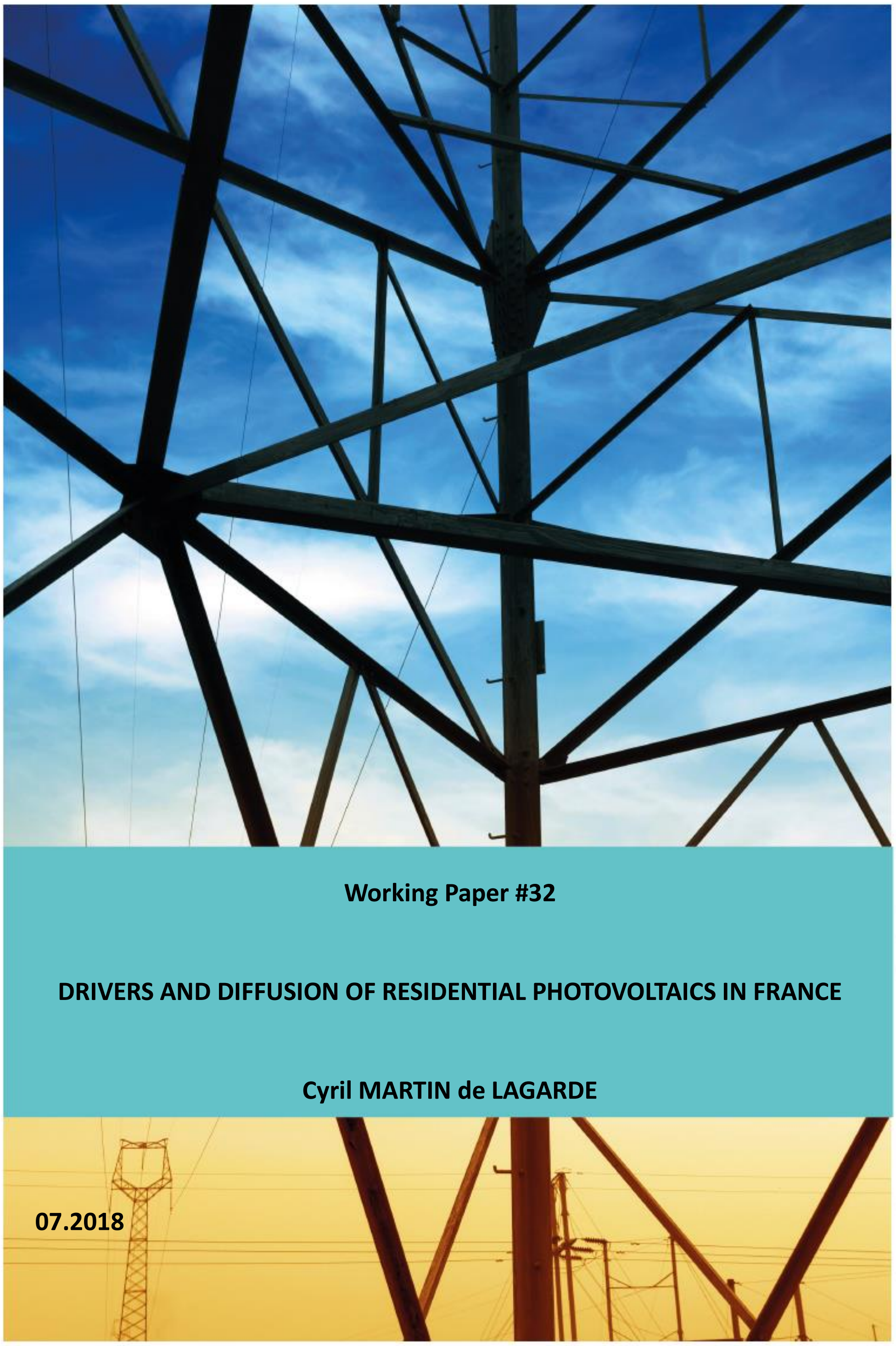
CHAIRE EUROPEAN ELECTRICITY MARKETS
Fondation Paris-Dauphine



Working Paper #32

DRIVERS AND DIFFUSION OF RESIDENTIAL PHOTOVOLTAICS IN FRANCE

Cyril MARTIN de LAGARDE

07.2018

DAUPHINE UNIVERSITÉ PARIS

Chaire de recherche soutenue par

RTE EDF EPEXSPOT CELEST

DRIVERS AND DIFFUSION OF RESIDENTIAL PHOTOVOLTAICS IN FRANCE

Cyril Martin de Lagarde*[1,2,3]

[1]Université Paris-Dauphine, PSL Research University
[2]Ecole des Ponts ParisTech
[3]IFPEN

Abstract

This paper analyses the diffusion of residential solar panels in France, and the impact of financial incentives (feed-in tariffs and local subsidies) on this dynamics. We use a unique database provided by Enedis, the main French DSO, giving the number of connection requests for 33,842 municipalities and 31 quarters, from the end of 2008 to mid-2016. Using solar irradiance, panel system costs, and national and local subsidies, we compute an internal rate of return per municipality and per quarter, which serves as an indicator for projects' profitability. Due to the high number of zero-installation data points, we adopt a two-stage ("hurdle") methodology. We first model the probability of having at least one installation, and then the (strictly positive) number of installations, the vast majority of which are 3-kW panels. Controlling for individual characteristics of the municipalities, we find that financial incentives have had a positive and significant effect on both the probability and the number of adoptions. Furthermore, we show that the diffusion process exhibits "epidemic" and "stock" effects, which are consistent with the "S"-shaped diffusion curve observed at the national and regional levels. Considering only epidemic effects, we show that an additional past installation in a city has the same effect as a one-point increase of the IRR, that is, an increase of the odds of installing at least one solar panel by roughly 10%. Hence, better informing households could help promote renewables at a lower cost. This could be done for example through more implication of citizens at the local level.

Keywords: Solar energy; feed-in tariffs; subsidies; diffusion; hurdle model

Acknowledgment: This paper has benefited from the support of the Chaire European Electricity Markets (CEEM) of the University Paris-Dauphine under the aegis of the Foundation Paris-Dauphine, supported by RTE, EDF, EPEX Spot and CELEST.

I also wish to thank Enedis for providing the data used in this article, and in particular Christophe Bonnery, François Dispot and Hervé Lextrait. Finally, I would like to thank Anna Creti, Jan Horst Keppler, Frédéric Lantz and Roland Ortt for their useful remarks.

Disclaimer: The views and opinions expressed in this article are those of the authors and do not necessarily reflect those of the partners of the CEEM.

*cyrildelagarde@gmail.com

I. INTRODUCTION

In the fight against global warming, many governments have chosen to develop renewable energy sources (RES). Indeed, these should hopefully help reduce greenhouse gas emissions associated with energy production, if they can be substituted to fossil fuels. In the electricity sector, renewables have been developing at high pace in many developed countries, mainly thanks to public subsidies. If these are often justified by the necessity to reduce CO_2 emissions, they also aim at internalising the learning effect (or "learning-by-doing"), i.e. the decrease in costs associated with the development of the technology. Indeed, without subsidies, RES would be installed in too little quantities as their cost would be too high compared to conventional power plants. In the long run, RES supporters expect that they will be competitive with other means of production.

A rather common way of subsidising electric renewables has been to use feed-in tariffs (FITs), that guarantee a price ex-ante for each unit of energy (e.g. kilowatt-hour - kWh) produced over the duration of the contract (usually fifteen to twenty years). This remuneration does not depend on the time at which the electricity is produced, and hence on the spot market price. As many other European countries such as Germany (who was a pioneer for RES in Europe), France has had such FIT for over a decade, before switching to ex-post feed-in premiums (FIPs) in 2015 for most of the new installations. However, the cost of these subsidies is passed on to consumers through taxes in the electricity bill[1]. Furthermore, residential solar panels have been subsidised at the national level through other instruments such as reduced VAT or tax credits, and several regions, departments, or cities have decided to subsidise them as well. These other subsidies are on the contrary borne by taxpayers.

Distributed[2] production units such as small-scale residential photovoltaic (PV) solar panels play an increasing role in this energy transition. Indeed, in France more than 95% of newly installed renewable production is connected to the distribution grid, and household solar panels represent the most widespread technology in number of installations (but not in capacity). However, the distribution network was originally designed to deliver electricity to end-users, which were almost only consumers. Now that decentralised producers and "prosumers" (i.e. producing consumers) massively invest in RES, many reinforcements need to be made at the distribution level in order to adapt to this whole new environment. The transmission network also needs to be reinforced as RES generation is variable, which calls for more global system integration and balancing. The cost of these investments is also borne by consumers, through the network tariff. Therefore, it is crucial for both policy makers, distribution and transmission system operators (DSO and TSO) to understand the dynamics of adoption of RES, and in particular solar panels, in order to limit unnecessary costs and inefficiencies through the design of efficient subsidies, network tariffs and electricity pricing methodologies.

Although electric renewables would probably not have emerged so quickly without subsidies, financial issues are not expected to be the only drivers of renewable development. Indeed, considering them as relatively new products, their diffusion pattern is also likely to follow an "S"-shaped curve at the aggregate level, as famously modelled by Bass (1969) for durable

[1]The electricity bill of a French residential customer is almost equally divided between the cost of energy, the network tariff, and taxes, according to the French ministry in charge of energy (MTES, 2016). In 2016, support mechanisms for electric renewables accounted for approximately a quarter of the taxes, according to the French energy regulator (CRE, 2016).

[2]There are several definitions of distributed (or decentralised), but the most widespread is to call distributed all devices (panels, batteries...) that are connected to the distribution network. In France, the limit between electricity transmission and distribution is set at 50 kV, but this limit can differ between countries.

consumer goods. The diffusion process is hence expected to benefit from an "epidemic" or "contagion" effect (word-of-mouth): the most the technology spreads, the most is becomes known and accepted, and adopted in return. It might also exhibit a "stock" effect, i.e. a decrease of the epidemic effect in the long run, due to the finite size of the market[3]. However, to the best of our knowledge, the impact of subsidies on the diffusion of solar panels (or other products) has not been yet studied at a very local scale.

Thus, we model the demand for solar panels by households, using a unique data set of all connection requests for small (< 3 kW) solar panels received by Enedis, the DSO for 95% of distributed customers and more than 95% of installed PV projects in France. The contribution of this article is threefold. First, we show that households act in a rather rational way. Indeed, they tend to maximise their installed capacity, and they strongly react to quarterly changes in the FIT. Second, using public data for national and local subsidies, solar irradiance and panel costs, we are able to compute an internal rate of return (IRR) for each city whose distribution network is managed by Enedis (33,842) over 31 quarters, from the end of 2008 to mid-2016. We use this variable to measure the impact of public subsidies on the number of installations. In addition, we control for individual (municipal) characteristics such as population or the proportion of houses, which also affect the number of installations in a given city. Therefore, we are able to measure the effect of variables that are often omitted or averaged in large-scale studies, despite their strong heterogeneity. Third, we are able to analyse the "intrinsic" diffusion process by adapting the famous Bass (1969) model at a very fine scale, using data at the municipality and quarterly levels. This requires to take into account the very important number of zeros in the data set. We manage to do this using the two-part "hurdle" model of Mullahy (1986), which we apply to the number of installations, as most installations have the same nominal capacity of 3 kW. In the end, we show that an additional past installation and a one-percentage point increase of the IRR have a similar effect, which is an increase the odds of having an installation by almost 10%. In addition, the IRR also has a positive effect on the number of installed panels in a given city during a quarter. This result calls for more "advertising" on RES at the local level, as information is essential in a diffusion process. Given the very high value of the observed subsidies and IRR, increased information is likely to help promote renewables at a lower cost. This may be encouraged by a better implication of local residents in RES projects, for example through shared ownership.

The remainder of the paper is organised as follow: section II first presents a literature review on the subject. Then, we describe the data on connection requests in section III. Afterwards, we present the rest of the data and the computation of the IRR in section IV. In section V, we detail our modelling strategy of the diffusion process and the hurdle model. The results are then shown and discussed in section VI. Section VII concludes and presents future research avenues.

II. LITERATURE REVIEW

Many articles have analysed the development of RES in several countries or regions. Although support schemes have been studied and compared theoretically (Marschinski and Quirion, 2014; Mir-Artigues and Río, 2014; Boomsma and Linnerud, 2015; Bauner and Crago, 2015) and empirically (Kilinc-Ata, 2016; Aguirre and Ibikunle, 2014), we will focus on empirical studies in the rest of the review. In particular, the impact of support schemes on the development of

[3] A more detailed description of the different diffusion effects can be found in Karshenas and Stoneman (1993), along with an empirical application to the computer numerically controlled control machine tools market.

RES has been widely discussed: feed-in tariffs, as one of the most widespread support scheme, but other mechanisms such as renewable portfolio standards, green certificates or carbon taxes have also been analysed. Most studies find a positive and significant effect of these on the development of solar PV, while there is less evidence concerning wind power. For example, Zhang et al. (2011) investigate the impact of regional subsidies on the diffusion of solar PV in Japan using panel data at the municipal level. Their results indicate that regional policies tend to promote the adoption of solar panels, while installation costs have a negative impact, and housing investment and environmental awareness a positive one. Jenner et al. (2013) use panel data in 26 European countries for the period 1992-2008 and construct a return-on-investment indicator for FIT strength in order to take into account in particular variability in tariff level and contact duration between countries. They conclude that FITs have driven solar PV but find no evidence concerning wind energy, and also argue that the influence of FITs highly depends on several factors such as production costs or electricity prices. Dijkgraaf et al. (2018) perform a similar analysis in 30 OECD countries for the period 1990-2011 and find a positive effect of FITs on the adoption of solar panels per capita. They also show that some features of the contract, such as the duration of the absence or presence of a cap have an impact on the adoption as well. Many authors have also studied other factors such as political ones (Cadoret and Padovano, 2016; Strunz et al., 2016), location, socio-economic drivers, local subsidies and spatial spillovers (Schmidt et al., 2013; Ek et al., 2013; Müller and Rode, 2013; Graziano and Gillingham, 2015; Schaffer and Brun, 2015; Balta-Ozkan et al., 2015; Schaffer and Düvelmeyer, 2016; Carfora et al., 2017; Allan and McIntyre, 2017).

In order to disentangle various effects influencing the development of residential PV, we also need to take into account the "intrinsic" diffusion process of these relatively new technologies. Indeed, several studies have modelled the deployment of electric renewables as following an "S-curve" and/or using a diffusion model in the line of Bass (1969): Schilling and Esmundo (2009), Guidolin and Mortarino (2010). A review of several models incorporating external influence, such as price or advertising, is provided in Radas (2006). However, these are often highly non-linear, which can make their estimation difficult. In order to study RES development, most authors have used linearised versions of diffusion models. For example, Liu and Wei (2016) derive a linear regression equation from a logistic growth function defined as in Benthem et al. (2008). They show that the development of wind power in China has been driven by financial incentives as well as by epidemic effects, using panel data from clean development mechanism projects in China. Using data at the zip code and street levels and a linear probability panel model, Bollinger and Gillingham (2012) show that peer effects have had a strong impact on PV adoption in California.

A drawback of these studies, and of diffusion models in general, is that they require relatively "smooth" data to be fit, as they predict a strictly positive number of sales (except in the long run: when the market is fully reached, sales drop to zero). This issue challenges their applicability to low-level data such as ours, where the number of installations is very low and often equal to zero. Indeed, we want to study the impact of both national and local subsidies, at a quarterly time step. Hence, our data contains many zeros (it is said to be "zero-inflated"). This peculiarity automatically rules out the use of a linear regression model, which would predict many negative values, and in particular not enough zeros. Fortunately, there exist several possibilities to deal with this kind of data, and the interested reader can have a look for example at Min and Agresti (2002) for a quick survey. Such models can be decomposed in several categories, depending on the data being (semi-)continuous or discrete, on the model proposing a single underlying mechanism for the creation of zeros or two separate mechanisms for zero and non-zero values, on the type of distribution (single, compound, finite mixture...), etc.

One of the first and most famous of these is the Tobit model (Tobin, 1958), which assumes a unique Gaussian data generating process for semi-continuous data. It has been used for example by Hitaj (2013) to analyse the drivers of wind development in the United States. However, the Tobit model is more adapted to censored data, as the same process determines whether the dependent variable is zero or is value when it is positive. Furthermore, it assumes semi-continuous data, which would not be adapted to small-scale PV projects. Indeed, we we will see in section III why using count data (i.e. the number of installations) is more adequate. For instance, Kwan (2012) use a zero-inflated negative binomial (ZINB) model to analyse the number of PV installations at the ZIP code level in the United States. Although this enables to model accurately to large number of zeros, the interpretation can be difficult, as zeros are produced twice: first by a binomial regression, and then by a count model. In our case, we apply the two-part "hurdle" model of Mullahy (1986) to the number of installations. Unlike zero-inflated models, hurdle models deal separately with zeros and non-zero values, hence leading to an easier interpretation. We will further motivate this choice in section V, as it is due to some characteristics of the data that will be presented in the following section.

III. CONNECTION REQUESTS

1. Overview

In order to study the dynamics of household PV adoptions, we use a unique database provided by Enedis, which is the distribution system operator (DSO) for 95% of French clients. We have received from Enedis the register of all received connection requests for PV projects until mid-2016[4]. We restricted ourselves to the small-scale (< 3 kW) segment, as it mostly corresponds to connection requests from households. It also has the advantage of having benefited from relatively stable subsidies, compared to other segments. Indeed, other categories have seen their capacity boundaries changed several times and tariffs have sometimes been dependent on the type of building or building integration, for which we do not have information. In particular, this segment was the only one that was not concerned by the three-month moratorium on feed-in tariffs from December 9 2010 to March 9 2011.

The original data being at the municipal and daily levels, we had to aggregate it in order to avoid having too many zeros (we still have a lot of zeros at the end of the aggregation process). Due to the high heterogeneity between municipalities in terms of solar irradiance, local subsidies, but also population characteristics, we find preferable to keep the data at this scale. Unfortunately, we are not able to consider household heterogeneity within a given city, due to the absence of personal data. In particular, the data set has been anonymised. Hence, our "individuals" will be municipalities rather than households. In the time domain, quarterly aggregation appears as a rather natural choice for small-scale solar PV, as FITs have been changing quarterly since 2011.

Hence, we end up with 31 quarters of observations for 33,842 municipalities, which represents 1,049,102 observations, only 62,909 (6%) of which are non-zeros. Mid-2016, there are still 13,011 municipalities (38%) with no domestic solar panel at all. Table 1 below displays some descriptive statistics for the demand of connection requests, in count and in capacity (kW), as well as for the cumulative data for the last quarter of the sample. Due to the very high number of zeros, we present statistics for the non-truncated data and for the strictly positive one, as

[4] In France, all generating capacity under 17 MW must be connected to the distribution network.

these will be modelled separately. These statistics show that the demand and count data are overdispersed (i.e. the variance is bigger than the mean), while the strictly positive count data is underdispersed. However, although most of the time there is only one connection request, this number can rise up to 18.

Statistic	N	Mean	St. Dev.	Min	Median	Max
Demand (kW)	1,049,102	0.213	0.999	0	0	51.6
Demand >0 (kW)	62,909	3.557	2.178	0.100	3	51.6
Cumulative demand (2016 Q2)	33,842	6.611	12.605	0	3	456
Count	1,049,102	0.076	0.356	0	0	18
Count >0	62,909	1.267	0.774	1	1	18
Cumulative count (2016 Q2)	33,842	2.356	4.516	0	1	162

Table 1: Descriptive statistics of the Enedis data set

To get an idea of the diffusion process and its heterogeneity, figures A.1 and B.1 in appendix A and B show the quarterly demand and cumulative demand for PV projects per region[5]. These graphs show that the dynamics is already highly heterogeneous between regions, but the ”S”-shaped curve can still be observed for all of them. It would also be observed at the national scale, but not at the scale of most municipalities, as there are usually very few installations, or not at all. Finally, the left map in figure 1 shows the regional cumulative capacity at the end of June 2016, while the right map presents the mean annual optimal irradiance (i.e. for the best possible angle). Quite interestingly, these two maps seem almost uncorrelated, or even negatively correlated. For example, Nord-Pas-de-Calais, which is the northernmost region, and Lorraine, a northeastern region, have very high installed capacities compared to their received irradiance.

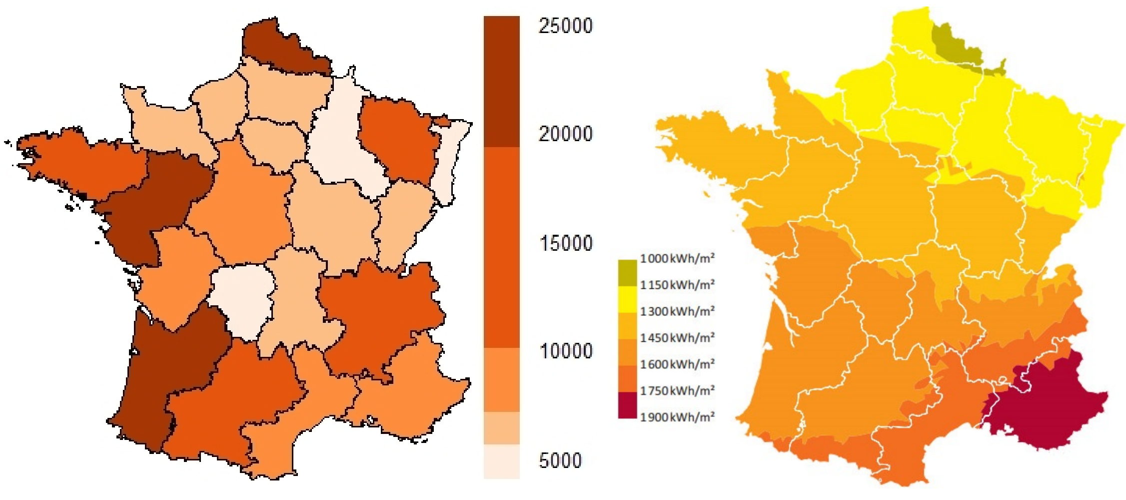


Figure 1: Left: Regional cumulative PV capacity (kW) of projects of less than 3 kW, mid-2016. Right: Annual optimal irradiance (source: PVGIS, JRC).

[5]Although there has been a change in the number of regions in January 2016 (from 20 to 12 in metropolitan France, Corsica excluded - Enedis is not the DSO there), we have decided to keep the initial regions as there have been several regional subsidies before this reform.

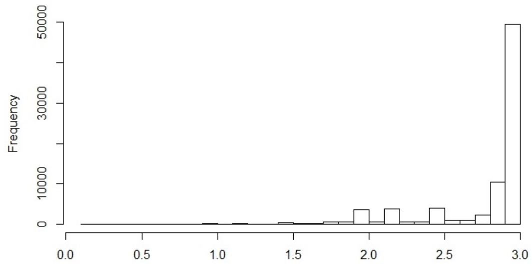


Figure 2: Histogram of individual capacity demands (kW)

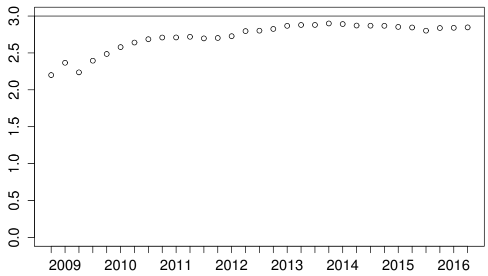


Figure 3: Average capacity per connection request (kW)

2. Analysis of households behaviour

When modelling the effect of financial incentives on the adoption of solar panels by households, we hope to find that these latter are rather rational, in that they react positively to subsidies. Without any modelling, the raw data can already give some confirmation that the behaviour of the households is somewhat rational.

First, assuming that a PV project is profitable for the investor, and considering the fact that FITs are designed by segments of power, agents should maximise their installed capacity while staying under the tariff change power threshold. We expect only physical constraints such as the available area on the roof to limit the installed capacity. As a matter of fact, most roofs are large enough for 3-kW panels (approximately 10 m^2 per kW are required). Figure 2 shows that most installations are indeed 3-kW ones. Furthermore, figure 3 illustrates that the average capacity per connection request has slightly increased towards 3 kW over time. This could be the sign of a "learning" process, of more confidence of investors, or maybe of a standardisation of panels. Besides, this characteristic of the installations will allow us to use the number of installations rather than the demand in kW, as we will explain in section V.

Another feature of FITs for solar PV is that they have been revised quarterly since the end of the moratorium on March 10, 2011 (which coincides with the first observed peak in figure 4)[6]. Indeed, at the end of each quarter the tariff is adjusted by a coefficient decreasing with the accumulated capacity. As a consequence, we expect connection requests to rise at the end of each quarter, in what we could call a "deadline effect", as agents would rather hurry in order to benefit from the highest possible FIT before it decreases. Figure 4 shows that this is indeed the case, as the spikes correspond to the end of each quarter. Thus, it is preferable that the minimum temporal aggregation be at least at the quarterly level, to avoid the difficulty to model this behaviour. Also, this comforts us in thinking that agents are price-sensitive, hence legitimating the question of quantifying the role of incentives in the diffusion process.

[6]There has also been a change at the beginning of February 2013, with a huge decrease in the number of FIT categories, but it did not concern small PV installations.

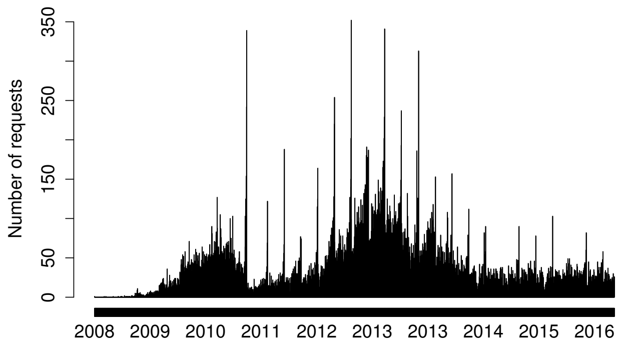


Figure 4: Illustration of the "deadline effect"

IV. ADDITIONAL DATA

In order to analyse the diffusion of residential PV, we need data on costs and subsidies for such installations, but also control variables, as the number of installations is likely to depend on socio-economic factors.

1. Costs and feed-in tariffs

While FIT data is publicly published on official websites, data for system costs (panels, power inverter[7], installation...) can be more difficult to find. Given the heterogeneity of sources, we have chosen annual data[8] provided by the French Environment and Energy Management Agency (ADEME). Figure 5 below shows the evolution of FITs and gross system costs from the end of 2008 to mid-2016. The graph shows that despite an overall decreasing trend for both costs and FITs, the latter have remained relatively constant until the end of 2010, while cost were declining. This explains at least partly the "bubble" observed at the end of 2010, which led to the three-month moratorium, and later an important decrease and a quarterly revision of FITs.

In addition to the FIT, the State has also decided to subsidise RES in several ways. Indeed, some installations give right to tax credits under certain conditions (e.g. income, type of installation, etc.). Also, VAT is reduced to 10% (instead of 20%) for solar panels, and to 5% for projects eligible to tax credits. Finally, the National agency for housing (Agence national de l'habitat) can also subsidise RES installations in case of renovation works for low-income households. Unfortunately, we are not able to determine whether projects were eligible to some of these subsidies or not. Thus, we will not consider them, but we believe they have had a smaller impact on the diffusion than FITs and local subsidies.

2. Local subsidies

As already mentioned, many local entities have decided to subsidise RES, and in particular small-scale PV, in addition to the existing national subsidies. In France, the main sub-national

[7]Electric device used to change direct current (DC) produced by the panel, to alternative current (AC).
[8]In order to get quarterly data, we used linear interpolation.

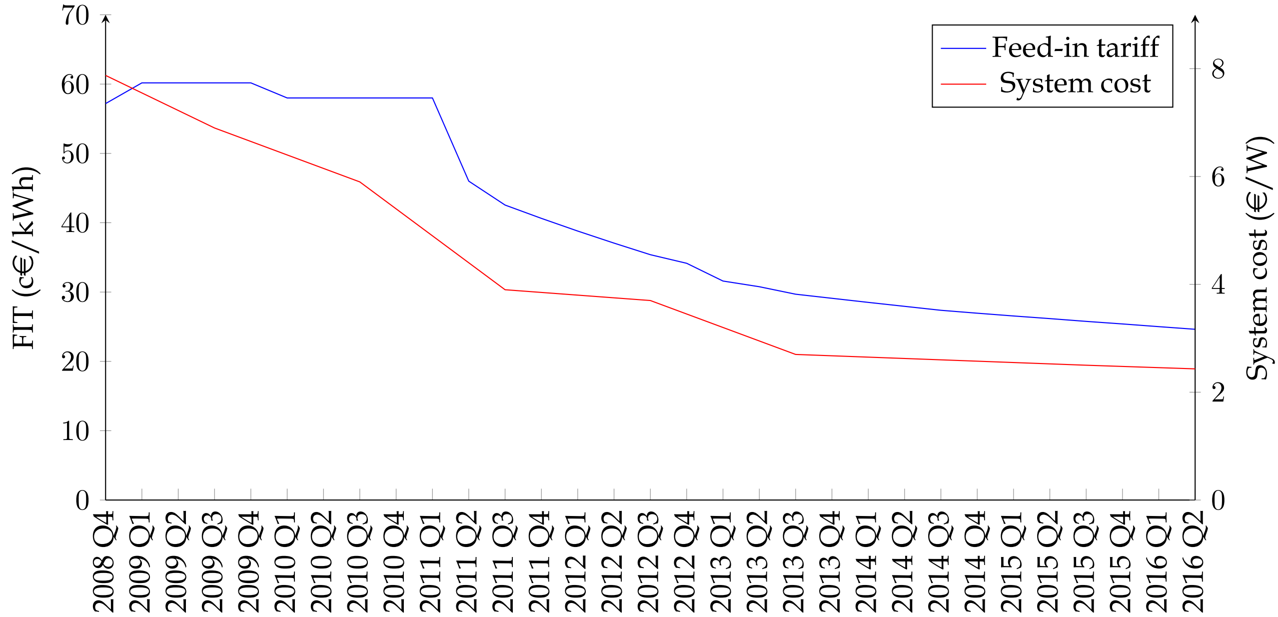


Figure 5: Evolution of feed-in tariffs and system costs for small-scale PV (< 3 kW). Sources: ADEME (costs) and Photovoltaique.info (tariffs).

administrative entities are the 21 regions, 94 departments and roughly 36,000 communes (municipalities) of Metropolitan France (Corsica excluded). Several have decided to subsidise RES in their own way, for example through: a fixed or proportional subsidy for the entire system or for the installation only (usually with a certified installer only); a subsidy per kilowatt-peak (sometimes with an upper limit); a premium on the FIT for a few years; etc. Due to the heterogeneity of local subsidies, and as most of panels have a capacity of 3 kW, we consider for simplicity reasons that all projects have this capacity, so that we can compare them. A summary of these subsidies is displayed in table 2 below.

Entity	N. obs. (entity level)	N. obs. (municipality level)	Mean	St. Dev.	Min	Max
Regions	651	1,049,102	368	970	0	6,000
Regions (>0)	120	183,578	1,998	1,364	300	6,000
Departments	2,914	1,049,102	53.9	384	0	6,396
Departments (>0)	97	36,016	1,619	1,384	300	6,396
Municipalities	1,049,102	1,049,102	5.78	92.1	0	3,690
Municipalities (>0)	6,935	6,935	874	723	75	3,690

Table 2: Descriptive statistics of local subsidies (in euros) for a 3-kW PV project, from 2008 Q1 to 2016 Q2, for several administrative levels.

As can be seen from the table, positive regional subsidies concern 17.5% of all observations, with relatively high mean an maximum values (the price of 3-kW modules have varied between roughly 24,000€and 7,300€). Subsidies at the department and municipality levels concern much less observations (3.4% and 0.66%, respectively), but attained values are rather high as well.

3. Internal rate of return

3.1. Motivation

To analyse the influence of subsidies on the demand for solar panels, we need in fact to assess the profitability of PV projects, taking into account costs and subsidies in this profitability measure. Indeed, using only costs and subsidies as covariates would be inappropriate: first, they are likely to evolve simultaneously over time (see figure 5) and hence be somewhat correlated; second, they would not entirely capture the profitability of a project, as they would ignore key local factors such as sun irradiation. The latter could be used as a proxy for profitability, but only in a cross-section analysis, as in Davidson et al. (2014). Knowing these local factors, an alternative would be to use the net present value (NPV), as in Liu and Wei (2016). For a given FIT (possibly enhanced by a premium), discount rate i, investment cost I (possibly reduced through various subsidies) and annual electricity output q, the NPV is defined as follow[9]:

$$\text{NPV} = -I + \sum_{t=1}^{T} \frac{\text{FIT} \times q}{(1+i)^t} \tag{1}$$

Despite its simplicity of use, we think that a NPV can be difficult to interpret, in particular for people who are not familiar with the currency (although the euro is relatively well-known, the use of the same methodology elsewhere could lead to a more difficult interpretation for foreigners). In addition, its computation requires to make an assumption regarding the discount rate. Other authors have used a return-on-investment (ROI) variable, but then an discount rate is either needed (as in Dharshing, 2017), or ignored (as in Jenner et al., 2013). Hence, we find it more convenient to compute the rate which sets the NPV to zero, namely the internal rate of return (IRR). Furthermore, the IRR enables to compare projects independently of the initial investment, while higher investments will give higher NPVs for the same IRR. Thus, we believe that it is a more adequate tool to make an investment decision. Finally, note that some authors use several profitability metrics such as NPV, IRR, payback period or discounted cash flows (e.g. Campoccia et al., 2014; De Boeck et al., 2016; Dusonchet and Telaretti, 2015; Dusonchet and Telaretti, 2010), but these studies aim more at comparing policy instruments in different countries than explaining the installation rates quantitatively.

3.2. Assumptions

In order to compute the IRR for each city, we need to make several assumptions concerning the characteristics of the PV systems and their installation. First of all, although the efficiency of a solar panel decreases with age, we consider for simplicity reasons that it stays constant over time (H1). In practice, solar panel manufacturers guarantee a minimum performance, for example 90% for the 10 first years, and 80% up to 25 years (Energy Informative, 2013).

In addition, we neglect operation and maintenance (O&M) costs (H2). In reality, although panels are usually guaranteed for a certain duration (e.g. 25 years) and do not necessitate any

[9] The choice of having the payments made at the end of each year is arbitrary, and another choice might be more appropriate, but would have a rather small impact on the computed value.

particular maintenance, the inverter usually needs to be replaced every 8-10 years, with a cost varying between 0.3€/kW and 0.6€/kW, i.e. between 900€and 1800€for a 3-kW panel (photovoltaique.info, 2018). Furthermore, the producer has to pay the distribution network for managing its production contract and for the metering of the electricity production. These cost respectively 14.88€and 19.8€(plus VAT) in 2018 (ibid.).

In France, FIT contracts for small-scale PV have a duration of 20 years. However, solar panels usually have a longer lifetime, so that their production becomes a substitute to the electricity sold by the retailer. The profitability of a panel which is no longer subsidised by FIT will depend on the future retail prices and cost of inverter (if it needs to be replaced). It will also depend on how synchronised production and consumption are. It may also be possible to sell the produced output on the wholesale market, for example via an aggregator. Given these uncertainties, we assume that the solar panels have no residual value at the end of the contract (H3), although it probably underestimates their profitability. On the contrary, assumptions H1 and H2 lead to an overestimation of the IRR.

Then, in order to compute the annual production q of a solar panel, we have to make additional hypotheses. Indeed, this output depends on several factors, such as solar irradiance, panel efficiency and orientation, temperature, etc. Solar irradiance (or irradiation) is the power per unit area which is received from the sun by a surface at a particular location. It can be given for example in kW/m^2 or in kWh/m^2.year, depending on the use we make of it. Although the orientation of the panel changes its received irradiance, we do not have this information and thus we use irradiance data for horizontal panels (H4). This assumption may lead to an underestimation of the IRR, as panels are usually set on the “best” (i.e. south, if possible) side of an inclined roof, or face the optimal direction when on a horizontal roof. In some cases however it could lead to an overestimation of the IRR, if the orientation is constrained, or if there are for example trees that shade the panel.

We downloaded irradiance data from PVGIS, the Photovoltaic Geographical Information System of the Joint Research Centre (JRC) of the European Commission[10]. It consists in the yearly sum of global irradiance (in kWh/m^2.year) on a horizontal surface, averaged over ten years (1981-1990), with a grid cell size of 1000 meters, for all Europe. Using a Geographical Information System software, we managed to average this solar irradiation for each French municipality.

Knowing the irradiance, the efficiency characterises how much of the received energy is converted to electricity by the panel. For simplicity, we assume that panels have an efficiency of 15% (H5), which is a frequently encountered value (construction21.org, 2012; solarpanelsphotovoltaic.net, 2014). Another characteristic of the panel is its size. Indeed, in reality the power produced by a panel depends on the received irradiance, on its efficiency and on its area. The output is then just a product of these three variables. However, solar panels have a nominal power output (or nameplate capacity), sometimes given in kilowatt-peak (kWp), which corresponds to the maximum measured power output of the panel under some Standard Test Conditions (STC)[11]. The ratio between the nominal and the real power output is the well-known capacity factor. The nameplate capacity is the information requested by the DSO to anyone asking to connect its panel to the grid, and it is thus what we have in our data. For simplicity reasons again, we assume that the area of all solar panels is 10 m^2/kWp (H6), which is also a common value (Engie, 2017).

[10] Data available at http://re.jrc.ec.europa.eu/pvgis/download/solar_radiation_classic_laea_download.html.

[11] STCs are usually defined by a particular irradiance, temperature, and light spectrum.

Unfortunately, we do not have precise information on the characteristics of the installed panels, apart from their nameplate capacity. Hence, H4-H6 may lead to an under- or overestimation of the IRR, depending on the situation, so that their impact on the IRR is ambiguous. In any case, we think that our assumptions will not change the results dramatically, and that the heterogeneity of the IRR across municipalities will still be captured. We summarise them in table 3 below.

	H1	H2	H3	H4	H5	H6
Assumption	constant efficiency	no O&M costs	no residual value	horizontal panels	common efficiency	common area
Impact on the computed IRR	+	+	-	ambiguous	ambiguous	ambiguous

Table 3: Assumptions made for the computation of the IRR, and their impact.

Given assumptions H1-H6, we are able to compute an IRR per municipality and per quarter. Although the values are not to be taken with certainty given these assumptions, we believe they are still quite reasonable. In any case, these values are a good indicator of the heterogeneity of the profitability of a PV project, and we have discussed how they may influence the IRR. Table 4 below presents some descriptive statistics of the computed IRR, and to illustrate the effect of local subsidies on the IRR, figure 6 shows the average IRR per region, with and without them, in the fourth quarter of 2010. We clearly see on these maps that local subsidies can strongly influence the IRR, and that some northern regions can see their IRR dramatically increased thanks to these.

Statistic	Mean	St. Dev.	Min	Max
IRR	16.006	2.559	7.525	26.039

Table 4: Descriptive statistics of the computed IRR

4. Cities characteristics

In order to control for some of the heterogeneity between municipalities, we use publicly available data on socioeconomic characteristics. These are available at an infra-municipal level called IRIS[12], and are provided by the National Institute of Statistics and Economic Studies (INSEE - Institut national de la statistique et des études économiques). Unfortunately, this data is not available for the most recent years, and so we will only use the 2013 data set to control for these characteristics. This will ***de facto*** prevent us from using individual fixed effects in our model, but these would lead to inconsistent estimates anyway, due to the use of lagged dependent variables (Nickell, 1981). In addition, with only 62,909 positive counts for 33,842 cities, the risk of overfitting the data with as many fixed effects would be relatively high. In the end, we believe that most of the heterogeneity can be captured by the chosen explanatory variables. The 2013 IRIS data set consists of more than a hundred variables, but we will only need a few of them. Indeed, we have tested several specifications using several variables or

[12]Îlots Regroupés pour l'Information Statistique, or "aggregated units for statistical information".

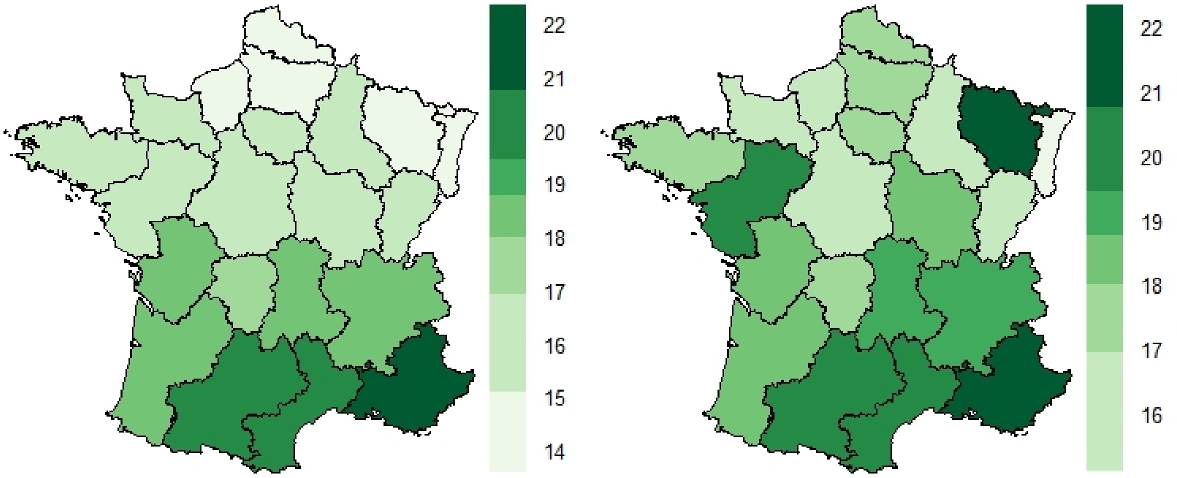


Figure 6: Left: average IRR without local subsidies for a project of less than 3 kW, Q4 2010. Right: average IRR with local subsidies for a project of less than 3 kW, Q4 2010.

combinations of them, but we are only going to present the ones that we finally kept in the model.

Ceteris paribus, the number of installations in a city should be proportional to the number of available sites, i.e. essentially the number houses and residential buildings roofs. Unfortunately, this data is not publicly available, but we can use proxies. To this end, we use the number of dwellings, which does not differentiate between apartments and houses and that can hence capture installations on houses and on buildings. To add another degree of control, we use the proportion of houses (the number of houses would be too collinear with the number of dwellings). We have tried using only the number of houses, but the fit was not as good.

Statistic	Mean	St. Dev.	Min	Median	Max
Number of dwellings	918	4,566	0	234	273,699
Number of primary residences	760	3,956	0	179	244,352
Proportion of primary residences	0.784	0.158	0.000	0.831	1.000
Number of houses	512	1,137	0	219	46,443
Proportion of houses	0.902	0.144	0.003	0.955	1.000

Table 5: Descriptive statistics of the control variables

We also use the proportion of primary residences, which we expect to have a positive effect for at least four reasons (that are debatable). First, we think that people are more likely to get more information about local subsidies where they live than where they go on vacation. Second, even though self-consumption (or “prosumption”) is still little developed in France, some people might value the fact of consuming their own electricity, or at least electricity produced locally. Then it makes more sense for people to invest in a solar panel where they consume the most, i.e. their primary residence. Third, we could argue that people would more likely make an investment that brings value to their principal residence rather than to a secondary one. Finally, we can expect people to want to be able to be present if there is a problem with the installation.

Some descriptive statistics for these variables are shown in table 5.

V. MODELLING STRATEGY

1. Diffusion

The original work of Bass (1969) models the diffusion of innovations by considering that purchases (or sales) S of a new durable good come from "innovators", in fixed proportion p in the remaining market of size $m - Y$ (m is the market size and Y is the stock of cumulative sales), and from "imitators", whose number is a fixed proportion q of the attained market share Y/m. In continuous time, the model then writes:

$$\forall t > 0 \quad S(t) = \text{Max}\left(0, p(m - Y(t)) + q\frac{Y(t)}{m}(m - Y(t)) = pm + (q-p)Y(t) - \frac{q}{m}[Y(t)]^2\right) \tag{2}$$

In discrete time, assuming $S > 0$ for the sake of simplicity, we obtain:

$$\forall t \in \mathbb{N}\backslash\{0\} \quad S_t = a + bY_{t-1} + cY_{t-1}^2 \tag{3}$$

where it is possible to identify coefficients a, b and c to m, p and q:

$$m = \frac{-b \pm \sqrt{b^2 - 4ca}}{2c},\ p = \frac{a}{m},\ q = -mc \tag{4}$$

In practice, b and c must be of opposite sign in order to have $m > 0$, and more precisely we expect to have $b > 0$ (epidemic effect) and $c < 0$ (stock effect): the influence of past sales is positive and decreasing with time. Equation 3 has the advantage of being linear, while equation 2 can be solved analytically and estimated by non-linear least squares. As we will use a non-linear model for which convergence can be difficult to achieve, we will add the Y_{t-1} and Y_{t-1}^2 terms of the linearised equation 3 to our set of covariates and keep a linear combination of covariates.

Although the Bass (ibid.) model does not originally included additional covariates, we keep the assumption of "additivity" of the terms in the equation, in order to keep it linear. Such a modelling of RES diffusion has been done for example by Bollinger and Gillingham (2012) (only for the epidemic effect) and Liu and Wei (2016).

2. Model

As seen before, we are dealing with "zero-inflated" data, which requires the use of an appropriate model to take into account the non-negativity and the important proportion of zeros in the data. In our case, as almost all non-negative requests are for projects of 3 kW, it would be rather difficult to directly use demand as dependent variable, because its distribution is multi-modal, with peaks at 3 kW, 6 kW, etc. For this reason, we prefer to model instead the number of connection requests as zero-inflated count data rather than demand as semi-continuous data. Furthermore, the observed zeros are "real" ones, i.e. are the result of a choice of non-installation, or of non-awareness of the technology (which is linked to the idea of diffusion and adoption of the technology). For this reason, we find it more realistic to model separately the number of zeros.

Hence, we use the "hurdle" model of Mullahy (1986), which consists in two parts: a discrete choice regression first models the proportion of zeros based on the distribution f_{zero} with covariates z; the proportion of strictly positive values is then estimated using a truncated count model. The non-truncated density probability of the count model is f_{count} and depends on covariates x, which need not be the same as z. Hence, the number of installation requests in municipality i during quarter t, S_{it}, follows the following "hurdle" distribution:

$$f_{hurdle}(S_{it}|z_{it}, x_{it}) = \begin{cases} f_{zero}(0|z_{it}) & \text{if } S_{it} = 0 \\ (1 - f_{zero}(0|z_{it}))\dfrac{f_{count}(S_{it}|x_{it})}{1 - f_{count}(0|x_{it})} & \text{if } S_{it} > 0 \end{cases} \tag{5}$$

We choose a logistic distribution[13] for the first part of the model, and a negative binomial distribution (which is common generalisation of the Poisson distribution) for the second part. Hence, we first assume that $S_{it} > 0$ with probability $p_{it} = \mathbb{P}[S_{it} > 0|z_{it}]$ such that:

$$\text{logit}(p_{it}) = \ln\left(\frac{p_{it}}{1 - p_{it}}\right) = z'_{it}\gamma \tag{6}$$

and then that $f_{count}(S_{it}|x_{it}) \rightsquigarrow NegBin(\theta, \mu_{it})$, such that:

$$\begin{cases} f_{count}(S_{it}|x_{it}) = \dfrac{\Gamma(S_{it} + \theta)}{S_{it}!\Gamma(\theta)}\left(\dfrac{\theta}{\theta + \mu_{it}}\right)^{\theta}\left(\dfrac{\mu_{it}}{\theta + \mu_{it}}\right)^{S_{it}} \\ \ln(\mu_{it}) = x'_{it}\beta \end{cases} \tag{7}$$

where Γ is the gamma function. Concerning the covariates, we used for both parts of the model:

- the IRR[14];
- the logarithm of the number of dwellings;
- the proportion of primary residences;
- the proportion of houses.

For the logistic regression, we also added the diffusion variables Y_{t-1} and Y_{t-1}^2, to take into account epidemic and stock effects. This is similar to El Zarwi et al. (2017), who use the cumulative number of adopters at the previous period and other explanatory variables to study the diffusion of new transportation services, using a discrete choice model. However their model differs from ours in that they use three latent classes, namely "imitators", "innovators" and "non-adopters", and use the cumulative number of adopters as a covariate only for the first class, which is consistent with the idea of contagion. Note that we did not use these variables in the count model, as the estimation produces ***NA*** coefficients for these variables.

The negative binomial distribution has the advantage of generalising the Poisson distribution for over-dispersed data, while keeping the same interpretation for the coefficients in case of a log link. Indeed, the conditional mean is then equal to the distribution parameter μ, whose logarithmic dependence with respect to the covariates enables to interpret the coefficients as semi-elasticities (if the covariates are in level) or elasticities (if the covariates are in logarithmic

[13] Other specifications could also be appropriate, for example a probit one. However, a logistic regression leads to more interpretable coefficients (logarithm of odds ratios) and is more consistent with heavy tails (recall that we have a large amount of zeros). Furthermore, the results are usually little affected by the choice of the distribution.

[14] What should really drive the investment decision of a rational agent is the difference between the IRR and a reference rate of return, for example a risk-free one. However, this is captured by the constant.

form). Then, whereas the variance of the Poisson distribution is also equal to μ, it is equal to $\mu(1+\mu/\theta) > \mu$ for the negative binomial distribution.

In our case, the non-truncated count data is over-dispersed (as it is usually the case with zero-inflated data), while the strictly positive count data is under-dispersed. Concerning the zero-truncated distributions, the Poisson distribution leads to under-dispersion while the negative binomial can give both under and over-dispersion, depending on the value of θ, which makes it more flexible and adapted to our case. Finally, although the interpretation of the coefficients is not as straightforward as in the non-truncated case, we will show in the next section how we can still interpret them in our case.

Finally, the use of the logarithm of the number of housing units is justified by the fact that we would expect the number of installations to be rather proportional to the number of dwellings. In a non-truncated model with log link, this is called an exposure variable, whose logarithm is used as an offset, i.e. a covariate with coefficient (elasticity) equal to one. In a zero-truncated negative binomial model however, the expected count is not equal to μ anymore, so that the coefficient of the exposure variable should not be constrained to one.

VI. RESULTS AND INTERPRETATION

Equations 6 and 7 are estimated for the whole sample. We kept the model with the highest Akaike information criterion (AIC), but the results are robust when changing control variables for others (e.g. number of houses instead of number of dwellings).

1. Binomial model (hurdle part)

The estimation results of the hurdle part are presented in table 6 below. We see that all coefficients are highly significant and have the expected sign. Although they cannot be interpreted directly, it is well known that in a logistic regression the exponential of a coefficient can be interpreted as an odds ratio:

$$e^{\beta_j} = OR_j = \frac{\dfrac{\mathbb{P}[S>0|x_j+1]}{1-\mathbb{P}[S>0|x_j+1]}}{\dfrac{\mathbb{P}[S>0|x_j]}{1-\mathbb{P}[S>0|x_j]}} \tag{8}$$

Hence, table 7 shows the exponential of the coefficients (we have omitted the intercept). From this, we can determine that a one-point increase of the IRR increases the odd of having at least one installation by 9.8%. The mean odds of installation would then increase from $6/94 \simeq 0.064$ on average to 0.070, i.e. there would be 93.5% of zeros instead of 94%. This represents a 0.5 percentage point increase of the probability of adoption over the whole period. Furthermore, we observe epidemic and stock effects. In particular, at the first order (i.e. considering only Y_{t-1}) we find that an additional installation in the past periods has almost the same impact as an additional point of IRR. This highlights the importance of word-of-mouth in the diffusion mechanism. Also, although residential solar panels are not widespread in France yet, we do find a stock effect, i.e. a decrease in the epidemic effect with the number of installations. Although this effect is very small and hence plays a minor role in the diffusion at this stage, it is likely to become more important in the future. Finally, the proportions of primary residences and houses have a much more stronger influence. However, recall that these are only control variables, on which public policies have little to no impact in the short run.

	Estimate	Std. Error	z value	p-value
(Intercept)	-12.0792	0.0714	-169.16	$< 2e\text{-}16$
IRR	0.0938	0.0018	53.22	$< 2e\text{-}16$
ln(Dwellings)	0.7270	0.0053	137.15	$< 2e\text{-}16$
Proportion of primary residences	1.6525	0.0351	47.05	$< 2e\text{-}16$
Proportion of houses	2.1300	0.0361	59.07	$< 2e\text{-}16$
Y_{t-1}	0.0920	0.0016	57.24	$< 2e\text{-}16$
Y_{t-1}^2	-0.0013	0.0000	-30.91	$< 2e\text{-}16$

Table 6: Estimation results of the hurdle part

IRR	ln(Dwellings)	Prop. of primary residences	Prop. of houses	Y_{t-1}	Y_{t-1}^2
1.098	2.069	5.220	8.415	1.096	0.999

Table 7: Exponentiated coefficients of the hurdle part

Concerning the explanatory power of the covariates, we use the AIC as means of comparison. Table C.1 in appendix C shows the value of the AIC for several specifications of the binomial model. We see that the IRR improves the IRR almost as well as the diffusion variables. We can hence consider that their explanatory power is quite similar.

2. Count model

The estimation results of the count part are shown in table 8 below. As for the logistic regression, all coefficients are significant (the IRR a little less) and have the expected sign. Unfortunately, convergence could not be achieved using diffusion variables Y_{t-1} and Y_{t-1}^2. However, as the main outcome is one, the first part of the model explains most of the adoptions, while the second part can be seen as a second-order refinement. Nevertheless the count model is expected to have an increasing explanatory power as the diffusion progresses, and to become influenced by the diffusion variables as well.

	Estimate	Std. Error	z value	p-value
(Intercept)	-11.052377	0.211384	-52.286	$< 2e\text{-}16$
IRR	0.009411	0.004591	2.050	0.0404
ln(Dwellings)	0.814709	0.012397	65.717	$< 2e\text{-}16$
Proportion of primary residences	1.289737	0.092136	13.998	$< 2e\text{-}16$
Proportion of houses	2.225016	0.080191	27.746	$< 2e\text{-}16$
$\ln(\theta)$	-1.126851	0.128216	-8.789	$< 2e\text{-}16$

Table 8: Estimation results of the count part

For a standard Poisson or negative binomial regression with log link, coefficients are semi-elasticities of the conditional mean (an increase of x_j by one unit leads to an increase of the dependent variable Y by $100 \times \beta_j$ percent:

$$\mathbb{E}[Y|X] = \lambda = e^{X'\beta} \Rightarrow \frac{1}{\mathbb{E}[Y|X]} \frac{\partial \mathbb{E}[Y|X]}{\partial x_j} = \beta_j \tag{9}$$

Alternatively, exponential of coefficients can be seen as a multiplicative impact when there is an increase of one unit:

$$\mathbb{E}[Y|x_j + 1] = e^{\beta_j} \mathbb{E}[Y|x_j] \tag{10}$$

At the difference of the non-truncated distribution, the zero-truncated negative binomial of parameters μ and θ have a mean equal to $\mathbb{E}[S|X, S > 0] \underset{not.}{=} \nu = \dfrac{\mu}{1 - (1 + \mu/\theta)^{-\theta}} > \mu$, so that the exponential of the coefficients cannot longer be interpreted as multiplicative incremental effects on the conditional mean. However, when $\mu \to 0$ i.e. when the main outcome is $S = 1$ (which is our case), we have:

$$\nu \underset{\mu \to 0}{\sim} 1 + \frac{\theta + 1}{2\theta} \mu \tag{11}$$

This enables to compute the following partial derivatives:

$$\frac{1}{\nu - 1} \frac{\partial(\nu - 1)}{\partial x_j} \simeq \beta_j \tag{12}$$

or alternatively:

$$(\mathbb{E}[S|x_j + 1, S > 0] - 1) \simeq e^{\beta_j} (\mathbb{E}[S|x_j, S > 0] - 1) \tag{13}$$

Coefficients are hence semi-elasticities (or elasticities when covariates are in logarithmic form) for the conditional mean minus one. Although these is not as easily interpretable as a "real" (semi-)elasticities, they nevertheless help to compute the effect of a given variable, as we will see now. Exponentiated coefficients are even more easily interpretable because they act in a multiplicative way on the conditional mean minus one. In our case, the main strictly positive outcome is 1, and the expected number of strictly positive counts is 1.267. Thus, we are in the case described above and we can interpret exponentiated coefficients presented in table 9 (we have removed the intercept).

IRR	ln(Dwellings)	Prop. of primary residences	Prop. of houses	θ
1.009	2.259	3.632	9.254	0.3240521

Table 9: Exponentiated coefficients of the zero-truncated count model

For example, starting from the mean (1.267), a one-point increase of the proportion of primary residences will drive the average strictly positive count up to $3.632 \times 0.267 + 1 = 1.97$. Concerning the IRR, a one-point increase leads to an average positive count of $1.009 \times 0.267 + 1 = 1.269$ (+0.16%). Finally, as for the hurdle part, the impact of the control variables is much stronger. In particular, the coefficient associated with the number of dwellings is quite close to one (which would indicate proportionality in the case of a non-truncated model, as explained in section V.

VII. CONCLUSION

This paper has investigated the diffusion of domestic solar panels in France. Thanks to a novel modelling methodology, we are able to exhibit "epidemic" and "stock" effects down to the local (municipal) level, even though the data is highly disaggregated and presents a very high proportion of zeros. Furthermore, we manage to take into account all kinds of support mechanisms (national feed-in tariffs and local subsidies) by computing an internal rate of return per quarter, for all 33,842 municipalities whose distribution network is managed by Enedis. We show that local subsidies can deeply influence the IRR, which has reached values in the north of France as high as in the south. Furthermore, the IRR has a positive and significant impact on both the probability of PV adoption and the strictly positive number of installations. This, along with other evidence from the data, indicates that households act in a rather rational way. It is thus the combination of financial incentives and an intrinsic diffusion process that enabled the development of residential solar panels in France. In particular, we show that an additional past installation in a city has the same impact as a one-point increase of the IRR. This result calls for more information and "advertising" of RES projects, as it can help promote RES at a lower cost than direct public subsidies. This could be done for example by better informing people, but also by encouraging local ownership and funding of such projects (this is still underdeveloped in France, whereas it is much more common in Germany for example). Hence, our conclusions support the findings of other authors, who have pointed out the important role of neighbours and installers (Rai et al., 2016) as well as local organisations promoting PV and local utilities (Palm, 2016).

Future avenues of research could aim at analysing the spatial dependence in the diffusion of residential solar panels. Indeed, several authors have found positive spatial correlation (Müller and Rode, 2013; Balta-Ozkan et al., 2015) or spatial clustering (Bollinger and Gillingham, 2012) of PV deployment. Although these authors have highlighted the presence of local peer effects, they did not analyse stock effects or looked into local subsidies. To the best of our knowledge, no study analyses both diffusion and the impact of all applicable subsidies at a fine spatio-temporal scale. However, we might need higher resolution data to be able to perform such an analysis, as in Müller and Rode (2013), who used building-level data in the city of Wiesbaden, Germany. Also, the French market might not be mature enough, with an insufficient number of installations.

Another continuation of this work could include forecasts based on our model, in order to compare the trajectories with national and regional targets. Such a study could also help answer the question of the end of subsidies, as policy makers would like RES to be competitive with other means of production as soon as possible. When a technology reaches the grid parity (i.e. a levelised cost of electricity - LCOE - equal to the retail price), it becomes interesting for a consumer to invest in order to consume its own production. However, this behaviour might create cross-subsidies from "regular" consumers to "prosumers" through the variable energy-related component of the network tariff. In turn the latter tends to increase, in order to compensate for the incurred loss in revenue. This might lead to the so-called "death spiral", i.e. disconnection of most consumers, due to very high network charges. Hence, the incorporation of retail electricity prices and network tariff in a diffusion model of "prosumption" could help forecast future revenues for the DSO, and hence design efficient tariffs and electricity pricing methodologies that enable the DSO to recover its costs while ensuring an efficient use of the network. This may also be done using for example system dynamics, which has proven to be adapted to describe the diffusion of RES (e.g. Bildik et al., 2015), but could nevertheless use our model as a valuable input. Finally, our methodology could also be used to describe the diffusion of other relatively new technologies, such as alternative-fuel vehicles or batteries, at a high spatio-temporal resolution.

Bibliography

Aguirre, Mariana and Gbenga Ibikunle (2014). “Determinants of renewable energy growth: A global sample analysis”. ***Energy Policy*** 69, pp. 374–384.

Allan, Grant J. and Stuart G. McIntyre (2017). “Green in the heart or greens in the wallet? The spatial uptake of small-scale renewable technologies”. ***Energy Policy*** 102, pp. 108–115.

Balta-Ozkan, Nazmiye, Julide Yildirim, and Peter M. Connor (2015). “Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach”. ***Energy Economics*** 51, pp. 417–429.

Bass, Frank M. (1969). “A New Product Growth for Model Consumer Durables”. ***Management Science*** 15.5, pp. 215–227.

Bauner, Christoph and Christine L. Crago (2015). “Adoption of residential solar power under uncertainty: Implications for renewable energy incentives”. ***Energy Policy*** 86, pp. 27–35.

Benthem, Arthur van, Kenneth Gillingham, and James Sweeney (2008). “Learning-by-Doing and the Optimal Solar Policy in California”. ***The Energy Journal*** 29.3, pp. 131–151.

Bildik, Y, CE Van Daalen, G Yü, JR Ortt, and WAH Thissen (2015). “Modelling Wind Turbine Diffusion: A comparative study of California and Denmark 1980-1995”. ***International Conference of the System Dynamics Society***, pp. 1–25.

Bollinger, Bryan and Kenneth Gillingham (2012). “Peer Effects in the Diffusion of Solar Photovoltaic Panels”. ***Marketing Science*** 31.6, pp. 900–912.

Boomsma, Trine Krogh and Kristin Linnerud (2015). “Market and policy risk under different renewable electricity support schemes”. ***Energy*** 89, pp. 435–448.

Cadoret, Isabelle and Fabio Padovano (2016). “The political drivers of renewable energies policies”. ***Energy Economics*** 56, pp. 261–269.

Campoccia, A, L Dusonchet, E Telaretti, and G Zizzo (2014). “An analysis of feed’in tariffs for solar PV in six representative countries of the European Union”. ***Solar Energy*** 107, pp. 530–542.

Carfora, Alfonso, Antonio A. Romano, Monica Ronghi, and Giuseppe Scandurra (2017). “Renewable generation across Italian regions: Spillover effects and effectiveness of European Regional Fund”. ***Energy Policy*** 102, pp. 132–141.

construction21.org (2012). ***Efficiency of Solar energy harvesting***. url: `https://www.construction21.org/articles/h/efficiency-of-solar-energy-harvesting.html` (visited on 04/13/2018).

CRE (2016). ***Délibération de la Commission de régulation de l’énergie du 13 juillet 2016 relative à l’évaluation des charges de service public de l’énergie pour 2017***. In French. url: `http://www.cre.fr/documents/deliberations/decision/cspe-2017/consulter-la-deliberation` (visited on 04/09/2018).

Davidson, Carolyn, Easan Drury, Anthony Lopez, Ryan Elmore, and Robert Margolis (2014). “Modeling photovoltaic diffusion: an analysis of geospatial datasets”. ***Environmental Research Letters*** 9.7, p. 074009.

De Boeck, Liesje, Siemen Van Asch, Philippe De Bruecker, and Amaryllis Audenaert (2016). “Comparison of support policies for residential photovoltaic systems in the major EU markets through investment profitability”. ***Renewable Energy*** 87, pp. 42–53.

Dharshing, Samdruk (2017). “Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany”. ***Energy Research & Social Science*** 23, pp. 113–124.

Dijkgraaf, Elbert, Tom P. van Dorp, and Emiel Maasland (2018). “On the effectiveness of feed-in tariffs in the development of solar photovoltaics”. ***The Energy Journal*** 39.1.

Dusonchet, L and E Telaretti (2015). "Comparative economic analysis of support policies for solar PV in the most representative EU countries". ***Renewable and Sustainable Energy Reviews*** 42, pp. 986–998.
Dusonchet, Luigi and Enrico Telaretti (2010). "Economic analysis of different supporting policies for the production of electrical energy by solar photovoltaics in western European Union countries". ***Energy Policy*** 38.7, pp. 3297–3308.
Ek, Kristina, Lars Persson, Maria Johansson, and Åsa Waldo (2013). "Location of Swedish wind power—Random or not? A quantitative analysis of differences in installed wind power capacity across Swedish municipalities". ***Energy Policy*** 58, pp. 135–141.
El Zarwi, Feras, Akshay Vij, and Joan L Walker (2017). "A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services". ***Transportation Research Part C: Emerging Technologies*** 79, pp. 207–223.
Energy Informative (2013). ***Solar Panel Warranty Comparison***. url: http://energyinformative.org/solar-panel-warranty-comparison/ (visited on 06/27/2018).
Engie (2017). ***Puissance d'un panneau solaire photovoltaïque : le b.a.-ba***. In French. url: https://particuliers.engie.fr/electricite/conseils-electricite/photovoltaique/puissance-panneau-solaire.html (visited on 04/13/2018).
Graziano, Marcello and Kenneth Gillingham (2015). "Spatial Patterns of Solar Photovoltaic System Adoption: The Influence of Neighbors and the Built Environment". ***Journal of Economic Geography*** 15.4, pp. 815–839.
Guidolin, Mariangela and Cinzia Mortarino (2010). "Cross-country diffusion of photovoltaic systems: Modelling choices and forecasts for national adoption patterns". ***Technological Forecasting and Social Change*** 77.2, pp. 279–296.
Hitaj, Claudia (2013). "Wind power development in the United States". ***Journal of Environmental Economics and Management*** 65.3, pp. 394–410.
Jenner, Steffen, Felix Groba, and Joe Indvik (2013). "Assessing the strength and effectiveness of renewable electricity feed-in tariffs in European Union countries". ***Energy Policy***. Special Section: Transition Pathways to a Low Carbon Economy 52, pp. 385–401.
Karshenas, Massoud and Paul L. Stoneman (1993). "Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model". ***The RAND Journal of Economics*** 24.4, pp. 503–528.
Kilinc-Ata, Nurcan (2016). "The evaluation of renewable energy policies across EU countries and US states: An econometric approach". ***Energy for Sustainable Development*** 31, pp. 83–90.
Kwan, Calvin Lee (2012). "Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States". ***Energy Policy*** 47, pp. 332–344.
Liu, Yang and Taoyuan Wei (2016). "Market and Non-market Policies for Renewable Energy Diffusion: A Unifying Framework and Empirical Evidence from China's Wind Power Sector". ***The Energy Journal*** 37.1.
Marschinski, Robert and Philippe Quirion (2014). "Tradable renewable quota vs. feed-in tariff vs. feed-in premium under uncertainty".
Min, Yongyi and Alan Agresti (2002). "Modeling nonnegative data with clumping at zero: a survey". ***Journal of the Iranian Statistical Society*** 1.1, pp. 7–33.
Mir-Artigues, Pere and Pablo del Río (2014). "Combining tariffs, investment subsidies and soft loans in a renewable electricity deployment policy". ***Energy Policy*** 69, pp. 430–442.
MTES (2016). ***Commercialisation de l'électricité***. In French. url: https://www.ecologique-solidaire.gouv.fr/commercialisation-lelectricite (visited on 04/09/2018).

Mullahy, John (1986). “Specification and testing of some modified count data models”. ***Journal of econometrics*** 33.3, pp. 341–365.
Müller, Sven and Johannes Rode (2013). “The adoption of photovoltaic systems in Wiesbaden, Germany”. ***Economics of Innovation and New Technology*** 22.5, pp. 519–535.
Nickell, Stephen (1981). “Biases in dynamic models with fixed effects”. ***Econometrica: Journal of the Econometric Society***, pp. 1417–1426.
Palm, Alvar (2016). “Local factors driving the diffusion of solar photovoltaics in Sweden: A case study of five municipalities in an early market”. ***Energy Research & Social Science*** 14, pp. 1–12.
photovoltaique.info (2018). ***Coûts de fonctionnement***. In French. url: http://www.photovoltaique.info/Couts-de-fonctionnement.html (visited on 06/27/2018).
Radas, Sonja (2006). “Diffusion Models in Marketing: How to Incorporate the Effect of External Influence?” ***Economic Trends and Economic Policy*** 15.105, pp. 30, 30–51, 51.
Rai, Varun, D Cale Reeves, and Robert Margolis (2016). “Overcoming barriers and uncertainties in the adoption of residential solar PV”. ***Renewable Energy*** 89, pp. 498–505.
Schaffer, Axel and Claudia Düvelmeyer (2016). “Regional drivers of on-farm energy production in Bavaria”. ***Energy Policy*** 95, pp. 361–369.
Schaffer, Axel J. and Sebastian Brun (2015). “Beyond the sun—Socioeconomic drivers of the adoption of small-scale photovoltaic installations in Germany”. ***Energy Research & Social Science*** 10, pp. 220–227.
Schilling, Melissa A. and Melissa Esmundo (2009). “Technology S-curves in renewable energy alternatives: Analysis and implications for industry and government”. ***Energy Policy*** 37.5, pp. 1767–1781.
Schmidt, J., G. Lehecka, V. Gass, and E. Schmid (2013). “Where the wind blows: Assessing the effect of fixed and premium based feed-in tariffs on the spatial diversification of wind turbines”. ***Energy Economics*** 40, pp. 269–276.
solarpanelsphotovoltaic.net (2014). ***Polycrystalline Solar Panels – Everything you possibly would want to know***. url: http://solarpanelsphotovoltaic.net/polycrystalline-solar-panels-everything-possibly-want-know/ (visited on 04/13/2018).
Strunz, Sebastian, Erik Gawel, and Paul Lehmann (2016). “The political economy of renewable energy policies in Germany and the EU”. ***Utilities Policy*** 42, pp. 33–41.
Tobin, James (1958). “Estimation of relationships for limited dependent variables”. ***Econometrica: journal of the Econometric Society***, pp. 24–36.
Zhang, Yu, Junghyun Song, and Shigeyuki Hamori (2011). “Impact of subsidy policies on diffusion of photovoltaic power generation”. ***Energy Policy*** 39.4, pp. 1958–1964.

Appendices

A. Demand for 3-kW PV projects per region

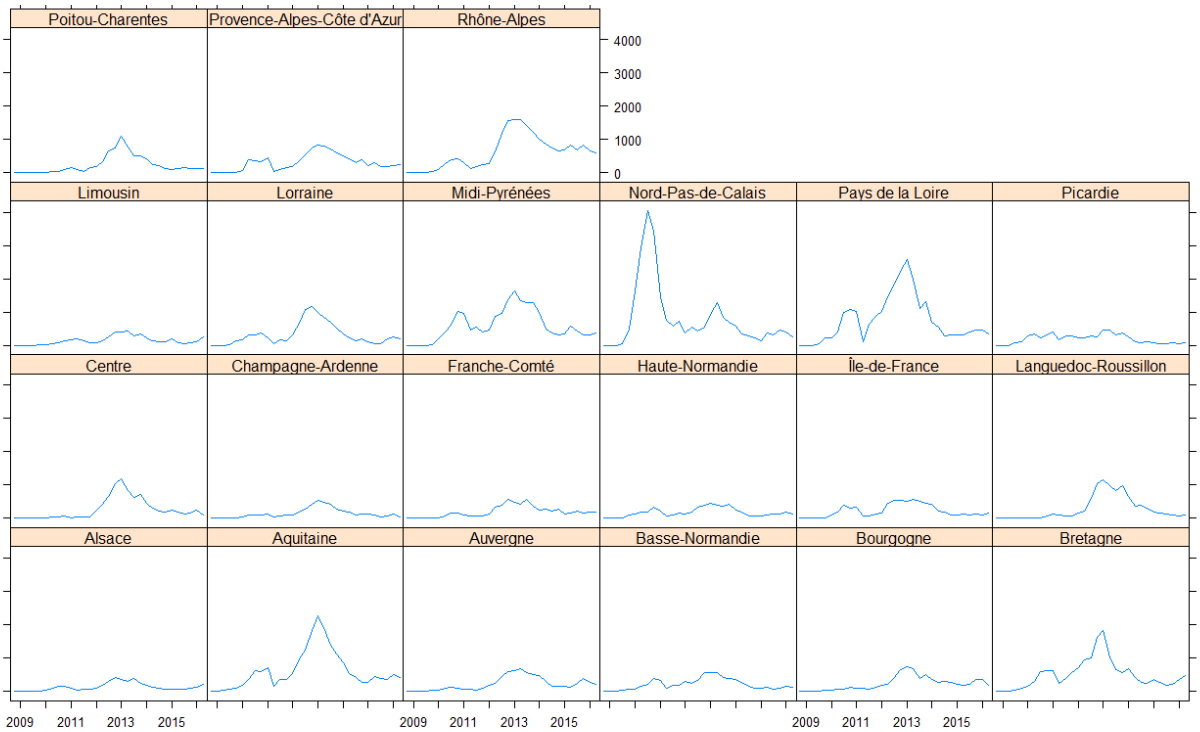


Figure A.1: Quarterly demand (kW) for PV projects of less than 3 kW

B. Cumulative demand for 3-kW PV projects per region

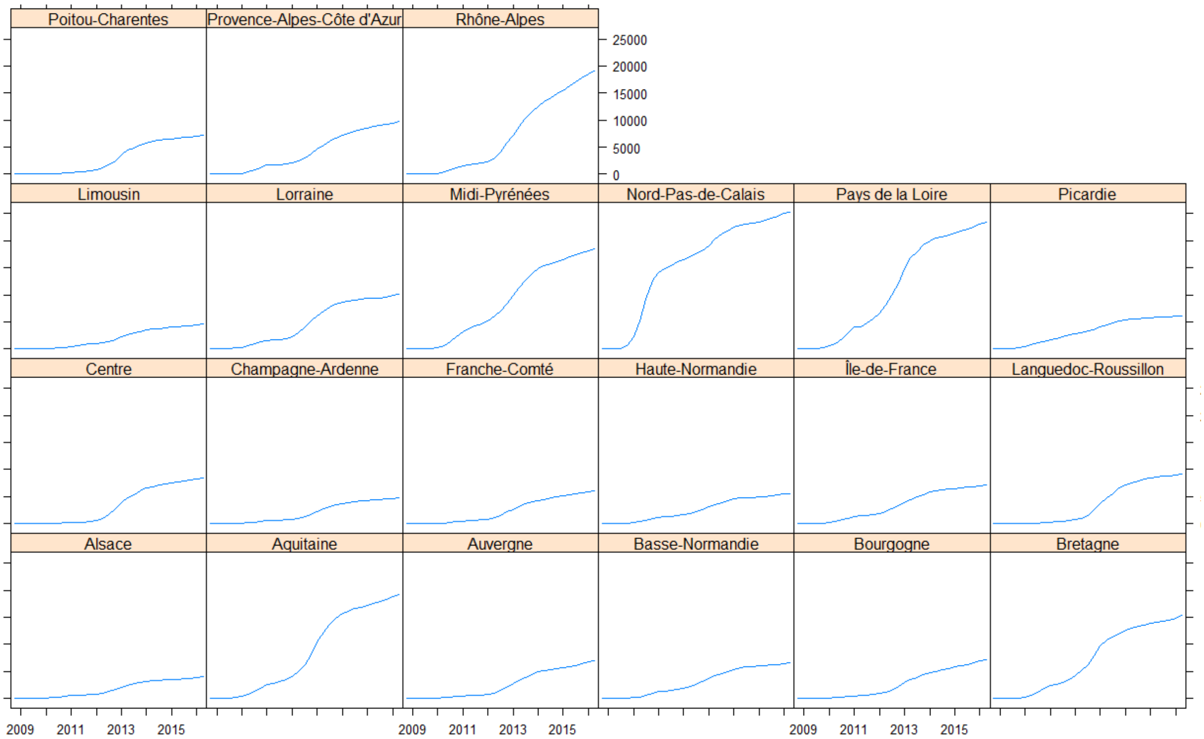


Figure B.1: Quarterly cumulative demand (kW) for PV projects of less than 3 kW

C. AIC of several binomial models

Constant only	Control variables only	Diffusion and control variables	IRR and control variables	All variables
474,579	416,033	411,995	412,947	409,171

Table C.1: AIC for several binomial model specifications