UNIVERSITY OF CAMBRIDGE | Energy Policy Research Group

Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK

EPRG Working Paper 1332

Cambridge Working Paper in Economics 1357

Laura-Lucia Richter

Abstract The main research question in this paper is whether the installation rate of solar PV technology is affected by social spillovers from spatially close households. The installed base, defined as the cumulative number of solar PV installations within a neighbourhood by the end of a particular month, serves as a measure for the social effects of interest. Motivated by the technology-specific time lag between the decision to adopt a solar PV panel and the completion of the installation, the third lag of the installed base serves as main regressor of interest in the panel data model employed. The results suggest small, but positive and significant social effects that can be exploited to promote adoption: at the average installation rate of 0.7 installations per 1,000 owner-occupied households, one more solar PV panel in the postcode district increases the installation rate three months later by one percent. At the average number of 6,629 owner-occupied households within a postcode district, this implies an increase in the number of new installations in the neighbourhood by 0.05. Projects involving a high number of installations could hence promote diffusion. A major limitation of the model is that social spillovers are assumed to spread within defined neighbourhoods, only. Spatial econometric methods could allow for social effects across these borders.

Keywords social effects, installed base, product adoption, diffusion, solar PV technology, micro-generation

JEL Classification C19, D12, D83, Q21, Q42

Contact llr23@cam.ac.uk
Publication December 2013
Financial Support DAAD – German Academic Exchange Service

www.eprg.group.cam.ac.uk

EPRG WORKING PAPER

EPRG 1332

Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK*

Laura-Lucia Richter†‡
University of Cambridge

April 10, 2014

Abstract

The main research question in this paper is whether the installation rate of solar PV technology is affected by social spillovers from spatially close households. The installed base, defined as the cumulative number of solar PV installations within a neighbourhood by the end of a particular month, serves as a measure for the social effects of interest. Motivated by the technology-specific time lag between the decision to adopt a solar PV panel and the completion of the installation, the third lag of the installed base serves as main regressor of interest in the panel data model employed. The results suggest small, but positive and significant social effects that can be exploited to promote adoption: at the average installation rate of 0.7 installations per 1,000 owner-occupied households, one more solar PV panel in the postcode district increases the installation rate three months later by one percent. At the average number of 6,629 owner-occupied households within a postcode district, this implies an increase in the number of new installations in the neighbourhood by 0.05. Projects involving a high number of installations could hence promote diffusion. A major limitation of the model is that social spillovers are assumed to spread within defined neighbourhoods, only. Spatial econometric methods could allow for social effects across these borders.

*My most profound thanks to Dr. Michael Pollitt and Dr. Melvyn Weeks for their support and consult as supervisors of this paper. I am grateful for valuable suggestions of Professor Oliver Linton and Dr. Alexei Onatski in context of the PhD Econometrics Workshop. And I thank Alexander Ross, Julien Gagnon and seminar participants of the University of Cambridge for their helpful suggestions and comments.

†PhD Candidate, Faculty of Economics, University of Cambridge, UK. *Email*: llr23@cam.ac.uk.

‡This is a shortened version of the paper. The original paper is available online as EPRG Working Paper 1332 and Cambridge Working Paper in Economics 1357: http://www.eprg.group.cam.ac.uk/.

Contents

List of Abbreviations

cdf	Cumulative distribution function
CMT	Continuous Mapping Theorem
DECC	Department of Energy and Climate Change
FD	First Difference
FE	Fixed Effect
FIT	Feed-in-Tariff
HH	Household
IEA	International Energy Agency
IMD	Index of Multiple Deprivation
LA	Local Authority
LSOA	Lower Level Super Output Area
GLS	Generalized Least Squares
NEED	National Energy Efficiency Database
OLS	Ordinary Least Squares
ONS	Office for National Statistics
OVB	Omitted Variable Bias
PCD	Postcode District
POLS	Pooled Ordinary Least Squares
PV	Photovoltaic
RE	Random Effects
ST	Slutsky Theorem
WG	Within-Group
WLLN	Weak Law of Large Numbers
zip	zip code (first four digits of a full UK postcode)

List of Figures

List of Tables

1 Introduction

The determinants of new technology adoption have been addressed in the economic literature for several decades. Among others, they matter for political, economic and business reasons. A recent example is the market for micro-generation technologies that can be installed by households, communities and small commercial sites. The installed capacity of those small-scale installations goes up to 50kW for electricity and 300kWth for heat generation (The Green Energy Act 2009). In a context of the EU target to increase the share of renewable electricity generation beyond 15 percent by 2020 and given the legally binding domestic energy policy goals to decrease national carbon emissions by 80 percent by 2050 compared to 1990, the UK government intends to encourage households to adopt micro-generation technologies and produce their own low-carbon electricity. Beyond these political reasons the analysis of the diffusion of micro-generation technologies is relevant for economic and business reasons: decentralised electricity generation has the potential to change the (energy) consumer - producer relationship, to alter the economic relations between customers and energy suppliers and to lead to new ownership and business models (Snape and Rynikiewicz, 2012, Watson and Devine-Wright, 2011).

So far, feed-in-tariffs (FIT) are the major instrument to promote adoption of small-scale electricity generation. In the UK they have been paid since April 2010 to mitigate the relatively high costs and uncertainties of solar PV, wind, hydro and anaerobic digestion technology. The FIT is paid for each kWh of electricity generated and the rate paid depends on the size of the system, the technology and the date of completion of the installation[1]. In addition, there is an export tariff that is paid if the micro-generated electricity is exported to the grid.

This paper focuses on the most established micro-generation technology in the UK: solar PV technology.[2] However, as the government's 2015 Micro-generation Strategy claims, financial incentives are not enough for sustained growth of micro-generation technologies. Besides non-financial barriers (e.g. related to insurance and warranties) non-financial *drivers* of growth should be in focus and exploited in future policy and market strategy design towards a low-carbon decentralized electricity system. In particular, social effects from others might impact the adoption decision and hence drive diffusion. As suggested by Weber and Rode (2012), solar PV panel installations in a neighbourhood are visible for passers-by, reducing uncertainty surrounding this technology. Observational learning from spatially close households might thus lead to a correlation of adoption decisions within neighbourhoods. If the adoption behaviour of others significantly affects the adoption of solar PV technology, targeted interventions could serve as attention catching projects that could promote diffusion effectively.

The main research question in this paper therefore is, whether the installation rate of solar PV technology is affected by social effects as measured by the installed base in the immediate local environment. The installed base thereby refers to the cumulative number of solar PV installations within a neighbourhood at the end of a particular month and serves as a measure for social effects, such as observational learning. The analysis is based on installation data

[1] The completion date is crucial for the determination of the FIT. The FIT valid at the installation date is paid for 20 years.

[2] By March 2013, 98.55 percent of all micro-generation installations were solar PV systems. Wind contributed only 1.22 percent, micro CHP 0.12 percent and anaerobic digestion and hydro even smaller shares of 0.01 and 0.11 percent.

that has been collected by Ofgem since the introduction of the FIT in April 2010. The main econometric panel data model specifies the postcode district-month as the smallest unit of observation. Identification of the suggested social effects within neighbourhoods is challenged by multiple endogeneity issues. To address these and to consistently estimate the installed base effect as a measure of social effects from spatially close households, the installation rate within a postcode district is modelled as being affected by the installed base three months prior. This is motivated by the technology-specific time lag between the *decision* to adopt a solar PV panel and the *completion* of the installation. Besides the lagged installed base and month dummies, the main panel equation includes time-varying fixed effects to account for potential self-selection into neighbourhoods and for correlated unobservable neighbourhood characteristics that are constant within a neighbourhood but vary over time. A first difference estimation strategy then yields unbiased and consistent estimates. Further model specifications allow for a time-varying installed base effect, consider different lags of the installed base as well as different geographical areas for robustness. In a last specification differences of the social effects across distinct groups of the population are analysed.

The paper contributes to previous literature in performing the first econometric analysis of the diffusion of solar PV technology within the UK. In particular, it delivers empirical evidence, in how far the adoption behaviour of others drives diffusion. The analysis is based on a remarkably recent and granular solar PV installation dataset of the UK. The results can be exploited for targeted marketing and resource allocations for the stimulation of future adoption. Nevertheless, the analysis has its limitations. Firstly, social effects are assumed to spread within neighbourhoods as defined by the neighbourhoods, only, while spillovers across neighbourhood borders are ignored. Spatial econometric methods, for example, could be employed to allow for more diverse spillover effects. However, it is a useful first highly disaggregated approach to explore the impact of social effects on solar PV adoption, that can be extended in future research. Another limitation is the aggregation to the neighbourhood level. Future research should make use of household level covariate data to further analyse the mechanisms underlying the adoption behaviour. In context of the National Energy Efficiency Database (NEED) the Department for Energy and Climate Change (DECC) is currently creating a database that matches the solar PV installation data with household characteristics.

This paper is organised as follows. Section 2 reviews literature on social effects in the diffusion of solar PV technology and previous applications and challenges of the installed base approach. Section 3 specifies the econometric model and justifies the third lag of the installed base as measure for social effects in solar PV technology adoption. Section 4 presents endogeneity issues that can challenge the estimation of the social effects as measured by the installed base and highlights the identification strategy exploited in the paper. Section 5 introduces the estimation strategy and derives conditions for consistent estimation of the social effects. Section 6 presents the data. The main results and the results of several robustness checks are presented in section 7 and section 8, respectively. To get an idea to what extent social effects vary across neighbourhood characteristics, the impact of contextual factors on social effects is considered in section 9. Finally, section 10 points out the major limitations of the analysis and section 11 concludes.

2 Literature Review

Existing literature has consistently found that cost reductions and subsidy policies are critical for the diffusion of solar PV technology, research on the role of non-financial barriers (such as insurance and warranty issues) and non-financial drivers (such as social effects, supply-side customer service programmes or a structural awareness change) is still scarce. This paper focuses on the role of social effects such as observational learning from the adoption behaviour of others, as one potential driver of solar PV technology adoption. The distinction between causal social effects and contextual effects (such as neighbourhood demographics) as sources of spurious correlation is thereby important: while social effects lead to a social multiplier, contextual effects do not.

Due to limited data availability it is often hard to pin down the precise channel of social spillovers, but multiple economic models of social effects are consistent with observational learning. In particular, Rogers' traditional theory of 'Diffusion of Innovations' (1962) considers observability as one of five major influential factors of an individual′s decision to adopt.

Two main studies have previously addressed observational learning as a potential driver of solar PV technology adoption: Bollinger and Gillingham (2012) estimate peer effects in the diffusion of residential solar PV systems in California and Weber and Rode (2012) analyse information and knowledge spillovers in the diffusion of solar PV technology in Germany. Both papers use the so-called installed base as a measure for social spillovers: the installed base refers the cumulative number of solar PV systems within a defined reference group at a specific point in time and arguably impacts the adoption behaviour of others within this group. Bollinger and Gillingham (2012) model the propensity of solar PV technology adoption in a Californian zip code as a function of the installed base in that zip code. Within this framework, their identifying assumption is that social interactions do not have an effect until the solar PV installations have actually been completed. The time lag between adoption decision and the beginning of the social spillovers is hence crucial in their model. They find that peer effects are increasing in magnitude over time, stronger for larger installations and on a more localized level. Weber and Rode (2011) also find support for the argument of social effects in the adoption of solar PV technology. They establish an epidemic diffusion model that includes a spatial dimension and argue that solar PV systems are easily visible to passers-by and that learning is possible without direct social interaction. As such, they assume information flows between spatially close neighbours and find that imitative adoption behaviour is highly localized and plays a significant role in the diffusion of solar PV technology.

Even if previous studies suggest the existence of social effects in solar PV technology adoption, further empirical evidence is needed. In particular, analyses should be based on more granular data and should be pursued for different countries.

3 Econometric Model

Consider the following linear dynamic panel model to analyse the impact of social effects on the adoption and diffusion of solar PV installation.

$$y_{zt} = \alpha_t + \beta \cdot b_{zt-3} + \underbrace{\alpha_{zq} + \epsilon_{zt}}_{u_{ztq}} \tag{1}$$

$$\text{where} \quad b_{zt-3} = \sum_{\tau=1}^{t-3} Y_{z\tau} \forall z, t \tag{2}$$

The model features three dimensions: firstly, there are the neighbourhoods z in the cross-section dimension. These are geographically defined neighbourhoods, as solar PV systems are visible to passers-by and learning is possible for anyone spatially close. By defining the reference groups in a geographical sense rather than on the basis of family or friends, the social effects identified are consistent with the idea of observing solar PV installations of spatially close households. Secondly, there are the months t in a first time-dimension and thirdly the quarters q in a second time-dimension. The months $t = 1, 2, 3$ pertain to quarter $q = 1$, the months $t = 4, 5, 6$ to $q = 2$ and so forth.

The outcome variable y_{zt} is an aggregation of individual choices. It is defined as the number of new solar PV installations Y_{zt} per owner-occupied household within a neighbourhood z in month t, $y_{zt} = \frac{Y_{zt}}{N_{zt}}$, and measures the degree of adoption and diffusion of solar PV technology. This installation rate as a measure of diffusion is preferred over a count variable Y_{zt}, as the number and the tenure type of households within a neighbourhood might vary considerably across neighbourhoods and a positive correlation between the number of owner-occupied households and the number of installations within a neighbourhood is likely. Dividing the count variable Y_{zt} by the respective number of owner-occupied households in the neighbourhood, N_{zt}, can control for this variation. y_{zt} is hence fractional ($y_{zt} \in [0, 1]$) and in the given application of solar PV technology likely to be very small.

For the normalisation the number of *owner-occupied* households is preferred over the total number of households in the neighbourhood, as the decision to install solar PV is likely to be made by households that own their property rather than by households that rent it.

Equation 2 defines the third lag of the installed base, b_{zt-3}, as the cumulative number of solar PV installations within a neighbourhood z by the end of time period $t - 3$.

The (lagged) installed base of domestic solar PV panels within a neighbourhood is used as a measure of social effects within this neighbourhood: installed solar PV systems are visible for passers-by and reduce uncertainty regarding the technology. Via the installed base, social effects such as observational learning can thus impact the adoption behaviour within the neighbourhood. The third lag of the installed base captures the technology specific time lag between the *decision* to adopt a solar PV panel and the *completion* of the installation: according to quotas of the main solar PV suppliers in the UK, the lead time between the first contact with the supplier (which can be seen as a measure of a household′s decision to adopt) and the completion of the installation usually lies between two and three months. Hence, the inclusion of the third *lag* of the installed base, $b_{zt-3} = \sum_{\tau=1}^{t-3} Y_{zt}$, is justified. All adoptions of solar PV panels up to $t - 3$, i.e. $Y_{zt-\tau} \forall \tau \geq 3$ thus enter the right hand side of the equation. The parameter of interest is β.

Apart from the installed base no further regressors are explicitly included in the equation. Instead, the right hand side of equation 1 includes three kinds of unobserved variables to capture the factors impacting the adoption rate within a neighbourhood, including in particular those characteristics that are correlated with the installed base. There are two types of fixed effects, α_t and α_{qz}, and there is an unobservable error term ϵ_{zt}:

The α_t are 35 month dummies[3], capturing month specific effects that are constant across neighbourhoods and could confound the estimation of the social effects. An example for such month-specific effects are nationwide policy announcements regarding the FIT. Historically, announcements to decrease the FIT have often been correlated with higher numbers of adoptions right before and with a lower number of adoptions after the effectiveness date of the policy. Such events should be controlled for.

Compared to a standard panel model that includes time-constant fixed effects, α_z, to control for unobserved heterogeneity on the neighbourhood level, equation 1 is more general: the model is specified with time-varying fixed effects[4]. However, since the installed base b_{zt-3} is defined on the neighbourhood-month level, i.e. is by definition the same for all households within neighbourhood in a particular month, a specification with neighbourhood-month effects, α_{zt}, is not feasible. These would be perfectly collinear with the installed base and hence prohibit identification of the installed base effect β. Neighbourhood-quarter effects, α_{zq}, on the other hand, avoid this collinearity problem and allow to control for neighbourhood characteristics that are constant within a neighbourhood and a quarter but vary over time, i.e. across quarters. Those fixed effects control for time-varying neighbourhood specific characteristics that are relevant for the installation rate in the neighbourhood. In particular they control for factors that are correlated with the lagged installed base in the neighbourhood *and* with the outcome variable y_{zt}. Those factors, if omitted from the equation, would lead to omitted variable bias of the installed base effect β. The neighbourhood-quarter effects could capture time and location-specific activities such as advertising or marketing campaigns. Changes of neighbourhood characteristics within a quarter are assumed to be negligible in the sense that existing variations do not significantly impact the decision to install solar PV. As an example, marginal changes in average income within a neighbourhood-quarter are assumed to be negligible for the installation rate, since the decision to invest in a solar PV panel is rather a question of accumulated capital than of marginally higher income.

Finally, ϵ_{zt} is an *i.i.d.* unobserved error term that captures random neighbourhood and month specific effects. In particular $E(b_{zt-3}\epsilon_{zt}) = 0$

[3] April 2010 being the reference category

[4] The model has also been specified with time-constant fixed-effects, α_z, only. The results are consistent with the theory of OVB and are available upon request to the author.

4 Endogeneity and Identification

The econometric panel model relies on the third lag of the installed base, b_{zt-3}, to measure social effects within geographically defined neighbourhoods. The large number of fixed effects controls for observable and unobservable characteristics that could confound the estimation.

First of all, there might be self-selection into neighbourhoods. This problem has its origins on the household level, but implies non-randomness on the neighbourhood level as well: if factors driving households to live in a specific neighbourhood are also correlated with the adoption rate, y_{zt}, spurious correlation results. For example, energy conscious households might prefer to live in environmentally friendly neighbourhoods and this unobserved preference might also make them behave similarly regarding the adoption of solar PV technology. If so, an observed correlation of the installed base with the adoption decisions within the neighbourhood can be misleading. On the neighbourhood level this self-selection into neighbourhoods implies that c.p. a higher share of households with 'green' preferences lives in the same neighbourhood: pronounced 'green', environmental awareness within the neighbourhoods might increase the installed base *and* the adoption rate within the neighbourhood. Hence, some of the observed correlation between installed base and adoption rate is likely to be spurious, resulting from the correlation of unobserved tastes with the adoption rate rather than from social effects.

Secondly, unobservable neighbourhood characteristics, such as solar PV supplier activities or local advertising campaigns, can result in a correlation of adoption behaviour within the neighbourhoods and lead to spatial clustering, i.e. to spurious correlation of the installed-base with the adoption rate.

Manski′s (1993) well-known 'reflection problem' on the other hand is *not* a problem in the case of solar PV technology. The reflection problem refers to a phenomenon that frequently challenges the identification of social effects. It occurs if an individual′s adoption decision depends on the behaviour of others within her reference group and simultaneously impacts the adoption decisions of the others in that group. In case of solar PV technology adoption, this characteristic simultaneity is *not* a problem: the decision to adopt takes place on average three months *before* the completion of the solar PV system. Households installing solar PV panels in month t made their decision to adopt three months prior and it was the installed base in month $t-3$, b_{zt-3}, that affected their behaviour.

Summing up, the technology specific time-lag between the *decision* to adopt a solar PV panel and the *completion* of the installation is crucial for identification and consistent estimation of the installed base effect. While the reflection problem is not an issue in the application of solar PV technology, self-selection and correlated unobservables are likely to be problematic and must be considered for identification and consistent estimation of the social effects. This is done via the fixed effect specification of the model.

5 Estimation Strategy

5.1 Pooled OLS Estimation

Consider equation 1. To consistently estimate the installed base effect β by pooled OLS (POLS), the following contemporaneous exogeneity assumption would need to hold:

$$E(b_{zt-3}u_{ztq}) = 0 \; \forall z, t, q \tag{3}$$

As ϵ is $i.i.d$ by assumption, this implies in particular that:

$$E(b_{zt-3}\alpha_{zq}) = 0 \; \forall z, t, q \tag{4}$$

As equation 4 states, for consistent estimation all neighbourhood-quarter specific characteristics, α_{zq}, must be uncorrelated with the installed base. This condition is unlikely to hold: firstly, because the model omits all neighbourhood characteristics apart from the installed base and controls for them with the α_{zq}. There are almost surely neighbourhood characteristics that are correlated with both, the dependent variable and the included installed base - omitting them would lead to omitted variable bias (OVB). Secondly, there are diverse endogeneity concerns, such as self-selection and correlated unobservables, that would justify the inclusion of fixed effects even if all relevant observable neighbourhood characteristics were included. It is therefore likely that there exist observable and unobservable neighbourhood and neighbourhood time specific characteristics that are correlated with the installed base *and* the installation rate. This can be summarized as follows:

$$\exists \; \alpha_{zq} \; s.t. \; E(b_{zt-3}\alpha_{zq}) \neq 0 \cap E(y_{zt}\alpha_{zq}) \neq 0$$

If so, then the exogeneity condition in 3 breaks down and POLS results in biased and inconsistent estimates of the installed base effect, β.

5.2 Within-Group Estimator

Since equation 1 includes neighbourhood-quarter effects α_{z_iq}, mean differencing must be performed on the neighbourhood-quarter level to eliminate these effects and estimate the equation by a within-group estimator. More precisely, these are the means taken over all observations of α_t, y_{zt}, b_{zt-3} and ϵ_{zt} within reference group z and quarter q. Let $\overline{\alpha}_q$, $\overline{y}_{zq}$, $\overline{b}_{z3q}$ and $\overline{\epsilon}_{zq}$ be the neighbourhood-quarter means of the time-dummies, the installation rate, the lagged installed base and the idiosyncratic error, respectively. For example, $\overline{b}_{z3q}$ is the mean of the third lag of the installed base, taken over the three months that pertain to quarter q of month t. Mean differencing on the neighbourhood-quarter level then yields:

$$(y_{zt} - \overline{y}_{zq}) = (\alpha_t - \overline{\alpha}_q) + \beta \cdot (b_{zt-3} - \overline{b}_{z3q}) + (\epsilon_{zt} - \overline{\epsilon}_{zq}) \tag{5}$$

The resulting within-group estimator is given by:

$$\hat{\beta}_{WG} = \frac{\sum_z \widetilde{(b_{zt-3} - \overline{b}_{z3q})(y_{zt} - \overline{y}_{zq})}}{\sum_z \widetilde{(b_{zt-3} - \overline{b}_{z3q})^2}} \tag{6}$$

Where $(\widetilde{b_{zt-3}-\bar{b}_{z3q}})$ is the residual from a regression of the mean-differenced installed base on the mean-differenced time-dummies. Further, substituting in the equation for y_{zt} and rearranging yields:

$$(\hat{\beta}_{WG}-\beta)=\frac{\sum_z(\widetilde{b_{zt-3}-\bar{b}_{z3q}})(\epsilon_{zt}-\bar{\epsilon}_{zq})}{\sum_z(\widetilde{b_{zt-3}-\bar{b}_{z3q}})^2} \tag{7}$$

It follows by the Slutsky Theorem, the Continuous Mapping Theorem and the Weak Law of Large Numbers that:

$$\lim_{Z\to\infty}(\hat{\beta}_{WG}-\beta)=\frac{E[(\widetilde{b_{zt-3}-\bar{b}_{z3q}})(\epsilon_{zt}-\bar{\epsilon}_{zq})]}{E[(\widetilde{b_{zt-3}-\bar{b}_{z3q}})^2]}=\frac{A}{B} \tag{8}$$

Where Z refers to the number of neighbourhoods. Since the denominator B is non-zero by assumption, the nominator A must be considered carefully. In contrast to the case illustrated by Narayanan and Nair (2011) that is based on the first lag of the installed base, the neighbourhood-quarter mean of the installed base $\bar{b}_{z3q}$ includes future and past observations of adoptions within quarter $q-1$. On the contrary, the neighbourhood-quarter mean of the errors $\bar{\epsilon}_{zq}$ includes all future and past errors within quarter q.

$$\lim_{Z\to\infty}(\hat{\beta}_{WG}-\beta)=\frac{E[b_{zt-3}\epsilon_{zt}]-E[\bar{b}_{z3q}\epsilon_{zt}]-E[b_{zt-3}\bar{\epsilon}_{zq}]+E[\bar{b}_{z3q}\bar{\epsilon}_{zq}]}{E[(b_{zt-3}-\bar{b}_{z3q})^2]}=\frac{A}{B}=0 \tag{9}$$

Overall, if there is no autocorrelation across quarters, the mean differenced lagged installed base, $(b_{zt-3}-\bar{b}_{z3q})$, is uncorrelated with the mean differenced error, $(\epsilon_{zt}-\bar{\epsilon}_{zq})$, and A converges to zero.[5] More generally, as can easily be shown, for lags of a length larger than 2, i.e. exceeding the length of a quarter, within-estimators aiming to eliminate neighbourhood-quarter fixed effects, can yield consistent estimates.

Proposition 1. *For lags of the installed base that exceed the length of a quarter, i.e. lags larger than 2, mean-differencing on the neighbourhood-quarter level eliminates the neighbourhood quarter effects and allows to consistently estimate the installed base effect by POLS on the mean-differenced equation.*

This contrasts with the model including the *first* lag of the installed base as considered by Narayanan and Nair (2011). They formalise that in the presence of self-selection and correlated unobservables, standard within-group estimators are biased and inconsistent, if there are installed base effects. In their model this is due to the correlation of the mean differenced installed base with the mean differenced errors, a problem relating closely to standard random or fixed effects estimators in dynamic panels. Due to a correlation of the transformed regressor(s) with the transformed errors, random effects GLS and fixed effects estimators (within-group estimators and first difference estimators) are biased and inconsistent. In dynamic panels, estimators as suggested by Anderson and Hsiao (1982) or by Arellano and Bond (1991) are required.

For the within-group estimator presented above, the time-lag of three months is essential for consistency. The estimator is not robust regarding the lag length of the installed base: for

[5] For the full proof of bias and inconsistency in case of the *first* lag of the installed base see Narayanan and Nair (2011). They assume no autocorrelation of the error terms and suggest that the bias is negative.

a lag length of two, $l = 2$, for example, the estimator is inconsistent due to a correlation of the mean-differenced installed base with the mean-differenced error. Hence, as consistency of the within-group estimator is based on relatively demanding conditions, such as a lag length of at least three months, this paper follows a first differencing estimation strategy as proposed by Bollinger and Gillingham (2012). For both estimation strategies, however, the natural time lag between adoption decision and installation is essential for identification and consistent estimation of the social effects.

5.3 First Difference Estimator

Comparable to the case of the within-group estimator, first differencing on the month level does not fully eliminate the neighbourhood-quarter effects. Rather, first differencing of the first month of a quarter yields a residual term of the size of the change of the neighbourhood-quarter effects from one quarter to the next. To eliminate this residual and hence the neighbourhood-quarter effects, the first month of each quarter is dropped after first differencing (see Bollinger and Gillingham, 2012). This leads to equation 10 and POLS on the differenced equation can follow.

$$(y_{zt} - y_{zt-1}) = (\alpha_t - \alpha_{t-1}) + \beta \cdot (b_{zt-3} - b_{zt-4}) + (\epsilon_{zt} - \epsilon_{zt-1}) \tag{10}$$

$$\Delta y_{zt} = \Delta\alpha_t + \beta\Delta b_{zt-3} + \Delta\epsilon_{zt} \tag{11}$$

The required exogeneity condition for consistent estimation of β is hence:

$$E(\Delta b_{t-3}\Delta\epsilon_{zt}) = 0 \tag{12}$$

As $\Delta b_{zt-3} = \Delta \sum_{\tau=1}^{t-3} Y_{zt} = (\sum_{\tau=1}^{t-3} Y_{zt} - \sum_{\tau=1}^{t-4} Y_{zt}) = Y_{zt-3}$, condition 12 can be written as:

$$E(Y_{zt-3}\Delta\epsilon_{zt}) = E(Y_{zt-3}\epsilon_{zt}) - E(Y_{zt-3}\epsilon_{zt-1}) = 0 \tag{13}$$

By construction of equation 1 Y_{zt-3} is not only correlated with its contemporaneous error ϵ_{zt-3}, but also with all previous errors, i.e. $E(Y_{zt-3}\epsilon_{zt-\tau}) \neq 0 \forall\tau \geq 3$. However, if ϵ_{zt-3} is uncorrelated with ϵ_{zt-1}, i.e. if the order of autocorrelation of ϵ_{zt}, is smaller than 2 ($\nu < 2$) the consistency condition given in 12, holds.

Taking this to the example of solar PV technology adoption, the suggested first differencing estimation strategy thus yields consistent estimates, if the lead time between adoption decision and installation is large enough. More specifically, given the natural time lag between adoption decision and installation of three months, the inclusion of the third lag of the installed base allows for a first differencing estimation strategy that yields a consistent estimate of the installed base effect, as long as the order of autocorrelation is smaller than 2, i.e. $\nu < 2$. This argument can be generalized: if the considered lag of the installed base is l, then $E(Y_{zt-l}\epsilon_{zt-\tau}) \neq 0 \forall\tau \geq l$ and consistent estimation is feasible as long as $E(\epsilon_{zt-l}\epsilon_{zt-1}) = 0$, i.e. the order of autocorrelation ν of the error ϵ must be smaller than $(t-1)-(t-l) = l-1$.

Proposition 1. *Let the lag of the installed base be l and let the order of autocorrelation of ϵ be ν. Then, if $\nu + 1 < l$, i.e. the order of autocorrelation ν is smaller than $l - 1$, consistent estimation of the installed base effect is feasible.*

6 Data

This paper focuses on social effects within geographically defined reference groups. Since solar PV panels are visible for passers-by, no social bonding in the sense of friendship, family, culture or religion is required for social effects such as observational learning. This motivates the spatial definition of the reference group.

6.1 Postcodes and Postcode Districts

To highlight the granularity of the analysis, consider the different geographic area definitions in the UK. A postcode district is identified by the first three to four digits of a full UK postcode, e.g. by CB3 in case of the postcode CB3 9DD. Each full UK postcode is divided by a space into two parts: the first part indicates the postcode area and the postcode district, e.g. CB and CB3 respectively. The second part begins with a single digit for the postcode sector within each district, e.g. 9 in case of CB3 9DD. In August 2012, there were 124 postcode areas, 2,987 postcode districts, approximately 11,000 postcode sectors and 1.78 million unit postcodes in the UK (including Channel Islands and Isle of Man)(Royal Mail, 2013).

6.2 Neighbourhood Statistics

For England and Wales the ONS provides Census statistics, such as on tenure for example, for multiple geographical aggregation levels, among others for the postcode district level. There are 2,269 postcode districts in England and Wales. Each postcode district consists, on average, of 10,297 households, with just under 65 percent of these owner-occupied (ONS, 2013). The number of owner-occupied households is of particular relevance for the installation of solar PV panels, as the decision to install a solar PV panel is more likely to be made by households who own their property than by households who rent it.[6] Even if tenants can benefit from the panels as well (e.g. by being provided with some free electricity), it is the landlord who receives the FIT and who decides whether to install the panel. The number of tenants inducing their landlords to install solar PV are assumed negligible for this analysis.

The ONS only provides neighbourhood statistics for England and Wales. Future research, however, could match the neighbourhood statistics of the Scottish Census with the postcode districts, using the pseudo LSOA code and the look up tables provided by the Scottish government.

6.3 Solar PV Installation Data

The econometric analysis in this paper is based on solar PV installation data from April 2010 through March 2013. The dataset is published by Ofgem and updated quarterly. It includes all domestic, commercial and communal micro-generation systems in the UK that have been registered to receive the FIT since its introduction in April 2010 (i.e. systems with up to 50kW capacity). By the end of March 2013 379,531 installations were registered. Besides an individual identifier for each installation, the dataset includes geographic location of the

[6] ONS tenure data provides information about whether a household rents or owns the accommodation that it occupies and, if rented, combines this with information about the type of landlord who owns or manages the accommodation.

installation down to the LSOA level, the corresponding postcode district (i.e. the first four digits of the postcode the LSOA code is associated with), the day at which the installation was commissioned, i.e. completed, as well as the installed capacity (in kW).

Of all 379,531 installations that were registered in April 2013, just under 83 percent were commissioned in England, 7.3 percent in Wales and 6.9 percent in Scotland. Solar PV technology dominates the market for micro-generated energy: 98.6 percent of all micro-generation installations are solar PV systems. Wind contributes only 1.2 percent, micro CHP 0.12 percent and anaerobic digestion and hydro even smaller shares of 0.01 and 0.11 percent. Moreover, the share of domestic installations lies above 96 percent (96.5 of all micro-generation technologies and 96.9 percent of the solar PV installations are domestic installations). As this paper aims to analyse the diffusion of a single technology rather than the choice between different kinds of technologies and focuses on social effects within neighbourhoods, all non-solar and non-domestic PV technology installations are excluded from the analysis. 4,945 solar PV installations were not associated with any locational information on LSOA or postcode district and hence excluded from the analysis, too. Finally, for the reasons mentioned above, the analysis considers solar PV installations in England and Wales, only. The cleaned data set counts 332,216 domestic solar PV installations in England and Wales by the end of March 2013. These are associated with 2,239 postcode districts with an average of 10,427 households each, with on average 63.5 percent of this (6,629) owner-occupied. Only districts with at least one domestic solar PV system commissioned between March 2010 and 2013 are considered, which are striking 98.7 percent of all postcode districts in England and Wales. The 30 postcode districts that had not installed any domestic solar PV panels were mainly districts in the centre of big cities such as Manchester (M1, M17, M50, M60), Liverpool (L2), Birmingham (B1, B2, B3), Leeds (LS1) and London (20 postcode district areas around the Tower Bridge, Westminster, Leicester Square for example). It is thus not at all surprising that there are no domestic solar PV installations in those areas as these non-adopting postcode districts obviously differ significantly in major neighbourhood characteristics.

The main variable of interest is the installed base of solar PV panels on a postcode district level. The installed base, as suggested above, serves as a measure of social effects and refers to the cumulative number of solar PV installations within a reference group z at a particular point in time t. For the remainder of this paper, z refers to the postcode district (interchangeably used with 'zip code') and t refers to the month. More precisely, the installed base, b_{zt} in postcode district z in month t is defined in equation 14. Where Y_{zt} refers to the number of new installations Y in zip code z in month t.

$$b_{zt} = \sum_{\tau=1}^{t} Y_{zt} \tag{14}$$

These variables are derived exploiting the postcode district identifier and the commission date of each installation. The resulting postcode district level panel data set consists of 80,604 data points: for 2,239 zip codes in England and Wales and for each of the 36 months there is an observation for the installed base and the count of new installations .

Figure 1 illustrates the increase in the average installed base from 2.3 solar PV systems within a postcode district in April 2010 to 148.4 domestic solar PV installations by the end of March 2013.

Figure 2 exemplarily illustrates how the installed base varies across CB postcode districts.

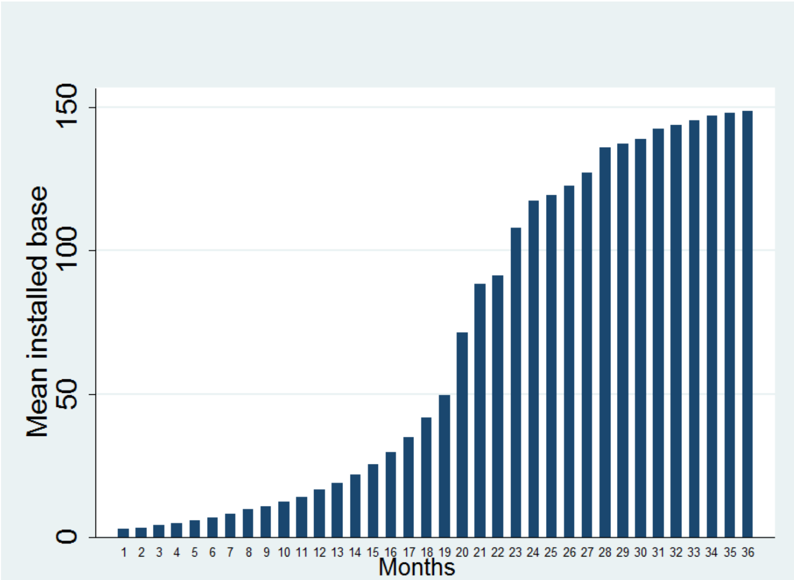

Figure 1: Average installed base in the postcode district (April 2010 - March 2013). $\frac{1}{2,239}\sum_z^{2,239} b_{zt} \forall\, t = 1 \cdots 36$. Source: Ofgem, own calculations.

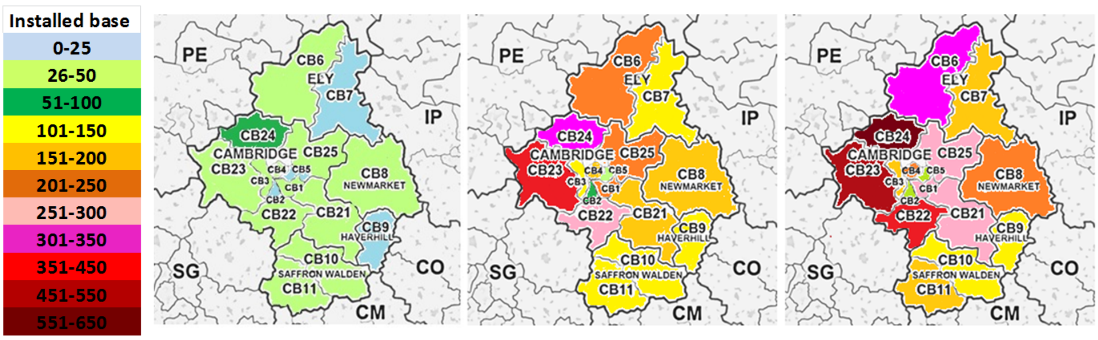

Figure 2: Installed base in CB postcode districts (March 2011 (l), March 2012 (m), March 2013 (r)). Source: Ofgem, ONS, own calculations.

Clustering in the west is clearly visible and increasing in time. Figure 3 plots the average number of monthly installations per owner-occupied household within the 2,239 considered postcode districts for the 36 considered months. The average adoption rate increased from 0.071 installations per 1000 households to 0.192, i.e. almost tripled within the considered two years. Moreover, the adoption rate has never fallen below the rate of March 2010 and peaked in month 20, i.e. in November 2011, with 4.065 installations per 1000 owner-occupied households.[7] The peaks of the adoption rate are most pronounced during the months before cuts of the subsidies came into effect. In December 2011, for example, the UK Government announced rigorous cuts of the FIT, which led to a visible demand response shortly before this change became effective. This reflects the fact that it is the commission date (i.e. the date of completion of the panel) that matters for the level of FIT paid. If a panel is installed before the official date of the subsidy change, the old FIT is paid throughout the 20 years of support. Such kinds of policy announcements must be controlled for, e.g. by the inclusion of time fixed effects in the econometric equation, as they impact the installed base as well as the adoption rate and can confound the estimation of the installed base effect.

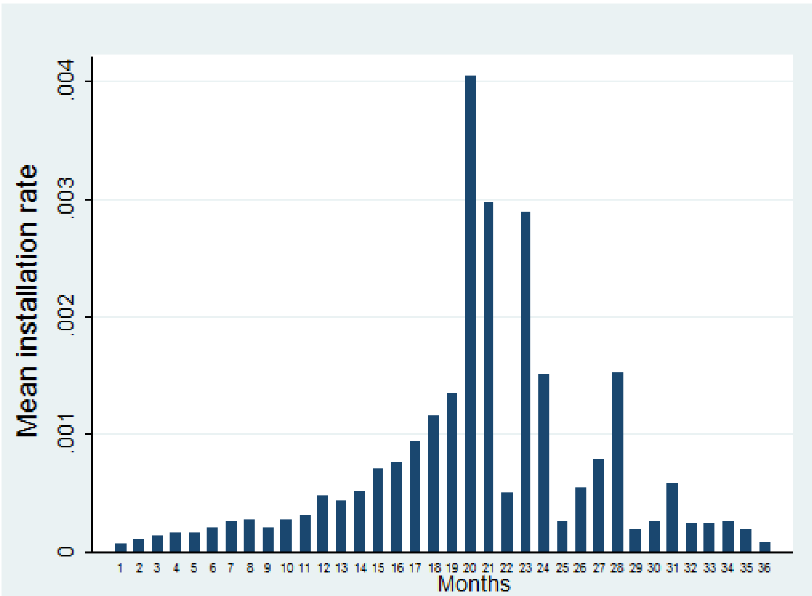

Figure 3: Average adoption rate (April 2010 - March 2013), i.e. $\frac{1}{2,239}\sum_{z=1}^{2,239} y_{zt} \forall t = 1 \ldots 36$. To control for the variation of owner-occupied households across postcode districts, the installation counts are normalised with the number of owner-occupied households in the postcode district. The resulting adoption rate serves as the main outcome variable of interest, y_{zt}. Source: Ofgem, ONS, own calculations.

[7] One could argue that these peaks in adoption rates lead to a decreasing pool of potential adopters within the zip code. However, as multiple installations of a household are possible, this paper assumes the pool of potential adopters to be unaffected by the number of installations. I.e. the number of owner-occupied households within a zip code as reported in the Census 2011 (ONS, 2013) is used for normalisation throughout.

7 Results

Table 1 lists the estimates of the social effects resulting from the suggested pooled OLS ($POLS_{cl}$), within-group (WG_{cl}) (de-meaned on the postcode district-quarter) and first difference estimations (FD_{cl}) based on equation 1. To account for serial correlation within a postcode district over time, the estimates allow for clustering on the postcode district level. The results suggest small, but positive and significant social effects.

The estimates are consistent with the theory: the POLS estimate ($POLS_{cl}$) is strongly downwards biased. This corresponds to an expected omitted variable bias, as the only regressor included in equation 1, apart from the month dummies, is the lagged installed base. The within-group estimator (WG_{cl}) is slightly lower than the first difference estimate (FD_{cl}). As illustrated above, for a lag of at least three, the within-group estimator of the installed base coefficient β can yield consistent estimates. For shorter lag lengths (i.e. shorter lead time between adoption decision and installation) the within-group estimator is inconsistent due to a correlation of the mean-differenced error with the mean-differenced installed base. As consistency of the first difference estimate (FD_{cl}) is conditional on weaker assumptions than those required for a consistent within-group estimate, the first difference estimate is the one of interest.

According to the first difference estimate (see FD_{cl} in table 1) one more solar PV panel in a postcode district increases the number of new installations per owner-occupied households three months later by $7.48e^{-06}$. At the average installation rate of 0.0007, this increase is equivalent to a one percent increase of the average installation rate. This is obviously and as expected a small effect. At the average number of 6,629 owner-occupied households within a postcode district, it implies that one more solar PV panel in the neighbourhood increases the number of new installations in the neighbourhood by 0.05.[8] It would hence require around 20 additional solar PV panels in a postcode district, for social effects alone to be strong enough to cause one further installation within the neighbourhood three months later. The installed base elasticity at the average installed base of 68 installations within a neighbourhood and the average installation rate of 0.0007 is 0.71, implying a rather inelastic demand response. These results illustrate that the social effects as measured by the installed base are small, but exist and can promote adoption. Particularly larger scale projects might hence lead to observational learning from spatially close installations. It is tempting to compare the results to the study of Bollinger and Gillingham (2012), who find an increase of the installation rate of $1.567e^{-06}$ for any additional solar PV installation in a neighbourhood.[9] However, due to differences in the definition of neighbourhoods, the granularity of the data and the model assumptions, such comparisons should be treated with caution.

[8] The average third lag of the installed base was 60.93 and the average number of new adoptions within a neighbourhood was 4.07.

[9] Their neighbourhoods contain on average 4,959 households and are geographically larger.

Table 1: The table shows the estimates of the installed base effect on the installation rate within a postcode district. Estimations are clustered on the postcode district level. The first difference estimate suggests small, but positive and significant social effects. One more solar PV panel in a postcode district increases the number of installations per owner occupied household (that is on average at 0.0007) by significant $7.48e^{-06}$.

Variable	$POLS_{cl}$	WG_{cl}	FD_{cl}
Installed Base (L.3)	1.93e-06***	6.59e-06***	7.48e-06***
	(1.20e-07)	(2.44e-06)	(2.66e-06)
Observations	73,887	73,887	49,258
R-squared	0.145	0.080	0.075

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

8 Robustness Checks

For robustness, the analysis has additionally been performed for model specifications that allow for a non-constant installed base coefficient, for different lags of the installed base and for a distinct geographical level, namely the local authority level.

8.1 Heterogeneous Installed Base Effect

To allow for non-constant social-effects, the model in equation 1 is extended by the squared lagged installed base.

$$y_{zt} = \alpha_t + \beta_1 \cdot b_{zt-3} + \beta_2 \cdot (b_{zt-3})^2 + \underbrace{\alpha_{zq} + \epsilon_{zt}}_{u_{ztq}} \tag{15}$$

Estimating equation 15 by POLS, within-group and first difference strategies, the bias of the POLS and the within-group estimator remain their order and direction. All three estimates suggest, however, a positive installed base effect that decreases with the size of the installed base. The resulting estimates ($POLS_{cl}$, WG_{cl} and FD_{cl}) are listed in table 2. This positive but decreasing effect of the installed base on the installation rate is consistent with the idea of satiation within neighbourhoods. When the installed base is low, the solar PV panels might be particularly attention catching as they are perceived as especially innovative. If so, social spillovers are likely to be most pronounced in the early stages and decrease with the number of installations then.

$$y_{zt} = \alpha_t + \beta_1 \cdot b_{zt-3} + \beta_2 \cdot (D_q \cdot b_{zt-3}) + \underbrace{\alpha_{zq} + \epsilon_{zt}}_{u_{ztq}} \tag{16}$$

Another, comparable, specification allows for a variable installed base effect over time. Equation 16 specifies the model with interactions of 12 quarter dummies, D_q, with the lagged installed base. The results suggest a time-varying installed base effect that overall decreases over time. The coefficient varies throughout the year.[10] Table A.1 in the Appendix

[10]Testing the coefficients of the first differenced equation for equality, the H_0 must be rejected, which is

Table 2: Testing for a heterogeneous installed base effect suggests positive and significant social effects that decreases with the size of the installed base.

Variable	$POLS_{cl}$	WG_{cl}	FD_{cl}
Installed Base (L.3)	4.65e-06***	6.78e-06***	1.45e-05***
	(3.34e-07)	(2.45e-06)	(3.88e-06)
Installed Base Squared (L.3)	-5.56e-09***	-4.83-11***	-1.17e-08***
	(6.73e-10)	(1.45-11)	(3.10e-09)
Observations	73,887	73,887	49,258
R-squared	0.148	0.080	0.075

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

lists the respective coefficients. Interestingly, the social effects are particularly pronounced during times of policy announcements to decrease the FIT. In the seventh quarter (Q_7) for example, i.e. from October to December 2011, the social effects were significantly higher than in the initial months. This could be due to higher media presence that caused higher awareness of the technology and higher perception of the installed panels, for example: in October 2011 the UK government announced a remarkable decrease of the FIT for December 2011. This policy change was then postponed until March 2011. Since all registrations until March 2011 were eligible to the higher FIT, strong and positive social spillovers before the effectivity date do make sense. Towards the end of the considered 12 quarters, the marginal impact of the installed base decreased from the tenth quarter onwards. These are known to be months of relative policy uncertainty, during which adoptions in general were particularly low. Social effects not only decreased, but turned negative as well during this time. These regressions show that the linear model is misspecified and a richer nonlinear model could improve the analysis. However, the main findings regarding the installed base effect are preserved.

8.2 Testing Different Lags

The estimates of social effects are based on the presumption that the lead time between the adoption decision and the completion of the installation is three months. This is obviously just an average. On the one hand the lead may vary across suppliers, on the other hand the lead may vary over time. In particular, during times before the effectiveness dates of the FIT cuts, the time between decision to adopt and actual installation might have varied. It makes thus sense to consider different time-lags of the installed base for robustness.

The coefficients of the contemporaneous and the first lag of the installed base show the expected downwards bias. This is due to the correlation of the first differenced installed base with the first differenced error (see section on identification). The second and third lag have a positive and significant impact on the installation rate. Given the quoted two to three months regarding the lead time between adoption decision and installation, sign and significance of the coefficient of the second and third lag make sense. The coefficient of

consistent with a time varying installed base effect. The test results in an F-statistic of 32.23 and a p-value of 0.000.

the fourth lag is negative, but not significant. This could indicate a 'time threshold' for the installed base effect. It might be, that even if households already perceive and learn from the panels, they are not actually *deciding* that early. These findings might suggest that social effects exist, but are only effective in a rather narrow time window. Table 3 shows the first difference estimates (FD_{cl}) for the specifications with different lags of the installed base.

Table 3: First difference estimates for different lags of the installed base. The results are consistent with the quoted time lead between adoption decision and installation. The estimates suggest that social effects exist, but are only effective in a rather narrow time window.

Variable	FD_{cl}	$FD_{cl}L.1$	$FDb_{cl}L.2$	$FD_{cl}L.3$	$FD_{cl}L.4$
Installed Base	5.75e-05***	-3.63e-05***	3.81e-06**	7.48e-06***	-2.15e-06
	(3.18e-06)	(3.78e-06)	(1.78e-06)	(2.66e-06)	(1.69e-06)
Observations	51,497	51,497	49,258	49,258	47,019
R-squared	0.104	0.082	0.074	0.075	0.074

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

8.3 Installed Base Effect on the Local Authority (LA) Level

For robustness the main regressions are run at the Local Authority (LA) level. This approach tests for social effects within broader defined geographical neighbourhoods than in the case of the postcode districts. The average number of owner-occupied households in the considered local authorities is 42,602, the average installed base 409 solar panel installations. The analysis on the LA level allows for spillovers within larger geographical units and hence across the borders of the postcode districts. This can matter, for instance, if the mobility within a LA is rather high. The estimations based on equation 15, but with z referring to the local authority, yield the results listed in table 4. As expected, the installed base effect

Table 4: Installed base effect in local authoritie. The results suggest that the impact of social effects, as proxied by the installed base, are stronger on a more local level than in broader defined neighbourhoods.

	$POLS_{cl}$	WG_{cl}	FD_{cl}
Installed Base (L.3)	3.14e-07***	1.60e-06*	1.77e-06*
	(5.71e-08)	(8.24e-07)	(9.21e-07)
Observations	11,451	11,451	7,634
R-squared	0.396	0.341	0.346

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

on the LA level is smaller than on a postcode district level. The first difference estimates are 1.77^{e-06}* for the installed base effect within local authorities and 7.63^{e-06}*** for the effect within postcode districts. This is consistent with the argument that social effects are stronger on a more localised level. The coefficient 1.77 implies that at the average number of owner-occupied households within a local authority (42,602), one more domestic solar PV installation increases the number of new installations in the local authority by 0.075. This implies that about 13 additional panels in the LA could induce a new adoption 3 months later.

9 Contextual Factors and the Installed Base Effect

9.1 The Use of ONS Neighbourhood Statistics

To allow for heterogeneous installed base effects across different types of neighbourhoods, interactions of contextual factors with the installed base are incorporated into the model. However, while the installation data is available in panel form, the postcode district neighbourhood characteristics are only available for the cross-section of March 2011, the date of the last Census in England and Wales. It is assumed that any changes of the postcode district characteristics over the considered time frame (2010-2013) are negligible for the installed base effect. This is clearly a strong assumption and simplification, but allows for first insights into heterogeneous installed base effects. Furthermore, most of the ONS neighbourhood characteristics are continuous variables. To allow for more meaningful interpretation of the coefficients of the interaction terms (i.e. the interactions of contextual factors with the installed base), the postcode districts are grouped into quintiles, i.e. categories. Dummy variables indicate whether the respective postcode district characteristic lies above the 60th percentile or not. The estimation strategy is equivalent to the first differencing strategy above.

The results suggest significantly stronger peer effects in neighbourhoods with a high population (in the sense of a population size above the 60th percentile as explained above). In neighbourhoods with a high share of unshared houses social effects are less pronounced, which might result from the higher likelihood to interact with neighbours in the immediate living environment (e.g. in the case of attached houses). Interestingly, relatively affluent (non-deprived) neighbourhoods and neighbourhoods with a high share of high social class households, show a less pronounced installed base effect. As Rogers (1962) suggests, this might result from the fact that those households are early adopters and hence the learning from others is less important. Higher educated neighbourhoods on the other hand show stronger installed base effects than neighbourhoods with on average lower educated populations. In neighbourhoods with a high ratio of economically active people, social effects are relatively low, which can result from the fact that those people spend less time in their neighbourhood. Finally, social effects in neighbourhoods with a high share of people travelling to work by bike or foot are less strong while private transport in a car increases the installed base effect. The estimation results are available upon request to the author.

10 Limitations and Suggestions for Further Research

The presented analysis, which is at a relatively disaggregated level, has several limitations. Still, the results serve as a useful first indicator for the impact of social effects on solar PV technology adoption in England and Wales. The model could be extended in future research.

The first constraint of the model employed in this paper is that social effects are assumed to spread within predefined neighbourhoods only, while spillovers across the neighbourhood borders are ignored. Spatial econometric methods, for example, could be employed to allow for more diverse spatial effects. Another related limitation is the aggregation of the installations to the neighbourhood level. Household level data could improve the analysis and, in particular, shed light on the exact mechanisms of social interactions. Unfortunately, such data is currently not available. Moreover, the approach to measure social effects with the lagged installed base has some drawbacks: as has been emphasised throughout this paper, the quarter lagged specification is important for the estimation method. The lag reflects the assumption that social spillovers such as observational learning only come into effect once the solar PV panel is installed. This is a reasonable and justifiable assumption, but one can imagine some social effects coming into effect before the completion already. For example, there might already be social effects once the solar PV panel supplier van shows up at the neighbour′s house to discuss the further procedure, or there might be word-of-mouth through the neighbours once they have set-up the contract with the supplier. These kind of spillovers are not captured by the lagged installed base. Additionally, friends, family and colleagues might contribute to social spillovers other than those captured by the lagged installed base. The consideration of different lags and the wider definition of neighbourhoods are an attempt to get an idea of the relevance of such effects. A further concern could be that the model obscures any possible partial adjustment processes resulting from inertia in the decision process. An increase in the lagged installed base b_{zt-3} might have an effect on the installation rate in t, but due to adjustment costs, a part of the intended increase might be postponed to the next period. If so, the causal interpretation of the installed base effect comes into question. As an example, there might be households who decide to install in a particular month, but are not able to do so in time due to lacking financial resources. If this is the case, then the observed installation rate, y_{zt}, might be a compromise between the previous installation rate, y_{zt-1}, and the lagged installed base, b_{zt-1}, (and other exogenous variables). To investigate this concern, one could compare the diffusion of solar PV technology with the diffusion process of a technology with a comparable time-lag between adoption decision and installations but of which observability of the installation per se is not a main feature. An example could be the installation of domestic internet connections in the early days of this technology. In summary, if there is partial adjustment in the application of solar PV technology adoption, the estimated coefficients of the lagged installed base might be biased and likely to indicate correlation rather than causation. A final concern regards the specification as a linear probability model. The regression results show that although the main findings regarding the installed base effect are preserved, the linear model is misspecified and a richer non-linear model could improve the analysis. Future research could test the model against a dynamic specification that takes some kind of dynamic adjustment into account. Also, it could be considered to use transform the outcome variable to $F^{-1}(y_{zt}) \in \mathbb{R}$ where F is a cdf that potentially could have heavy tails like a Cauchy distribution.

11 Conclusion

The main research question in this paper is, whether the installation rate of solar PV technology is affected by social effects as measured by the installed base in the immediate local environment. The installed base thereby refers to the cumulative number of solar PV installations within a neighbourhood by the end of a particular month and serves as a measure of social effects from spatially close households. The analysis is based on installation data that has been collected by Ofgem since the introduction of the FIT in April 2010. The econometric panel data model specifies the postcode district-month as the smallest unit of observation. Besides the lagged installed base and month dummies, the main panel equation includes time-varying fixed effects to account for potential self-selection into neighbourhoods and for correlated unobservables that are constant within neighbourhood, but vary over time (i.e. across quarters). Exploiting the time lag between adoption decision and installation, a first difference estimate yields unbiased and consistent estimates of the social effects of interest. Further model specifications allow for a time-varying installed base effect and consider different lags of the installed base as well as different outcome variables and different geographical areas for robustness. In a last specification differences of the social effects across distinct groups of the population are analysed.

The results suggest small, but positive and significant social effects: one more solar PV panel in a postcode district increases the number of new adoptions per owner occupied households in a given month by $7.48e^{-06}$. At the average installation rate within the neighbourhoods, this is equivalent to a one percent increase in the rate to install a solar PV panel. At the average number of 6,629 owner-occupied households within a postcode district, it implies that one more solar PV panel in the neighbourhood increases the number of new installations in the neighbourhood by 0.05. This is obviously and as expected a very small effect. It would require around 20 additional solar panels in a postcode district, for social effects alone to be strong enough to cause one further installation within the neighbourhood. The installed base elasticity at the average installed base of 68 installations within a neighbourhood and the average installation rate of 0.0007 is 0.71. These results illustrate that the social effects as measured by the installed base are small, but exist and can promote adoption. Especially community projects that involve a high number of installations could hence promote diffusion.

The social effects vary across months and overall diminish over time. Moreover, social spillovers on the postcode district level are found to be stronger than on a higher geographical level, the local authority level. Remarkably, relatively affluent (non-deprived) neighbourhoods show a less pronounced installed base effect. This might result from the fact that those households are early adopters and hence the learning from others is less important. Higher educated neighbourhoods on the other hand show stronger installed base effects than neighbourhoods with on average lower educated populations.

This paper contributes to previous literature in performing the first econometric analysis of the diffusion of solar PV technology within the UK. In particular, it delivers empirical evidence, in how far the adoption behaviour of others drives diffusion. The analysis is based on a remarkably recent and granular solar PV installation dataset of the UK. The results can be exploited for targeted marketing and resource allocations for the stimulation of future adoption. For example, increasing the visibility of the panels could increase the rate of adoption and the use of demonstration sites could have positive effects on the adoption of green technologies.

Nevertheless, the analysis has its limitations. Firstly, social effects are assumed to spread within neighbourhoods as defined by the postcode districts or local authorities, only, while spillovers across the neighbourhood borders are ignored. Spatial econometric methods, for example, could be employed to allow for more diverse spillover effects. Another limitation is the aggregation to the neighbourhood level. Future research should make use of household level covariate data to further analyse the mechanisms underlying the adoption behaviour and thus the installed base effect. Lastly, if there is inertia in the decision process, the consideration of a partial adjustment process in the model might be useful. However, overall, this paper delivers a useful first highly disaggregated analysis of the impact of social effects on solar PV adoption, that can be extended in future research.

References

BANDIERA, O. AND I. RASUL (2002): "Social Networks and Technology Adoption in Northern Mozambique," STICERD - Development Economics Papers 35, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

BASS, F. M. (1969): "A New Product Growth for Model Consumer Durables," *Management Science*, 15, 215–227.

BEISE, M. (2004): "The International Adoption of Photovoltaic Energy Conversion Is Japan a Lead Market?" Discussion Paper Series 153, Research Institute for Economics & Business Administration, Kobe University.

BOLLINGER, B. AND K. GILLINGHAM (2012): "Peer Effects in the Diffusion of Solar Photovoltaic Panels." *Marketing Science*, 31, 900–912.

DUFLO, E. AND E. SAEZ (2002): "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," NBER Working Papers 8885, National Bureau of Economic Research, Inc.

GUIDOLIN, M. AND C. MORTARINO (2010): "Cross-country diffusion of photovoltaic systems: Modelling choices and forecasts for national adoption patterns," *Technological Forecasting and Social Change*, 77, 279–296.

MANCHANDA, P., Y. XIE, AND N. YOUN (2008): "The Role of Targeted Communication and Contagion in Product Adoption." *Marketing Science*, 27, 961–976.

MANSKI, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60, pp. 531–542.

NARAYANAN, S. AND H. S. NAIR (2011): "Estimating Causal Installed-Base Effects: A Bias-Correction Approach," Research papers, Stanford University, Graduate School of Business.

RODE, J. AND A. WEBER (2011): "Knowledge Does Not Fall Far from the Tree - A Case Study on the Diffusion of Solar Cells in Germany," ERSA conference papers ersa11p497, European Regional Science Association.

ROGERS, E. M. (1962): *Diffusion of Innovations.*, Glencoe: Free Press.

RYNIKIEWICZ, C. (2012): "Gouvernance participative et choix energetiques : initiatives emergentes et institutions necessaires pour des territoires viables," *CNRS*.

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

SNAPE, R. AND C. RYNIKIEWICZ (2012): "Peer effect and social learning in micro-generation adoption and urban smarter grids development?" *Network Industries Quarterly, 14 (2&3).*, ISSN ISSN 1662-6176, pp. 24–27.

Watson, J. and P. Devine-Wright (2011): "Centralisation, decentralisation and the scales in between." *The Future of Electricity Demand: Customers, Citizens and Loads.*, Cambridge University Press, pp. 542–577.

Wstenhagen, R. and M. Bilharz (2006): "Green energy market development in Germany: effective public policy and emerging customer demand," *Energy Policy*, 34, 1681 – 1696.

Zhang, Y., J. Song, and S. Hamori (2011): "Impact of subsidy policies on diffusion of photovoltaic power generation," *Energy Policy*, 39, 1958 – 1964.

A Appendix

Table A.1: Time-Varying Installed Base Effect (interactions with quarter dummies). The results suggest a time-varying installed base effect. While social effects appear to increase initially, they decrease towards the end of the considered period and eventually even get negative

Variable	FD_{cl}
Installed Base (L.3)	-2.32e-05**
	(9.12e-06)
Q3 x Installed Base (L.3)	-1.46e-06
	(9.99e-06)
Q4 x Installed Base (L.3)	2.42e-05*
	(1.28e-05)
Q5 x Installed Base (L.3)	3.21e-05***
	(9.04e-06)
Q6 x Installed Base (L.3)	4.96e-05***
	(9.55e-06)
Q7 x Installed Base (L.3)	4.81e-05***
	(1.09e-05)
Q8 x Installed Base (L.3)	4.87e-05***
	(9.11e-06)
Q9 x Installed Base (L.3)	2.78e-05***
	(9.09e-06)
Q10 x Installed Base (L.3)	-3.39e-05***
	(1.02e-05)
Q11 x Installed Base (L.3)	-1.01e-05
	(1.12e-05)
Q12 x Installed Base (L.3)	-1.53e-05
	(1.12e-05)
Constant	6.19e-05***
	(1.20e-05)
Observations	49,258
R-squared	0.007

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1