$	2.14
$Dummy_{2003}$	2.15
$Dummy_{2004}$	2.21
$Dummy_{2005}$	2.26
$Dummy_{2006}$	2.36
$Dummy_{2007}$	2.49
$Dummy_{2008}$	2.66
$Dummy_{2009}$	2.95

Additionally, we tested other proxies for the economic wealth, the population density and the newly available roof space of a given spatial unit. However, they were less influential as the factors shown above.[10]

The regression runs under the raster with step width $20\,km$ are not shown, but were roughly the same and found a lower influence of PV_{t-1}. Our interpretation is that a scanning raster of $20\,km$ and a radius of $30\,km$ are to coarse: according to our results, knowledge exchange seems to be localized as suggested by the correctly specified regression models with a fine raster of $10\,km$ and radius $15\,km$.

To sum up, the analysis of PV data suggests that spatial proximity facilitates knowledge exchange. In line with the theoretic background introduced in chapter 2, we argue that interactions that favor the exchange of tacit knowledge thanks to face-to-face contact stimulate trust between people and a fast diffusion of knowledge. We interpret that proximity is – according to our findings – vital since investment opportunities are only shared, diffused and finally made when trust in the investment has been built. Indeed, knowledge about investments in PV systems seems to be localized – and therefore tacit – to some degree. Observing a PV system at work and talking about it with a person of trust, seems to increase the likelihood of installing PV.

5 Summary and Outlook

We set out to study the spatial diffusion of photovoltaics installations in Germany with an econometric model and to test if pre-existing photovoltaic systems stimulate further installations nearby, which we took as a proxy for knowledge diffusion about a highly profitable investment opportunity. Our model accounts for spatial variations in global radiation, available installation space, the income level and the population density of the spatial units studied. Further, we analyze if the knowledge and experience

[10] The proxies tested were the income tax payments, the GDP per capita and GDP per employee instead of household income. Further, we replaced the population density by a proxy for rural and urban areas. Finally, the single and double family homes were substituted by a measure of new built flats per capita, the share of flats in single and double family homes and the amount of multi-family home.

level regarding PV installations within a certain geographical unit, which we assume to be measured by previously installed PV systems, determine the amount of new PV installations.

Our study is based on a dataset of some $550,000$ PV systems installed in Germany until 2009. We geocoded the dataset and developed an econometric model to describe the heterogeneous geographical diffusion of PV installations. The model is discrete in time. The level of geographical agglomeration is discrete, but adjustable in arbitrarily small steps. Our model may be used to study the diffusion of geographical point processes in general. Applying our modeling approach allows us to reach more detailed results in comparison to ordinary discrete spatial models since the usual boundary problems are evaded. Besides, our model can be used to study overlapping spatial units and very different levels of geographical agglomeration.

We find that the installation process of PV systems is highly dependent on previously installed PV systems nearby. Thus, the likelihood of installing a PV system increases with the amount of installations that can be found in the neighborhood. In conclusion, our results are in line with the theory that knowledge exchange decreases with distance.

Further, our analysis shows a positive relation between economic wealth and global radiation and the amount of PV systems installed within a spatial unit. The algebraic sign of the relation between available roof space – measured by single and double family homes – and the population density is also found to be positive.

Our analysis leaves room for improvements: A lower level of geographical agglomeration for the proxies of the economic wealth, the population density and the share of single and double family homes should be beneficial. Using a spatial aggregation function, which values spatial point process nearby higher than those far away should also be a promising step.

References

Altrock, M., Oschmann, V. and Theobald, C. (2006), EEG: Erneuerbare-Energien-Gesetz; Kommentar, Beck, Munich.

Anselin, L. (1988), Spatial Econometrics: Methods and models, Kluwer Academic Publishers, Boston.

Anselin, L., Varga, A. and Acs, Z. (1997), 'Local geographic spillovers between university research and high technology innovations', Journal of Urban Economics **42**, 422–448.

Audretsch, D. B. and Feldman, M. P. (1996), 'R&D spillovers and the geography of innovation and production', The American Economic Review **86**(3), 630–640.

Autant-Bernard, C., Billand, P., Frachisse, D. and Massard, N. (2007), 'Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies', Papers in Regional Science **86**(3), 495–519.

Autant-Bernard, C., Mairesse, J. and Massard, N. (2007), 'Spatial knowledge diffusion through collaborative networks', Papers in Regional Science **86**(3), 341–350.

Baptista, R. (1999), 'The diffusion of process innovations: A selective review', International Journal of the Economics of Business **6**(1), 107–129.

Bivand, R. S., Pebesma, E. J. and Gómez-Rubio, V. (Eds.) (2008), Applied Spatial Data Analysis with R, Springer, New York.

Breschi, S. and Lissoni, F. (2001), 'Knowledge Spillovers and Local Innovation Systems: A Critical Survey', Industrial and Corporate Change **10**(4), 975–1005.

Breusch, T. S. and Pagan, A. R. (1979), 'A simple test for heteroscedasticity and random coefficient variation', Econometrica **47**(5), 1287–1294.

Bundesverband Solarwirtschaft (2011), 'Solarbranche in Deutschland - statistisches Kurzprofil Ende 2010', website. Acc. Jan. 27, 2011.
URL: *http://www.solarwirtschaft.de/medienvertreter/marktdaten.html*

Cameron, A. C. and Trivedi, P. K. (1998), Regression Analysis of Count Data, Cambridge Univ.Press, New York.

DESTATIS (2010a), 'Regional statistics (Regionalstatistik), Subject areas (Sachgebiete): 11111 Festsetzung des Gebietsstandes, Table (Tabelle): 171-01-4 Gebietsstand: Gebietsfläche in qkm - Stichtag 31.12. - regionale Tiefe: Kreise und krfr. Städte', Database, Statistische Ämter des Bundes und der Länder. Acc. Aug. 26, 2010.
URL: *http://www.regionalstatistik.de/*

DESTATIS (2010b), 'Regional statistics (Regionalstatistik), Subject areas (Sachgebiete): 12411 Fortschreibung des Bevölkerungsstandes, Table (Tabelle): 173-21-4 Bevölkerungsstand: Bevölkerung nach Geschlecht und Altersgruppe - Stichtag 31.12. - regionale Tiefe: Kreise und krfr. Städte', Database, Statistische Ämter des Bundes und der Länder. Acc. Aug. 26, 2010.
URL: *http://www.regionalstatistik.de/*

DWD (2010), '1-km-Rasterdaten der mittleren jährlichen Globalstrahlung (kWh/m^2, Zeitraum 1981-2000) der gesamten BRD', Data was provided after personal request, Deutscher Wetterdienst. Acc. July 26, 2010.

Eaton, J. and Kortum, S. (1996), 'Trade in ideas patenting and productivity in the OECD', Journal of International Economics **40**(3-4), 251–278. Symposium on Growth and International Trade: Empirical Studies.

Feldman, M. P. (2002), 'The internet revolution and the geography of innovation', International Social Science Journal **54**(171), 47–56.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and Shleifer, A. (1992), 'Growth in cities', The Journal of Political Economy **100**(6), 1126–1152.

Griliches, Z. (1979), 'Issues in assessing the contribution of research and development to productivity growth', Bell Journal of Economics **10**(1), 92–116.

INKAR (2010), 'Indikatoren und Karten zur Raumentwicklung in Deutschland und Europa (CD-ROM)', Database, Bundesamt für Bauwesen und Raumordnung, Bonn.

Jaffe, A. B. (1986), 'Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value', American Economic Review **76**(5), 984–1001.

Jaffe, A. B. (1989), 'Real effects of academic research', American Economic Review **79**(5), 957–970.

Jaffe, A. B., Trajtenberg, M. and Henderson, R. (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citations', The Quarterly Journal of Economics **108**(3), 577–598.

Keller, W. (2002), 'Geographic localization of international technology diffusion', The American Economic Review **92**(1), 120–142.

Krugman, P. (1991a), Geography and Trade, MIT Press, Cambridge, MA.

Krugman, P. (1991b), 'Increasing returns and economic geography', The Journal of Political Economy **99**(3), 483–499.

Marquardt, D. W. (1970), 'Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation', Technometrics **12**(3), 591–612.

Marshall, A. (1890), Principles of Economics, Macmillan, London.

Newey, W. K. and West, K. D. (1987), 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', Econometrica **55**(3), 703–708.

O'brien, R. (2007), 'A caution regarding rules of thumb for variance inflation factors', Quality & Quantity **41**, 673–690.

Polanyi, M. (1967), The Tacit Dimension, Anchor Books, New York.

Ponds, R., van Oort, F. and Frenken, K. (2007), 'The geographical and institutional proximity of research collaboration', Papers in Regional Science **86**(3), 423–443.

R Development Core Team (2010), R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org/*

Redding, S. and Venables, A. J. (2004), 'Economic geography and international inequality', Journal of International Economics **62**(1), 53–82.

Ripley, B. D. (1977), 'Modelling spatial patterns', Journal of the Royal Statistical Society. Series B (Methodological) **39**(2), 172–212.

Utts, J. (1982), 'The rainbow test for lack of fit in regression', Communications in Statistics - Theory and Methods **11**, 1801–1815.

Appendix

Table 5: *Regression results OLS (step: 10 km, radius: 15 km)*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Constant	−323.608***	−867.997***	0.512***	−232.870***	−643.879***	−13.527***	−112.868***	−170.544***	−44.770***	0.353	−100.265***
	(9.910)	(16.896)	(0.023)	(3.888)	(12.584)	(0.676)	(10.138)	(6.833)	(2.950)	(0.331)	(4.590)
PV_{t-1}	0.233***		0.252***					0.241***	0.242***	0.252***	0.252***
	(0.005)		(0.004)					(0.004)	(0.005)	(0.004)	(0.004)
INC_{t-1}	0.013***	0.126***		0.222***					0.043***		
	(0.003)	(0.004)		(0.004)					(0.003)		
GR	0.155***	0.518***			0.626***			0.166***			
	(0.006)	(0.011)			(0.012)			(0.007)			
PD_{t-1}	0.025***	0.064***				0.060***				0.001	
	(0.002)	(0.003)				(0.003)				(0.001)	
$HOUS_{t-1}$	1.675***	2.147***					1.308***				1.159***
	(0.077)	(0.125)					(0.116)				(0.053)
$Dummy_{1994}$	−0.659†	−3.548**	−0.405***	−6.266***	−0.184	−0.199	−0.119	−0.395†	−1.574***	−0.405***	−0.347†
	(0.377)	(1.243)	(0.032)	(1.068)	(0.872)	(0.432)	(0.232)	(0.234)	(0.221)	(0.032)	(0.199)
$Dummy_{1995}$	−1.109**	−6.932***	−0.598***	−12.366***	−0.203	−0.233	−0.074	−0.580*	−2.939***	−0.598***	−0.483*
	(0.393)	(1.233)	(0.040)	(1.037)	(0.875)	(0.426)	(0.234)	(0.237)	(0.242)	(0.039)	(0.199)
$Dummy_{1996}$	−0.351	−9.110***	0.419***	−17.262***	0.983	0.938*	1.177***	0.444†	−3.094***	0.419***	0.591**
	(0.431)	(1.229)	(0.100)	(1.014)	(0.882)	(0.413)	(0.262)	(0.254)	(0.289)	(0.099)	(0.226)
$Dummy_{1997}$	−0.151	−11.572***	0.853***	−22.441***	1.886***	1.826***	2.144***	0.898***	−3.820***	0.853***	1.083***
	(0.351)	(0.520)	(0.139)	(0.524)	(0.165)	(0.166)	(0.166)	(0.140)	(0.327)	(0.139)	(0.141)
$Dummy_{1998}$	−0.398	−13.393***	0.764***	−26.229***	2.493***	2.464***	2.773***	0.840***	−4.734***	0.765***	1.013***
	(0.405)	(0.605)	(0.138)	(0.599)	(0.193)	(0.195)	(0.195)	(0.150)	(0.372)	(0.137)	(0.149)
$Dummy_{1999}$	−0.581	−15.173***	0.854***	−29.922***	3.433***	3.414***	3.652***	0.967***	−5.508***	0.855***	1.051***
	(0.550)	(1.305)	(0.124)	(1.089)	(0.897)	(0.397)	(0.322)	(0.258)	(0.427)	(0.123)	(0.230)
$Dummy_{2000}$	14.015***	−3.393*	15.855***	−21.363***	19.522***	19.585***	19.667***	16.016***	8.077***	15.858***	15.989***
	(0.730)	(1.337)	(0.513)	(1.182)	(0.925)	(0.607)	(0.608)	(0.508)	(0.666)	(0.513)	(0.538)
$Dummy_{2001}$	35.733***	19.154***	37.737***	−1.924	46.550***	46.600***	46.601***	38.124***	28.691***	37.743***	37.792***
	(1.004)	(1.398)	(0.967)	(1.299)	(1.061)	(1.062)	(1.110)	(0.893)	(0.960)	(0.968)	(0.974)
$Dummy_{2002}$	13.160***	2.073	14.979***	−23.477***	35.759***	35.777***	35.722***	15.891***	4.320***	14.992***	14.970***
	(0.851)	(1.490)	(0.672)	(1.375)	(0.864)	(0.834)	(0.827)	(0.646)	(0.713)	(0.680)	(0.676)
$Dummy_{2003}$	7.561***	4.149**	8.880***	−21.918***	38.903***	38.906***	38.775***	10.198***	−1.722*	8.898***	8.801***
	(0.962)	(1.513)	(0.808)	(1.352)	(0.953)	(0.974)	(0.956)	(0.826)	(0.765)	(0.825)	(0.814)
$Dummy_{2004}$	46.772***	49.391***	47.782***	20.730***	87.843***	87.869***	87.624***	49.542***	36.357***	47.807***	47.635***
	(1.320)	(1.954)	(1.266)	(1.866)	(1.481)	(1.680)	(1.665)	(1.231)	(1.174)	(1.279)	(1.242)
$Dummy_{2005}$	62.177***	82.786***	61.905***	52.097***	124.353***	124.384***	124.013***	64.648***	50.367***	61.944***	61.676***
	(1.801)	(2.487)	(1.843)	(2.450)	(2.252)	(2.501)	(2.492)	(1.859)	(1.607)	(1.872)	(1.818)
$Dummy_{2006}$	26.229***	72.806***	23.987***	39.587***	118.035***	118.076***	117.599***	28.117***	12.496***	24.045***	23.709***
	(1.676)	(2.193)	(1.756)	(1.989)	(1.927)	(1.940)	(1.992)	(1.819)	(1.341)	(1.822)	(1.741)
$Dummy_{2007}$	26.001***	96.438***	21.969***	60.119***	146.022***	146.084***	145.499***	27.417***	10.217***	22.046***	21.649***
	(2.084)	(2.541)	(2.189)	(2.336)	(2.269)	(2.318)	(2.366)	(2.266)	(1.722)	(2.275)	(2.173)
$Dummy_{2008}$	62.523***	163.877***	56.080***	124.895***	217.201***	217.281***	216.604***	63.156***	44.551***	56.180***	55.738***
	(2.961)	(3.542)	(2.950)	(3.418)	(3.190)	(3.456)	(3.492)	(3.063)	(2.532)	(3.043)	(2.941)
$Dummy_{2009}$	40.279***	189.144***	29.995***	147.436***	246.145***	246.254***	245.504***	39.487***	19.396***	30.130***	29.677***
	(2.456)	(4.949)	(2.353)	(4.682)	(4.515)	(4.622)	(4.580)	(2.486)	(2.354)	(2.332)	(2.328)
N	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558
R^2	0.726	0.516	0.717	0.465	0.456	0.403	0.388	0.721	0.719	0.717	0.720
adj. R^2	0.726	0.516	0.717	0.465	0.456	0.403	0.388	0.721	0.719	0.717	0.720
Resid. sd	65.335	86.876	66.425	91.314	92.084	96.452	97.656	65.907	66.130	66.425	66.074
Rainb. (p-val)	0.9784	0.8835	0.9737	0.854	0.7357	0.6876	0.6991	0.9737	0.9774	0.974	0.9745

Newey-West standard errors in parentheses

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 6: *Regression results ML poisson (step: 10 km, radius: 15 km)*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Constant	−11.102***	−14.043***	−0.136***	−3.576***	−9.157***	−0.304***	−2.251***	−4.629***	−2.259***	−0.192***	−2.178***
	(0.027)	(0.025)	(0.020)	(0.020)	(0.022)	(0.020)	(0.022)	(0.023)	(0.020)	(0.020)	(0.022)
PV_{t-1}	0.000***		0.001***					0.001***	0.001***	0.001***	0.001***
	(0.000)		(0.000)					(0.000)	(0.000)	(0.000)	(0.000)
INC_{t-1}	0.001***	0.002***		0.003***					0.002***		
	(0.000)	(0.000)		(0.000)					(0.000)		
GR	0.004***	0.006***			0.009***			0.004***			
	(0.000)	(0.000)			(0.000)			(0.000)			
PD_{t-1}	0.001***	0.001***				0.001***				0.000***	
	(0.000)	(0.000)				(0.000)				(0.000)	
$HOUS_{t-1}$	0.059***	0.064***					0.024***				0.023***
	(0.000)	(0.000)					(0.000)				(0.000)
$Dummy_{1994}$	−0.263***	−0.268***	−0.237***	−0.302***	−0.236***	−0.236***	−0.235***	−0.237***	−0.282***	−0.237***	−0.236***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
$Dummy_{1995}$	−0.317***	−0.329***	−0.265***	−0.398***	−0.264***	−0.264***	−0.261***	−0.265***	−0.356***	−0.266***	−0.263***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
$Dummy_{1996}$	0.673***	0.655***	0.751***	0.550***	0.754***	0.754***	0.758***	0.752***	0.614***	0.751***	0.756***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
$Dummy_{1997}$	1.041***	1.017***	1.146***	0.874***	1.150***	1.150***	1.155***	1.147***	0.960***	1.145***	1.152***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
$Dummy_{1998}$	1.213***	1.183***	1.342***	1.009***	1.349***	1.349***	1.355***	1.344***	1.117***	1.341***	1.348***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
$Dummy_{1999}$	1.424***	1.386***	1.584***	1.182***	1.595***	1.596***	1.600***	1.587***	1.316***	1.584***	1.591***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
$Dummy_{2000}$	2.929***	2.881***	3.135***	2.637***	3.150***	3.152***	3.154***	3.139***	2.803***	3.134***	3.140***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2001}$	3.707***	3.658***	3.958***	3.370***	3.994***	3.996***	3.996***	3.965***	3.561***	3.957***	3.962***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2002}$	3.341***	3.301***	3.651***	2.950***	3.735***	3.737***	3.737***	3.664***	3.162***	3.651***	3.653***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2003}$	3.386***	3.369***	3.694***	3.014***	3.818***	3.819***	3.817***	3.711***	3.198***	3.695***	3.694***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2004}$	4.116***	4.117***	4.451***	3.726***	4.620***	4.622***	4.618***	4.471***	3.907***	4.453***	4.448***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2005}$	4.362***	4.413***	4.692***	3.995***	4.965***	4.967***	4.961***	4.726***	4.123***	4.697***	4.682***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2006}$	4.176***	4.301***	4.482***	3.843***	4.913***	4.915***	4.907***	4.544***	3.888***	4.494***	4.466***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2007}$	4.251***	4.441***	4.534***	3.943***	5.124***	5.126***	5.117***	4.630***	3.907***	4.550***	4.515***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2008}$	4.496***	4.779***	4.714***	4.242***	5.519***	5.521***	5.511***	4.858***	4.084***	4.737***	4.691***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$Dummy_{2009}$	4.413***	4.848***	4.481***	4.271***	5.644***	5.646***	5.635***	4.713***	3.896***	4.519***	4.451***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
N	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558
AIC	1687978.771	1806828.255	2188911.928	2373117.190	2469320.408	3098549.787	3217298.383	2044085.699	1960883.577	2170346.980	2135220.044
BIC	1688755.888	1807570.049	2189547.751	2373753.013	2469956.231	3099185.611	3217934.206	2044756.846	1961554.724	2171018.127	2135891.190
$\log L$	−843901.385	−903330.128	−1094383.964	−1186486.595	−1234588.204	−1549202.894	−1608577.191	−1021966.850	−980365.789	−1085097.490	−1067534.022

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 7: *Regression results quasi ML poisson (step: 10 km, radius: 15 km)*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Constant	−11.102***	−14.043***	−0.136	−3.576***	−9.157***	−0.304†	−2.251***	−4.629***	−2.259***	−0.192	−2.178***
	(0.151)	(0.152)	(0.128)	(0.141)	(0.165)	(0.160)	(0.186)	(0.146)	(0.123)	(0.127)	(0.143)
PV_{t-1}	0.000***		0.001***					0.001***	0.001***	0.001***	0.001***
	(0.000)		(0.000)					(0.000)	(0.000)	(0.000)	(0.000)
INC_{t-1}	0.001***	0.002***		0.003***					0.002***		
	(0.000)	(0.000)		(0.000)					(0.000)		
GR	0.004***	0.006***			0.009***			0.004***			
	(0.000)	(0.000)			(0.000)			(0.000)			
PD_{t-1}	0.001***	0.001***				0.001***				0.000***	
	(0.000)	(0.000)				(0.000)				(0.000)	
$HOUS_{t-1}$	0.059***	0.064***					0.024***				0.023***
	(0.001)	(0.001)					(0.001)				(0.001)
$Dummy_{1994}$	−0.263	−0.268	−0.237	−0.302	−0.236	−0.236	−0.235	−0.237	−0.282	−0.237	−0.236
	(0.168)	(0.179)	(0.193)	(0.208)	(0.221)	(0.241)	(0.252)	(0.188)	(0.180)	(0.190)	(0.195)
$Dummy_{1995}$	−0.317†	−0.329†	−0.265	−0.398†	−0.264	−0.264	−0.261	−0.265	−0.356*	−0.266	−0.263
	(0.169)	(0.180)	(0.194)	(0.210)	(0.223)	(0.243)	(0.254)	(0.189)	(0.182)	(0.192)	(0.196)
$Dummy_{1996}$	0.673***	0.655***	0.751***	0.550**	0.754***	0.754***	0.758***	0.752***	0.614***	0.751***	0.756***
	(0.135)	(0.144)	(0.155)	(0.168)	(0.178)	(0.194)	(0.203)	(0.151)	(0.145)	(0.153)	(0.157)
$Dummy_{1997}$	1.041***	1.017***	1.146***	0.874***	1.150***	1.150***	1.155***	1.147***	0.960***	1.145***	1.152***
	(0.128)	(0.136)	(0.147)	(0.159)	(0.168)	(0.184)	(0.192)	(0.143)	(0.137)	(0.145)	(0.149)
$Dummy_{1998}$	1.213***	1.183***	1.342***	1.009***	1.349***	1.349***	1.355***	1.344***	1.117***	1.341***	1.348***
	(0.125)	(0.133)	(0.144)	(0.155)	(0.165)	(0.180)	(0.188)	(0.140)	(0.134)	(0.142)	(0.145)
$Dummy_{1999}$	1.424***	1.386***	1.584***	1.182***	1.595***	1.596***	1.600***	1.587***	1.316***	1.584***	1.591***
	(0.123)	(0.130)	(0.140)	(0.152)	(0.161)	(0.175)	(0.184)	(0.137)	(0.131)	(0.139)	(0.142)
$Dummy_{2000}$	2.929***	2.881***	3.135***	2.637***	3.150***	3.152***	3.154***	3.139***	2.803***	3.134***	3.140***
	(0.114)	(0.121)	(0.131)	(0.141)	(0.150)	(0.163)	(0.171)	(0.127)	(0.122)	(0.129)	(0.132)
$Dummy_{2001}$	3.707***	3.658***	3.958***	3.370***	3.994***	3.996***	3.996***	3.965***	3.561***	3.957***	3.962***
	(0.113)	(0.120)	(0.129)	(0.140)	(0.148)	(0.161)	(0.169)	(0.126)	(0.121)	(0.128)	(0.131)
$Dummy_{2002}$	3.341***	3.301***	3.651***	2.950***	3.735***	3.737***	3.737***	3.664***	3.162***	3.651***	3.653***
	(0.113)	(0.120)	(0.130)	(0.140)	(0.148)	(0.162)	(0.170)	(0.126)	(0.121)	(0.128)	(0.131)
$Dummy_{2003}$	3.386***	3.369***	3.694***	3.014***	3.818***	3.819***	3.817***	3.711***	3.198***	3.695***	3.694***
	(0.113)	(0.120)	(0.129)	(0.140)	(0.148)	(0.162)	(0.169)	(0.126)	(0.121)	(0.128)	(0.131)
$Dummy_{2004}$	4.116***	4.117***	4.451***	3.726***	4.620***	4.622***	4.618***	4.471***	3.907***	4.453***	4.448***
	(0.112)	(0.120)	(0.129)	(0.139)	(0.147)	(0.161)	(0.168)	(0.125)	(0.121)	(0.127)	(0.130)
$Dummy_{2005}$	4.362***	4.413***	4.692***	3.995***	4.965***	4.967***	4.961***	4.726***	4.123***	4.697***	4.682***
	(0.112)	(0.119)	(0.128)	(0.139)	(0.147)	(0.161)	(0.168)	(0.125)	(0.120)	(0.127)	(0.130)
$Dummy_{2006}$	4.176***	4.301***	4.482***	3.843***	4.913***	4.915***	4.907***	4.544***	3.888***	4.494***	4.466***
	(0.112)	(0.120)	(0.129)	(0.139)	(0.147)	(0.161)	(0.168)	(0.125)	(0.120)	(0.127)	(0.130)
$Dummy_{2007}$	4.251***	4.441***	4.534***	3.943***	5.124***	5.126***	5.117***	4.630***	3.907***	4.550***	4.515***
	(0.112)	(0.120)	(0.128)	(0.139)	(0.147)	(0.160)	(0.168)	(0.125)	(0.120)	(0.127)	(0.130)
$Dummy_{2008}$	4.496***	4.779***	4.714***	4.242***	5.519***	5.521***	5.511***	4.858***	4.084***	4.737***	4.691***
	(0.112)	(0.120)	(0.128)	(0.139)	(0.147)	(0.160)	(0.168)	(0.125)	(0.120)	(0.127)	(0.130)
$Dummy_{2009}$	4.413***	4.848***	4.481***	4.271***	5.644***	5.646***	5.635***	4.713***	3.896***	4.519***	4.451***
	(0.113)	(0.120)	(0.129)	(0.139)	(0.147)	(0.160)	(0.168)	(0.125)	(0.120)	(0.127)	(0.130)
N	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 8: *Regression results ML negative binomial (step: 10 km, radius: 15 km)*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Constant	-12.163^{***}	-14.234^{***}	-0.146^{***}	-6.186^{***}	-4.945^{***}	-0.861^{***}	2.434^{***}	-1.952^{***}	-5.039^{***}	-0.752^{***}	1.941^{***}
	(0.143)	(0.133)	(0.029)	(0.048)	(0.119)	(0.032)	(0.096)	(0.119)	(0.050)	(0.030)	(0.087)
PV_{t-1}	0.001^{***}		0.002^{***}					0.002^{***}	0.001^{***}	0.002^{***}	0.002^{***}
	(0.000)		(0.000)					(0.000)	(0.000)	(0.000)	(0.000)
INC_{t-1}	0.003^{***}	0.004^{***}		0.006^{***}					0.004^{***}		
	(0.000)	(0.000)		(0.000)					(0.000)		
GR	0.004^{***}	0.005^{***}			0.005^{***}			0.002^{***}			
	(0.000)	(0.000)			(0.000)			(0.000)			
PD_{t-1}	0.002^{***}	0.002^{***}				0.002^{***}				0.002^{***}	
	(0.000)	(0.000)				(0.000)				(0.000)	
$HOUS_{t-1}$	0.045^{***}	0.050^{***}					-0.030^{***}				-0.024^{***}
	(0.001)	(0.001)					(0.001)				(0.001)
$Dummy_{1994}$	-0.272^{***}	-0.281^{***}	-0.238^{***}	-0.299^{***}	-0.260^{***}	-0.254^{***}	-0.274^{***}	-0.246^{***}	-0.288^{***}	-0.253^{***}	-0.265^{***}
	(0.040)	(0.040)	(0.042)	(0.041)	(0.044)	(0.046)	(0.045)	(0.042)	(0.040)	(0.043)	(0.043)
$Dummy_{1995}$	-0.585^{***}	-0.624^{***}	-0.272^{***}	-0.694^{***}	-0.247^{***}	-0.446^{***}	-0.322^{***}	-0.265^{***}	-0.606^{***}	-0.418^{***}	-0.314^{***}
	(0.042)	(0.042)	(0.042)	(0.043)	(0.044)	(0.047)	(0.046)	(0.042)	(0.042)	(0.044)	(0.043)
$Dummy_{1996}$	0.241^{***}	0.197^{***}	0.739^{***}	-0.021	0.765^{***}	0.541^{***}	0.695^{***}	0.740^{***}	0.110^{**}	0.569^{***}	0.693^{***}
	(0.037)	(0.038)	(0.038)	(0.039)	(0.040)	(0.042)	(0.042)	(0.038)	(0.037)	(0.039)	(0.039)
$Dummy_{1997}$	0.505^{***}	0.461^{***}	1.095^{***}	0.208^{***}	1.197^{***}	0.883^{***}	1.079^{***}	1.110^{***}	0.347^{***}	0.892^{***}	1.041^{***}
	(0.036)	(0.036)	(0.038)	(0.037)	(0.039)	(0.041)	(0.041)	(0.037)	(0.036)	(0.038)	(0.038)
$Dummy_{1998}$	0.625^{***}	0.585^{***}	1.247^{***}	0.336^{***}	1.373^{***}	1.061^{***}	1.262^{***}	1.255^{***}	0.464^{***}	1.043^{***}	1.184^{***}
	(0.036)	(0.036)	(0.037)	(0.037)	(0.039)	(0.040)	(0.041)	(0.037)	(0.036)	(0.038)	(0.038)
$Dummy_{1999}$	0.885^{***}	0.843^{***}	1.470^{***}	0.582^{***}	1.588^{***}	1.423^{***}	1.551^{***}	1.467^{***}	0.701^{***}	1.378^{***}	1.446^{***}
	(0.035)	(0.035)	(0.037)	(0.036)	(0.039)	(0.040)	(0.040)	(0.037)	(0.035)	(0.037)	(0.037)
$Dummy_{2000}$	2.501^{***}	2.440^{***}	3.084^{***}	2.116^{***}	3.017^{***}	3.277^{***}	3.212^{***}	3.027^{***}	2.255^{***}	3.218^{***}	3.141^{***}
	(0.033)	(0.034)	(0.036)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.034)	(0.036)	(0.036)
$Dummy_{2001}$	3.204^{***}	3.139^{***}	3.870^{***}	2.708^{***}	3.819^{***}	4.149^{***}	4.057^{***}	3.804^{***}	2.864^{***}	4.040^{***}	3.926^{***}
	(0.033)	(0.034)	(0.036)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.034)	(0.036)	(0.036)
$Dummy_{2002}$	2.693^{***}	2.636^{***}	3.409^{***}	2.134^{***}	3.519^{***}	3.923^{***}	3.812^{***}	3.346^{***}	2.267^{***}	3.651^{***}	3.481^{***}
	(0.034)	(0.034)	(0.036)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.034)	(0.036)	(0.036)
$Dummy_{2003}$	2.691^{***}	2.665^{***}	3.320^{***}	2.167^{***}	3.569^{***}	4.038^{***}	3.913^{***}	3.263^{***}	2.236^{***}	3.618^{***}	3.408^{***}
	(0.034)	(0.034)	(0.036)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.034)	(0.036)	(0.036)
$Dummy_{2004}$	3.458^{***}	3.463^{***}	4.012^{***}	2.886^{***}	4.431^{***}	4.851^{***}	4.722^{***}	3.991^{***}	2.921^{***}	4.334^{***}	4.105^{***}
	(0.034)	(0.034)	(0.036)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.034)	(0.036)	(0.036)
$Dummy_{2005}$	3.709^{***}	3.776^{***}	4.113^{***}	3.132^{***}	4.799^{***}	5.185^{***}	5.065^{***}	4.123^{***}	3.076^{***}	4.466^{***}	4.204^{***}
	(0.034)	(0.034)	(0.036)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.034)	(0.036)	(0.036)
$Dummy_{2006}$	3.517^{***}	3.654^{***}	3.781^{***}	2.943^{***}	4.783^{***}	5.086^{***}	4.995^{***}	3.825^{***}	2.790^{***}	4.147^{***}	3.859^{***}
	(0.035)	(0.035)	(0.037)	(0.035)	(0.038)	(0.038)	(0.039)	(0.036)	(0.035)	(0.036)	(0.037)
$Dummy_{2007}$	3.577^{***}	3.768^{***}	3.719^{***}	3.002^{***}	4.987^{***}	5.299^{***}	5.205^{***}	3.776^{***}	2.768^{***}	4.125^{***}	3.786^{***}
	(0.035)	(0.035)	(0.037)	(0.036)	(0.038)	(0.038)	(0.039)	(0.037)	(0.035)	(0.037)	(0.037)
$Dummy_{2008}$	3.804^{***}	4.067^{***}	3.821^{***}	3.261^{***}	5.372^{***}	5.712^{***}	5.608^{***}	3.890^{***}	2.915^{***}	4.278^{***}	3.886^{***}
	(0.036)	(0.036)	(0.038)	(0.036)	(0.038)	(0.038)	(0.039)	(0.038)	(0.036)	(0.037)	(0.038)
$Dummy_{2009}$	3.769^{***}	4.145^{***}	3.664^{***}	3.288^{***}	5.537^{***}	5.906^{***}	5.770^{***}	3.764^{***}	2.766^{***}	4.199^{***}	3.758^{***}
	(0.038)	(0.036)	(0.039)	(0.036)	(0.038)	(0.038)	(0.039)	(0.039)	(0.037)	(0.039)	(0.039)
Dispersion parameter	1.169^{***}	1.141^{***}	0.746^{***}	0.922^{***}	0.631^{***}	0.707^{***}	0.598^{***}	0.756^{***}	0.994^{***}	0.860^{***}	0.750^{***}
	(0.009)	(0.009)	(0.006)	(0.007)	(0.005)	(0.005)	(0.004)	(0.006)	(0.008)	(0.007)	(0.006)
N	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558	50558
AIC	364094.527	365114.310	381955.720	373279.089	388905.059	381831.296	389966.587	381666.457	370535.890	374592.132	381179.223
BIC	364906.967	365891.427	382626.867	373950.236	389576.206	382502.443	390637.734	382372.927	371242.360	375298.602	381885.693
$\log L$	−181955.263	−182469.155	−190901.860	−186563.545	−194376.530	−190839.648	−194907.293	−190753.229	−185187.945	−187216.066	−190509.612

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$