

Rode, Johannes; Müller, Sven

**Conference Paper**

## Spatio-temporal variation in peer effects - The case of rooftop photovoltaic systems in Germany

56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Rode, Johannes; Müller, Sven (2016) : Spatio-temporal variation in peer effects - The case of rooftop photovoltaic systems in Germany, 56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/174663>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Spatio-Temporal Variation in Peer Effects

—

## The Case of Rooftop Photovoltaic Systems in Germany

Johannes Rode\*, Sven Müller†

April 25, 2016

### Abstract

We study spatio-temporal variation of peer effects in rooftop photovoltaic adoption of households. Our investigation employs geocoded data on all potential adopters and on all grid-connected photovoltaic systems set up in Germany through 2010. The detailed locational data allows us to construct an individual measure of peer effects for each potential adopter across Germany. Using a discrete choice model with panel data, we find evidence that the impact of peers on adoption decisions is non-linearly decreasing in distance to a location. The pattern is most pronounced for distances up to 200 meters. We also find that peer effects in photovoltaics adoption decrease over time and that German solar system adopters are of high socio-economic status.

**Keywords:** Peer effects, installed base, discrete choice, technology adoption, technology diffusion, imitation, photovoltaics, solar, Germany

*JEL classification:* O33, C35, Q55, R10

---

\*Corresponding author is Johannes Rode: Technische Universität Darmstadt, Chair of International Economics, Bleichstraße 2, 64283 Darmstadt, Germany, phone: 0049 6151 1657263, email: rode@vwl.tu-darmstadt.de

†Sven Müller: Karlsruhe University of Applied Sciences, Moltkestraße 30, 76133 Karlsruhe, Germany, email: sven.mueller@hs-karlsruhe.de

# 1 Introduction

New technologies often diffuse more slowly than would be optimal (Rogers, 1983; Geroski, 2000; Oster & Thornton, 2012). The diffusion of new technologies (in space and time) results from a series of individual decisions to adopt (i.e., to begin using the new technology). Understanding factors driving the adoption decision and identifying early adopter characteristics may help to foster the diffusion.

We analyze the spatio-temporal diffusion of all rooftop photovoltaic systems set up by households in Germany through 2010. Photovoltaics (PV) are solar cell systems for producing electric power. Several studies revealed that peer effects influence individual technology adoption decisions in general (Brock & Durlauf, 2010; Conley & Udry, 2010; Oster & Thornton, 2012) and PV adoption in particular (Bollinger & Gillingham, 2012; Müller & Rode, 2013; Islam, 2014; Rai & Robinson, 2013; Noll et al., 2014; Graziano & Gillingham, 2015; Richter, 2013; Rode & Weber, 2016). As rooftop PV systems have a high visibility, peers are – in this context – considered as proximate adopters of preceding periods. If peer effects indeed drive adoption, installation seeds could be used by firms and political decision-makers to raise the diffusion speed by steering adoption to locations where adoption is most intended and efficient (Müller & Rode, 2013; Islam, 2014; Graziano & Gillingham, 2015; Rode & Weber, 2016).

Most studies on PV adoption aggregate data to regions or local spatial units. Aggregating the data does not allow us to build a specific measure of the peer effect for each adopter and may therefore deliver biased results.

Bollinger & Gillingham (2012) analyze PV adoption in zip code areas in California. Their data set includes some 85,000 systems installed between 2001 and 2011.<sup>1</sup> In contrast, Müller & Rode (2013) study explicitly the individual decisions to adopt PV. They use an individual measure of the peer effect for each potential adopter but only consider about 300 PV systems installed in the city of Wiesbaden, Germany, through 2009.

Since the decision to adopt (in a certain period of time) is a discrete one (Karshenas & Stoneman, 1992), we, e.g., follow Gowrisankaran & Stavins (2004) and Müller & Rode (2013) and employ a discrete choice model (McFadden, 2001) to analyze all potential adopters' individual decisions to adopt a PV system across a whole country. I.e., we study the adoption of PV systems in Germany using individual (choice maker) level panel data through 2010. Exact locational data on all photovoltaics adopters (almost 900,000) and potential adopters (some 18 million) allows us to be the first to build a specific measure of the peer effect for each potential adopter per time period across a whole country. By doing so, we can confidently identify peer effects and find out how peer effects in PV system adoption vary over time and space.

The more previous technology adopters exist, the lower uncertainty regarding the technology may be, e.g. in terms of reliability, investments costs and operating costs. In consequence, valuable experiences from previous

---

<sup>1</sup>There are several other studies on PV adoption employing aggregate data: Graziano & Gillingham (2015) analyze a data set of almost 4,000 PV systems and aggregate it to block group level in Connecticut. Richter (2013) studies some 330,000 PV installations aggregated to some 2,200 zip codes in England and Wales. Rode & Weber (2016) analyze 580,000 PV systems installed in Germany through 2009. They aggregate the data to a 500 meter grid of rings which results in 1.4 million spatial units across Germany.

adopters may be less relevant during later stages of the diffusion path. We hypothesize that the influence of peer effects may decrease over time.

Further, we are interested in the determinants of districts with many PV adopters and PV adopter characteristics. Rogers (1983) describes early adopters as being of high socio-economic status. We hypothesize that in our study PV adopters may be associated with measures indicating high income and low population density (Graziano & Gillingham, 2015; Müller & Rode, 2013) as PV diffusion may still be in an early diffusion phase in Germany.

Photovoltaic systems are a sustainable energy technology. Due to a strong subsidy system, Germany had the highest PV capacity installed per capita in the world through 2012 (PVPS, 2013). Figure 1 illustrates the spatio-temporal dimension of our data. The figure shows that in many years there are substantially fewer new PV systems in the east than in the west of Germany, which may be due to a general economic development lag in the east (Redding & Sturm, 2008). This observation makes us hypothesize that the peer effect in PV adoption may also vary over space. There may be differences in the peer effect's level in the east and the west of Germany. Comin & Rode (2013) illustrate that more PV systems per capita are installed in rural areas compared to cities. Therefore, we also investigate differences in the peer effect in PV adoption between rural and urban areas.

Our analysis reveals that peer effects in PV system adoption are largely localized. The peer effect's impact on the decision to adopt decreases over time. We find different scales of the effects in rural and non-rural areas, and a larger peer effect in the east compared to the west of Germany. Our study indicates that changes in the subsidy system may have increased uncertainty

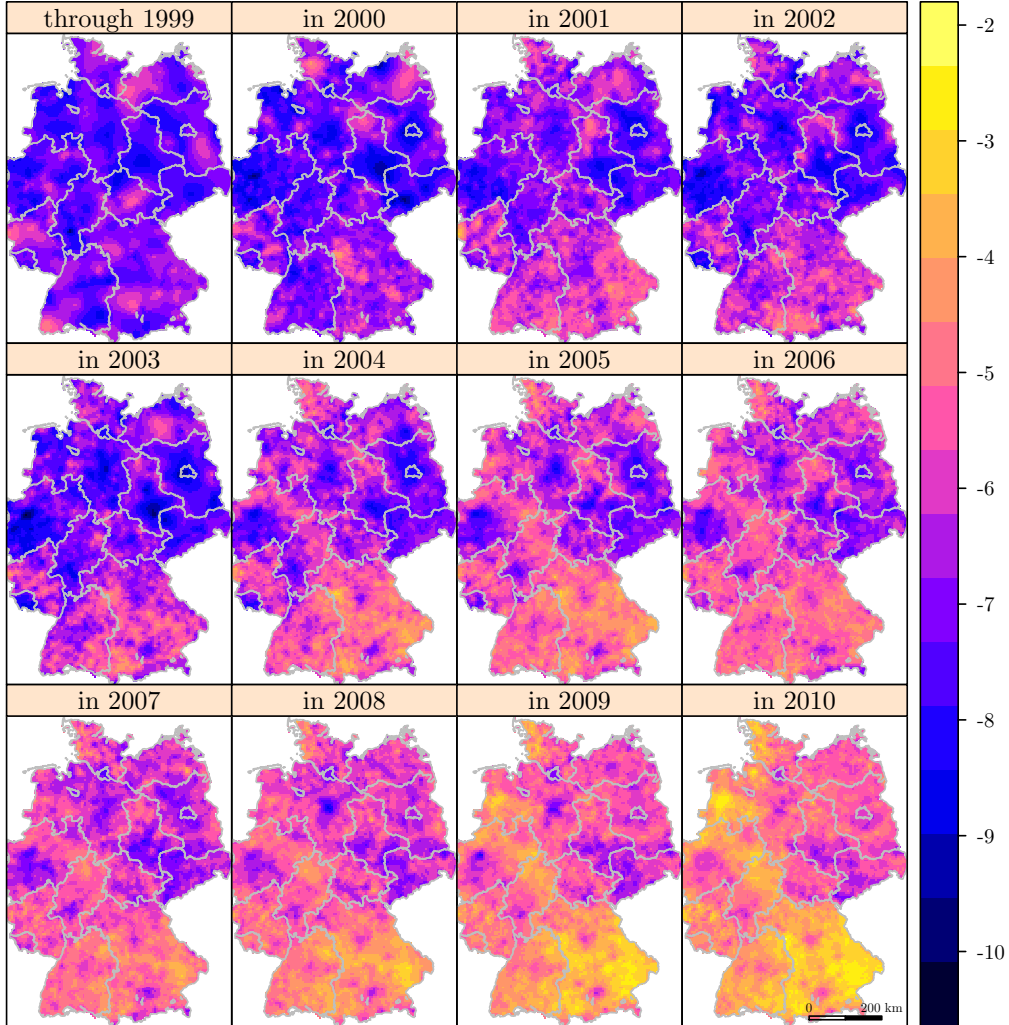


Figure 1: Natural logarithm of yearly annex of PV installations divided by number of potential adopters across Germany. The lighter a region is colored in the figure, the more PV systems are installed in the corresponding year while controlling for the number of buildings.

for potential adopters. In consequence, the relative importance of peer effects may have grown again.

Identifying peer effects comes with well-known challenges. We use spatial fixed effects (for 77,847 spatial districts) for each year to control for

unobserved heterogeneity. This procedure controls for different adoption probabilities during different stages of the common S-shaped technology diffusion path (Rogers, 1983) in every single district-year combination. This approach also controls for different subsidy levels per year (and the subsidy’s district-specific impact per year) or any other district-year-specific adoption shock. Robustness tests with individual adopter fixed effects and placebo regressions with randomly allocated PV systems across Germany confirm our findings. We also consider Manski’s (1993) ‘reflection problem’. The reflection problem refers to situations in which the adoption decision of an individual depends on others in her reference group and vice versa. Further, a case study with building-specific data on global radiation indicates that the peer effect in PV adoption may in fact be larger than our Germany-wide analysis reveals. Finally, we analyze a second data set on solar energy system adoption in Germany: an analysis of (representative) survey panel data from the socio-economic panel (SOEP, 2013) indicates that solar energy adopters are of high socio-economic status.

In Section 2, we introduce our modeling approach and our data. The results are shown in Section 3. Section 4 summarizes the paper and provides an outlook on further research.

## **2 Modeling approach and data**

In parts, we follow Gowrisankaran & Stavins (2004), West (2004) and Müller & Rode (2013) when we introduce our discrete choice (logit) model with panel data. Thereafter, we describe our data.

## 2.1 Estimation model

Logit regression analysis is a well-established approach in the literature on technology adoption (e.g., Feder et al. (1985) and Gowrisankaran & Stavins (2004)). In making the choice whether or not to adopt the technology, the choice maker weighs up the marginal advantages (and disadvantages) of adoption. Let us denote the (to the analyst) unobservable utility from adoption to choice maker  $n$  in period  $t$  as

$$u_{n,t} = v_{n,t} + \epsilon_{n,t} \quad (1)$$

with  $v_{n,t}$  as the deterministic, i.e., observable, utility of  $n$  to adopt in  $t$  and error term  $\epsilon_{n,t}$  that contains unobserved attributes. The actual choice,  $y_{n,t}$ , is a discrete (binary) choice. The observed values of  $y_{n,t}$  are related to  $u_{n,t}$  as follows:

$$y_{n,t} = \begin{cases} 1 & \text{if } u_{n,t} > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Because a choice maker adopts once at most within the considered time span the property  $\sum_t y_{n,t} \leq 1$  holds. Since  $u_{n,t}$  in (1) is a stochastic quantity, the probability that choice maker  $n$  chooses to install a photovoltaic system in year  $t$  is

$$P_{n,t} = \Pr(y_{n,t} = 1) = \Pr(u_{n,t} > 0). \quad (3)$$

If we now assume that  $\epsilon_{n,t}$  are independent and identically extreme value



distributed (Train, 2009), then

$$P_{n,t} = \frac{e^{v_{n,t}}}{1 + e^{v_{n,t}}}. \quad (4)$$

Accordingly, the odds ratio which defines the probability of adoption relative to non-adoption is given as:

$$\text{OR} = \frac{P_{n,t}}{1 - P_{n,t}} = e^{v_{n,t}}. \quad (5)$$

To measure the impact of the peer effect on the decision to adopt, i.e., to install, a photovoltaic system we consider

$$v_{n,t} = \beta \text{Ibase}_{n,t-1} + \gamma X_{n,t} + \delta_t + \alpha_{i_n} + \eta_{i_n,t} + \zeta_n. \quad (6)$$

In the following we explain the various variables and parameters of the deterministic utility (6). The installed base (Farrell & Saloner, 1986; Bollinger & Gillingham, 2012; Graziano & Gillingham, 2015)

$$\text{Ibase}_{n,t} = \sum_{\substack{m \in N, \\ m \neq n \\ d_{n,m} \leq D}} \sum_{l=0}^t y_{m,l} f(d_{n,m}) \quad (7)$$

is a spatio-temporal lag variable that measures the dependencies between choice makers  $n$  and  $m$ :  $\text{Ibase}_{n,t}$  considers the influence of the decision of choice maker  $m$  in periods through period  $t$  on the decision of choice maker  $n$  in period  $t$ .  $d_{n,m} > 0$  denotes the Euclidean distance in meters between the location of  $n$  and the location of  $m$ .  $D$  is a cut-off parameter to be set

by the analyst. We may assume that there is no remarkable influence of PV installations farther away from location  $n$  than  $D$ . Hence, the installed base measures the number of preexisting PV installations within radius  $D$  around the location of choice maker  $n$ , with the importance of each location weighted by  $f(d_{nm})$  (here, the importance declines in distance  $d_{n,m}$ ).

In our analysis (Section 3), we employ different values of  $D$  and functional forms of  $f(d_{n,m})$ . In our baseline specification, we set  $D = 200\text{m}$  and  $f(d_{n,m}) = 1/d_{n,m}$ . Of course, choice maker  $n$  might only be influenced by peers who adopted in the immediately preceding period, i.e., in  $t - 1$ . The corresponding measure is denoted as *Ibase non-cumulative* $_{n,t}$  and is defined in Appendix A.

Further, we denote the  $k$ -dimensional vector  $X_{n,t}$  as attributes related to choice maker  $n$  with vector  $\gamma$  to measure the respective effects. Based on (5) are

$$\text{OR}(\beta) = e^\beta, \text{ and} \quad (8)$$

$$\text{OR}(\gamma_k) = e^{\gamma_k} \quad (9)$$

the odds ratios associated with a one-unit increase in *Ibase* $_{n,t}$  (8) or a one-unit increase in the  $k$ th control variable (9).

$\delta_t$  are temporal fixed-effects common across choice makers and  $\alpha_{i_n}$  are locational effects fixed over time with  $i$  as the location related to choice maker  $n$ .  $\eta_{i_n,t}$  are locational fixed effects for every year  $t$  and  $\zeta_n$  are choice maker  $n$  fixed effects. Of course,  $\eta_{i_n,t}$ ,  $\alpha_{i_n}$  and  $\zeta_n$  may be numerous. Since our data are cross-sectional time-series (panel) data (see Section 2.2), we employ

the fixed-effects logit model for panel data as described in Greene (2012, pp. 721–724).<sup>2</sup> Using the corresponding conditional likelihood, we condition  $\delta_t$ ,  $\alpha_{i_n}$ ,  $\eta_{i_n,t}$  and  $\zeta_n$  out from our model (6).  $\beta$  is the parameter of focal interest since it measures the association of peer decisions on the odds of adopting, i.e., the peer effect.

## 2.2 Data

Since 80% of the PV systems in Germany are installed on roofs (BMU (2011), Dewald & Truffer (2011) and Rode & Weber (2016)), we consider buildings as the predominantly potential sites for PV systems. Therefore, we are only interested in PV systems on buildings. At some point, the owner of the building chooses to install (i.e., adopt) in a certain period or not.

We do not observe the choice directly. However, we obtain the location (in WGS84 coordinates) of 18,413,514 buildings in Germany in 2009 (Infas, 2009a).<sup>3</sup> We assume that each building is owned by someone. Of course, whether the building is owned by a private household or a house cooperation (or a firm) makes a difference in terms of adoption. Unfortunately, we do not obtain information on ownership. Still, to the best of our knowledge, no other study on PV system adoption, particularly those using aggregate approaches, accounts for ownership (see Bollinger & Gillingham (2012), Müller & Rode (2013), Graziano & Gillingham (2015) and Rode & Weber (2016)).

---

<sup>2</sup>We also show standard logit estimates for the pooled data in Section 3.1, 3.2 and 3.3.

<sup>3</sup>Note that Infas (2009a) only contains buildings with an address. We drop buildings which have the same coordinates assigned and end up with 18,413,514 buildings. We assume the number of buildings to be constant through our period of study. This assumption is reasonable as the number of residential building only increased by 7.4%: from 16,977,662 in 2001 to 18,234,580 in 2010 (DESTATIS, 2016).

We further assume that each building in Germany can be equipped with one PV system. We observe whether there is a PV system at a given building in a given year or not. Hence, we consider the buildings in Germany as the choice maker  $n$ . We study the years  $t = 2000, \dots, 2010$ .<sup>4</sup> As our measure of Ibase is lagged in time, our estimation sample contains the years 2001 to 2010.

We obtain a data set that covers all 879,020 grid-connected PV systems installed in Germany through the end of the year 2010 (I-TSO, 2012). The data set includes the location (we geocoded the address information to WGS84 coordinates) and the year of installation.

We first drop solar systems that are obviously solar parks (by excluding those which have an indicating key word, such as ‘Solarpark’, in their address). Then, each PV system is assigned to its closest building. See Appendix B for details. Finally, we end up with 877,114 PV systems, each allocated to a mutually exclusive building. As a result we obtain the observed choices on adoption, i.e., the values of  $y_{n,t}$  in (2).<sup>5</sup> The corresponding frequencies are given in Table 7 in Appendix D.1.

Figure 2 visualizes the spatio-temporal dimension of the dependent variable  $y_{n,t}$  by a small scale example. The example provides evidence that the spatio-temporal diffusion process might be influenced by preceding PV installations nearby. Our variable of interest  $Ibase_{n,t}$  (7) can be derived straightforward from this data.

---

<sup>4</sup>We start in 2000 since very few systems were set up before this date (see Figure 1 and Figure 3). In 2000, the Renewable Energy Sources Act (“Erneuerbare-Energien-Gesetz”) introduced a country-wide, high level feed-in tariff for electricity from PV. The level of the feed-in tariff mainly changed on a yearly basis. E.g., see Agnolucci (2006), Altrock et al. (2008) or Rode & Weber (2016) for details on the feed-in tariff.

<sup>5</sup>Note that  $y_{n,t}$  is only 1 if a PV system is installed by choice maker  $n$  in year  $t$ . In consequence,  $y_{n,t+1}$  will be 0 if a PV system was installed in  $t$ .

We presume that the peer effect might be spatially non-stationary (Section 1). Therefore, we consider a dummy variable  $East_{i_n}$  that indicates whether building  $n$  is located in a statistical district  $i$  which lies in the acceded territories of the former German Democratic Republic.



Figure 2: Example of the PV adoption process. Hollow circles represent potential adopters. Filled circles are PV installations. Newer PV installations are colored in light gray.

We also take into account the dummy variable  $Rural_{j_n}$  denoting whether building  $n$  is located in rural area  $j$ . Therefore, we use 2006's CORINE Land Cover (CLC) data set (CLC, 2009).<sup>6</sup> The interaction of these dummy variables with the installed base accounts for spatial-non-stationarity in peer effects.

<sup>6</sup>This data set comprises vector data on a scale of 1:100,000. The minimum mapping unit for the polygons is 0.25 sqkm. If  $n$  is located in a polygon that is not classified as *urban fabric* (CLC 111, 112), *industrial, commercial and transport units* (CLC 121, 123), *mine, dump and construction sites* (CLC 133), or *sport and leisure facilities* (CLC 142), then  $Rural_n$  equals one (zero, otherwise).

Of course, the utility a choice maker gains by adopting in a certain year or not may also be influenced by factors ( $X_{n,t}$ ) other than the installed base. Unfortunately, we do not have additional information on the building level. Instead, we consider time-invariant data on the location of building  $n$ .<sup>7</sup> In particular, *Global Radiation* $_{o_n}$  denotes the average yearly global radiation in 10 kilowatt hours per square meter (kWh/sqm) according to  $n$ 's location in one kilometer raster cells  $o$  provided by DWD (2010). A higher level of global radiation indicates a higher potential to produce electricity and therefore a higher remuneration potential to the owner of a PV system at a given location. Hence, we expect the higher global radiation is, the higher the utility from installing a PV system.

The elevation of  $n$ 's location might indirectly impact the choice to adopt as well. *Elevation* $_{p_n}$  denotes  $n$ 's elevation in 100m provided by 0.1 kilometer raster cells  $p$  from Jarvis et al. (2008). If building  $n$  exhibits low values of *Elevation* $_{p_n}$ , i.e., close to sea level, the propensity for shadowing may be lower compared to higher values of *Elevation* $_{p_n}$ . On the other hand, there may also be exposed sites on hills. Thus, we are unsure about the expected relation between *Elevation* $_{p_n}$  and the probability of adopting a PV system.

*Population Density* $_{i_n}$  denotes population density in sqm times 100 of the statistical district  $i$  where  $n$  is located in 2009. We use 77,847 statistical districts provided by Infas (2009b).<sup>8</sup> Low values of *Population Density* $_{i_n}$  may refer to places with a high share of single- and double-family homes. For

---

<sup>7</sup>Unfortunately, we lack time-variant data on a reasonably high spatial resolution.

<sup>8</sup>The average area of a statistical district is about 4.6 sqkm and the average number of buildings in such a statistical district is 280.

choice makers located in these places the decision to install a PV system may be easier as fewer parties have to agree upon the installation on a certain building. As a consequence, we expect the probability of PV adoption to decline in *Population Density<sub>i<sub>n</sub></sub>*.

We expect that buildings located in statistical districts with high values of *Firm Density<sub>i<sub>n</sub></sub>* – denoting the number of firms per sqkm – exhibit low probabilities of adoption. Of course, it is reasonable to expect that the share of single- and double-family homes is low in statistical districts with high firm densities.

*Purchasing Power<sub>i<sub>n</sub></sub>* is the purchasing power index of the statistical district  $i$  according to building  $n$ . An index value of 10 corresponds to the median purchasing power of German households in 2009 (Infas, 2009b). The propensity of high purchasing power of a choice maker located in a wealthy statistical district is higher compared to a choice maker located in a statistical district that exhibits low purchasing power. Since PV installations are expensive (e.g. see Comin & Rode (2013)), we expect that the probability of PV adoption increases with *Purchasing Power<sub>i<sub>n</sub></sub>*. The summary statistics of *Purchasing Power<sub>i<sub>n</sub></sub>* and all the other variables can be found in Table 1.

### 3 Results and discussion

Estimating peer effects comes with well-known challenges. Failing to control for unobserved heterogeneity can result in biased estimates. We control for temporal (i.e., year) fixed effects capturing time-varying factors that have a similar effect on the increase in the adoption decision across all spatial

Table 1: Descriptive statistics of full data set (specification M1, M2, M3 and M10).

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.0047	.068	0	1
$Ibase_{n,t-1}$	$1/d_{n,m}$	.0093	.021	0	1
$Rural_{j_n}$		.2	.4	0	1
$East_{i_n}$		.18	.38	0	1
$Global\ Radiation_{o_n}$		103	5.6	93	121
$Elevation_{p_n}$		2.1	1.9	-.08	19
$Firm\ Density_{i_n}$		.1	.27	0	14
$Population\ Density_{i_n}$		1.9	3.5	0	134
$Purchasing\ Power_{i_n}$		9.9	1.6	0	32
N	184,135,140				

*Notes:* 184,135,140 observations come from 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states.

units: for example, changes in federal legislation fostering PV adoption or changes in their installation costs. The first column in Table 2 illustrates the estimates for specification M1 with year fixed effects. The relevant descriptive statistics are shown in Table 1 in Section 2.2. Specification M1 in Table 2 shows the exponentiated coefficient, which we can interpret as odds ratio (see equation (8)).

Specification M1 reveals that the installed base has an odds ratio significantly greater than one. In consequence, the installed base has a positive influence on the decision to install a PV system. That is, the more proximate PV systems in the preceding years, the higher the propensity of a potential adopter to obtain a PV system in the current year. Potential adopters might be influenced by the decisions of their peers. Imitation of spatially close precursors might indeed be an explanatory factor in PV system adoption; i.e., our results provide evidence for localized peer effect in the adoption of PV systems.



Table 2: Odds ratio of spatio-temporal variation of peer effects in Germany.

$f(d_{n,m})$ Cut-off	(1) M1 $1/d_{n,m}$ 200m	(2) M2 $1/d_{n,m}$ 200m	(3) M3 $1/d_{n,m}$ 200m	(4) M4 $1/d_{n,m}$ 200m	(5) M5 $1/d_{n,m}$ 200m
$Ibase_{n,t-1}$	93,901.7*** (18.39)	383,469.7*** (21.86)	201,922.4*** (18.98)	213.1*** (4.85)	8424.2*** (10.67)
$Ibase_{n,t-1} \times Year_{2002}$	0.501 (-0.89)	0.192* (-2.16)	0.0695*** (-3.42)	0.199 (-1.27)	0.00184*** (-6.78)
$Ibase_{n,t-1} \times Year_{2003}$	0.143** (-2.72)	0.0489*** (-4.32)	0.0171*** (-5.76)	0.0475* (-2.44)	0.000656*** (-8.36)
$Ibase_{n,t-1} \times Year_{2004}$	0.804 (-0.30)	0.306 (-1.68)	0.0205*** (-5.56)	0.00705*** (-4.25)	0.0000713*** (-11.08)
$Ibase_{n,t-1} \times Year_{2005}$	1.731 (0.82)	0.552 (-0.94)	0.0150*** (-6.26)	0.00731*** (-4.33)	0.00000656*** (-14.01)
$Ibase_{n,t-1} \times Year_{2006}$	0.263* (-2.07)	0.0800*** (-4.12)	0.00232*** (-9.23)	0.00640*** (-4.46)	0.000000439*** (-17.14)
$Ibase_{n,t-1} \times Year_{2007}$	0.137** (-3.14)	0.0433*** (-5.22)	0.00149*** (-9.98)	0.00320*** (-5.13)	0.000000464*** (-17.22)
$Ibase_{n,t-1} \times Year_{2008}$	0.0900*** (-3.84)	0.0307*** (-5.87)	0.00124*** (-10.32)	0.00270*** (-5.32)	0.000000670*** (-16.88)
$Ibase_{n,t-1} \times Year_{2009}$	0.0341*** (-5.41)	0.0121*** (-7.47)	0.000658*** (-11.34)	0.00168*** (-5.77)	0.000000772*** (-16.74)
$Ibase_{n,t-1} \times Year_{2010}$	0.00986*** (-7.41)	0.00353*** (-9.58)	0.000271*** (-12.74)	0.00176*** (-5.73)	0.000000876*** (-16.59)
$Ibase_{n,t-1} \times Rural_{j_n}$		0.0399*** (-53.80)	0.0316*** (-50.11)	0.143*** (-20.50)	1.855*** (5.07)
$Ibase_{n,t-1} \times East_{i_n}$		5.379*** (7.86)	95.46*** (20.92)	3.670*** (3.96)	1.970 (1.54)
$Ibase_{n,t-1} \times Rural_{j_n} \times East_{i_n}$		36.87*** (8.24)	50.68*** (8.79)	20.86*** (4.42)	1.774 (0.62)
$East_{i_n}$		0.446*** (-151.84)	0.819*** (-16.00)	1.154 (0.34)	
$Rural_{j_n}$		2.385*** (327.61)	2.186*** (283.68)	1.371*** (77.89)	
$Rural_{j_n} \times East_{i_n}$		0.589*** (-52.35)	0.608*** (-48.58)	0.811*** (-15.58)	
Observations	184,135,140	184,135,140	184,135,140	91,145,630	8,581,810
DF <sub>M</sub>	19	25	39	16	22
Final log-likelihood $\mathcal{L}$	-5,165,155	-5,078,625	-4,997,443	-4,087,285	-1,709,292
LR: $\chi^2$ (DF)		173,061 (6)	162,363 (14)		
LR: p-value		0	0		
LR test against		M2-> M1	M3-> M2		
Year <sub>t</sub> fixed effects	Yes	Yes	Yes	No	Yes
Federal state <sub>n</sub> fixed effects	No	No	Yes	No	No
District <sub>t</sub> $\times$ Year <sub>t</sub> fixed effects	No	No	No	Yes	No
Choice maker <sub>n</sub> fixed effects	No	No	No	No	Yes

Exponentiated coefficients; robust  $t$  statistics in parentheses\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

Notes: The dependent variable is always  $y_{n,t}$ , i.e., choice maker  $n$ 's observed choice in terms of (newly) installing a PV system in  $t$ . Specification M1, M2 and M3 are estimated via logit. The panel of 18,413,514 choice makers over 10 years  $t$  (2001-2010) results in the 184,135,140 observations. We estimate specification M4 and M5 with the conditional logit estimator (see footnote 16). The sample is the same as in specification M1-M3. However, the conditional logit estimator drops all positive (or all negative) outcomes (in terms of District<sub>t</sub>  $\times$  Year<sub>t</sub> groups for specification M4, and in terms of choice maker<sub>n</sub> groups for specification M5). I.e., for specification M4 District<sub>t</sub>  $\times$  Year<sub>t</sub> groups with no adoption (or if all choice makers adopt at once), and, for specification M5, all non-adopters are dropped. This procedure results in less observations.

### 3.1 Temporal variation of peer effects

As we want to study temporal differences in the peer effect, specification M1 provides insights on the peer effect per year. An obvious interpretation is in terms of a 0.01 unit increase in  $Ibase_{n,t-1}$  which, e.g., corresponds to one previously installed PV system 100 meters away. Specification M1 shows that for a 0.01 unit increase in  $Ibase_{n,t-1}$  in 2001, we expect an increase of  $(\exp(\ln(93,902)/100) - 1) \times 100\% = 12\%$  in the odds of installing a PV system.

According to specification M1, the peer effect may decrease over time. For example, multiplying  $Ibase_{n,t-1} \times Year_{2002}$ 's exponentiated coefficient with the exponentiated coefficient of  $Ibase_{n,t-1}$  indicates the odds ratio for 2002. The level in the odds of installing in year 2002, 2004 and 2005 does not significantly vary in comparison to 2001. However, a 0.01 increase (one system 100 meters away) in the installed base, e.g, indicates an increase of  $(\exp(\ln(93,902 \times 0.09)/100) - 1) \times 100\% = 9.5\%$  in the odds of installing a PV system in 2008. Further, a 0.01 unit increase in  $Ibase_{n,t-1}$  boosts the odds of installing PV by 8.3%<sup>9</sup> in 2009 and by 7.1%<sup>10</sup> in 2010. In consequence, when policy-makers want to take advantage of peer effects to stimulate diffusion, they should bear in mind that peer effects are more relevant during the very early phase of technology diffusion.

---

<sup>9</sup> $(\exp(\ln(93,902 \times 0.03)/100) - 1) \times 100\% = 8.3\%$ .

<sup>10</sup> $(\exp(\ln(93,902 \times 0.01)/100) - 1) \times 100\% = 7.1\%$ .

### 3.2 Spatial variation of peer effects

The purpose of this paper is not only to identify the temporal variation of peer effects but also to study their spatial variation. Specification M2 (column (2) in Table 2) includes the time-varying measure of the installed base and also controls for the east and rural parts of Germany as well as their interaction. Since we include these interaction terms, M2's odds ratio of 383,470 for  $Ibase_{n,t-1}$  refers to non-rural areas in the west of Germany in 2001. Specification M2 in Table 2 reveals that for a 0.01 increase in the installed base, we expect an increase of  $(\exp(\ln(383,470)/100) - 1) \times 100\% = 14\%$  in the odds of installing a PV system (in non-rural areas in the west in 2001). A likelihood ratio test confirms, M2 is superior to M1.

Interestingly, we find evidence for spatial non-stationarity: the odds ratio for the interaction between  $Ibase_{n,t-1}$  and  $East_{i_n}$  is significantly greater than one, i.e., the peer effect may be more important (in non-rural areas) in the east of Germany. For example, for M2, multiplying  $Ibase_{n,t-1} \times East_n$ 's exponentiated coefficient with the exponentiated coefficient of  $Ibase_{n,t-1}$  and considering a 0.01 increase in the installed base yields an increase of 16%<sup>11</sup> in the odds of installing a PV system (for non-rural areas) in the east (in 2001). Since the fraction of buildings with PV is lower in the east (see Figure 3), the east may still be in a very early stage of the S-shaped diffusion path. Our estimates on temporal variation of peer effects indicate that the peer effect may be more important in the early stage of diffusion. During this stage, uncertainty regarding the reliability of a PV system may be higher: therefore,

---

<sup>11</sup> $(\exp(\ln(383,470 \times 5.4)/100) - 1) \times 100\% = 16\%$ .

information from peers might be more important in the east compared to the west of Germany. Similarly, the interaction between  $Ibase_{n,t-1}$  and  $Rural_{j_n}$  is significantly less than 1, i.e., the peer effect may be more important in non-rural than in rural areas. In 2001, the increase in odds from a 0.01 increase in  $Ibase_{n,t-1}$  is 10%<sup>12</sup> for rural areas in the west, whereas it is 16%<sup>13</sup> for rural areas in the east.

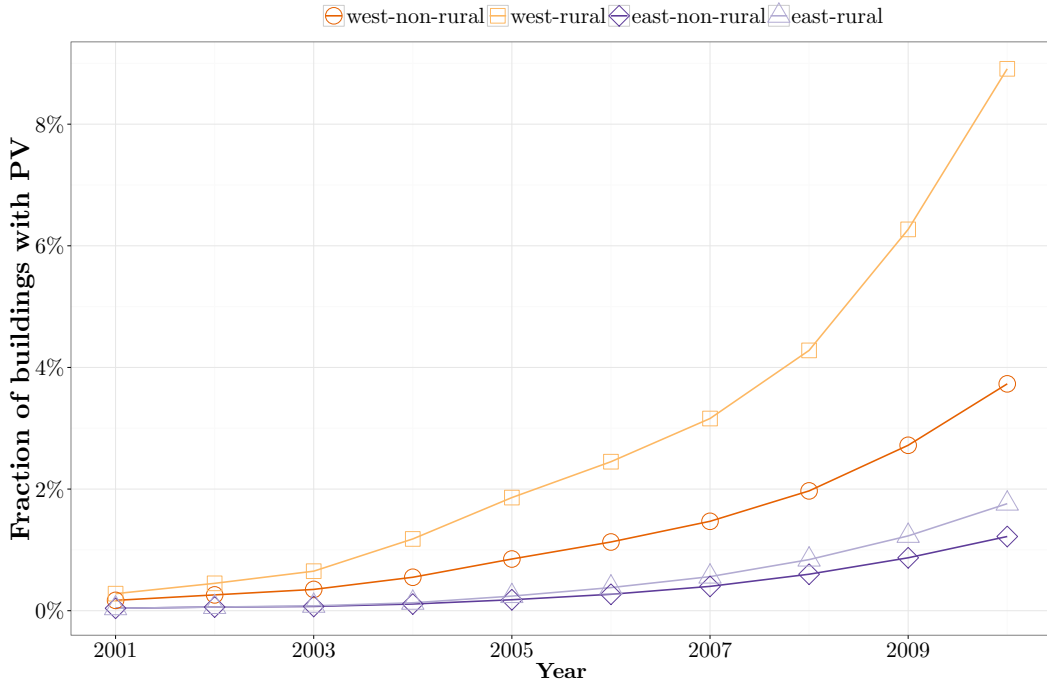


Figure 3: Cumulative fraction of buildings with PV system by east and west, and rural and non-rural.

### 3.3 Omitted variables

*Federal state and year fixed effects:* To measure peer effects more confidently, we should consider omitted variables. We can control for spatial fixed effects,

<sup>12</sup> $(\exp(\ln(383,470 \times 0.04)/100) - 1) \times 100\% = 10\%$ .

<sup>13</sup> $(\exp(\ln(383,470 \times 0.04 \times 5.4 \times 36.9)/100) - 1) \times 100\% = 16\%$ .

e.g., on the federal state (NUTS-1) level.<sup>14</sup> There are 16 federal states in Germany. Federal state fixed effects absorb federal state-specific effects in the adoption decision and could be caused by federal state-specific characteristics affecting the usability of the technology: e.g., additional time-invariant incentives to install PV in a specific federal state. M3 in Table 2 accounts for federal state and year fixed effects and confirms our previous findings. Again, although of lower magnitude, there is a significantly positive peer effect, which is lower in rural areas and higher in the east. A likelihood ratio test confirms that we should prefer M3 over M2.

*District  $\times$  year fixed effects:* Omitted variables on the local scale may drive our results. We consider 77,847 districts taken from Infas (2009b). To be very accurate, we take into account time-variant adoption shocks on the district level. Such a shock could be, for example, a local advertisement campaign by a PV seller, a new local subsidy fostering PV installations, a housing development in which new local regulations force residents to install PV, or simply different propensities to adopt at different stages on the S-shaped diffusion path. We can control for such shocks by including time-variant district specific fixed effects: district  $\times$  year fixed effects.<sup>15</sup>

Table 2 shows the estimates for our *baseline specification* M4 with district  $\times$  year fixed effects. The relevant descriptive statistics are shown in Table 8 of Appendix D.1.<sup>16</sup> Specification M4 in Table 2 reveals that for a

---

<sup>14</sup>NUTS stands for “Nomenclature des unités territoriales statistiques”. It is the European Union’s Nomenclature of territorial units for statistics: a hierarchical system which divides the economic territory (in case of NUTS-1 into the major socio-economic regions).

<sup>15</sup>A district fixed effect for every year of study, i.e.  $77,847 \times 10 = 778,470$  fixed effects.

<sup>16</sup>We use the conditional (fixed effects) logit model for panel data as described in Greene (2012, pp. 721-724). Note, when estimating a conditional logit model, groups with all

0.01 increase in  $Ibase_{n,t-1}$  we expect an increase of 5.5%<sup>17</sup> in the odds of installing a PV system (in non-rural areas in the west, in 2001). In line with the previous estimates, the peer effect is lower in rural areas and higher in the east.

As before, specification M4 confirms the peer effect's decrease over time. For a 0.01 increase in the installed base, we expect an increase of 3.8%<sup>18</sup> in the odds of installing a PV system in 2002 (in non-rural areas in the west). A 0.01 unit increase in  $Ibase_{n,t-1}$  increases the odds of installing PV only by 2.3%<sup>19</sup> in 2003 and by 0.4%<sup>20</sup> in 2004 (for non-rural areas in the west). Figure 4 illustrates the diminishing odds to install from a 0.01  $Ibase_{n,t-1}$  increase over time. The figure also illustrates differences between rural vs. non-rural and east vs. west.<sup>21</sup> In consequence, when policy-makers want to take advantage of peer effects to stimulate diffusion, they should have in mind that peer effects are mainly relevant during the very early phase of technology diffusion (here, through 2003).

The figure also indicates a break in the diminishing odds of installing from increases in  $Ibase_{n,t-1}$  in 2004. This break may be linked to changes in the feed-in tariff by the Amendment of the Renewable Energy Sources Act in 2004 (e.g., see Agnolucci (2006) and Altrock et al. (2008)). Between 2004

---

positive (or all negative) outcomes are dropped, i.e., the number of observations is lower in specification M4 (and M5, which we discuss below) than for specification M1-M3 (and M10, which we discuss below). I.e., for specification M4  $District_i \times Year_t$  groups with no adoption (or if all choice makers adopt at once) are dropped. For M1, M2, M3 (and M10) the number of observations corresponds to the number of buildings in Germany times the number of years under study.

<sup>17</sup> $(\exp(\ln(213)/100) - 1) \times 100\% = 5.5\%$ .

<sup>18</sup> $(\exp(\ln(213 \times 0.2)/100) - 1) \times 100\% = 3.8\%$ .

<sup>19</sup> $(\exp(\ln(213 \times 0.05)/100) - 1) \times 100\% = 2.3\%$ .

<sup>20</sup> $(\exp(\ln(213 \times 0.007)/100) - 1) \times 100\% = 0.4\%$ .

<sup>21</sup>We should not be concerned by odds ratios smaller than zero as these naturally come with saturation.

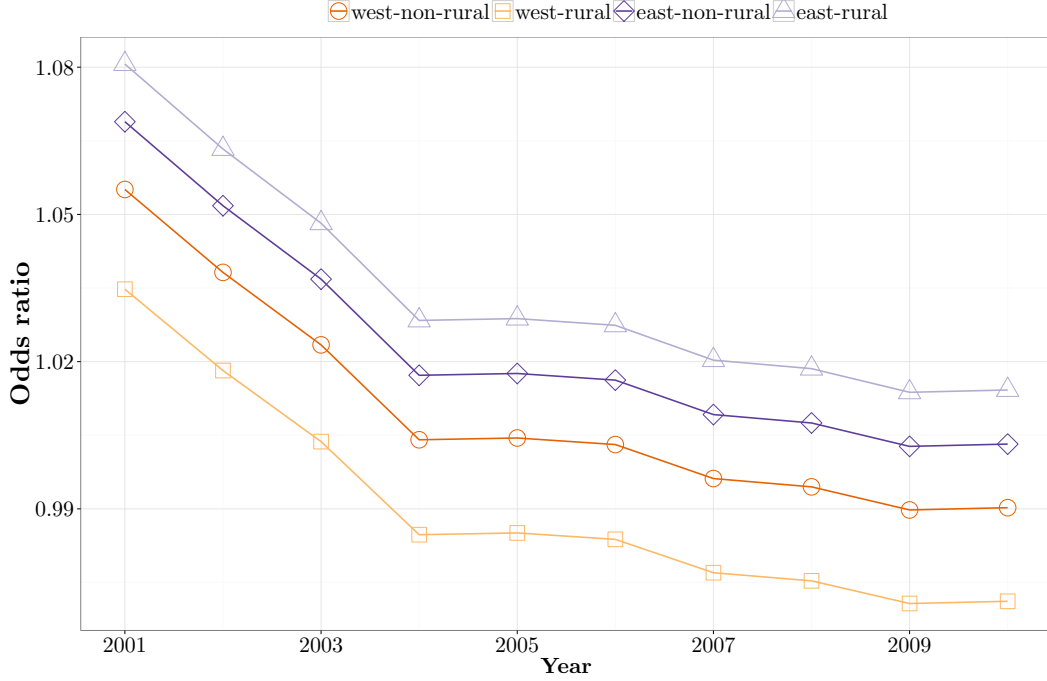


Figure 4: Time-diminishing odds ratio to install a PV system due to increase in the installed base by one system 100 meters away for M4.

and 2009 the odds if installing from increases in  $I_{base_{n,t-1}}$  decreased again. In 2009, we observe another break. This break may be linked to changes in the feed-in tariff by 2009's Amendment of the Renewable Energy Sources Act (e.g., see Altrock et al. (2008)). 2009's amendment put in place an incentive for self consumption of PV electricity. Both breaks indicate that changes in the subsidy system may have increased uncertainty for potential adopters. In consequence, the relative importance of peers may have grown again.

*Individual-specific and year fixed effects:* Our baseline regression includes district  $\times$  year fixed effects. We could also think of individual-specific effects which may influence the decision to install a PV system. Therefore, specification M5 (column (5) in Table 2) includes choice maker (building)

fixed effects and year fixed effects. Table 9 of Appendix D.1 contains the corresponding descriptive statistics. When including choice maker fixed effects, only those choice makers who installed a PV system during our period of study are considered in the analysis. M5 confirms a significantly positive peer effect for 2001, 2002 and 2003 (in non-rural areas in the west). In the following years the peer effect becomes negative, due to saturation.<sup>22</sup> Specification M5 confirms that the peer effect is higher in the east. In contrast to the previous results, M5 indicates a higher peer effect in rural areas than in non-rural areas. Consequently, we are unsure how the peer effect differs between rural and non-rural areas.

Rode & Weber (2016) find evidence for a larger peer effect at higher levels of remaining non-adopters. This is in line with our finding of a higher peer effect in the very early stages of diffusion. In contrast, Bollinger & Gillingham’s (2012, p. 905) analysis indicates that the peer effect may increase over time. Possibly, reasons for PV adoption largely vary in Bollinger & Gillingham’s sample from California and ours from Germany, where feeding PV electricity into the grid is highly subsidized: E.g. Bollinger & Gillingham mention that marketing efforts leveraged peer effects in the later periods of their study in California.

### 3.4 Robustness checks

In the following, we conduct several robustness tests. We return to our baseline specification M4 with district  $\times$  year fixed effects.

---

<sup>22</sup>In the east and in rural areas in the west, saturation starts in 2005 (according to specification M5).



*Different cut-off distances:* In Table 12 of Appendix D.2, we show estimates with adjusted peer effect measures (and district  $\times$  year fixed effects). The results from our baseline regression M4 remain unaffected from reducing the cut-off distance to 100m (M6, see column (1) in Table 10 in Appendix D.2 for the descriptives) or increasing the cut-off distance to 400m (M7, column (2) in Table 11 in Appendix D.2 contains the descriptives). According to these results, the peer effect is highly localized: considering installations farther away than 400 meters (or only those closer than 100 meters) significantly decreases the explanatory power (see Horowitz test between M4 and M7 (or M6) in Table 12 of Appendix D.2).<sup>23</sup> We conclude that studies concentrating on peer effects at a high level of geographical aggregation may fail to represent the peer effect appropriately. This in line with previous studies on PV adoption. Bollinger & Gillingham (2012) find stronger peer effects on the street level than on the zip code level. Rode & Weber (2016)’s analysis also indicates a peer effect which occurs between adopters and potential adopters who live within circles of 1 km radius or less around a 500 m grid of points across Germany. Our analysis indicates that peer effects in PV adoption may be even more localized. We find no additional explanatory power from PV installations located farther away than 200 meters from a potential adopter.

*Only last year’s installations:* Incorporating PV installations from the last year instead of all previous installations in the installed base measure – as outlined in Appendix A – does not affect our core findings (in Appendix D.2, see M8 in Table 12, and Table 8 for the descriptives). This result indicates

---

<sup>23</sup>E.g., Ben-Akiva & Lerman (1985) describe the Horowitz non-nested hypothesis test. The Horowitz test checks the hypothesis if the model with the lower adjusted likelihood ratio index is the true model.

that – besides the most recent installations – older ones are also relevant for the peer effect. Still, not shown estimates indicate that the most recent previous installations are the most important for the peer effect. This finding is in line with Graziano & Gillingham’s (2015, p. 19) evidence for a “*diminishing neighbor effect over time since prior installations*”. M9 reveals that redefining  $f(d_{n,m})$  to  $1/d_{n,m}^2$  confirms our previous results. In any case, Horowitz tests confirm that M4’s explanatory power is significantly higher than M6-M9’s.

*Capacity-adjusted samples:* So far, we have taken into account all PV installations across Germany. We can use the capacity – a measure of size – of each PV system to conduct a natural robustness test. A peer effect should only exist for (small) household systems. In contrast, we do not expect industrial investors to be affected by their neighbors. An investor who wants to install a (large) industrial PV system should search for the best location across a large region or even the whole country.<sup>24</sup> In line with our previous findings, M4<sub>≤30kW<sub>p</sub></sub> (column (1) of Table 16 in Appendix D.3) confirms a time-decreasing peer effect for household systems (below a capacity of 30kW<sub>p</sub>). Table 13 (in Appendix D.3) contains the corresponding descriptive statistics. As expected, M4<sub>>1MW<sub>p</sub></sub> (column (2) in Table 16 of Appendix D.3) rejects a significantly positive peer effect for industrial PV systems (above a capacity of 1MW<sub>p</sub>). Table 14 (in Appendix D.3) contains the corresponding descriptive statistics. Of course, the installed base measure’s mean is smaller for industrial PV systems than for household

---

<sup>24</sup>Also see Comin & Rode (2013) for a similar approach to distinguishing household and industrial PV adoption.

systems. The fact that we can only confirm a peer effect for household PV systems also supports our approach of considering buildings as choice makers.

*Placebo test:* We conduct a placebo test to verify that our estimations do not by definition find a positive peer effect in PV system adoption. We randomly allocate the same number of PV installations, which were in fact installed in Germany per year, to the existing buildings. In Appendix D.3,  $M4_{\text{Placebo}}$  (column (3) in Table 16) shows estimates for the resulting data set and Table 15 the corresponding descriptive statistics.  $M4_{\text{Placebo}}$  reveals no significantly positive peer effect in PV adoption. I.e., the placebo test supports the validity of our findings.<sup>25</sup>

*The reflection problem:* Our findings remain unaffected by lagging our peer effect measure by two years (e.g., see Richter (2013) and Rode & Weber (2016)).<sup>26</sup> We can therefore confidently rule out that the ‘reflection problem’ described by Manski (1993) biases our results. The reflection problem applies to situations where the adoption decision of an individual  $n$  depends on others in  $n$ ’s reference group and  $n$ ’s adoption also affects other group members. Bearing in mind that we study PV adoption on a yearly basis, it is reasonable to assume that an individual who adopts in  $t$  also decided to adopt in  $t$  or  $t-1$ . If so,  $i$  was affected by the behavior of reference group members at  $t-2$  or before. In consequence, the time lag is very likely to rule out the possibility that the individual could have affected the adoption decision of reference group members, which in turn influenced the individual’s adoption.<sup>27</sup>

---

<sup>25</sup>Note that the difference in the descriptives is due to dropping districts with all zero outcomes in a year (see footnote 16).

<sup>26</sup>The results are available upon request.

<sup>27</sup>We also conduct a similar test where we neglect previous installations within a radius of 100m. Then, identifying the peer effect only relies on previous installations in a ring

*Neglecting randomly allocated systems:* In Section 3, we outline our data processing approach. One relevant step is that we allocate PV systems to their nearest building but allocate a PV system randomly to another building (in the same district) if the nearest building is already occupied. We confirm the robustness of our findings to neglecting the PV systems which are allocated randomly to another building.<sup>28</sup>

*Neglecting PV adopters after adopting:* So far, we have analyzed a balanced panel. The dependent variable  $y_{n,t}$  is the choice of  $n$  in terms of (newly) installing a PV system in  $t$ . In consequence, we have treated PV adopters in the years after they have adopted PV in the same way as non-adopters.<sup>29</sup> However, there are no more choice options after adopting. We can simply take this into account by constructing an unbalanced panel which excludes the PV adopters in the years after they have adopted PV (but still consider the PV adopters when calculating  $I_{base_{n,t-1}}$ ). This procedure confirms our previous results.<sup>30</sup>

### 3.5 Case study

We conduct a case study for the cities of Darmstadt, Karlsruhe, Marburg, and Wiesbaden. We study these cities because building-specific global radiation data [Global radiation (Building-specific) $_n$ ] is available. The case study helps us to check if we can indeed disregard building-specific data on global radiation for our analysis of Germany. Building-specific data on global

---

between 100m and 200m distance to the choice maker (see Graziano & Gillingham (2015) for a similar analysis). This approach confirms our previous findings.

<sup>28</sup>The results are available upon request.

<sup>29</sup>In consequence, we have systematically underestimated the peer effect.

<sup>30</sup>The results are available upon request.

radiation indicates that certain buildings are not appropriate or less appropriate for PV systems, e.g. due to shadowing from neighboring buildings or a building’s roof orientation and inclination. The case study allows us to test whether neglecting these factors in the Germany-wide analysis is appropriate. Table 3 shows the case study data.

Table 3: Characteristics for case study cities.

Characteristic	Darmstadt	Karlsruhe	Marburg	Wiesbaden	Sum
Residents in 2010	144,402	294,761	80,656	275,976	795,795
Number of buildings	100,004	83,819	40,872	113,548	338,243
Number of PV systems through 2010	374	1,058	450	463	2,345
Districts	109	302	56	177	644

Data on residents in 2010 is taken from DESTATIS (2013).

Darmstadt (2008), KEK (2010), Marburg (2011) and Wiesbaden (2009) provide spatial data from the cities including information on buildings and detailed data on global radiation.<sup>31</sup> Appendix D.4 gives details on the case study data.

Due to the low number of PV systems installed in the four cities through 2010, we estimate a time-invariant coefficient for the peer effect. Table 4 again shows odds ratios. Specification MA, MB, and MC (column (1), (2), (3) in Table 4) include district  $\times$  year fixed effects. Employing building-specific data on global radiation confirms a significantly positive peer effect. Comparing specification MA with MC indicates that we may have underestimated the peer effect when studying all installations across Germany: the

<sup>31</sup>Certainly, the level of detail is different for the specific cities. For example, in 2010, Karlsruhe had more than twice as many residents than Darmstadt (see Table 3). Nevertheless, in our data set the number of buildings is larger in Darmstadt compared to Karlsruhe. Since we obtain similar results from the following study for each city individually, we neglect the different level of detail (which we also control for by our fixed effects).

odds ratio for  $\text{Ibase}_{n,t-1}$  increases if we control for building-specific data on global radiation. A likelihood ratio test shows that using building-specific data on global radiation (MB->MC) significantly improves the explanatory power of our case study estimations.<sup>32</sup>

Table 4: Odds ratio of peer effects for Case Study.

	(1) MA $1/d_{n,m}$	(2) MB $1/d_{n,m}$	(3) MC $1/d_{n,m}$
$f(d_{n,m})$			
Cut-off	200m	200m	200m
$\text{Ibase}_{n,t-1}$	17.44*** (3.72)	16.98*** (3.68)	27.92*** (4.12)
Global Radiation $_{on}$		0.891 (-0.62)	0.881 (-0.68)
Global Radiation (Building-specific) $_n$			1.014*** (23.77)
Observations	1,089,105	1,089,105	1,089,105
DF <sub>M</sub>	1	2	3
Final log-likelihood $\mathcal{L}$	-13,655	-13,655	-13,248
LR: $\chi^2$ (DF)		1 (1)	814 (1)
LR: p-value		.45	4.5e-179
LR test against		MB-> MA	MC-> MB
Period $_t$ fixed effects	No	No	No
Federal state $_n$ fixed effects	No	No	No
District $_i \times$ Period $_t$ fixed effects	Yes	Yes	Yes
Address $_n$ fixed effects	No	No	No

Exponentiated coefficients; robust  $t$  statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: The dependent variable is always  $y_{n,t}$ , i.e., choice maker  $n$ 's observed choice in terms of (newly) installing a PV system in  $t$ . The panel of 338,242 choice makers in the four cities over 10 years  $t$  (2001-2010) results in the 3,382,420 observations. We estimate specification MA, MB and MC with the conditional logit estimator (see footnote 16). Note that the conditional logit estimator drops all positive (or all negative) outcomes (here, in terms of District $_n \times$  Year $_t$  groups). I.e., District $_i \times$  Year $_t$  groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in 1,089,105 observations.

### 3.6 District and adopter characteristics

*District characteristics:* We are interested in the characteristics of districts with many PV systems. If we take into account district characteristics, we cannot include fixed effects at the local level. Besides controlling for federal state and year fixed effects, specification M10 (column (4) in Table 16, Ap-

<sup>32</sup>The insignificance of Global Radiation $_{on}$  (specification MB) makes sense as we control for district  $\times$  year fixed effects and variation in Global Radiation $_{on}$  is close to zero within each district-year.

pendix D.3) accounts for spatial differences in global radiation, elevation, firm density, population density and purchasing power (Table 1 contains the corresponding descriptives). These controls improve the model fit significantly (see likelihood ratio test against M3) and indicate an even larger effect for our measure of the installed base (in comparison with specification M3). M10 confirms a significant, diminishing peer effect over time, which is lower in rural areas and higher in the east.

Regarding district characteristics, we observe that a unit increase in global radiation increases the odds of installing a PV system by  $(1.05 - 1) \times 100\% = 5\%$  (in non-rural areas in the west).<sup>33</sup> This finding is appropriate as a high level of solar radiation indicates a higher potential to produce electricity and confirms a large income potential from a PV system. If we compare the area with the lowest value of global radiation ( $93 \times 10 \text{ kWh/sqm}$ ) with the one with the maximum amount ( $121 \times 10 \text{ kWh/sqm}$ ), we observe that the odds of installing is  $(\exp(\ln(1.05) \times 28) - 1) \times 100\% = 292\%$  higher in the German region with the maximum global radiation (in non-rural areas in the west). Although in both cases the odds of installing PV rise, increases in global radiation are associated with lower increases in the odds of installing PV in the east and with higher increases in rural-areas.

A unit increase in elevation – referring to 100m – increases the odds of installing PV by 4% (in non-rural areas in the west). In contrast, a unit increase in elevation decreases the odds of installing PV between 6% and 7% in rural areas and the east. Indeed, if choice maker (building)  $n$  exhibits a

---

<sup>33</sup>A one unit increase refers to an increase in  $10 \text{ kWh/sqm}$  of yearly averaged global radiation.

low value of  $Elevation_{p_n}$  the propensity for shadowing by hills may be lower compared to higher values of  $Elevation_{p_n}$ . Similarly, the odds of installing PV decreases by 22% with every unit increase in firm density (in non-rural areas in the west). This decrease is even larger in rural areas but does not differ significantly between the east and the west. The negative association of firm density and PV adoption makes sense as most PV systems are private.

According to M10, disproportionately many PV systems can be found in sparsely populated areas. Graziano & Gillingham (2015) also find more PV systems in less densely populated areas. A one unit increase (0.3 standard deviations) in population density decreases the odds of installing PV by 13% (in non-rural areas in the west). The decrease in the odds is higher in rural areas but lower in the east. We expect that the propensity of choice makers located in areas with low population density to own a house is high. This finding indicates again that most PV installations are residential small-scale systems. Graziano & Gillingham (2015) find weak evidence for a positive association between household income and increased adoption. Our estimates may indicate that more PV systems can be found in areas with low purchasing power. However, measures of income (or purchasing power) may be heterogeneous within a statistical district  $i$ . To account for this income heterogeneity, we consider individual survey panel data from the Socio-Economic Panel (SOEP, 2013) representative for Germany.

*Adopter characteristics:* About 30,000 individuals from 11,000 German households take part in the SOEP every year. Interestingly, the SOEP includes yearly data on solar energy system adoption for the years 2007-2012 in Germany. Solar energy refers to PV systems and solar thermal systems. The



difference between both is that while solar thermal systems produce energy that can only be used to heat water, PV systems produce electricity that can be either used or sold to the electric grid. In terms of an analysis like ours, there is no relevant difference between both types of solar energy systems (Comin & Rode, 2013).<sup>34</sup>

There is no small scale locational information included in the public SOEP data, i.e., it does not allow us to study peer effects. However, analyzing the individual level data allows us to investigate whether PV adopters are of high socio-economic status.

A choice maker  $n$ 's household is  $h$ . Some variables are available on the household level  $h$ , other variables are on the choice maker level  $n$ , both for each year  $t$ .  $Solar_{h,n,t}$  is one if household  $h$  adopts a new solar energy system in year  $t$  (otherwise zero). See Appendix E for details on the variable definitions and their descriptives.

The estimates in Table 5 illustrate that solar energy adopters are of high socio-economic status. Choice makers living in large dwellings, who are employed and own the dwelling they live in have higher odds of adopting a solar energy system (see first column in Table 5). The estimates also indicate a negative association of PV adoption and population density, e.g. measured by a significantly positive odds ratio of PV adoption for households with a garden. If we focus on home owners who have a garden (second column in Table 5), we see that solar energy adoption is also associated with high levels

---

<sup>34</sup>Since 2007, most solar energy systems installed in Germany have been PV systems. In 2007 there were 1 million solar thermal systems installed in Germany while there were only 360,000 PV systems (BSW-Solar, 2014). By 2012, the number of PV systems was 1.3 million while the number of solar thermal systems was 1.8 million.

of real household income. Focusing on home owners (with a garden) makes sense as these will probably have taken the decision to adopt a solar energy system themselves (Comin & Rode, 2013). In contrast, tenants may be less involved in the adoption decision.<sup>35</sup>

Table 5: Odds ratio of solar energy adopter characteristics for SOEP data.

	All	Home owners with garden
	(1)	(2)
	$\Delta Solar_{h_{n,t}}$	$\Delta Solar_{h_{n,t}}$
$\ln(\text{Real household income}_{h_{n,t}})$	1.114 (1.51)	1.338*** (2.67)
Dwelling size $_{h_{n,t}}$	1.004*** (5.16)	1.003*** (2.68)
Vocational training $_{n,t}$	0.987 (-0.17)	1.020 (0.18)
College $_{n,t}$	0.963 (-0.50)	0.813** (-2.03)
Working $_{n,t}$	1.237*** (3.30)	
Garden $_{h_{n,t}}$	2.436*** (5.81)	
Home owner $_{h_{n,t}}$	1.264*** (2.60)	
Time $\times$ NUTS-1 fixed effects	Yes	Yes
Observations	85,568	39,671
DF <sub>M</sub>	7	8
Final log-likelihood $\mathcal{L}$	-6684.8	-4059.9

Exponentiated coefficients; robust  $t$  statistics in parentheses clustered on federal state-year groups

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Notes: The dependent variable is always  $y_{n,t}$ , i.e., choice maker  $n$ 's observed choice in terms of (newly) installing a solar system in  $t$ . We estimate specification MI and MII with the conditional logit estimator (see footnote 16). The conditional logit estimator drops all outcomes (here, in terms of Federal state $_i \times$  Year $_t$  groups). I.e., Federal state $_i \times$  Year $_t$  groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in less observations.

Not shown estimates indicate that the socio-economic status of solar energy adopters is even higher in the east than in the west. Solar energy adopters from the east have a significantly higher real household income and live in larger dwellings. This finding confirms a previous result illustrated in Figure 3: Since the fraction of buildings with PV installations is lower in the east, the east may still be in a very early stage of the S-shaped diffusion path.

<sup>35</sup>We only take into account respondents who do not claim to have removed their solar system. There is strong evidence that this group best illustrates the effects under study. Comparing the SOEP data set with reported data from the German transmission system operators, shows that disproportionately many solar systems were removed according to the SOEP data, see Comin & Rode (2013).

## 4 Conclusion

To the best of our knowledge, this is the most comprehensive study on PV adoption conducted so far (in terms of the number of adopters). We built upon a geocoded data set of all PV systems in Germany and all potential adopters through 2010. This paper is the first which uses an individual peer effects measure for each potential PV adopter across a whole country.

In line with Bollinger & Gillingham (2012), Müller & Rode (2013), Graziano & Gillingham (2015) and Rode & Weber (2016), we confirm peer effects in PV adoption. However, our study indicates that peer effects mainly impact PV adoption during the very early phases of the diffusion path in Germany. Further, the spatial range of peer effects is very limited (only to about 200m). Several robustness test confirm these findings.

Our results also indicate that changes in the subsidy system may have increased uncertainty for potential adopters. Due to the increased uncertainty, the relative importance of peer effects may have grown again. Finally, we find that solar energy adopters are of high socio-economic status.

The policy implications of our study are straightforward. Our results reveal that it would be efficient to influence the adoption of PV systems using seed installations in the very early periods of diffusion since the importance of peer effects decreases over time. Seeding is most promising at locations with low population density and high global radiation with many households of high socio-economic status.<sup>36</sup>

---

<sup>36</sup>Certainly, we study PV adoption after political decision makers decided to foster PV. We do not evaluate which energy technology or which mix of energy technologies is the best choice for a certain country.

Further disentangling of the peer effect may be rewarding. For example, it may be interesting to find out if the identified effect is driven by communication from face to face or whether seeing a PV system at work is sufficient.

## Acknowledgments

We express our deep gratitude for the friendly support of Bing’s Spatial Data Service that allowed us to geocode our data set. We gratefully acknowledge that the land surveying offices in Darmstadt, Karlsruhe, Marburg and Wiesbaden provided base maps including information on solar radiation. We want to thank Harry Korn, Birgit Groh, Marion Kühn and Jutta-Maria Braun at the respective institutions. Further, we are very grateful for the kind support of Martina Klärle and Katharina Englbrecht from SUN-AREA. Additionally, our sincere gratitude goes to *Deutscher Wetterdienst*, which provided data on global radiation across Germany. Furthermore, we are grateful for helpful comments from colleagues. Johannes Rode acknowledges funding from the Graduate School for Urban Studies (URBANgrad) at Technische Universität Darmstadt (kindly supported by LOEWE-Schwerpunkt “Eigenlogik der Städte”) and the support of the Chair of International Economics at Technische Universität Darmstadt. An early version of this paper was part of Johannes Rode’s doctoral thesis.

## References

- Agnolucci, P. (2006). Use of economic instruments in the German renewable electricity policy. Energy Policy 34(18), 3538–3548. doi:10.1016/j.enpol.2005.08.010.
- Altrock, M., V. Oschmann, & C. Theobald (2008). EEG: Erneuerbare-Energien-Gesetz; Kommentar (2nd ed.). Munich: Beck.
- Ben-Akiva, M. & S. Lerman (1985). Discrete choice analysis, theory and applications to travel demand. Cambridge, MA: MIT Press.
- BMU (2011). Vorbereitung und Begleitung der Erstellung des Erfahrungsberichtes 2011 gemäß § 65 EEG – Vorhaben II c – Solare Strahlungsenergie – Endbericht. Technical report, Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Berlin. Im Auftrag des Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU), Projektleitung: Matthias Reichmuth – Leipziger Institut für Energie GmbH. Accessed April 10, 2013. Available from: [http://www.erneuerbare-energien.de/fileadmin/ee-import/files/pdfs/allgemein/application/pdf/eeg\\_eb\\_2011\\_solare\\_strahlung\\_bf.pdf](http://www.erneuerbare-energien.de/fileadmin/ee-import/files/pdfs/allgemein/application/pdf/eeg_eb_2011_solare_strahlung_bf.pdf).
- Bollinger, B. & K. Gillingham (2012). Environmental preferences and peer effects in the diffusion of solar photovoltaic panels. Marketing Science 31(6), 900–912. doi:10.1287/mksc.1120.0727.
- Brock, W. A. & S. N. Durlauf (2010). Adoption curves and social interactions. Journal of the European Economic Association 8(1), 232–251. doi:10.1111/j.1542-4774.2010.tb00501.x.
- BSW-Solar (2014). Statistische Zahlen der deutschen Solarwärmebranche (Solarthermie). Technical report, Bundesverband Solarwirtschaft e.V. Acc. Oct. 13, 2014. Available from: [http://www.solarwirtschaft.de/fileadmin/media/pdf/2014\\_03\\_BSW\\_Solar\\_Faktenblatt\\_Solarwaerme.pdf](http://www.solarwirtschaft.de/fileadmin/media/pdf/2014_03_BSW_Solar_Faktenblatt_Solarwaerme.pdf).

- CLC (2009). CORINE Land Cover from the year 2006, Umweltbundesamt, deutsches Zentrum für Luft- und Raumfahrt (DLR), Deutsches Fernerkundungsdatenzentrum (DFD). Accessed June 13, 2012. Available from: [http://www.corine.dfd.dlr.de/datadescription\\_2006\\_de.html](http://www.corine.dfd.dlr.de/datadescription_2006_de.html).
- Comin, D. & J. Rode (2013). From Green Users to Green Voters. Working Paper #19219, National Bureau of Economic Research. doi:10.3386/w19219.
- Conley, T. G. & C. R. Udry (2010). Learning about a new technology: Pineapple in Ghana. American Economic Review 100(1), 35–69. doi: 10.1257/aer.100.1.35.
- Darmstadt (2008). AEROSOLAR 3D Solarpotential. Vermessungsamt Darmstadt. Data was provided after personal request. Acc. March 29, 2012.
- DESTATIS (2013). Regionalstatistik, Sachgebiete: 12411 Fortschreibung des Bevölkerungsstandes, Tabelle: 173-01-5-b Bevölkerungsstand: Bevölkerung nach Geschlecht – Stichtag 31.12. – regionale Ebenen; data for 2010: 06411000 Darmstadt, krsfr. Stadt; 08212000 Karlsruhe, Stadtkreis; 06534014 Marburg; 06414000 Wiesbaden, krsfr. Stadt. Database, Statistische Ämter des Bundes und der Länder. Acc. Jan. 17, 2013.
- DESTATIS (2016). Fortschreibung Wohngebäude- und Wohnungsbestand, Tabelle: 31231-0001 Anzahl der Wohnungen in Wohngebäuden - Stichtag 31.12. – deutschland insgesamt. Database, Statistische Ämter des Bundes und der Länder. Acc. Feb. 25, 2016.
- Dewald, U. & B. Truffer (2011). Market formation in technological innovation systems – Diffusion of photovoltaic applications in Germany. Industry & Innovation 18(3), 285–300. doi:10.1080/13662716.2011.561028.
- DWD (2010). 1-km-Rasterdaten der mittleren jährlichen Globalstrahlung (kWh/qm, Zeitraum 1981-2000) der gesamten Bundesrepublik Deutschland. Data provided after personal request, July 26, 2010. Available from: <http://www.dwd.de/>.

- Farrell, J. & G. Saloner (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. American Economic Review 76(5), 940–955. Available from: <http://www.jstor.org/stable/1816461>.
- Feder, G., R. E. Just, & D. Zilberman (1985, Jan.). Adoption of agricultural innovations in developing countries: A survey. Economic Development and Cultural Change 33(2), 255–298. doi:10.1086/451461.
- Geroski, P. A. (2000). Models of technology diffusion. Research Policy 29(4–5), 603–625. doi:10.1016/S0048-7333(99)00092-X.
- Gowrisankaran, G. & J. Stavins (2004). Network externalities and technology adoption: Lessons from electronic payments. The RAND Journal of Economics 35(2), 260–276. Available from: <http://www.jstor.org/stable/1593691>.
- Graziano, M. & K. Gillingham (2015). Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment. Journal of Economic Geography 15(4), 815–839. doi:10.1093/jeg/1bu036.
- Greene, W. (2012). Econometric Analysis (7 ed.). New Jersey: Pearson Education.
- I-TSO (2012). EEG-Anlagenstammdaten. Informationsplattform der Deutschen Übertragungsnetzbetreiber, Acc. Mar. 4, 2012. <http://www.eeg-kwk.net/de/Anlagenstammdaten.htm>.
- Infas (2009a). Gebäudeverzeichnis, Hauskoordinaten, KGS44. Data set, Infas Geodaten GmbH, Bonn. Purchased data, provided after personal request, December, 2010. Available from: <http://www.infas-geodaten.de/>.
- Infas (2009b). Wohnquartiere, Kreis-Gemeinde-Schlüssel (KGS) 22, 2009. Purchased data, provided after personal request, December, 2010. Available from: <http://www.infas-geodaten.de/>.



- Islam, T. (2014). Household level innovation diffusion model of photo-voltaic (PV) solar cells from stated preference data. Energy Policy 65, 340–350. doi:10.1016/j.enpol.2013.10.004.
- Jarvis, A., H. Reuter, A. Nelson, & E. Guevara (2008). Hole-filled seamless shuttle radar topographic mission (SRTM) data V4, International Centre for Tropical Agriculture (CIAT). Acc. July 30, 2012. <http://srtm.csi.cgiar.org>.
- Karshenas, M. D. & P. Stoneman (1992). A flexible model of technological diffusion incorporating economic factors with an application to the spread of colour television ownership in the UK. Journal of Forecasting 11(7), 577–601. doi:10.1002/for.3980110702.
- KEK (2010). GIS-gestützte Standortanalyse für Photovoltaik- und thermische Solaranlagen mittels Laserscannerdaten, SUN-AREA. Data was provided after personal request, Karlsruher Energie- und Klimaschutzagentur, Karlsruhe. Acc. April 11, 2012.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60(3), 531–542. Available from: <http://www.jstor.org/stable/2298123>.
- Marburg (2011). GIS-gestützte Standortanalyse für Photovoltaik- und thermische Solaranlagen mittels Laserscannerdaten, SUN-AREA. Data was provided after personal request, Stadt Marburg,. Marburg. Acc. July 19, 2012.
- McFadden, D. (2001). Economic choices. American Economic Review 91(3), 351–378. doi:10.1257/aer.91.3.351.
- Müller, S. & J. Rode (2013). The adoption of photovoltaic systems in Wiesbaden, Germany. Economics of Innovation and New Technology 22(5), 519–535. doi:10.1080/10438599.2013.804333.

- Noll, D., C. Dawes, & V. Rai (2014). Solar community organizations and active peer effects in the adoption of residential PV. Energy Policy 67, 330–343. doi:10.1016/j.enpol.2013.12.050.
- Oster, E. & R. Thornton (2012). Determinants of technology adoption: Peer effects in menstrual cup take-up. Journal of the European Economic Association 10(6), 1263–1293. doi:10.1111/j.1542-4774.2012.01090.x.
- PVPS (2013). A Snapshot of Global PV 1992-2012. Report IEA-PVPS T1-22:2013, International Energy Agency (IEA) – Photovoltaic Power Systems Programme (PVPS). Available from: <http://www.iea-pvps.org/index.php?id=92>.
- Rai, V. & S. A. Robinson (2013). Effective information channels for reducing costs of environmentally-friendly technologies: evidence from residential PV markets. Environmental Research Letters 8(1), 014044. Available from: <http://stacks.iop.org/1748-9326/8/i=1/a=014044>.
- Redding, S. J. & D. M. Sturm (2008). The costs of remoteness: Evidence from german division and reunification. American Economic Review 98(5), 1766–97. doi:10.1257/aer.98.5.1766.
- Richter, L.-L. (2013). Social effects in the diffusion of solar photovoltaic technology in the UK. Technical report, Energy Policy Research Group. EPRG Working Paper 1332. Available from: <http://www.eprg.group.cam.ac.uk/eprg1332/>.
- Rode, J. (2014). Renewable Energy Adoption in Germany - Drivers, Barriers and Implications. Doctoral thesis, TU Darmstadt, Darmstadt.
- Rode, J. & A. Weber (2016). Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany. Journal of Environmental Economics and Management 78, 38–48. doi:10.1016/j.jeem.2016.02.001.

- Rogers, E. M. (1983). Diffusion of Innovations (3rd ed.). New York: Free Press.
- SOEP (2013). Socio-Economic Panel, data for years 1984-2012. version 29. doi:10.5684/soep.v29.
- Train, K. E. (2009). Discrete choice methods with simulation. Cambridge University Press.
- West, S. E. (2004). Distributional effects of alternative vehicle pollution control policies. Journal of Public Economics 88(3-4), 735–757. doi: 10.1016/S0047-2727(02)00186-X.
- Wiesbaden (2009). GIS-gestützte Standortanalyse für Photovoltaik- und thermische Solaranlagen mittels Laserscannerdaten, SUN-AREA. Data was provided after personal request, Stadt Wiesbaden. Wiesbaden, Acc. March 29, 2012.

The following is not intended to be included in the journal version of the article, but as an online appendix.

## Appendix

### A Non-Cumulative Peer Effects

Since the subsidy system for PV systems changes during the period of study, only peers who adopted in the preceding year ( $t - 1$ ) may pass on reliable information regarding, e.g., the reliability, initial costs, and the net present value of PV systems. Therefore, we consider the non-cumulative installed base as

$$\text{Ibase non-cumulative}_{n,t} = \sum_{\substack{m \in N, \\ m \neq n, \\ d_{n,m} \leq D}} o_{m,t,1} f(d_{n,m}). \quad (10)$$

### B Allocation of PV systems to buildings

Table 6 in Appendix C shows the geocoding accuracy of the address data on 879,020 PV installations. We neglect the fact that a PV system may be uninstalled because this is the case for only 0.35% of the systems under study. Each of the 879,020 PV installations is assigned to its closest building. 269,752 buildings end up with more than one allocated PV system.<sup>37</sup> This is due to inaccuracy in geocoding, missing address information in the PV

---

<sup>37</sup>These do not drive our results. We confirm our findings with a data set which neglects the PV systems allocated to these 269,752 buildings.

system data set and the possibility that more than one PV system has been installed on one building or the PV system is installed on a building with no address. We randomly allocate the PV systems from these observations to another building located in the same statistical district.<sup>38</sup> The 77,847 statistical districts are taken from Infas (2009b). This procedure results in 877,114 PV systems, each allocated to a mutually exclusive building.

## C Geocoding

For only 28,862 – thus, 3.28% – of the installations the entity type of the geocoding accuracy is unknown and the confidence less than medium, see Table 6. However, 86.59% of the installations have a high geocoding confidence, and 86.54% have a high geocoding confidence and are allocated to addresses or road block entities.

Table 6: Geocoding accuracy.

Entity type	Confidence	Frequency in category
Address	High	599,746
Address	Medium	25,786
Neighborhood	Medium	62
PopulatedPlace	High	8
PopulatedPlace	Medium	3,201
Postcode1	High	18
Postcode1	Medium	19,092
RoadBlock	High	160,929
RoadBlock	Medium	40,524
RoadIntersection	High	402
RoadIntersection	Medium	390
Unknown		28,862
Sum		879,020

<sup>38</sup>Later on, we conduct robustness checks which confirm that this procedure does not drive our results.

## D Tables

### D.1 Frequencies and descriptives

Table 7: Frequencies of choices ( $\Delta y_{n,t}$ ).

Year <sub>t</sub>	Alternative	Frequency in category
$\leq 1999$	0: No (or no new) PV system installed	18,404,144
	1: New PV system installed	9,370
2000	0: No (or no new) PV system installed	18,403,951
	1: New PV system installed	9,563
2001	0: No (or no new) PV system installed	18,388,390
	1: New PV system installed	25,124
2002	0: No (or no new) PV system installed	18,394,739
	1: New PV system installed	18,775
2003	0: No (or no new) PV system installed	18,393,381
	1: New PV system installed	20,133
2004	0: No (or no new) PV system installed	18,366,533
	1: New PV system installed	46,981
2005	0: No (or no new) PV system installed	18,347,253
	1: New PV system installed	6,6261
2006	0: No (or no new) PV system installed	18,349,923
	1: New PV system installed	63,591
2007	0: No (or no new) PV system installed	18,336,697
	1: New PV system installed	76,817
2008	0: No (or no new) PV system installed	18,298,322
	1: New PV system installed	115,192
2009	0: No (or no new) PV system installed	18,231,899
	1: New PV system installed	181,615
2010	0: No (or no new) PV system installed	18,169,822
	1: New PV system installed	243,692
Sum of PV installations over all years		877,114

Table 8: Descriptive statistics specification M4, M8 and M9.

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.0094	.097	0	1
Ibase <sub>n,t-1</sub>	$1/d_{n,m}$	.015	.026	0	1
Ibase non-cumulative <sub>n,t</sub>	$1/d_{n,m}$	.0041	.011	0	.9
Ibase <sub>n,t</sub>	$1/d_{n,m}^2$	.00032	.0011	0	.08
Rural <sub>j<sub>n</sub></sub>		.22	.42	0	1
East <sub>i<sub>n</sub></sub>		.13	.33	0	1
N	91,145,630				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of District <sub>$i$</sub>   $\times$  Year <sub>$t$</sub>  groups). I.e., District <sub>$i$</sub>   $\times$  Year <sub>$t$</sub>  groups with no adoption (or if all choice makers adopt at once), are dropped. This procedure results in 91,145,630 observations.

Table 9: Descriptive statistics specification M5.

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.1	.3	0	1
Ibase $_{n,t-1}$	$1/d_{n,m}$	.015	.03	0	.99
Rural $_{j_n}$		.34	.47	0	1
East $_{i_n}$		.069	.25	0	1
N	8,581,810				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of choice maker $_n$ ). I.e., choice makers with no adoption are dropped. This procedure results in 8,581,810 observations.

## D.2 Adjustments to peer effects measure

Table 10: Descriptive statistics specification M6.

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.0094	.097	0	1
$\text{Ibase}_{n,t-1}$	$1/d_{n,m}$	.0087	.021	0	1
$\text{Rural}_{j_n}$		.22	.42	0	1
$\text{East}_{i_n}$		.13	.33	0	1
N	91,145,630				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of  $\text{District}_i \times \text{Year}_t$  groups). I.e.,  $\text{District}_i \times \text{Year}_t$  groups with no adoption (or if all choice makers adopt at once), are dropped. This procedure results in 91,145,630 observations.

Table 11: Descriptive statistics specification M7.

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.0094	.097	0	1
$\text{Ibase}_{n,t-1}$	$1/d_{n,m}$	.0048	.017	0	.9
$\text{Rural}_{j_n}$		.22	.42	0	1
$\text{East}_{i_n}$		.13	.33	0	1
N	91,145,630				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of  $\text{District}_i \times \text{Year}_t$  groups). I.e.,  $\text{District}_i \times \text{Year}_t$  groups with no adoption (or if all choice makers adopt at once), are dropped. This procedure results in 91,145,630 observations.



Table 12: Odds ratio of spatio-temporal variation of peer effects in Germany.

$f(d_{n,m})$	(1) M6 $1/d_{n,m}$	(2) M7 $1/d_{n,m}$	(3) M8 $1/d_{n,m}$	(4) M9 $1/d_{n,m}^2$
Cut-off	100m	400m	200m	200m
Ibase $_{n,t-1}$	193.7*** (4.79)	58.33** (3.27)		1.12314e + 18* (2.42)
Ibase $_{n,t-1} \times \text{Year}_{2002}$	0.157 (-1.46)	0.271 (-0.92)		0.00218 (-0.31)
Ibase $_{n,t-1} \times \text{Year}_{2003}$	0.0473* (-2.42)	0.107 (-1.58)		3.79e - 10 (-1.09)
Ibase $_{n,t-1} \times \text{Year}_{2004}$	0.00881*** (-4.05)	0.0183** (-3.03)		2.34e - 21* (-2.57)
Ibase $_{n,t-1} \times \text{Year}_{2005}$	0.00775*** (-4.29)	0.0157** (-3.24)		5.38e - 21** (-2.63)
Ibase $_{n,t-1} \times \text{Year}_{2006}$	0.00557*** (-4.58)	0.00980*** (-3.62)		1.29e - 25** (-3.23)
Ibase $_{n,t-1} \times \text{Year}_{2007}$	0.00414*** (-4.91)	0.0105*** (-3.61)		7.57e - 25** (-3.17)
Ibase $_{n,t-1} \times \text{Year}_{2008}$	0.00302*** (-5.23)	0.00704*** (-3.95)		3.15e - 26*** (-3.39)
Ibase $_{n,t-1} \times \text{Year}_{2009}$	0.00184*** (-5.70)	0.00450*** (-4.33)		1.96e - 29*** (-3.83)
Ibase $_{n,t-1} \times \text{Year}_{2010}$	0.00200*** (-5.63)	0.00506*** (-4.24)		1.91e - 28*** (-3.71)
Ibase $_{n,t-1} \times \text{Rural}_{j_n}$	0.158*** (-16.73)	0.194*** (-12.52)		8.93e - 10*** (-9.14)
Ibase $_{n,t-1} \times \text{East}_{i_n}$	2.803** (2.93)	2.461* (2.32)		4589362.9** (2.79)
Ibase $_{n,t-1} \times \text{Rural}_{j_n} \times \text{East}_{i_n}$	16.16*** (3.88)	13.74*** (3.42)		5.49223e + 14** (3.29)
Rural $_{j_n}$	1.362*** (80.86)	1.352*** (81.95)	1.364*** (82.62)	1.347*** (80.48)
Rural $_{j_n} \times \text{East}_{i_n}$	0.818*** (-15.28)	0.824*** (-14.86)	0.819*** (-15.26)	0.827*** (-14.60)
Ibase non-cumulative $_{n,t-1}$			452.6*** (4.30)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2002}$			0.123 (-1.29)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2003}$			0.0302* (-2.04)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2004}$			0.00911** (-3.00)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2005}$			0.0174** (-2.74)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2006}$			0.0185** (-2.71)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2007}$			0.00542*** (-3.54)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2008}$			0.00581*** (-3.55)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2009}$			0.00170*** (-4.44)	
Ibase non-cumulative $_{n,t-1} \times \text{Year}_{2010}$			0.00216*** (-4.30)	
Ibase non-cumulative $_{n,t-1} \times \text{Rural}_{j_n}$			0.0461*** (-15.19)	
Ibase non-cumulative $_{n,t-1} \times \text{East}_{i_n}$			6.213** (3.07)	
Ibase non-cumulative $_{n,t-1} \times \text{Rural}_{j_n} \times \text{East}_{i_n}$			53.87** (3.21)	
Observations	91,145,630	91,145,630	91,145,630	91,145,630
DF <sub>M</sub>	16	16	16	16
Final log-likelihood $\mathcal{L}$	-4,087,505	-4,087,571	-4,087,894	-4,087,522
Horowitz test statistic (signific. level)	-4 (4.3e-06)	-5 (2.0e-07)	-7 (7.0e-14)	-5 (2.0e-06)
Horowitz test against	M4-> M6	M4-> M7	M4-> M8	M4-> M9
Year <sub>t</sub> fixed effects	No	No	No	No
Federal state <sub>n</sub> fixed effects	No	No	No	No
District <sub>i</sub> $\times$ Year <sub>t</sub> fixed effects	Yes	Yes	Yes	Yes
Choice maker <sub>n</sub> fixed effects	No	No	No	No

Exponentiated coefficients; robust  $t$  statistics in parentheses\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

Notes: The dependent variable is always  $y_{n,t}$ , i.e., choice maker  $n$ 's observed choice in terms of (newly) installing a PV system in  $t$ . We estimate specification M6-M9 with the conditional logit estimator (see footnote 16). The sample is the same as in specification M1-M3 of Table 2. However, the conditional logit estimator drops all positive (or all negative) outcomes (here, in terms of District <sub>$i$</sub>   $\times$  Year <sub>$t$</sub>  groups). I.e., District <sub>$i$</sub>   $\times$  Year <sub>$t$</sub>  groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in less observations.

### D.3 Capacity-specific estimates, placebo test and district characteristics

Table 13: Descriptive statistics specification  $M4_{\leq 30kW_p}$ .

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.0086	.092	0	1
$Ibase_{n,t-1}$	$1/d_{n,m}$	.014	.026	0	1
$Rural_{j_n}$		.22	.42	0	1
$East_{i_n}$		.12	.33	0	1
N	88,156,268				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of  $District_i \times Year_t$  groups). I.e.,  $District_i \times Year_t$  groups with no adoption (or if all choice makers adopt at once), are dropped. This procedure results in 88,156,268 observations.

Table 14: Descriptive statistics specification  $M4_{>1MW_p}$ .

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.003	.055	0	1
$Ibase_{n,t-1}$	$1/d_{n,m}$	.0011	.0052	0	.16
$Rural_{j_n}$		.23	.42	0	1
$East_{i_n}$		.3	.46	0	1
N	4,206,326				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of  $District_i \times Year_t$  groups). I.e.,  $District_i \times Year_t$  groups with no adoption (or if all choice makers adopt at once), are dropped. This procedure results in 4,206,326 observations.

Table 15: Descriptive statistics specification  $M4_{placebo}$ .

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.0074	.086	0	1
$Ibase_{n,t-1}$	$1/d_{n,m}$	.048	.046	0	.96
$Rural_{j_n}$		.19	.39	0	1
$East_{i_n}$		.18	.38	0	1
N	115,408,422				

*Notes:* There are 18,413,514 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 77,847 districts  $i$  and 16 federal states. The conditional logit estimator drops all positive (or all negative) outcomes (in this case, in terms of  $District_i \times Year_t$  groups). I.e.,  $District_i \times Year_t$  groups with no adoption (or if all choice makers adopt at once), are dropped. This procedure results in 115,408,422 observations.

Table 16: Odds ratio of spatio-temporal variation of peer effects in Germany.

	(1) M4 <sub>≤30kW<sub>P</sub></sub> 1/ $d_{n,m}$ 200m	(2) M4 <sub>&gt;1MW<sub>P</sub></sub> 1/ $d_{n,m}$ 200m	(3) M4Placebo 1/ $d_{n,m}$ 200m	(4) M10 1/ $d_{n,m}$ 200m
$f(d_{n,m})$				
Cut-off				
Ibase <sub>n,t-1</sub>	250.4*** (4.96)	991,193.7 (0.82)	0.0948*** (-12.55)	823,275.3*** (20.62)
Ibase <sub>n,t-1</sub> × Year <sub>2002</sub>	0.169 (-1.39)	8236.0 (0.39)	1.110 (0.37)	0.0210*** (-4.76)
Ibase <sub>n,t-1</sub> × Year <sub>2003</sub>	0.0293** (-2.79)	1.01486e + 13 (1.22)	1.427 (1.27)	0.00517*** (-7.18)
Ibase <sub>n,t-1</sub> × Year <sub>2004</sub>	0.00807*** (-4.09)	2245.8 (0.39)	1.383 (1.40)	0.00275*** (-8.37)
Ibase <sub>n,t-1</sub> × Year <sub>2005</sub>	0.00812*** (-4.20)	8.07e - 11 (-1.10)	1.218 (0.90)	0.000838*** (-10.44)
Ibase <sub>n,t-1</sub> × Year <sub>2006</sub>	0.00811*** (-4.21)	9.95e - 14 (-1.61)	1.032 (0.14)	0.0000790*** (-13.97)
Ibase <sub>n,t-1</sub> × Year <sub>2007</sub>	0.00386*** (-4.91)	0.000122 (-0.52)	1.342 (1.36)	0.0000494*** (-14.78)
Ibase <sub>n,t-1</sub> × Year <sub>2008</sub>	0.00361*** (-5.01)	1.76e - 09 (-1.16)	1.241 (1.04)	0.0000402*** (-15.20)
Ibase <sub>n,t-1</sub> × Year <sub>2009</sub>	0.00286*** (-5.24)	2.94e - 09 (-1.16)	1.215 (0.97)	0.0000210*** (-16.25)
Ibase <sub>n,t-1</sub> × Year <sub>2010</sub>	0.00307*** (-5.18)	2.07e - 09 (-1.18)	1.182 (0.85)	0.00000859*** (-17.63)
Ibase <sub>n,t-1</sub> × Rural <sub>j<sub>n</sub></sub>	0.197*** (-16.47)	0.125 (-0.43)	0.859 (-1.42)	0.0980*** (-31.30)
Ibase <sub>n,t-1</sub> × East <sub>i<sub>n</sub></sub>	4.135*** (3.85)	160556.8* (2.57)	0.867 (-1.41)	442.5*** (27.58)
Ibase <sub>n,t-1</sub> × Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>	12.86*** (3.35)	4.25e - 08 (-1.15)	1.719* (2.14)	22.67*** (6.92)
East <sub>i<sub>n</sub></sub>	1.412 (0.80)	0.523 (-1.47)	0.940 (-0.26)	2.731*** (3.35)
Rural <sub>j<sub>n</sub></sub>	1.305*** (63.55)	1.445*** (12.41)	0.957*** (-8.86)	1.028 (0.48)
Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>	0.841*** (-12.11)	0.752*** (-4.72)	0.992 (-0.74)	19.31*** (5.60)
Global Radiation <sub>o<sub>n</sub></sub>				1.052*** (114.43)
Global Radiation <sub>o<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub>				1.006*** (11.04)
Global Radiation <sub>o<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				0.976*** (-8.26)
Global Radiation <sub>o<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				0.966*** (-6.77)
Elevation <sub>p<sub>n</sub></sub>				1.036*** (36.13)
Elevation <sub>p<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub>				0.942*** (-37.88)
Elevation <sub>p<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				0.933*** (-14.70)
Elevation <sub>p<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				1.053*** (8.12)
Firm Density <sub>i<sub>n</sub></sub>				0.783*** (-17.42)
Firm Density <sub>i<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub>				0.614** (-2.85)
Firm Density <sub>i<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				0.945 (-1.14)
Firm Density <sub>i<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				2.516 (1.77)
Population Density <sub>i<sub>n</sub></sub>				0.869*** (-118.95)
Population Density <sub>i<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub>				0.815*** (-19.56)
Population Density <sub>i<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				1.080*** (24.50)
Population Density <sub>i<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				1.154*** (4.37)
Purchasing Power <sub>i<sub>n</sub></sub>				0.834*** (-159.99)
Purchasing Power <sub>i<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub>				1.002 (1.09)
Purchasing Power <sub>i<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				1.086*** (13.34)
Purchasing Power <sub>i<sub>n</sub></sub> × Rural <sub>j<sub>n</sub></sub> × East <sub>i<sub>n</sub></sub>				1.004 (0.32)
Observations	88,156,268	4,206,326	115,408,422	184,130,380
DF <sub>M</sub>	16	16	16	59
Final log-likelihood $\mathcal{L}$	-3,690,752	-71,922	-4,363,174	-4,928,134
LR: $\chi^2$ (DF)				138,618 (20)
LR: p-value				0
LR test against				M10-> M3
Year <sub>t</sub> fixed effects	No	No	No	Yes
Federal state <sub>n</sub> fixed effects	No	No	No	Yes
District <sub>i</sub> × Year <sub>t</sub> fixed effects	Yes	Yes	Yes	No
Choice maker <sub>n</sub> fixed effects	No	No	No	No

Exponentiated coefficients; robust  $t$  statistics in parentheses; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: The dependent variable is always  $y_{n,t}$ , i.e., choice maker  $n$ 's observed choice in terms of (newly) installing a PV system in  $t$ . We estimate specification M4<sub>≤30kW<sub>P</sub></sub>, M4<sub>>1MW<sub>P</sub></sub> and M4Placebo with the conditional logit estimator (see footnote 16). The sample is the same as in specification M1-M3 of Table 2. However, the conditional logit estimator drops all positive (or all negative) outcomes (here, in terms of District<sub>i</sub> × Year<sub>t</sub> groups). I.e., District<sub>i</sub> × Year<sub>t</sub> groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in less observations. For M10, we use the standard logit estimator. A small number of observations is dropped in comparison to M1-M3 as for these small number of observations data on some adopter characteristics are missing.

## D.4 Case study

Taking the four cities – Darmstadt, Karlsruhe, Marburg and Wiesbaden – together we end up with 338,242 observations (i.e., buildings and therefore potential places for PV installations). As not all PV systems fall inside a building-polygon we allocate 743 out of 2,345 to their nearest building. Of these 743, the median distance to the nearest building is 28 meter. Due to inaccuracy in geocoding and the possibility that more than one PV system could be installed on one building, 128 buildings have several PV systems installed. As we (also) assume one PV system per building in the case study, we allocate them to a random building in the same spatial unit. Table 3 shows that we have all in all 644 spatial districts for the four cities under study. Table 17 and Table 18 present the frequencies of the categories for the case study and the descriptive statistics.

Table 17: Frequency of choice alternatives for case study of four cities.

Year <sub>t</sub>	Alternative	Frequency in category
≤1999	0: No (or no new) PV system installed	338,217
	1: New PV system installed	25
2000	0: No (or no new) PV system installed	338,172
	1: New PV system installed	70
2001	0: No (or no new) PV system installed	338,128
	1: New PV system installed	114
2002	0: No (or no new) PV system installed	338,196
	1: PV system installed	46
2003	0: No (or no new) PV system installed	338,195
	1: New PV system installed	47
2004	0: No (or no new) PV system installed	338,169
	1: New PV system installed	73
2005	0: No (or no new) PV system installed	338,055
	1: New PV system installed	187
2006	0: No (or no new) PV system installed	338,043
	1: New PV system installed	199
2007	0: No (or no new) PV system installed	337,983
	1: PV system installed	259
2008	0: No (or no new) PV system installed	337,943
	1: New PV system installed	299
2009	0: No (or no new) PV system installed	337,784
	1: New PV system installed	458
2010	0: No (or no new) PV system installed	337,674
	1: New PV system installed	568
Sum of PV installations over all years		2,345

Table 18: Descriptive statistics specification MA, MB and MC.

	$f(d_{n,m})$	Mean	Std. Dev.	Min.	Max.
$y_{n,t}$		.00067	.026	0	1
$Ibase_{n,t}$	$1/d_{n,m}$	.0056	.015	0	.96
Global Radiation <sub><math>o_n</math></sub>		106	2.7	99	110
Global Radiation (Building-specific) <sub><math>n</math></sub>		60	50	0	130
N	3,382,420				

*Notes:* In the case study, the 3,382,420 observations come from 338,242 choice makers (buildings)  $n$  over 10 years  $t$  (2001-2010). The choice makers are distributed across 644 districts  $i$  in four cities.

## E SOEP

The socio-economic panel is a representative, yearly, longitudinal survey of German households and individuals. A choice maker  $n$ 's household is  $h$ . Some variables are available on the household level:  $Solar_{h,n,t}$  is one if household  $h$  adopts a new solar energy system in year  $t$  (otherwise zero).  $\ln(\text{Real household income}_{h,n,t})$  is the natural logarithm of household  $h$ 's real income in EUR in year  $t$ ,  $\text{Dwelling size}_{h,n,t}$  is household  $h$ 's housing unit size in sqm,  $\text{Home owner}_{h,n,t}$  is a dummy which is only one if the household owns the dwelling it lives in (in year  $t$ ), and  $\text{Garden}_{h,n,t}$  is a dummy which is only one if the household's dwelling has a garden (in year  $t$ ).

Other variables are on the choice maker level:  $\text{Working}_{n,t}$  is a dummy which is one if the choice maker's labor status is working.  $\text{Vocational training}_{n,t}$  is a dummy which is one if the choice maker holds a vocational degree. The dummy is set to one if the respondent states that she completed one of the following (zero otherwise): Lehre (Apprenticeship), Berufsfachschule, Gesundheitswesen (Vocational School), Schule Gesundheitswesen (Health Care School), Fachschule, Meister (Technical School), Beamtenausbildung (Civil Service Training), Sonstiger Abschluss (Other Training).  $\text{College}_{n,t}$  is a dummy which is one if the choice maker holds a college degree. The dummy is set to one if the respondent states that she completed one of the following (zero otherwise): Fachhochschule (Technical College), Universität, TH (University, Technical College), Hochschule im Ausland (College Not In Germany), Ingenieur-, Fachschule (Ost) (Engineering, Technical School (East)), Hochschule (Ost) (University (East)). Table 19 contains the descriptive statistics.

Table 19: Descriptive statistics specification MI and MII.

	Mean	Std. Dev.	Min.	Max.
$\Delta Solar_{h_n,t}$	.017	.13	0	1
$\ln(\text{Real household income}_{h_n,t})$	7.8	.59	3.7	12
Dwelling size $_{h_n,t}$	108	48	8	938
Working $_{n,t}$	.57	.5	0	1
Garden $_{h_n,t}$	.68	.47	0	1
Home owner $_{h_n,t}$	.55	.5	0	1
Vocational training $_{n,t}$	.67	.47	0	1
College $_{n,t}$	.23	.42	0	1
N	85,568			

*Notes:* 85,568 observations come from an unbalanced panel of choice makers  $n$  over 6 years  $t$  (2007-2012). The choice makers are distributed across 16 federal states.