I spot, I adopt!

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

Johannes Rode, Sven Müller *

March 8, 2021

Abstract

We study peer effect variation 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 Baden-Württemberg, 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: measures for their average inclination and orientation. Using different models of adoption with panel data, we only find evidence for causal peer effects for visible systems. We show that visible PV systems cause an increase in the probability of installing which is around 8 times higher in comparison to not visible PV systems.

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

*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; orcid: 0000-0002-3134-0759;
Sven Müller: Otto-von-Guericke-University Magdeburg, Center for Operations Research and Business Analytics, Universitätsplatz 2, 39106 Magdeburg, Germany, email: sven.mueller@ovgu.de, orcid: 0000-0002-3406-2133

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 (Bass 1969; 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 Baden-Württemberg, 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). Within Germany, Baden-Württemberg is among the regions with most PV systems.

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 government – active in the energy market are interested in understanding the PV adoption process.

Both economic and ecological concerns impact the choice behavior of households towards PV adoption (Wittenberg, Blöbaum, and Matthies 2018; Islam and Meade 2013). The purchase of a PV system by a household is associated with uncertainties regarding actual energy savings, power production, remuneration, to name but a few. Households may reduce such uncertainties by gathering information from various sources.

One way to reduce uncertainties is information transfer from previous adopters. Households update their beliefs about costs and benefits of PV adoption after discussing it with previous adopters or observing the outcomes of adoption. The impact of former adopters on the adoption decision of a household is denoted as social contagion or *peer effects*. Van den Bulte and Lilien (2001) describe the adoption behavior as a function of the actor's exposure to other actors' knowledge, attitude, or behavior towards an innovation. Besides information transfer, other theoretical mechanisms such as social normative pressures, competitive concerns, and performance network effects also result in peer effects (Van den Bulte and Lilien 2001; Risselada, Verhoef, and Bijmolt 2014).

The existence of peer effects is widely accepted in the literature (Conley and Udry 2010; Manchanda, Y. Xie, and Youn 2008; Katona, Zubcsek, and Sarvary 2011). Bollinger and Gillingham (2012), Müller and Rode (2013), and Graziano and Gillingham (2015) were among the first to study peer effects in the adoption process of PV systems. There is evidence on the relevance of social interactions between consumers on purchase decisions. The opportunities to exploit these interactions as marketing instruments have increased while the effectiveness of traditional marketing instruments has declined (Risselada, Verhoef, and Bijmolt 2014). Therefore, it is of general interest to shed light on the processes and structures underlying the observed peer effects (Godes 2011).

In this paper, we focus on information transfer as the theoretical concept of peer effects in PV adoption. We denote peers as proximate households who adopted PV in preceding periods. Gardete (2015) assumes that consumers often pay attention to the decisions of other consumers before making their own choice, through either direct observation or indirect observation. In the context of PV adoption, a direct observation is considered as a neighbor's PV system while indirect observation would comprise total PV power production, online ratings etc. We focus on direct observation and the corresponding peer effect on households' adoption choices.

As McShane, Bradlow, and Berger (2012) suggest, researchers should use pinpointed addresses to track how the effects of an adoption propagate throughout a neighborhood. By doing so, the ties between adopters can be measured by geographical distance (Risselada, Verhoef, and Bijmolt 2014). We follow McShane, Bradlow, and Berger (2012) and further break down these ties by *visibility*. From a given household's location an existing PV system nearby might be directly visible or not. We consider a PV system to be visible if no building lies within the line-of-sight between the household's location and the PV system *and* the PV system is exposed to the household's location (Section Visibility Measures). Besides distance, our visibility measure works as an additional dosage variation of peer influence. The dosage of peer influence is assumed to be higher for visible, close-by PV systems compared to non-visible PV systems nearby. A proximate and visible PV system may make a potential adopter more aware of – and thus more likely to adopt – the new technology. The every-day visual trigger and the generally low burden of interaction with a proximate neighbor yields a high propensity of information transfer (Section Installed Base). As such, we disentangle the impact of visibility and, as alternative hypothesis, word-of-mouth on peer effects (Chen, Wang, and J. Xie 2011).

Our granular data covers all potential and actual PV adopters up to the year 2010 in the State of Baden-Württemberg, Germany with approximately 3 million buildings (Section Data). We analyze the visibility of a PV system from the location of any potential adopter (any building). We deepen the understanding of peer effects in technology adoption in general and in PV adoption in particular. Neither Bollinger and Gillingham (2012) nor Graziano and Gillingham (2015), who have studied peer effects in PV adoption before, have used individual-level data with direct visibility measures.[1] Instead they consider zones or streets as their observational unit. In contrast, Iyengar, Jedidi, et al. (2011) and Akerlof (1997) point out that a deeper understanding of peer effects demands individual-level adoption data.

Even if individual-level data and direct measures of visibility are available, as in our case, the identification of causal peer effects remains challenging, because many competing forces (promotions, for instance) and unobserved adopter characteristics are likely to hinder a thorough identification of social effects (Atkin et al. 2017). Endogenous group formation (homophily), correlated unobservables, simultaneity, and endogeneity may bias the analysis. Fortunately, tailored methods exist to tackle these sources of bias. Simultaneity (Manski's reflection problem) occurs in situations where the adoption decision of an actor depends on others in the actor's reference group. A household's adoption may affect other proximate households and the direction of impact may not be clear. As out-

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

lined by Manski (1993), panel data helps to control for this source of bias. If we observe a new PV system at a given location and year, we may assume that the corresponding household also made the decision to adopt in this year or in the preceding year. Then, the actual contagion took place in the years before. Nair, Manchanda, and Bhatia (2010) argue that a time lag is expected to control for the impact of the actor on her reference group (Section Installed Base).

We expect that unobserved exogenous variables are correlated with PV adoption and so are our error terms. This again may yield biased estimates. The sources of correlated unobservables may, for instance, stem from local advertisement campaigns by a PV seller, new local subsidies, housing developments in which new local regulations force residents to install PV, different propensities to adopt at different stages along the diffusion path, or media exposure. Unobserved household characteristics may yield a homophily bias: similar households tend to make similar decisions in terms of location and adoption choice. Correlated unobservables may therefore create a sham peer effect. To control for the mean effect of correlated unobservables, we consider a rich pattern of fixed effects (Nair, Manchanda, and Bhatia 2010; Van den Bulte and Lilien 2001). We include year fixed effects, granular spatial fixed effects, combinations of year and granular spatial fixed effects, and household-specific fixed effects (Section Fixed Effects).

To identify peer effects in PV adoption, we seek to exclude bias from endogeneity. Our measure of peer effects – the *installed base* (Farrell and Saloner 1986; Narayanan and Nair 2013) – may be endogenous (Nair, Manchanda, and Bhatia 2010; McShane, Bradlow, and Berger 2012; Risselada, Verhoef, and Bijmolt 2014). In fact, unobservables may induce an increase in PV installations. Our estimates may therefore be biased through a correlation between the error term and the *installed base*. We use measures of the average roof inclination and orientation over previous installations nearby as exogenous variation in photovoltaic system roof appropriateness. We assume that the construction of buildings is in almost all cases not driven by the installation of a roof-top PV system and we expect our instruments to be exogenous to the installation decision of a focal household (Section Instrumental Variables). 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 granular spatial fixed effects on the street-level, which control for potentially similar PV appropriateness (or potentially similar inclination and orientation) of roofs in the same street. Our results provide strong evidence for the validity of our instruments.

Our modeling approach comprises a high-dimensional fixed effects panel model with instrumental variables (Section Individual Adoption and Modeling Approaches). We consider two models of adoption. First, since the decision to purchase a PV system in a given period is made by an individual (household), an intuitive way to model the choice behavior (adopt or not adopt) is a binary-choice panel regression (Van den Bulte and Lilien 2001; Iyengar, Jedidi, et al. 2011; Müller and Rode 2013; Iyengar, Van den Bulte, and Lee 2015). Because we include high-dimensional fixed effects in conjunction with instrumental variables, we employ a linear probability model as suggested by Grinblatt, Keloharju, and Ikaheimo (2008), Gardete (2015), and Narayanan and Nair (2013). Since we are mainly interested in the inference on peer effects and less on predicting probabilities of adoption, the benefits in terms of reduced complexity of the linear probability model outweigh its cost in terms of predictive power (J. Wooldridge 2012). Second, we estimate a hazard model of adoption (Risselada, Verhoef, and Bijmolt 2014). We fit a proportional hazard model with a piece-wise exponential baseline. We include high-dimensional fixed effects. To deal with endogeneity, we again exploit the exogenous variation from measures of the average roof inclination and orientation over previous installations nearby in a control function. We follow the procedures for non-linear models suggested by Lin and J. M. Wooldridge (2019).

In line with Bollinger and Gillingham (2012), Graziano and Gillingham (2015), and Rode and Weber (2016), we find evidence for peer effects in PV adoption. Our instrumental variable approach indicates that peer effects in PV adoption are, in general, not causal. In contrast, our estimates based on exogenous variation from average roof inclination and orientation over previous installations indicate that peer effects from directly visible PV systems indeed drive PV adoption. Our measure for directly visible PV systems reveals an eight times higher increase in the probability of installing than our measure with all non-visible PV systems. A battery of robustness tests confirms our results.

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. Because visible PV systems induce adoption, the visible usage of other products and technologies has the potential to foster diffusion. Marketers should identify potential users that make the usage of a product or technology visible to others or encourage visibility by, for instance, choosing striking designs. Our results allow to come up with efficient marketing strategies by exploiting social interaction to increase sales and new technology diffusion. By focusing on visible actors marketers are likely to reduce scatter of their promotion budget.

Modeling

Individual Adoption and Modeling Approaches

We assume that a household, proxied by a building, faces the decision to install or not to install a PV system in a given period. We denote the building (or the building's owner) by n and the period by t. The endogenous variable

$$y_{n,t} = \begin{cases} 1 & \text{if } n \text{ chooses to adopt in } t \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

covers the actual and observable choice of PV technology adoption. Hence, the valid inequality $\sum_t y_{n,t} \leq 1 \; \forall \; n$ always holds. We observe in which period t building n (or the building's owner) *uses the technology for the first time* (i.e., adoption).

One intuitive way of modeling the adoption choice process would be by panel-logistic regression or discrete-time hazard models. However, identification issues yield a high-dimensional fixed effects model with instrumental variables. In particular, time-invariant fixed effects for building n in conjunction with instrument variables support the use of ordinary least squares (Bai 2009). William Greene (2004) finds that the linear probability model (LPM) allows for straightforward consideration of unobserved heterogeneity on the level of the observations n, which is not the case in non-linear models such as logistic regression. Caudill (1988) points to problems with individual-specific dummy variables (here, building fixed effects) in logistic regression as well. We employ a two stage estimation approach with instrument variables to account for endogeneity. This approach is straightforward and computationally cheap for LPM, while for non-linear models strong assumptions are made (see J. Wooldridge 2012, p. 472). Therefore, we employ linear probability models of the form

$$y_{n,t} = \boldsymbol{\beta}' \boldsymbol{X}_{n,t} + \epsilon_{n,t} \tag{2}$$

with exogenous variables $\boldsymbol{X}_{n,t}$ that are expected to impact the choice of n to (not) install a PV system in t, $y_{n,t}$. The corresponding effects are denoted by $\boldsymbol{\beta}$, i.e.,

$$\boldsymbol{\beta}' \boldsymbol{X}_{n,t} = \beta_0 + \beta_1 X_{n,t,1} + \beta_2 X_{n,t,2} + \cdots . \tag{3}$$

The error term $\epsilon_{n,t}$ is assumed to be iid normal, and hence the probability of adoption ($y_{n,t} = 1$) is given by

$$P(y_{n,t} = 1 \,|\, \boldsymbol{X}_{n,t}) = \beta_0 + \beta_1 X_{n,t,1} + \beta_2 X_{n,t,2} + \cdots , \tag{4}$$

which is a linear function of the exogenous variables $\boldsymbol{X}_{n,t}$. The LPM is given by (4) and the coefficients $\boldsymbol{\beta}'$ are estimated via ordinary least squares. The estimates $\hat{\boldsymbol{\beta}}'$ are unbiased and consistent (see J. Wooldridge 2012, pp. 248 for details). The linear probability model directly links the observed endogenous variable $y_{n,t}$ to the exogenous variables $\boldsymbol{X}_{n,t}$. Although LPM is intuitively not the first choice when modeling a binary dependent variable, it is often applied when high-dimensional fixed effects are considered (Bernard and Jensen 2004; Jimenez et al. 2014). Since our analyses comprise many observations, the computational simplicity of the LPM is advantageous. However, the convenience of the LPM comes at a cost: predicted values $\hat{y}_{n,t}$ are not bounded between zero and one, as demanded for probabilities. Because we are interested in inference, we can neglect this problem.[2] Further, the probability in (4) cannot be linearly related to the exogenous variables $\boldsymbol{X}_{n,t}$ for all possible values (Caudill 1988). Fortunately, the LPM approximates the true adoption probability fairly well for common values of the explanatory variables, i.e., values around the center of the distribution of $\boldsymbol{X}_{n,t}$ (J. Wooldridge 2002, pp. 454). Finally, the LPM suffers from heteroscedasticity by construction. We use robust standard errors to account for this problem. To increase confidence in our results, we illustrate the robustness of our LPM estimates with those from a logistic panel regression with fewer fixed effects and a hazard model of adoption.

Fixed Effects

A major challenge for the identification of peer effects is homophily: building owners living next to each other may make similar decisions not because of peer effects but because they share attitudes, characteristics, roof size or roof orientation, for instance. We lack information on most of these characteristics and attitudes. Instead, we consider high-dimensional fixed effects to account for unobserved heterogeneity (Nair, Manchanda, and Bhatia 2010; Van den Bulte and Lilien 2001). In particular, for every period t we consider the spatially invariant effect α_t. The temporal fixed effects α_t absorb time specific adoption shocks that could for instance be caused by changes in the subsidy system, buyback prices, asset cost of PV systems, or simply different propensities to adopt at different stages along the diffusion path. The time invariant, building specific fixed effects γ_n account for innate factors of building owner n towards the decision to adopt or not. Such factors might be attitudes, beliefs, or income, for example. To account for unobservables that apply to several buildings n in period t in the same way, in particular homophily, we consider fine grained spatial units K. Here, set K is defined as a combination of statistical districts and street segments. For a given $k \in K$ all buildings n located in k are denoted by the set N_k such that $\bigcup_k N_k = N$. The corresponding fixed effect is $\eta_{k,t}$. It allows us to control, for instance, for different k-specific impact from a subsidy by year or any other k-specific yearly adoption shock (on the net present value of PV systems, for example). Then, for

[2] If the LPM should be used for predictions, we refer to W.H. Greene (2012) for a bounded LPM.

all spatial units $k \in K$, buildings $n \in N_k$, and periods t, equation (3) becomes

$$\boldsymbol{\beta}'\boldsymbol{X}_{n,t} = \beta_0 + \beta_1 X_{n,t,1} + \beta_2 X_{n,t,2} + \cdots + \alpha_t + \gamma_n + \eta_{k,t}. \tag{5}$$

Fixed effects α_t, γ_n, and $\eta_{k,t}$ are eliminated by demeaning the variables using an appropriate transformation and the unbiased coefficients of interest $\boldsymbol{\beta}'$ are obtained by ordinary least squares (OLS). A detailed discussion of fixed effects estimators can be found in W.H. Greene (2012) and J. Wooldridge (2012).

Data

Buildings.– Each building in Baden-Württemberg is geocoded (longitude, latitude) and has a unique identifier (LUBW 2016). There are 2,957,332 buildings in Baden-Württemberg in 2019. For each building we have specific information on its roof.

PV systems.– We use public domain data of PV installations that contains the address and year of installation from I-TSO (2012).[3] Periods are in our case years $t = 2000$ and before, 2001,..., 2010.[4] We do not consider PV systems with more than 30 kW because such large systems are industrial systems mostly (Dewald and Truffer 2011). To merge building data and PV installations, we geocode the address data on PV installations. Geocoding accuracy is high for 83.8% of the systems. A high geocoding accuracy revers to the address or the street - census block level. Due 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 building-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. We successfully allocate 71,432 PV systems to unique roofs (denoted $y_{n,t}$). We take a conservative approach and neglect

[3] We 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.

[4] During our period of study, the subsidy system for PV mostly changed by year. At the same time, the cost of PV installation declined continuously. In consequence, most PV installations were installed at the end of a year.

PV systems in our baseline analyses if more than one PV system is allocated to the same roof. However, we also make use of (and show the robustness of our results for) $y_{n,t}^{\text{All}}$, which refers to all PV systems. In this case, the PV systems, which are not uniquely allocated are then randomly assigned to another roof in the same statistical district.[5] Table 5 in Appendix A.3 contains the frequency of PV installations in the baseline ($y_{n,t}$) by year.

Merging the building information and PV installations in the baseline yields $N =$ 29,580,810 observations. See Figure 1 for a small scale excerpt of the data. Buildings without a PV system are considered as potential adopters, while those with a PV systems are considered as actual adopters.

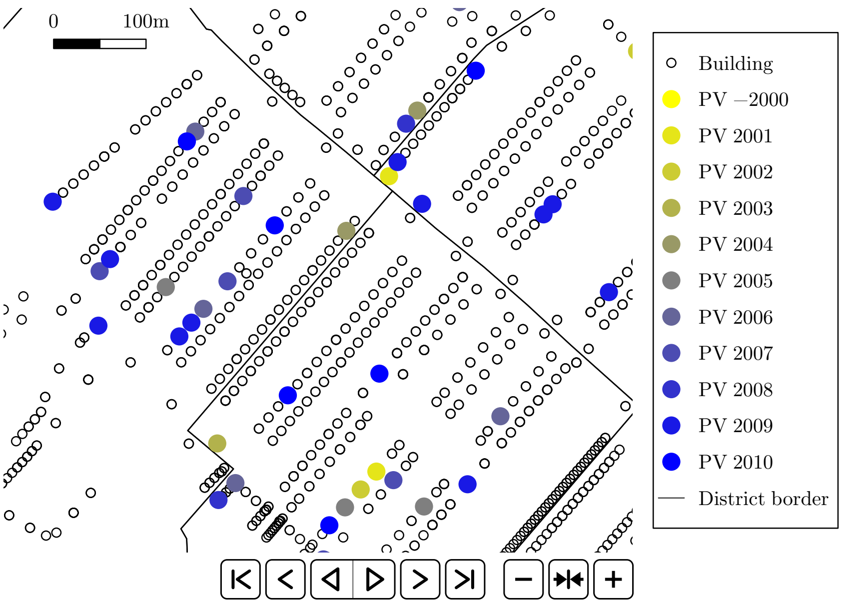


Figure 1: Map of selected districts, building locations, and PV installations in the city of Freiburg, Baden-Württemberg (Germany).

Notes: Hollow circles represent potential adopters $y_{n,t} = 0$. Filled circles are PV installations (actual adopters). 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.

Suitability of buildings for PV.– Our information on buildings comes from a roof census (LUBW 2016). The roof census includes 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,

[5] Statistical districts are granular regional entities. Our buildings are located across 8988 statistical districts in Baden-Württemberg. The average area of a statistical district is about 3.98 sq km and the average number of buildings per district is 374. For instance, the largest city in Baden-Württemberg Stuttgart (with population of about 600,000 people in 2010) consists of 656 statistical districts.

and roof-specific information on inclination and orientation. We normalize the potential electricity production by roof area and denote *potential electricity production per sq m of n* = potential electricity production of n / maximum module area of n. *Potential electricity production per sq m of n* is measured in MWh/a. *Maximum module area of n* is the roof area feasible for a PV system in sq m.

Granular spatial fixed effects for each year.– The state of Baden-Württemberg is partitioned into 8,988 statistical districts. We further obtain data on street segments from Geofabrik (2018). There are 544,243 street segments in our data set. Based on districts and streets, we construct granular spatial units K as spatial intersections of street polygons and district polygons. We end up with 579,832 District-Street units $k \in K$. Buildings $n \in N_k$ are located in the same District-Street unit (Figure 2). In our baseline data set, the average cardinality of N_k is 5.6. We interact District-Street units k with periods t, which results in 5,785,748 District-Street-Year fixed effects $\eta_{k,t}$.

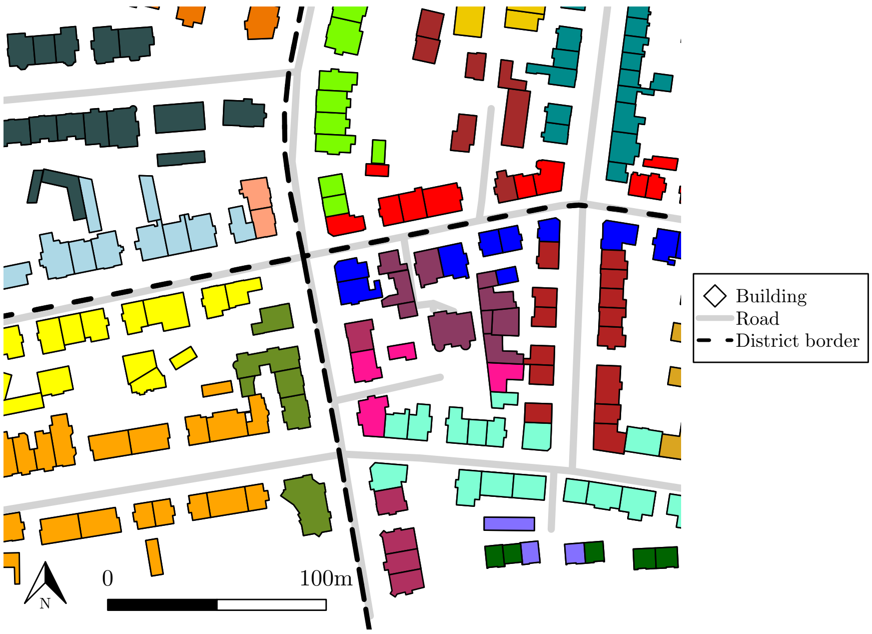


Figure 2: Illustration of District-Street-Year fixed effects $\eta_{k,t}$ for a given year t.

Notes: The map illustrates District-Street units for buildings n. Buildings that belong to the same District-Street unit are colored identical. In this example, we have 4 different districts and 9 streets, which results in 21 District-Streets k because longer streets are sliced by district borders.

Installed Base

Generally, the peer effect is the impact of former adopters on adoption choices of actors. Adoption of building owner n in period t, $y_{n,t}$, depends on adoptions in previous periods $t' < t$, i.e., $y_{m,t'}$ ($t' = t-1$, for instance) with $m \neq n$. As suggested by Risselada, Verhoef, and Bijmolt (2014), ties between adopters should be measured by small scale distances. In particular, McShane, Bradlow, and Berger (2012) point to the importance of small scale neighborhoods to trickle down the peer effect in technology adoption. We therefore assume that the closer a former adopter m the larger is the impact on n's choice to adopt. To identify peer effects in PV adoption, we introduce the spatio-temporal installed base measure (Farrell and Saloner 1986)

$$\mathrm{IB}_{n,t} = \sum_{m \in \bar{N}_n} \sum_{l=2000}^{t} f\left(y_{m,l}^{\mathrm{All}},\ d_{n,m}\right) \tag{6}$$

which accounts for the weighted number of installed PV systems. The set $\bar{N}_n = \{m \in N \,|m \neq n, d_{n,m} < D\}$ with $d_{n,m}$ as the euclidean distance between n and m in meters. The parameter D is a cut-off parameter, denoting the distance up to which locations $m \in \bar{N}_n$ are considered as potential peers. Defining potential peers on geographic distance only is of course limited. Due to missing social measures and a strong positive correlation between geographic distance and social ties as outlined by Hipp, Faris, and Boessen (2012), we are confident that $d_{n,m}$ is a sufficient proxy for peer ties. However, the setting of D is up to the analyst and as such somewhat ad hoc. Hipp and Perrin (2009) find that the probability of strong ties drops from over 0.9 to 0.5 within the first mile between two households. Bailey et al. (2018) find an even stronger effect: the social connectedness index drops from 100 to 1 within the first mile. Hence, selecting $D = 100, 200, 400$ meters seem reasonable. The function $f\left(y_{m,l}^{\mathrm{All}},\ d_{n,m}\right)$ represents a spatial weighting of prior installations, for instance a distance decay can be represented by $f\left(y_{m,l}^{\mathrm{All}},\ d_{n,m}\right) = y_{m,l}^{\mathrm{All}}/d_{n,m}$. In the baseline, we stick to the simplest form without distance decay, $f\left(y_{m,l}^{\mathrm{All}},\ d_{n,m}\right) = y_{m,l}^{\mathrm{All}}/1$. Further, (6) is flexible in terms of the size of the temporal lag: we can consider all previous periods or just a subset of previous periods up to $t-1$, for instance.

This approach allows to account for simultaneity issues (Manski 1993). The reflection problem (simultaneity) 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. A lagged installed base $\mathrm{IB}_{n,t-1}$ is very likely to rule out the possibility that the individual affected the adoption decision of reference group members, which in turn influenced the individual's adoption (Iyengar, Van den Bulte, and Lee 2015; Iyengar, Van den Bulte, and Valente 2011).

We follow Nair, Manchanda, and Bhatia (2010) and Van den Bulte and Stremersch (2004) and also construct the installed base relative to all potential and actual adopters denoted as

$$\mathrm{IB}^{\text{relative}}_{n,t} = \frac{\mathrm{IB}_{n,t}}{|\bar{N}_n|}, \tag{7}$$

with $\bar{N}_n \neq \emptyset$. Table 1 contains the descriptive statistics of (6) and (7).

Visibility Measures

We are interested in breaking down the dosage variation of peer influence besides geographical distance. We consider whether an existing PV installation is visible from a potential adopter's location. Duntley (1948) finds that an object of 25 square meter is visible from a maximum of 750 meter if the person stands on the ground.[6] This result indicates that geographic distance does not hinder visibility in our study. Further, we consider a PV system at location m to be visible to n if both of the following conditions hold:

1. m is located in the 90° north view-shed of n: $m \in N_n^{90^\circ}$,

2. no building lies within the line-of-sight between n and m: $m \in N_n^{\text{los}}$.

Ad 1) Baden-Württemberg (Germany) is located in the northern hemisphere. Most radiation is expected on the south-side of a building (Figure 3). Hence, we assume that PV systems are almost always installed with exposition to the south. Therefore, we assume

[6] Duntley (1948) refers to liminal visibility. Liminal means that a human – if forced to judge – is more likely to distinguish the object from the background than to classify background and object to be the same.

Table 1: Descriptive statistics.

	Mean	Std. Dev.	Min.	Max.
Panel A: Baseline				
New PV installation: $y_{n,t}$	.0024	.049	0	1
Installed base:				
$\text{IB}_{n,t-1}$	.015	.023	0	.44
Instruments (average ratio over previous installations nearby):				
Avg. inclination ratio: $\text{AvgIncRatio}_{n,t-1}$	.004	.004	0	.01
Avg. orientation ratio: $\text{AvgOrRatio}_{n,t-1}$	.0043	.0043	0	.01
Panel B: Visibility and relative measures				
Installed base (direct visibility):				
No building in-between: $\text{IB}^{\triangleleft}_{n,t-1}$	.00046	.0022	0	.07
No building in-between (rel.): $\text{IB}^{\text{relative}\triangleleft}_{n,t-1}$	.00057	.0044	0	.67
Installed base (no direct visibility):				
Building in-between: $\text{IB}^{\ntriangleleft}_{n,t-1}$	.0032	.0075	0	.34
Rest: $\text{IB}^{\text{Rest}}_{n,t-1}$	.012	.019	0	.4
Building in-between: $\text{IB}^{\text{relative},\ntriangleleft}_{n,t-1}$	.0028	.0079	0	.6
Rest: $\text{IB}^{\text{Rest, relative}}_{n,t-1}$	.011	.023	0	1
Installed base (relative):				
$\text{IB}^{\text{relative}}_{n,t-1}$	.015	.026	0	1
Instruments (average ratio over previous visible PV installations nearby):				
Avg. inclination ratio: $\text{AvgIncRatio}^{\triangleleft}_{n,t-1}$	.00031	.0016	0	.01
Avg. orientation ratio: $\text{AvgOrRatio}^{\triangleleft}_{n,t-1}$	.00034	.0017	0	.01
N	29,380,453			

Notes: The 29,380,453 observations come from 2,957,332 buildings n over 10 years t (2001-2010). In total, we have 29,580,810 observations. We censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. The buildings are distributed across 5,785,748 District-Street-Year groups in Baden-Wuerttemberg. Note that an increase of the installed base $(IB_{n,t-1})$ of 0.01 refers to 1 additional previously installed PV system nearby.

a PV systems on a building m, which is located north to n is visible to n. We employ the concept of view-shed and only expect PV systems within a 90° angle as visible (Llobera 2003). Consider, locations (a) and (c) of potential adopters in Figure 4. The closest PV installations are located to the north and are expected to be visible from (a) and (c), respectively. In contrast, potential adopter (b) is located to the north of the most proximate PV system. We consider this PV system not visible to (b). Of course, some PV systems may be installed exposed to west or east and may be visible from other directions. We take a conservative approach and only consider PV systems visible if this is most likely the case.

Ad 2) Of course, if there are sufficiently large obstacles within the line-of-sight (straight line) between two locations n and m, the PV system at m is not visible to n. Unfor-

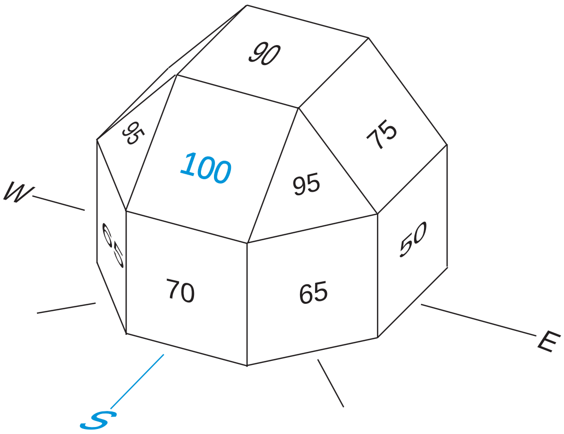


Figure 3: Roof orientation and inclination shape the PV appropriateness of a roof (Stark et al. 2005, p. 20).

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°. 0° inclination refers to a flat roof. 100% refers to the maximum radiation potential.

tunately, we do not have terrain and vegetation data. However, we obtain the shape of buildings. We expect that buildings are mainly sufficiently large to hinder visibility. We assume that if there is a building in the line-of-sight, the PV system is not visible. Figure 4 illustrates the situation. The PV system north to the potential adopter located at (c) is considered visible because no building lies within the line-of-sight. In contrast, the potential adopter at (a) does not see the closest PV system because there are buildings within the line-of-sight.

Based on the two conditions, we construct several visibility measures. We assume that condition 1 and 2 always hold for the directly visible installed base

$$\mathrm{IB}^{\triangleleft}_{n,t} = \sum_{m \in N^{\triangleleft}_n} \sum_{l=2000}^{t} f\left(y^{\mathrm{All}}_{m,l},\ d_{n,m}\right). \tag{8}$$

This is ensured by $N^{\triangleleft}_n = \bar{N} \bigcap N^{45^\circ}_n \bigcap N^{\mathrm{los}}_n$. We also consider (8) in relative terms in the fashion of (7) denoted as $\mathrm{IB}^{\mathrm{relative},\triangleleft}_{n,t}$. To identify the effect of not visible PV systems, we consider

$$\mathrm{IB}^{\ntriangleleft}_{n,t} = \sum_{m \in N^{\ntriangleleft}_n} \sum_{l=2000}^{t} f\left(y^{\mathrm{All}}_{m,l},\ d_{n,m}\right) \tag{9}$$

for which condition 2 does not hold (building in between), i.e., $N^{\ntriangleleft}_n = \bar{N} \bigcap N^{90^\circ}_n \setminus N^{\mathrm{los}}_n$.

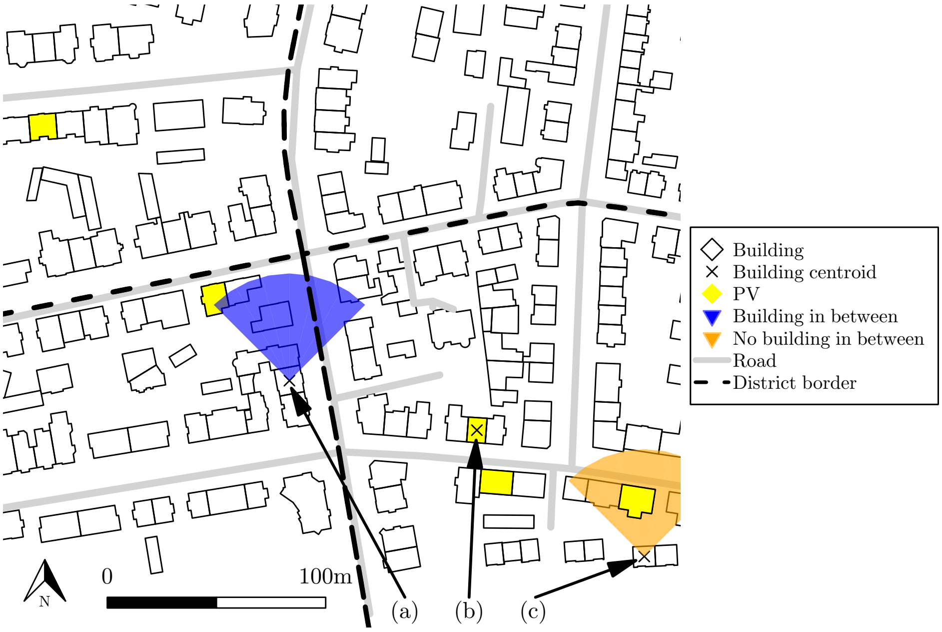


Figure 4: Examples of measuring visibility of PV systems.

Notes: The map displays examples of visibility measures. The yellow PV system north to location (a) is within the view shed but not visible because there is at least one building within the line-of-sight between the PV system and location (a). The PV system south to location (b) is not visible to (b) because we assume that the PV system is installed on the south side of the roof. The PV system is, therefore, not within the view-shed of location (b). In contrast, we consider the PV system north to location (c) as visible, because it lies within the view shed and there is no building in the line-of-sight.

Again, we also compute a relative measure denoted as $\text{IB}_{n,t}^{\text{relative},\not\sphericalangle}$. We also compute

$$\text{IB}_{n,t-1}^{\text{Rest}} = \text{IB}_{n,t-1} - \text{IB}_{n,t-1}^{\sphericalangle} - \text{IB}_{n,t-1}^{\not\sphericalangle} \tag{10}$$

and the corresponding relative measure, which captures the remaining PV systems that are not classified as visible or not visible. Table 1 reveals the descriptive statistics of our visibility measures.

The rationale of (8) and (9) is to disentangle the proximity-driven peer effect. (8) and (9) allow us to isolate the effect of visibility on the choice to adopt PV. Peer effects may exist for both (8) and (9), but (8) explicitly adds visibility (in contrast to (9)) to the peer effect. We expect the visibility-driven peer effect to be stronger than the peer effect of non-visible PVs.

Instrumental Variables

Endogeneity is a well known problem in studying social contagion (Nair, Manchanda, and Bhatia 2010; McShane, Bradlow, and Berger 2012; Risselada, Verhoef, and Bijmolt 2014). Unobservables may induce PV adoption and may be correlated with the error term $\epsilon_{n,t}$ in (3). Our estimates of the peer effect may therefore be biased. An instrumental variable estimation with two stage least squares is a common approach to account for endogeneity (J. Wooldridge 2012, Ch. 15). To do so, we need exogenous variation correlated with our measure of the installed base but not driving the actual adoption decision of a given adopter.

The (average) inclination and orientation over previous installations nearby are good candidates. First, a neighbor's roof inclination and roof orientation are important drivers 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 buildings were constructed without considering a PV system.[7]

Both, inclination and orientation contribute to the appropriateness of a roof for producing electricity from PV. In Western Europe an inclination of 37° and an orientation to the south corresponds to an exploitation of 100% of the energy potential. Figure 3 illustrates that a deviation from these optimal values yields a decline in electricity production. We employ high-resolution laser scanner data – collected 2000 - 2005 – to determine roof inclination and orientation (LUBW 2016). To be included in our data set, inclined roofs have to allow for more than 10 sq m of PV modules while flat roofs must have at least 25 sq m available to install a PV system (on stilts).

All roofs in our data set range between 0° and 74° inclination.[8] According to Figure 3, the optimal inclination is 37°. We normalize the inclination to range between 0 and 37,

[7]The number of new buildings is very low. The average yearly increase of residential buildings was 0.73% between 31.12.1999 and 31.12.2010.

[8]Table 6 in Appendix A.4 contains the descriptive statistics for the raw data. Next to the table we describe the raw data in detail.

i.e.,

$$\text{Inclination}_n = \begin{cases} \text{Inclination}^*_n & \text{if Inclination}^*_n \leq 37 \\ 74 - \text{Inclination}^*_n & \text{otherwise,} \end{cases} \tag{11}$$

with Inclination^*_n from LUBW (2016). Roof orientation has values between 0° and 359° with 180° representing the optimal orientation to the south. We apply the same concept as for inclination, i.e.,

$$\text{Orientation}_n = \begin{cases} \text{Orientation}^*_n & \text{if Orientation}^*_n \leq 180 \\ 180 - \text{Orientation}^*_n & \text{otherwise.} \end{cases} \tag{12}$$

For flat roofs, $\text{Inclination}_n = 37$ and $\text{Orientation}_n = 180$ is always true. We employ Inclination_n and Orientation_n to build our instruments, *average inclination ratio over previous installations nearby*

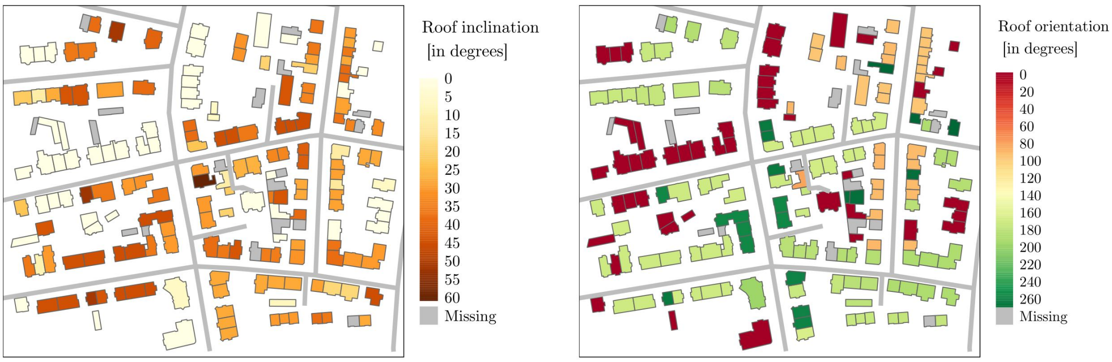
$$\text{AvgIncRatio}_{n,t} = \begin{cases} \frac{\sum_{m \in \bar{N}} \sum_{l=2000}^{t} f(y^{\text{All}}_{m,l}, d_{n,m}) \frac{\text{Inclination}_m}{37}}{\sum_{m \in \bar{N}_n} \sum_{l=2000}^{t} y^{\text{All}}_{m,l}} & \text{if } \sum_{m \in \bar{N}_n} \sum_{l=2000}^{t} y^{\text{All}}_{m,l} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

and *average orientation ratio over previous installations nearby*

$$\text{AvgOrRatio}_{n,t} = \begin{cases} \frac{\sum_{m \in \bar{N}} \sum_{l=2000}^{t} f(y^{\text{All}}_{m,l}, d_{n,m}) \frac{\text{Orientation}_m}{180}}{\sum_{m \in \bar{N}_n} \sum_{l=2000}^{t} y^{\text{All}}_{m,l}} & \text{if } \sum_{m \in \bar{N}_n} \sum_{l=2000}^{t} y^{\text{All}}_{m,l} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

as spatio-temporal lag variables. (13) and (14) work in a similar way as for the installed base measure (7). Preexisting, proximate PV systems are weighted by an exogenous measure of roof inclination $\left(\frac{\text{Inclination}_m}{37}\right)$ and orientation $\left(\frac{\text{Orientation}_m}{180}\right)$ on a scale between 0 and 1. Our instruments (13) and (14) do not rely on maximum available PV module area on a given roof since we do not want to favor larger roofs: larger roofs may be associated with confounding factors such as higher income which in turn may be related to PV adoption decisions. Additionally, we consider instruments with respect to visibility, i.e., we replace $\bar{N}$ in (13) and (14) by $N^{\triangleleft}_n$ and $N^{\not\triangleleft}_n$ accordingly. Table 1 contains the corresponding descriptive statistics.

A potential threat to our instruments may be identical inclination and orientation of neighboring roofs. Fortunately, this is not the case here. Buildings are often aligned according to the course of a street and streets are often not straight in Germany. PV appropriateness of roofs, which is determined by their inclination and orientation, varies along a street. Figure 5 illustrates the raw data for roof inclination and orientation and confirms that roofs of buildings in the same street located next to each other do not necessarily have the same roof inclination and orientation.



(a) Inclination$_n^*$ (b) Orientation$_n^*$

Figure 5: Illustration of roof inclination and roof orientation from roof-census in Baden-Württemberg. Both panels show the same excerpt of Freiburg as Figure 4.

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 consider District-Street-Year fixed effects $\eta_{k,t}$ in our analysis. We thereby control for time-variant adoption shocks on the District-Street-Year level. These fixed effects also control for similar PV suitability (or potentially similar inclination and orientation) of buildings n in the same District-Street unit k. High-dimensional fixed effects like $\eta_{k,t}$ may bias peer effects. However, as shown by Narayanan and Nair (2013), strong instruments control for this bias. On top of that, the correlation of inclination of 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 relative 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.

Controls

Besides a rich set of fixed effects, we include building-specific controls i in our analysis. They account for the suitability of a building n to produce electrical power from PV. We include: potential electricity production, potential electricity production times year, maximum module area, maximum module area times year and a rural area dummy. The maximum module area and the potential electricity production control for the capacity of n to produce solar energy. They are also proxies for a series of income related variables. The larger the feasible module area and the larger the potential electricity production, the larger may be the building and the higher may be the income (of a building owner). During our period of study, PV installations got cheaper and the subsidy system changed. Both may have a time-variant influence on adoption: if installation cost for PV systems are high then high income households (high maximum module area) are more likely to adopt compared to low income households. We therefore include the interaction between maximum module area and t as well as potential electricity production with t as time-variant controls. Finally, we employ Corine Land Cover data (CLC 2009) to identify whether building n is located in a rural area or not. It may be easier to make a decision about PV adoption with less individuals involved in the decision process. In consequence, there may be more PV systems on single family houses than on apartment houses and the share of single family houses is larger in rural areas than in densely populated urban areas.

Estimation Results

We first discuss our baseline findings for peer effects in PV adoption including some robustness checks. Then, we turn to identify causal peer effects in the adoption process and study the visibility channel for peer effects in PV adoption. In a final step, we illustrate the robustness of our findings for a proportional hazard model with a piece-wise exponential baseline.

Baseline Peer Effects

Our baseline regression equation (column (4) in Table 2) takes the following form:

$$y_{n,t} = \beta_1 \mathrm{IB}_{n,t-1} + \sum_i \beta_i X_{i,n,t}^{\text{controls}} + \alpha_t + \gamma_n + \eta_{k,t} + \epsilon_{n,t}. \tag{15}$$

Recall, $y_{n,t} = 1$ if we observe a PV installation on building n in year t for the first time (0, otherwise). Because usually only one PV system can be installed, we censor the data in the way that we ignore buildings once they have installed a PV system. β_1 measures the peer effect based on all PV systems installed until the preceding period (lag $t-1$) within radius D.[9]

Baseline results.– Column (1) in Table 2 illustrates the estimated coefficients from the linear probability model with year fixed effects. The estimated coefficient for the installed base is positive and statistically significant at all relevant statistical levels. In consequence, the lagged installed base $\mathrm{IB}_{n,t-1}$ is positively linked to PV system adoption. 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.

District-Street-Year fixed effects $\eta_{k,t}$.– To measure peer effects more confidently, Column (2) in Table 2 illustrates that our estimates are robust to controlling for granular spatial fixed effects by year. An obvious interpretation of the level of our results is in terms of a 0.01 unit increase in $\mathrm{IB}_{n,t-1}$, which roughly corresponds to a half standard deviation increase (see Table 1) and is equal to one previously installed PV system within 200m distance. The estimates from column (2) in Table 2 reveal that for a 0.01 unit increase in the installed base, we expect the probability of installing a PV system to increase by $(0.0098 \times 0.01) \times 100 \approx 0.01$ percentage points. Similarly, 10 previously installed PV systems within 200m distance correspond to a 0.1 unit increase in the installed base (in each time period after the 10 PV systems were installed). In this case, we expect the probability of installing a PV system to increase by 0.1 percentage points. This increase may seem to be low. However, we have to consider that only about 5% of the buildings

[9] We use REGHDFE in stata (Correia 2017), which allow us to obtain robust standard errors clustered at the street level.

Table 2: Estimates for peer effects.

	No FE	FE		More FE	IV (OLS)	
	OLS	OLS	CLogit	OLS	1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)	(6)
	$y_{n,t}$	$y_{n,t}$	$y_{n,t}$	$y_{n,t}$	$\mathrm{IB}_{n,t-1}$	$y_{n,t}$
Installed base:						
$\mathrm{IB}_{n,t-1}$	0.017^{**}	0.0098^{**}	2.23^{**}	0.012^{**}		
	(22.9)	(4.86)	(4.69)	(4.13)		
Predicted installed base:						
$\widehat{\mathrm{IB}}_{n,t-1}$						-0.0059
						(-0.56)
Instruments (average ratio over previous PV installations nearby):						
Avg. inclination ratio: $\mathrm{AvgIncRatio}_{n,t-1}$					0.26^{**}	
					(19.5)	
Avg. orientation ratio: $\mathrm{AvgOrRatio}_{n,t-1}$					0.50^{**}	
					(39.4)	
Observations	29,380,453	27,643,593	769,963	27,640,461	27,640,461	27,640,461
$\mathrm{DF_M}$	15	6	6	3	4	3
Final log-likelihood	46,947,707		-142,374			
Adj. R^2	0.00	0.02		0.06		0.00
F	1782.7	208.5		351.8		346.7
Hansen J (p-value)						0.3 (0.61)
Kleibergen-Paap rk Wald F statistic						16,375.0
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	No	Yes	Yes	Yes	Yes	Yes
Building fixed effects	No	No	No	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We show coefficients and, in parentheses, t statistics from standard errors clustered at the street level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

In columns (1-4, 6), the dependent variable is always $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. In column (5), the dependent variable is the installed base ($\mathrm{IB}_{n,t-1}$) for building n, which includes all PV systems installed through $t-1$. We estimate columns (1-2) and (4-6) via OLS and column (3) via conditional logit. The panel of 2,957,332 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. Across columns, the sample is always the same. However, the conditional logit estimator drops all positive (or all negative) outcomes in terms of District-Street-Year groups for column (2). I.e., for column (2) District-Street-Year groups with no adoption (or if all buildings in that group adopt at once) are dropped. These procedures result in fewer observations. Similarly, singleton observations are dropped in columns (3-6), which also results in fewer observations. Across columns, the Cut-off distance D = 200m. Column (6) uses the predicted values for the installed base from the first stage in column (5). We include n-specific controls: potential electricity production, potential electricity production times year, maximum module area, maximum module area times year and a rural area dummy. With address fixed effects the time-invariant controls (potential electricity production, maximum module area and the rural dummy) are excluded. Note that an increase of the installed base ($\mathrm{IB}_{n,t-1}$) of 0.01 refers to 1 additional previously installed PV system nearby.

have a PV system installed at the end of our period of analysis (2010), which implies that the overall probability to install a PV system per individual and time period is low.

To better understand the relevance of the estimated effect, we consider conditional logit estimates (section Appendix A.1 for details). These allow for a relative interpretation of the estimated probabilities.[10] As before, we account for District-Street-Year fixed effects in column (3) of Table 2. The estimated coefficient is statistically significant and positive. One additional, previously installed PV system within 200m distance, is linked to an increase in the odds of installing a PV system by $(\exp(2.23 \times 0.01) - 1) \times 100 = 2.3\%$. Similarly, 10 previously installed PV systems within 200m distance correspond to a 0.1 unit increase in the installed base. In this case, we expect the odds of installing a PV system to increase by $(\exp(2.23 \times 0.1) - 1) \times 100 = 25\%$.[11]

[10] See W.H. Greene (2012, pp. 721-724) for details on the conditional logit estimator.

[11] We cannot directly compare the estimated coefficients from the linear probability model and the ones from conditional logit.

District-Street-Year fixed effects $\eta_{k,t}$ and building fixed effects γ_n.– The estimates of our LPM with high dimensional fixed effects $\eta_{k,t}$ and γ_n are given in column (4) Table 2. The results are within the ballpark of the other OLS estimates (columns (1) and (2)). In particular, a z-test (Clogg, Petkova, and Haritou 1995) reveals a statistically significant difference between the coefficients of $\text{IB}_{n,t-1}$ in columns (1) and (2). However, the z-test reveals no difference between the coefficients in column (1) and (4) as well as between column (2) and (4). Hence, we are confident that our results are robust in terms of the fixed effects strategy. If not mentioned otherwise, we only show results including District-Street-Year $\eta_{k,t}$ and building fixed effects γ_n from now on.

Robustness.– We conduct a battery of robustness and sensitivity tests. In Appendix B.1 we provide further details. First, we illustrate that we do not find evidence for time varying peer effects. There is no difference between the estimates for our first periods of study and later periods (see column (1), Table 11). Second, we show that our estimates of peer effects are localized. In column (2), Table 11, we use a cut-off distance of 100m for the installed base $\text{IB}_{n,t-1}^{D\leq 100}$ and also include the installed base that includes PV systems farther away than 100m but still closer than 200m $\text{IB}_{n,t-1}^{100<D\leq 200}$ as well as those farther away than 200m and closer than 400m $\text{IB}_{n,t-1}^{200<D\leq 400}$. Table 7 contains the descriptive statistics. We observe a statistically significant and positive coefficient for the installed base measure with cut-off $D = 100$ $\left(\text{IB}_{n,t-1}^{D\leq 100}\right)$. The estimated coefficients for the remaining measures are smaller in absolute terms. We conclude that a cut-off distance of 200m is most likely to capture most of the peer effect and, therefore, stick to a cut-off distance of 200m in the following. Third, we check that our estimates are robust to a distance-weighted measure of the installed base in column (3), Table 11. We include both our baseline measure and a distance-weighted version $\left(IB_{n,t-1}^{1/d}\right)$ and find a statistically significant and positive coefficient for the baseline measure. We conclude that our results remain unaffected from controlling for distance. Fourth, Nair, Manchanda, and Bhatia (2010) and Van den Bulte and Stremersch (2004) propose relative measures of the installed base. In column (4), Table 11, we include the relative installed base $\text{IB}_{n,t-1}^{\text{Relative}}$ (7), which is the installed base normalized by the number of buildings nearby. The corresponding

coefficient is positive and statistically significant as expected. Fifth, we randomly allocate the same number of PV installations, which were in fact installed in Baden-Württemberg per year, to existing buildings. This procedure confirms that our estimations do not find a positive peer effect in PV system adoption by definition (column (1) in Table 12 in Appendix B.1). Sixth, we classify PV systems with more than 30kW$_p$ as industrial. When focusing on industrial PV systems we do not find evidence for positive peer effects. If at all, there is a negative association between the installed base and the decision to install an industrial PV system, see column (2) in Table 12 in Appendix B.1. This makes sense as industrial investors should search for the best location across a large region rather than install a PV system at their headquarters only. Seventh, our findings remain unaffected by lagging our peer effect measure by two years (Rode and Weber 2016), see column (3) in Table 12 in Appendix B.1. We can therefore confidently rule out that the 'reflection problem' described by Manski (1993) biases our results. Eighth, we illustrate that incorporating the installed base with a one year lag and the installed base with a two year lag does not affect our core findings (column (4) in Table 12, Appendix B.1). The estimates indicate that the installed base including last year's PV installations is 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"*. Of course, recent adopters may be more contagious than those who adopted less recently because they may be more credible or enthusiastic. Further, as more common knowledge on PV become available, information of additional adopters becomes less important (Grinblatt, Keloharju, and Ikaheimo 2008; Risselada, Verhoef, and Bijmolt 2014).

Causal Peer Effects?

So far, we have only reported a positive association between the installed base and adoption decisions. 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 which 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.

We regress the installed base $\text{IB}_{n,t}$ on our exogenous instrumental variables in the first stage. Then, we regress the actual PV adoption decision $y_{n,t}$ on the corresponding predicted installed base $\widehat{\text{IB}_{n,t}}$ as an exogenous variable in the second stage. This procedure allows us to identify causal peer effects.[12]

Causal estimates.– In column (5) of Table 2, we illustrate the results for the ordinary least squares first stage regression. In this case, we regress the lagged installed base on the average inclination (13) and orientation (14) ratio over previous PV installations nearby. Both instruments are strong, since the corresponding coefficients are statistically significant different from zero. As we would expect, we find a positive association of both instruments with the installed base. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. We now turn to the second stage regression in column (6) in which we include the predicted installed base from the first stage. Interestingly, there is no effect of the instrumented installed base on the adoption decision.

From the previous robustness analysis (Section Baseline Peer Effects), we learned that the relative installed base may be an alternative measure capturing peer effects. Columns (5) and (6) of Table 11 (Appendix B.1) reveal that the average inclination and orientation ratio over previous PV installations nearby are also strong instruments for the relative installed base but again, there is no positive statistically significant effect of the instrumented (relative) installed base on the adoption decision.[13] We conclude that we do not find evidence for a general peer effect in PV adoption, which is in contrast to the findings in the literature on PV adoption (Bollinger and Gillingham 2012; Bollinger, Gillingham, et al. 2019; Müller and Rode 2013). Risselada, Verhoef, and Bijmolt (2014) and Van den Bulte and Iyengar (2011) discuss that peer effects may be spurious due to incorrect models and/or aggregate data. Hence, the general peer effects in PV adoption

[12] We use IVREGHDFE in stata (Correia 2018), which allow us to obtain robust standard errors clustered at the street level.

[13] In contrast to the OLS regression with a positive and statistically significant coefficient for the relative installed base (Column (4) of Table 11), the coefficient for the predicted relative installed base is negative and statistically significant to the 10% level. Because the coefficient turns sign, we do not consider the instrumental variables result as evidence for negative peer effects.

reported in the literature so far, may be driven by the use of aggregate data or insufficient identification.

Peer Effects from Visible Systems

Visibility from buildings.– We exploit installed base measures that rely on visibility between buildings. As before, we include the full set of high-dimensional fixed effects α_t, γ_n, and $\eta_{k,t}$.

The measure $\mathrm{IB}^{\triangleleft}_{n,t-1}$ (8) contains PV systems visible to n, $\mathrm{IB}^{\not\triangleleft}_{n,t-1}$ (9) contains PV systems not visible to n and $\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}$ (10) captures the remaining PV systems. The corresponding estimates in Table 3 column (1) indicate a positive association between PV adoption and previously installed visible PV systems nearby. In contrast, there is no association between PV adoption and previously installed non-visible PV systems nearby. The estimated coefficient for $\mathrm{IB}^{\triangleleft}_{n,t-1}$ is statistically significant different from zero and positive. In contrast, the coefficient of $\mathrm{IB}^{\not\triangleleft}_{n,t-1}$ is not statistically different from zero. The estimated coefficient for $\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}$ is statistically significant different from zero and positive. The coefficients of $\mathrm{IB}^{\triangleleft}_{n,t-1}$ and $\mathrm{IB}_{n,t-1}$ are statistically significant different from each other as indicated by a corresponding Wald test. The coefficients of $\mathrm{IB}^{\triangleleft}_{n,t-1}$ and $\mathrm{IB}^{Rest}_{n,t-1}$ also differ in a statistically significant way.

Again, we are interested in causal peer effects. We use the average inclination and orientation ratio over previous visible PV systems nearby to instrument for $\mathrm{IB}^{\not\triangleleft}_{n,t-1}$. Column (2) in Table 3 contains the first stage regression. Both instruments are strong because they are statistically significant different from zero. As expected, we find a positive association of both instruments on the visible installed base $\mathrm{IB}^{\triangleleft}_{n,t-1}$. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. In the second stage regression in column (3), we use the predicted installed base from the first stage. The effect of the instrumented visible installed base on the adoption decision is strong and statistically significant at all confidence levels. Comparing the OLS specification in column (1) with the instrumented specification in column (3) shows that the estimates are in the same ballpark.

Table 3: Estimates for peer effects visibility measures.

	Absolute			Relative		
	OLS	IV		OLS	IV	
		1st stage	2nd stage		1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)	(6)
	$y_{n,t}$	$\mathrm{IB}^{\triangleleft}_{n,t-1}$	$y_{n,t}$	$y_{n,t}$	$\mathrm{IB}^{\mathrm{Relative},\triangleleft}_{n,t-1}$	$y_{n,t}$
Installed base (direct visibility):						
No building in-between: $\mathrm{IB}^{\triangleleft}_{n,t-1}$	0.048^{**}					
	(4.61)					
No building in-between (rel.): $\mathrm{IB}^{\mathrm{relative},\triangleleft}_{n,t-1}$				0.035^{**}		
				(5.50)		
Predicted installed base (direct visibility):						
No building in-between: $\widehat{\mathrm{IB}}^{\triangleleft}_{n,t-1}$			0.039^{**}			
			(3.50)			
No building in-between (rel.): $\widehat{\mathrm{IB}}^{\mathrm{relative},\triangleleft}_{n,t-1}$						0.079^{**}
						(4.08)
Instruments (average ratio over previous visible PV nearby):						
Avg. inclination ratio: $\mathrm{AvgIncRatio}^{\triangleleft}_{n,t-1}$		0.41^{**}			0.21^{**}	
		(146.9)			(19.2)	
Avg. orientation ratio: $\mathrm{AvgOrRatio}^{\triangleleft}_{n,t-1}$		0.81^{**}			0.48^{**}	
		(314.7)			(47.0)	
Installed base controls (no direct visibility):						
Building in-between: $\mathrm{IB}^{\not\triangleleft}_{n,t-1}$	0.0056	-0.020^{**}	0.0046			
	(1.19)	(−86.4)	(0.98)			
Rest: $\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}$	0.013^{**}	-0.010^{**}	0.012^{**}			
	(4.12)	(−72.0)	(3.94)			
Building in-between: $\mathrm{IB}^{\mathrm{relative},\not\triangleleft}_{n,t-1}$				0.018^{**}	0.45^{**}	-0.021^{*}
				(4.30)	(92.3)	(−2.31)
Relative, rest: $\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}$				0.014^{**}	0.28^{**}	-0.0099^{+}
				(7.17)	(55.6)	(−1.77)
Observations	27,640,461	27,640,461	27,640,461	27,630,743	27,630,743	27,630,743
$\mathrm{DF_M}$	5	6	5	5	6	5
Adj. R^2	0.06		0.00	0.06		0.00
F	213.7		212.1	220.5		217.1
Hansen J (p-value)			0.3 (0.55)			0.4 (0.51)
Kleibergen-Paap rk Wald F statistic			1,319,495.0			6,887.1
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
In columns (1,3,4,6), the dependent variable is $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. The panel of 2,957,332 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. We estimate all columns via OLS. Singleton observations are dropped. This procedure results in fewer observations. Column (3) uses the predicted values for the installed base $\left(\mathrm{IB}^{\triangleleft}_{n,t-1}\right)$ from the first stage in column (2). Column (6) uses the predicted values for the relative installed base $\left(IB^{\mathrm{Relative},\triangleleft}_{n,t-1}\right)$ from the first stage in column (5). Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of the installed base measure $\mathrm{IB}^{\triangleleft}_{n,t-1}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n.

We conduct an overidentifying restrictions test. As both instruments are statistically significant in the first stage regression, the overidentifying restrictions test has power. The p-value for the Hansen statistic is 0.55. The test indicates that our instruments are uncorrelated with the error term at all significance levels. Therefore, the instruments pass the overidentifying restrictions test, which supports the validity of our instruments: i.e., the effect of the average inclination and orientation ratio over previous visible PV systems nearby, on the adoption decision of an individual is through the peer effect (of visible PV systems).

We test the equality of the estimated coefficients for the predicted, visible installed base $\widehat{\mathrm{IB}}^{\triangleleft}_{n,t-1}$ and the one for the installed base baseline measure $\mathrm{IB}_{n,t-1}$ (in column (3)). A Wald test confirms that we can reject the null that they are equal. The coefficients of $\widehat{\mathrm{IB}}^{\triangleleft}_{n,t-1}$ and $\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}$ also differ in a statistically significant way. We learn from the trade-off between visible PV systems and not visible PV systems nearby, that one prior visible PV system may substitute $0.039/0.0046 \approx 8$ prior, not visible PV systems nearby.

In Table 13 (Appendix B.1), we illustrate that the effect of the remaining PV systems does not drive PV adoption. In this case, we regress the lagged installed base of the remaining PV systems $\left(\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}\right)$ on the average inclination (13) and orientation (14) ratio over previous PV installations nearby. Both instruments are strong, since the corresponding coefficients are statistically significant different from zero. As we would expect, we find a positive association of both instruments with the installed base. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. We now turn to the second stage regression in column (2) in which we include the predicted installed base from the first stage. There is no effect of the instrumented installed base of the remaining PV systems on the adoption decision.

Our results provide evidence that visibility is an important channel for peer effects in PV adoption. This finding is in line with Bollinger, Gillingham, et al. (2019), who study the U.S. and find that PV systems visible from roads induce higher peer effects than systems that are not visible.

Robustness.– We conduct several robustness and sensitivity tests. First, from our previous analyses, we know that the relative installed base measure may also capture peer effects. We build relative installed base measures for visible $\text{IB}_{n,t-1}^{\text{Relative},\lhd}$ and non visible PV systems $\text{IB}_{n,t-1}^{\text{relative},\ntriangleleft}$. The OLS estimates in column (4) of Table 3 indicate a positive association between PV adoption and previously installed visible PV systems nearby. We also instrument for $\text{IB}_{n,t-1}^{\text{relative},\lhd}$ using the average inclination and orientation ratio over previous visible PV systems nearby. According to the first stage regression in column (5), Table 3, both instruments are strong because they are highly significant to all relevant statistical levels. The association of both instruments with the relative visible installed base $\text{IB}_{n,t-1}^{\text{Relative},\lhd}$ is positive and statistically significant. The Kleibergen-Paap Wald F-test rejects the null of insignificance of the instruments. In the second stage regression in column (6), we use the predicted relative installed base from the first stage. The effect of the instrumented, relative visible installed base measure on the adoption decision is strong and significant at all confidence levels.

Second, we illustrate that our results are robust to also including other installed base measures at the same time, see Table 14 in Appendix B.2. Here, we control for our baseline installed base measures $\text{IB}_{n,t-1}$ and $\text{IB}_{n,t-1}^{\text{Relative}}$. We include the installed base from flat roofs and the one from roofs that are not flat. We also include a measure that only includes roofs with PV systems that are close to a street (at most 50 meters away from the closest street). None of these measures robustly indicates peer effects no matter if we build absolute or relative installed base measures. This finding makes sense because none of the measures can clearly distinguish visible from non-visible PV systems. In the regressions, we, of course, also include our measure for visible PV systems nearby. Both, for absolute and for relative measures, the peer effect for directly visible PV systems nearby is statistically significant and positive and in levels almost similar to the estimates shown in Table 3. As before, the results are robust against instrumentation.

Third, we modify our sample. The results shown so far are based only on those adoption decisions, $y_{n,t}$, for which we can be very sure that the allocation of a PV system is to the correct roof. We can however also include those PV systems for which no unique allocation was possible. The PV systems without unique allocation are randomly assigned to another roof in the same statistical district. We show the descriptive statistics in Table 8 in Appendix A.5. Table 15 in Appendix B.2 illustrates that our results remain unaffected from analyzing the full sample. In Table 15, we show the robustness of our results while controlling for the other installed base measures (introduced in Table 14), both for OLS and for instrumental variables estimation as well as for absolute and for relative installed base measures.

Fourth, we show that our results on visibility remain unaffected from lagging our installed base measure by two years for absolute installed base measures in column (1) of Table 16 and for the relative measure in column (1) of Table 17 in Appendix B.2.[14] Columns (2) and (3) in both tables, show that the analysis with two year lags is robust to our instrumentation approach. We can therefore confidently rule out that the 'reflection problem', described by Manski (1993), biases our results.

Fifth, we show that our results remain unaffected from including a placebo-lead, visible installed base. Column (4) of Table 16 (Appendix B.2) contains the estimates for the absolute measures and column (4) of Table 17 (Appendix B.2) for the relative measures.[15] In both cases, the placebo-lead has a statistically significant and negative coefficient and the lagged visible installed base has a positive and statistically significant coefficient. The placebo-lead test again indicates that our installed base measures do not indicate positive peer effects by construction.

Sixth, we randomly allocate the same number of PV installations, which were in fact installed in Baden-Württemberg per year, to existing buildings. This procedure confirms that our estimates from our visibility measure (both, for the absolute and relative measures) do not find a positive peer effect in PV system adoption by definition (column (1) and (2) in Table 18 in Appendix B.1).

Seventh, we classify PV systems with more than 30kW$_p$ as industrial. When focusing on industrial PV systems, we, again, do not find evidence for positive peer effects. If at all, there is a negative association between the installed base and the decision to install an industrial PV system, see column (3) and (4) in Table 18 in Appendix B.1.

[14] Table 9 in Appendix A.5 contains the descriptive statistics.
[15] Table 10 in Appendix A.5 reveals the descriptive statistics.

Hazard Model of Adoption

Our results remain unaffected from using a different modeling approach to technology adoption. We fit a proportional hazard model with a piece-wise exponential baseline where the hazard changes for each District-Street-Year unit (see Appendix A.2).

Table 4 indicates almost the same coefficient as in the conditional logit model shown in Table 2 column (3). Endogeneity may bias the results and hence Lin and J. M. Wooldridge (2019) suggest to use a control function in non-linear models if the endogenous variable (installed base) is continuous (as in our case). We follow this approach and, as before, use the average orientation and the average inclination of neighboring roofs with PV to instrument for the baseline installed base in column (2), Table 4. Then, in column (3), we control for the residual obtained from the first stage regression denoted as $\hat{\nu}$. We follow Lin and J. M. Wooldridge (2019) and bootstrap standard errors in column (3). Squared parenthesis contain the resulting p-values. The insignificance of the first stage residuals ($\hat{\nu}$) indicate that the installed base is indeed exogenous (Lin and J. M. Wooldridge 2019). The estimated coefficient for the installed base is positive and statistically significant (at the 10% level). Our hazard model therefore points to causal peer effects in PV adoption.

Column (4) in Table 4 contains the estimates of the installed base measures for visibility, while controlling for the non visible installed base and the installed base for the remaining systems. We observe a positive and statistically significant coefficient for the visible installed base. The control function approach (columns (5) and (6)) again reveal causal peer effects for visible PV systems.

Conclusion and Managerial Implications

In this paper, we empirically analyze the building-specific adoption of rooftop PV installations in Baden-Württemberg, 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

Table 4: Hazard model of adoption.

	Baseline			Visibility		
	Poisson	Control function		Poisson	Control function	
		1st stage OLS	2nd stage Poisson		1st stage OLS	2nd stage Poisson
	(1)	(2)	(3)	(4)	(5)	(6)
	$h_{n,t}$	$\mathrm{IB}_{n,t-1}$	$h_{n,t}$	$h_{n,t}$	$\mathrm{IB}^{\triangleleft}_{n,t-1}$	$h_{n,t}$
Installed base:						
$\mathrm{IB}_{n,t-1}$	2.19^{**}		2.92^{+}			
	(4.69)		(1.94)[0.07]			
First stage residuals for installed base:						
$\mathrm{IB}_{n,t-1}$: $\hat{\nu}$			-0.80			
			(-0.51)[0.68]			
Instruments:						
Avg. inclination of nearby roofs with $\mathrm{PV}_{n,t-1}$		0.64^{**}				
		(10.5)				
Avg. orientation of nearby roofs with $\mathrm{PV}_{n,t-1}$		1.09^{**}				
		(18.7)				
Installed base (direct visibility):						
No building in-between: $\mathrm{IB}^{\triangleleft}_{n,t-1}$				7.03^{**}		6.52^{**}
				(4.66)		(3.92)[0.00]
First stage residuals installed base (dir. visibil.):						
No building in-between $\left(\mathrm{IB}^{\triangleleft}_{n,t-1}\right)$: $\hat{\nu}$						2.79
						(0.73)[0.33]
Instruments (average ratio over previous visible PV nearby):						
Avg. inclination ratio: $\mathrm{AvgIncRatio}^{\triangleleft}_{n,t-1}$					0.46^{**}	
					(54.0)	
Avg. orientation ratio: $\mathrm{AvgOrRatio}^{\triangleleft}_{n,t-1}$					0.80^{**}	
					(103.1)	
Installed base controls (no direct visibility):						
Building in-between: $\mathrm{IB}^{\ntriangleleft}_{n,t-1}$				1.18	-0.013^{**}	1.15
				(1.47)	(-21.2)	(1.43)[0.36]
Rest: $\mathrm{IB}^{\mathrm{Rest}}_{n,t-1}$				2.32^{**}	-0.0069^{**}	2.30^{**}
				(4.63)	(-18.2)	(4.58)[0.00]
Observations	770,075	770,075	770,075	770,075	770,075	770,075
$\mathrm{DF_M}$	6	7	7	8	9	9
Final log-likelihood	-206,650	2,493,380	-206,650	-206,643	4,157,534	-206,642
Adj. R^2		0.89			0.86	
F		2028.3			28571.7	
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Building fixed effects	No	No	No	No	No	No
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We show coefficients and, in parentheses, t statistics from standard errors clustered at the street level, and in squared parentheses p-values from bootstrapping standard errors (clustered at the street level); $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
In columns (1, 3-4, 6), we estimate the piecewise hazard model of adoption. We use the Poisson regression model with multiple high-dimensional fixed effects (Correia, Guimarães, and Zylkin 2019). In column (2), the dependent variable is the installed base $(\mathrm{IB}_{n,t-1})$ for building n, which includes all visible PV systems installed through $t-1$. In column (5), the dependent variable is the directly visible installed base $\left(IB^{\triangleleft}_{n,t-1}\right)$ for building n, which includes all PV systems installed through $t-1$. We estimate columns (2) and (5) via OLS with high-dimensional fixed effects. In column (3), we use the control function approach described in Lin and J. M. Wooldridge (2019) and control for the residuals from the first stage in column (2). In column (6), we use the control function approach described in Lin and J. M. Wooldridge (2019) and control for the residuals from the first stage in column (5). The panel of 2,957,332 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. Across columns, the sample is always the same. However, the Poisson estimator with high-dimensional fixed effects drops all positive (or all negative) outcomes in terms of District-Street-Year groups for columns (1, 3, 4, 6). I.e., for columns (1, 3, 4, 6) District-Street-Year groups with no adoption (or if all buildings in that group adopt at once) are dropped. These procedures result in fewer observations. In column (2), we use the sample from column (1) and, in column (5), the sample from column (4). Across columns, the Cut-off distance $D = 200$m. We include n-specific controls: potential electricity production, potential electricity production times year, maximum module area, maximum module area times year and a rural area dummy. Note that an increase of the installed base $(\mathrm{IB}_{n,t-1})$ of 0.01 refers to 1 additional previously installed PV system nearby.

using disaggregate adoption data to document social contagion. We confirm the existence of a peer effect in PV adoption for visible installations. Of course, our study has limitations. We mainly focus on visibility measures from one building to another. Visibility of PV systems during commuting or traveling may also have an impact.

Our instrumental variable estimation provides strong evidence for a cause-and-effect relationship between prior visible 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. Visible PV systems nearby may be seen as a large diverse pool of information which reduces uncertainty about the technology.

Our findings have important managerial implications. First, fostering technology diffusion by pilot projects should focus on the most efficient ones – projects which are highly visible. We find that the probability of installing a PV system is around eight times higher if there is a visible PV systems nearby in comparison to not visible PV systems.

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.

Our findings are comparable to other technology adoption processes, such as solar thermal systems, which are much like PV systems but generate hot water. It may also be relevant for 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 to establish efficient (i.e., visible) seeds which may cause further adoptions. Further, designs that allow us to easily identify electric vehicles may support diffusion.

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. Furthermore, we are grateful for helpful comments from colleagues. Johannes Rode acknowledges the support of the Chair of International Economics at Technische Universität Darmstadt.

References

Akerlof, George A. (1997). "Social Distance and Social Decisions". *Econometrica* 65(5), 1005–1027. DOI: 10.2307/2171877.

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.

Bai, Jushan (2009). "Panel Data Models With Interactive Fixed Effects". *Econometrica* 77(4), 1229–1279. DOI: 10.3982/ECTA6135.

Bailey, Michael et al. (Aug. 2018). "Social Connectedness: Measurement, Determinants, and Effects". *Journal of Economic Perspectives* 32(3), 259–80. DOI: 10.1257/jep.32.3.259.

Bass, Frank M. (1969). "A New Product Growth Model for Consumer Durables". *Management Science* 15(5), 215–227. URL: http://www.jstor.org/stable/2628128.

Bernard, Andrew B. and J. Bradford Jensen (2004). "Why Some Firms Export". *The Review of Economics and Statistics* 86(2), 561–569. DOI: 10.1162/003465304323031111.

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 (September 8, 2020).* SSRN Working Paper. DOI: 10.2139/ssrn.3409420.

Caudill, Steven B. (1988). "PRACTITIONERS CORNER: An Advantage of the Linear Probability Model over Probit or Logit". *Oxford Bulletin of Economics and Statistics* 50(4), 425–427. DOI: 10.1111/j.1468-0084.1988.mp50004005.x.

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.

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.

Clogg, Clifford C., Eva Petkova, and Adamantios Haritou (1995). "Statistical methods for comparing regression coefficients between models." *American Journal of Sociology* 100(5), 1261–1293. DOI: 10.1086/230638.

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.

Correia, Sergio (July 2017). *REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects*. Statistical Software Components, Boston College Department of Economics. URL: https://ideas.repec.org/c/boc/bocode/s457874.html.

— (Sept. 2018). *IVREGHDFE: Stata module for extended instrumental variable regressions with multiple levels of fixed effects*. Statistical Software Components, Boston College Department of Economics. URL: https://ideas.repec.org/c/boc/bocode/s458530.html.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin (2019). *ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects.*

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.

Duntley, Seibert Q. (1948). "The Visibility of Distant Objects". *Journal of the Optical Society of America* 38(3), 237–249.

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.

Gardete, Pedro M. (2015). "Social Effects in the In-Flight Marketplace: Characterization and Managerial Implications". *Journal of Marketing Research* 52(3), 360–374. DOI: 10.1509/jmr.13.0527.

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.

Godes, David (2011). "Commentary—Invited Comment on "Opinion Leadership and Social Contagion in New Product Diffusion"". *Marketing Science* 30(2), 224–229. DOI: 10.1287/mksc.1100.0605.

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/lbu036.

Greene, W.H. (2012). *Econometric Analysis*. 7th ed. New Jersey: Pearson Education.

Greene, William (June 2004). "The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects". *The Econometrics Journal* 7(1), 98–119. DOI: 10.1111/j.1368-423X.2004.00123.x.

Grinblatt, Mark, Matti Keloharju, and Seppo Ikaheimo (2008). "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors". *The Review of Economics and Statistics* 90(4), 735–753. DOI: 10.1162/rest.90.4.735.

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.

Hipp, John R., Robert W. Faris, and Adam Boessen (2012). "Measuring 'neighborhood': Constructing network neighborhoods". *Social Networks* 34(1). Capturing Context: Integrating Spatial and Social Network Analyses, 128–140. DOI: https://doi.org/10.1016/j.socnet.2011.05.002.

Hipp, John R. and Andrew J. Perrin (2009). "The Simultaneous Effect of Social Distance and Physical Distance on the Formation of Neighborhood Ties". *City & Community* 8(1), 5–25. DOI: https://doi.org/10.1111/j.1540-6040.2009.01267.x.

I-TSO (2012). "EEG-Anlagenstammdaten". Informationsplattform der Deutschen Übertragungsnetzbetreiber, Acc. Mar. 4, 2012. URL: http://www.eeg-kwk.net/de/Anlagenstammdaten.htm.

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/.

Islam, Towhidul and Nigel Meade (2013). “The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation”. *Energy Policy* 55. Special section: Long Run Transitions to Sustainable Economic Structures in the European Union and Beyond, 521–530. DOI: 10.1016/j.enpol.2012.12.041.

Iyengar, Raghuram, Kamel Jedidi, et al. (2011). “The Impact of Tariff Structure on Customer Retention, Usage, and Profitability of Access Services”. *Marketing Science* 30(5), 820–836. DOI: 10.1287/mksc.1110.0655.

Iyengar, Raghuram, Christophe Van den Bulte, and Jae Young Lee (2015). “Social Contagion in New Product Trial and Repeat”. *Marketing Science* 34(3), 408–429. DOI: 10.1287/mksc.2014.0888.

Iyengar, Raghuram, Christophe Van den Bulte, and Thomas W. Valente (2011). “Rejoinder—Further Reflections on Studying Social Influence in New Product Diffusion”. *Marketing Science* 30(2), 230–232. DOI: 10.1287/mksc.1100.0614.

Jimenez, Gabriel et al. (2014). “Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk-Taking?” *Econometrica* 82(2), 463–505. DOI: 10.3982/ECTA10104.

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.

Lin, Wei and Jeffrey M. Wooldridge (2019). “Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models”. In: *Panel Data Econometrics*. Ed. by Mike Tsionas. Academic Press, 21–43. DOI: https://doi.org/10.1016/B978-0-12-814367-4.00002-2.

Llobera, Marcos (2003). “Extending GIS-based visual analysis: the concept of visualscapes”. *International journal of geographical information science* 17(1), 25–48.

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.

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. DOI: 10.2307/2298123.

McShane, Blakeley B., Eric T. Bradlow, and Jonah Berger (2012). "Visual Influence and Social Groups". *Journal of Marketing Research* 49(6), 854–871. DOI: 10.1509/jmr.11.0223.

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.

Nair, Harikesh S., Puneet Manchanda, and Tulikaa Bhatia (2010). "Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders". *Journal of Marketing Research* 47(5), 883–895. DOI: 10.1509/jmkr.47.5.883.

Narayanan, Sridhar and Harikesh S. Nair (2013). "Estimating Causal Installed-Base Effects: A Bias-Correction Approach". *Journal of Marketing Research* 50(1), 70–94. DOI: 10.1509/jmr.11.0183.

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.

Stark, Thomas et al. (2005). *Architektonische Integration von Photovoltaik-Anlagen*. Tech. rep.

Van den Bulte, Christophe and Raghuram Iyengar (2011). "Tricked by Truncation: Spurious Duration Dependence and Social Contagion in Hazard Models". *Marketing Science* 30(2), 233–248. DOI: 10.1287/mksc.1100.0615.

Van den Bulte, Christophe and Gary L. Lilien (2001). "Medical Innovation Revisited: Social Contagion versus Marketing Effort". *American Journal of Sociology* 106(5), 1409–1435. DOI: 10.1086/320819.

Van den Bulte, Christophe and Stefan Stremersch (2004). "Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test". *Marketing Science* 23(4), 530–544. DOI: 10.1287/mksc.1040.0054.

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.

Wittenberg, Inga, Anke Blöbaum, and Ellen Matthies (2018). "Environmental motivations for energy use in PV households: Proposal of a modified norm activation model for the specific context of PV households". *Journal of Environmental Psychology* 55, 110–120. DOI: 10.1016/j.jenvp.2018.01.002.

Wooldridge, Jeffrey (July 2002). *Introductory Econometrics: A Modern Approach*. 2nd ed. South-Western College Pub.

— (July 2012). *Introductory Econometrics: A Modern Approach*. 3rd ed. South-Western College Pub.

Web-Appendix

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

A Data and Modeling Approach

A.1 Logistic Panel Regression

We define the endogenous variable of (1) as

$$y_{n,t} = \begin{cases} 1 & \text{if } u_{n,t} > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

with the latent variable

$$u_{n,t} = v_{n,t} + \epsilon_{n,t} \tag{17}$$

denoting the utility of building owner n to adopt in period t. The random quantity $u_{n,t}$ is decomposed in $v_{n,t} = \boldsymbol{\beta'} \boldsymbol{X}_{n,t}$ (see equation 3) as the deterministic, i.e., observable, utility of n to adopt in t and error term $\epsilon_{n,t}$ that contains unobserved information. Without loss of generality, we normalize the utility of not installing a PV system in period t to 0. Hence, (16) is according to the utility maximization principle (McFadden 1986; McFadden 2001). Since $u_{n,t}$ in (17) is a stochastic variable, the probability that building owner n chooses to install a PV system in period t is

$$P_{n,t} = \Pr\left(y_{n,t} = 1\right) = \Pr\left(u_{n,t} > 0\right). \tag{18}$$

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

$$P_{n,t} = \frac{e^{v_{n,t}}}{1 + e^{v_{n,t}}}. \tag{19}$$

The coefficients $\boldsymbol{\beta'}$ are estimated using a maximum likelihood procedure (Greene 2012). The advantages of (19) over the LPM (4) are that the probabilities are strictly $0 < P_{n,t} < 1$

and the relationship between $\boldsymbol{X}_{n,t}$ and $P_{n,t}$ is valid for the whole domain of $\boldsymbol{X}_{n,t}$. However, high-dimensional fixed effects can not be considered in (17) as outlined in Wooldridge (2012).

A.2 Hazard

We define the hazard of adoption of building owner n in year t as a function of explanatory variables:

$$y_{n,t} = 1 - \exp\left[-\exp\left(\boldsymbol{\beta}'\boldsymbol{X}_{n,t}\right)\right], \tag{20}$$

with $y_{n,t}$ defined in (1) and $\boldsymbol{\beta}'\boldsymbol{X}_{n,t}$ defined in equation (5) (without γ_n because the exponential model does not converge if included. We use the PPMLHDFE package in stata (Correia, Guimarães, and Zylkin 2019) for estimation.

A.3 Raw data on PV systems

Table 5: Frequencies of adoption ($y_{n,t}$) for household PV systems in Baden-Württemberg, 2000-2010.

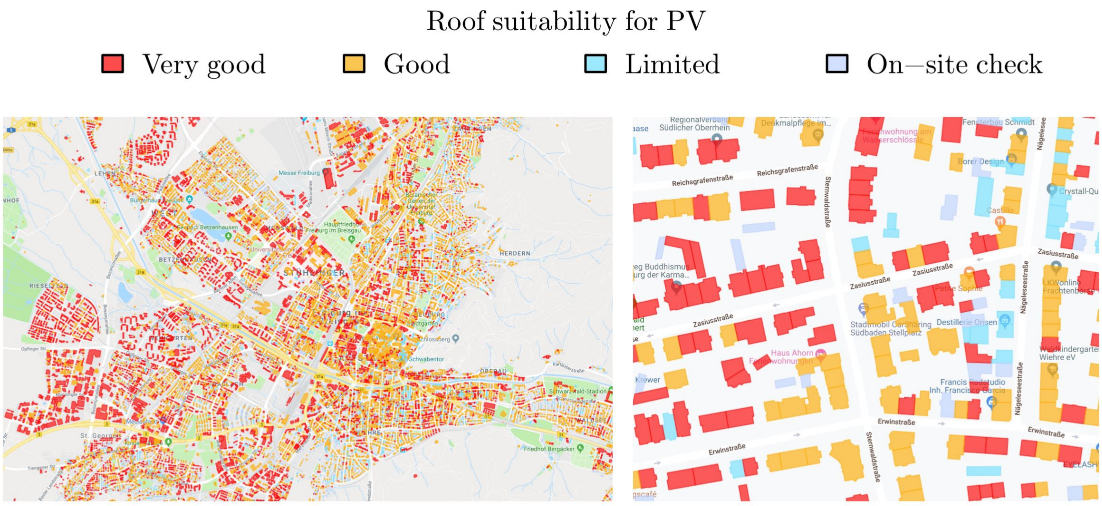
Year$_t$	Alternative	Frequency in category
2000	0: No (or no new) PV system installed	2,956,583
	1: New PV system installed	749
2001	0: No (or no new) PV system installed	2,954,352
	1: New PV system installed	2,980
2002	0: No (or no new) PV system installed	2,952,464
	1: PV system installed	1,888
2003	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.

A.4 Raw data from 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 6 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.



(a) Freiburg (b) Details for Freiburg

Figure 6: Screenshots of information from roof-census for Freiburg in Baden-Württemberg. Panel (b) shows the same excerpt of Freiburg as Figure 4.

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/dachflachen/potenzial-dachflachenanlagen. 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 6, 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.[16] The following website shows the high-resolution data in general: https://www.energieatlas-bw.de/sonne/dachflachen/potenzial-dachflachenanlagen (last visit: 02.10.2019). Zooming and clicking on a specific roof indicates the roof's appropriate PV area (in German: Mögliche geeignete Modulfläche).

Orientation and inclination determine roof suitability for PV. Table 6 contains the descriptive statistics for the raw data on orientation and inclination.

Table 6: Descriptive statistics for roof inclination and orientation.

	Mean	Std. Dev.	Min.	Max.
Inclination^*_n	30	18	0	74
Orientation^*_n	152	81	-1	359
N	3,006,675			

Notes: We have information on roof inclination and roof orientation for 3,006,675 buildings 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.

[16] However, 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.

A.5 Data

Table 7: Descriptive statistics for other installed base measures.

	Mean	Std. Dev.	Min.	Max.
Installed base (other baseline measures):				
$IB^{1/d}_{n,t-1}$	.00016	.00027	0	.011
$IB^{D\leq 100}_{n,t-1}$	.0047	.0091	0	.33
$IB^{100<D\leq 200}_{n,t-1}$	.011	.017	0	.39
$IB^{200<D\leq 400}_{n,t-1}$	.03	.043	0	.69
Installed base ($t-2$ lag):				
$IB_{n,t-2}$	.012	.019	0	.43
Other installed base measures:				
Road 50m: $IB^{\text{NR50}}_{n,t-1}$	.015	.023	0	.44
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	.0012	.0043	0	.24
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	.013	.02	0	.36
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$	.015	.026	0	1
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$	.0013	.0075	0	1
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$	.012	.023	0	1
N	29,380,453			

Notes: The 29,380,453 observations come from 2,957,332 buildings n over 10 years t (2001-2010).

Table 8: Descriptive statistics, all PV systems.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $\Delta y^{\text{All}}_{n,t}$	.0052	.072	0	1
Installed base (direct visibility):				
No building in-between: $IB^{\sphericalangle}_{n,t-1}$	.00045	.0022	0	.07
No building in-between (rel.): $IB^{\text{Relative},\sphericalangle}_{n,t-1}$	.00056	.0044	0	.67
Installed base (no direct visibility):				
Building in-between: $IB^{\not\sphericalangle}_{n,t-1}$	.0032	.0074	0	.34
Building in-between: $IB^{\text{Relative},\not\sphericalangle}_{n,t-1}$	.0028	.0078	0	.5
Other installed base measures:				
Road 50m: $IB^{\text{NR50}}_{n,t-1}$	.015	.023	0	.44
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	.0011	.0043	0	.24
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	.013	.02	0	.36
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$	.015	.026	0	1
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$	.0013	.0074	0	1
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$	.012	.023	0	1
Installed base (baseline measures):				
$IB_{n,t-1}$	.015	.023	0	.44
$IB^{\text{Relative}}_{n,t-1}$	.015	.026	0	1
Instruments:				
Avg. inclination of nearby roofs with $\text{PV}^{\sphericalangle}_{n,t-1}$	.00031	.0016	0	.01
Avg. orientation of nearby roofs with $\text{PV}^{\sphericalangle}_{n,t-1}$	.00033	.0017	0	.01
N	29,141,406			

Notes: In this specification, the 29,141,406 observations come from 2,776,179 choice makers (buildings) n over 10 years t (2001-2010). In total, we have 29,580,810 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 29,141,406 observations. The choice makers are distributed across 8982 districts i, across 405,547 District-Street groups and across 4,028,910 District-Street-Year groups in Baden-Wuerttemberg. Note that an increase of the installed base measure $IB^{\sphericalangle}_{n,t-1}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n.

Table 9: Descriptive statistics for (lagged) peer effects.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $y_{n,t}$	.0026	.051	0	1
Installed base (direct visibility) [$t-2$ lag]:				
No building in-between: $IB^{\triangleleft}_{n,t-2}$	.00037	.002	0	.07
No building in-between: $IB^{\text{Relative},\triangleleft}_{n,t-2}$	.007	.014	0	.67
Installed base (no direct visibility):				
Building in-between: $IB^{\not\triangleleft}_{n,t-1}$	.0035	.0078	0	.34
Building in-between: $IB^{\text{Relative},\not\triangleleft}_{n,t-1}$	.0031	.0083	0	.6
Other installed base measures:				
Road 50m: $IB^{\text{NR50}}_{n,t-1}$	.017	.024	0	.44
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	.0013	.0045	0	.24
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	.014	.021	0	.36
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$	.016	.027	0	1
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$	.0014	.0079	0	1
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$	.013	.024	0	1
Installed base (baseline measures):				
$IB_{n,t-1}$	.017	.024	0	.44
$IB^{\text{Relative}}_{n,t-1}$	.016	.027	0	1
Instruments (direct visibility) [$t-2$ lag]:				
Avg. inclination of nearby roofs with $\text{PV}^{\triangleleft}_{n,t-2}$	.00025	.0014	0	.01
Avg. orientation of nearby roofs with $\text{PV}^{\triangleleft}_{n,t-2}$	.00027	.0015	0	.01
N	26,423,542			

Notes: In this specification, the 26,423,542 observations come from 2,777,799 buildings n over 9 years t (2002-2010) for the dependent variable. In total, we have 26,622,729 observations. We censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 26,423,542 observations. The choice makers are distributed across 405,665 District-Street groups and across 3,639,946 District-Street-Year groups in Baden-Wuerttemberg. Note that an increase of the installed base measure $IB^{\triangleleft}_{n,t-2}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n.

Table 10: Descriptive statistics for placebo-lead peer effects.

	Mean	Std. Dev.	Min.	Max.
New PV installation: $\Delta y_{n,t}$	.0017	.041	0	1
Installed base (direct visibility) [$t+1$ lead]:				
No building in-between: $IB^{\triangleleft}_{n,t+1}$	.00056	.0025	0	.07
No building in-between: $IB^{\text{Relative},\triangleleft}_{n,t+1}$	.0007	.0049	0	.67
No building in-between: $IB^{\triangleleft}_{n,t-1}$	.0003	.0018	0	.07
No building in-between (rel.): $IB^{\text{Relative},\triangleleft}_{n,t-1}$	.00036	.0035	0	.67
Installed base (no direct visibility):				
Building in-between: $IB^{\not\triangleleft}_{n,t-1}$	.0021	.0056	0	.34
Building in-between: $IB^{\text{Relative},\not\triangleleft}_{n,t-1}$	.0018	.006	0	.5
Other installed base measures:				
Road 50m: $IB^{\text{NR50}}_{n,t-1}$	.01	.016	0	.43
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	.00074	.0033	0	.24
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	.0082	.014	0	.25
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$	.0096	.019	0	1
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$	.00083	.0057	0	1
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$	.0079	.017	0	1
Installed base (baseline measures):				
$IB_{n,t-1}$	.01	.016	0	.43
$IB^{\text{Relative}}_{n,t-1}$	.0096	.019	0	1
N	23,560,023			

Notes: In this specification, the 23,560,023 observations come from 2,779,566 buildings \$n\$ over 9 years \$t\$ (2002-2010) for the dependent variable. We censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. The buildings are distributed across 8982 districts i, across 405,784 District-Street groups and across 3,240,667 District-Street-Year groups in Baden-Wuerttemberg. Note that an increase of the installed base measure $IB^{\triangleleft}_{n,t-2}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n.

B Robustness analyses

B.1 Baseline installed base

Table 11: Robustness of estimates for peer effects.

	Time-varying	Cut-off D	Distance-weighted	Relative IB		
	OLS	OLS	OLS	OLS	1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)	(6)
	$y_{n,t}$	$y_{n,t}$	$y_{n,t}$	$y_{n,t}$	$IB_{n,t-1}^{\text{Relative}}$	$y_{n,t}$
Installed base:						
$IB_{n,t-1}$	0.0086		0.011^{**}	0.0010	0.66^{**}	0.024^{**}
	(1.44)		(3.62)	(0.31)	(285.2)	(2.94)
$IB_{n,t-1} \times \text{Period}_t^{\text{Since 2005}}$	0.0032					
	(0.57)					
$IB_{n,t-1}^{D \leq 100}$		0.031^{**}				
		(6.49)				
$IB_{n,t-1}^{100<D \leq 200}$		0.0049				
		(1.54)				
$IB_{n,t-1}^{200<D \leq 400}$		-0.0038^{+}				
		(-1.90)				
Installed base (distance-weighted):						
$IB_{n,t-1}^{1/d}$			0.034			
			(0.21)			
Relative installed base:						
$IB_{n,t-1}^{\text{Relative}}$				0.015^{**}		
				(6.99)		
Predicted relative installed base:						
$\widehat{IB}_{n,t-1}^{\text{Relative}}$						-0.017^{+}
						(-1.70)
Instruments:						
Avg. inclination of nearby roofs with $\text{PV}_{n,t-1}$					0.29^{**}	
					(18.9)	
Avg. orientation of nearby roofs with $\text{PV}_{n,t-1}$					0.57^{**}	
					(38.3)	
Observations	27,640,461	27,640,461	27,640,461	27,630,743	27,630,743	27,630,743
DF_M	4	5	4	4	5	4
Final log-likelihood	49,523,950	49,523,997	49,523,950	49,505,797		
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level); $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
the dependent variable is always $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. All columns are estimated via OLS. The panel of 2,957,332 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. Across columns, the sample is always the same and singleton observations are dropped. If not indicated otherwise, the Cut-off distance D = 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of all the (non-distance-weighted, non-relative) installed base measures of 0.01 refers to 1 additional previously installed PV system nearby. For the distance-weighted measure $IB_{n,t-1}^{1/d}$, an increase of 0.001, e.g., refers to 1 previously installed PV system in 10m distance to the building n.

Table 12: Placebo estimates for peer effects and longer lag.

	Random allocation	Large systems	Two-year lag	Both lags
	OLS	OLS	OLS	OLS
	(1) $y_{n,t}^{\text{Random}}$	(2) $y_{n,t}^{\text{Large PV}}$	(3) $y_{n,t}$	(4) $y_{n,t}$
Installed base:				
$IB_{n,t-1}$	-0.0047^{+}	-0.00078		0.0091^{*}
	(-1.84)	(-1.21)		(2.16)
Installed base ($t-2$ lag):				
$IB_{n,t-2}$			0.012^{**}	0.0046
			(3.59)	(0.98)
Observations	24,858,632	27,839,032	24,858,172	24,858,172
DF_M	3	3	3	4
Adj. R^2	-0.02	0.10	0.06	0.06
F	1.1	34.4	334.3	251.2
Year fixed effects	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
In column (1), the dependent variable is $y_{n,t}^{\text{Random}}$, i.e., whether a (new) PV system was installed on building n in t. In this case, the same number of PV systems that were actually installed are randomly allocated to buildings and real PV systems are ignored. In column (2), the dependent variable is $y_{n,t}^{\text{Large PV}}$, i.e., whether a (new) large PV (non-household) system (larger than 30kW_p) PV system was installed on building n in t. In columns (3) and (4), the dependent variable is $y_{n,t}$, i.e., building n's observed choice in terms of (newly) installing a PV system in t. As always, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. We estimate all columns via OLS. Singleton observations are dropped. This procedure results in fewer observations. Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of the installed base measures of 0.01 refers to 1 additional previously installed PV system nearby.

B.2 Direct visibility

B.2.1 No causal effect from remaining PV systems

Table 13: Estimates for peer effects visibility measures.

	Absolute	
	IV	
	1st stage	2nd stage
	(1) $IB^{Rest}_{n,t-1}$	(2) $y_{n,t}$
Predicted installed base controls (no direct visibility):		
Rest: $\widehat{IB}^{Rest}_{n,t-1}$		-0.0085
		(-0.69)
Instruments (average ratio over previous PV nearby):		
Avg. inclination ratio: $\text{AvgIncRatio}_{n,t-1}$	0.22^{**}	
	(18.5)	
Avg. orientation ratio: $\text{AvgOrRatio}_{n,t-1}$	0.44^{**}	
	(38.7)	
Installed base (direct visibility):		
No building in-between: $IB^{\sphericalangle}_{n,t-1}$	-0.57^{**}	0.036^{**}
	(-186.3)	(2.98)
Installed base controls (no direct visibility):		
Building in-between: $IB^{\not\sphericalangle}_{n,t-1}$	-0.43^{**}	-0.0031
	(-144.1)	(-0.47)
Observations	27,640,461	27,640,461
DF_M	6	5
Adj. R^2		0.00
F		210.8
Hansen J (p-value)		0.3 (0.61)
Kleibergen-Paap rk Wald F statistic		14,587.8
Year fixed effects	Yes	Yes
District-Street-Year fixed effects	Yes	Yes
Building fixed effects	Yes	Yes
Controls	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ In columns (1,3), the dependent variable is $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. The panel of 2,957,332 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. We estimate all columns via OLS. Singleton observations are dropped. This procedure results in fewer observations. Column (2) uses the predicted values for the installed base $\left(IB^{Rest}_{n,t-1}\right)$ from the first stage in column (1). Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of the installed base measure $IB^{\sphericalangle}_{n,t-1}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n. $IB^{Rest}_{n,t-1} = IB_{n,t-1} - IB^{\sphericalangle}_{n,t-1} - IB^{\not\sphericalangle}_{n,t-1}$.

B.2.2 More installed base measures

The measure $IB^{\text{NR50}}_{n,t-1}$ comprises only PV systems which are located at most 50 meters away from the closest street.

Table 14: Estimates for peer effects visibility measures, more installed base measures.

	Absolute			Relative		
	OLS	IV		OLS	IV	
		1st stage	2nd stage		1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)	(6)
	$y_{n,t}$	$IB^{\triangleleft}_{n,t-1}$	$y_{n,t}$	$y_{n,t}$	$IB^{\text{Relative},\triangleleft}_{n,t-1}$	$y_{n,t}$
Installed base (direct visibility):						
No building in-between: $IB^{\triangleleft}_{n,t-1}$	0.034^{**}					
	(3.30)					
No building in-between (relative): $IB^{\text{Relative},\triangleleft}_{n,t-1}$				0.021^{**}		
				(3.35)		
Predicted installed base (direct visibility):						
No building in-between: $\widehat{IB}^{\triangleleft}_{n,t-1}$			0.025^{*}			
			(2.25)			
No building in-between (relative): $\widehat{IB}^{\text{Relative},\triangleleft}_{n,t-1}$						0.16^{**}
						(2.67)
Instruments (average ratio over previous visible PV nearby)::						
Avg. inclination ratio: $\text{AvgIncRatio}^{\triangleleft}_{n,t-1}$		0.42^{**}			0.037^{**}	
		(147.2)			(4.44)	
Avg. orientation of nearby roofs with $\text{PV}^{\triangleleft}_{n,t-1}$		0.81^{**}			0.18^{**}	
		(315.2)			(22.9)	
Installed base controls (no direct visibility):						
Building in-between: $IB^{\ntriangleleft}_{n,t-1}$	−0.0064	-0.021^{**}	−0.0074			
	(−1.32)	(−88.7)	(−1.52)			
Building in-between: $IB^{\text{Relative},\ntriangleleft}_{n,t-1}$				0.0037	0.14^{**}	-0.021^{*}
				(0.91)	(34.4)	(−2.30)
Other installed base measures:						
Next road 50m: $IB^{NR50}_{n,t-1}$	−0.073	0.0080	−0.072			
	(−0.50)	(1.55)	(−0.50)			
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	0.028^{*}	-0.0062^{**}	0.028^{*}			
	(2.10)	(−11.9)	(2.11)			
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	0.013	-0.0018^{**}	0.013			
	(1.50)	(−5.01)	(1.51)			
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$				0.016	0.43^{**}	−0.053
				(0.24)	(4.58)	(−0.69)
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$				0.0061	-0.21^{**}	0.040^{**}
				(0.91)	(−22.2)	(2.76)
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$				0.012^{*}	0.15^{**}	−0.011
				(2.36)	(15.2)	(−1.11)
Installed base (baseline measures):						
$IB_{n,t-1}$	0.062	0.0025	0.062	0.00089	0.23^{**}	-0.036^{*}
	(0.43)	(0.48)	(0.43)	(0.28)	(51.0)	(−2.54)
$IB^{\text{Relative}}_{n,t-1}$	0.015^{**}	0.00044^{**}	0.015^{**}	−0.013	-0.27^{**}	0.032
	(6.80)	(6.16)	(6.81)	(−0.19)	(−2.93)	(0.43)
Observations	27,630,743	27,630,743	27,630,743	27,630,743	27,630,743	27,630,743
DF_{M}	9	10	9	9	10	9
Adj. R^2	0.06		0.00	0.06		-0.00
F	123.1		122.7	123.0		121.9
Hansen J (p-value)			0.3 (0.57)			0.5 (0.47)
Kleibergen-Paap rk Wald F statistic			1,366,109.3			1,833.2
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
In columns (1, 3, 4, 6), the dependent variable is $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. The panel of 2,957,332 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,380,453 observations. We estimate all columns via OLS. Singleton observations are dropped. This procedure results in fewer observations. Column (3) uses the predicted values for the installed base $\left(IB^{\triangleleft}_{n,t-1}\right)$ from the first stage in column (2). Column (6) uses the predicted values for the relative installed base $\left(IB^{\text{Relative},\triangleleft}_{n,t-1}\right)$ from the first stage in column (5). Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of the installed base measure $IB^{\triangleleft}_{n,t-1}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n..

B.2.3 All PV systems

Table 15: Estimates for peer effects visibility measures, all PV systems (and more installed base measures).

	Absolute			Relative		
	OLS	IV		OLS	IV	
		1st stage	2nd stage		1st stage	2nd stage
	(1) $y^{\text{All}}_{n,t}$	(2) $IB^{\triangleleft}_{n,t-1}$	(3) $y^{\text{All}}_{n,t}$	(4) $y^{\text{All}}_{n,t}$	(5) $IB^{\text{Relative},\triangleleft}_{n,t-1}$	(6) $y^{\text{All}}_{n,t}$
Installed base (direct visibility):						
No building in-between: $IB^{\triangleleft}_{n,t-1}$	0.055^{**}					
	(3.65)					
No building in-between (relative): $IB^{\text{Relative},\triangleleft}_{n,t-1}$				0.024^{*}		
				(2.39)		
Predicted installed base (direct visibility):						
No building in-between: $\widehat{IB}^{\triangleleft}_{n,t-1}$			0.046^{**}			
			(2.81)			
No building in-between (relative): $\widehat{IB}^{\text{Relative},\triangleleft}_{n,t-1}$						0.30^{**}
						(3.08)
Instruments (average ratio over previous visible PV nearby):						
Avg. inclination ratio: $\text{AvgIncRatio}^{\triangleleft}_{n,t-1}$		0.42^{**}			0.035^{**}	
		(146.5)			(4.50)	
Avg. orientation ratio: $\text{AvgOrRatio}^{\triangleleft}_{n,t-1}$		0.81^{**}			0.17^{**}	
		(313.8)			(22.1)	
Installed base controls (no direct visibility):						
Building in-between: $IB^{\ntriangleleft}_{n,t-1}$	-0.0018	-0.021^{**}	-0.0029			
	(-0.25)	(-88.6)	(-0.39)			
Building in-between: $IB^{\text{Relative},\ntriangleleft}_{n,t-1}$				0.0069	0.13^{**}	-0.032^{*}
				(1.03)	(30.8)	(-2.43)
Other installed base measures:						
Next road 50m: $IB^{NR50}_{n,t-1}$	-0.059	0.0086^{+}	-0.058			
	(-0.31)	(1.70)	(-0.31)			
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	0.019	-0.0061^{**}	0.019			
	(0.87)	(-11.6)	(0.87)			
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	-0.00087	-0.0018^{**}	-0.00080			
	(-0.061)	(-4.83)	(-0.055)			
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$				0.046	0.40^{**}	-0.075
				(0.48)	(4.86)	(-0.67)
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$				0.0094	-0.24^{**}	0.082^{**}
				(0.57)	(-22.8)	(2.85)
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$				0.0071	0.18^{**}	-0.047^{*}
				(0.67)	(16.2)	(-2.26)
Installed base (baseline measures):						
$IB_{n,t-1}$	0.055	0.0016	0.055	-0.0030	0.18^{**}	-0.059^{**}
	(0.29)	(0.32)	(0.29)	(-0.54)	(34.9)	(-3.11)
$IB^{\text{Relative}}_{n,t-1}$	0.013^{**}	0.00069^{**}	0.013^{**}	-0.040	-0.21^{*}	0.023
	(2.69)	(7.01)	(2.70)	(-0.43)	(-2.53)	(0.22)
Observations	27,387,352	27,387,352	27,387,352	27,387,352	27,387,352	27,387,352
DF_{M}	9	10	9	9	10	9
Adj. R^2	0.06		0.00	0.06		-0.00
F	60.7		60.2	60.0		60.0
Hansen J (p-value)			0.2 (0.69)			0.3 (0.58)
Kleibergen-Paap rk Wald F statistic			1,354,668.5			1,598.6
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

In columns (1, 3, 4, 6), the dependent variable is $y^{\text{All}}_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. In this table, we include all PV systems no matter how well the allocation procedure to buildings worked out. The panel of 2,776,179 buildings over 10 years t (2001-2010) results in 29,580,810 observations. Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with 29,141,406 observations. We estimate all columns via OLS. Singleton observations are dropped. This procedure results in fewer observations. Column (3) uses the predicted values for the installed base $\left(IB^{\triangleleft}_{n,t-1}\right)$ from the first stage in column (2). Column (6) uses the predicted values for the relative installed base $\left(IB^{\text{Relative},\triangleleft}_{n,t-1}\right)$ from the first stage in column (5). Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of the installed base measure $IB^{\triangleleft}_{n,t-1}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n..

B.2.4 Two-years lag and placebo-lead

Table 16: Estimates for (lagged) peer effects (or with lead).

	Lagged			Lead
	OLS	IV		OLS
		1st stage	2nd stage	
	(1) $y_{n,t}$	(2) $IB^{\triangleleft}_{n,t-2}$	(3) $y_{n,t}$	(4) $y_{n,t}$
Installed base (direct visibility) [$t+1$ lead]:				
No building in-between: $IB^{\triangleleft}_{n,t+1}$				-0.29^{**}
				(-27.0)
Installed base (direct visibility):				
No building in-between: $IB^{\triangleleft}_{n,t-1}$				0.16^{**}
				(12.8)
Installed base (direct visibility) [$t-2$ lag]:				
No building in-between: $IB^{\triangleleft}_{n,t-2}$	0.023^{+}			
	(1.91)			
Predicted installed base (direct visibility) [$t-2$ lag]:				
No building in-between: $\widehat{IB}^{\triangleleft}_{n,t-2}$			0.021^{+}	
			(1.66)	
Instruments (average ratio over previous visible PV nearby) [$t-2$ lag]:				
Avg. inclination ratio: $\text{AvgIncRatio}^{\triangleleft}_{n,t-2}$		0.42^{**}		
		(134.7)		
Avg. orientation ratio: $\text{AvgOrRatio}^{\triangleleft}_{n,t-2}$		0.81^{**}		
		(287.5)		
Installed base control (no direct visibility):				
Building in-between: $IB^{\not\triangleleft}_{n,t-1}$	-0.0090^{+}	-0.012^{**}	-0.0091^{+}	-0.010^{+}
	(-1.75)	(-67.6)	(-1.76)	(-1.85)
Other installed base measures:				
Next road 50m: $IB^{NR50}_{n,t-1}$	-0.076	0.0037	-0.076	-0.046
	(-0.43)	(0.85)	(-0.43)	(-0.27)
Flat roof: $IB^{\text{Flat}}_{n,t-1}$	0.030^{*}	-0.0044^{**}	0.030^{*}	0.022
	(2.10)	(-10.8)	(2.10)	(1.43)
No flat roof: $IB^{\text{No flat}}_{n,t-1}$	0.014	-0.0014^{**}	0.014	0.013
	(1.44)	(-4.91)	(1.45)	(1.27)
Installed base (baseline measures):				
$IB_{n,t-1}$	0.066	0.0024	0.066	0.039
	(0.37)	(0.54)	(0.37)	(0.23)
$IB^{\text{Relative}}_{n,t-1}$	0.015^{**}	0.00017^{**}	0.015^{**}	0.012^{**}
	(6.56)	(3.36)	(6.56)	(4.86)
Observations	24,849,434	24,849,434	24,849,434	22,160,218
DF_M	9	10	9	10
Adj. R^2	0.06		0.00	0.07
F	116.5		116.4	138.0
Hansen J (p-value)			0.2 (0.68)	
Kleibergen-Paap rk Wald F statistic			1,187,197.9	
Year fixed effects	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and robust t statistics in parentheses; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ the dependent variable is always $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. We estimate estimate all columns via OLS. The panel of almost 2,8000,000 buildings over 10 years t (for the dependent variable, columns (1-3): 2002-2010, columns (4): 2001-2008) results in more than 20,000,000 observations. Table 9 contains details and descriptive statistics for columns (1-3) and Table 10 for column (4). Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with fewer observations. Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year. Note that an increase of the installed base measure $IB^{\triangleleft}_{n,t-1}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n.

Table 17: Estimates for (lagged) peer effects (or with lead).

	Lagged			Lead
	OLS	IV		OLS
		1st stage	2nd stage	
	(1)	(2)	(3)	(4)
	$y_{n,t}$	$IB^{\text{Relative},\triangleleft}_{n,t-2}$	$y_{n,t}$	$y_{n,t}$
Installed base (direct visibility) [t + 1 lead]:				
No building in-between: $IB^{\text{Relative}}_{n,t+1}$				-0.16^{**}
				(-21.5)
Installed base (direct visibility):				
No building in-between (relative): $IB^{\text{Relative},\triangleleft}_{n,t-1}$				0.095^{**}
				(11.4)
Installed base (direct visibility) [$t-2$ lag]:				
No building in-between (relative) [$t-2$ lag]: $IB^{\text{Relative},\triangleleft}_{n,t-2}$	0.019^{**}			
	(2.60)			
Predicted installed base (direct visibility) [$t-2$ lag]:				
No building in-between (relative) [$t-2$ lag]: $\widehat{IB}^{\text{Relative},\triangleleft}_{n,t-2}$			0.12^{*}	
			(1.97)	
Instruments (average ratio over previous visible PV nearby) [$t-2$ lag]:				
Avg. inclination ratio: $\text{AvgIncRatio}^{\triangleleft}_{n,t-2}$		0.044^{**}		
		(4.48)		
Avg. orientation ratio: $\text{AvgOrRatio}^{\triangleleft}_{n,t-2}$		0.21^{**}		
		(22.5)		
Installed base control (no direct visibility):				
Building in-between: $IB^{\text{Relative},\ntriangleleft}_{n,t-1}$	0.0021	0.086^{**}	-0.0090	0.00070
	(0.48)	(27.7)	(-1.39)	(0.15)
Other installed base measures:				
Road 50m: $IB^{\text{Relative, NR50}}_{n,t-1}$	0.010	0.34^{**}	-0.030	0.047
	(0.12)	(4.00)	(-0.34)	(0.61)
Flat roof: $IB^{\text{Relative, flat roof}}_{n,t-1}$	0.0060	-0.13^{**}	0.022^{*}	0.0037
	(0.84)	(-20.3)	(2.03)	(0.55)
No flat roof: $IB^{\text{Relative, no flat roof}}_{n,t-1}$	0.012^{*}	0.089^{**}	0.0015	0.0084
	(2.24)	(13.8)	(0.19)	(1.50)
Installed base (baseline measures):				
$IB_{n,t-1}$	0.00087	0.14^{**}	-0.016^{+}	0.0014
	(0.25)	(45.9)	(-1.75)	(0.38)
$IB^{\text{Relative}}_{n,t-1}$	-0.0059	-0.25^{**}	0.024	-0.042
	(-0.070)	(-2.93)	(0.28)	(-0.55)
Observations	24,849,434	24,849,434	24,849,434	22,160,218
DF_M	9	10	9	10
Adj. R^2	0.06		-0.00	0.07
F	116.5		116.3	112.8
Hansen J (p-value)			0.1 (0.76)	
Kleibergen-Paap rk Wald F statistic			2,815.8	
Year fixed effects	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and robust t statistics in parentheses; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ the dependent variable is always $y_{n,t}$, i.e., whether a (new) PV system was installed on building n in t. We estimate estimate all columns via OLS. The panel of almost 2,800,000 buildings over 10 years t (for the dependent variable, columns (1-3): 2002-2010, columns (4): 2001-2008) results in more than 20,000,000 observations. Table 9 contains details and descriptive statistics for columns (1-3) and Table 10 for column (4). Across columns, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. In consequence, we end up with fewer observations. Across columns, the Cut-off distance D is 200m. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year.

B.2.5 Random allocation and large systems

Table 18: Placebo estimates for peer effects visibility measures, more installed base measures.

	Random allocation		Large systems	
	Absolute	Relative	Absolute	Relative
	OLS	OLS	OLS	OLS
	(1) $y_{n,t}^{\text{Random}}$	(2) $y_{n,t}^{\text{Random}}$	(3) $y_{n,t}^{\text{Large PV}}$	(4) $y_{n,t}^{\text{Large PV}}$
Installed base (direct visibility):				
No building in-between: $IB_{n,t-1}^{\triangleleft}$	0.0095		0.0026	
	(1.12)		(1.13)	
No building in-between (relative): $IB_{n,t-1}^{\text{Relative},\triangleleft}$		0.0016		0.0012
		(0.35)		(0.48)
Instruments (direct visibility):				
Avg. inclination of nearby roofs with $\text{PV}_{n,t-1}^{\triangleleft}$				
Avg. orientation of nearby roofs with $\text{PV}_{n,t-1}^{\triangleleft}$				
Installed base controls (no direct visibility):				
Building in-between: $IB_{n,t-1}^{\ntriangleleft}$	0.0030		−0.00039	
	(0.70)		(−0.40)	
Building in-between: $IB_{n,t-1}^{\text{Relative},\ntriangleleft}$		−0.00079		−0.00021
		(−0.24)		(−0.14)
Other installed base measures:				
Next road 50m: $IB_{n,t-1}^{NR50}$	0.15		−0.0073	
	(1.06)		(−0.52)	
Flat roof: $IB_{n,t-1}^{\text{Flat}}$	0.018		0.0034	
	(1.39)		(0.92)	
No flat roof: $IB_{n,t-1}^{\text{No flat}}$	0.0032		−0.0015	
	(0.36)		(−0.74)	
Road 50m: $IB_{n,t-1}^{\text{Relative, NR50}}$		0.15		0.012
		(1.37)		(0.86)
Flat roof: $IB_{n,t-1}^{\text{Relative, flat roof}}$		0.010		0.0023
		(1.33)		(0.47)
No flat roof: $IB_{n,t-1}^{\text{Relative, no flat roof}}$		0.0058		−0.0021
		(1.02)		(−0.48)
Installed base (baseline measures):				
$IB_{n,t-1}$	−0.16	−0.0026	0.0060	−0.0023*
	(−1.11)	(−0.84)	(0.44)	(−2.44)
$IB_{n,t-1}^{\text{Relative}}$	−0.0026	−0.16	0.0022*	−0.0088
	(−1.10)	(−1.45)	(2.10)	(−0.63)
Observations	24,849,894	24,849,894	27,829,189	27,829,189
DF_M	9	9	9	9
Adj. R^2	-0.02	-0.02	0.10	0.10
F	1.0	0.9	12.9	12.7
Year fixed effects	Yes	Yes	Yes	Yes
District-Street-Year fixed effects	Yes	Yes	Yes	Yes
Building fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Across columns, we show coefficients and t statistics in parentheses (based on SE clustered at street-level; $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
In columns (1, 2), the dependent variable is $y_{n,t}^{\text{Random}}$, i.e., whether a (new) PV system was installed on building n in t. In this case, the same number of PV systems that were actually installed are randomly allocated to buildings and real PV systems are ignored. In columns (3, 4), the dependent variable is $y_{n,t}^{\text{Large PV}}$, i.e., whether a (new) large PV (non-household) system (larger than 30kW_p) was installed on building n in t. As always, we censor the data in the way that we ignore buildings once they have installed a PV system because usually only one PV system can be installed. We estimate all columns via OLS. Singleton observations are dropped. This procedure results in fewer observations. Across columns, the Cut-off distance D is 200m. Note that an increase of the installed base measure $IB_{n,t-1}^{\triangleleft}$ of 0.01 refers to 1 additional previously installed PV system nearby, which is directly visible from n. The installed base measures are the same as in Table 14: they are based on the actually installed household systems. We include n-specific time-variant controls: potential electricity production times year and maximum module area times year.

References Web-Appendix

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin (2019). *ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects.*

Greene, W.H. (2012). *Econometric Analysis.* 7th ed. New Jersey: Pearson Education.

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.

McFadden, D. (1986). "The choice theory approach to market research". *Marketing Science* 5(4), 275–297. DOI: 10.1287/mksc.5.4.275.

— (2001). "Economic choices". *American Economic Review* 91(3), 351–378. DOI: 10.1257/aer.91.3.351.

Train, Kenneth E. (2009). *Discrete choice methods with simulation.* Cambridge University Press.

Wooldridge, Jeffrey (July 2012). *Introductory Econometrics: A Modern Approach.* 3rd ed. South-Western College Pub.