

Rode, Johannes; Müller, Sven

Conference Paper

I spot, I adopt! Peer effects and visibility in solar photovoltaic system adoption of households.

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Rode, Johannes; Müller, Sven (2020) : I spot, I adopt! Peer effects and visibility in solar photovoltaic system adoption of households., Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/224644>

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.

I spot, I adopt!

Peer effects and visibility in solar photovoltaic system adoption of households.

Johannes Rode, Sven Müller*

March 1, 2020[†]

Abstract

We study variation of peer effects in rooftop photovoltaic adoption by households. Our investigation employs geocoded data on all potential adopters and on all grid-connected photovoltaic systems set up in Germany through 2010. We construct an individual measure of peer effects for each potential adopter. For identification, we exploit exogenous variation in two dimensions of photovoltaic system roof appropriateness of neighbors: their inclination and their orientation. Using discrete choice models with panel data, we find evidence for causal peer effects. However, the impact of one previously installed PV system on current adoption decreases over time. We also show that visible PV systems cause an increase in the odds of installing which is up to three times higher in comparison to all PV systems. At rural locations visibility may be less important, which indicates that word-of-mouth communication plays a stronger role.

Keywords: Causal peer effects, installed base, discrete choice, technology adoption and diffusion, solar photovoltaic panels, visibility

JEL classification: O33, C35, Q55, R10

*Corresponding author is Johannes Rode: Technische Universität Darmstadt, Chair of International Economics, Hochschulstrasse 1, 64289 Darmstadt, Germany, phone: 0049 6151 1657263, email: rode@vwl.tu-darmstadt.de; Sven Müller: Europa-Universität Viadrina, Department of Information and Operations Management, Große Scharrnstraße 59, 15230 Frankfurt (Oder), Germany, email: smueller@europa-uni.de

[†]This is a revised version of the working paper: “What drives peer effects in household technology adoption? Evidence from solar photovoltaic systems.” [Link to the current version.](#)

1 Introduction

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. Identifying factors driving the adoption decision is key to understanding the process of economic development and growth. Such knowledge helps organizations to foster diffusion, because new technologies often diffuse more slowly than desired (Rogers 1983; Hinz et al. 2011).

We contribute to the literature on technology adoption of consumers by analyzing 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. Germany has been among the countries with the most PV systems for several years, with total nominal PV power installed amounting to 46 gigawatt by the end of 2018 (Wirth 2019). This corresponds to the highest installed PV power per capita worldwide (IEA 2019).

The energy transition problem is a major challenge of the current century and PV technology has the potential to contribute noticeably to this process. PV can help to relieve energy dependency on fossil fuels (IPCC 2018). This is important from an ecological, economical, and political point of view. Organizations – firms and the state – active in the energy market are interested in understanding the PV adoption process.

Technology adoption is difficult to study empirically, because use of technology is seldom observed directly (Atkin et al. 2017). Adoption depends on the technology, but also on consumer behavior. Several studies reveal that peer effects influence consumers’ individual technology adoption decisions in general (Conley and Udry 2010; Manchanda, Y. Xie, and Youn 2008; Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011; Risselada, Verhoef, and Bijmolt 2014) and PV adoption in particular (Bollinger and Gillingham 2012). Peers are – in this context – considered as proximate adopters in preceding periods. Bollinger and Gillingham’s (2012) seminal paper shows that peer effects drive PV adoption. Their study and most others on (PV) adoption (Graziano and Gillingham 2015; Rode and Weber 2016) aggregate data to regions or local spatial units. Aggregate data does not allow us to build a specific measure of the peer effect

for individual adopters and may yield biased results (Albuquerque, Bronnenberg, and Corbett 2007; Angst et al. 2010). Our setting helps to better understand the technology adoption process, because we have access to exact locational panel data on all photovoltaic adopters (almost 800,000) and potential adopters (some 18 million). We determine an individual measure for peer effects for each potential PV adopter per year on the very local level. Since the decision to adopt in a certain period is discrete (Karshenas and Stoneman 1992), we employ a discrete choice approach (McFadden 2001) to analyze PV adoption (Section 2).

The purpose of this paper is, first, to identify causal peer effects in PV adoption, second, to test for spatio-temporal variation, and, third, to study visibility as a channel of peer effects. Identifying peer effects comes with well-known challenges. In a series of analyses, we use high-dimensional multi-level fixed effects to control for (local) unobserved heterogeneity, adoption shocks (on the net present value of PV systems, for example) and the tendency that neighbors may be alike and may have similar preferences for adoption (homophily). We also consider the reflection problem (Manski 1993). It refers to situations in which the adoption decision of an individual depends on others in her reference group and vice versa. The results provide strong evidence for the robustness of our analyses.

We conduct a Germany-wide analysis and a case study for the federal state Baden-Württemberg, with some 3 million potential adopters for whom we have even more detailed data (see Section 2). The Germany-wide analysis (Section 3.1) reveals that peer effects in PV system adoption are very localized. 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. To be precise, a single prior PV system within 200 meters distance increases the odds of installing a PV system by 9% (in rural areas in the west) to 26% (in rural areas in the east) in 2001. Interestingly, the odds decline to 16% (in rural areas in the east) and below in 2010.

One concern is that endogeneity may bias our estimates. To confidently argue that our estimates reflect a causal peer effect, we exploit exogenous source of variation in the installed base. In our case study of Baden-Württemberg, we utilize exogenous variation in

two dimensions of photovoltaic system roof appropriateness: inclination and orientation of roofs from previous PV adopters nearby. We use both to conduct an instrumental variable (IV) estimation (Section 3.2). The roof inclination and orientation of previous PV adopters nearby should be exogenous to the adoption decision of a given potential adopter since roofs were built in nearly all cases before PV system diffusion reached relevant levels: the houses were constructed without considering a PV system. Further, we show that the inclination and the orientation of the roof of a given adopter are not correlated to those of neighboring roofs. Nevertheless, we include fixed effects on the street-level, which controls for potentially similar PV appropriateness (or potentially similar inclination and orientation) of roofs in the same street. The instruments are strong in the first stage regression. An overidentifying restrictions test supports their validity. Our instrumental variable estimation indicates that peer effects indeed drive PV adoption.

We study visibility as one potential channel of peer effects.¹ We exploit the fact that PV systems are special in the sense that they may or may not be directly seen from a street or buildings nearby. We are able to identify different peer effects for PV systems nearby a street and those not nearby a street. We also construct a measure for peer effects considering PV systems directly visible from a given building or not. As such, we disentangle the impact of visibility and, as alternative hypotheses, for instance, proximity or word-of-mouth on peer effects (Chen, Wang, and J. Xie 2011). Our measures for visible PV systems indicate an up to three times higher increase in the odds of installing than our measures with all PV systems. Further, our results reveal that peer effects rely less on visibility in rural areas. Neighbors in rural areas may be more closely related to each other than in non-rural areas. In consequence, alternative channels such as word-of-mouth may be more important for peer effects in PV adoption in rural areas whereas in urban areas visibility seems to be dominant.

Our empirical findings give new managerial insights to achieve adoption targets that are also valuable for other markets, such as novel foods or electric vehicles (Section 4). Firstly, sales strategies should vary over time: in the early stages of diffusion seed instal-

¹Bollinger and Gillingham (2012) also discuss the visibility of PV systems but only exploit the size of a system. Larger systems are not necessarily better visible. We therefore use exact locational individual level data to identify whether a PV system is visible from a street and buildings.

lations – in particular, if visible – are promising. In later stages salespersons might be more likely to achieve a given adoption target. Secondly, sales strategies should also vary in space: While visibility seems to be sufficient in urban areas, more effort needs to be spent in rural areas (sales events, salespersons etc). Also, regions which are less developed (such as the east of Germany) are more sensitive to peer effects. This finding points to the potential of peripheral regions to contribute to the transition of the energy system, for instance due to PV adoption. Further, policy makers and firms may support peer effects to overcome a slowdown in adoption, which could for instance be induced by the phase out of a subsidy system.

2 Data and Modeling Approach

In Appendix A, we underpin our modeling approach with random utility theory and outline the corresponding econometric model specification. We conduct a Germany-wide analysis and a case study on a subset, using data from Baden-Württemberg.

Germany-wide analysis. – We consider 18,413,514 buildings n in Germany in 2009 and 769,525 PV systems, each allocated to a mutually exclusive building. We study the years $t = 2001, \dots, 2010$ since very few systems were set up before (see Figure 4 and Figure 5 in Appendix A.2). Figure 1 visualizes the spatio-temporal dimension of the choice variable $y_{n,t}$ by a small scale example. $y_{n,t}$ defines whether a building n is equipped with a PV system in year t or not.

To analyze the impact of peers on the decision to adopt we use the installed base (Farrell and Saloner 1986). The lagged installed base ($IB_{n,t-1}$) measures the number of preexisting PV installations within a radius of 200m around the location of a choice maker. As an alternative measure, the lagged installed base $\left(IB_{n,t-1}^{1/d} \right)$ measures the number of preexisting PV installations within a radius of 200m around the location of a choice maker, with the importance of each location weighted by the inverse distance.

Figure 4 in Appendix A.1 illustrates the spatio-temporal dimension of our data. The figure shows that there are differences in PV density in the east and in the west, which may be due to a general economic development lag in the east (Redding and Sturm

Figure 1: Excerpt of the PV adoption process for Germany-wide analysis.

Notes: Hollow circles represent potential adopters. Filled circles are PV installations. The color indicates the year of installation, i.e., the first year when we observe a PV system. Older PV installations are colored in yellow, more recent PV installations in blue.

2008). There are also differences in PV density in urban and rural areas. Since more previous adopters may reduce uncertainty for potential new adopters, valuable experience from previous adopters may be less relevant during later stages of the diffusion path. In consequence, we presume that the peer effect might be spatially non-stationary between east and west and rural and non-rural locations. We describe our data sources and the level of detail of the data in Appendix A.2. There, we also provide formal definitions and descriptive statistics.

Case Study Baden-Württemberg. – Detailed data on Baden-Württemberg allows us to identify causal peer effects and to build installed base measures that take into account the visibility of a PV system. This analysis is based on a sub-sample consisting of 71,432 household PV systems installed in Baden-Württemberg through 2010.

To identify causal peer effects, we need exogenous variation correlated with our measure of the installed base but not driving the actual adoption decisions of a given adopter. We employ high-resolution data on the appropriateness of roofs for PV systems in the federal state Baden-Württemberg. Figures 8 and 9 in Appendix A.3 illustrate the high-resolution data. The spatio-temporal lag variable *Inclination of neighboring roofs* $_{n,t-1}$

measures the number of preexisting PV installations within radius 200m around the location of building n , with the importance of each location weighted by an exogenous measure for the inclination of a roof on a linear scale between 0 and 100%. The electricity generation potential is best for an inclination of 37° , which corresponds to 100%, whereas the maximum value of 74° (in the data set) corresponds to 0 and a linear scale in-between.² Similarly, the spatio-temporal lag variable *Orientation of neighboring roofs* $_{n,t-1}$ measures the number of preexisting PV installations within radius 200m around the location of building n , with the importance of each location weighted by an exogenous measure for the orientation of a roof on a linear scale between 0 and 100%. A perfect south orientation corresponds to 100% whereas an orientation to the north corresponds to 0 and a linear scale in-between.³ Both instruments do not account for the roof area feasible for PV since we do not want to weight larger roofs more strongly: larger roofs may be associated with confounding factors such as higher incomes or levels of education which may be related to the PV adoption decision. In an instrumental variable estimation, we employ *Inclination of neighboring roofs* $_{n,t-1}$ and *Orientation of neighboring roofs* $_{n,t-1}$ as exogenous variables to predict the installed base and then, in a second step, use the predicted $\widehat{IB}_{n,t-1}$ to identify causal peer effects in PV adoption.

To analyze the impact of visibility on peer effects, we merge the roof-specific information on a building (orientation and inclination) with OpenStreetMap data on the course of roads. This enables us to include only PV systems which are within 50m distance to the next road ($IB_{n,t-1}^{NR50}$) since these are likely to be visible from this road. We assume that PV systems located farther away than 50 meters from the closest street are less likely to be visible. We do not account for obstacles like trees etc. In Figure 2, we illustrate $IB_{n,t-1}^{NR50}$ in green while the large red point indicates that the PV system (in yellow) with the green next road 50m corridor area would be visible from the street.

$IB_{n,t-1}^{NR50}$ includes the PV systems which are not located within 50m distance to the next road. They are less likely to be visible.

²Flat roofs also have a 100% inclination value since perfectly inclined PV systems (on stilts) can be installed there. The minimum inclination for non-flat roofs is 6° which corresponds to an inclination value of 6/37% and a linear scale between the optimal inclination of 100%.

³Flat roofs also have a 100% orientation value since perfectly oriented PV systems (on stilts) may be installed there.



Figure 2: Visibility measures of peer effects.

Notes: The map shows the construction principle of our installed base visibility measures $IB_{n,t-1}^{NR50}$, $IB_{n,t-1}^{NR50}$, $IB_{n,t-1}^{\triangleleft}$ and $IB_{n,t-1}^{\triangleleft}$. If the next road point lies within the green next road 50m corridor area, the PV system is included in the installed base measure $IB_{n,t-1}^{NR50}$. The example (in green) shows a next road point (in large, red) that is considered in the measure. $IB_{n,t-1}^{NR50}$ includes the remaining PV systems, i.e., the ones which are not located within the green next road 50m corridor area. The example in orange illustrates an installed base measure for which no building lies in between the choice maker's location and a (yellow) PV system within 50m distance (for which the south roof should be visible for the choice maker). In consequence, the PV system is directly visible from the building for which we calculate the installed base in that case ($IB_{n,t-1}^{\triangleleft}$). In blue, we see an example for an installed base measure for which a building lies in between the potential adopter and the (yellow) PV system nearby. Therefore, the PV system is not directly visible from the potential adopter's building, for which we calculate the installed base ($IB_{n,t-1}^{\triangleleft}$). The example is based on data from 2010.

The orientation of roof-top PV systems of households in Germany is mainly to the south. We use the orientation to construct two additional measures of visibility for the installed base. We also build installed base measures based on direct visibility of PV systems from one building to another. $IB_{n,t-1}^{\triangleleft}$ includes the PV systems directly visible to a choice-maker (see orange corridor area in Figure 2) whereas $IB_{n,t-1}^{\triangleleft}$ refers to the PV systems which are not directly visible because another building lies in between (see blue corridor area in Figure 2). In Appendix A.3, we show summary statistics, provide formal definitions and describe our data sources.

Modeling approach. – We consider the owner(s) of a building to make the choice to install a PV system in a certain period or not. As such we denote observation n as the

choice maker. We use random utility theory – as described in Appendix A.4 – to model the adoption decision, i.e., choice maker n ’s observed choice in terms of (newly) installing a PV system in t : $\Delta y_{n,t}$. We run logit and conditional logit regressions. We also make use of the linear probability model to include high-dimensional multi-level fixed effects. Because a choice maker adopts at most once within the considered time span, we censor observations. In our case study for Baden-Württemberg, we have choice maker n -specific information available. There, we control for choice maker n -specific potential electricity production per square meter (sqm) and maximum module area. PV installations got cheaper during our period of study and the subsidy system changed. Both may have an time-variant influence on adoption. We therefore also control for choice maker n -specific potential electricity production per sqm and maximum module area interacted with a time trend.

3 Results and Discussion

We first discuss our baseline findings for the Germany-wide analysis and some robustness checks (Section 3.1). Then, we turn to our case study of Baden-Württemberg to identify causal effects in the adoption process (Section 3.2). Finally, we study a potential channel for peer effects in PV adoption: visibility (Section 3.3).

3.1 Spatio-Temporal Variation of Peer Effects across Germany

Baseline results. – Column (1) in Table 1 illustrates the estimated (logit) odds ratios with temporal fixed effects and federal state fixed effects. Temporal fixed effects absorb time specific adoption shocks that could for instance be caused by changes in the subsidy system. Federal state fixed effects cover federal state specific effects, e.g. differences in sun intensity or additional subsidies in a federal state. The estimates reveal 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. This finding provides evidence for localized peer effects in the adoption of PV systems.

Table 1: Estimates for peer effects and their spatio-temporal variation in Germany with different types of fixed effects. Column (1) shows odds ratios from logit, columns (2, 4-5) odds ratios from conditional logit and column (3) coefficients from OLS.

	Year FE	District-Year FE	District-Year & Choice maker FE	District-Year FE	Choice maker FE
	(1) CLogit	(2) CLogit	(3) OLS	(4) CLogit	(5) CLogit
	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$y_{n,t}$
<i>Installed base:</i>					
$IB_{n,t-1}$	1.09** (110.6)	1.01** (4.55)	0.00016** (8.90)	1.14** (3.87)	1.16** (9.90)
$IB_{n,t-1} \times \text{Year}_{2002}$				0.97 (-0.84)	0.91** (-5.72)
$IB_{n,t-1} \times \text{Year}_{2003}$				0.89** (-3.02)	0.88** (-8.23)
$IB_{n,t-1} \times \text{Year}_{2004}$				0.88** (-3.77)	0.84** (-11.1)
$IB_{n,t-1} \times \text{Year}_{2005}$				0.90** (-2.98)	0.83** (-12.5)
$IB_{n,t-1} \times \text{Year}_{2006}$				0.92* (-2.37)	0.80** (-14.5)
$IB_{n,t-1} \times \text{Year}_{2007}$				0.90** (-2.99)	0.79** (-15.7)
$IB_{n,t-1} \times \text{Year}_{2008}$				0.90** (-2.97)	0.80** (-15.3)
$IB_{n,t-1} \times \text{Year}_{2009}$				0.90** (-2.95)	0.80** (-15.4)
$IB_{n,t-1} \times \text{Year}_{2010}$				0.90** (-2.94)	0.80** (-15.4)
$IB_{n,t-1} \times \text{East}_{i_n}$				1.06** (5.52)	0.97** (-3.01)
$IB_{n,t-1} \times \text{Rural}_{j_n}$				0.96** (-18.1)	0.97** (-10.2)
$IB_{n,t-1} \times \text{Rural}_{j_n} \times \text{East}_{i_n}$				1.12** (4.84)	1.02 (0.75)
<i>Controls:</i>					
Rural_{j_n}				1.51** (79.2)	
$\text{Rural}_{j_n} \times \text{East}_{i_n}$				0.75** (-15.1)	
East_{i_n}				1.43 (0.69)	
Observations	182,750,649	74,997,032	182,727,991	74,997,032	5,765,940
DF _M	25	1	1	16	22
Final log-likelihood	-3,347,984	-2,618,881		-2,614,787	-1,180,134
Adj. R^2			0.07		
Year _t fixed effects	Yes	Yes	Yes	No	No
Federal state _n fixed effects	Yes	Yes	Yes	No	No
District _i \times Year _t fixed effects	No	Yes	Yes	Yes	Yes
Choice maker _n fixed effects	No	No	Yes	No	No

Notes: Exponentiated coefficients (odds ratios) in columns (1-2) and (4-5) and coefficients in column (3); robust t statistics in parentheses (in column (3), t statistics are based on SE clustered on the district-level); * $p < 0.05$, ** $p < 0.01$

The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . Column (1) is estimated via logit. The panel of 18,413,514 choice makers over 10 years t (2001-2010) results in 184,135,140 observations. We censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 182,750,649 observations. We estimate column (1) with logit and columns (2,4-5) with the conditional logit estimator. The sample is the same across columns. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District_i \times Year_t groups for column (2) and (4) (choice makers for column (5)). I.e., for column (2) and (4) District_i \times Year_t groups with no adoption (or if all choice makers in that group adopt at once) are dropped. For column (5) choice makers with no adoption are dropped. Similarly, singleton observations are dropped in column (3). These procedures result in fewer observations. We estimate but do not show the constant for column (1). $f(d_{n,m}) = 1$ and the Cut-off distance D is 200m.

District \times year fixed effects $\eta_{i_n,t}$. – To measure peer effects more confidently, Column (2) in Table 1 illustrates that our estimates are robust to controlling for detailed spatial fixed effects. We consider 77,847 districts i .⁴ This approach takes time-variant

⁴The average area of a statistical district i is about 4.6 sqkm and the average number of buildings in such a statistical district i is 237. For instance, the largest German city Berlin (3.6 million inhabitants in 2010) consists of 3886 districts i .

(and invariant, e.g. caused by demographics) adoption shocks on the district_{*i*} level into account.⁵ Such a shock could be 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 along the diffusion path. The fixed effects also control for homophily at the district_{*i*} level. Homophily refers to the tendency that neighboring adopters share characteristics and preferences.

*District_{*i*} × year_{*t*} η_{*i,n,t*} and choice maker_{*n*} fixed effects α_{*n*}.* – We estimate column (3) in Table 1 with the linear probability model since it allows us to include two levels of high-dimensional fixed effects. Our results remain unaffected from also considering choice maker fixed effects. This approach also takes time-invariant individual effects into account. These could be caused by individual-specific characteristics as income or the suitability of a roof for a PV systems, e.g. determined by roof inclination, roof orientation or roof size.

In column (4), we focus on odds ratios from conditional logit again. This procedure does not allow us to include several levels of high-dimensional fixed-effects. According to column (4) with District_{*i*} × year_{*t*} fixed effects (Table 1), the peer effect varies across space. The odds ratio of installing in rural parts in the west is lower than in non-rural parts. The interaction of the installed base with East_{*n*} indicates that the peer effect is even larger for choice makers in the east of Germany. Since the fraction of buildings with PV is lower in the east (see Figure 5 in Appendix A.2), the east may still be in a very early stage of the 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, information from peers might be more important in the east compared to the west of Germany. Similarly, the interaction between $IB_{n,t-1}$ 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.

⁵A district_{*i*} fixed effect for every year of study, i.e. $77,847 \times 10 = 778,470$ fixed effects. District_{*i*} × Year_{*t*} fixed effects also control for spatial differences at higher levels of aggregation: e.g., additional time-variant incentives to install PV in a specific federal state. See Greene (2012, pp. 721-724) for details on the conditional logit estimator.

We now focus on the time-varying coefficients of the installed base in Column (4). An obvious interpretation of the level of our results is in terms of a one unit increase in $IB_{n,t-1}$, which roughly corresponds to a one standard deviation increase and is equal to one previously installed PV system within 200m distance. The estimates from column (4) in Table 1 reveal that for a one unit increase in the installed base, we expect the odds of installing a PV system (in non-rural areas in the west in 2001) to increase by 14%. The odds of installing in year 2002 does not statistically vary compared to 2001: it is insignificant to all relevant levels. However, a one unit increase in the installed base increases the odds of installing a PV system for instance in 2006 by $(1.14 \times 0.92 - 1) \times 100\% = 5\%$.

Figure 3 illustrates the diminishing odds of installing from a one unit $IB_{n,t-1}$ increase over time. The figure also illustrates differences between rural vs. non-rural and east vs. west. In consequence, peer effects are more relevant during the very early phase of technology diffusion. A decrease in odds – in western rural areas from 2003 on (excluding 2006), for example – provides evidence for a saturation effect. Our study reveals that PV diffusion has matured in western Germany from the mid 2000's on. The figure also indicates a break in the diminishing odds of installing from increases in $IB_{n,t-1}$ 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. This effect seems to last until 2006, when odds decline again until 2007 and from then on stay constant. The break in 2004 indicates that a change in the subsidy system may have increased uncertainty for potential adopters. In consequence, the importance of peers may have grown again.

Rode and 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 and Gillingham's (2012, p. 905) analysis indicates that the peer effect may increase over time. Possibly, reasons for PV adoption largely vary in Bollinger and Gillingham's sample from California and ours from Germany, where feeding PV electricity into the grid is highly subsidized: Bollinger and Gillingham mention that marketing efforts leveraged peer effects in the later periods of their study in California.

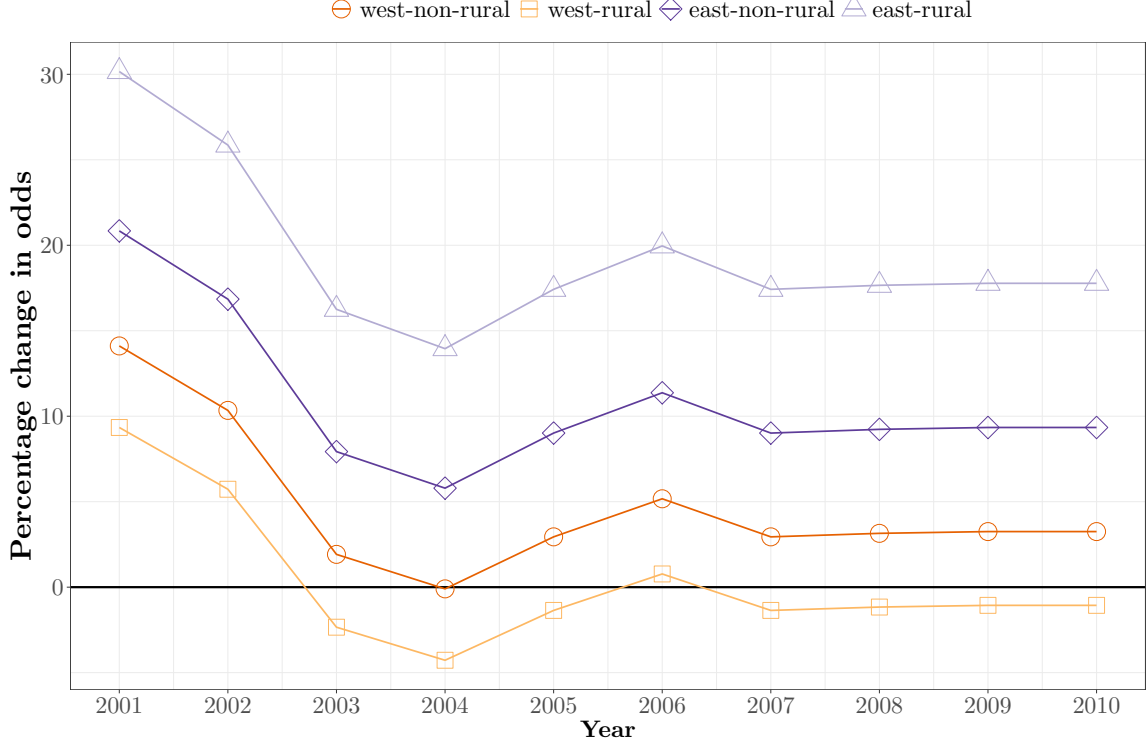


Figure 3: Spatio-temporal variation in peer effects.

Notes: The graph displays the increase in odds of installing a PV system in certain periods given an increase in the installed base by one system (for the specification of column (4) in Table 1). For example, an additional existing PV system yields an increase in the odds of installing a PV system by $(1.14 \times 0.92 \times 1.06 - 1) \times 100\% = 11.2\%$ in eastern non-rural areas in Germany in 2006.

Robustness. – We conduct several robustness tests. In Appendix C.1 we give details. First, we illustrate that our estimates are robust to a distance-weighted measure of the installed base. As the results for distance-weighted and non-distance-weighted measures of the installed base are in line, it is unlikely that proximity is the real channel driving of our results. Second, when focusing on industrial PV systems we do not find evidence for peer effects. This makes sense as industrial investors should search for the best location across a large region or even the whole country. Third, we randomly allocate the same number of PV installations, which were in fact installed in Germany per year, to existing buildings. This procedure confirms that our estimations do not find a positive peer effect in PV system adoption by definition. Fourth, we illustrate that incorporating PV installations from the last year instead of all previous installations in the installed base measure does not affect our core findings. The estimates indicate that last year’s PV installations are more relevant for the peer effect in comparison to the ones installed a longer time

ago. This finding is in line with Graziano and Gillingham’s (2015, p. 19) evidence for a “*diminishing neighbor effect over time since prior installations*”. Fifth, our findings remain unaffected by lagging our peer effect measure by two years (Rode and Weber 2016). 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. 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. Sixth, we discuss characteristics of district i with many PV systems. Seventh, we also illuminate adopter characteristics by exploiting the household panel survey Socio-Economic Panel. Eighth, we show that our results are robust to different cut-off radii and different functional forms for our distance weights.

3.2 Causality: Case Study of Baden-Württemberg

Baseline results. – We take a closer look at the federal state Baden-Württemberg because building-specific data on PV appropriateness is available. For convenience reasons, we estimate a time-invariant coefficient for the peer effect. In Table 2, we illustrate that our results for Baden-Württemberg are in the ballpark of those for Germany. Column (1) confirms a positive significant peer effect in PV adoption when we include $\text{District}_i \times \text{Year}_t$ fixed effects. They are robust to even more detailed spatial fixed effects: $\text{District}_i \times \text{Street}_k \times \text{Year}_t$ fixed effects (see column (2)).⁶ Also, the results are unaffected when using a distance-weighted installed base measure, see column (3). Column (4) illustrates that in Baden-Württemberg there is no evidence for a statistically differing peer effect in rural and non-rural areas. Robustness checks, which are not shown, confirm that the peer effect importance is decreasing over time, as in our Germany-wide analysis.

Instrumental variables. – Of course, the association between PV adoption and the installed base could be driven by some omitted variable that is the real driver of installing

⁶There are 544,243 streets in our data set for Baden-Württemberg and 8988 district i . As some streets only lie in one district i and as we include a fixed effect for every year of study, we end up with 5,785,981 $\text{District}_i \times \text{Street}_k \times \text{Year}_t$ fixed effects. The average number of buildings in such a $\text{District}_i \times \text{Street}_k$ is 5.

a PV system and that is correlated with the installed base, beyond our controls. To confidently argue that our estimates reflect a causal peer effect, we need some exogenous source of variation in the installed base. That is, variation in the installed base that is driven by factors that do not directly affect the adoption of a given potential installer or that are not correlated with factors that may drive adoption, other than the installed base.

Table 2: Odds ratios from conditional logit estimates for peer effects for Case Study of Baden-Wuerttemberg.

	Baseline	Street FE	Distance-weighted	Urban vs. rural
	(1)	(2)	(3)	(4)
	CLogit	CLogit	CLogit	CLogit
	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$y_{n,t}$
<i>Installed base:</i>				
$IB_{n,t-1}$	1.02** (8.98)	1.02** (4.87)		1.02** (4.34)
$IB_{n,t-1} \times \text{Rural}_{j_n}$				1.01 (1.64)
<i>Installed base (distance-weighted):</i>				
$IB_{n,t-1}^{1/d}$			1.79* (2.22)	
<i>Controls:</i>				
Potential electricity production per sqm _n	1.6E+9** (16.4)	5.2E+7** (10.8)	5.2E+7** (10.8)	5.2E+7** (10.8)
Potential electricity production per sqm _n $\times t$	1.24 (1.44)	0.90 (-0.56)	0.90 (-0.56)	0.90 (-0.57)
Maximum module area _n	1.14** (7.72)	1.49** (2.73)	1.48** (2.72)	1.49** (2.73)
Maximum module area _n $\times t$	0.98** (-5.93)	0.98 (-1.15)	0.98 (-1.14)	0.98 (-1.15)
Rural _{j_n}	1.18** (14.4)	1.03 (1.14)	1.02 (0.75)	1.00 (-0.058)
Observations	14,571,227	770,063	770,063	770,063
DF _M	6	6	6	7
Final log-likelihood	-385,919	-142,383	-142,392	-142,381
Year _t fixed effects	Yes	Yes	Yes	Yes
Federal state _n fixed effects	Yes	Yes	Yes	Yes
District _i \times Year _t fixed effects	Yes	Yes	Yes	Yes
District _i \times Street _k \times Year _t fixed effects	No	Yes	Yes	Yes
Choice maker _n fixed effects	No	No	No	No

Notes: Exponentiated coefficients (odds ratios); robust t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . All columns are estimated via logit. The panel of 2,903,730 choice makers over 10 years t (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. We estimate all columns with the conditional logit estimator. The sample is always the same. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District_i \times Year_t groups for columns (1) (3) (District_i \times Street_k \times Year_t groups for column (4)). I.e., for columns (1)-(3) District_i \times Year_t groups with no adoption (or if all choice makers in that group adopt at once) are dropped. For column (4), District_i \times Street_k \times Year_t groups with no adoption (or if all choice makers in that group adopt at once) are dropped. These procedures result in fewer observations. In columns (1), (3) and (4), $f(d_{n,m}) = 1$, and in column (2), $f(d_{n,m}) = 1/d_{n,m}$. The Cut-off distance D is 200m.

The inclination and orientation of roofs for PV from previous PV adopters nearby are good candidates. First, a neighbor's roof inclination and roof orientation are an important driver of the neighbors adoption of a PV system since they impact the amount of electricity the neighbor can produce with a PV system. Second, both measures are exogenous to the adoption decision of a given individual because the inclination and orientation of roofs were determined in nearly all cases before PV system diffusion reached

relevant levels: nearly all houses were constructed without considering a PV system.⁷ A critical reader may still expect that the inclination and orientation of neighboring roofs are alike. However, this is not the case in our study. First, we show that streets are often not straight in Germany and that the PV appropriateness of roofs, which is determined by their inclination and orientation, varies within a street (Figure 8 in Appendix A.3). Second, we include $\text{District}_i \times \text{Street}_k \times \text{Year}_t$ fixed effects in our regressions. We thereby control for time-variant adoption shocks on the $\text{District}_i \times \text{Street}_k \times \text{Year}_t$ level. These also control for similar PV appropriateness (or potentially similar inclination and orientation) of roofs in the same street. Third, we calculate the correlation of our instruments on the $\text{District}_i \times \text{Street}_k$ level with the the same measures for a given roof. The correlation of roof inclination measures for a given roof with a PV system and inclination measures for all other roofs in the same street is -0.27. Similarly, the correlation of our roof orientation measure for a given roof with a PV system and the roof orientation measure for all other roofs in the same street is -0.22. We conclude that the inclination and orientation of roofs for PV from previous PV adopters nearby are likely to be exogenous to the PV adoption decision of a potential adopter. We therefore use the inclination and orientation of roofs from previous PV adopters nearby to instrument for the installed base.

Causal estimates. – We include $\text{District}_i \times \text{Street}_k \times \text{Year}_t$ fixed effects in the estimations that follow. In column (2) of Table 3, we illustrate the results for the ordinary least squares first stage regression. In this case, we regress the lagged installed base on the lagged roof inclination and on the lagged roof orientation of a given choice maker n 's neighbors. Both instruments are strong and highly significant to all relevant statistical levels. Note that column (2) of Table 3 shows the estimated coefficients as we use ordinary least squares here (because the installed base is a continuous variable). As we would expect, we find a positive association of both instruments on the installed base. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. We

⁷The number of new buildings is very low. E.g., considering Baden-Württemberg the average yearly increase of residential buildings was 0.73% between 31.12.1999 and 31.12.2010. As described in Appendix A.3, the information for roof inclination and orientation comes from laser scans that were conducted between 2000 and 2005. I.e., we do not include newly constructed buildings for which a PV system may have been considered during construction our analysis. Also, see footnote 14 and the related text in Appendix A.3.

now turn to the second stage regression in column (3), which also comes from ordinary least squares (the linear probability model). In the second stage regression, we use the predicted installed base from the first stage. The effect of the instrumented installed base on the adoption decision is strong and significant at all confidence levels. Comparing the OLS specification in column (1) with the instrumented specification in column (3) shows that they are in line. They are also in the range of the results for Germany (column (3) in Table 1).

We conduct an overidentifying restrictions test. As both instruments are strongly significant in the first stage regression, the overidentifying restriction test will have power. The p-value for the Hansen statistic is 0.92. The test indicates that our instruments are uncorrelated with the error term at all significance levels. Therefore, the instruments pass the overidentifying restriction test, which supports the validity of our instruments: i.e., all the effect the inclination and orientation that neighboring roofs have on the individual adoption decision of an individual is through the peer effect.

In column (4), we use the instrumented installed base from column (2) in a second stage estimation with conditional logit. We observe that the results are also in the ballpark of the ones for Germany. We find a positive and significant odds ratio of 1.03 for the installed base. In consequence, we may have, if at all, underestimated the peer effect when studying Germany.

Robustness. – We conduct several robustness tests in Appendix D. We give a short summary here. First, we show that our results are robust to adding choice maker_n fixed effects α_n (while still considering District_i × Street_k × Year_t fixed effects). See Table 18 column (1). This approach also controls for time-invariant individual effects, which could be caused by individual-specific characteristics such as income. Second, our results also hold when instrumenting the installed base with our measures for the inclination and orientation of neighboring roofs with choice maker_n fixed effects (columns (2-3)). The instruments are still strong and the overidentifying restrictions test confirms their validity. Third, our results remain unaffected when using a distance-weighted installed base (columns (4-5)). The distance-weighted installed base is instrumented with distance-

Table 3: OLS and instrument variable estimates for peer effects for Case Study of Baden-Wuerttemberg.

	OLS	IV		
		1st stage (OLS)	2nd stage (OLS)	2nd stage (CLogit)
	(1) $\Delta y_{n,t}$	(2) $IB_{n,t-1}$	(3) $\Delta y_{n,t}$	(4) $\Delta y_{n,t}$
<i>Installed base:</i>				
$IB_{n,t-1}$	0.00010** (5.01)			
<i>Predicted installed base:</i>				
$\hat{IB}_{n,t-1}$			0.00013** (5.94)	1.03** (5.78)
<i>Instruments:</i>				
Inclination of neighboring roofs $_{n,t-1}$		0.46** (121.0)		
Orientation of neighboring roofs $_{n,t-1}$		0.77** (221.3)		
<i>Controls:</i>				
Potential electricity production per sqm $_n$	-0.015** (-7.08)	0.10** (7.29)	-0.015** (-7.08)	5.2E+7** (10.8)
Potential electricity production per sqm $_n \times t$	0.0082** (20.1)	-0.029** (-7.92)	0.0082** (20.1)	0.90 (-0.57)
Maximum module area $_n$	0.000094 (0.68)	0.0051** (7.49)	0.000093 (0.67)	1.49** (2.73)
Maximum module area $_n \times t$	0.000044 (1.78)	-0.0017** (-9.06)	0.000044 (1.79)	0.98 (-1.15)
Rural $_{j_n}$	0.000069 (1.04)	-0.037** (-28.8)	0.000079 (1.18)	1.03 (1.25)
Observations	27,646,497	27,646,497	27,646,497	770,063
DF _M	6	7	6	6
Final log-likelihood				-142,378
Adj. R^2	0.02		0.00	
F	208.7		210.3	
Hansen J (p-value)			0.0 (0.92)	
Kleibergen-Paap rk Wald F statistic			682,095.5	
Year $_t$ fixed effects	Yes	Yes	Yes	Yes
Federal state $_n$ fixed effects	Yes	Yes	Yes	Yes
District $_i \times$ Year $_t$ fixed effects	Yes	Yes	Yes	Yes
District $_i \times$ Street $_k \times$ Year $_t$ fixed effects	Yes	Yes	Yes	Yes
Choice maker $_n$ fixed effects	No	No	No	No

Notes: Columns (1)-(3) show coefficients, column (4) shows exponentiated coefficients (odds ratios); t statistics in parentheses (based on SE clustered at street-level; in column (4), we show robust t statistics); * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . Columns (1)-(3) are estimated via OLS. Column (3) uses the predicted values for the installed base from the first stage in column (2). Column (4) is estimated with the conditional logit estimator. Column (4) uses the predicted values for the installed base from the first stage in column (2). The panel of 2,903,730 choice makers over 10 years t (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. Across columns, the sample is always the same. However, the conditional logit estimator (column (4)) drops all positive (or all negative) outcomes in terms of District $_i \times$ Year $_t$ groups. I.e., for column (4) District $_i \times$ Street $_k \times$ Year $_t$ groups with no adoption (or if all choice makers in that group adopt at once) are dropped. Similarly, singleton observations are dropped in columns (1)-(3). These procedures result in fewer observations. $f(d_{n,m}) = 1$ and the Cut-off distance D is 200m.

weighted measures for the inclination and orientation of neighboring roofs. Again, these instruments are strong and the overidentifying restrictions test confirms their validity. Forth, we can modify our sample. The case study results shown so far are based only on those adoption decisions ($\Delta y_{n,t}$) for which we can be very sure that allocation of the PV system is to the correct roof. We can however also include those PV systems for which no unique allocation was possible. They are then randomly assigned to another roof in the same District $_i$. Table 20 illustrates that our results remain unaffected when analyzing the full sample. Fifth, we show that our results remain unaffected from lagging our installed base measure by two years in Table 22. We can therefore confidently rule out that the ‘reflection problem’ described by Manski (1993) biases our results.

3.3 Are Peer Effects Driven by Visibility?

Visibility from roads. – Given that peer effects indeed drive PV adoption, we are curious about their underlying mechanism. The measures $IB_{n,t-1}^{NR50}$ and $IB_{n,t-1}^{NR50,1/d}$ allow us to investigate the impact of the visibility of a PV system from the closest street on the odds of installing a new PV system nearby. The measure $IB_{n,t-1}^{NR50}$ comprises only PV systems which are located at most 50 meters away from the closest street. First, we compare the levels of the odds ratios in column (1) of Table 4 to those in column (2) of Table 2. We observe that excluding PV systems from our installed base measure that lie more than 50 meters away from the next road does not change the odds ratio of the installed base. If we use a distance-weighted measure of the installed base (column (2) in Table 4), the odds ratio for the installed base is a bit larger (but not statistically different) when only including PV systems which are located at most 50 meters away from the closest street to the installed base (comparing with column (3) of Table 2). However, if we construct an installed base only using PV systems that are more than 50 meters away from the next road ($IB_{n,t-1}^{NR50}$ in column (3), Table 4), the odds ratio increases (due to a much lower mean and lower standard deviation of $IB_{n,t-1}^{NR50}$ in comparison to $IB_{n,t-1}^{NR50}$, see descriptive statistics in Table 10, Appendix A.3) but turns out to be statistically insignificant. This finding indicates that visibility is important for peer effects in PV adoption since PV systems farther away than 50 meters from the next road are less likely to be visible.

In Appendix D.2, we include choice maker_n fixed effects α_n and District_i × Street_k × Year_t fixed effects again. In columns (1-2) of Table 23, we show that our results hold for instrumenting the installed base that only considers PV systems which are located at most 50 meters away from the closest street with the inclination and orientation of neighboring roofs that are also located at most 50 meters away from the closest street. Once again, the instruments are strong and the overidentifying restrictions test confirms their validity. We obtain similar results for a distance-weighted installed base that focuses on PV systems which are located at most 50 meters away from the closest street (columns (3-4)). We do not show the estimates but our results also hold for non-instrumented ordinary least squares for $IB_{n,t-1}^{NR50}$, $IB_{n,t-1}^{NR50,1/d}$ and $IB_{n,t-1}^{NR50}$.

Table 4: Visibility Results for Case Study of Baden-Wuerttemberg. Odds ratios from conditional logit estimates of peer effects by whether (or not) a road is within 50m distance (columns (1)-(3)) and whether (or not) a building is in the way (columns (4)-(7)).

	Visible						
	from all		not from all	from building			not from building
	All	All	All	All	Urban	Rural	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	CLogit	CLogit	CLogit	CLogit	CLogit	CLogit	CLogit
	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$
<i>Installed base (visibility measures):</i>							
Next road 50m: $IB_{n,t-1}^{NR50}$	1.02** (4.77)						
Next road 50m (dist.-wght.): $IB_{n,t-1}^{NR50,1/d}$		1.83* (2.29)					
No road 50m: $IB_{n,t-1}^{NR50}$			1.12 (1.70)				
No building in-between: $IB_{n,t-1}^{\triangleleft}$				1.06** (3.96)	1.05** (3.15)	1.07* (1.97)	
Building in-between: $IB_{n,t-1}^{\triangleleft}$							1.00 (-0.38)
<i>Controls:</i>							
Potential electricity production per sqm _n	5.2E+7 ** (10.8)	5.2E+7 ** (10.8)	5.1E+7** (10.8)	5.2E+7** (10.8)	4.8E+7** (9.25)	4.4E+7** (5.15)	5.1E+7** (10.8)
Potential electricity production per sqm _n × t	0.90 (-0.56)	0.90 (-0.56)	0.90 (-0.56)	0.90 (-0.57)	0.90 (-0.49)	0.88 (-0.32)	0.90 (-0.56)
Maximum module area _n	1.49** (2.73)	1.48** (2.72)	1.49** (2.73)	1.49** (2.74)	1.66** (2.66)	0.68 (-1.26)	1.49** (2.73)
Maximum module area _n × t	0.98 (-1.15)	0.98 (-1.14)	0.98 (-1.18)	0.98 (-1.19)	0.98 (-0.80)	1.02 (0.62)	0.98 (-1.17)
Rural _{j_n}	1.03 (1.13)	1.02 (0.76)	1.01 (0.63)	1.02 (0.66)			1.01 (0.59)
Observations	770,063	770,063	770,063	770,063	580,802	134,920	770,063
DF _M	6	6	6	6	5	5	6
Final log-likelihood	-142,383	-142,392	-142,393	-142,387	-106,127	-30,441	-142,395
Year _t fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Federal state _n fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District _i × Year _t fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District _i × Street _k × Year _t fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Choice maker _n fixed effects	No	No	No	No	No	No	No

Notes: Exponentiated coefficients (odds ratios); robust t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . All columns are estimated via logit. The panel of 2,903,730 choice makers over 10 years t (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. We estimate all columns with the conditional logit estimator. The sample is always the same. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District_i × Street_k × Year_t groups. I.e., District_i × Street_k × Year_t groups with no adoption (or if all choice makers in that group adopt at once) are dropped. These procedures result in fewer observations. In columns (1), (3)-(7), $f(d_{n,m}) = 1$, and in column (2), $f(d_{n,m}) = 1/d_{n,m}$. The Cut-off distance D is 200m.

Visibility from buildings. – Next, we exploit installed base measures that rely on direct visibility between buildings. The measure $IB_{n,t-1}^{\triangleleft}$ contains all PV systems visible from n and $IB_{n,t-1}^{\triangleleft}$ contains all PV systems not visible from n (because another building lies in-between). The corresponding estimates confirm our findings on visibility. According to Table 4 column (5), the odds of installing a PV system are higher if there are many directly visible PV systems nearby in comparison to all PV systems. Column (6) and (7) indicate that there is a significant peer effect in urban and rural areas. However, if we instrument the installed base that contains all PV systems visible from n ($IB_{n,t-1}^{\triangleleft}$) with the inclination and orientation of neighboring roofs that contains all PV systems visible from n and then again distinguish urban and rural locations, we can only confirm a causal

peer effect for this visibility measure for urban locations (see Table 24 in Appendix D.2). For urban locations, the instruments are strong and the overidentifying restrictions test confirms their validity.

We now focus on Table 4 column (7). The odds for PV adoption associated with the installed base that is based on PV systems, which are not directly visible from the choice maker’s building, are statistically insignificant. Table 25 in Appendix D.2 confirms this finding with the linear probability model and for distinguishing urban and rural locations.

Robustness. – Finally, we illustrate that our results are robust to including all visibility measures at once. Table 26 (in Appendix D.2) contains the OLS and the IV estimates. The specifications in Table 26 allow us to be even more confident that proximity is not the real channel driving of our results since (Next road 50m) $IB_{n,t-1}^{NR50}$ and (No road 50m:) $IB_{n,t-1}^{NR50}$ add up to $IB_{n,t-1}$. In consequence, the significantly positive estimate for (No building in-between:) $IB_{n,t-1}^{\leq}$ indicates the peer effect from visible previously installed PV systems nearby while controlling for all previously installed PV systems nearby.

In order to understand the relevance of the visibility effect, we can compare column (4) of Table 4 to column (1) of Table 2. We learn that one additional previously installed directly visible PV system nearby (No building in-between: $IB_{n,t-1}^{\leq}$) increases the odds of installing three times as much as one based on $IB_{n,t-1}$ (which does not distinguish visible and not visible systems).⁸

Our results also suggest that visibility is an important channel for peer effects in PV adoption but may be less important for peer effects in PV adoption in rural areas. Neighbors in rural areas may be more closely related to each other (in comparison to non-rural areas) and may therefore engage in word-of-mouth communication more often. In contrast, in non-rural areas peer effects may instead be driven by visibility.⁹ Assuming rural

⁸Our results are in line with Bollinger, Gillingham, et al.’s (2019) working paper on the importance of visibility in PV adoption in parts of Connecticut. They use one instrument for the visibility of neighboring PV systems, which is the visibility of neighboring PV systems from a road without considering obstruction. By construction, their instrument and their instrumented variable are very similar.

⁹We experiment with other measures of visibility that exploit the orientation of each single roof. They are in line with the results that we show.

population to be more interdependent compared to a more independent urban population, we enhance the results of Ma, Yang, and Mourali (2014) such that independent consumers are more sensitive to visibility than to alternative channels such as word-of-mouth.

4 Conclusion and Managerial Implications

In this paper, we empirically analyze the individual adoption decisions on rooftop PV installations in Germany through 2010. The small-scale, complete data set of locations of PV installations and locations of potential adopters enables us to generate a series of insights on spatio-temporal peer effects in PV adoption and new technology adoption in general. Our paper adds to the small but growing body of research that is using disaggregate choice data to document social contagion. We confirm the existence of a peer effect in PV adoption.

Our instrumental variable estimation provides strong evidence for a cause-and-effect relationship between prior installations and the individual decision to adopt. The positive peer effect is strongest on a very small scale – i.e., within distances up to 200m. PV systems nearby may be seen as a large diverse pool of information which reduces uncertainty about the technology. The decrease in the peer effect over time may be due to a decline in the marginal information contribution of every new PV system user in the number of previous adopters. Interestingly, modifications of subsidies may increase the importance of the peer effect again. We hypothesize that such a modification increases uncertainty for potential adopters. Further findings reveal that peer effects might be stronger in less developed regions (as the east of Germany) compared to regions of high economic activity. Finally, visibility drives peer effects in PV adoption in non-rural areas and may be less important in rural areas. For rural areas, the alternative (word-of-mouth) may be more relevant.

Our findings have important managerial implications. First, the decline of peer effects over time calls for sales strategies to differ over time as well. In the early stage of the diffusion process strategies to set up seed installations by pilot projects seem to be most efficient. In later stages other sales channels, such as employing more sales representatives,

might be more appropriate. We suppose that these stages can be determined by major changes in the feed-in tariffs (subsidy system) and the magnitude of the peer effect. This is an important insight for political decision makers: reducing subsidies increases demand uncertainty and hence reduces the decline of peer effect importance. This allows policy makers to steer PV adoption. Second, Seel, Barbose, and Wiser (2014) find a significant difference between customer acquisition cost in Germany and the US in the PV industry. The authors partly relate this finding to different marketing and sales processes. In particular, they discuss the fact that peer effects and word-of-mouth contribute to the lower cost in Germany compared to the US. With respect to our findings managers should regard the small-scale location of potential seed PV installations. If the location is visible to potential adopters nearby, the seed installation alone might be sufficient. If visibility is less present or word-of-mouth communication between neighbors is more likely, additional strategies as suggested by Bollinger and Gillingham (2012) might be more effective. Such a diverse sales strategy is likely to be effective and cost efficient. In particular, in non-rural areas seed installations alone might be sufficient while in rural areas setting incentives for early adopters to have personal conversations with potential adopters seems to be more appropriate. We find peer effects to be more important in less developed regions (as in the east of Germany). There may be a great potential for exploiting peer effects in PV adoption in developing countries.

Our findings are comparable to other technology adoption processes, such as the diffusion of electric vehicle systems, for example (Avci, Girotra, and Netessine 2014; Cohen, Lobel, and Perakis 2016). Buying an electric vehicle (EV) is an uncertain investment, as is buying a rooftop PV system. The uncertainty comes with the lack of experience on how to fulfill day-to-day mobility demands with an EV instead of a traditional gasoline car – in particular, in terms of range and recharging time (Lim, Mak, and Rong 2015; Kempton 2016). Although the hard numbers may be known, it might be difficult for households to interpret these numbers in terms of their mobility behavior. However, a neighbor with an EV in front of her house might reduce this uncertainty in terms of reporting how to manage her mobility needs with an EV. That being said, EV producing firms might

want to increase sales by discounts in markets at the beginning of the diffusion process to establish efficient seeds which may cause further adoptions. Further, designs that allow us to easily identify electric vehicles may support diffusion. In rural areas, incentives for EV adopters to stimulate conversations with neighbors about their EV may be beneficial. On top of that, prior PV installation might be a predictor for future EV purchase (Prins and Verhoef 2007).

References

- Albuquerque, Paulo, Bart J. Bronnenberg, and Charles J. Corbett (2007). “A Spatiotemporal Analysis of the Global Diffusion of ISO 9000 and ISO 14000 Certification”. *Management Science* 53(3), 451–468. DOI: 10.1287/mnsc.1060.0633.
- Angst, Corey M. et al. (2010). “Social Contagion and Information Technology Diffusion: The Adoption of Electronic Medical Records in U.S. Hospitals”. *Management Science* 56(8), 1219–1241. DOI: 10.1287/mnsc.1100.1183.
- Atkin, David et al. (2017). “Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan*”. *The Quarterly Journal of Economics* 132(3), 1101–1164. DOI: 10.1093/qje/qjx010.
- Avci, Buket, Karan Girotra, and Serguei Netessine (2014). “Electric Vehicles with a Battery Switching Station: Adoption and Environmental Impact”. *Management Science* 61(4), 772–794. DOI: 10.1287/mnsc.2014.1916.
- Bollinger, Bryan and Kenneth Gillingham (2012). “Peer Effects in the Diffusion of Solar Photovoltaic Panels”. *Marketing Science* 31(6), 900–912. DOI: 10.1287/mksc.1120.0727.
- Bollinger, Bryan, Kenneth Gillingham, et al. (2019). *Visibility and Peer Influence in Durable Good Adoption (June 14, 2019)*. SSRN Working Paper. DOI: 10.2139/ssrn.3409420.
- Chen, Yubo, Qi Wang, and Jinhong Xie (2011). “Online Social Interactions: A Natural Experiment on Word of Mouth versus Observational Learning”. *Journal of Marketing Research* 48(2), 238–254. DOI: 10.1509/jmkr.48.2.238.

- Cohen, Maxime C., Ruben Lobel, and Georgia Perakis (2016). “The Impact of Demand Uncertainty on Consumer Subsidies for Green Technology Adoption”. *Management Science* 62(5), 1235–1258. DOI: 10.1287/mnsc.2015.2173.
- Conley, Timothy G. and Christopher 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.
- Farrell, Joseph and Garth Saloner (1986). “Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation”. *American Economic Review* 76(5), 940–955. URL: <http://www.jstor.org/stable/1816461>.
- Goldenberg, Jacob et al. (2009). “The Role of Hubs in the Adoption Process”. *Journal of Marketing* 73(2), 1–13. DOI: 10.1509/jmkg.73.2.1.
- Graziano, Marcello and Kenneth 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.H. (2012). *Econometric Analysis*. 7th ed. New Jersey: Pearson Education.
- Hinz, Oliver et al. (2011). “Seeding Strategies for Viral Marketing: An Empirical Comparison”. *Journal of Marketing* 11, 55–71. DOI: 10.1509/jm.10.0088.
- IEA (2019). *IEA PVPS Snapshot of Global PV 2019 – Report IEA PVPS T1-35:2019*. Report. URL: http://www.iea-pvps.org/fileadmin/dam/public/report/statistics/IEA-PVPS_T1_35_Snapshot2019-Report.pdf.
- IPCC (2018). *Global warming of 1.5 Degree Celsius*. Special Report. URL: <http://www.ipcc.ch/report/sr15/>.
- Karshenas, Massou D. and Paul 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.
- Katona, Zsolt, Peter Pal Zubcsek, and Miklos Sarvary (2011). “Network Effects and Personal Influences: The Diffusion of an Online Social Network”. *Journal of Marketing Research* 48(3), 425–443. DOI: 10.1509/jmkr.48.3.425.

- Kempton, Willett (2016). “Electric vehicles: Driving range”. *Nature Energy* 1, 16131. DOI: 10.1038/nenergy.2016.131.
- Lim, Michael K., Ho-Yin Mak, and Ying Rong (2015). “Toward Mass Adoption of Electric Vehicles: Impact of the Range and Resale Anxieties”. *Manufacturing & Service Operations Management* 17(1), 101–119. DOI: 10.1287/msom.2014.0504.
- Ma, Zhenfeng, Zhiyong Yang, and Mehdi Mourali (2014). “Consumer Adoption of New Products: Independent versus Interdependent Self-Perspectives”. *Journal of Marketing* 78(2), 101–117. DOI: 10.1509/jm.12.0051.
- Manchanda, Puneet, Ying Xie, and Nara Youn (2008). “The Role of Targeted Communication and Contagion in Product Adoption”. *Marketing Science* 26(6), 961–976. DOI: 10.1287/mksc.1070.0354.
- Manski, Charles F. (1993). “Identification of Endogenous Social Effects: The Reflection Problem”. *The Review of Economic Studies* 60(3), 531–542. URL: <http://www.jstor.org/stable/2298123>.
- McFadden, D. (2001). “Economic choices”. *American Economic Review* 91(3), 351–378. DOI: 10.1257/aer.91.3.351.
- Prins, Remco and Peter C. Verhoef (2007). “Marketing Communication Drivers of Adoption Timing of a New E-Service among Existing Customers”. *Journal of Marketing* 71(2), 169–183. DOI: 10.1509/jmkg.71.2.169.
- Redding, Stephen J. and Daniel 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.
- Risselada, Hans, Peter C. Verhoef, and Tammo H.A. Bijmolt (2014). “Dynamic Effects of Social Influence and Direct Marketing on the Adoption of High-Technology Products”. *Journal of Marketing* 78(2), 52–68. DOI: 10.1509/jm.11.0592.
- Rode, Johannes and Alexander 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, Everett M. (1983). *Diffusion of Innovations*. 3rd. New York: Free Press.
- Seel, Joachim, Galen L. Barbose, and Ryan H. Wiser (2014). “An analysis of residential PV system price differences between the United States and Germany”. *Energy Policy* 69, 216–226. DOI: 10.1016/j.enpol.2014.02.022.
- Wirth, Harry (2019). *Recent Facts about Photovoltaics in Germany (version from: October 14, 2019)*. Report. URL: <https://www.ise.fraunhofer.de/content/dam/ise/en/documents/publications/studies/recent-facts-about-photovoltaics-in-germany.pdf>.

Online-Appendix

For online publication: The following is not intended to be included in the journal version of the article, but as an e-companion.

Acknowledgments

We express our deep gratitude for the friendly support of Bing’s Spatial Data Service that allowed us to geocode our PV data set. We gratefully acknowledge that David Gschwender at (LUBW – Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg) provided base maps with detailed information as PV appropriateness, inclination and orientation of roofs in Baden-Württemberg. 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. We gratefully acknowledge address data provided by Hamburg Business School. Johannes Rode acknowledges 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 (Rode 2014).

A Data and Modeling Approach

A.1 Spatio-temporal dimension of data

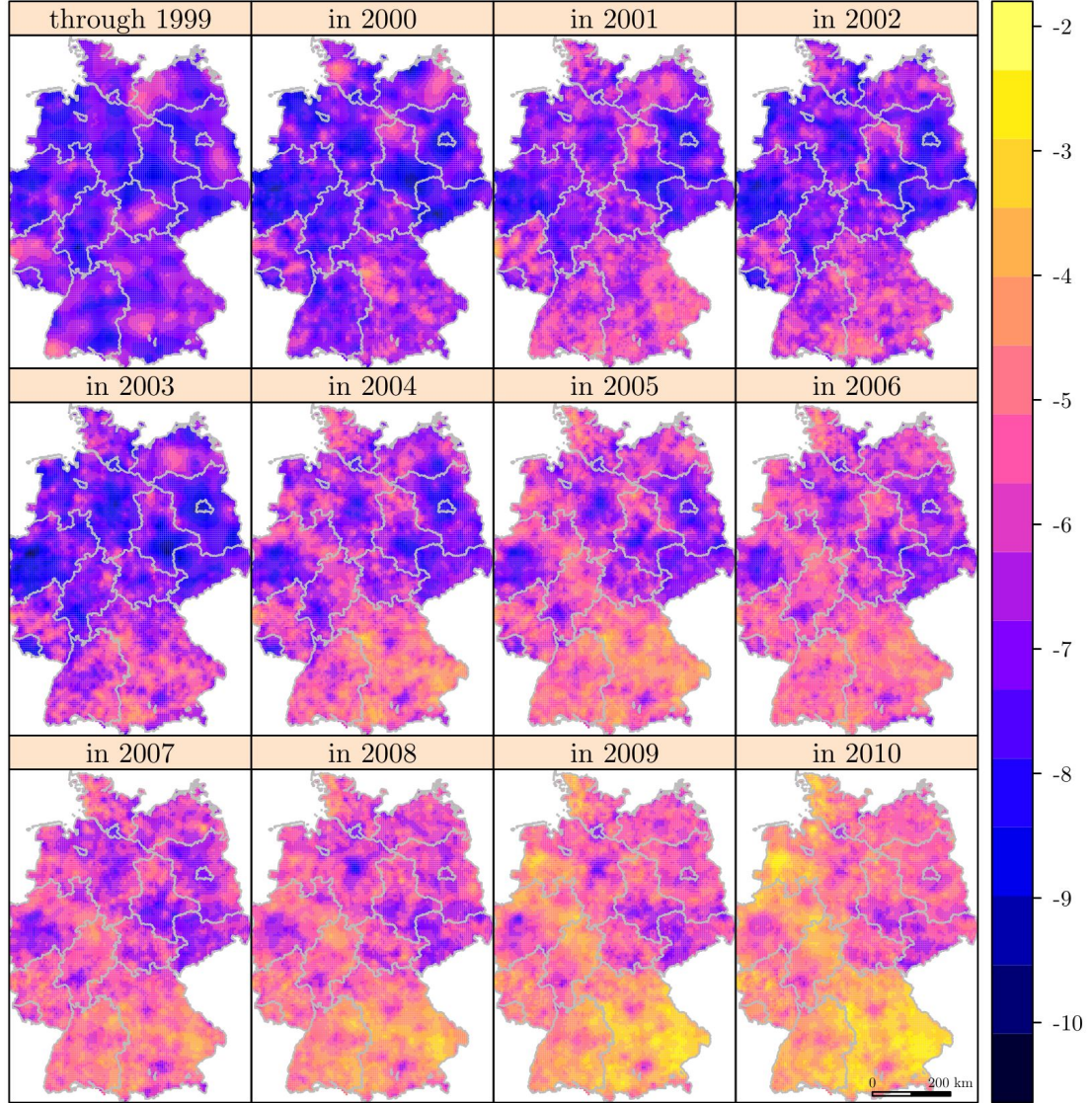


Figure 4: The spatio-temporal diffusion of PV systems in Germany through 2010.

Notes: The numbers in the legend (on the right-hand) show the natural logarithm of yearly annex of PV installations divided by the 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.

A.2 Germany-wide analysis

Potential adopters. – We obtain the location (in coordinates in the coordinate system WGS84) of 18,413,514 buildings in Germany in 2009 (Infas 2009a).¹⁰ We assume that each building has an owner. Certainly, whether the owner is a private household or a house cooperation (or a firm) may make 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 peer effects in PV adoption, particularly those using aggregate approaches, accounts for ownership (see Bollinger and Gillingham (2012), Müller and Rode (2013), Rai and Robinson (2013), Noll, Dawes, and Rai (2014), Graziano and Gillingham (2015), Rode and Weber (2016), and Graziano, Fiaschetti, and Atkinson-Palombo (2019)).

PV systems. – We 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. Buildings are denoted as observations n . We study the years $t = 2001, \dots, 2010$. We start in 2001 since very few systems were set up before this date (see Figure 4 in Appendix A.1 and Figure 5). 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, Oschmann, and Theobald (2008) or Rode and Weber (2016) for details on the feed-in tariff.

We obtain a data set that covers all 879,020 grid-connected PV systems installed in Germany through the end of the year 2010 from I-TSO (2012).¹¹ The data set includes the location (address and 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 (also see Appendix B.1). Information on capacity, a

¹⁰Note 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 buildings only increased by 7.4%: from 16,977,662 in 2001 to 18,234,580 in 2010 (DESTATIS 2016).

¹¹We downloaded the data on March 4, 2012. By now, a new platform was established by the German transmission set operators. Current data on renewable energy systems is available here <https://www.netztransparenz.de/EEG/Anlagenstammdaten> (last visit January 9, 2020). However, the current data set does not include addresses anymore. Postalcodes are still available.

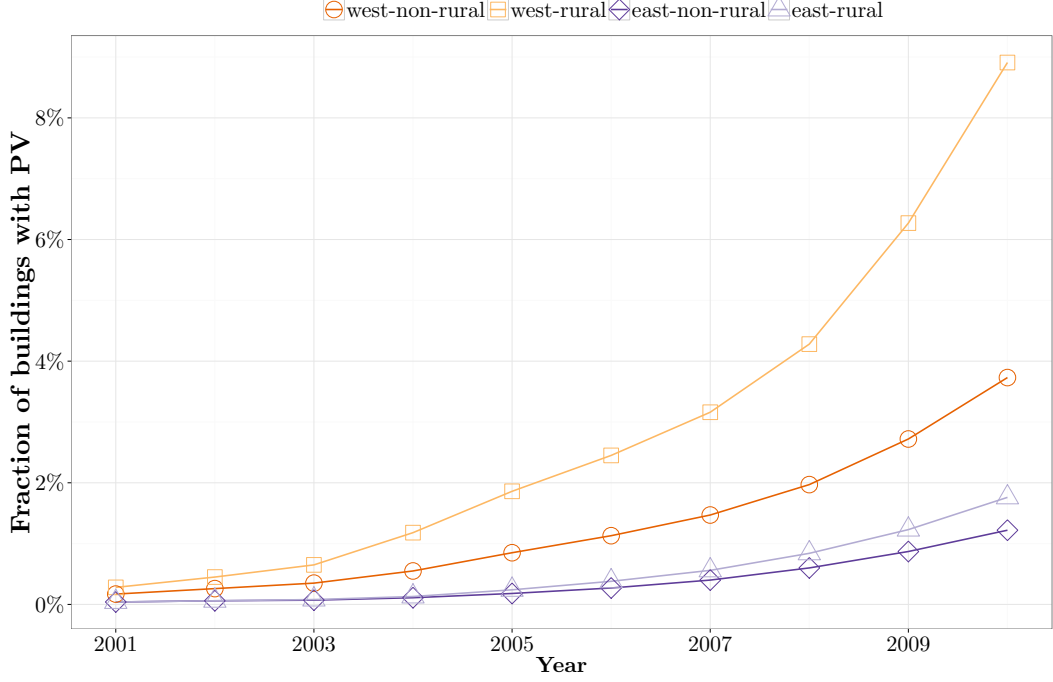


Figure 5: Cumulative fraction of buildings with PV system by east and west, and rural and non-rural.

measure of size of PV systems, allows us to select small-scale household systems (capacity ≤ 30 kilowatt). Finally, we end up with 769,525 PV systems, each allocated to a mutually exclusive building. As a result we observe the adoption of PV technology: the dummy variable $y_{n,t}$ indicates whether there is a PV system installed at building n in year t (or not). The corresponding frequencies are given in Table 5. Figure 1 (see main part of the paper) visualizes the spatio-temporal dimension of the $y_{n,t}$ by a small-scale example.

Installed base. – In order to analyze the impact of peers on the decision to adopt we introduce the installed base (Farrell and Saloner 1986; Bollinger and Gillingham 2012; Graziano and Gillingham 2015). The installed base measures the number of preexisting PV installations within a given radius around the location of a choice maker. We define

$$IB_{n,t} = \sum_{m \in \bar{N}_n} \sum_{l=2000}^t y_{m,l} \quad (1)$$

as a spatio-temporal lag variable ($IB_{n,t-1}$). It measures the dependencies between buildings n and m . We set the buildings nearby to $\bar{N}_n = \{m \in N | m \neq n, d_{n,m} \leq D\}$. The Euclidean distance in meters between the location of n and the location of m is denoted as $d_{n,m} > 0$. Radius D is a cut-off parameter to be set by the analyst. We may assume

Table 5: Frequencies of choices ($\Delta y_{n,t}$) for household PV systems in Germany 2000-2010.

Year _t	Alternative	Frequency in category
2000	0: No (or no new) PV system installed	18,404,093
	1: New PV system installed	9,421
2001	0: No (or no new) PV system installed	18,379,367
	1: New PV system installed	24,726
2002	0: No (or no new) PV system installed	18,361,025
	1: New PV system installed	18,342
2003	0: No (or no new) PV system installed	18,341,255
	1: New PV system installed	19,770
2004	0: No (or no new) PV system installed	18,297,922
	1: New PV system installed	43,333
2005	0: No (or no new) PV system installed	18,236,232
	1: New PV system installed	61,690
2006	0: No (or no new) PV system installed	18,177,152
	1: New PV system installed	59,080
2007	0: No (or no new) PV system installed	18,107,212
	1: New PV system installed	69,940
2008	0: No (or no new) PV system installed	18,004,135
	1: New PV system installed	103,077
2009	0: No (or no new) PV system installed	17,848,534
	1: New PV system installed	155,601
2010	0: No (or no new) PV system installed	17,643,989
	1: New PV system installed	204,545
Sum of household PV installations installed between 2000 and 2010		769,525

Notes: There are 18,413,514 choice makers (buildings) n over 10 years t (2001-2010) since we lag our installed base measures by one year. The choice makers are distributed across 77,847 districts i and 16 federal states. 18,555 PV systems were installed before our period of analysis (≤ 1999). Summing up the number of observations with choice 0 and 1 does not result in the same number of choice makers across years since, once a choice maker has installed a PV system, we will neglect the choice maker in the following years.

that there is no remarkable influence of PV installations farther away from location n than D . In our baseline specification, we set $D = 200\text{m}$.

We also define a distance-weighted installed base:

$$IB_{n,t}^{1/d} = \sum_{m \in \tilde{N}_n} \sum_{l=2000}^t y_{m,l} f(d_{n,m}). \quad (2)$$

In this case, we weight the importance of each location by the inverse distance: $f(d_{n,m}) = 1/d_{n,m}$. In robustness analyzes (Appendix C.4), we employ different values of D and functional forms of $f(d_{n,m})$.¹² Alternatively, n might also only be influenced by peers who adopted in the immediately preceding period, i.e., in $t - 1$. We define a corresponding non-cumulative measure for the installed base since the subsidy system for PV systems changed during the period of study and only peers who adopted in the preceding year ($t - 1$) may pass on reliable information regarding, e.g., the reliability, initial costs, and

¹²Note that the first period $l = 2000$ includes all PV systems installed in the year 2000 and before. In the estimations, we focus on newly installed PV systems between 2001, 2002,...,2010 by year (on the left-hand side). In consequence, the (lagged) installed base relies on the PV systems installed though 2000, through 2001,...,through 2009.

the net present value of PV systems:

$$\Delta IB_{n,t}^{1/d} = \sum_{\substack{m \in N, \\ m \neq n \\ d_{n,m} \leq D}} y_{m,t} f(d_{n,m}). \quad (3)$$

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

We employ 2006's CORINE Land Cover (CLC) data set (CLC 2009) to identify rural areas j .¹³ We take into account the dummy variable $Rural_{j_n}$ denoting whether building n is located in rural areas j . Figure 6 illustrates the high level of detail of $Rural_{j_n}$. The interaction of the dummy variables $East_{i_n}$ and $Rural_{j_n}$ with the installed base accounts for spatial-non-stationarity in peer effects. The corresponding descriptive statistics are provided in Table 6.

Table 6: Descriptive statistics for Germany.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $\Delta y_{n,t}$.0028	.053	0	1
<i>Installed base:</i>				
$IB_{n,t-1}$.57	1.2	0	43
<i>Installed base (distance-weighted):</i>				
$IB_{n,t-1}^{1/d}$.0062	.015	0	1.03
<i>Controls:</i>				
$East_{i_n}$.18	.38	0	1
$Rural_{j_n}$.2	.4	0	1
N	182,750,649			

Notes: The 182,750,649 observations come from 18,413,514 choice makers (buildings) n over 10 years t (2001-2010). We censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 182,750,649 observations. The choice makers are distributed across 7787 districts i in Baden-Wuerttemberg. $\Delta y_{n,t}$ depends on $y_{n,t}^*$ and will be defined in section A.4.

The decision to adopt 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

¹³The CLC data set comprises vector data on a scale of 1:100,000. The minimum mapping unit for the polygons is 0.25 sq-km. 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).

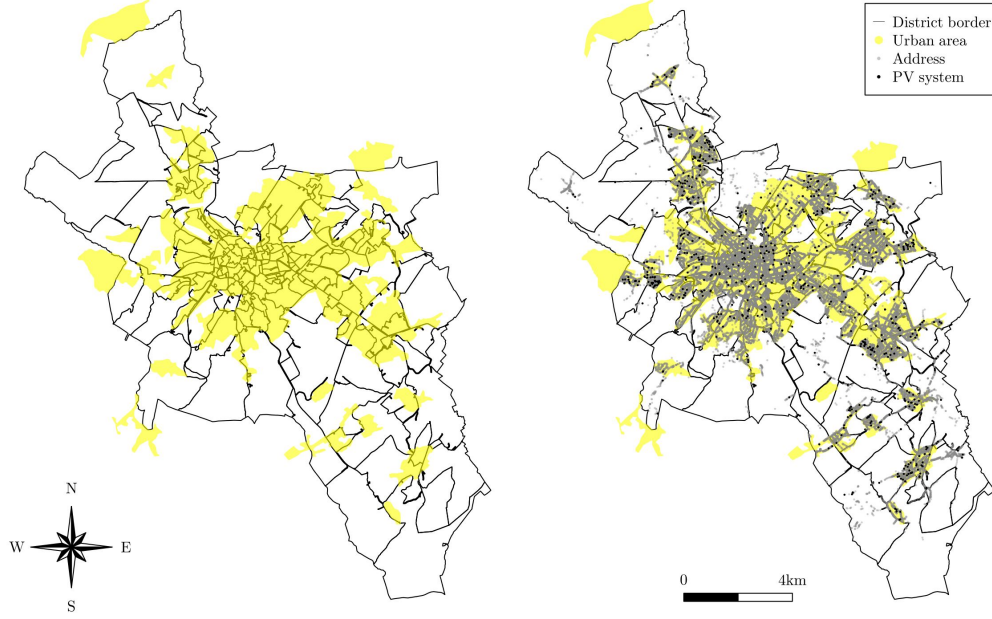


Figure 6: Level of detail for distinguishing rural and urban areas.

Notes: The figures illustrate the high level of detail of the CORINE Land Cover data set that allows us to identify rural areas. As an example, we show the city of Aachen with some 250.000 inhabitants. Even within this city, many addresses do not lie in a yellow colored urban area.

since we do not have access to time-variant data with a reasonably high spatial resolution. In a robustness check, we include information on global radiation, elevation, population density, firm density and purchasing power. Appendix B.2 contains details.

A.3 Case Study Baden-Württemberg

Detailed data on Baden-Württemberg allows us to identify causal peer effects and to build installed base measures that take into account the visibility of a PV system. This analysis is based on a sub-sample consisting of 71,432 household PV systems installed in Baden-Württemberg through 2010.

Roof appropriateness for PV. – To identify causal peer effects, we need exogenous variation correlated with our measure of the installed base (1) but not driving the actual adoption decisions. We employ high-resolution data on the inclination and orientation of roofs, which both shape the appropriateness of roofs for PV systems. We focus on the federal state Baden-Württemberg. Figure 7 illustrates that roof orientation and inclination shape the PV appropriateness of a roof. The figure shows approximate ratios relative

to the optimal orientation and inclination. In western Europe, the optimal orientation, corresponding to an exploitation of 100%, is to the south (180 degrees) with an inclination of about 37 degrees.

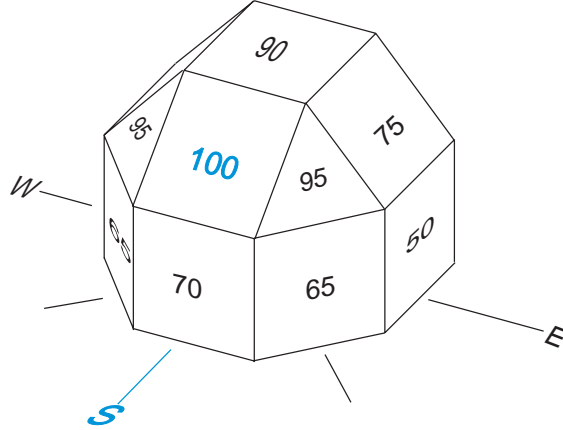


Figure 7: Roof orientation and inclination shape the PV appropriateness of a roof (Stark et al. 2005, p. 20).

Notes: The figure contains approximate ratios relative to the optimal orientation and inclination. In western Europe, the optimal orientation is to the south with an inclination of about 37 degrees. 0 degrees refers to a flat roof.

Roof census. – We exploit a roof census from (LUBW 2016). Based on laser scan data from overflights between 2000 and 2005, the data set builds on information on the roof inclination, orientation, area and solar radiation to calculate the potential area of PV systems on each roof for the 3 million buildings in Baden-Württemberg. This roof census (LUBW 2016) is based on the location of buildings from the cadastral land register of Baden-Württemberg as of 2012. The outer walls of a building define its contour. Overhanging roofs are not included. The census considers roof areas which have a solar energy potential between 75% and 100% of the maximum solar radiation in a region as appropriate for PV systems. For inclined roofs, areas have to allow for more than 10 sqm of PV modules to be included. Flat roofs are only included if they allow for more than 25 sqm of PV systems (assuming a PV system built on stilts).

Figure 8 illustrates the high-resolution data from (LUBW 2016) as an example for Freiburg. In the figure, red roofs have a very good suitability for PV. This corresponds to a solar energy potential between 95% and 100% of the maximum solar radiation. Orange roofs have a good suitability (solar energy potential between 80% and 94%), light blue a limited suitability (solar energy potential between 75 and 79%) and gray roofs have to be checked on-site (solar energy potential below 75%). Areas with a solar energy potential

below 75% are not considered as areas potentially appropriate for PV. Panel (a) illustrates that most roofs have a very good or good PV suitability in Freiburg.

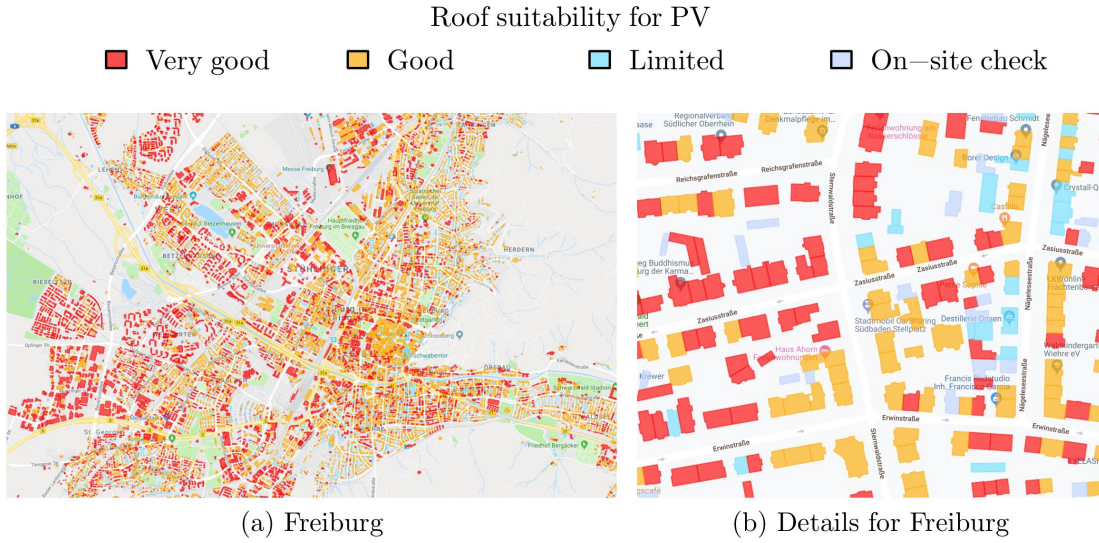


Figure 8: Screenshots of information from roof-census for Freiburg in Baden-Württemberg. Panel (b) shows the same excerpt of Freiburg as Figure 2.

Notes: The city Freiburg is the fourth largest in Baden-Württemberg, located in the south-east. We took the screenshot on 02.10.2019 from the roof-census website: <https://www.energieatlas-bw.de/sonne/dachflächen/potenzial-dachflächenanlagen>. Red roofs have a very good suitability for PV, which corresponds to a solar energy potential between 95% and 100% of the maximum solar radiation. Orange roofs have a good suitability (solar energy potential between 80% and 94%), light blue a limited suitability (solar energy potential between 75 and 79%) and gray roofs have to be checked on-site (solar energy potential below 75%).

In panel (b) of Figure 8, we zoom in and observe the accurateness of the roof information in detail. From panel (b) we learn several things. First, roofs of buildings in the same street located next to each other do not necessarily have the same suitability for PV. Second, there are roofs for which no suitability assessment is given (see lower part of panel (b) and (c)). This can have different reasons. As outlined earlier, for inclined roofs, areas have to allow for more than 10 sqm of PV modules to be included. Further, flat roofs are only included if they allow for more than 25 sqm of PV systems (assuming a PV system built on stilts). Finally, buildings that were constructed before 2012 (date of cadastral maps) but after the laser scan (between 2000 and 2005) are all classified as inappropriate for PV.¹⁴ The following website shows the high-resolution data in general: <https://www.energieatlas-bw.de/sonne/dachflächen/potenzial-dachflächenanlagen> (last visit: 02.10.2019). Zooming and clicking on a specific roof indicates the roof's appropriate PV area (in German: Mögliche geeignete Modulfläche).

¹⁴However, out of the 3.36 million buildings only 10.5% are inappropriate. Some of them will indeed be inappropriate, for others no information on appropriateness is available since the laser scan was conducted (2000-2005) before they were constructed. We focus on the ones appropriate for PV.

In our case study data set, we have 2,957,332 observations (i.e., roofs and therefore potential places for PV installations). The 2,957,332 observations are distributed across 8,988 districts i . Geocoded PV systems fall within a roof polygon or are (as in our Germany-wide analysis) allocated to the next roof.¹⁵ We successfully allocate 71,432 PV systems to unique roofs. If more than one PV system is allocated to the same roof, we neglect the PV systems in our baseline results.¹⁶ Table 7 and Table 9 present the frequencies of the categories for the case study and the descriptive statistics.

Table 7: Frequencies of choices ($\Delta y_{n,t}$) for household PV systems in Baden-Württemberg, 2000-2010.

Year _{t}	Alternative	Frequency in category
2000	0: No (or no new) PV system installed	2,956,583
	1: New PV system installed	749
2001	0: No (or no new) PV system installed	2,954,352
	1: New PV system installed	2,980
2002	0: No (or no new) PV system installed	2,952,464
	1: PV system installed	1,888
2003	0: No (or no new) PV system installed	2,950,594
	1: New PV system installed	1,870
2004	0: No (or no new) PV system installed	2,946,514
	1: New PV system installed	4,080
2005	0: No (or no new) PV system installed	2,940,436
	1: New PV system installed	6,078
2006	0: No (or no new) PV system installed	2,934,109
	1: New PV system installed	6,327
2007	0: No (or no new) PV system installed	2,926,974
	1: PV system installed	7,135
2008	0: No (or no new) PV system installed	2,916,838
	1: New PV system installed	10,136
2009	0: No (or no new) PV system installed	2,903,730
	1: New PV system installed	13,108
2010	0: No (or no new) PV system installed	2,886,649
	1: New PV system installed	17,081
Sum of PV installations over all years		71,432

Notes: There are 2,957,332 choice makers (buildings) n over 10 years t (2001-2010) since we lag our installed base measures by one year. The choice makers are distributed across 8,988 districts i and one federal state. Summing up the number of observations with choice 0 and 1 does not result in the same number of choice makers across years since, once a choice maker has installed a PV system, we will neglect the choice maker in the following years.

Besides the PV appropriateness in 4 categories, the roof census shape-files from LUBW (2016) include building-specific information on how much area (in square meters) is appropriate for PV systems, how much electricity can potentially be generated from a standardized PV system (with 15% power efficiency) on each building and roof-specific information on inclination and orientation. We normalize the potential electricity

¹⁵Due to inaccuracy in geocoding and the possibility that more than one PV system could be installed on one building, not all geocoded PV systems fall inside a roof-polygon. In consequence, we have to allocate about 1/3 of the PV systems to their nearest building. Of these, the median distance to the nearest building to which the PV system is allocated is 20 meters.

¹⁶However, robustness tests confirm that our results are robust to including PV systems that are allocated to the same roof (see Table 20 and the corresponding descriptive statistics in Table 19).

production by roof area and denote Potential electricity production per sqm_n = Potential electricity production $_n$ / Maximum module area $_n$. Potential electricity production per sqm_n is measured in MWh/a. Maximum module area $_n$ is the roof area feasible for a PV system in sqm .

A.3.1 Measures for the inclination and orientation of neighboring roofs

Table 8 shows the descriptive statistics for our raw data on roof inclination and orientation. The optimal inclination for a PV system is 37 degrees. Inclination $_n^*$ is a natural number with values between 0 and 74 degrees, which we transform to values between 0 and 37 degrees. If the inclination is larger than 37, we transform it to the corresponding value between 0 and 37, i.e. Inclination $_n$ = 74 - Inclination $_n^*$.

Table 8: Descriptive statistics for roof inclination and orientation.

	Mean	Std. Dev.	Min.	Max.
Inclination $_n^*$	30	17	0	74
Orientation $_n^*$	152	81	-1	359
N	2,957,332			
<i>Notes:</i> We have information on roof inclination and roof orientation for 2,957,332 choice makers (buildings) in Baden-Wuerttemberg. The choice makers are distributed across 8,982 districts i in Baden-Wuerttemberg. Inclination and orientation are measured in degrees. An inclination of zero and an orientation of -1 corresponds to a flat roof. An orientation of 180 indicates perfect south orientation.				

The optimal orientation for a PV system is 180 degrees. Orientation $_n^*$ is a natural number with values between -1 and 359 degrees, which we transform to values between 0 and 180. If the orientation is actually larger than 180, we transform it to the corresponding value between 0 and 180, i.e. Orientation $_n$ = 360 - Orientation $_n^*$. Flat roofs have Inclination $_n^*$ = 0 and Orientation $_n^*$ = -1. We assign the optimal inclination and the optimal orientation value to them since perfectly inclined and perfectly oriented PV systems (on stilts) can be installed there.

In Figure 9, we illustrate the raw data for roof inclination and orientation. We observe again that roofs of buildings in the same street located next to each other do not necessarily have the same suitability for PV, i.e. the same roof inclination and orientation.

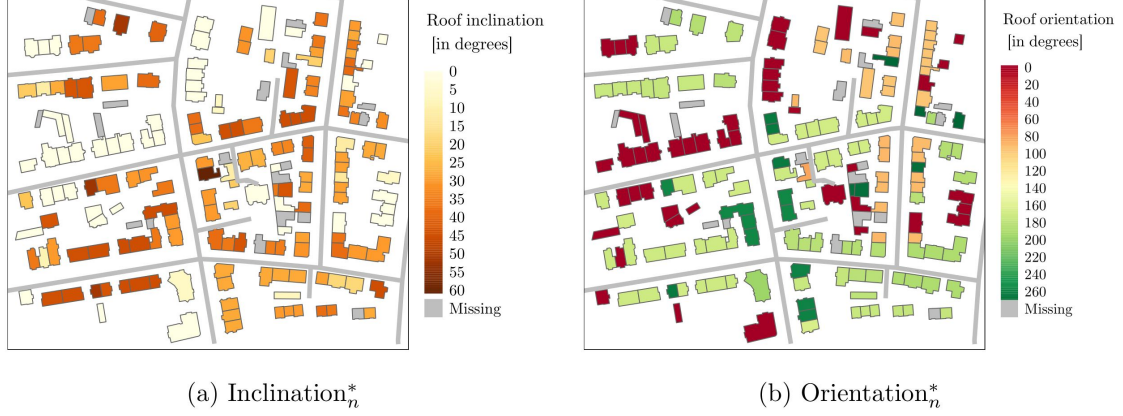


Figure 9: Illustration of roof inclination and roof orientation from roof-census in Baden-Württemberg. Both panels show the same excerpt of Freiburg as Figure 2.

Notes: We show the raw data on roof inclination and roof orientation. Gray lines are streets. The optimal inclination is 37 degrees. Flat roofs have an inclination of 0. In panel (b), dark red illustrates flat roofs. We assign, the optimal inclination and the optimal orientation to flat roofs since perfectly inclined and perfectly oriented PV systems (on stilts) can be installed there. The optimal orientation for a PV systems is 180 degrees, which corresponds to the south.

We define the measures for the inclination and orientation of neighboring roofs as follows:

$$\text{Inclination of neighboring roofs}_{n,t} = \sum_{m \in \bar{N}} \sum_{l=2000}^t y_{m,l} \frac{\text{Inclination}_m}{\text{Optimal inclination}}, \quad (4)$$

$$\text{Orientation of neighboring roofs}_{n,t} = \sum_{m \in \bar{N}} \sum_{l=2000}^t y_{m,l} \frac{\text{Orientation}_m}{\text{Optimal orientation}}. \quad (5)$$

As before, m defines the buildings neaby n . We also define distance-weighted versions of our instruments with $f(d_{n,m}) = 1/d_{n,m}$:

$$\text{Inclination of neighboring roofs}_{n,t}^{1/d} = \sum_{m \in \bar{N}} \sum_{l=2000}^t y_{m,l} f(d_{n,m}) \frac{\text{Inclination}_m}{\text{Optimal inclination}}, \quad (6)$$

$$\text{Orientation of neighboring roofs}_{n,t}^{1/d} = \sum_{m \in \bar{N}} \sum_{l=2000}^t y_{m,l} f(d_{n,m}) \frac{\text{Orientation}_m}{\text{Optimal orientation}}. \quad (7)$$

Our instruments are spatio-temporal lag variables. They measure the number of preexisting PV installations within radius D around the location of building n , with the

importance of each location weighted by an exogenous measure for the inclination of a roof on a linear scale between 0 and 100% or an exogenous measure for the orientation of a roof on a linear scale between 0 and 100%. Buildings (roofs) not appropriate for PV, which still have a PV installation, have a low impact, while PV systems installed on very appropriate roofs are strongly weighted. Note that the instruments do not rely on the maximum module area of a roof since we do not want to weight larger roofs stronger: larger roofs may be associated with confounding factors such as higher incomes or levels of education which may be related to the PV adoption decision.

We assume the inclination and the orientation of a neighboring building m should be exogenous to the adoption decision of a decision maker located in n since in almost all cases the buildings were constructed without considering a PV system.¹⁷ In an IV estimation we employ (4) and 5 as exogenous variables to predict $IB_{n,t-1}$ – denoted as $\widehat{IB_{n,t-1}}$ – and then, in a second step, use $\widehat{IB_{n,t-1}}$ to identify causal peer effects in PV adoption.

A.3.2 Measures for visibility

Visibility from roads. – To analyze the impact of visibility on peer effects, we merge the roof-specific information of a building (orientation and inclination) with OpenStreetMap data on the course of roads (Geofabrik 2018). This enables us to measure whether a PV system is visible from the nearest street. We assume that PV systems located farther away than 50 meters from the closest street are not likely to be visible. Consider the parameter dr_n denoting the distance in meters of a building n to the closest road and the set $N_n^{\text{nr}} = \{m \in N \mid m \neq n, d_{n,m} \leq D, dr_m \leq 50\}$. Now, the variable

$$IB_{n,t}^{\text{NR50}} = \sum_{m \in N_n^{\text{nr}}} \sum_{l=2000}^t y_{m,l} f(d_{n,m}) \quad (8)$$

is based on our baseline installed base measure (1) but only considers PV systems which are within 50 meters distance to a road. In the baseline, we set $f(d_{n,m}) = 1$. In robustness tests, we use $IB_{n,t}^{\text{NR50}}$ for which we set $f(d_{n,m}) = 1/d_{n,m}$.

¹⁷The number of new buildings is very low. E.g., considering Baden-Württemberg the average yearly increase of residential buildings was 0.73% between 31.12.1999 and 31.12.2010.

Table 9: Descriptive statistics for Case Study of Baden-Wuerttemberg.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $\Delta y_{n,t}$.0024	.049	0	1
<i>Installed base:</i>				
$IB_{n,t-1}$	1.5	2.3	0	44
Predicted: $\widehat{IB}_{n,t-1}$	1.5	2.2	-.13	45
Instrument: Inclination of neighboring roofs $_{n,t-1}$	1.1	1.7	0	34
Instrument: Orientation of neighboring roofs $_{n,t-1}$	1.2	1.8	0	38
<i>Installed base (distance-weighted):</i>				
$IB_{n,t-1}^{1/d}$.016	.027	0	1.1
Instrument: Inclination of neighboring roofs $_{n,t-1}^{1/d}$.012	.02	0	.84
Instrument: Orientation of neighboring roofs $_{n,t-1}^{1/d}$.012	.022	0	.95
<i>Controls:</i>				
Potential electricity production per sqm $_n$.14	.011	.1	.16
Potential electricity production per sqm $_n \times t$.88	.4	.21	1.8
Maximum module area $_n$.08	.18	.01	29
Maximum module area $_n \times t$.52	1.3	.02	321
Rural $_n$.21	.41	0	1
N	29,383,343			

Notes: In the case study, the 29,383,343 observations come from 2,957,332 choice makers (buildings) n over 10 years t (2001-2010). The choice makers are distributed across 8988 districts i in Baden-Wuerttemberg. $\Delta y_{n,t}$ depends on $y_{n,t}^*$ and will be defined in section A.4.

The orientation of roof-top PV systems in Germany are mainly to the south since this guarantees the highest efficiency (Dewald and Truffer 2011). We use this orientation to construct several measures of visibility for the installed base. The variable $IB_{n,t}^{NR50}$ is based on (8), but only considers PV systems (i.e., buildings m) with no road located within 50 meters distance. Such PV systems should be less visible than those with the next road within 50 meters distance (included in $IB_{n,t}^{NR50}$). Therefore, we assume peer effects based on $IB_{n,t}^{NR50}$ are less related to visual exposure, while $IB_{n,t-1}^{NR50}$ indicates that visual exposure of PV systems to peers might be the main channel driving the peer effect in the adoption process. In Figure 2 (main part of the paper), we illustrate $IB_{n,t-1}^{NR50}$ in green while the large red point indicates that the PV system (in yellow and marked with an x, lying in the middle of the green area) would be visible from the street. In Table 10, we show the summary statistics of the corresponding variables.

Visibility from buildings. – We also build installed base measures based on direct visibility of PV systems from one building to another. We define the incidence parameter $pos_{n,m} = 1$, if m is located in the north of n (0, otherwise) and we define the incidence parameter $vis_{n,m} = 1$, if building m is visible from location of n (0, otherwise). Now the set $N_n^{\text{vis}} = \{m \in N | m \neq n, vis_{n,m} \times pos_{n,m} = 1, d_{n,m} \leq D\}$

Table 10: Descriptive statistics for visibility measures for Case Study of Baden-Wuerttemberg.

	Mean	Std. Dev.	Min.	Max.
<i>Installed base (visibility measure):</i>				
Next road 50m: $IB_{n,t-1}^{NR50}$	1.5	2.3	0	44
(next road 50m): Inclination of neighboring roofs $_{n,t-1}^{NR50}$	1.1	1.7	0	34
(next road 50m): Orientation of neighboring roofs $_{n,t-1}^{NR50}$	1.2	1.8	0	38
<i>Installed base (visibility measure, distance-weighted):</i>				
Next road 50m: $IB_{n,t-1}^{NR50,1/d}$.016	.027	0	1.1
Instrument: Inclination of neighboring roofs $_{n,t-1}^{NR50,1/d}$.012	.02	0	.84
Instrument: Orientation of neighboring roofs $_{n,t-1}^{NR50,1/d}$.012	.021	0	.95
<i>Installed base (visibility measure):</i>				
No building in-between: $IB_{n,t-1}^{\triangleleft}$.046	.22	0	7
Instrument: Inclination of neighboring roofs $_{n,t-1}^{\triangleleft}$.033	.17	0	4.9
Instrument: Orientation of neighboring roofs $_{n,t-1}^{\triangleleft}$.036	.19	0	7
<i>Installed base (visibility measure):</i>				
No road 50m: $IB_{n,t-1}^{NR50}$.0077	.093	0	6
Building in-between: $IB_{n,t-1}^{\nabla}$.32	.75	0	34
N	29,383,343			

Notes: In this case study specification, the 29,383,343 observations come from 2,957,332 choice makers (buildings) n over 10 years t (2001-2010). The choice makers are distributed across 8988 districts i in Baden-Wuerttemberg. $\Delta y_{n,t}$ depends on $y_{n,t}^*$ and will be defined in section A.4.

includes all visible buildings north of n within distance D . We include all buildings north of n within distance D that are **not** visible from the location of n in the set $N_n^{\text{notvis}} = \{m \in N \mid m \neq n, pos_{n,m} = 1, vis_{n,m} = 0, d_{n,m} \leq D\}$. N_n^{vis} is then used to calculate $IB_{n,t-1}^{\triangleleft}$ (see orange corridor area in Figure 2 in the main part of the paper) and N_n^{notvis} to calculate $IB_{n,t-1}^{\nabla}$ (see blue corridor area in Figure 2 in the main part of the paper).

We also denote instrumental variables that are based on the same logic as the installed base visibility measures: the next road within 50m distance installed base ($IB_{n,t-1}^{NR50}$) corresponds to the instruments Inclination of neighboring roofs $_{n,t-1}^{NR50}$ and Orientation of neighboring roofs $_{n,t-1}^{NR50}$; the distance-weighted next road within 50m distance installed base ($IB_{n,t-1}^{NR50,1/d}$) relates to the instruments Inclination of neighboring roofs $_{n,t-1}^{NR50,1/d}$ and Orientation of neighboring roofs $_{n,t-1}^{NR50,1/d}$; the no building in-between installed base ($IB_{n,t-1}^{\triangleleft}$) comply with the instruments Inclination of neighboring roofs $_{n,t-1}^{\triangleleft}$ and Orientation of neighboring roofs $_{n,t-1}^{\triangleleft}$.

A.4 Econometric Model

We employ a discrete choice model (McFadden 2001; Gowrisankaran and Stavins 2004; Chintagunta and Nair 2011) to analyze all potential adopters' individual decisions to adopt a PV system (Müller and Rode 2013). So far, the considered observations are buildings and their locations. Since every building is owned by someone (e.g., private household or legal person), we consider this individual to make the choice to install a PV system in a certain period or not (Müller and Rode 2013). As such we denote observation n as the choice maker from hereon. In making the choice whether to adopt the technology, the choice maker weighs up the marginal advantages (and disadvantages) of adoption. We denote the (to the analyst) unobservable utility of adoption to choice maker n in period t as the latent variable

$$u_{n,t} = v_{n,t} + \epsilon_{n,t} \quad (9)$$

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 adoption choice, $\Delta y_{n,t}$, is a discrete (binary) choice. The observed values of the endogenous variable $\Delta y_{n,t} = y_{n,t} - y_{n,t-1}$ are related to $u_{n,t}$ according to random utility theory (Manski 1977):

$$\Delta y_{n,t} = \begin{cases} 1 & \text{if } u_{n,t} > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

We censor observations, because a choice maker adopts at most once within the considered time span. Observation n is censored for $t > t^*$ if $y_{n,t^*} = 1$, i.e. n has adopted in t^* (that is, to not consider the same adoption choice again). Since $u_{n,t}$ in (1) is a stochastic variable, the probability that choice maker n chooses to install a photovoltaic system in year t is

$$P_{n,t} = \Pr(\Delta y_{n,t} = 1) = \Pr(u_{n,t} > 0). \quad (11)$$

If we assume that $\epsilon_{n,t}$ are independent and identically extreme value distributed (Train 2009), then

$$P_{n,t} = \frac{e^{v_{n,t}}}{1 + e^{v_{n,t}}}. \quad (12)$$

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

$$v_{n,t} = \beta_{j,t}IB_{n,t-1} + \gamma'X_{n,t} + \eta_{i_n,t}(\times\theta_{k_n}) + (\delta_t + \alpha_n). \quad (13)$$

We denote the l -dimensional vector $X_{n,t}$ as attributes related to choice maker n with vector γ' to measure the respective effects. In our baseline specification for Germany, we include $\eta_{i_n,t}$: locational fixed effects for every year t .¹⁸ In robustness checks, we instead use δ_t , i.e. temporal fixed-effects common across choice makers, and α_n , i.e. choice maker n fixed effects invariant over time with n as the location related to a choice maker. Since our data are cross-sectional time-series (panel) data, we mostly employ the fixed-effects logit model for panel data as described in Greene (2012, pp. 721–724). Using the corresponding conditional likelihood function, we can condition δ_t , α_n , and $\eta_{k_n,t}$ out from our model (13) and get estimates of β and γ' using StataCorp (2015). Parameter β is of focal interest since it measures the association of peer decisions on the odds of adopting, i.e., the peer effect. Accordingly, the odds ratio which defines the probability of adoption relative to non-adoption is given as

$$\text{odd}_{n,t} = \frac{P_{n,t}}{1 - P_{n,t}} = e^{v_{n,t}}. \quad (14)$$

Based on (14)

$$\text{odd}_{IB}(\beta) = e^\beta, \text{ and} \quad (15)$$

$$\text{odd}_l(\gamma_l) = e^{\gamma_l} \quad (16)$$

are the odds ratios associated with a one-unit increase in $IB_{n,t-1}$ (15) or a one-unit increase in the l th control variable (16).

¹⁸ i refers to the 77,847 districts across Germany, t to the year 2001-2010. In the case-study of Baden-Württemberg, we also include θ_{k_n} , fixed effects on for 544,243 streets k .

We also estimate linear probability models with multi-level fixed effects:

$$\Delta y_{n,t} = \beta_{j,t} IB_{n,t-1} + \gamma' X_{n,t} + \eta_{i_n,t} (\times \theta_{k_n}) + \alpha_n + \epsilon_{n,t}, \quad (17)$$

where $\epsilon_{n,t}$ is an error term. This approach allows us to include time-specific locational $\eta_{i_n,t} \times \theta_{k_n}$ fixed effects and choice maker α_n fixed effects at the same time. For estimation, we use Correia (2014) in StataCorp (2015).

In our case-study of Baden-Württemberg, we extend our analysis with an instrumental variable approach. We first estimate

$$\begin{aligned} IB_{n,t-1} = & \alpha_1 \text{Inclination of neighboring roofs}_{n,t-1} \\ & + \alpha_1 \text{Orientation of neighboring roofs}_{n,t-1} \\ & + \gamma X_{n,t-1} + \eta_{i_n,t-1} \times \theta_{k_n} + \alpha_n + \varepsilon_{n,t-1} \end{aligned} \quad (18)$$

with ordinary least squares. $\eta_{i_n,t} \times \theta_{k_n}$ are locational fixed effects for every year $t - 1$ and $\varepsilon_{n,t-1}$ is an iid normal error term. In a second step, we use the predicted values of the installed base $\left(\widehat{IB_{n,t-1}} \right)$ from (18) in (13) or (17) to identify causal peer effects in PV adoption. For estimation, we use Correia (2018) in StataCorp (2015).

B Raw data

B.1 Geocoding and Allocation of PV systems to buildings

Table 11 shows the geocoding accuracy of the address data on 879,020 PV installations. 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 11. 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 11: 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

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.¹⁹ This is due to inaccuracy in geocoding, missing address information in the PV 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 i . We conduct robustness checks which confirm that this procedure does not drive our results. The 77,847 statistical districts i are taken from Infas (2009b). This procedure results in 877,114 PV systems, each allocated to a mutually exclusive building.

¹⁹These do not drive our results. We confirm our findings with a data set which neglects the PV systems allocated to these 269,752 buildings.

B.2 Time-invariant controls

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 i provided by Infas (2009b). Low values of $Population\ Density_{i_n}$ may refer to places with a high share of single- and double-family homes. For 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 $Population\ Density_{i_n}$.

We expect that buildings located in statistical districts i with high values of $Firm\ Density_{i_n}$ – 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.

$Purchasing\ Power_{k_n}$ 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 and Rode (2013)), we expect that the probability of PV adoption increases with *Purchasing Power* _{k_n} .

C Robustness for Germany

C.1 Main specification with distance-weight and robustness

Table 12 corresponds to Table 1 but refers to distance-weighted measures of the installed base. The results in Table 12 are in line with the results in the main text. For instance, column (4) in Table 12 includes $\text{District}_i \times \text{Year}_t$ fixed effects. We estimate the time-varying measure of the installed base and control for the east and rural parts of Germany as well as their interaction. Since we include these interaction terms, the odds ratio of 8745 for $IB_{n,t-1}$ corresponds to non-rural areas in the west of Germany in 2001. An obvious interpretation of the level of our results is in terms of a 0.02 unit increase in $IB_{n,t-1}$, which corresponds roughly to a one standard deviation increase and is equal to one previously installed PV system 50 meters away. The estimates from column (2) in Table 12 reveal that for a 0.02 increase in the installed base, we expect the odds of installing a PV system (in non-rural areas in the west in 2001) to increase by $(\exp(\ln(8745)/50) - 1) \times 100\% = 19.9\%$. A 0.02 increase in the installed base increases the odds of installing a PV system for instance in 2006 by $(\exp(\ln(259 \times 0.0018)/50) - 1) \times 100\% = 5.7\%$. Interestingly, we still find evidence for spatial non-stationarity: the odds ratio for the interaction between $IB_{n,t-1}$ and East_{in} is significantly greater than one, i.e., the peer effect may be more important (in non-rural areas) in the east of Germany.

We conduct several other robustness tests. For these, we consider (if not mentioned otherwise) the specification with $\text{District}_i \times \text{Year}_t$ fixed effects.

Capacity-adjusted samples: We 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, which we analyze in the baseline specification. 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.²⁰ Specification Industrial PV in column (1) in Table 15 rejects a significantly positive peer effect for industrial PV systems (PV systems larger

²⁰Also see Comin and Rode (2013) for a similar approach to distinguishing household and industrial PV adoption.

Table 12: Estimates for (distance-weighted) peer effects and their spatio-temporal variation in Germany with different types of fixed effects. Column (1) shows odds ratios from logit, columns (2, 4-5) odds ratios from conditional logit and column (3) coefficients from OLS.

	Year FE	District-Year FE	District-Year & Choice maker FE	District-Year FE	Choice maker FE
	(1) CLogit	(2) CLogit	(3) OLS	(4) CLogit	(5) CLogit
	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$y_{n,t}$
<i>Installed base (distance-weighted):</i>					
$IB_{n,t-1}^{1/d}$	232.8** (105.3)	1.38** (3.87)	0.010** (8.96)	8745.0** (6.05)	1877.9** (7.81)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2002}$				0.025* (-2.14)	0.00099** (-6.85)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2003}$				0.00094** (-3.98)	0.00019** (-8.75)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2004}$				0.00019** (-5.30)	0.000027** (-10.6)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2005}$				0.00064** (-4.73)	0.0000084** (-12.0)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2006}$				0.0018** (-4.11)	0.0000010** (-14.1)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2007}$				0.00065** (-4.80)	0.00000030** (-15.4)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2008}$				0.00060** (-4.90)	0.00000050** (-15.1)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2009}$				0.00070** (-4.82)	0.00000053** (-15.1)
$IB_{n,t-1}^{1/d} \times \text{Year}_{2010}$				0.00092** (-4.65)	0.00000055** (-15.1)
$IB_{n,t-1}^{1/d} \times \text{East}_{i_n}$				19.6** (5.32)	0.27 (-1.79)
$IB_{n,t-1}^{1/d} \times \text{Rural}_{j_n}$				0.034** (-21.3)	0.43** (-4.14)
$IB_{n,t-1}^{1/d} \times \text{Rural}_{j_n} \times \text{East}_{i_n}$				98.4** (3.94)	6.56 (1.22)
<i>Controls:</i>					
East_{i_n}				1.42 (0.68)	
Rural_{j_n}				1.49** (81.0)	
$\text{Rural}_{j_n} \times \text{East}_{i_n}$				0.76** (-14.6)	
Observations	182,750,649	74,997,032	182,727,991	74,997,032	5,765,940
DF _M	25	1	1	16	22
Final log-likelihood	-3,349,011	-2,618,886		-2,614,867	-1,179,735
Adj. R^2			0.07		
Year _t fixed effects	Yes	Yes	Yes	No	No
Federal state _n fixed effects	Yes	Yes	Yes	No	No
District _t × Year _t fixed effects	No	Yes	Yes	Yes	Yes
Choice maker _n fixed effects	No	No	Yes	No	No

Notes: Exponentiated coefficients (odds ratios) in columns (1-2) and (4-5) and coefficients in column (3); robust t statistics in parentheses (in column (3), t statistics are based on SE clustered on the district-level); * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . Column (1) is estimated via logit. The panel of 18,413,514 choice makers over 10 years t (2001-2010) results in 184,135,140 observations. We censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 182,750,649 observations. We estimate column (1) with logit and columns (2,4-5) with the conditional logit estimator. The sample is the same across columns. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District_t × Year_t groups for column (2) and (4) (choice makers for column (5)). I.e., for column (2) and (4) District_t × Year_t groups with no adoption (or if all choice makers in that group adopt at once) are dropped. For column (5) choice makers with no adoption are dropped. Similarly, singleton observations are dropped in column (3). These procedures result in fewer observations. We estimate but do not show the constant for column (1). $f(d_{n,m}) = 1/d_{n,m}$ and the Cut-off distance D is 200m.

than 1 megawatt total nominal power). Table 13 contains the corresponding descriptive statistics. Of course, the installed base measure's mean is smaller for industrial PV systems than for household 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 (Appendix A.4).

Table 13: Descriptive statistics of Industrial PV data set for Germany.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $y_{n,t}$.003	.055	0	1
<hr/> <i>Industrial installed base:</i> <hr/>				
$IB_{n,t-1}$.0011	.0052	0	.16
<hr/> <i>Controls:</i> <hr/>				
Rural $_{jn}$.23	.42	0	1
East $_{in}$.3	.46	0	1
<hr/>				
N	4,200,439			
<hr/>				
<i>Notes:</i> 4,200,439 observations come from 18,413,514 choice makers (buildings) n over 10 years t (2001-2010). The number of observations is much lower than in our baseline specification since the number here refers to the observations that are left after conditioning out the fixed effects. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District $_i$ \times Year $_t$ groups. I.e., for the specification in column (1) of Table 15 District $_i$ \times Year $_t$ groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in fewer observations.				

Placebo test: We conduct a placebo test to verify that our estimations do not find a positive peer effect in PV system adoption by definition. We randomly allocate the same number of PV installations, which were in fact installed in Germany per year, to existing buildings. Table 14 contains the corresponding descriptive statistics. Column (2) in Table 15 reveals no significantly positive peer effect in PV adoption, supporting the validity of our findings.

Table 14: Descriptive statistics of data set with Placebo allocations for Germany.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $y_{n,t}$.0075	.087	0	1
<hr/> <i>Placebo installed base:</i>				
$IB_{n,t-1}$.048	.046	0	.96
<hr/> <i>Controls:</i>				
Rural $_{jn}$.19	.39	0	1
East $_{in}$.18	.38	0	1
<hr/>				
N	113,761,840			
<hr/>				

Notes: 113,761,840 observations come from 18,413,514 choice makers (buildings) n over 10 years t (2001-2010). The number of observations is much lower than in our baseline specification since the number here refers to the observations that are left after conditioning out the fixed effects. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of $\text{District}_i \times \text{Year}_t$ groups. I.e., for the specification in column (2) of Table 15 $\text{District}_i \times \text{Year}_t$ groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in fewer observations.

Last year’s installations only: Incorporating PV installations from the last year instead of all previous installations in the installed base measure – as outlined in Equation 3 Appendix A.2 – does not affect our core findings (see column (3) of Table 15). The estimates indicate that last year’s PV installations are more relevant for the peer effect in comparison to the ones installed a longer time ago. This finding is in line with Graziano and Gillingham’s (2015, p. 19) evidence for a “*diminishing neighbor effect over time since prior installations*”.

The reflection problem: Our findings remain unaffected by lagging our peer effect measure by two years (Richter 2013; Rode and Weber 2016). The results are available upon request. 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, n 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.²¹

District and adopter characteristics: In Appendix C.2 and C.3, we discuss characteristics of districts with many PV systems. We also illuminate adopter characteristics by exploiting the household panel survey Socio-Economic Panel (SOEP 2013). We control for global radiation, elevation, firm density, population density, and purchasing power on the district level i . These controls confirm a significant, diminishing peer effect over time, which is lower in rural areas and higher in the east. The estimates provide evidence that high global radiation and high elevation yield higher odds of installing PV while firm density and population density decrease these odds. We found that adopters are of

²¹We 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 between 100m and 200m distance to the choice maker (see Graziano and Gillingham (2015) for a similar analysis). This approach confirms our previous findings.

high socio-economic status – in particular owning a house strongly increases the odds of adopting a PV system.

In Appendix C.4, we discuss the fact that our results are robust to different cut-off distances. We also discuss findings with different functional forms of the distance decay function $f(d_{nm})$.

Table 15: Odds ratio of spatio-temporal variation of peer effects in Germany.

	Industrial PV	Placebo allocation	Installed base last year
	(1)	(2)	(3)
	CLogit	CLogit	CLogit
	$\Delta y_{n,t}$	$\Delta y_{n,t}$	$\Delta y_{n,t}$
<i>Installed base measures (distance-weighted):</i>			
$IB_{n,t-1}$	1000158.4 (0.82)	0.0947*** (-12.55)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2002}$	7808.8 (0.38)	1.107 (0.36)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2003}$	1.16534e + 13 (1.22)	1.419 (1.25)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2004}$	2140.7 (0.39)	1.373 (1.37)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2005}$	8.25e - 11 (-1.10)	1.202 (0.84)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2006}$	8.30e - 14 (-1.62)	1.011 (0.05)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2007}$	0.000113 (-0.52)	1.305 (1.23)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2008}$	1.55e - 09 (-1.17)	1.197 (0.87)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2009}$	2.58e - 09 (-1.17)	1.155 (0.72)	
$IB_{n,t-1}^{1/d} \times \text{Year}_{2010}$	1.74e - 09 (-1.19)	1.099 (0.48)	
$IB_{n,t-1}^{1/d} \times \text{Rural}_{jn}$	0.109 (-0.45)	0.859 (-1.41)	
$IB_{n,t-1}^{1/d} \times \text{East}_{in}$	160337.5* (2.56)	0.867 (-1.41)	
$IB_{n,t-1}^{1/d} \times \text{Rural}_{jn} \times \text{East}_{in}$	4.41e - 08 (-1.14)	1.740* (2.17)	
$\Delta IB_{n,t-1}^{1/d}$			452.5** (4.27)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2002}$			0.121 (-1.29)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2003}$			0.0212* (-2.22)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2004}$			0.0132** (-2.74)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2005}$			0.0253* (-2.45)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2006}$			0.0362* (-2.23)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2007}$			0.00843** (-3.20)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2008}$			0.0105** (-3.11)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2009}$			0.00469** (-3.70)
$\Delta IB_{n,t-1}^{1/d} \times \text{Year}_{2010}$			0.00576** (-3.58)
$\Delta IB_{n,t-1}^{1/d} \times \text{Rural}_{jn}$			0.0682** (-12.35)
$\Delta IB_{n,t-1}^{1/d} \times \text{East}_{in}$			6.403** (2.78)
$\Delta IB_{n,t-1}^{1/d} \times \text{Rural}_{jn} \times \text{East}_{in}$			68.74** (3.21)
East_{in}	0.523 (-1.47)	0.937 (-0.27)	1.407 (0.79)
Rural_{jn}	1.446** (12.44)	0.956** (-9.02)	1.295** (66.24)
$\text{Rural}_{jn} \times \text{East}_{in}$	0.751** (-4.73)	0.992 (-0.75)	0.848** (-11.78)
Observations	4,200,439	113,761,840	88,156,268
DF _M	16	16	16
Final log-likelihood	-71,899	-4,346,150	-3,690,897
Year _t fixed effects	Yes	Yes	Yes
District _i × Year _t fixed effects	Yes	Yes	Yes
Choice maker _n fixed effects	No	No	No

Notes: Exponentiated coefficients (odds ratios); robust t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . Across columns, we use the conditional logit estimator. The panel of 18,413,514 choice makers over 10 years t (2001-2010) results in the 184,135,140 observations. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District _{i} × Year _{t} groups. I.e., District _{i} × Year _{t} groups with no adoption (or if all choice makers adopt at once) are dropped. This procedure results in fewer observations. $f(d_{n,m}) = 1/d_{n,m}$ and the Cut-off distance D is 200m.

C.2 District characteristics

We are interested in the characteristics of districts i with many PV systems. If we take into account district characteristics, we cannot include fixed effects at the local level. In an unshown specification, we account for spatial differences in global radiation, elevation, firm density, population density and purchasing power (besides controlling for federal state and year fixed effects). These controls improve the model fit significantly (in comparison to column (1) Table 1) and indicate an even larger effect for our measure of the installed base. The peer effect is significant and diminishes over time. Again it 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).²² 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 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 household systems.

²²A one unit increase refers to an increase in 10 kWh/sqm of yearly averaged global radiation.

According to this specification with district characteristics, disproportionately many PV systems can be found in sparsely populated areas. Graziano and 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 and 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 SOEP (2013) representative for Germany, see Appendix C.3.

C.3 Adopters characteristics

C.3.1 SOEP Data

The socio-economic panel (SOEP) 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), Universitaet, TH (University, Technical College), Hochschule im Ausland (College Not In Germany), Ingenieur-, Fachschule (Ost) (Engineering, Technical School (East)), Hochschule (Ost) (University (East)). Table 16 contains the descriptive statistics.

Table 16: Descriptive statistics for SOEP.

	Mean	Std. Dev.	Min.	Max.
$\Delta \text{Solar}_{h_n,t}$.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.

C.3.2 SOEP Results

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. There is no relevant difference between both types of solar energy systems for our analysis (Comin and Rode 2013).²³

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.

²³Since 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.

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 C.3.1 for details on the variable definitions and their descriptives.

The estimates in Table 17 illustrate that solar energy adopters are of high socioeconomic 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 17). 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 17), we see that solar energy adoption is also associated with high levels 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 and Rode 2013). In contrast, tenants may be less involved in the adoption decision.²⁴

Table 17: 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×NUTS-1 fixed effects	Yes	Yes
Observations	85,568	39,671
DF _M	7	8
Final log-likelihood \mathcal{L}	-6684.8	-4059.9

Notes: Exponentiated coefficients; robust t statistics in parentheses clustered on federal state-year groups; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, 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. 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 fewer observations.

²⁴We 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 and Rode (2013).

Unshown 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 5: 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.

C.4 Cut-off distances

We conduct estimations with adjusted peer effect measures (and $\text{District}_i \times \text{Year}_t$ fixed effects). The results are available upon request.

The results from our baseline regression remain unaffected from reducing the cut-off distance to 100m or increasing the cut-off distance to 400m. 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 (confirmed by a Horowitz test).²⁵

We conclude that studies concentrating on peer effects at a high level of geographical aggregation may fail to represent the peer effect appropriately. This is in line with previous studies on PV adoption. Bollinger and Gillingham (2012) find stronger peer effects on the street level than on the zip code level. Rode and 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.

²⁵E.g., Ben-Akiva and 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.

D Robustness for case study of Baden-Württemberg

D.1 Robustness for Causality

We conduct several robustness tests for our case study of Baden-Württemberg. First, we show that our results are robust to adding choice maker_{*n*} fixed effects α_n (while still considering District_{*i*}×Street_{*k*}×Year_{*t*} fixed effects). Table 18 column (1) shows the estimates which are in line with those obtained in the baseline. This approach also controls for time-invariant individual effects, which could be caused by individual-specific characteristics as income.

Table 18: OLS and instrument variable estimates for peer effects (with additional Choice Maker fixed effects) for Case Study of Baden-Wuerttemberg.

	No distance-weighting			With distance-weighting	
	OLS	IV		IV	
		1st stage (OLS)	2nd stage (OLS)	1st stage (OLS)	2nd stage (OLS)
	(1)	(2)	(3)	(4)	(5)
	$\Delta y_{n,t}$	$IB_{n,t-1}$	$\Delta y_{n,t}$	$IB_{n,t-1}^{1/d}$	$\Delta y_{n,t}$
<i>Installed base:</i>					
$IB_{n,t-1}$	0.00012** (4.21)				
<i>Predicted installed base:</i>					
$\widehat{IB}_{n,t-1}$			0.00016** (5.20)		
<i>Instruments:</i>					
Inclination of neighboring roofs _{<i>n,t-1</i>}		0.45** (137.8)			
Orientation of neighboring roofs _{<i>n,t-1</i>}		0.77** (250.2)			
<i>Predicted installed base (distance-weighted):</i>					
$\widehat{IB}_{n,t-1}^{1/d}$					0.0055** (3.35)
<i>Instruments (distance-weighted):</i>					
Inclination of neighboring roofs _{<i>n,t-1</i>} ^{1/d}				0.46** (124.4)	
Orientation of neighboring roofs _{<i>n,t-1</i>} ^{1/d}				0.76** (222.8)	
<i>Controls:</i>					
Potential electricity production per sqm _{<i>n</i>} × <i>t</i>	0.014** (32.1)	−0.029** (−7.94)	0.014** (32.1)	−0.0017** (−25.1)	0.014** (32.1)
Maximum module area _{<i>n</i>} × <i>t</i>	0.00012** (4.84)	−0.0017** (−9.19)	0.00012** (4.86)	−0.000050** (−19.2)	0.00013** (4.88)
Observations	27,643,514	27,643,514	27,643,514	27,643,514	27,643,514
DF _M	3	4	3	4	3
Adj. <i>R</i> ²	0.06		0.00		0.00
F	351.9		354.6		349.0
Hansen J (p-value)			0.2 (0.65)		0.5 (0.48)
Kleibergen-Paap rk Wald F statistic			834,123.1		814,637.6
Year _{<i>t</i>} fixed effects	Yes	Yes	Yes	Yes	Yes
Federal state _{<i>n</i>} fixed effects	Yes	Yes	Yes	Yes	Yes
District _{<i>i</i>} × Year _{<i>t</i>} fixed effects	Yes	Yes	Yes	Yes	Yes
District _{<i>i</i>} × Street _{<i>k</i>} × Year _{<i>t</i>} fixed effects	Yes	Yes	Yes	Yes	Yes
Choice maker _{<i>n</i>} fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients; *t* statistics in parentheses (based on SE clustered at street-level); * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker *n*'s observed choice in terms of (newly) installing a PV system in *t*. All columns are estimated via OLS. Column (3) uses the predicted values for the installed base from the first stage in column (2). Column (5) uses the predicted values for the installed base from the first stage in column (4). The panel of 2,903,730 choice makers over 10 years *t* (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. Across columns, the sample is always the same and singleton observations are dropped. In columns (1)-(3), $f(d_{n,m}) = 1$, and in columns (4)-(5), $f(d_{n,m}) = 1/d_{n,m}$. The Cut-off distance *D* is 200m.

Second, our results also hold when instrumenting the installed base with our measures for the inclination and orientation of neighboring roofs with choice maker_{*n*} fixed effects (columns (2-3)). The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. Thus, the instruments are still strong (column (2)). We now turn to the second stage regression in column (3), which also come from ordinary least squares (the linear probability model). In the second stage regression we use the predicted installed base from the first stage. The effect of the instrumented installed base on the adoption decision is strong and significant at all confidence levels.

We again conduct an overidentifying restrictions test. As both instruments are strongly significant in the first stage regression, the overidentifying restriction test will have power. The p-value for the Hansen statistic is 0.65, which indicates that our instruments are uncorrelated with the error term at all significance levels. The overidentifying restrictions test supports the validity of our instruments. Further, the OLS specification in column (1) and the instrumented in column (3) are in line.

Third, our results remain unaffected when using a distance-weighted installed base (columns (4-5), Table 18). The distance-weighted installed base is instrumented with distance-weighted measures for the inclination and orientation of neighboring roofs. Again, these instruments are strong and the overidentifying restrictions test confirms their validity.

Forth, we can modify our sample. The case study results shown so far are based only on those adoption decisions ($\Delta y_{n,t}$) for which we can be very sure that allocation of the PV system is to the correct roof. We can, however, also include those PV systems for which no unique allocation was possible. They are then randomly assigned to another roof in the same District_{*i*}. Table 19 contains the descriptive statistics for $\Delta y_{n,t}^{\text{All}}$.

Table 19: Descriptive statistics for Case Study of Baden-Wuerttemberg, all PV systems.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $\Delta y_{n,t}^{\text{All}}$.0052	.072	0	1
N	29,145,636			

Notes: In this case study specification, the 29,145,636 observations come from 2,956,938 choice makers (buildings) *n* over 10 years *t* (2001-2010). The choice makers are distributed across 8988 districts *i* in Baden-Wuerttemberg. $\Delta y_{n,t}$ depends on $y_{n,t}^*$, which is defined in section A.4.

In Table 20, we show that analyzing the full sample does not change our results. For reasons of clarity, we focus on the IV estimates here with choice maker_{*n*} fixed effects α_n and District_{*i*}×Street_{*k*}×Year_{*t*} fixed effects. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. Thus, the instruments are strong (column (1)). In the second stage regression (column (2)), we use the predicted installed base from the first stage. As before, the effect of the instrumented installed base on the adoption decision is strong and significant at all confidence levels. We again conduct an overidentifying restrictions test. The test supports the validity of our instruments. Further, the estimates with the full sample are in the ballpark of the estimates of the baseline sample.

Table 20: Instrument variable estimates for peer effects for Case Study of Baden-Wuerttemberg, all PV installations.

	1st stage (OLS)	2nd stage (OLS)
	(1)	(2)
	IB _{<i>n,t-1</i>}	$\Delta y_{n,t}^{\text{All}}$
<i>Predicted installed base:</i>		
$\bar{I}B_{n,t-1}$		0.00010* (2.24)
<i>Instruments:</i>		
Inclination of neighboring roofs _{<i>n,t-1</i>}	0.45** (135.8)	
Orientation of neighboring roofs _{<i>n,t-1</i>}	0.77** (246.6)	
<i>Controls:</i>		
Potential electricity production per sqm _{<i>n</i>} × <i>t</i>	-0.057** (-15.3)	0.015** (22.9)
Maximum module area _{<i>n</i>} × <i>t</i>	-0.0019** (-10.1)	0.00010* (2.38)
Observations	27,401,397	27,401,397
DF _M	4	3
Adj. R ²		0.00
F		176.6
Hansen J (p-value)		1.7 (0.20)
Kleibergen-Paap rk Wald F statistic		810,036.8
Year _{<i>t</i>} fixed effects	Yes	Yes
Federal state _{<i>n</i>} fixed effects	Yes	Yes
District _{<i>i</i>} × Year _{<i>t</i>} fixed effects	Yes	Yes
District _{<i>i</i>} × Street _{<i>k</i>} × Year _{<i>t</i>} fixed effects	Yes	Yes
Choice maker _{<i>n</i>} fixed effects	Yes	Yes

Notes: We show coefficients; *t* statistics in parentheses (based on clustered SE at street-level); * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker *n*'s observed choice in terms of (newly) installing a PV system in *t*. All columns are estimated via OLS. Column (2) uses the predicted values for the installed base from the first stage in column (1). The panel of 2,956,938 choice makers over 10 years *t* (2001-2010) results in 29,569,380 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,145,636 observations. Across columns, the sample is always the same and singleton observations are dropped. $f(d_{n,m}) = 1$ and the Cut-off distance *D* is 200m.

Fifth, as for the Germany-wide analysis, we can lag our installed base measure by two years in the case study. Table 21 contains the descriptive statistics.

In Table 22, we illustrate that our estimates are robust to lagging our installed base measure by two years. We include choice maker_{*n*} fixed effects α_n and District_{*i*}×Street_{*k*}×Year_{*t*} fixed effects and show OLS and IV estimates in Table 22.

Table 21: Descriptive statistics for (lagged) peer effects for Case Study of Baden-Wuerttemberg.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $\Delta y_{n,t}$.0026	.051	0	1
<hr/> <i>Installed base (lagged by $t - 2$):</i>				
$IB_{n,t-2}$	1.3	1.9	0	43
<hr/> <i>Instruments (lagged by $t - 2$):</i>				
Inclination of neighboring roofs $_{n,t-2}$.89	1.4	0	33
Orientation of neighboring roofs $_{n,t-2}$.96	1.5	0	37
<hr/>				
N	26,426,011			
<hr/>				
<i>Notes:</i> The 26,426,011 observations come from 2,954,352 choice makers (buildings) n over 9 years t (2002-2010). The choice makers are distributed across 8,982 districts i in Baden-Wuerttemberg. $\Delta y_{n,t}$ depends on $y_{n,t}^*$, which is defined in section A.4.				

The OLS and the IV results are in line. Here, we describe the IV estimates. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. Therefore, the instruments are strong (column (2)). In the second stage regression (column (3)), we use the predicted installed base from the first stage. As before, the effect of the instru-

Table 22: OLS and instrument variable estimates for (lagged) peer effects for Case Study of Baden-Wuerttemberg.

	OLS	IV	
		1st stage (OLS)	2nd stage (OLS)
	(1) $\Delta y_{n,t}$	(2) $IB_{n,t-2}$	(3) $\Delta y_{n,t}$
<i>Installed base (lagged by $t - 2$):</i>			
$IB_{n,t-2}$	0.00013** (3.72)		
<i>Predicted installed base (lagged by $t - 2$):</i>			
$\widehat{IB}_{n,t-2}$			0.00016** (4.37)
<i>Instruments (lagged by $t - 2$):</i>			
Inclination of neighboring roofs $_{n,t-2}$		0.45** (133.4)	
Orientation of neighboring roofs $_{n,t-2}$		0.77** (244.2)	
<i>Controls:</i>			
Potential electricity production per sqm $_n \times t$	0.017** (31.3)	-0.025** (-6.86)	0.017** (31.3)
Maximum module area $_n \times t$	0.00014** (4.67)	-0.0015** (-8.68)	0.00014** (4.68)
Observations	24,860,810	24,860,810	24,860,810
DF _M	3	4	3
Adj. R^2	0.06		0.00
F	334.5		336.3
Hansen J (p-value)			0.0 (0.92)
Kleibergen-Paap rk Wald F statistic			778,180.6
Year $_t$ fixed effects	Yes	Yes	Yes
Federal state $_n$ fixed effects	Yes	Yes	Yes
District $_i \times$ Year $_t$ fixed effects	Yes	Yes	Yes
District $_i \times$ Street $_k \times$ Year $_t$ fixed effects	Yes	Yes	Yes
Choice maker $_n$ fixed effects	Yes	Yes	Yes

Notes: Across columns, we show coefficients; t statistics in parentheses (based on SE clustered at street-level); * $p < 0.05$, ** $p < 0.01$

The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . All columns are estimated via OLS. Column (3) uses the predicted values for the installed base from the first stage in column (2). The panel of 2,954,352 choice makers over 10 years t (2002-2010) results in 29,543,520 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 26,426,011 observations. Across columns, the sample is always the same and singleton observations are dropped. $f(d_{n,m}) = 1$. The Cut-off distance D is 200m.

mented installed base on the adoption decision is strong and significant at all confidence levels. We again conduct an overidentifying restrictions test. The test supports the validity of our instruments. We can therefore confidently rule out that the ‘reflection problem’ described by Manski (1993) biases our results.

D.2 Robustness for visibility measures

Here, we focus on the robustness of our estimates for visibility measures. In Table 23, we include choice maker_{*n*} fixed effects α_n and District_{*i*}×Street_{*k*}×Year_{*t*} fixed effects and estimate linear probability models.

Table 23: Instrument variable estimates for peer effects visibility measure (next road 50m) without (columns (1)-(2)) and with distance-weighting (columns (3)-(4)) for Case Study of Baden-Wuerttemberg.

	IV visible, no distance-weighting		IV visible, with distance-weighting	
	1st stage (OLS)	2nd stage (OLS)	1st stage (OLS)	2nd stage (OLS)
	(1)	(2)	(3)	(4)
	$IB_{n,t-1}^{NR50}$	$\Delta y_{n,t}$	$IB_{n,t-1}^{NR50,1/d}$	$\Delta y_{n,t}$
<i>Predicted installed base:</i>				
Next road 50m: $\widehat{IB}_{n,t-1}^{NR50}$		0.00016** (5.12)		
<i>Instruments (next road 50m):</i>				
Inclination of neighboring roofs $_{n,t-1}^{NR50}$	0.46** (138.2)			
Orientation of neighboring roofs $_{n,t-1}^{NR50}$	0.77** (249.2)			
<i>Predicted installed base (distance-weighted):</i>				
Next road 50m: $\widehat{IB}_{n,t-1}^{NR50,1/d}$				0.0056** (3.42)
<i>Instruments (next road 50m, distance-weighted):</i>				
Inclination of neighboring roofs $_{n,t-1}^{NR50,1/d}$			0.46** (124.3)	
Orientation of neighboring roofs $_{n,t-1}^{NR50,1/d}$			0.76** (222.0)	
<i>Controls:</i>				
Potential electricity production per sqm _{<i>n</i>} × <i>t</i>	-0.029** (-7.89)	0.014** (32.1)	-0.0017** (-25.0)	0.014** (32.1)
Maximum module area _{<i>n</i>} × <i>t</i>	-0.0016** (-9.15)	0.00012** (4.86)	-0.000050** (-19.1)	0.00013** (4.88)
Observations	27,643,514	27,643,514	27,643,514	27,643,514
DF _M	4	3	4	3
Adj. R ²		0.00		0.00
F		354.3		349.1
Hansen J (p-value)		0.1 (0.73)		0.4 (0.50)
Kleibergen-Paap rk Wald F statistic		833,257.0		811,436.5
Year _{<i>t</i>} fixed effects	Yes	Yes	Yes	Yes
Federal state _{<i>n</i>} fixed effects	Yes	Yes	Yes	Yes
District _{<i>i</i>} × Year _{<i>t</i>} fixed effects	Yes	Yes	Yes	Yes
District _{<i>i</i>} × Street _{<i>k</i>} × Year _{<i>t</i>} fixed effects	Yes	Yes	Yes	Yes
Choice maker _{<i>n</i>} fixed effects	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients; *t* statistics in parentheses (based on SE clustered at street-level); * $p < 0.05$, ** $p < 0.01$

The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker *n*'s observed choice in terms of (newly) installing a PV system in *t*. All columns are estimated via OLS. Column (2) uses the predicted values for the installed base from the first stage in column (1). Column (4) uses the predicted values for the installed base from the first stage in column (3). The panel of 2,903,730 choice makers over 10 years *t* (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. Across columns, the sample is always the same and singleton observations are dropped. In columns (1)-(2), $f(d_{n,m}) = 1$, and in columns (3)-(4), $f(d_{n,m}) = 1/d_{n,m}$. The Cut-off distance *D* is 200m.

In columns (1-2) of Table 23, we show that our results hold for instrumenting the installed base that only considers PV systems which are located at most 50 meters away from the closest street ($IB_{n,t-1}^{NR50}$). As instruments, we use the inclination and orientation of neighboring roofs that are also located at most 50 meters away from the closest street. Once again, the instruments are strong and the overidentifying restrictions test

confirms their validity. We obtain similar results for a distance-weighted installed base that focuses on PV systems which are located at most 50 meters away from the closest street (columns (3-4)). We do not show the results but they remain the same for non-instrumented ordinary least squares for $IB_{n,t-1}^{NR50}$, $IB_{n,t-1}^{NR50,1/d}$ and $IB_{n,t-1}^{NR50}$.

Next, we show the robustness of our results for instrumented installed base measures that rely on direct visibility between buildings. Again, we include choice maker_n fixed effects α_n and District_i×Street_k×Year_t fixed effects and estimate linear probability models in Table 24.

Table 24: Instrument variable estimates for peer effects visibility measure (visible from choice maker building) for all locations (columns (1)-(2)), for urban locations (columns (3)-(4)) and for rural locations (columns (5)-(6)) for Case Study of Baden-Wuerttemberg.

	All		Urban		Rural	
	1st stage (OLS)	2nd stage (OLS)	1st stage (OLS)	2nd stage (OLS)	1st stage (OLS)	2nd stage (OLS)
	(1) $IB_{n,t-1}^{\triangleleft}$	(2) $\Delta y_{n,t}$	(3) $IB_{n,t-1}^{\triangleleft}$	(4) $\Delta y_{n,t}$	(5) $IB_{n,t-1}^{\triangleleft}$	(6) $\Delta y_{n,t}$
<i>Predicted installed base (visible from n):</i>						
No building in-between: $\widehat{IB}_{n,t-1}^{\triangleleft}$		0.00037** (3.43)		0.00033** (2.79)		0.00047 (1.65)
<i>Instruments (visible from n):</i>						
Inclination of neighboring roofs $_{n,t-1}^{\triangleleft}$	0.40** (163.7)		0.40** (146.7)		0.42** (74.2)	
Orientation of neighboring roofs $_{n,t-1}^{\triangleleft}$	0.78** (353.8)		0.79** (321.9)		0.77** (148.4)	
<i>Controls:</i>						
Potent. electr. product. per sqm _n × t	0.0051** (5.84)	0.014** (32.0)	0.0049** (5.05)	0.013** (27.2)	0.0077** (3.95)	0.019** (15.5)
Maximum module area _n × t	0.00025** (5.07)	0.00012** (4.76)	0.00026** (4.95)	0.00019** (6.69)	0.00029* (2.18)	-0.00018* (-2.46)
Observations	27,643,514	27,643,514	21,879,866	21,879,866	5,530,110	5,530,110
DF _M	4	3	4	3	4	3
Adj. R ²		0.00		0.00		0.00
F		349.7		260.5		85.1
Hansen J (p-value)		0.6 (0.45)		1.3 (0.25)		0.2 (0.63)
Kleibergen-Paap rk Wald F statistic		2,255,656.4		1,850,294.5		419,124.7
Year _t fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Federal state _n fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District _i × Year _t fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District _i × Street _k × Year _t fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Choice maker _n fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients; *t* statistics in parentheses (based on SE clustered at street-level); * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker *n*'s observed choice in terms of (newly) installing a PV system in *t*. All columns are estimated via OLS. Column (2) uses the predicted values for the installed base from the first stage in column (1). Column (4) uses the predicted values for the installed base from the first stage in column (3). Column (6) uses the predicted values for the installed base from the first stage in column (5). The panel of 2,903,730 choice makers over 10 years *t* (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. In columns (1)-(2), we include the full sample, in columns (3)-(4), we focus on urban locations, in columns (5)-(6) on rural locations. Across columns, singleton observations are dropped, $f(d_{n,m}) = 1$ and the Cut-off distance *D* is 200m.

The measure $IB_{n,t-1}^{\triangleleft}$ contains all PV systems visible from *n* and $IB_{n,t-1}^{\triangleleft}$ contains all PV systems not visible from *n* (because another building lies in-between). We instrument the installed base that contains all PV systems visible from *n* ($IB_{n,t-1}^{\triangleleft}$) with the inclination and orientation of neighboring roofs that contains all PV systems visible from *n* and then again distinguish urban and rural locations. Columns (1-2) indicate a causal

peer effect for this visibility measure in general. However, we can only confirm a causal peer effect for $IB_{n,t-1}^{\triangleleft}$ at urban locations (see Table 24, columns (3-4)). For urban locations, the instruments are strong and the overidentifying restrictions test confirms their validity. In contrast, we cannot confirm a causal peer effect based on the installed base that contains all PV systems visible from n ($IB_{n,t-1}^{\triangleleft}$) for rural locations. The estimated coefficient for the predicted installed base in columns (6) is statistically insignificant.

We illustrate the robustness of our results for the installed base on PV systems that are not directly visible from the choice maker's building. Table 25 shows odds ratios from conditional logit (columns (2) and (4)). In columns (1), (3) and (5) we include estimates with choice maker $_n$ fixed effects α_n and District $_i \times$ Street $_k \times$ Year $_t$ fixed effects from the linear probability model.

Table 25: Estimates for peer effects visibility measure (not visible from choice maker building) for all locations (column (1)), for urban locations (columns (2)-(3)) and for rural locations (columns (4)-(5)) for Case Study of Baden-Wuerttemberg.

	All	Urban		Rural	
	OLS	CLogit	OLS	CLogit	OLS
	(1) $\Delta y_{n,t}$	(2) $\Delta y_{n,t}$	(3) $\Delta y_{n,t}$	(4) $\Delta y_{n,t}$	(5) $\Delta y_{n,t}$
<i>Installed base (not visible from n):</i>					
Building in-between: $\widehat{IB}_{n,t-1}^{\triangleleft}$	-0.000033 (-0.74)	1.00 (-0.28)	-0.000023 (-0.48)	0.99 (-0.33)	-0.000082 (-0.57)
<i>Controls:</i>					
Potential electricity production per sqm $_n$		4.7E+7** (9.24)		4.4E+7** (5.15)	
Potential electricity production per sqm $_n \times t$	0.014** (32.0)	0.90 (-0.48)	0.013** (27.2)	0.88 (-0.32)	0.019** (15.5)
Maximum module area $_n$		1.66** (2.66)		0.68 (-1.27)	
Maximum module area $_n \times t$	0.00012** (4.78)	0.98 (-0.79)	0.00019** (6.71)	1.02 (0.63)	-0.00018* (-2.46)
Observations	27,643,514	580,802	21,879,866	134,920	5,530,110
DF $_M$	3	5	3	5	3
Adj. R^2	0.06		0.06		0.05
F	346.5		258.4		84.4
Hansen J (p-value)					
Kleibergen-Paap rk Wald F statistic					
Year $_t$ fixed effects	Yes	Yes	Yes	Yes	Yes
Federal state $_n$ fixed effects	Yes	Yes	Yes	Yes	Yes
District $_i \times$ Year $_t$ fixed effects	Yes	Yes	Yes	Yes	Yes
District $_i \times$ Street $_k \times$ Year $_t$ fixed effects	Yes	Yes	Yes	Yes	Yes
Choice maker $_n$ fixed effects	Yes	No	Yes	No	Yes

Notes: In columns (1), (3) and (5), we show coefficients, in columns (2) and (4) exponentiated coefficients; t statistics in parentheses (based on SE clustered at street-level; in columns (2) and (4), we show robust t statistics); * $p < 0.05$, ** $p < 0.01$

The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . The panel of 2,903,730 choice makers over 10 years t (2001-2010) results in 29,037,300 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,383,343 observations. Columns (1), (3) and (5) are estimated via OLS. We estimate columns (2) and (4) with the conditional logit estimator. Note that the conditional logit estimator drops all positive (or all negative) outcomes in terms of District $_i \times$ Street $_k \times$ Year $_t$ groups. I.e., District $_i \times$ Street $_k \times$ Year $_t$ groups with no adoption (or if all choice makers in that group adopt at once) are dropped. Similarly, singleton observations are dropped in columns (1), (3) and (5). These procedures result in fewer observations. In column (1), we include the full sample, in columns (2)-(3), we focus on urban locations, in columns (4)-(5) on rural locations. Across columns, $f(d_{n,m}) = 1$ and the Cut-off distance D is 200m.

No matter if we focus on all, on urban or on rural locations, across specification, the estimates for $\widehat{IB}_{n,t-1}^{\triangleleft}$ are statistically insignificant. In summary, there is less evidence

for peer effects from PV systems that are not directly visible from the choice maker's building.

Finally, we show that our results are robust to including all visibility measures in one regression. As before, we see in column (1) of Table 26 that the estimated coefficients for $IB_{n,t-1}^{NR50}$ and for $IB_{n,t-1}^{\triangleleft}$, which refers to visible PV systems, are significantly positive and similar in of similar size. In contrast, the estimates for $IB_{n,t-1}^{NR50}$ $IB_{n,t-1}^{\ntriangleleft}$, which cover non-visible system, are insignificant. The regressions include choice maker_n fixed effects α_n and District_i×Street_k×Year_t fixed effects. These results remain unaffected from instrumenting $IB_{n,t-1}^{NR50}$ (see columns (2) and (3)) and $IB_{n,t-1}^{\triangleleft}$ (see columns (4) and (5)) with their respective measure of the inclination and orientation of neighboring roofs. In columns (2-5), we also include the not instrumented visibility measures in the regressions. The specifications in Table 26 also reveal that proximity is not the real channel driving our results since (Next road 50m) $IB_{n,t-1}^{NR50}$ and (No road 50m:) $IB_{n,t-1}^{NR50}$ add up to $IB_{n,t-1}$. The significantly positive estimate for (No building in-between:) $IB_{n,t-1}^{\triangleleft}$ is the peer effect from visible previously installed PV systems nearby (while controlling for all previously installed PV systems nearby).

Table 26: OLS and instrument variable estimates for all visibility measures included at once for Case Study of Baden-Wuerttemberg.

	OLS	IV (next road 50m)		IV (no building in-between)	
		1st stage (OLS)	2nd stage (OLS)	1st stage (OLS)	2nd stage (OLS)
	(1) $\Delta y_{n,t}$	(2) $IB_{n,t-1}^{NR50}$	(3) $\Delta y_{n,t}$	(4) $IB_{n,t-1}^{\triangleleft}$	(5) $\Delta y_{n,t}$
<i>Installed base (visibility measures):</i>					
Next road 50m: $IB_{n,t-1}^{NR50}$	0.00013** (4.12)			0.0043** (51.1)	0.00013** (4.17)
No road 50m: $IB_{n,t-1}^{NR50}$	0.00058 (1.48)	0.0076 (0.90)	0.00058 (1.48)	0.0039** (3.94)	0.00058 (1.48)
No building in-between: $IB_{n,t-1}^{\triangleleft}$	0.00035** (3.37)	0.036** (40.4)	0.00033** (3.18)		
Building in-between: $IB_{n,t-1}^{\triangleleft}$	-0.000074 (-1.52)	0.062** (67.6)	-0.000100* (-2.04)	-0.011** (-71.0)	-0.000078 (-1.61)
<i>Predicted installed base (visibility measures):</i>					
Next road 50m: $\widehat{IB}_{n,t-1}^{NR50}$			0.00017** (5.17)		
No building in-between: $\widehat{IB}_{n,t-1}^{\triangleleft}$					0.00031** (2.80)
<i>Instruments (next road 50m):</i>					
Inclination of neighboring roofs $_{n,t-1}^{NR50}$		0.45** (138.5)			
Orientation of neighboring roofs $_{n,t-1}^{NR50}$		0.76** (249.2)			
<i>Instruments (visible from n):</i>					
Inclination of neighboring roofs $_{n,t-1}^{\triangleleft}$				0.40** (163.7)	
Orientation of neighboring roofs $_{n,t-1}^{\triangleleft}$				0.78** (353.5)	
<i>Controls:</i>					
Potential electricity production per sqm $_n \times t$	0.014** (32.0)	-0.027** (-7.47)	0.014** (32.0)	0.0049** (5.68)	0.014** (32.0)
Maximum module area $_n \times t$	0.00012** (4.79)	-0.0014** (-8.10)	0.00012** (4.81)	0.00023** (4.72)	0.00012** (4.79)
Observations	27,643,514	27,643,514	27,643,514	27,643,514	27,643,514
DF _M	6	7	6	7	6
Adj. R^2	0.06		0.00		0.00
F	178.5		179.9		178.0
Hansen J (p-value)			0.1 (0.75)		0.6 (0.44)
Kleibergen-Paap rk Wald F statistic			823,268.5		2,231,494.1
Year _t fixed effects	Yes	Yes	Yes	Yes	Yes
Federal state _n fixed effects	Yes	Yes	Yes	Yes	Yes
District _i \times Year _t fixed effects	Yes	Yes	Yes	Yes	Yes
District _i \times Street _k \times Year _t fixed effects	Yes	Yes	Yes	Yes	Yes
Choice maker _n fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients; t statistics in parentheses (based on SE clustered at street-level); * $p < 0.05$, ** $p < 0.01$. The dependent variable is always $\Delta y_{n,t}$, i.e., choice maker n 's observed choice in terms of (newly) installing a PV system in t . All columns are estimated via OLS. Column (3) uses the predicted values for the installed base (next road 50m) from the first stage in column (2). Column (5) uses the predicted values for the installed base (no building in-between) from the first stage in column (4). The panel of 2,956,938 choice makers over 10 years t (2001-2010) results in 29,569,380 observations. Across columns, we censor the data in the way that we ignore choice makers once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,145,636 observations. Across columns, the sample is always the same and singleton observations are dropped. $f(d_{n,m}) = 1$ and the Cut-off distance D is 200m.

References Online-Appendix

- Agnolucci, Paolo (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, and C. Theobald (2008). *EEG: Erneuerbare-Energien-Gesetz; Kommentar*. 2nd. Munich: Beck.
- Ben-Akiva, M. and S.R. Lerman (1985). *Discrete choice analysis, theory and applications to travel demand*. Cambridge, MA: MIT Press.
- Bollinger, Bryan and Kenneth Gillingham (2012). “Peer Effects in the Diffusion of Solar Photovoltaic Panels”. *Marketing Science* 31(6), 900–912. DOI: 10.1287/mksc.1120.0727.
- BSW-Solar (2014). Tech. rep. Acc. Oct. 13, 2014. Bundesverband Solarwirtschaft e.V. URL: http://www.solarwirtschaft.de/fileadmin/media/pdf/2014_03_BSW_Solar_Faktenblatt_Solarwaerme.pdf.
- Chintagunta, Pradeep K. and Harikesh S. Nair (2011). “Discrete-Choice Models of Consumer Demand in Marketing”. *Marketing Science* 30(6), 977–996. URL: <http://www.jstor.org/stable/41408413>.
- CLC (2009). “CORINE Land Cover from the year 2006, Umweltbundesamt, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Deutsches Fernerkundungsdatenzentrum (DFD)”. Database. Accessed June 13, 2012. Weßling. URL: http://www.corine.dfd.dlr.de/datadescription_2006_de.html.
- Comin, Diego and Johannes Rode (2013). *From Green Users to Green Voters*. Working Paper #19219 (Revised version November 2015). National Bureau of Economic Research. DOI: 10.3386/w19219.
- Correia, Sergio (2014). *REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects, revised 18 Nov 2019,*. Tech. rep. Statistical Software Components S457874, Boston College Department of Economics. URL: <https://ideas.repec.org/c/boc/bocode/s457874.html>.

- Correia, Sergio (2018). *IVREGHDFE: Stata module for extended instrumental variable regressions with multiple levels of fixed effects*. Tech. rep. Statistical Software Components S458530, Boston College Department of Economics. URL: <https://ideas.repec.org/c/boc/bocode/s458530.html>.
- DESTATIS (2016).
- Dewald, Ulrich and Bernhard 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. Offenbach. URL: <http://www.dwd.de/>.
- Farrell, Joseph and Garth Saloner (1986). “Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation”. *American Economic Review* 76(5), 940–955. URL: <http://www.jstor.org/stable/1816461>.
- Geofabrik (2018). *Download OpenStreetMap data for this region: Germany, Baden-Württemberg*. Database. Acc. and downloaded Jan. 15, 2018. Geofabrik. URL: <http://download.geofabrik.de/europe/germany.html>.
- Gowrisankaran, Gautam and Joanna Stavins (2004). “Network Externalities and Technology Adoption: Lessons from Electronic Payments”. *The RAND Journal of Economics* 35(2), 260–276. URL: <http://www.jstor.org/stable/1593691>.
- Graziano, Marcello, Maurizio Fiaschetti, and Carol Atkinson-Palombo (2019). “Peer effects in the adoption of solar energy technologies in the United States: An urban case study”. *Energy Research & Social Science* 48. DOI: 10.1016/j.erss.2018.09.002.
- Graziano, Marcello and Kenneth 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.
- I-TSO (2012). “EEG-Anlagenstammdaten”. Informationsplattform der Deutschen Übertragungsnetzbetreiber, Acc. Mar. 4, 2012. URL: <http://www.eeg-kwk.net/de/Anlagenstammdaten.htm>.

- Infas (2009a). Database. Bonn. URL: <http://www.infas-geodaten.de/>.
- (2009b). “Wohnquartiere, Kreis-Gemeinde-Schlüssel (KGS) 22, 2009”. Database. Bonn. URL: <http://www.infas-geodaten.de/>.
- Jarvis, A. et al. (2008). “Hole-filled seamless Shuttle Radar Topographic Mission (SRTM) data V4, International Centre for Tropical Agriculture (CIAT)”. Acc. July 30, 2012. URL: <http://srtm.csi.cgiar.org>.
- LUBW (2016). *Ermitteltes Solarpotential auf Dachflächen*. Last visit: 03.09.2018. Data freely available upon personal request. Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg. URL: <https://www.energieatlas-bw.de/sonne/dachflachen/potenzial-dachflachenanlagen>.
- Manski, Charles F. (1977). “The structure of random utility models”. *Theory and Decision* 8(3), 229–254. DOI: 10.1007/BF00133443.
- (1993). “Identification of Endogenous Social Effects: The Reflection Problem”. *The Review of Economic Studies* 60(3), 531–542. URL: <http://www.jstor.org/stable/2298123>.
- McFadden, D. (2001). “Economic choices”. *American Economic Review* 91(3), 351–378. DOI: 10.1257/aer.91.3.351.
- Müller, Sven and Johannes 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, Daniel, Colleen Dawes, and Varun 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.
- Rai, Varun and Scott A Robinson (2013). “Effective information channels for reducing costs of environmentally-friendly technologies: evidence from residential PV markets”. *Environmental Research Letters* 8(1), 014044. URL: <http://stacks.iop.org/1748-9326/8/i=1/a=014044>.

- Richter, Laura-Lucia (2013). *Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK*. Tech. rep. EPRG Working Paper 1332. Energy Policy Research Group. URL: <http://www.eprg.group.cam.ac.uk/eprg1332/>.
- Rode, Johannes (2014). “Renewable Energy Adoption in Germany – Drivers, Barriers and Implications.” Doctoral thesis. Darmstadt: TU Darmstadt.
- Rode, Johannes and Alexander 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.
- SOEP (2013). *Socio-Economic Panel, data for years 1984-2012*. version 29. DOI: 10.5684/soep.v29.
- Stark, Thomas et al. (2005). *Architektonische Integration von Photovoltaik-Anlagen*. Tech. rep.
- StataCorp (2015). *Stata Statistical Software: Release 14*. StataCorp LP. College Station, TX.
- Train, Kenneth E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.