UNIVERSITY OF MANNHEIM

University of Mannheim / Department of Economics

Working Paper Series

Does social interaction make bad policies even worse? Evidence from renewable energy subsidies

Justus Inhoffen Christoph Siemroth Philipp Zahn

Working Paper 16-09

June 2016

Does social interaction make bad policies even worse? Evidence from renewable energy subsidies*

Justus Inhoffen
University of Amsterdam

Christoph Siemroth
University of Mannheim

Philipp Zahn
University of St. Gallen

June 15, 2016

Abstract

Minimum prices above the market level can lead to inefficient production and oversupply. We investigate whether this effect is even more pronounced when decision makers are influenced by their social environment. Using data of minimum prices for renewable energy production in Germany, we analyze if individual decisions to install solar panels are affected by the investment decisions of others. We implement a propensity score matching routine on municipality level and estimate that existing panels in the municipality increase the probability and number of further installations considerably, even in areas with minimal solar potential. This social effect is stronger in areas with more solar potential and less unemployment. A higher number of existing panels and more concentrated installations increase the social effect further. We discuss policy implications of these social effects.

Keywords: EEG, Minimum Prices, Peer Effects, Public Policy, Renewable Energy, Social Interaction, Social Effect, Social Multiplier, Solar Power, Solar Panels, Subsidy
JEL Classification: H23, L14, Q42, Q48, Q58

*We thank Sebastian Siegloch and Max Loeffler for providing us with German municipality structure data. We are grateful to Andreas Peichl and to seminar participants at the University of Mannheim and ZEW for very helpful comments. This research was supported by the German Research Foundation (DFG) via SFB 884. Contact: j.h.inhoffen[at]tinbergen.nl; christoph.siemroth[at]gess.uni-mannheim.de; philipp.zahn[at]unisg.ch

1 Introduction

> Failing to include network structure can lead one to a deficient understanding of an observed behavior and can lead to poor policy design. — Jackson (2016)

Social influence by family, friends, or colleagues on decisions have been empirically documented and theoretically analyzed in many settings. However, in the design of new policies neither empirical evidence nor theoretical insights are typically considered, even if policy success depends on the effect of social interactions. In this paper, we investigate the consequences of ignoring social interactions for one particular policy instrument: minimum prices. We study minimum prices for a large government program in Germany with the goal of expanding private solar (photovoltaic) panel installations. We find that the participation rate was substantially increased due to social interactions even in areas with minimal solar potential. Thus, from a welfare perspective, social interactions may exacerbate the known allocative deficiencies of minimum prices.

In the context of power production, due to its spatial and temporal dimension, there are additional, direct consequences of social effects for the allocative efficiency of individual investment decisions. Local clustering of renewable energy generators due to social interaction makes diversification against weather shocks less effective, requiring a higher overall capacity. Moreover, it strains the local power grid if clustering leads to local excess supply of energy. In this case, clustering might make more long-range power transmission lines necessary, which are very costly, imply higher energy loss, and are often met with a strong 'not in my backyard'-type of local opposition, citing the significant change in the landscape among other things (e.g., WSJ, 2009; FT, 2015).

The program we study is one of the largest renewable energy subsidy schemes in the world, the German renewable energy act (Erneuerbare Energien Gesetz, EEG). With the EEG, the share of green energy in Germany increased from 6% to 32% within 15 years (BMWi, 2016); it is estimated that around 24.6 billion EUR will be spent in 2016 alone (P3 Group, 2015). The EEG regulation guaranteed (i) a fixed compensation rate over 20 years for every kWh of solar energy produced (feed-in tariff), which was above the market compensation, and (ii) required energy network operators to feed in green energy independent of time, location, demand, supply, or energy prices.

According to our estimates, the effect of social interaction is large and increases the base rate probability of solar panel installations by about 25% and the number of installations by about 50%. These estimates increase with more panels in the area. Hence, we provide evidence that neighbors or other households in the same municipality influence peers in the decision to install a solar panel and use the EEG public support scheme. We also find that social influences have immediate effect and weaken over time. Moreover, we find that social influence has a stronger effect in municipalities with more solar radiation and less unemployment. This finding suggests that the social effect is strongest if it is attractive, based on the fundamentals, to install a panel. The social effect is positive even in the areas with minimal solar radiation, which suggests that social interactions can exacerbate inefficiencies. In our policy implications, we therefore argue that social influences can interact with the non-competitive design of the policy, so either subsidies in areas with little solar potential should be limited or a competitive allocation system such as auctions should be used. Moreover, the regulation should add provisions to prevent more installations in areas that already exhibit excess power supply.

The EEG regulation is ideal to test for the effect of social interaction in a policy context, because—unlike in markets—the individual decision whether to claim the renewable energy subsidies to install a generator does not depend on (the expectations of) the actions of other agents. Thus, viewing solar panels supported by the EEG as a financial investment with a payment stream that depends on local characteristics (such as solar radiation), the decision to install solar panels should only be determined by the individual maximization problem of the producers. But we show that this simple rationale of isolated investment decisions does not account for the actual investment patterns.

We compiled a rich dataset that contains all solar panels which received a subsidy from 2000 until 2012 in Germany (over one million). The data contains information on the capacity of each generator, when it was installed, and its exact location. We also added data on solar radiation for each square kilometer in Germany to determine how profitable a solar panel at each location is. Finally, we added several political and socio-economic variables of the municipalities in which the panels are located.

To identify and estimate the effect of social interaction, we employ a propensity matching procedure on the municipality level. When the EEG legislation started, every municipality was untreated—nobody could be influenced by someone else who has a panel supported by the EEG. The first panel installation under the EEG in a municipality changed this status; starting with the first panel, others may be influenced by the presence of panels. For example, neighbors may talk to each other and convince the non-adopters to get a panel as well, since it is simple and profitable. Or households in the same municipality may seek to copy the behavior of their friends or betters for status reasons. We use the term social influence or social effect, because the effect on others is more general than mere peer effects, which are usually expected to work through personal interaction. Here, personal interaction may not be necessary to affect others; the high visibility of a solar panel may be enough to influence the neighborhood. While we cannot distinguish the different channels of social influence in our aggregate data analysis, we can determine whether they exist and quantify them.

Whenever a municipality is treated, i.e., the first panel is installed, we match it with a municipality where no panel is installed yet which is comparable in size, number of inhabitants, solar radiation, income, voting history etc. Comparing the probability whether at least one more panel is installed over the next year between treatment and control municipalities captures the effect due to social interaction, i.e., the effect due a panel already being present. In a robustness test, where we exclude all municipalities that installed solar panels *before* the introduction of the EEG (past outcomes), we find very similar results. Hence, our matching procedure appears to control for all relevant variables and our findings are not driven by selection on unobservables.

1.1 Related Literature

Social influences have long been known to affect individual behavior (e.g., Becker and Murphy, 2000). In the presence of social influence, the effectiveness of policy interventions may change. Interestingly, the EEG is not so much a policy *given* existing social effects, but social effects are a consequence of the policy itself. Thus, it is crucial to recognize the effects of social interaction when designing the policy. In the context of energy, where social influence may lead to a spatially inefficient allocation, this side-effect of policy is particularly important. Note also that we are interested in the aggregate effect and not in individual elasticities. Thus, the identification problems from the "social multipliers" literature (Glaeser et al., 2003) are not an issue here.

Peer effects. Our paper is related to a relatively new literature on the existence of peer effects. Among many others, personal interaction with friends and peers may affect decisions on financial investments (Bursztyn et al., 2014), retirement plans (Duflo and Saez, 2003), may affect behavior such as crime (Glaeser et al., 1996) or smoking (Gaviria and Raphael, 2001), and outcomes such as obesity (Trogdon et al., 2008), grades (Sacerdote, 2001), work performance (Mas and Moretti, 2009) or well-being (Katz et al., 2001). Regarding the exact mechanism underlying peer effects, there is a quickly growing theoretical and empirical literature with an upcoming handbook devoted solely to the economics of networks (Bramoullé et al., 2016).

The paper closest to ours in terms of research question is Dahl et al. (2014), who investigate the presence and strength of peer effects in the decision to take paternity leave, i.e., in social policy, using Norwegian data. They use a regression discontinuity design to exploit the introduction of a reform which allowed one extra month of government paid leave to fathers. The main finding is that fathers who take leave affect immediate peers in the workplace and family (brothers) positively. Their focus differs from ours in that they investigate personal relationships, whereas we investigate social influence more broadly in the community, possibly without personal contact. Moreover, their data, methodology and policy differ substantially from ours. In particular, the parental leave policy is motivated by equality concerns and does not have consequences for allocative efficiency as in the case of renewable energy subsidies.

Innovation adoption and solar panels. Our paper is also related to the interdisciplinary literature on innovation adoption. Evidence suggests that clustering of innovation adopters occur regularly—often due to social learning (Young, 2009). This link has also been made for adoption of photovoltaic cells (Bollinger and Gillingham, 2012; Richter, 2014).[1] Several papers use spatial methods to investigate patterns and shows that photovoltaic cells tend to be clustered (Graziano and Gillingham, 2015; Rode and Weber, 2016).

Rode and Weber (2016) investigate peer effects also with EEG solar panel data. Their methodology differs considerably from ours. They estimate a spatial diffusion model from epidemiology, which models adoption of a new technology (solar panels) as a function of the adoption rate in the surrounding area. Rode and Weber (2016) conclude that adoption choices are affected by others, and that this influence quickly diminishes with distance. A strength of their approach is the novel estimation in the spatial dimension. A downside is that they have to allocate buildings randomly within a NUTS-3 region when the building location data is incomplete. Moreover, except for time and region fixed effects, they use no controls, so it is unclear whether their results may reflect common characteristics or actual peer effects.

Methodologically closest to our study is Sianesi (2004), who studies a Swedish labor market program. She matches the unemployed who have undergone training with the unemployed who have not in order to estimate the effect of job training. As in our study, treatment status is time dependent and an unemployed citizen at time t may be part of the control group while being treated at some later time $t' > t$. The crucial step for the identification strategy to be valid in both of these papers is that the matching procedure controls for all factors that determine the outcomes absent (before) treatment.

2 Theory and background

2.1 Description of the German Renewable Energy Act

The German Renewable Energy Act (EEG) was adopted in 2000 with the goal of increasing the share of renewable energy production based on water, wind, solar radiation, and biomass. The regulation had two main dimensions. First, energy grid operators were required to provide access to producers of renewable energy and to feed in the produced electricity into the grid before any conventional electricity. Second, producers of renewable energy were guaranteed a fixed payment per produced kWh (often called feed-in tariff), which was set above the market compensation. These minimum compensation rates were viewed as necessary to increase the share of renewable energy at the time of the enactment of the law, when most renewable energy sources could not operate economically. Since its introduction in 2000, the regulation has been overhauled several times, the last time in 2014.

In this paper, we focus on the regulation concerning solar energy production.[2] A sched-

[1] See also the extended literature review on diffusion of PV technology in Richter (2014).

[2] Solar energy production encompasses different methods of energy production such as direct conversion

ule of minimum rates is guaranteed for a fixed time period, typically 20 years, and rates decrease by a fixed amount every year. The compensation per produced kWh varies with characteristics of the solar panel. Two main factors determine the compensation rate: the size of the panels (measured by capacity in kW) and installations on buildings versus installations on ground structures. In general, panels on the ground as well as larger panels receive smaller compensation rates.

Partly due to large increases in solar energy production at the end of the 2000s (implying large financial obligations for the following 20 years), the minimum rates have been reduced several times by the regulator. Moreover, measures[3] have been taken to prevent sudden or excessive increases in solar panel installations (e.g., due to large price decreases for solar panels) beginning in 2009. First, in an attempt to track future solar panel price reductions, the compensation rate for new installations is set to be reduced by a fixed proportion each year. Second, the compensation rate for new panels decreases in the overall number of installations in the previous year. Thus, larger expansions reduce the compensation available for new installations even more. As these changes took effect in 2012 and our sample only includes observations up to the year 2012, our analysis is not affected for the bulk of observations made between 2000 and 2011.[4]

2.2 Incentives for the investment decision

The incentives that are provided by the Renewable Energy Act to invest in solar panels are central to our analysis. What investment patterns should we observe over regions if agents make their decisions in isolation and if social influences do not play a role?

Given the guaranteed compensation, a rational agent's investment decision depends solely on the costs to install a solar panel and the return stream over time, which depends on the capacity to produce electricity at his location, which in turn is mainly determined by the solar radiation and roof size. As the subsidy per kWh is fixed, demand and supply by other renewable energy plants or conventional energy plants do not affect the compensation and therefore do not play a role in the individual investment decision.

To fix ideas, assume that a single investor maximizes his expected wealth and decides

into electricity (photo-voltaic) or heating of materials to propel turbines. In this article, solar energy production refers to typically small photo-voltaic solar panels mounted on roofs of buildings or panels mounted on ground structures.

[3]Cf. `https://www.clearingstelle-eeg.de/eeg2009` for the 2009, 2010, and 2011 amendments of the EEG; in German; accessed 06/08/15.

[4]There have been further overhauls of the regulation not affecting our data. One important change regarding the incentive structure was introduced in the amendment of 2014. Instead of a fixed compensation rate, the overall compensation for a plant is supposed to be determined by competitive procurement procedures rather than posted feed-in tariffs. So far this is limited to solar panels mounted on ground structures, but shall be extended in the future. See `http://www.erneuerbare-energien.de/EE/Navigation/DE/Gesetze/Das_EEG/das_eeg.html` for an overview of the German Renewable Energy Act; legal texts; and a comparison of different amendments that were introduced over time. Most texts are available only in German. Accessed 06/06/15.

whether to install a solar panel of a fixed size or not. Then his simple investment rationale can be represented in the following way:

$$\max_{x\in\{0,1\}} x\left[\sum_{t=1}^{T=20}\delta^t q_t^k m_t^k - C_i^k\right]. \tag{1}$$

The variables have the following interpretation:

- $x \in \{0,1\}$: Choice variable. Install the panel or not.
- C_i^k: Cost of installation of a panel of size k for household i.
- T: Time in years for which the EEG guarantees payments.
- $1 \geq \delta > 0$: Discount factor.
- q_t^k: Expected energy production during year t in kWh for a panel of size k.[5]
- $m_t^k > 0$: EEG remuneration for a panel of size k in year t per kWh.

A household installs the solar panel if discounted expected returns $\sum_{t=1}^{T=20}\delta^t q_t^k m_t^k$ outweigh the costs C_i^k. Now consider a specific municipality j. In this municipality, we have a joint distribution of δ, q_t^k, C_i^k on support Δ, Q, C, denoted by $f_j(\delta, q_t^k, C_i^k)$. If all households in the municipality decide in isolation whether to install the panel based on (1), then the probability of at least one solar panel installation in the municipality is

$$P = 1 - \left(\int_\Delta\int_Q\int_C \mathbf{1}\left\{\sum_{t=1}^{T=20}\delta^t q_t^k m_t^k < C_i^k\right\} f_j(\delta, q_t^k, C_i^k)\mathrm{d}\delta\mathrm{d}q_t^k\mathrm{d}C_i^k\right)^{N_j}, \tag{2}$$

where N_j is the number of households. Thus, we can empirically estimate a (reduced form) version of (2) as a measure of how attractive it is to install solar panels in a municipality given observables x. Clearly, bigger municipalities or those with higher solar radiation (large q_t^k) should have a higher probability of installations. As we argue in the next section, the empirical version of (2) is in fact the propensity score of a municipality, which is a measure of how comparable municipalities are in terms of their propensity to install panels. This measure allows us to compare pairs of municipalities j and l with a predicted probability of installations $P(x_j) = P(x_l)$, where social influences are active in j but inactive in l. Thus, if we observe systematic differences between these municipalities, then the difference must be due to social influences, because the fundamentals (propensity scores) are the same in

[5]Note that there exist various online tools for people to calculate the expected energy production (as well as costs) before installing a panel. See, for instance, `https://www.solarworld.de/service/solarstromrechner/`, `http://www.renewable-energy-concepts.com/german/sonnenenergie/basiswissen-solarenergie/pv-solar-rechner.html`, accessed 2/4/2016. Moreover, the local company installing the panels will provide projections of yield in their offer.

the pairs by construction. The next section shows how this comparison is done using a matching approach.

3 Method and data

3.1 Matching approach

Denote the outcome of municipality j absent treatment by Y_{jt}^0, the outcome of municipality j with treatment by Y_{jt}^1, and the treatment status by $d_{jt} \in \{0, 1\}$. Outcomes may be the probability of at least one installation or the number of installations over one year starting at t.

Treatment in our case means that municipality j already has a solar panel installed under the EEG legislation, hence others in the municipality may be influenced by it (i.e., social effects are active). Consequently, treatment status is time dependent in our setting. Municipality j may be untreated at t, $d_{jt} = 0$ (no panel installed; social effects are inactive), but it may be treated at some $t' > t$, $d_{jt'} = 1$, because a panel was installed. The econometric problem is identical in Sianesi (2004), where job seekers may be untreated at t but treated at $t' > 0$. Consequently, as in Sianesi (2004), untreated units may be used at first as control and later as a treatment observations.

We want to estimate the average treatment effect of the treated (ATT),

$$\Delta_{\text{ATT}} := \mathbb{E}[Y_{jt}^1 - Y_{jt}^0 | d_{jt} = 1].$$

As is well known, a direct estimate of the ATT is not possible, since the outcome for a municipality at the same time is not observed both when treated and untreated. If social effects are active in municipality j at t, then we observe Y_{jt}^1, but we do not observe what the outcome would have been if it had not been treated (Y_{jt}^0). The matching approach relies on finding a comparable municipality l that is untreated at t, so that its observed outcome Y_{lt}^0 equals the counterfactual Y_{jt}^0 on average.

For the identification of the ATT with the matching procedure, the standard conditional independence assumption (CIA) is sufficient. Formally, the CIA requires that, conditional on a vector of observable characteristics x, the potential outcome without treatment is independent from the treatment status,

$$Y_{jt}^0 \perp d_{jt} \mid x_{jt}.$$

Thus, the CIA requires that treated and untreated municipalities sharing the same characteristics $x_{jt} = x_{lt} = x$ draw Y^0 from the same distribution. This is fulfilled for the characteristics x that determine whether a municipality installs panels without social effects, i.e., factors that form a kind of municipality 'solar panel production function' *absent* *treatment*. Among those variables are the fundamentals that determine how attractive installing a solar panel in municipality j is (e.g., solar radiation), but also size variables and income that controls for the possibility of budget constraints in municipalities (a list of matching variables is discussed below).

Informally, our identification strategy exploits the fact that municipality outcomes Y^0_{jt} given municipality characteristics x_{jt} are not deterministic, so that differences in outcomes and hence treatment status (social influences active/inactive) are random given x_{jt}. That is, given the observables x_{jt} and x_{lt}, municipalities j and l are comparable in that they predict a similar distribution of outcomes over the next year, but the actual (realized) outcomes may differ. By matching two with different treatment status realization, we observe both the outcome of the treated municipality and the outcome of the comparable untreated municipality.

As is standard in the literature, we solve the curse of dimensionality problem by matching on the propensity score. If the CIA holds based on x_{jt}, then it also holds with propensity score matching based on x_{jt} (e.g., Rosenbaum and Rubin, 1983). Thus, we estimate the probability that at least one solar panel is installed in municipality j at t over the next year given a vector of characteristics x_{jt}. Then we match all municipalities where the first panel has been installed with another comparable municipality at the same time where no panel has been installed yet. The difference in the probability of further installations over the next year is the effect of social influence, i.e., the causal effect of having one panel in the municipality.

One critical point in the context of social interaction is the stable unit treatment value assumption, which requires that the treatment status of one municipality does not affect potential outcomes of other municipalities. Although we cannot rule out spillovers in single cases where a panel is installed close to the border to another municipality, we expect that social influences are active only locally, i.e., affect others only in proximity (see, e.g., Rode and Weber, 2016 for prior evidence). Since the municipality level (our unit of observation) is much larger/coarser than this, we expect social effects, if they exist, to affect others mostly in the same municipality. Note also that, if control municipalities are somehow affected by panels in a neighboring municipality, then (given positive social effects) our treatment effect estimates will be biased downwards, since we overestimate the counterfactual.

3.2 Matching variables

As the previous discussion shows, a valid matching procedure requires variables that explain the municipality production function of solar panels for a given year. In this section, we discuss the crucial variables that explain the number of new solar panel installations on the municipality level. Among the most important variables, we consider the solar radiation, number of inhabitants, and mean income.

Solar radiation. Holding capacity, angle towards the sun etc. of a solar panel constant, energy production and thus compensation under the EEG is increasing in solar radiation. Thus, it is more attractive to install a solar panel in locations with higher solar radiation. We obtained spatial data of yearly solar radiation means over the years 1981-2010 for all 1 km×1 km cells in Germany from the German weather service, which are measured in kWh/m^2.[6] This data is displayed on a map in Figure 1. It is the best publicly available estimate of the solar panel potential in Germany. Many consumer guides or websites on solar panels provide similar solar radiation data to aid the decision whether to install a panel, although these website data may be less precise (e.g., location determined via postal code is typically less accurate than our 1 km×1 km grid; see footnote 5 for two examples).

Inhabitants. All else equal, a larger municipality in terms of inhabitants should install more solar panels.

Number of (residential) buildings. Most of the solar panels in our data set are roof-mounted solar panels, and more buildings (i.e., roofs) may lead to more solar panels, all else equal.

Mean income. Solar panels under the EEG are essentially a financial investment for which additionally requires a roof (or ground structure) and whose return on investment depends on the location (see solar radiation above). Especially in the early years of the EEG, solar panel prices have been comparatively high, and the installation of a solar panel required a large upfront investment.[7] Consequently, all else equal, inhabitants of higher income municipalities are more likely to be able to afford the investment and to install more solar panels.

Political and ideological factors. We expect environmentally friendly citizens to be more likely to install solar panels in order to support the government efforts to expand the share of renewable energy in the energy production. On the municipality level, we have federal election results every four years, and we use the German green party[8] vote share as our measure of environmental awareness among municipality inhabitants. All else equal, a higher green party vote share should lead to more solar panels. The vote shares of other parties (Conservative, Social-democrats, Left etc.) proxy for other ideological and political factors.

We further control for the area of the municipality as a measure of size, for population

[6]These data also take into account local weather conditions, since a large number of rain days will reduce solar radiation, all else equal. kWh (kilowatt hour) is a unit of energy and kWh/m^2 is therefore the energy potential for a normalized area.

[7]Loans for the purpose of buying solar panels at favorable conditions are available from the government owned German Credit Bank (KfW) since 1999 (Schwarz, 2014, footnote 39). Still, the bank requires sufficiently good credit scores to approve the loan application, so it does not completely negate wealth constraints.

[8]Unlike in other countries, the German Green party has a strong presence in the political landscape, and has been part of governing coalitions on the federal level in two legislative periods (1998-2005) since its foundation in 1980. Moreover, the EEG subsidy system was enacted with the Greens as part of the governing coalition and driving force.

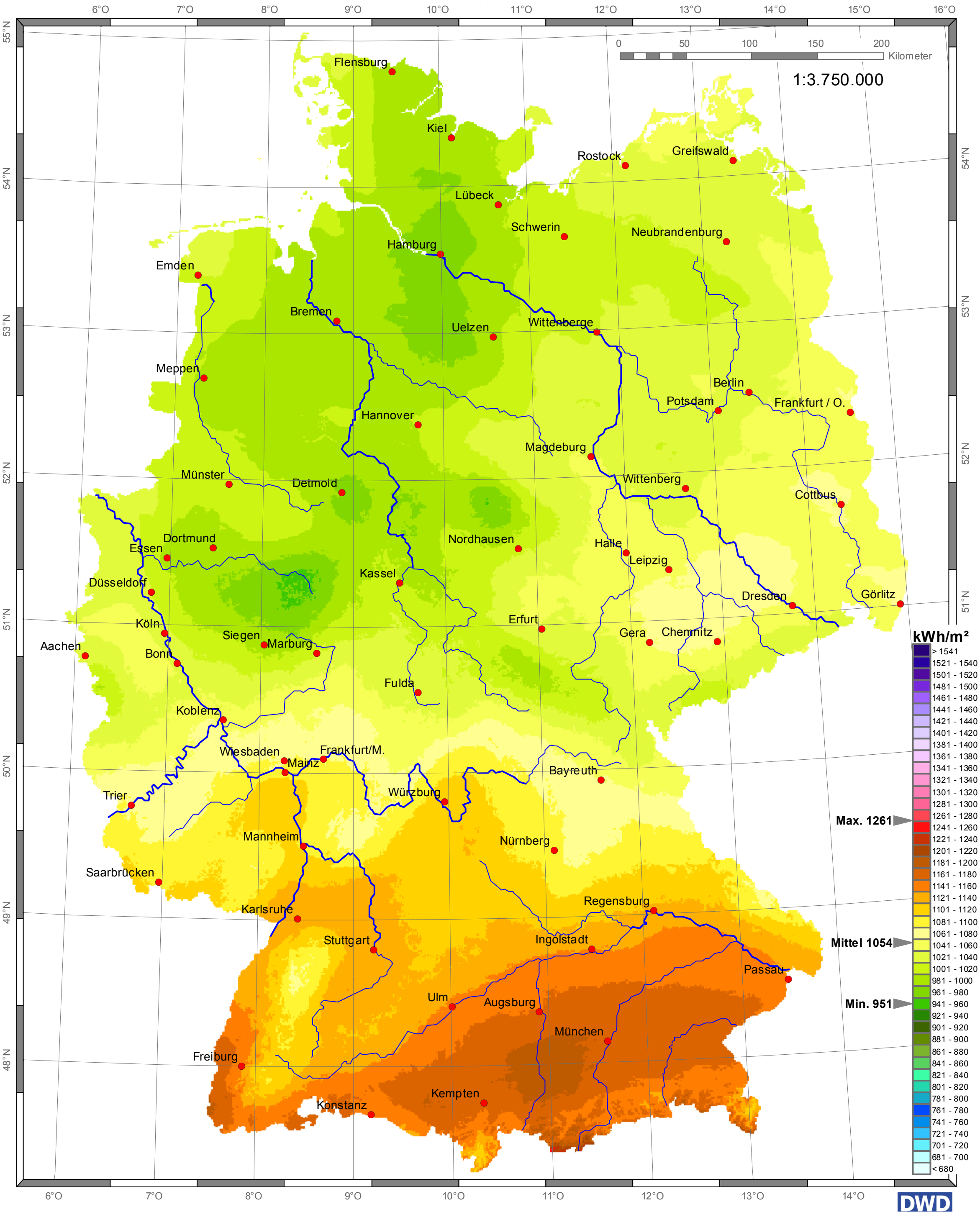


Figure 1: Map of solar radiation / solar energy potential in Germany; yearly means 1981-2010 in kWh/m^2; Source: German weather service (DWD)

density as it might affect social interaction, for the unemployment rate as further control for the financial situation besides income, for the share of seniors as a measure of age composition, and for the federal election vote turnout. Moreover, we have various interactions between these variables.

We also include state fixed effects in the propensity score estimation to control for state-

Table 1: Municipality Level Summary Statistics

	Mean	SD	Min	Max
Total area (in ha)	2,703.61	3,127.61	42.00	40,515.00
Residential buildings	1,398.33	3,963.34	8.00	126,349.00
Inhabitants	6,453.20	27,174.61	19.00	1,234,692.00
Mean income (in 1000 €)	30.31	7.56	8.00	272.57
Unemployment rate	0.04	0.03	0.00	0.28
Global radiation (in kWh/m^2)	1,059.39	54.81	958.00	1,192.00
Population density (number/ha)	1.88	2.67	0.03	39.77
Share seniors	0.17	0.04	0.00	0.41
Votes Green	0.06	0.03	0.00	0.38
Votes Conservative	0.44	0.14	0.08	0.96
Votes Soc-Dem	0.36	0.11	0.02	0.73
Votes Liberals	0.07	0.03	0.00	0.52
Votes Left	0.04	0.06	0.00	0.37
Votes others	0.03	0.02	0.00	0.17
Voter turnout	0.75	0.08	0.37	0.98
Observations	10905			

Note: Summary statistic refers to the universum of municipalities in 2000.

wide factors. Moreover, since the time intervals for which we calculate the outcomes are the same for each matched control and treatment municipality, we effectively include time fixed effects as well. This is important since the EEG regulation was changed several times, and by comparing outcomes in the same time intervals we make sure that these changes impact treatment and control municipalities equally.

Table 1 shows summary statistics of the characteristics of all municipalities we have in our sample in year 2000 (the first year of our sample). Section B.1 in the appendix lists all data sources with explanations.

3.3 Implementation of the matching procedure

In this section, we explain how we implemented the main matching procedure; robustness checks and alternative specifications are discussed in section 4.2.

In our main matching procedure, the propensity score is estimated separately each year with a logit model. Hence, we estimate the probability of at least one panel installation within year t in municipality j given observables x_{jt} with observations from the pool of all untreated municipalities at the beginning of years $t = 2000, 2001, \ldots, 2010$. The advantage of estimating the propensity score by year rather than over all years is that we not only allow for the constant in the probability of treatment to vary by year (like a year fixed effect), but we also allow all slope parameters of the covariates to vary by year, thus greatly improving the flexibility of our specification.

In the propensity score logit specification for each year, we use all variables discussed above as second degree polynomials, and add interactions of the number of inhabitants with mean income, the population density, the voter turnout, and the share of seniors. Different specifications using the same set of variables yield the same conclusions. In addition to these variables and interactions, we use state fixed effects to capture possible state-wide factors that might influence aggregate solar panel installations.

We use a nearest neighbor matching procedure with replacement on the estimated propensity scores, with one control municipality matched to each treatment municipality, and each control municipality is allowed to be matched to several different treatment municipalities. Given our large sample size and given that we match with replacement, we use an ambitious caliper of 0.01, i.e., if there is no control municipality with a propensity score difference of one percentage point or less, then we drop the treatment observation. This ensures almost exact matches in terms of the propensity score. Moreover, we enforce a common support among treatment and control municipalities. The common support ensures that the pools of treatment and control municipalities are similar. The combination of the 1-nearest-neighbor match, replacement, the ambitious caliper, and the common support are all aimed at minimizing bias rather than variance. Since we have a large number of observations despite the strict caliper, variance reduction is not a priority for us.

3.4 Matching quality

Appendix A assesses the quality of the match produced by the main model in detail. In short, while the set of treatment and the set of control municipalities differs considerably *prior to matching*, there are no significant differences left in terms of the propensity score distribution or the distribution of single covariates after matching. Figure 2 displays the distribution of the covariates for both treatment and control municipalities after matching. The two graphs are very similar for all variables. Tests for differences in means yield no significiant differences (see Appendix A). Thus, the quality of the match is very good, as we would expect given the ambitious caliper with replacement. Based on observable characteristics, our matching procedure was able to produce a balanced sample.

3.5 Propensity scores: What drives solar panel installations on municipality level?

Which factors determine whether a municipality installs at least one solar panel per year? Table 2 shows the marginal effects of the logit propensity score estimations for all years.[9] As expected, most size variables—area of the municipality, number of residential buildings, population density—have a positive effect on the probability of installing at least one solar

[9] Table 11 in the appendix shows the estimated coefficients with all interactions.

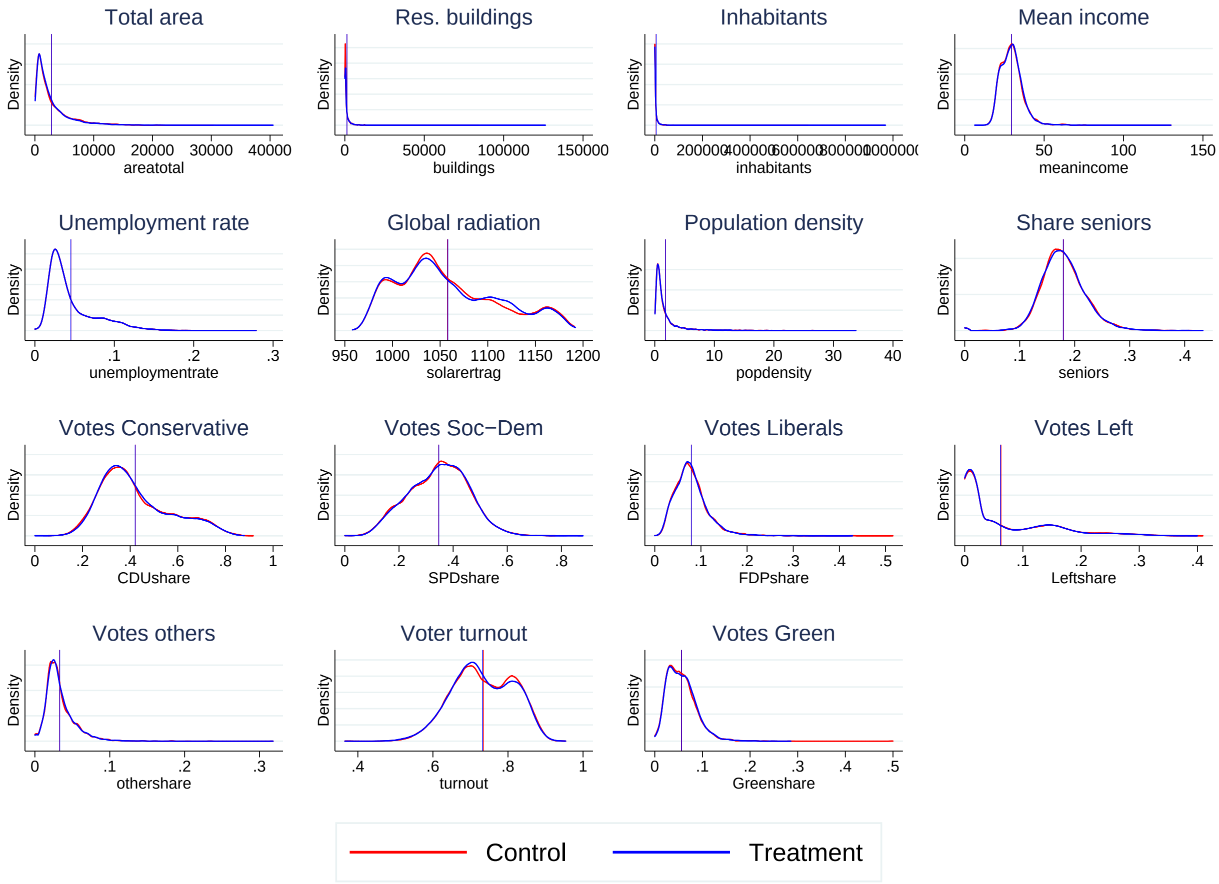


Figure 2: Distribution of covariates after matching for treatment and control group

panel per year, only the number of inhabitants is sometimes negatively related holding all other factors constant. The potential for solar energy production—solar radiation—is positively related, as expected. Regarding the variables capturing income and wealth, mean income has a positive effect, suggesting that budget constraints may inhibit expansion of the relatively expensive solar panels. The unemployment rate typically has a negative, albeit not significant effect. The vote shares of the non-green parties usually have negative point estimates, which is also as expected (the base category is the green party vote share, which is not displayed). The share of seniors in the municipality typically does not have a significant impact.

4 Results

4.1 The main specification

4.1.1 The average treatment effect on the treated

What role did social interactions play in the EEG-subsidized expansion of solar energy in Germany? We compute the ATT for three outcome variables: the probability of installing

Table 2: Estimation of Propensity Score by Year (Marginal Effects)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Total area	0.000***	0.000***	0.000***	0.000*	0.000***	0.000***	0.000**	0.000	-0.000	0.000**	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Res. buildings	0.000***	0.000***	0.000***	0.000**	0.000	0.001***	0.000	-0.000	0.001*	0.000	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inhabitants	-0.000***	-0.000***	-0.000	-0.000	0.000*	-0.000	0.000	0.000***	0.000	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Mean income	0.002**	0.002	0.001	0.001	-0.001	0.000	0.002	0.005*	0.001	-0.003	0.009
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Population density	0.016***	0.023***	0.017***	0.008	-0.021	0.007	0.074***	0.001	-0.008	0.157	-0.005
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)	(0.06)	(0.10)	(0.15)
Share seniors	-0.133	0.039	-0.158	-0.154	-0.459***	-0.352**	-0.287	0.076	-0.090	-0.464	1.140*
	(0.12)	(0.12)	(0.12)	(0.11)	(0.16)	(0.17)	(0.19)	(0.26)	(0.31)	(0.40)	(0.61)
Voter turnout	-0.163*	-0.028	-0.114	-0.038	-0.234**	0.008	0.029	-0.267	0.332	-0.117	-0.034
	(0.10)	(0.09)	(0.09)	(0.08)	(0.11)	(0.13)	(0.16)	(0.18)	(0.23)	(0.30)	(0.41)
Unemployment rate	-0.492	-0.026	-0.470	0.068	-1.135**	-0.321	-0.186	0.599	-1.215	-1.467*	0.428
	(0.46)	(0.40)	(0.35)	(0.27)	(0.44)	(0.45)	(0.56)	(0.73)	(0.84)	(0.88)	(1.48)
Global radiation	0.001***	0.000	0.000	0.001***	0.002***	0.002***	0.003***	0.002***	0.002*	0.001	0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Votes Conservative	-1.318***	-0.750***	-0.295*	-0.157	-0.157	-0.124	0.237	0.145	0.371	-0.159	-2.251***
	(0.15)	(0.17)	(0.15)	(0.15)	(0.22)	(0.26)	(0.32)	(0.42)	(0.47)	(0.58)	(0.87)
Votes Soc-Dem	-1.212***	-0.794***	-0.252	-0.193	-0.284	-0.133	0.172	0.255	0.426	-0.114	-2.478***
	(0.17)	(0.18)	(0.16)	(0.15)	(0.22)	(0.27)	(0.32)	(0.42)	(0.48)	(0.57)	(0.89)
Votes Liberals	-1.336***	-0.674**	-0.142	0.096	-0.027	0.472	0.075	0.935*	0.765	1.892**	-0.943
	(0.30)	(0.28)	(0.25)	(0.23)	(0.30)	(0.37)	(0.43)	(0.53)	(0.59)	(0.79)	(1.22)
Votes Left	-1.480**	-0.404	-0.120	-0.478	-0.628**	-0.983***	0.172	0.536	0.473	-0.083	-1.667
	(0.65)	(0.48)	(0.41)	(0.32)	(0.31)	(0.37)	(0.41)	(0.51)	(0.57)	(0.72)	(1.08)
Votes others	-1.350***	-1.030***	0.119	0.006	-0.721*	0.143	0.310	0.631	1.497*	-0.124	-0.714
	(0.37)	(0.37)	(0.33)	(0.28)	(0.39)	(0.47)	(0.55)	(0.67)	(0.81)	(0.96)	(1.41)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10873	9113	6396	5643	5178	3968	2881	2078	1498	943	430
Pseudo R^2	0.350	0.349	0.191	0.172	0.119	0.099	0.089	0.069	0.075	0.066	0.120

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parentheses.

at least one solar panel in the municipality, the number of solar panels installed, and the number of solar panels per 1000 inhabitants within one year.[10] Recall that we compare the probability/number of *further* solar panels in the treatment group (since treatment implies that one panel was installed) within a year with the probability/number of solar panel installations in the matched control group in the same time interval. The average treatment effects on the treated are displayed in Table 3. On all outcome variables, the estimated treatment effect (effect of social interaction) is positive and statistically significant.

The probability of at least one panel installation approximately increases by about a quarter due to social influence, from about 49% in the matched control group to about 60% in the treatment group. Next, the number of installations increase by about a half, increasing the number from about 1.6 to about 2.4 solar panels within a year on average. Finally, the number of installations per 1000 inhabitants increases by more than half from 0.85 to about 1.44 solar panels within a year.

To foreshadow the results of the robustness section, we did not find a single specification without positive and significant treatment effects. Thus, social effects had an economically

[10]Note that the number of solar panels measure underestimates the effect if social influences are positive, since the control municipality is untreated only until the first panel is installed, and afterwards social influences are active. The probability outcome measure, on the other hand, does not underestimate the effect, since the indicator does not distinguish between one or more panels.

Table 3: Average treatment effects on the treated over one year

	(1) OLS	(2) OLS	(3) OLS
Dependent variable	ProbInstall	NumInstall	NumInstall per 1000
Treatment	0.118***	0.800***	0.587***
	(0.007)	(0.065)	(0.040)
Constant	0.487***	1.575***	0.845***
	(0.006)	(0.059)	(0.034)
R^2	0.01	0.01	0.01
Observations	22284	22284	22284
Clusters	10712	10712	10712

Note: ProbInstall is the probability of at least one new panel within a year, NumInstall is the number of new panels within a year, and NumInstall per 1000 is the number of new panels within a year per 1000 inhabitants. Standard errors are clustered at the municipality level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

large and statistically significant impact on the investment decision and thereby on the expansion of renewable energy. It therefore appears not to be the case that individuals make their investment decisions in isolation, but are affected by neighbors or spatially close households. In particular, social interaction can explain a large part of the high participation rate in the EEG-program.

Result 1. *Social effects increase the number and probability of further solar panel installations.*

4.1.2 Heterogeneous treatment effects

Are social effects stronger in municipalities with particular characteristics? For example, are social effects stronger in small close-knit municipalities? Or are they stronger in southern areas with more solar radiation, where it is more attractive to install panels? We investigate these questions by comparing how the outcomes change by characteristic in both treatment and control group. The regression results for all outcome variables are given in Table 4. The interactions of the treatment dummy with the municipality characteristics are the average changes of the treatment effect in the characteristic (holding other characteristics constant).

We will discuss only the most important results. First, more solar radiation increases the treatment effect on the number of installations, but not for the other outcome variables. This result suggests that solar radiation works on the intensive margin, but not on the extensive margin of the social effects. Second, the treatment effect decreases in the vote share for the green party. This is surprising, since the probability and number of solar panel installations increases significantly in the vote share for the green party. The result might suggest that municipalities with more inhabitants motivated by environmental concerns (rather than financial reasons) are not influenced as much by others. Third, the effect

Table 4: Average treatment effects on the treated over one year by municipality characteristics

	(1) OLS	(2) OLS	(3) OLS
Dependent variable	ProbInstall	NumInstall	NumInstall per 1000
Treatment	0.568***	-4.245***	3.143***
	(0.148)	(1.382)	(1.140)
Treatment× Solar radiation	-0.000	0.006***	-0.001
	(0.000)	(0.001)	(0.001)
Treatment× Voteshare green	-0.689***	-3.918*	-14.914***
	(0.266)	(2.219)	(4.978)
Treatment× Total area	-0.000	0.000*	-0.000***
	(0.000)	(0.000)	(0.000)
Treatment× Inhabitants	0.000*	0.000	0.000**
	(0.000)	(0.000)	(0.000)
Treatment× Buildings	-0.000*	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)
Treatment× Population density	0.000	-0.105	-0.051**
	(0.004)	(0.105)	(0.024)
Treatment× Unemployment	-1.142***	-14.554***	-16.976***
	(0.249)	(1.725)	(2.506)
Treatment× Mean income	-0.007***	-0.014	-0.008
	(0.001)	(0.011)	(0.023)
Treatment× Share seniors	-0.149	-1.230	3.599
	(0.172)	(0.802)	(2.491)
Solar radiation	0.002***	0.016***	0.001
	(0.000)	(0.001)	(0.001)
Voteshare green	0.322*	4.632***	7.934*
	(0.183)	(1.749)	(4.298)
Total area	0.000***	0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Inhabitants	-0.000***	0.000	0.000***
	(0.000)	(0.000)	(0.000)
Buildings	0.000***	0.000**	-0.000
	(0.000)	(0.000)	(0.000)
Population density	0.017***	0.038	-0.184***
	(0.003)	(0.068)	(0.017)
Unemployment	-1.437***	-1.592*	-5.456***
	(0.156)	(0.889)	(1.257)
Mean income	0.006***	0.003	-0.012
	(0.001)	(0.009)	(0.010)
Share seniors	0.372***	0.074	3.548**
	(0.126)	(0.760)	(1.466)
Constant	-1.782***	-16.341***	0.016
	(0.112)	(1.451)	(0.740)
R^2	0.17	0.29	0.05
Observations	22284	22284	22284
Clusters	10712	10712	10712

Note: ProbInstall is the probability of at least one new panel within a year. NumInstall is the number of installations within a year, and NumInstall per 1000 is the number of installation per 1000 inhabitants within a year. Standard errors are clustered at the municipality level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

of social interaction decreases in the unemployment rate for all outcome variables. Since solar panels are comparatively expensive investments, it is not surprising that households in municipalities with high unemployment are influenced as much by others, as they may not be able to afford the investment. Size variables such as area, inhabitants and number of residential buildings do not unambiguously change the magnitude of the social effects.

The number of installations increasing with solar radiation and decreasing with unemployment might suggest that social effects are stronger where it is more attractive to install them (more solar radiation, more likely to be able to afford the panel).

Result 2. *The social effect is stronger in municipalities with lower unemployment and a lower vote share for the green party. More solar radiation increases the social effect on the number of installations.*

From a policy perspective, it is interesting to ask whether social influences account for installations even in areas with minimal solar radiation, where it might be inefficient to install them. After all, these panels could be installed in municipalities with more solar radiation, and, all else equal, the power output of a panel increases linearly in solar radiation.

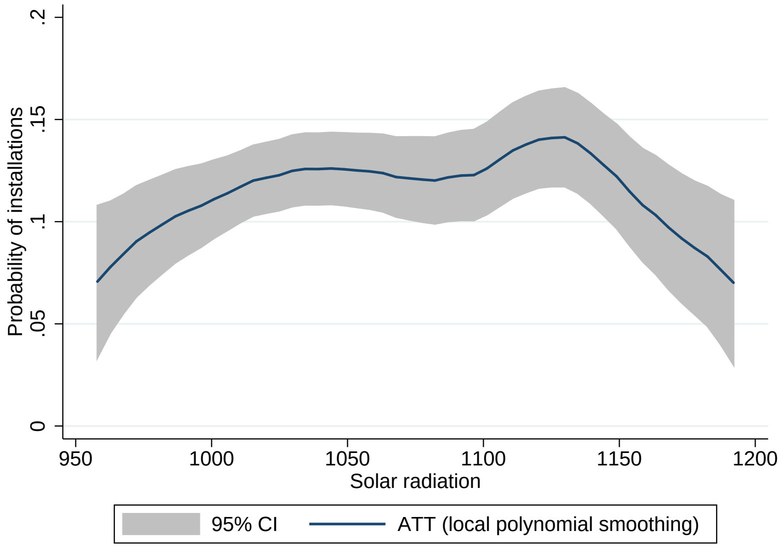
An advantage of the matching approach is that it allows computation of 'local treatment effects' rather than one average effect. Thus, we can compute the treatment effect for each treated municipality by taking the outcome difference of the treatment municipality and the matched control municipality. We can then locally average the results with nonparametric methods. Figures 3a and 3b display the local average treatment effects by solar radiation of the treatment municipality with nonparametric smoothing. First, for the probability of installations, Figure 3a shows that the treatment effect is significantly positive for the entire range of solar radiation levels in the sample. Second, for the number of installations, Figure 3b shows that the social effects are never significantly negative, and significantly positive for most of the solar radiation range (except close to the minimum and maximum sample values). Third, the effect of others on installations appears to be strongest in the upper range of solar radiation levels between 1100 and 1150 kWh/m^2 for both outcome variables.

Thus, social effects are responsible for installations in areas with low solar radiation where it might be inefficient. Consequently, the presence of social effects may make an inefficient policy even worse. We elaborate on this possible interaction of social effects with the non-competitive policy design in the policy implications below.

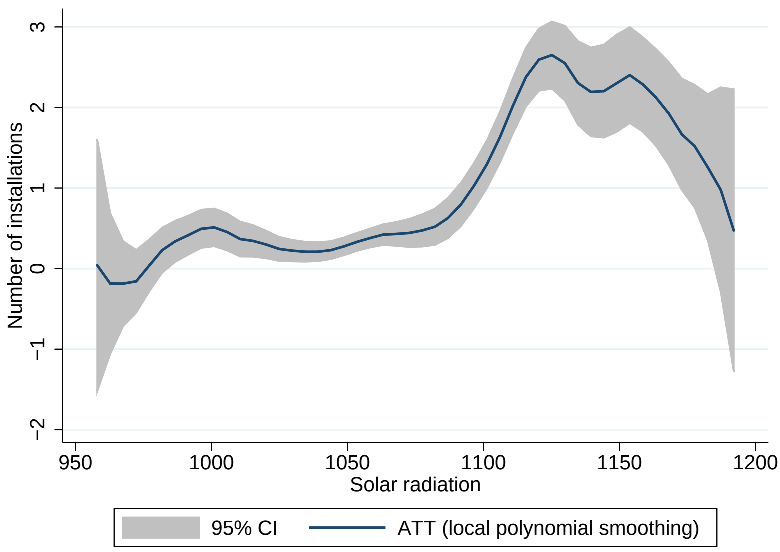
Result 3. *The social effect on the probability of installations is positive at all sample values of solar radiation, and positive on the number of installations in the interior range of solar radiation levels.*

4.1.3 The ATT over time

Does the effect of social interactions weaken over time? We might expect a new panel in the municipality to influence others only for a limited amount of time until the focus in the



(a) Probability to install at least one solar panel

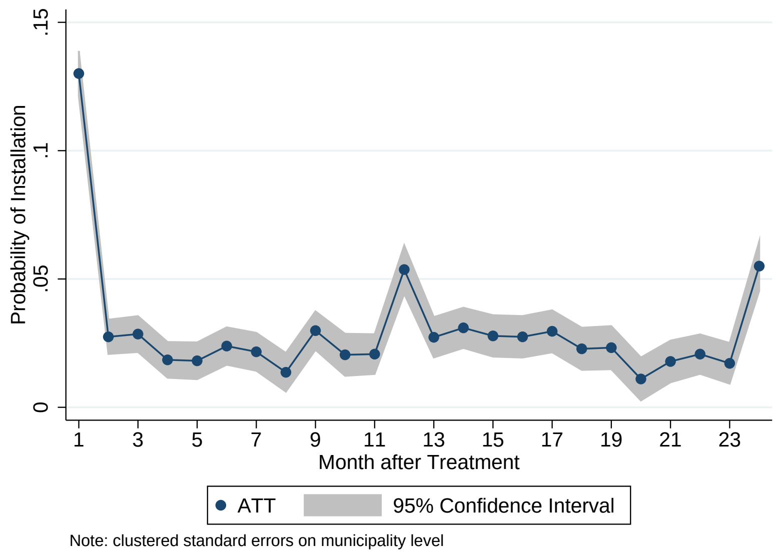


(b) Number of solar panel installations

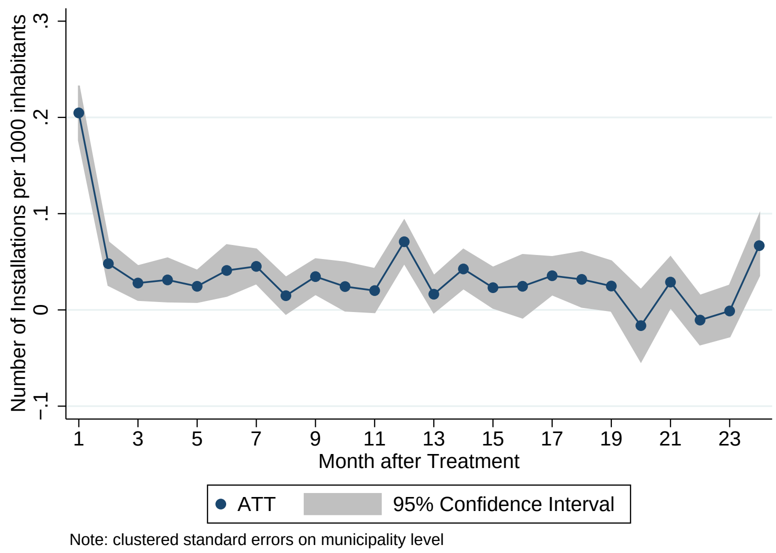
Figure 3: The ATT by solar radiation (in kWh/m^2)

neighborhood shifts to something else. On the other hand, it might take a while for social influences to affect outcomes in the neighborhood, because it takes time for others to plan and install panels.

The effects of social interactions are displayed in Figures 4a and 4b for each month after the first panel installation (treatment) on the probability of at least one more panel and on the number of more panels per 1000 inhabitants. Since both outcome measures are a function of the number of new panels, the two graphs look similar. In the first month after treatment, social influences are the strongest, and do not exhibit a noticeable time trend afterwards except for two upward spikes in month 12 and month 24 after treatment. The effects on the probability of installations are significantly positive for all months, and for



(a) Probability to install at least one solar panel



(b) Number of solar panel installations per 1000 inhabitants

Figure 4: The ATT for each of the 24 months after treatment

the number per 1000 inhabitants the effect is positive in most months. For the number, the total treatment effect is the sum of all monthly treatment effects, which is therefore weakly increasing over time. The big estimates for the first month show that social influences have almost immediate effect.

The reason for the spikes in months 12 and 24 is that a lot of panel installations are bunched on January 1 in the years 2001, 2002 and 2003.[11] Thus, many of the municipalities

[11]This bunching occurs because the EEG pays the compensation rates for 20 years plus the remainder of the year where the panel was installed. Thus, a panel installation on January 1 maximizes the time in which energy production is subsidized. We observe almost none of this bunching on January 1 in the later years, since the compensation rates where lowered at the beginning of each of these later years.

Table 5: Average treatment effects on the treated: Monthly time trends

	(1) OLS	(2) OLS	(3) OLS
Dependent variable	ProbInstall	NumInstall	NumInstall per 1000
Treatment	0.041***	0.074***	0.067***
	(0.002)	(0.006)	(0.005)
Treatment$\times$ t	-0.001***	-0.001***	-0.003***
	(0.000)	(0.000)	(0.000)
Constant	0.113***	0.298***	0.139***
	(0.003)	(0.012)	(0.008)
Month dummies	yes	yes	yes
R^2	0.01	0.00	0.00
Observations	534816	534816	534816
Clusters	10712	10712	10712

Note: The table displays estimates of a linear monthly time trend for the treatment effect. ProbInstall is the probability of at least one new panel within a year, NumInstall is the number of new panels within a year, and NumInstall per 1000 is the number of new panels within a year per 1000 inhabitants. Standard errors are clustered at the municipality level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

that were treated on January 1, 2001—which are about 6% of the treated municipalities—have a spike in installations exactly one year later, on January 1, 2002 and again in 2002.

To formally test for time trends, we estimate a linear time trend β on the matched sample for the treatment effect / social influences of the form

$$y_{it} = \alpha \text{Treatment}_{it} + \beta \text{Treatment}_{it} \times t + \gamma' \mathbf{Month}_{it} + \delta + \varepsilon_{it},$$

where y_{it} is the outcome of municipality i in month t after treatment, Treatment_{it} is the treatment dummy, $\mathbf{Month}_{it}$ is a column vector of dummies for month $2, 3, \ldots, 24$ (month 1 is the reference category), and δ is the constant. Table 5 shows the estimates for all outcome variables.[12] The slope coefficient β on Treatment$\times t$ is significantly negative for all outcome variables. Hence, the effect of social interaction decreases over time on average, and it remains positive for most months we consider (see Figures 4a and 4b). Note, however, that the time trend is driven mostly by the strong effect in the first month. Estimating a time trend when excluding the first month from the sample, we only retain a significant time trend for the last outcome variable, the number of installations per 1000 inhabitants (not displayed).

Result 4. *The social effect is strongest in the first month after treatment and decreases over time. The social effect is positive on the probability of installations for at least two years after the first panel installation.*

[12]The estimates are virtually identical if we exclude the municipalities treated on January 1, 2001.

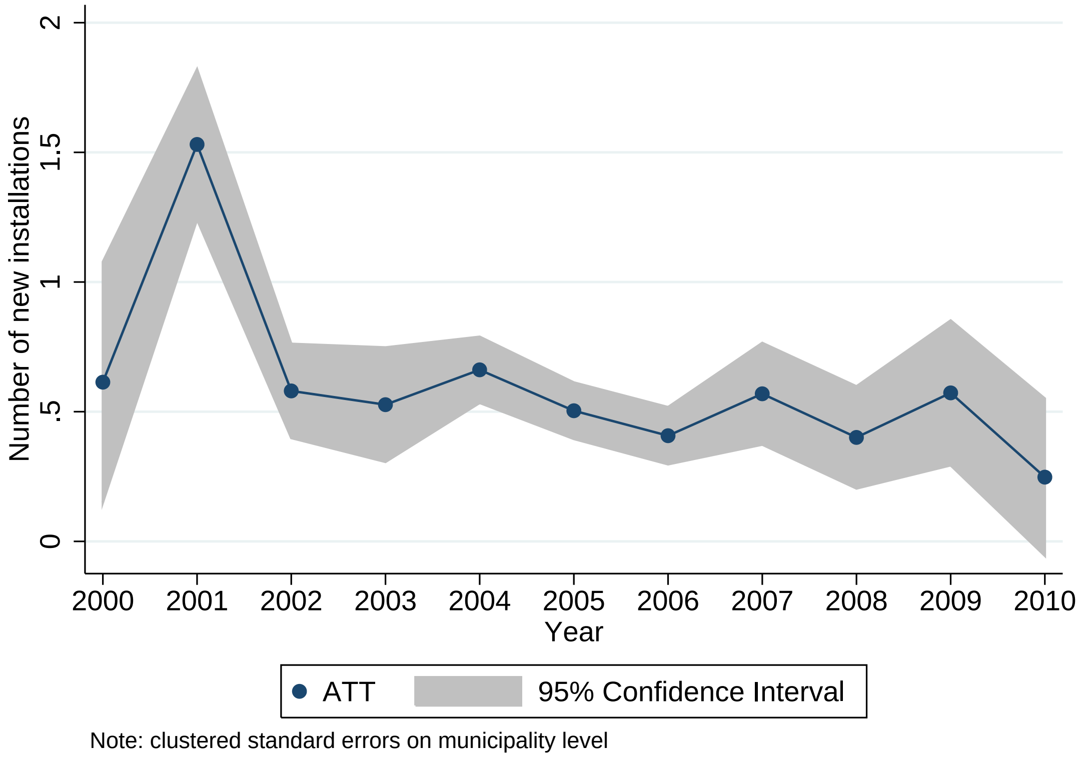


Figure 5: The ATT on the number of new installations by year

4.1.4 Treatment effects by year

Are social effects stronger in the early years of the regulation, or do they pick up only after a while? To address this question, we calculate the ATT separately for all municipalities that were treated in year 2000, year 2001, and so on. Figure 5 displays the ATT on the number of installations within one year, since this measure directly captures the expansion of solar panels. The ATT is strongest in the year following the introduction of the EEG. In the years thereafter, the effect appears to be very slightly decreasing. The strong year 2001 might suggest that social interaction helped spread the word about the profitable solar panel subsidy when it was not yet well known. Or it might be strongest in this early year simply because sunnier municipalities are overrepresented—since they have a larger propensity score—and the effects are stronger in these municipalities (see previous paragraphs on heterogeneous treatment effects).

To check this alternative possibility, we control for the propensity score of the municipalities when calculating the ATT by year, which accounts for the probability of treatment:

$$y_{it} = \alpha' \text{Treatment}_{it} \times \mathbf{Year}_{it} + \beta \text{Treatment}_{it} \times \text{PScore}_{it} + \gamma' \mathbf{Year}_{it} + \delta \text{PScore}_{it} + \varepsilon_{it}, \quad (3)$$

where $\mathbf{Year}_{it}$ is a column vector of dummies for years 2000, 2001, ..., 2010 (which absorbs the constant), and PScore_{it} is the propensity score of municipality i in year t. Thus, in this

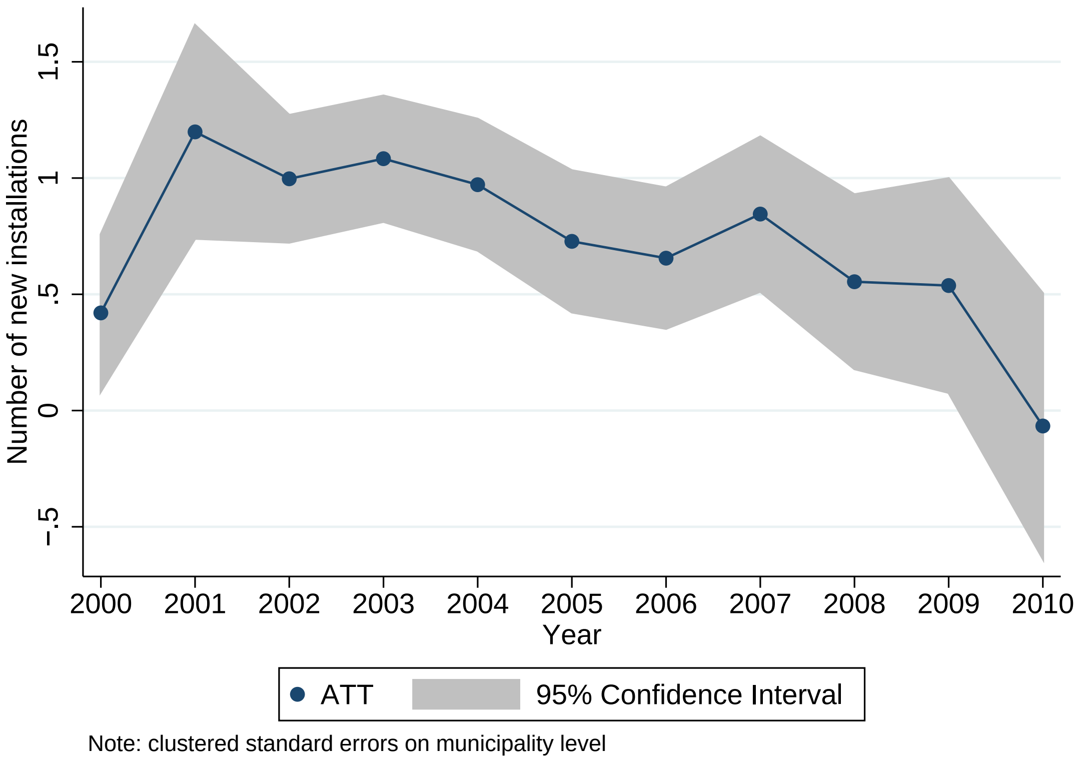


Figure 6: The ATT on the number of new installations by year, controlled for the propensity score of the municipality, evaluated at the mean propensity score 0.4468 (see (3) for the estimated equation)

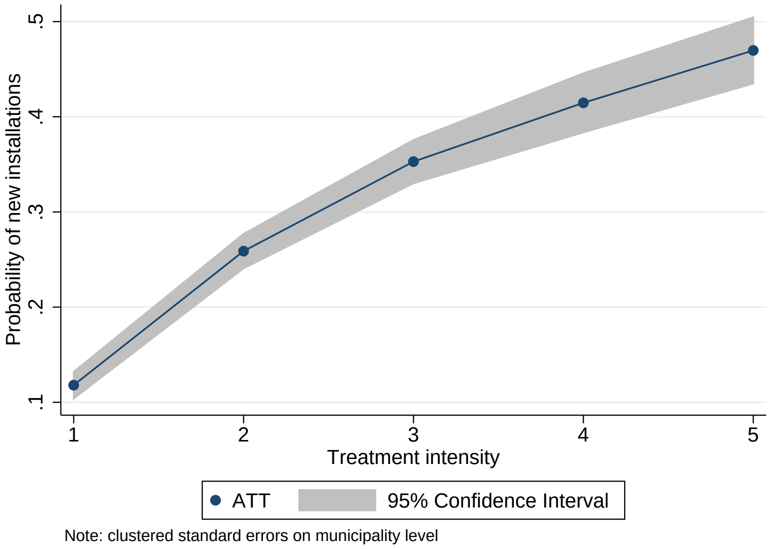
specification, the treatment effect may vary by year and by propensity score, which allows us to determine how the treatment effect varies by year (α) after controlling for propensity scores.

The resulting ATT estimates are displayed in Figure 6. The graph shows that the spike in 2001 is considerably smaller once we control for the propensity to install new panels. There appears to be a slightly decreasing time trend after controlling for the propensity score. However, a formal test for a linear time trend gives a significant negative trend only at the 10% significance level. Our conclusion is that social influence appears to be stronger in the early years, but a lot of this annual difference is due to "better" municipalities with larger propensity score—where social effects are stronger—being treated in the early years.

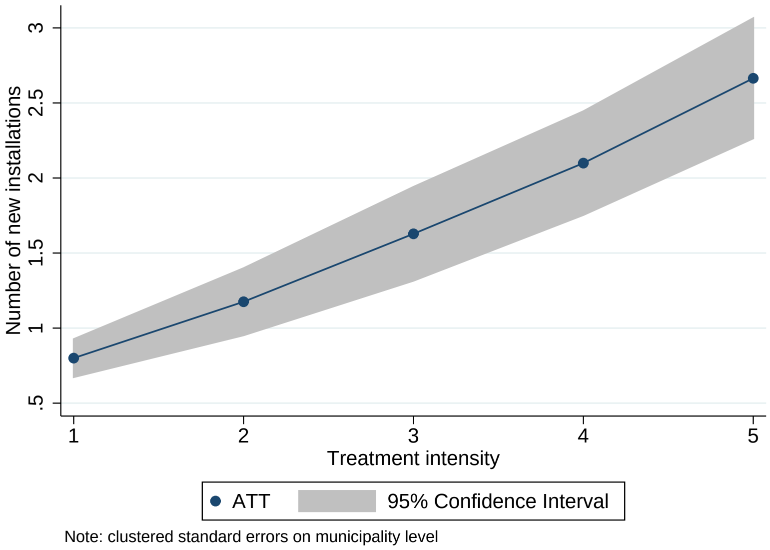
Result 5. *The social effect was strongest in 2001 shortly after the introduction of the policy, but this difference can be explained by sunnier and bigger municipalities getting treated earlier.*

4.1.5 Treatment intensity

So far we assumed that 'treatment' is the installation of one panel in a municipality, hence the treatment effect reflected the effect of one panel in the municipality versus no panel in



(a) Probability to install at least one solar panel



(b) Number of solar panel installations

Figure 7: The ATT by treatment intensity (number of existing panels in the municipality)

the municipality (control group). We now ask how the effects change if treatment intensity varies, i.e., if there are two, three, four or five instead of one panel in the municipality (versus none in the control group). Thus, we run our matching routine with the same specification as before, except we define treatment as starting once the second, third, etc. panel was installed.

Figures 7a and 7b display the average treatment effect on the treated by treatment intensity. As expected, the treatment effects increase in treatment intensity both for the probability of further installations within a year as well as for the number of further installations. For the probability of further installations, the treatment effect appears to be concave in treatment intensity. By modeling the treatment effect as second degree polynomial, we can carry out a simple formal test for this observation when pooling the matching data sets for all intensities:

$$y_{it} = \alpha \text{Treatment}_{it} + \beta \text{Treatment}_{it} \times \text{Intensity}_{it} + \gamma \text{Treatment}_{it} \times \text{Intensity}_{it}^2 \\ + \delta' \textbf{Intensity}_{it} + \varepsilon_{it},$$

where $\textbf{Intensity}_{it}$ is a dummy indicating treatment intensity of the matched pair (which absorbs the constant), and Intensity_{it} is the scalar value of the intensity. The test for concavity is $\gamma < 0$. We estimate $\gamma = -.0148$ ($t = -6.62$), so the treatment effect on the probability of further installations is concave in treatment intensity in this second order polynomial approximation.

For the number of further panel installations in Figure 7b, the treatment effect appears to be slightly convex in treatment intensity. The test for convexity is $\gamma > 0$, however, we obtain $\gamma = 0.028$ ($t = 1.64$), so we cannot reject linearity of the treatment effect in intensity on the number of further installations at the 5% significance level.

Interestingly, for the number of installations, the marginal increase is the largest for the first panel. This has important implications for policy makers that wish to encourage installations in certain areas. The first panel may be considered a "seed" panel, and these results may give a justification to introduce "encouragement" policies that are targeted with no or only a few installations.

Result 6. *The social effect is stronger with more existing panels in the municipality. While the social effect increases approximately linearly for the number of installations in intensity, it is concave for the probability of further installations.*

Accounting for history. If treatment starts not at the first installed panel in the municipality, but with installation of the K*th* panel ($K > 1$), then the history of panel installations will differ between treatment municipalities. For example, if treatment starts with the installation of the second panel, then the installation of the first panel may have occurred immediately before the second, or years before. The estimates in Figures 7a and 7b give the average effect over all histories in our sample. Now we want to investigate if the social effect is stronger if several panels have been recently installed, i.e., how the social effect changes by installation history.

The history variables of interest are the times between the panel installations and the treatment start. For treatment start at installation of K panels, the history is $\{d_k\}_{k=1,\dots,K}$, where d_k is the time (distance) between installation of the k*th* panel and time of treatment. Because treatment starts at installation of the K*th* panel, we have $d_K = 0$, and the relevant history that may differ between municipalities is $\{d_1, \dots, d_{K-1}\}$. To account for the history, we could use interactions with the treatment dummy:

$$y_{it} = \alpha \text{Treatment}_{it} + \beta' \text{Treatment}_{it} \times \mathbf{d}_{it} + \gamma + \varepsilon_{it}, \qquad (4)$$

Table 6: Average treatment effects on the treated by panel installation history

	(1) OLS $K = 2$	(2) OLS $K = 2$	(3) OLS $K = 3$	(4) OLS $K = 3$	(5) OLS $K = 4$	(6) OLS $K = 4$
Dependent variable	ProbInstall	NumInstall	ProbInstall	NumInstall	ProbInstall	NumInstall
Treatment$\times d_1$	-0.090***	-1.259***	-0.056***	-1.214***	-0.032***	-1.205***
	(0.008)	(0.094)	(0.008)	(0.107)	(0.009)	(0.108)
Treatment$\times d_2$			-0.034**	-0.769***	-0.021	-0.517***
			(0.015)	(0.193)	(0.016)	(0.194)
Treatment$\times d_3$					-0.033	-0.766***
					(0.022)	(0.265)
FE PS group	yes	yes	yes	yes	yes	yes
R^2	0.01	0.01	0.01	0.02	0.00	0.03
Observations	20886	20886	19456	19456	17832	17832

Note: ProbInstall is the probability of at least one new panel within a year and NumInstall is the number of new panels within a year. d_k is the time since installation of the kth panel at treatment start in 1000 days. Standard errors are clustered at the propensity score group level; group $g = 1, \ldots, 100$ has a propensity score in $(g-1, g]$. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

where $\mathbf{d}_{it} = (d_1, \ldots, d_{K-1})'$ is the vector of the times between installation of the $K-1$ old panels and treatment start. The effect of the history β is estimated solely from variation between treatment municipalities, since control municipalities all have the same history (no panel installations). However, treatment municipalities may not necessarily be comparable, since our matching approach aimed at making treatment and control municipalities comparable. Thus, large sunny treatment municipalities likely have more concentrated installations (small values of d_k) and also likely have more installations in the future, compared to smaller treatment municipalities with less sun, which might give a negative effect of history ($\beta < 0$) due to the between-treatment-municipality comparison.

To make treatment municipalities comparable, we use the propensity scores we already have from matching with control municipalities. Ideally, we would want to run regression (4) only for treatment municipalities in a narrow propensity score interval, say for municipalities with propensity score in $[0.9, 0.91]$, which ensures that the effect of history is estimated from variation within comparable treatment municipalities. Since it is impractical to run many different regressions on different subintervals, we use the within estimator, i.e., a single fixed effects regression with a fixed effect for each propensity score interval. This allows us to estimate the effect of history within propensity score groups, rather than between groups as in (4).

Table 6 reports the results of the within estimator for treatment intensity $K = \{2, 3, 4\}$. We use 100 propensity score groups with a 1 percentage point interval for the treatment municipalities and 100 groups for the control municipalities. Thus, the treatment dummy is constant within each group, and we can only estimate the parameters for the history that vary within groups.

If treatment starts at two panels in the municipality ($K = 2$), the only relevant history is how long ago the first panel was installed when treatment started (d_1). Table 6 shows that the effect of distance is significantly negative for all outcome variables. On average, if the first panel was installed 1000 days prior to the second panel ($d_1 = 1$), then the social effect on the probability of further installations after the second panel is 9 percentage points lower compared to the case of $d_1 = 0$. Moreover, the number of further panel installations in the municipality is reduced by 1.26 panels. For $K = 3$, the installation history of the first two panels may differ between treatment municipalities. Again, installations farther in the past imply a weaker social effect. For $K = 4$, the point estimates for the distance interactions are all negative. While these effects are all significantly negative on the number of installations, they are not all statistically significant for the probability of installations. Still, overall, the effect of distance on the social effect is negative, and temporally more concentrated panel installations affect others more.

Result 7. *The social effect is stronger with temporally more concentrated installations, i.e., if there are more installations in the recent past rather than the distant past.*

4.2 Alternative specifications and robustness checks

In this section we present the treatment effect estimates using different specifications, parameters, or sets of observations. The biggest concern with any matching approach is whether it really matches treatment municipalities with comparable control municipalities to obtain an estimate of the counterfactual.

Exclude municipalities with pre-EEG solar panels. Our first robustness check directly addresses the concern that the matching does not account for all relevant variables. In the EEG solar panel data, we observe solar panels that have been installed before the EEG went into force and thus before the EEG guaranteed minimum prices to solar panel owners.[13] In other words, we observe past outcomes, which allows us to see whether our matching approach is able to predict current outcomes correctly. Municipalities with pre-EEG panels are very likely to also install panels under the EEG, for example because they are bigger, solar radiation is stronger, or other factors we might not have considered are more favorable. Consequently, if there are unobserved differences we do not account for, then we should underestimate counterfactual outcomes for these municipalities in particular. This implies that if our matching procedure does not account for all relevant factors, so that part of the treatment effect we estimate is driven by unobserved differences, then we should obtain a considerably smaller treatment effect when excluding these municipalities.

Specification 2 in Table 7 reports the ATTs once we exclude all of the municipalities with

[13] The EEG allowed older panels to be subsidized from 2000 onwards only if these panels had a major refit after the EEG went into force, i.e., only with substantial investment due to the EEG. The pre-EEG panel data is thus not exhaustive like the post-EEG panel data, since not all pre-EEG panel owners chose to modernize, but it is still very informative about a municipality's propensity to install solar panels.

Table 7: ATT using different specifications and parameters

Description	ProbInstall	NumInstall	NumInstall per 1000	N	Clusters
1. Main matching model	0.118*** (0.007)	0.800*** (0.065)	0.587*** (0.040)	22284	10712
2. Exclude obs with pre-EEG solar panels[a]	0.130*** (0.008)	0.662*** (0.056)	0.650*** (0.050)	17842	8557
3. Fixed effects on district level[b]	0.225*** (0.007)	1.179*** (0.041)	0.692*** (0.044)	20832	10312
4. Exclude multiple panels on same street[c]	0.078*** (0.007)	0.513*** (0.040)	0.144*** (0.029)	21902	10666
5. Without replacement[d]	0.157*** (0.008)	0.659*** (0.065)	0.517*** (0.034)	12898	9862
6. Pooled propensity score estimation[e]	0.137*** (0.007)	0.798*** (0.054)	0.636*** (0.040)	22306	10711

Note: The table reports the average treatment effect on the treated on three outcome variables, the probability, number, and number (per 1000 inhabitants) of solar panel installations within one year using different specifications and parameters in the matching procedure.
[a] All municipalities with solar panels installed before inactment of EEG are dropped from the sample.
[b] Fixed effects on district rather than state level in the propensity score estimation.
[c] Count only the first panel if several panels are installed on the same street.
[d] Main matching model except without replacement of control municipalities.
[e] One propensity score regression over all years with year fixed effects.
***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

pre-EEG panels. Comparing with the main specification that includes those, the estimated treatment effects barely change, even though the sample is reduced by about 20%. In other words, we exclude 20% of the observations which—based on past outcomes—are most prone to install solar panels, and yet our estimates barely change. This result strongly suggests that our matching procedure captures all factors determining treatment and outcomes, and the results are not driven by selection on unobservables.

Fixed effects on district level. Another way to address the concern of unobserved differences is by using fixed effects on a finer geographical level. In the main specification we use state fixed effects, specification 3 in Table 7 uses district (*Kreis*) level fixed effects.[14] Hence, if there are relevant factors at the district level that are not picked up in our main matching procedure, then we should observe different treatment effects once we use these fixed effects on this finer geographical level. However, the table shows that the treatment effect increases, if anything.

[14] We cannot use fixed effects on an even finer level, e.g., on municipality level, because the fixed effect dummies would perfectly predict the outcomes of the municipality. We already lose some observations by using district fixed effects in specification 3, because there is no variation in treatment for some districts in some years. The matching approach relies on the x-vector used for matching not perfectly predicting treatment, otherwise one cannot find municipalities with similar propensity scores that are treated and untreated.

Exclude multiple panels on the same street. There might be the problem that some households install multiple panels. If a household installs several panels in the same year in the same municipality, then it looks as if these new panels are a result of social interaction, while it is actually the same decision maker installing more panels.

Unfortunately we do not observe the identity of the panel installers in the anonymized data. However, since we observe the street and zip code where each panel was installed, we can run robustness tests on the basis of these address information. More specifically, we exclude all solar panels but the first that are installed on the same street. This radical approach will also exclude panels installed due to social influence (e.g., neighbor influenced by peer effects), hence we will obtain overly conservative treatment effect estimates, but we can be sure that all multiples installed by the same household at the same address are removed. Thus, if we still find a positive treatment effect, then it cannot be explained by the same households installing multiple solar panels. Specification 4 in Table 7 reports the results of this robustness check. Although the estimated ATTs are noticeably smaller, the effects are still positive on all outcome measures. Thus, even if we can reasonably exclude the possibility of the same individuals installing multiple panels, the effects remain economically and statistically significant.

Without replacement. We want to determine the impact of allowing replacement of control municipalities on the estimates, which can give single control municipalities a lot of weight. Specification 5 in Table 7 matches with the same ambitious caliper as the main model, but does not allow the same untreated municipality to be matched with different treatment municipalities. We lose almost half of the sample by not allowing replacement, but the resulting ATT estimates are very similar to the main model. Consequently, the results we find are not driven by a few control municipalities that have been matched repeatedly in the main model.

Pooled propensity score estimation. Finally, we replace the year specific propensity score regressions with a single regression spanning all years. If there are problems, for example due to a small sample size in later years, then we should remove those in this robustness check. Specification 6 in Table 7 gives the results, and they are very close to the main model.

In conclusion, reasonable robustness checks confirm the magnitude of the social effects obtained in the main specification. In fact, we have not been able to find a single specification where we did not find positive social effects on any of the outcome variables. Thus, the matching method employed in this paper is very robust to changes in the specification or sample.

5 Policy implications

For the EEG regulation, social interactions that lead to spatial clustering of panels have policy implications along at least three dimensions. First, the EEG regulation offers a compensation rate above market levels. As such, it may pay for households to install solar panel in areas where it would not pay in a competitive setting, for example because of insufficient solar radiation potential in northern Germany.[15] Social effects can exacerbate these inefficiencies if it becomes fashionable to install panels in such areas. And we indeed find that there are positive social effects even in the municipalities with the lowest solar potential (see section 4.1.2). Thus, if the policy also aims at efficient energy production, then it should shift from the posted price mechanism that it currently is to a more competition-based mechanism such as auctions. This way, solar panels would be installed in areas where they can produce the most energy or where energy is needed most, so that social influence in inefficient areas is unlikely to be a problem. Recently, steps have been taken to explore auctions as allocation mechanisms for the small group of solar panels on ground-mounted structures (e.g., BMWi, 2015).[16]

Second, social interaction may lead to local excess power supply. Currently, the regulation does not have any safeguards against excessive installations in areas where more power production is not needed. Thus, the energy that is produced in these areas is either wasted or has to be transported off at a cost, possibly requiring investments in the long-range power transmission grid as the share of renewable energy increases. The latter gives rise to serious NIMBY-problems (e.g., WSJ, 2009; FT, 2015). Consequently, the EEG and other policies whose success and efficiency depends on a spatial distribution should add constraints. For example, the EEG may deny subsidies in areas where a lot of panels already exist or where an excess supply of energy is common. The funds are then better used to install panels where energy is in short supply (e.g., close to major cities or industrial areas).

Third, the large social effects we find—which increase the number of installations by about a half—can in part explain why the participation rate and the associated costs are so large for the EEG program. Anecdotal evidence suggests that a lot of households joined the program after solar panel costs dropped and the regulator did not adjust compensation rates quickly enough, which made participation very attractive. But our findings show these cost decreases are not the only reason.[17] Given that positive social effects increase the covariance of installations locally, we could expect either considerably more or considerably fewer installations by area compared to a model based on isolated decisions. The regulation should

[15]See Figure 1. According to the data, relocating a panel from the area with minimal to the area with maximal solar energy potential within Germany gives a theoretical energy production increase of about 33%. For wind energy potential, differences within Germany are even larger.

[16]As of 6/8/16, the federal cabinet has approved a proposal for extending procurement mechanisms more widely. However, the planned changes will only apply for large panels.

[17]Recall that we compare outcomes in control and treatment municipalities in the same time interval, where panel costs were identical. Note also that there could be an interaction between social influence and price decreases, so that word spreads more quickly if a price decrease made participation more attractive.

therefore add safeguards against excessive claiming of the subsidy to protect consumers from the 20 year payment obligations, which they have to pay via taxes, and also to make the expansion of renewable energy more predictable. It may be surprising that such an obvious recommendation is necessary, but even now the EEG does not have hard constraints in the sense that it stops subsidizing at a threshold capacity or number of panels. Only some soft constraints have been added in the last few years, for example that the compensation rate declines as more panels are installed (see section 2.1). Controlling the number of new solar panels is also important because there is evidence from the US that the carbon emissions offset by each new panel decreases as the stock increases (Novan, 2015). The regulation could also encourage the first installations in areas without panels if panels are desirable there (high power net demand or a lot of sun), which would enhance spatial diversification and thus insure against local weather shocks, by taking advantage of social interactions.

In summary, while the disincentives induced by minimum prices are well known, our results show these effects may be even more pronounced due to social interactions in the case of the EEG. Thus, measures working towards a more competitive allocation mechanism might be even more beneficial than previously thought.

6 Conclusion

In this paper, we analyzed one of the largest renewable energy support programs, the German EEG regulation. This regulation allows households to install solar panels and receive a minimum price for each kWh that is produced, which is above the market compensation. This new environmental policy shifts energy production away from energy companies and towards smaller units such as households. We asked whether the decision to become a producer and install a panel on one's roof is influenced by others.

Our results suggest that social effects play a significant role in explaining if and where solar panels are installed. We estimate that social influence, i.e., the effect of having at least one other solar panel in the municipality, increase the probability and number of new solar panel installations considerably by up to 50%. These large estimates can explain part of the spatial clustering of solar panels that we observe in Germany. In the context of energy, such clustering is crucial to recognize and understand because energy cannot be stored or transported without cost. We also find that the social effect accounts for installations at all levels of solar radiation in Germany, which suggests that social influence can interact with the inefficient policy design. We derive several policy recommendations based on our findings. In particular, a multi-unit auction that induces competition between potential producers would lower compensation rates and ensure production in areas with the largest solar potential or power demand. Thus, more competitive compensation rates would remove the basis for inefficient social effects. Moreover, the regulation should add safeguards against local clustering in areas with excess supply of power and should add an upper threshold on the capacity or number of panels to limit costs. Because social interactions induce a higher covariance in the household installation decisions, the absence of such thresholds may lead to unexpectedly high installation rates, as has been observed in the past.

The EEG renewable energy regulation is one of many which aims at promoting some behaviors while dissuading others, and where the success of the policy depends on how agents respond to the regulation. Another policy, cash for clunkers programs, may aim at stimulating the economy or increasing the share of low-emission cars, and goal-attainment depends on how many households participate and which cars they choose. Both, especially the choice of the car, may very well depend on friends, family and peers, and thus influence the success of the policy. For social programs, it has already been established that peer effects influence program participation (Dahl et al., 2014). Thus, the findings we present here are relevant for policies in other countries and areas.

In future work, it would be interesting to directly investigate the allocative efficiency of the EEG remuneration scheme with data on spatial demand and supply in the power grid, and quantify the costs of offsetting carbon emissions. Previous program evaluations found that subsidies for renewables or energy-efficient consumption goods can reduce emissions, but are rather costly (e.g., Cullen, 2013; Davis et al., 2014). It is also important to investigate the exact social channels that drive the solar panel expansion further. Finally, our results imply the theoretical question how social influence should be included in the design of a policy. 'Mechanism design with social interactions', i.e., mechanism design which recognizes that decisions are not made in isolation, may therefore be a fruitful tool for these kinds of policies.

References

Becker, G. S. and K. M. Murphy (2000): *Social economics. Market behavior in a social environment.*, Harvard University Press.

BMWi (2015): "Designing renewable energy tenders for Germany," `http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/ausschreibungen-eeg-en,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf`, accessed 5/5/2016, Federal Ministry for Economic Affairs and Energy.

——— (2016): "Zeitreihen zur Entwicklung der erneuerbaren Energien in Deutschland," `http://www.erneuerbare-energien.de/EE/Redaktion/DE/Downloads/zeitreihen-zur-entwicklung-der-erneuerbaren-energien-in-deutschland-1990-2015.pdf?__blob=publicationFile&v=6`, accessed 19/4/2016, Federal Ministry for Economic Affairs and Energy.

Bollinger, B. and K. Gillingham (2012): "Peer effects in the diffusion of solar photovoltaic panels," *Marketing Science*, 31, 900–912.

Bramoullé, Y., A. Galeotti, and B. Rogers, eds. (2016): *The Oxford Handbook of the Economics of Networks*, Oxford University Press.

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2014): “Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions,” *Econometrica*, 82, 1273–1301.

Cullen, J. (2013): “Measuring the environmental benefits of wind-generated electricity,” *American Economic Journal: Economic Policy*, 5, 107–133.

Dahl, G. B., K. V. Løken, and M. Mogstad (2014): “Peer Effects in Program Participation,” *The American Economic Review*, 104, 2049–2074.

Davis, L. W., A. Fuchs, and P. Gertler (2014): “Cash for coolers: evaluating a large-scale appliance replacement program in Mexico,” *American Economic Journal: Economic Policy*, 6, 207–238.

Duflo, E. and E. Saez (2003): “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *The Quarterly Journal of Economics*, 815–842.

FT (2015): “Compromise and opposition on road to Germany’s renewable energy,” `http://www.ft.com/intl/cms/s/2/6a254194-47e9-11e5-af2f-4d6e0e5eda22.html`.

Gaviria, A. and S. Raphael (2001): “School-based peer effects and juvenile behavior,” *Review of Economics and Statistics*, 83, 257–268.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996): “Crime and Social Interactions,” *The Quarterly Journal of Economics*, 111, 507–548.

Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman (2003): “The social multiplier,” *Journal of the European Economic Association*, 1, 345–353.

Graziano, M. and K. Gillingham (2015): “Spatial Patterns of Solar Photovoltaic System Adoption : The Influence of Neighbors and the Built Environment,” *Journal of Economic Geography*, 15, 815–839.

Jackson, M. O. (2016): “The Past and Future of Network Analysis in Economics,” in *The Oxford Handbook of the Economics of Networks*, ed. by Y. Bramoullé, A. Galeotti, and B. Rogers, Oxford University Press.

Katz, L. F., J. R. Kling, and J. B. Liebman (2001): “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment,” *The Quarterly Journal of Economics*, 116, 607–654.

Mas, A. and E. Moretti (2009): “Peers at Work,” *American Economic Review*, 99, 112–145.

Novan, K. (2015): “Valuing the wind: renewable energy policies and air pollution avoided,” *American Economic Journal: Economic Policy*, 7, 291–326.

P3 Group (2015): "Entwicklung des Ausbaus der erneuerbaren Energien in den Jahren 2016 bis 2020," http://www.netztransparenz.de/de/file/20151015_EE-Prognose_2016_-_2020.pdf, accessed 19/4/2016.

Richter, L. (2014): "Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK," mimeo.

Rode, J. and A. Weber (2016): "Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany," *Journal of Environmental Economics and Management*, 78, 38–48.

Rosenbaum, P. R. and D. B. Rubin (1983): "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41–55.

Sacerdote, B. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *The Quarterly Journal of Economics*, 116, 681–704.

Schwarz, A. (2014): "Die Förderung der Stromerzeugung aus erneuerbaren Energien in Deutschland," Wissenschaftliche Dienste des Deutschen Bundestags (German Parliament).

Sianesi, B. (2004): "An evaluation of the Swedish system of active labor market programs in the 1990s," *Review of Economics and Statistics*, 86, 133–155.

Trogdon, J. G., J. Nonnemaker, and J. Pais (2008): "Peer effects in adolescent overweight," *Journal of Health Economics*, 27, 1388–1399.

WSJ (2009): "Renewable Energy, Meet the New Nimbys," http://www.wsj.com/articles/SB125201834987684787.

Young, P. H. (2009): "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *American Economic Review*, 99, 1899–1924.

A Matching quality of the main specification

It is crucial for the validity of the matching approach to reconstruct a comparable control group to obtain the counterfactual outcomes of all treatment observations. This is done by matching the treatment observations to similar control observations conditional on covariates. Although the identifying assumption (CIA) of the ATT is not directly testable, the similarity of treatment and control group can be directly assessed. If groups are not sufficiently similar, then we might also expect the outcomes of the control group to differ from the counterfactual outcomes of the treatment group.

Figure 8 displays the distribution of propensity scores before matching, i.e., the propensity score distribution among all municipalities treated in a given year and among all untreated municipalities. There are large differences in the distribution, especially in the early

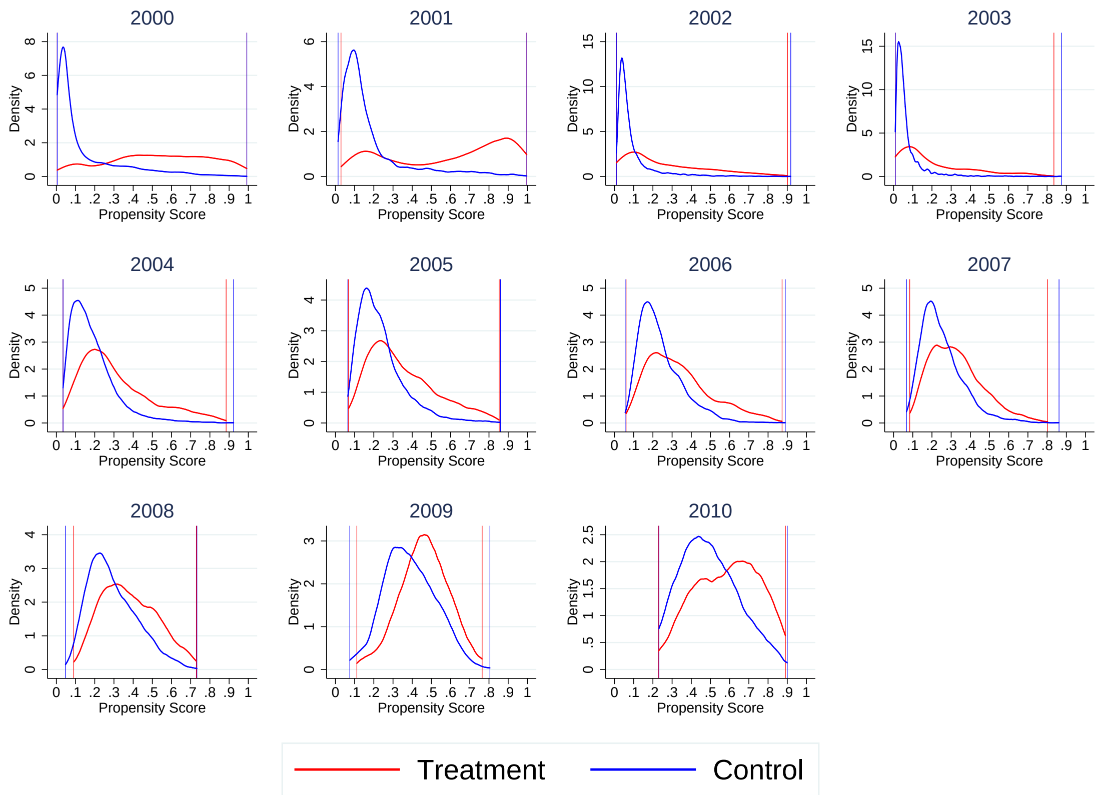


Figure 8: Distribution of propensity scores *before* matching among treated and untreated municipalities

years, so it is clearly necessary to control for the differences between these two groups. Simply comparing outcomes of control and treatment municipalities, even when controlling for other covariates, will not be sufficient.

Figure 9 shows the distribution of propensity scores after matching, i.e., among all matched treated and among all matched untreated municipalities. There are no visible differences left in the distribution of propensity scores after matching.

The propensity score maps all covariates into a single scalar. Thus, a similar propensity score distribution may still show imbalances on individual covariates. To investigate this possibility, Figure 2 (in the main text) displays the distribution of all covariates for all municipalities in treatment and control group after matching. For all variables, the distribution is very similar between treatment and control municipalities; for a few variables such as voter turnout there are small differences on subsets of the support.

Finally, for a more formal test of similarity between the groups, Table 8 reports t-tests for mean differences of all the variables. None of the variable means are significantly different between treatment and control group.

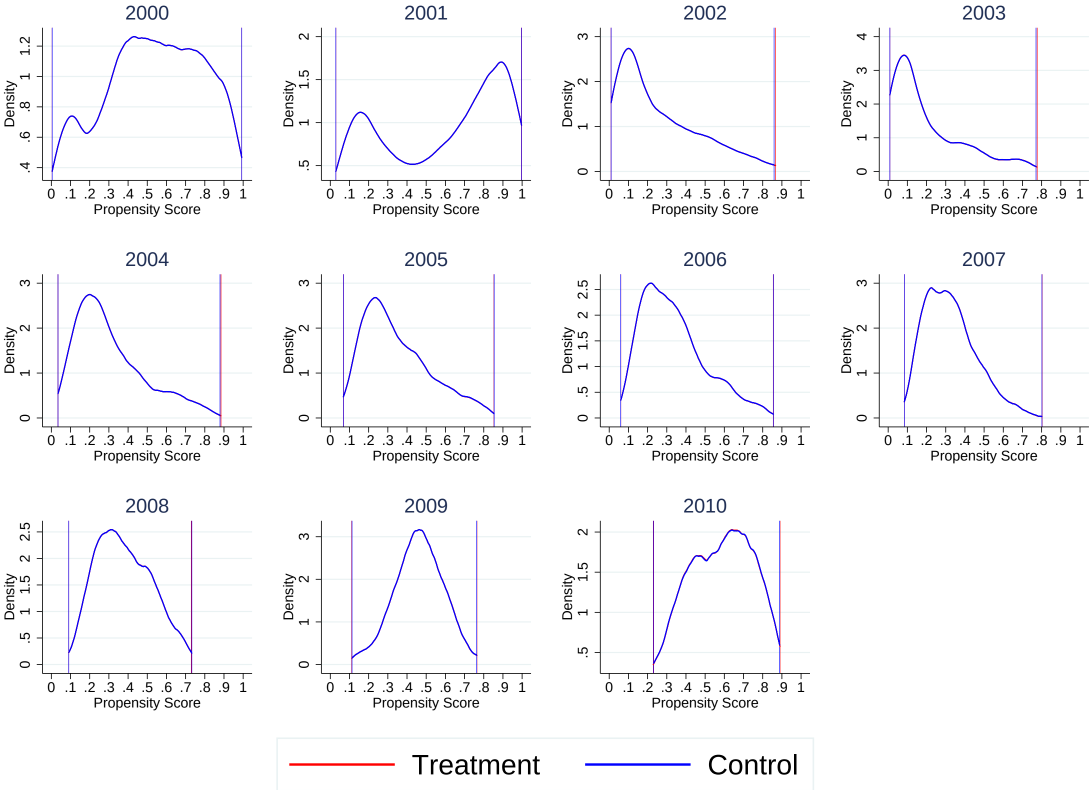


Figure 9: Distribution of propensity scores after matching among matched treated and matched untreated municipalities

B Data and imputation

B.1 Data sources

In this paper we use data from various sources, see Table 9. First, we obtained the solar radiation data for each square kilometer in Germany. In order to match the radiation data to municipalities in Germany, we—second—obtained spatial data containing the exact municipality partitions for each year from the German Federal Agency for Cartography and Geodesy, which allows us to match any coordinate pair in Germany to a municipality. Third, we obtained municipality characteristics such as the number of inhabitants or mean income from all statistical state offices in Germany. Not all of the variables are collected or published every year, so we used a straightforward data imputation procedure to fill in the gaps (see Appendix B.2 for details). Fourth, we obtained lists of all EEG subsidized solar panels in Germany including size, data of installation, and address from the four transmission system providers in Germany. These lists of all solar panels are very reliable, since compensation is determined on this basis. Overall, we observe 1.29 million panels in these lists. Fifth, in order to obtain the coordinates of the panel locations and to match the panel addresses with municipalities and spatial solar radiation data, we used open street map, a service similar to Google maps but free of charge.

Table 8: Balance Check

	Control	Treatment	p-value
Total area	2818.1	2799.8	0.65
Res. buildings	1282.0	1308.7	0.51
Inhabitants	5605.8	5812.0	0.41
Mean income	29.41	29.47	0.54
Unemployment rate	0.0453	0.0453	1.00
Global radiation	1057.6	1058.1	0.51
Population density	1.793	1.794	0.97
Share seniors	0.180	0.179	0.33
Votes Green	0.0557	0.0564	0.09
Votes Conservative	0.422	0.421	0.73
Votes Soc-Dem	0.346	0.347	0.56
Votes Liberals	0.0792	0.0791	0.83
Votes Left	0.0624	0.0613	0.29
Votes others	0.0330	0.0332	0.34
Voter turnout	0.735	0.733	0.06
Observations	22284		

Note: The matching algorithm imposes common support of estimated propensity scores , replacement of control municipalities , a caliper of 0.01.

Table 9: Used data and sources

Data	Source
Lists of all photovoltaic panels in Germany subsidized under the EEG	The 4 energy net providers, `http://www.netztransparenz.de`
Municipality characteristics (income, unemployment, etc.)	Federal Statistical Office of Germany, and state offices, `https://www.regionalstatistik.de`
Spatial data on partition and location of all German municipalities	German Federal Agency for Cartography and Geodesy, `http://www.geodatenzentrum.de`
Solar energy generation potential by location, calculated from records 1981 to 2010	German Weather Service (DWD), Office Hamburg
Spatial data to geocode panel addresses	Open Street Map, `https://www.openstreetmap.org/`

B.2 Imputation

Since the municipality variables we rely on to implement the matching are not gathered every year in every municipality, we would lose a majority of the observations without imputation. We have data for most municipalities but have gaps for certain variables in certain years. We use a straightforward imputation procedure to fill these gaps. For a given variable of a given municipality, we fill a gap in a year by using the value of the closest year. If there is a municipality for which we have no data in any year, then we cannot impute this way and do not use the municipality. Table 10 gives a simple example how the imputation works. Clearly, there are other ways to implement the imputation, for example as weighted averages, but slightly different specifications did not change our results.

Table 10: Imputation procedure illustrated on a hypothetical example

year	municipality	inhabitants	year	municipality	inhabitants
2000	1	10000	2000	1	10000
2001	1		2001	1	**10000**
2002	1		2002	1	**12000**
2003	1		2003	1	**12000**
2004	1	12000	2004	1	12000
without imputation			with imputed values		

Table 11: Estimation of Propensity Score by Year

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Total area	0.000***	0.000***	0.000***	0.000*	0.000***	0.000**	0.000**	0.000	-0.000	0.001**	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total area × Total area	-0.000***	-0.000***	-0.000***	-0.000	-0.000***	-0.000	-0.000	-0.000	0.000	-0.000	0.000**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Res. buildings	0.000***	0.002***	0.002***	0.002**	0.001	0.004***	0.001	-0.002	0.007**	0.001	0.007
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Res. buildings × Res. buildings	-0.000***	-0.000***	-0.000***	-0.000	-0.000	-0.000***	-0.000	-0.000	-0.000***	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inhabitants × Mean income	-0.000***	-0.000	0.000	-0.000	0.000	-0.000	0.000	0.000	-0.000	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inhabitants × Population density	-0.000**	-0.000	0.000	-0.000	-0.000**	0.000	0.000	-0.000	-0.000*	0.001**	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inhabitants × Share seniors	0.000	-0.000	0.001	-0.000	0.001	-0.001	-0.001	-0.001	-0.004	-0.003	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Inhabitants × Voter turnout	0.000***	0.000	0.000	0.000	-0.000	0.000	-0.001	-0.000	-0.003	-0.002	-0.005
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Unemployment rate	-4.893	7.073	-2.991	6.963	-5.938	-0.845	8.250	8.824	-12.819	-7.367	2.499
	(6.42)	(6.09)	(8.19)	(8.85)	(5.96)	(5.50)	(7.23)	(8.43)	(8.92)	(4.79)	(13.41)
Unemployment rate × Unemployment rate	8.329	-75.509**	-27.974	-53.246	-16.102	-7.751	-77.590*	-60.536	70.501	8.187	-4.537
	(43.16)	(36.31)	(47.01)	(50.64)	(30.89)	(26.70)	(40.33)	(57.47)	(58.13)	(10.60)	(74.47)
Inhabitants	-0.000**	-0.000**	-0.000*	-0.000	0.000	-0.000	0.001	0.002	0.004*	0.002	0.004
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inhabitants × Inhabitants	0.000***	0.000*	0.000**	0.000	-0.000	0.000	0.000	-0.000	0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Mean income	0.048***	0.036	0.058	0.001	0.034	0.017	0.000	0.043	0.016	-0.003	-0.025
	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)	(0.03)	(0.02)	(0.04)	(0.03)	(0.05)	(0.10)
Mean income × Mean income	-0.000*	-0.000	-0.001	0.000	-0.001	-0.000	0.000	-0.001	0.000	0.000	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Global radiation	0.005	-0.157***	-0.163***	-0.062	0.106	-0.043	-0.010	0.286*	0.030	-0.185	0.687*
	(0.03)	(0.04)	(0.06)	(0.08)	(0.07)	(0.09)	(0.11)	(0.16)	(0.18)	(0.24)	(0.37)
Global radiation × Global radiation	0.000	0.000***	0.000***	0.000	-0.000	0.000	0.000	-0.000*	-0.000	0.000	-0.000*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Population density	0.160***	0.225***	0.237***	0.155	-0.161	0.020	0.484**	0.087	0.128	0.517	0.102
	(0.04)	(0.06)	(0.08)	(0.11)	(0.12)	(0.13)	(0.20)	(0.24)	(0.38)	(0.53)	(0.93)
Population density × Population density	-0.003*	-0.013**	-0.013**	0.001	0.029**	-0.008	-0.067	0.021	0.032	-0.217	-0.214
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)	(0.10)	(0.17)	(0.27)
Share seniors	13.797***	14.435***	4.034	12.313**	8.389***	2.920	-0.719	12.130*	10.435	12.700	13.100
	(4.82)	(4.05)	(4.65)	(5.40)	(3.05)	(2.82)	(3.14)	(7.22)	(9.01)	(10.39)	(13.42)
Share seniors × Share seniors	-45.306***	-41.268***	-21.273	-40.783**	-35.453***	-10.255	-0.425	-27.134*	-20.617	-32.812	-18.471
	(14.14)	(11.51)	(13.73)	(16.13)	(9.47)	(7.90)	(9.04)	(15.91)	(19.55)	(23.76)	(27.11)
Votes Conservative	-18.158***	-6.546**	-7.276**	-9.246**	2.344	-7.041**	4.115	-2.456	-1.674	0.025	-9.154
	(2.61)	(2.57)	(3.70)	(4.34)	(2.98)	(2.93)	(3.67)	(3.99)	(4.57)	(5.45)	(8.27)
Votes Conservative × Votes Conservative	6.793***	0.679	4.406	8.175**	-4.789	9.012***	-4.032	4.792	5.148	-1.083	-2.202
	(2.52)	(2.42)	(3.65)	(4.10)	(3.21)	(3.15)	(4.10)	(4.32)	(5.03)	(5.98)	(8.99)
Votes Soc-Dem	-5.896*	-6.031**	-1.258	-0.175	-8.246***	5.866*	-2.566	2.772	3.674	-3.380	-15.657**
	(3.29)	(3.01)	(4.60)	(5.29)	(2.85)	(3.42)	(3.35)	(4.21)	(4.29)	(4.79)	(7.56)
Votes Soc-Dem × Votes Soc-Dem	-7.153*	-0.388	-2.494	-3.935	8.719***	-9.173**	4.873	-1.885	-2.147	4.006	5.595
	(3.79)	(3.44)	(5.20)	(5.96)	(3.07)	(4.05)	(3.82)	(4.41)	(4.59)	(5.38)	(7.79)
Votes Liberals	-2.822	-3.354	-1.890	11.852	2.968	4.026	-0.809	8.337*	2.104	14.775**	3.185
	(6.42)	(3.82)	(4.14)	(8.09)	(3.79)	(3.74)	(3.91)	(4.92)	(4.58)	(6.33)	(9.52)
Votes Liberals × Votes Liberals	-68.239*	-14.094	0.857	-68.180*	-16.177	-6.669	6.518	-16.975	9.413	-33.895*	-43.875
	(36.81)	(15.85)	(10.99)	(38.36)	(12.69)	(11.67)	(11.70)	(15.10)	(13.20)	(20.13)	(31.99)
Votes Left	-12.682*	-2.497	0.579	-11.688*	-4.638	-5.405*	1.769	5.801	3.797	-2.086	-2.213
	(6.54)	(5.20)	(7.48)	(6.96)	(3.05)	(3.15)	(3.46)	(4.10)	(4.33)	(4.93)	(7.36)
Votes Left × Votes Left	-6.575	-6.996	-18.511	34.799*	1.683	-1.396	-3.441	-11.864	-5.640	6.917	-22.122
	(20.57)	(15.51)	(25.07)	(19.53)	(7.69)	(7.79)	(8.16)	(9.01)	(9.63)	(11.25)	(15.70)
Votes others	-13.204**	-3.770	10.323	-0.026	-5.227	0.644	5.584	7.959	18.213**	2.684	0.342
	(6.00)	(6.09)	(8.21)	(5.52)	(3.81)	(3.78)	(5.33)	(6.08)	(8.65)	(6.74)	(10.81)
Votes others × Votes others	23.320	-72.144	-140.679*	1.831	3.781	2.449	-50.328	-59.341	-138.948*	-42.770	-44.018
	(62.82)	(68.76)	(84.35)	(26.83)	(23.65)	(21.16)	(42.54)	(42.44)	(75.73)	(46.85)	(69.12)
Voter turnout	-11.253	9.278	-5.287	-7.387	14.606	-2.143	8.260	9.406	-2.459	20.015	26.550
	(7.72)	(7.01)	(9.91)	(11.43)	(9.39)	(8.21)	(11.38)	(13.26)	(12.87)	(15.94)	(21.44)
Voter turnout × Voter turnout	5.979	-6.535	2.297	4.203	-11.340*	1.431	-5.568	-7.664	4.243	-14.117	-17.662
	(5.17)	(4.83)	(6.99)	(7.97)	(6.72)	(5.91)	(8.26)	(9.39)	(8.96)	(11.00)	(14.85)
Constant	5.730	79.402***	86.547***	29.367	-67.122*	16.670	-9.262	-161.901**	-25.650	82.728	-362.833*
	(16.09)	(19.53)	(31.12)	(42.77)	(37.38)	(47.45)	(57.93)	(81.49)	(92.51)	(124.45)	(194.74)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10873	9113	6396	5643	5178	3968	2881	2078	1498	943	430
Pseudo R^2	0.350	0.349	0.191	0.172	0.119	0.099	0.089	0.069	0.075	0.066	0.120

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parentheses.