The impact of renewable energy facilities and traditional power plants on local housing prices: a study from the UK.

Contents

1. Introduction

Although restrictions, economic closures, and lockdowns due to the Covid-19 pandemic continues to subdue industrial and social activities worldwide, global energy demand is expected to increase by 4.6% in 2021, more than offsetting the 4% contraction in 2020 and driving demand 0.5% above 2019 pre pandemic levels (IEA, 2021). Despite growing interest in renewable energy sources, most of this growth is projected to be met with increased production from fossil fuels. Coal demand alone is expected to expand by 60% more than all renewables combined, approaching global coal demand peak in 2014 and registering a rise in emissions of almost 5% with respect to the past year.

The 2021 Energy Review by the International Energy Agency also reports that electricity demand is expected to increase by 4.5% in 2021 and almost 80% will come from emerging markets and developing countries. In order to meet this increasing demand of electricity, since 2000, coal-fired power capacity in the world has doubled to around 2,045 gigawatts (GW) after exceptional growth in China and India. Along with natural gas, renewable energy (mainly hydropower, wind, solar and biofuels) has also increasingly grown in importance in terms of installed capacity and investments in many countries in order to halt the pace of climate change (IRENA, 2021). For these reasons, new gas and coal power plants are scheduled to be opened worldwide, and new and progressively advanced renewables facilities such as photovoltaic parks and wind farms are becoming widespread over the planet.

The majority of coal and gas-based power plants are large structures that alter the visual landscape and release hazardous air pollutants. As a consequence, and despite the fact that over time new and more environmentally friendly techniques and technologies have been developed, fossil fuel power plants are still bad for human health.

Along with traditional fossil fuel power plants, renewables energy facilities too have increased their size over time. Wind turbines have become taller over the years: the height of the first turbines built in the 1980s did not surpass 30 metres, while the newest generation of wind turbines goes well beyond 100 metres. Similarly, the first solar park was constructed during the 1980s in California. However, with advances in technology, it was only in the last decade that the costs of photovoltaic panels declined dramatically and the commercial exploitation of solar parks has become economically attractive.

The size of these photovoltaic stations has also become bigger over time: the largest solar park currently active is 40 km^2 and located in Bhadla, India.

For these reasons, it is increasingly more difficult to obtain the approval of local communities in siting new plants. Households may value the new power plant as a disamenity and it is expected that these claims about negative external effects would capitalise into local house prices. This dissertation seeks to test this, using data on residential property sales tracked by HM Land Registry in England and Wales, matched with information on the times and locations of 86 new energy facilities' openings between 1995 and 2021. The existing literature has analysed this relationship using relatively small samples or inspecting the effect of a single facility on the local housing market.

Moreover, very few studies to date have examined a comparison of several technologies, be them green energy or fossil fuel power. This study instead wants to provide a clear and encompassing assessment of the effect of different technologies through the analysis of different types facilities, including more traditional ones such as gas-fired power plants, combined cycle power plants, and steam power plants, as well as newer stations based on renewable sources such as wind, solar, biomass and waste.

Following the hedonic price method literature for environmental valuation, I use a difference-in-differences (DID) research design that compares price changes before and after the new power station becomes operational between postcodes within 3 kilometres of facility and postcodes that are more than 3 kilometres away from the facility. My data consists of a panel of transactions, whereby each house is observed to change owner multiple times over time. I aggregate the observations at the postcode level and use postcode fixed effects in the analysis.

The estimated effect of plant openings on the local housing prices varies starkly across plant technologies. Openings of more traditional power stations, especially CCGT and conventional steam facilities, determine a significant reduction in nearby house prices, of around 3% and 12%, respectively. Among renewable energy facilities, openings of biomass and waste sites significantly decrease housing prices by around 9%, displaying a similar effect in sign and magnitude to that of traditional plant openings. Openings of wind farms have no significant effect on housing prices. Conversely, openings of solar parks increase nearby house prices by around 2%. Furthermore, the effect of introducing new energy sites on house prices is found to vary across house as well as plant characteristics. As one might expect, larger stations tend to have a bigger impact on their vicinity.

This dissertation is structured as follows. Section 2 describes briefly the energy policy worldwide and the trends in the use of alternative sources such as nuclear and renewable energy. After that, the focus shifts to the United Kingdom and the country's plans and developments for meeting the increasing energy demand. In the last subsection of Section 2, the alleged externalities of different types of power stations are presented within the context of the existing literature, whose main results are reviewed. Section 3 introduces the most common methods used to evaluate the externalities of electricity generation based on real estate data. These include stated preferences and revealed preferences methods. Special attention is devoted to the hedonic price method. Then, a wide literature review about the effects of different kinds of energy facilities on housing values is presented. Section 4 describes the data used for the analysis and the empirical strategy deployed, also delineating the difference in differences frameworks and the assumptions needed for the approach to yield causal estimates. Section 5 reports the results and robustness checks. Section 6 discusses about some policy implications and concludes.

2. Background

Modern economies depend on the reliable, affordable, and certain delivery of electricity. At the same time, climate change is real and is driving a dramatic transformation of power systems globally. According to the 2020 Statistical Review of World Energy by bp, in 2019 renewable energy accounted for 41% of the annual growth in primary energy. This exceptional result is majorly led by wind and solar power. Simultaneously, coal consumption fell in a pre pandemic situation for the fourth time in the past six years. Still, coal is the largest source of power generation, accounting for 36% of global power, against the 10% provided by renewable energy.

Nevertheless, the International Renewable Energy Agency (IRENA) has estimated that renewables will provide the majority of global energy at an impressive 86% by 2050. In this sense, the majority of countries in the world are reconsidering their energy policies, as technology and innovation disrupt traditional models. With regards to nuclear power, nuclear technology has lost its appeal. Research funding has dropped dramatically and some countries have decided to definitively exit nuclear power. For instance, Germany has responded to the Fukushima nuclear disaster by proclaiming an “energy transition plan” that entails an accelerated phase-out of nuclear power by the end of 2022 and a generic goal for phasing-in renewable energy (RE). Many others have decided that at the end of life of their reactors they will not build new ones (Sweden, Belgium, Spain, Switzerland). A once-leading country like Japan expressed strong doubts on whether or not it should try to reopen most of its reactors and complete those under construction. On the other hand, nuclear’s share in French electricity production, the IEA reports, was 70.6% in 2019, the highest in the world, and France still promotes nuclear as the cornerstone of the country’s energy mix.

In terms of renewable energy, Europe was the world leader for a long time, but China and other developing countries such as India and Brazil have now become central in wind and solar energy’s development. Germany’s renewables use is the highest in the world followed by other European countries (Sweden, Spain, Italy) due to massive investments made as part of efforts by the governments to meet EU renewable energy and climate action goals.

2.1 Context in the United Kingdom

The United Kingdom's energy system has witnessed an exceptionally rapid growth in the share of low-carbon energy, which accounted for over 50% of the electricity mix in 2017 (IEA, 2019a): natural gas (41%), nuclear (21%), wind (15%, up from 3% in 2010), solar (3%), bioenergy and waste (11%), coal (7%, down from 29% in 2010) and hydro (2%).

The UK is one of the global leaders in decarbonising energy supply. As Figure 1 shows, coal supply has decreased since 1990 and the government has committed to phasing out all remaining unabated coal-fired power generation by 2025. In fact, since 1970 the UK has already implemented a dramatic decrease of 96% in its use of coal, demonstrating its effort to using different energy sources. As far as climate change is concerned, the country has led the way in the transition to a low-carbon economy by taking ambitious climate action at international and national levels. These efforts are consistent with its goal to reduce greenhouse gas emissions by at least 80% by 2050 from 1990 levels, as decided by the 2008 Climate Change Act.

In term of nuclear energy, the UK has 13 operational nuclear reactors at six locations (12 advanced gas-cooled reactors and one pressurised water reactor), as well as nuclear reprocessing plants at Sellafield and the Tails Management Facility (TMF) operated in Capenhurst.

To date, the Capacity Market -the mechanism introduced by the government to ensure that electricity supply continues to meet demand as more volatile and unpredictable renewable generation plants come on stream- has secured over 5.4 GW of new gas-fired build capacity for delivery between 2018/19 and 2021/22. Table 1 shows the new electricity generating capacity from 2017 to 2022 by fuel types. As it is clear from figures shown in Table 1, investments to increment the power capacity involving oil and natural gas facilities have been made in the last few years and are planned also for 2021/2022. Plans indeed have been made to build two new combined cycle gas turbine (CCGT) plants at Kings Lynn and Carrington, a new OCGT plant at Spalding, and a significant number of small-scale gas engines. However, renewables are by far

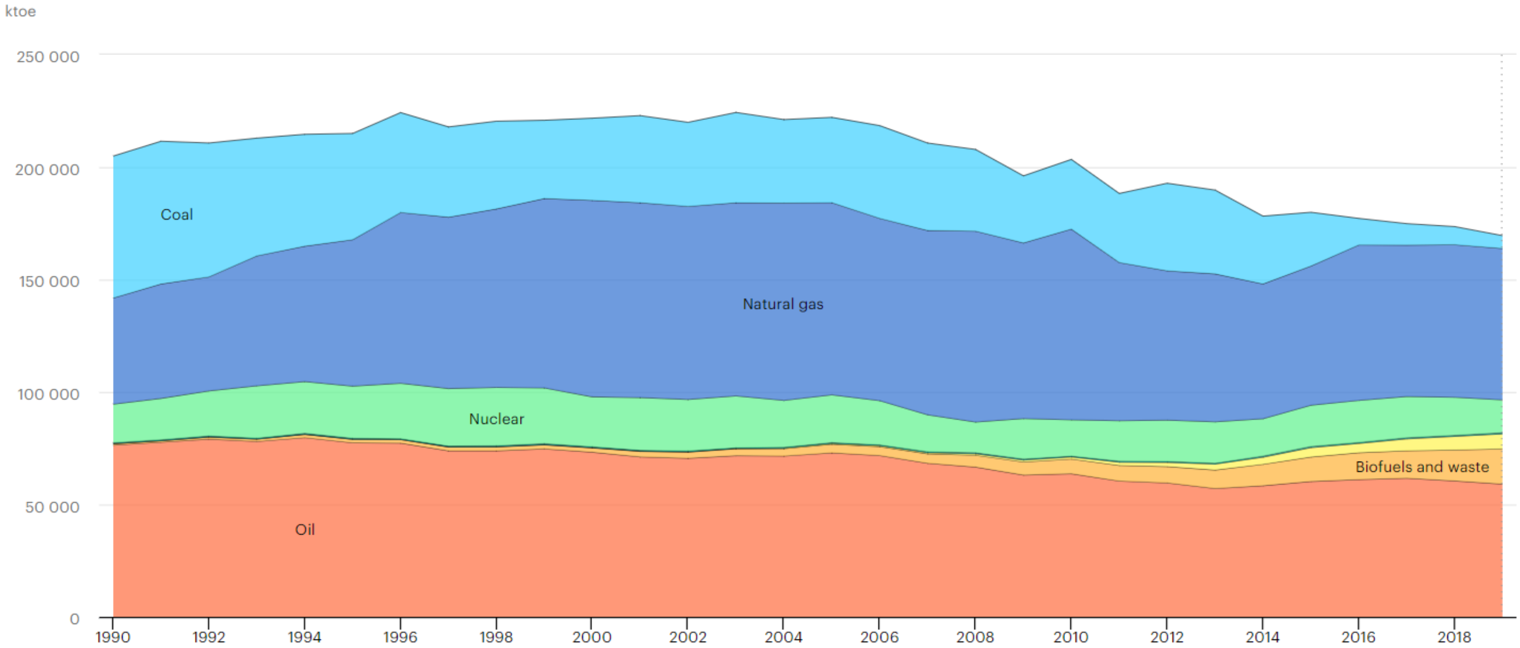


Figure 1: Total energy supply, United Kingdom 1990-2019. Source: iea.org

the largest new investments in terms of power production capacity. This includes wind farms and solar parks which are potentially attractive businesses for producers and small landowners because the electricity they generate is eligible for Renewables Obligation Certificates, which are issued by the sector regulator (Ofgem) and guarantee a price at premium above the market rate. This premium price is subsidised by a tariff on consumer energy bills. In fact, in the last ten years, the UK has massively increased the production of wind and solar energy, which accounted for 29 percent of

Table 1: Cumulative new capacity (GW) for all power producers by fuel type

Fuel type	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
Coal	0.00	0.00	0.00	0.00	0.00	0.00
Coal and natural gas CCS	0.00	0.00	0.00	0.00	0.00	0.00
Oil	0.00	0.20	0.35	0.82	0.94	0.94
Natural gas	0.00	1.02	1.73	2.27	3.83	3.83
Nuclear	0.00	0.00	0.00	0.00	0.00	0.00
Other thermal	0.00	0.00	0.00	0.00	0.00	0.00
Renewables	3.59	6.93	8.30	9.76	11.60	15.14
Interconnectors	0.00	0.00	1.00	1.00	3.00	5.80
Storage	0.00	0.21	0.21	0.21	0.60	0.60
Total cumulative new build	**3.59**	**8.36**	**11.58**	**14.06**	**19.96**	**26.31**

Source: NG (2017), Reference Scenario, also available under: www.gov.uk/government/uploads/system/uploads/attachment_data/file/666266/Annex-k-total-cumulative-newcapacity.xls.

Britain's electricity production last year, putting the country in the sixth place on the global league table. To make a comparison, wind power alone surpassed coal in 2016 and nuclear in 2018 in the electricity production in the country.

2.2 The negative impact of power plants

Electric energy cannot be directly used to fulfil man's needs, especially on an industrial scale. Therefore, it needs to be generated by power plants through the use of other forms of energy. These plants can be classified according to the method used to generate electricity into the following macro categories (Bielecki et al. 2020):

- **Conventional thermal power plants**, where thermal energy is obtained through chemical combustion of fossil fuels: coal, crude oil, or natural gas. One example is a simple cycle gas-turbine, also known as open cycle gas-turbine (OCGT). Due to their high level of emissions of toxic substances and carbon dioxide, newer, more efficient and cleaner technologies are being deployed, especially in developed countries. One of the most used is the combined-cycle power plant. Using both gas and steam turbines, they produce higher amounts of electricity from a single fuel source compared to a traditional power plant, while reduce health impact costs.
- **Nuclear power plants**, where electricity is produced in the process of nuclear fission.
- **Power plants based on the combustion of biomass** or its decomposition production. Biomass sources for energy include:
    - Wood and wood processing wastes;
    - Agricultural crops and waste materials;
    - Biogenic materials in municipal solid waste.
- **Hydroelectric power plants**, which generate electricity either by using the kinetic energy of flowing water or by the potential energy of water gathered in dams in the Earth's gravitational field.
- **Geothermal power plants**, which use the Earth's thermal energy.
- **Photovoltaic power plants,** also called solar parks, large-scale solar photovoltaic arrays which convert direct sunlight and solar power into electricity.
- **Wind farms,** which transform the kinetic energy of wind into electricity, through a group of wind turbines in the same location**.** They can be onshore or offshore.

One of the biggest challenge to the cost-benefit analysis of alternative electricity supply options consists of the externalities created by electricity generation. This challenge involves two different layers. One is the overall level of negative externalities created by all the options available. And this overall level is crucial to the choice of the electricity mix in a country that is preoccupied for climate change and environmental pollution. The other dimension is the incidence of those externalities from power plants that have an impact only at a local or regional level.

In fact, one of the greatest obstacle in siting new plants is opposition from local communities. Citizen groups argue that power plants are a source of numerous negative local externalities including pollution, visual disamenities and noise. This happens not only with traditional fossil fuel based plants. It may happen to face frequent resistance from local homeowners even when hydropower projects, wind turbines and photovoltaic farms are commissioned. There are often tensions among citizens due to their unwillingness to accept the construction of large-scale projects by corporations or governmental entities nearby (the so-called Not in My Backyard phenomenon), which, they claim, may affect their quality of life and the value of their property. The main negative externalities associated with different types of power plant are now briefly described.

2.2.1 Fossil fuel power plants

There are several local externalities from power plants that are important for households living in their immediate proximity.

First and foremost, pollution can cause birth defects and irreversible developmental and neurological disabilities, and immune system damage. It can also cause various cancers, heart and lung diseases, to name a few (WHO, 2019). Pollution mainly involves the emission of sulphur dioxide (SO_2), nitrogen oxides (NOx), and particulate matter (PM), which can cause significant contamination of air, water, and soil. While primary PM is mainly a local or regional problem, SO_2 and NOx can travel large distances, affecting downwind areas several kilometres away from the place of origin (Welsch, 2016). CO_2, SO_2, and NOx are generically related to fossil fuel power generation, especially coal, oil and gas, although the combustion of natural gas is a relatively clean technology in terms of both climate change and air pollution whereas coal is the most polluting. The European Environment Agency published in 2008 a report documenting and trying to quantify the emission of pollution produced by conventional power plants within the EU. The report concluded that, at that time, the NOx and SO_2 emissions from the large combustion plants could have been lowered by at least 60% if all existing plants used the best available techniques.

Even though they do not entail the combustion of fossil fuels, waste-to-energy incinerators may present even more varieties of pollutants, beyond CO_2, SO_2, NOx, and PM. Incinerators are indeed perceived as a health risk facility, due to toxic pollutants like dioxins (Kiel and McClain, 1995a).

Moreover, in most cases power plants are big facilities that can be seen from a distance because of their tall stacks. These visual disamenities are especially severe for particularly large plants.

Noise pollution from power plants is another issue, although available technologies can help mitigate this. Fossil fuel plants use giant machineries which generate high levels of noise and vibration. Similarly, natural gas plants utilize turbine engines, which can be extremely noisy. It is not uncommon for power plants to be detected and heard far away, especially when plants are being tested or cleaned.

Finally, another potential source of negative externalities is traffic from fuel deliveries, especially for coal-based plants, where coal typically arrives by train, truck, or barge (Davis, 2011).

2.2.2 Nuclear Power

The externalities from a nuclear power plant are related to the potential radiation coming from the normal operation of the plant and the risk of a nuclear disaster. Isotopes of types Iodine-131, Caesium-137, Strontium-90 and Plutonium-239, are all linked to increased risk of several kinds of cancer (Welsch, 2016). Although rare, failures can be very dangerous. Some accidents at nuclear power plants have taken place since nuclear power started to be used. Maybe two of the most common known accidents are the one in Chernobyl in 1986, and the accident after the tsunami in 2011 at the Fukushima site, in Japan. These disasters may affect people's perceived or subjective risks of nuclear plants and, through these, people's valuation of nuclear power externalities. Also, the waste from the production has to be taken care of in a proper way. The waste is still radioactive and therefore dangerous and nuclear waste sites are built and similar considerations can be applied.

2.2.3 Renewable Energy

Hydropower projects affect negatively the landscape, vegetation and wildlife as well as artifacts of cultural and historical value. Moreover, a large reservoir greatly increases evaporation, which alters the air humidity in a large area (Mattman et al. 2016).

Moving to wind power, public support for wind energy has increased over the years and there is more understanding and recognition of the benefits of renewable energy such as wind turbines. Public Attitudes Tracker, published in May 2020 by the UK Department for Business, Energy and Industrial Strategy (BEIS), shows that 82% of people surveyed support the use of renewable energy, while only 2% opposing it. Offshore wind’s popularity is high at 81%, while onshore wind has a support level of 77%. However, the unwillingness to have this kind of facilities in their neighbourhood is still very strong and the NIMBY (not-in-my-backyard) issue applies firmly. Indeed, wind turbines cast shadows, cause flickering, and visually pollute the landscape with their increasingly taller and larger structures. In addition, people living in the vicinity of wind turbines are at risk of being annoyed by the noise, which in turn could lead to sleep disturbance and psychological distress (Bakker et al. 2012).

At the same time, wind farms have potential local economic benefits of various types (Munday et al., 2011). These include the use of locally manufactured inputs and local labour, discounted electricity supplies, payments into community funds, sponsorship of local events, environmental enhancement projects, and tourism facilities.

Similar disamenities to wind turbines are present for ground-mounted solar panels, as they reflect ambient sound and sunlight causing glare risks, create a buzzing sound, and also affect landscape aesthetics because of their spatially very large units. The panels, once installed, can also alter local habitats and affect wildlife in negative ways (Lovich and Ennen, 2011).

Biogas plants affect residents through odour nuisance and noise caused by fuel deliveries (Soland et al. 2013). Moreover, despite a common conviction that the combustion of biomass is more environmentally friendly than the combustion of fossil fuels, this fact is not confirmed by scientific researches. Biomass either in a solid, liquid, or gaseous state, can also emit other pollutants and particulate matter into the air, including carbon monoxide, volatile organic compounds, and nitrogen oxides (IEA, 2019b).

3. Valuation of externalities from electricity generation

3.1 Conceptual Framework

Externalities arise when a real variable (not a price) chosen by one economic agent enters the utility or production function of other economic agents. Thus, if an electric generating plant is established and allegedly the smoke produced by it causes people to become ill, it is worth to assess its impact in order to guide policymaking to solve this kind of market failures.

As a general framework (Welsch and Ferreira, 2014), two basic underlying economic assumptions are needed:

1. Economic agents, when they have to choose between two or more bundles of goods (e.g. residential property ownership options), have preference for one over another;
2. Economic agents attempt to maximise overall satisfaction or utility in such choice behaviour.

Additionally, the Appraisal Foundation, which is the United States organisation responsible for setting standards for the real estate valuation profession, gives commonly accepted conditions when dealing with real estate data: both parties are well informed or well advised, and acting in what they consider their own best interests; the price is unaffected by special or creative financing or sales concessions granted by anyone associated with the sale or by undue stimulus.

Assume that decision makers are risk-neutral and derive their utility from income and housing and disutility from electricity related risks. The utility function is:

$$u = v(R, p, y, \theta)$$

where R is the person's perceived risk of externalities, p is the price of housing, y is the person's income, and θ is a vector of personal characteristics. Clearly, the utility function should be decreasing in the first two variables, and increasing in the third one.

Perceived risk can be conceptualized as the expected value of damage from the externality, that is $R = \pi * D$, where π is the (subjective) probability of that a damage will occur, and D is the actual (physical or psychological) damage caused by the externality. In the case of the opening of a new power plant, for example, D may be interpreted as the air pollution, visual, acoustic or odour nuisance (or, in the case of a nuclear plant, the expected damage associated with a nuclear accident), while π is the individual specific susceptibility to externalities produced by the power plant.

Being the damage distance-dependent, agents can choose their favourite house according to their subjective π. This means that more susceptible people choose to live in more distant places. This concept is formalised by setting π as a function of the distance from the power plant that people desire to live. Thus, $R = \pi(dist) * D(dist)$, and the previous equation can be rewritten as:

$$u = V(\theta, p, y, dist)$$

Hence, the derivative of V with respect to $dist$, $\partial v/\partial dist$, captures the externality in different places. This should be positive, as, after controlling for income and costs, utility increases with distance due to the externalities that are distance-dependent. Subsequently, in order to convert the marginal utility of distance into monetary terms, you divide it by the marginal utility of income. This is the marginal rate of substitution of income for distance, the utility-constant trade-off between the two and what the different valuation methods seek to measure:

$$MRS(y, dist) = \frac{(\partial v/\partial dist)}{(\partial v/\partial y)}$$

In this formulation, I considered for simplicity distance as a continuous variable. In the following analysis using UK data instead, I will consider distance as a "discrete" category, more specifically with radius rings around the facilities.

In the following sub-sections I briefly analyse the most common methods used to evaluate environmental goods and electricity-related externalities, with a particular focus on the hedonic price method, since it is the method used in this analysis.

3.2 Revealed Preference Methods

The revealed preference method involves determining the value that consumers hold for an environmental good by observing their purchase of goods in the market that directly or indirectly relate to environmental quality. Considering housing price, the willingness to pay for housing and, hence, the price of housing is a decreasing function of risk where houses are located: $p = p(R)$. With the same principle, also local wages (and thus income) increase in risk: $y = y(R)$. Therefore, the utility function becomes:

$$u = v(p(R), y(R), R, \theta),$$

where, as before, $R = R(dist)$. Agents choose their location such that the disutility from risk and the utility from less expensive housing and higher income are balanced. If this is not the case, individuals would rather move in a different location. Formally,

$$\frac{dv}{d\ dist} = \frac{\partial v}{\partial p}\frac{dp}{d\ dist} + \frac{\partial v}{\partial y}\frac{dy}{d\ dist} + \frac{\partial v}{\partial\ dist} = 0$$

Externalities from power plants are capitalised in housing prices and income such that the marginal disutility from risk in closer houses, is just offset by the marginal utility from lower housing prices in riskier and closer places. Symmetrically, the marginal utility from lower risk in more distant places equals the marginal disutility from higher housing prices and lower income in more distant (less risky) places.

The most widely used revealed preference technique is the hedonic method.

3.2.1 Hedonic Price Method

As mentioned above, if households evaluate the presence of power stations as a disamenity, this should be incorporated in the housing values and ultimately in their prices. Indeed, house prices have become a way to use in investigating values of environmental non-marketable goods.

As described in Freeman et al. (2014), one of the most used methods is the hedonic price method which treats a marketed good, usually a house, as a sum of individual characteristics or attributes that cannot be sold separately in the market. The main objective of a hedonic pricing method is to test whether or not households value nonmarket amenities such as air pollution or noise pollution.

In the hedonic method, a differentiated good is described by a vector of characteristics $(z_1, z_2, \ldots, z_n)$. For a house, these attributes include structural characteristics (different number of rooms, size), neighbourhood amenities (distance to a city, schools and so on), local environmental quality, and other factors. All these characteristics affect the price, determining the so called hedonic price schedule:

$$P(z) = P(z_1, z_2, \ldots, z_n) \tag{1}$$

This relationship can be linear in a characteristic if repackaging of that characteristic is possible, if consumers are able to untie and repackage bundles of attributes. For example, if individuals are indifferent between owning two two-door cars and one four-door car, other things being equal, they can create equivalents of four-door cars by repackaging smaller units. If both sizes exist on the market, the larger size must sell at twice the price of the smaller one, and the hedonic price of a car will be a linear function of the number of doors. However, in general for houses this is not likely to be possible. For this reason, equation (1) has to be nonlinear, usually by considering the logarithm of the price, and not the price itself.

A consumer is assumed to always maximise utility, subject to her budget constraint. Utility functions within this framework are simplified as

$$u = u(x, z_1, z_2, \ldots, z_n),$$

where x represents all other goods (and held constant) and z_i is an attribute of a good y. The derivative of the hedonic price function $P(z)$, with respect to any of its arguments, $\partial P(z)/\partial z_i$, gives the implicit marginal price of that characteristic - that is, the additional amount that must be paid by any household to move to a housing bundle with a higher level of that characteristic, other things being equal. An individual maximizes utility by simultaneously moving along each marginal price schedule until she reaches a point where her marginal willingness to pay for an additional unit of that characteristic equals the marginal implicit price of that characteristic. If an individual is in equilibrium, the marginal implicit prices associated with the housing bundle actually chosen must be equal to the corresponding marginal willingness to pay for those characteristics. Therefore, the hedonic price of a characteristic can be interpreted as the willingness to pay of households for a marginal increase in that particular characteristic.

In the case of an externality affecting houses, if, for instance, a characteristic considered is a measure of environmental quality, say z_1, such as distance to a power plant, and this is valued by buyers, a low level of environmental quality determines lower prices of houses in locations with this attribute than equivalent houses in other locations in order to attract households to these locations.

Typically the second step of the hedonic method combines the quantity and implicit price information, with restrictions on the structure of preferences and it estimates the willingness to pay function for the characteristic z_i. It thus additionally accounts for different incomes and tastes (usually based on age, race, social background, family size etc). For example, a decrease in environmental quality will lead to a change in the composition of local neighbourhoods. Consider a decline in environmental quality from z_1 to z_1' as indicated in Figure 2 and that the preferences of buyers are represented by the bid function θ. In the figure two different households are depicted with two different bid functions, θ_1 and θ_2, while $p(z_1)$ is the hedonic price schedule for the characteristic accounting for environmental quality. After the decrease of z_1 to z_1', type 1 households find themselves in a neighbourhood with a sub optimally low level of environmental quality and these households can increase utility by moving to a neighbourhood with the level of quality that they selected originally. After the decrease, households in the neighbourhood will be those with relatively weaker tastes for environmental quality. Therefore, if environmental quality is a normal good then household income is likely to decline after a reduction in environmental quality (Davis, 2011).

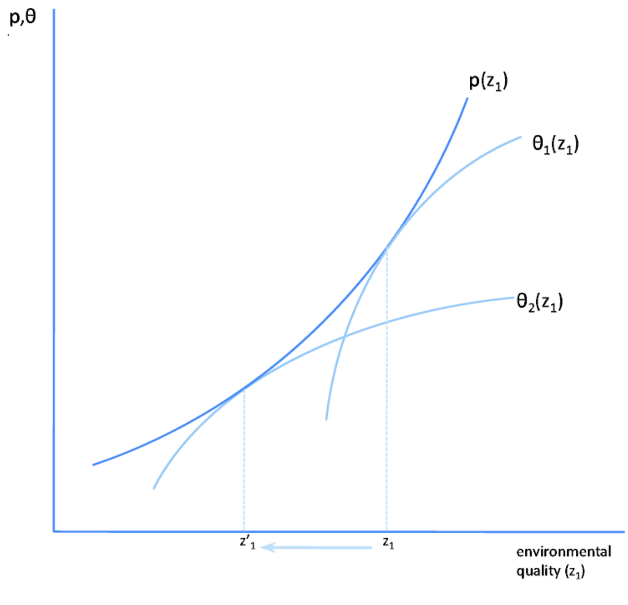


Figure 2: A Hedonic Market for Environmental Quality. Source: Davis (2011)

To conclude, in a classic hedonic price method the dependent variable is often the sale price of residential houses. In some papers other variables such as tax assessor, are used for the dependent variable, due to the difficulties of collecting sale prices. However, this could lead to a non-correlation of the independent variables and does not correspond to the same extent. Plus, it is good to have as many variables as possible that can explain variation, but it is not always feasible to get data for all factors that explain variation. In addition, factors that do not affect the price should not be included, there could be characteristics that does vary between products but does not influence price. Finally, most papers divide characteristics into three different categories: actual characteristics of the house, of the neighbourhood and distance to a recreation area.

3.3 Stated Preference Methods

While revealed preference methods rely on real choices (of houses or jobs), stated preference methods establish hypothetical choice situations and they are particularly used in the transport sector, quantitative marketing and environmental economics. They use individual respondents' statements about preferences in a set of options in order to estimate their utility functions. The options are typically descriptions of situations or context constructed by the researcher (Kroes and Sheldon, 1988). One of the most used method is the contingent valuation, that is a survey-based technique for the valuation of non-market resources, such as environmental preservation or the impact of externalities like pollution. The respondents are typically asked to state the amount of money they would pay to avoid, or request to accept, certain hypothetical risks or disamenities.

Contingent valuation is now widely accepted as a real estate appraisal technique, particularly in contaminated property (i.e. residences in or near areas on the National Priorities List, that is the priority list of hazardous waste sites in the United States eligible for long-term remedial investigation and remedial action) or other situations where revealed preference methods fail due to disequilibrium in the market (Mundy and McLean, 1998).

A general advantage of stated preference methods in comparison to other approaches is their independence of real-world observations and, hence, their ability of valuing hypothetical or actual options ex ante. This is especially relevant for technologies that are not yet widespread and real observations and data are difficult to be collected. On the other hand, contingent valuation method (CVM) cannot be considered an appropriate approach to value of real estate unless in specific situations involving special purpose or limited market properties for which there are few housing transactions that can be analysed. According to Roddewig and Frey (2006), the inaccuracy of the real estate values' prediction based on hypothetical surveys depends on several reasons: the CVM questionnaires provide less information about the good than the ones generally available in a real market; the opinions of both sellers and buyers are likely to be affected by the intermediaries that manage the transaction; the survey format usually does not include those factors generally affecting real estate purchase and sale decisions as urban context's characters. Furthermore, contingent valuation technique's surveys only consider one side of transaction, either buyer or seller. Finally, the model disregards that the price is the result of a negotiation process between the two subjects involved, buyers and sellers, who in most cases agree on a level of price that is a compromise as to their initial positions. In addition, CVM can generate a series of biases:

- Hypothetical market bias: there are no incentives to avoid overestimated or underestimated responses since respondents are aware it is not a real market;
- Strategic bias: if respondents believe that their response will affect the determination of the price or costs actually applied, they can over or underestimate their real willingness to pay;
- Embedding effect: it arises because of the difficulty for individuals of identifying the particular value that they attach to one particular good which is embedded in a collection of similar goods.

In short, contingent valuation technique does not usually respect some assumptions I listed at the beginning of this chapter: that the price is totally unaffected by external factors and that agents always seek to maximise their utility functions.

To conclude, revealed preference methods and stated preference methods have both inherent weaknesses. In order to exploit the strengths of the different approaches while minimising their weaknesses, combining the two methods has become an important methodological option, even in valuing environmental amenities (Whitehead et al. 2008). The amount of information increases, and findings can be cross-validated. On the one hand, use of revealed preference data ensures that estimation is based on observed behaviour. On the other hand, inclusion of stated preference responses to hypothetical changes enables identification of parameters that otherwise would not be identified. Revealed preference methods generate cross-sectional or panel data about the observed, current choices of consumers facing real market constraints, while stated preference data record the options the same consumers might choose at another point in time.

After this brief introduction to the principal methods to evaluate externalities from electricity generation, I next review the relevant literature that uses stated preference methods and the hedonic models to assess externalities caused by energy facilities on housing values. As it will be explained more deeply in next sections, given the nature of my data, real housing sale prices, my study belongs to the framework of an hedonic model, and seeks to give an idea of how the disamenity of having a power plant affects prices of houses, and how households are willing to pay to avoid it. In particular, my study is based on a difference-in-differences design implemented via fixed effects regression. Kuminoff, Parmeter and Pope (2010) examine the advantages of quasi-experimental approaches of this type in the context of hedonic methods for environmental valuation. They provide evidence that large gains in accuracy can be realised by moving from the standard linear specifications for the price function to a framework that uses a combination of spatial fixed effects, quasi-experimental identification, and temporal controls for housing market adjustment.

3.4 Literature Review

There is an extensive literature using especially the hedonic method but also stated preference method that examines the impact of locally undesirable facilities on housing values. The list that will follow does not intend to be exhaustive.

Davis (2011) studies 92 power plants opened in the U.S. during the 1990s. The research is the first large-scale effort to assess the impact of power plants on local housing markets. He considers only those plants that had capacities greater than 100 MW, 85 of them being gas fired. Compared to neighbourhoods with similar housing and demographic characteristics, neighbourhoods within 2 miles of plants experienced 3 to 7 percent decline in housing values, which corresponds to about $2770 to $6470 lower than average house prices. There is evidence of even larger decreases within 1 mile. Instead, beyond 4 miles, no effects on housing values are found. The small number of coal plant openings in the sample do not allow significant differentiation between coal plants and gas plants. A moderately bigger effect is found for large plants (capacity > 380 megawatts) for housing value, though the differences are not statistically significant.

Muehlenbachs et al. (2015), by analysing a very long panel of property transactions spanning 1995 to 2012 from Pennsylvania, studies the home value impacts from nearby shale gas development. Shale gas has become an increasingly important source of natural gas in the United States. Shale gas development involves constructing a well to drill into the ground vertically and horizontally to reach the shale rock layer. Properties surrounding shale gas wells can face growth or decline in value depending on whether the benefits of the activity outweigh the costs. They demonstrate that groundwater-dependent homes are negatively affected by nearby shale gas development, indicating that the risk of groundwater contamination has indeed materialized into a real impact. Similarly, proximate homes that have access to publicly supplied piped water, on the other hand, appear to receive small benefits from that development.

Kiel and McClain (1995a, 1995b) investigate price-distance effects of a municipal solid waste-to-energy incinerator for several stages of the siting process from rumour, construction phase, online stage and finally the fully operation phase. No significant price change is found prior to construction, while adverse effects for the subsequent stages appear. In addition, housing appreciation rates were 2 to 3.5 percent lower during these phases in the impacted neighbourhood compared to non-impacted neighbourhoods.

A large literature has estimated the impact of nuclear power plants on house prices. Clark et al. (1997), using an hedonic method mixed with geographic information system, estimate housing price near two nuclear power plants in California. Surprisingly and in opposition to most papers that observe property prices to be lower at greater proximity to NPPs, they find that house nuclear power plants or stored nuclear waste do not have a significant detrimental effect on residential home prices but in fact, especially for one of the two plants, home prices rise with proximity to the plant. This, they argue, may be due to the uncongested nature of the area around the plant and the recreational opportunities nearby. These amenities overcome the possible detrimental impacts from the plant and the nuclear waste. Clark and Allison (1999) instead, analyse the effect of the spent nuclear fuel storage, that is the nuclear fuel that has been irradiated in a nuclear reactor and stored in structures of water pools, usually built near a no longer active NPP. They discover that the proximity and visual reminders of the plant have some influence on local property markets, and that this effect weakened over time, possibly due to relocation, or to preference adaptation attenuating the initial price decrease.

A number of papers were concerned with effects of nuclear accidents on houses in the proximity of nuclear power plants. Fink and Stratmann (2015) for the USA, Coulomb and Zylberberg (2020) for the UK and Bauer et al. (2017) for Germany, analyse an external shock, the Fukushima nuclear accident in March 2011, that did not in any physical way affect individual houses in the area considered for the analysis. The research in USA do not find evidence in support of the hypothesis that individuals reappraised the risks associated with nuclear power plants. On the other hand, Coulomb and Zylberberg (2020) and Bauer et al. (2017) find that the Fukushima accident reduced housing prices near nuclear power plants that were in operation before Fukushima by 4.9% and 4.2%, respectively, albeit the latter is also related to the expected closure or phasing down of nuclear plants and the associated employment effects.

Several studies have analysed the effect of wind turbines on property values, with conflicting results. Hoen et al. (2011) apply cross-sectional hedonic analysis, based on 24 wind farms across US states. They make the comparison between price effects at places where turbines are visible compared to places where turbines are non-visible, finding no differences. A very similar approach is used by Gibbons (2015), with data from the UK, but including as control groups places close to wind farms that became visible in the past, or where they will become operational in the future and places close to wind farms sites but where the turbines are hidden by the terrain. He uses a quasi-experimental difference in differences method, with postcode fixed effects methodology and found that the housing price reduction is around 5-6% within 2 km, falling to less than 2% between 2 and 4 km, and less than 1% by 14 km which is at the limit of likely visibility, with an even bigger effect for larger farms, while no impact is found beyond 4 km for the small ones.

Studies on renewables other than wind power are much fewer in number. Koster and Droes (2020) examine the effects of wind turbines and solar farms on residential property values using a housing transactions dataset covering the Netherlands since 1985. Making use of difference in difference methodology, they compare changes to house prices in areas that will receive a turbine or solar farm in the future to areas in which a turbine or a solar farm has already been built. The house prices within 2 km from a wind turbine decrease on average by 2% relative to comparable properties. Moreover, tall wind turbines (>150 m height) determine a negative price effect of about 5% within 2 km, while no significant effects for turbines below 50 m are found. An other interesting finding, which contrasts with Gibbons (2015), was that only the first turbine within 2 km has an effect on property prices, implying that, from a policy standpoint, it would be better to cluster the turbines in a large wind farm rather than building scattered turbines. As far as solar farms are concerned, the results are less convincing because the number of solar farms is much lower, making the estimated coefficients less precise. The presence of solar farms led to a house price decrease of about 2-3%. The effect is more localized than the effect of turbines and confined to 1 km.

Using geospatial analysis and a state preference method's survey to residential property assessors, Al-Hamoodah et al. (2018) seek to assess whether utility-scale solar facilities is a potential amenity or disamenity in the US. They find that very few homes are located around utility-scale solar installations and therefore few properties could be potentially affected. However, the survey of residential home assessors indicate that the majority of respondents believe that proximity to a solar installation has either no impact or a positive impact on property values. Conversely, regression analyses suggest that closer proximity to an installation is associated with more negative estimates of property value impacts, as well as larger installation size.

Lipscomb (2011) examines the effects of a proposed biomass facility on prospective property values in the Midwestern United States using the contingent valuation method. He uses a web-based survey presenting two different scenarios: i) the status quo of the area of study and (ii) the proposal of a biomass facility within one-half a mile, highlighting the stack height of the biomass plant (265 feet) as a major attribute. The paper found no statistically significant difference in respondents' willingness to pay for a house based on whether they read about the existing conditions (no mention of the biomass plant) or the proposed biomass plant.

To conclude, two works analyse the willingness to pay of consumers for several sources of energy, by using stated preferences methods. Borchers et al. (2007), focusing on green energy options, report a choice experiment involving the development of several types of renewable electricity programs (wind, solar, farm methane, and biomass) in one county in Delaware, USA. They find positive mean WTP for all sources, except for biomass in one of their treatments. Furthermore, individuals have a preference for solar energy over a generic green and wind, while biomass and farm methane are found to be the least preferred sources.

Cicia et al. (2012) report a latent class choice experiment in Italy that involves not only three types of renewable energy (wind, solar and agricultural biomass), but also fossil fuel electricity and nuclear power. They find a positive WTP for replacing fossil fuel electricity by wind and solar power and a negative WTP for replacing it by nuclear power and electricity from biomass. Latent class analysis provides some heterogeneity in terms of household socio-economic characteristics with respect to whether solar or wind power was most preferred and whether nuclear or biomass was least preferred.

4. Data and Empirical Strategy

4.1 Data

Housing data comes from the Land Registry Price Paid transactions data, from January 1995 to April 2021, and it refers to England and Wales. Under the Land Registration Act 2002 and the Land Registration Rules 2003, Land Registry registers all sales and changes in ownership rights (mortgage, lease or right of way) in England and Wales. It is open to public and freely available for commercial and non-commercial reuse purposes. The 'price paid' data include information on sales price, the address including postcode, the date when the sale was completed, property types – detached, semi-detached, terraced, flat/maisonette, or others – whether the property is a newly built property or an established residential building, and whether it is sold on freehold or leasehold basis. The housing transactions were geocoded using the address postcode and aggregated to mean values in postcode-by-day, to create an unbalanced panel of postcodes observed at daily intervals, with gaps in the series for a postcode when there are no transactions in a given day.

Information on power plants was provided by Digest of UK Energy Statistics from the Department for Business, Energy & Industrial Strategy. It contains characteristics of all power stations in the United Kingdom as at end of May 2020: the name of the producer, station name, the type of power plant (the technology used), installed capacity, the year of commission or year generation began and the location. Unfortunately, these public data provided by the government do not provide a complete record of the history for a given facility, because there is only one date (the year to be precise) for each power station. Therefore, for operational sites, either the date of commencement of operation or the commissioning year is known, but not the date when planning applications were submitted, approved or when construction began. This shortage of information limits the scope of investigation of the impact of different events and stages in the planning and operation process, and eventually can affect the validity of the analysis.

I selected only those sites with > 20 MW capacity, whose year of opening was ≥ 1996 (since the housing dataset starts from 1995 and only the years of power plants are known), and located only in England and Wales due to housing data limitations. Also, existing facilities that increase the number of generators on site and plants that change their primary energy source (e.g. switch from coal to natural gas or becoming combined cycle plants) are excluded. Changes in capacity, technologies and emissions may influence the local effect of power plants, but including these in the analysis would make the results complicated to interpret. Furthermore, reorganisations of plants often occur simultaneously with other modifications at the plant, making the interpretation of results even more difficult. More specifically as an example, the Drax Power Station, whose capacity is the highest of any power station in the United Kingdom, providing about 6% of the UK's electricity supply, is excluded as the original power plant was constructed in 1967 and several conversions from coal-fired to CCGT and biomass occurred over the years. Similar to this, other plants have been converted from coal-fired based to natural gas or CCGT, in accordance with the government's 2025 deadline to ban all coal-fired electricity in the United Kingdom.

I then geolocated (latitude, longitude) the power stations and excluded those whose location was unclear or unsure. In order to make sure that what I found was the exact and precise location of facilities, I matched my findings with the coordinates provided by the global power plant database (Global Energy Observatory, 2018) and check whether or not locations coincided. I further selected stations whose technology is: Wind (onshore), Solar, Bioenergy, Conventional Steam, CCGT (Combined Cycle Gas Turbine), OCGT (Open Cycle Gas Turbine). As far as bioenergy plants are concerned, this definition includes biomass power plants (wood pellets, sunflower/oat husk pellets or straw, litter or woodchip) or waste (municipal solid waste). The final list contains 86 power stations: 34 Wind, 23 CCGT, 19 Solar, 8 Bioenergy, 2 Conventional Steam and 1 OCGT.

Finally, I found all the postcodes (and their distances from the plant) inside the radius of each geographic coordinates of the power plants using the online resource freemaptools.com, and linked those to the plants characteristics. The limit of 15 km is set to keep the dataset at a manageable size. Ultimately, the housing transactions and plants data (with postcodes within the radius) were linked by postcode to create an end product which is an unbalanced panel, with information on housing prices and characteristics, and information about the power station nearby.

Figure 3 shows the location of the power plants included in the sample. By looking at the map, it is clear that especially for wind farms, many sites are very close to one another. For instance, in southern Wales in a radius of approximately 10 kilometres eight wind farms are operating; in the English region of Yorkshire and Humber, five wind farms are located in a range of 9 kilometres. This is likely to occur because the wind, altitude and natural conditions are particularly favourable in specific areas. The majority of solar farms are located in the Cambridgeshire area and South East England due to the higher amounts of sunlight available. Combined-cycle power plants and bioenergy sites are quite sparse but particularly present in the north of England, as many coal-based power plants built in those areas full of coal during the XIX and XX centuries have been demolished or reconverted and later new power station have been constructed in the same site. In addition, some plants are part of a bigger site that produce energy also using different fuels and technology and therefore are located near to each other.

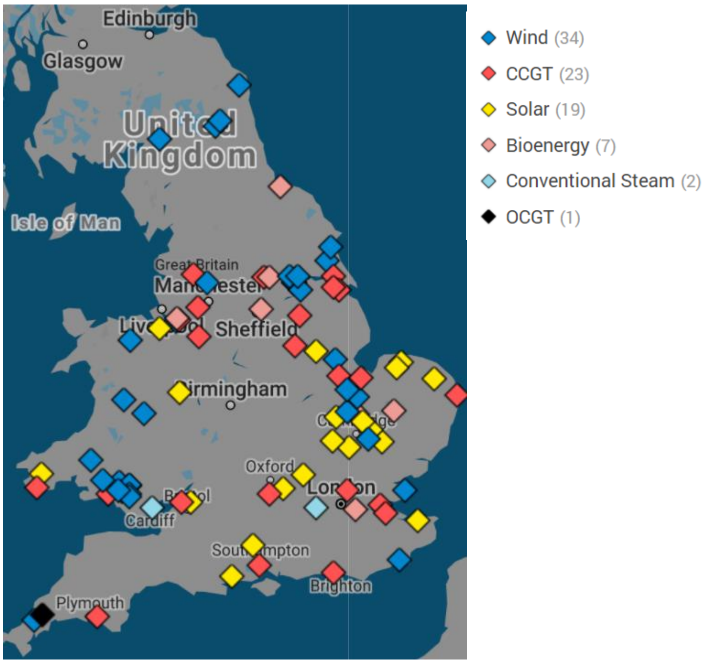


Figure 3. Map of selected power plants in England and Wales. Note that many other power plants are absent due to restrictions to the dataset.

4.2 Difference in differences Framework and Assumptions

I now explain the technique used in this dissertation to study the effect of power plants on local housing prices. I first denote the abstract formulation of the difference-in-differences technique within the potential outcome framework and then I briefly introduce the difference-in-differences design applied to the variables in my study. Finally, in the next sub-section I define the empirical strategy more precisely.

Conceptually, the difference-in-differences design takes the before-after difference in mean outcomes between treatment and control groups. This is the first difference. By firstly doing this, the first difference controls for factors that are constant over time in that group. Then, to capture time-varying factors, difference-in-differences takes the before-after difference in the control group, which was exposed to the same set of environmental conditions as the treatment group. This is the second difference. Finally, difference-in-differences cleans all time-varying factors from the first difference by subtracting the second difference from it. This shows the impact estimation that is the difference-in-differences.

More specifically, in a difference in differences estimation with only two periods, while my specification has multiple periods but the baseline idea is the same, we typically have (see Cunningham 2021):

- two periods: $t = B$ is before treatment; $t = A$ is after treatment ($B < A$);
- time treatment dummy: $\tau_t = 1$ if $t = A$;
- two groups: $r = 1$ is the treatment group, $r = 0$ the control group;
- treatment dummy $d_{rt} = r_t \tau_t$;
- y_{dt} : outcome of treated/untreated in time t. This outcome might be observed or not.

The treatment therefore is given if $d = 1$:

	t = B	**t = A**
r = 1	d = 0	d = 1
r = 0	d = 0	d = 0

What we can observe is thus:

	t = B	**t = A**
r = 1	y_{0B}	y_{1A}
r = 0	y_{0B}	y_{0A}

It is relevant to note that y_{0B} is in both treated and control groups and that we cannot observe y_{0A} in $r = 1$. Within this notation, the equation for diff-in-diff is:

$$DD = E(y_{1A} - y_{0B} | r = 1) - E(y_{0A} - y_{0B} | r = 0)$$

There are two key assumptions with a standard DID model. First, there is no time-variant group specific unobservables, namely there is nothing unobserved in the treated group that is changing over time that also determines the outcome. The second key assumption of difference in differences estimate, the parallel trends assumption, is that the change in outcomes from pre- to post-intervention in the control group is a good proxy for the counterfactual change in untreated potential outcomes in the treated group, more formally:

$$E(y_{0A} - y_{0B} | r_A = 1) = E(y_{0A} - y_{0B} | r_A = 0)$$

The left-hand side expectation in the equation is a counterfactual object because y_{0A} is unobserved among treated units which have $r_A = 1$. Whereas the right-hand side is completely observed. In other words, in the counterfactual state in which treated units had not been treated, their mean change in outcomes between the before and after times should be the same as the observed mean change in outcomes among the control units.

Coming back to my analysis, I seek to estimate the effect of the openings of power plants on house prices. In a DID framework, this means that my outcome y is housing prices. I define my treatment and control groups as follows: a housing transaction is in the treated group ($r = 1$) if the house is within three kilometres of where the power plant opened between 1995 and 2021; a housing transaction is in the control group ($r = 0$) if the house is over three kilometres away of where the power plant opened between 1995 and 2021. I rewrite the DID estimate:

$$DD = E(y_{1A} - y_{0B} | r = 1) - E(y_{0A} - y_{0B} | r = 0)$$

Thus, y_{1A} in treated group ($r = 1$) represents the price of houses within 3 km after the opening of power plant; y_{0B} is the price of houses within 3 km before the opening. This first difference is the before-after the opening of power plant for houses within 3 kilometres. Similarly, y_{0A} in control group ($r = 0$) is the price of houses more than 3 km away after the opening; y_{0B} is the price of houses more than 3 km away before the opening. This difference instead represents the before-after the opening of power plant for houses more than 3 km away. The difference between these two differences is the average treatment effect that is the DID estimate and it amounts to comparing the change in price in close houses to the change in prices in houses further away from the energy facility.

Since I cannot empirically observe prices in treated houses in the counterfactual world in which the power plant never became operational and check whether their mean change between the before and after times are the same as the observed mean change among the control houses, the parallel trends assumption must hold. However, this is not something that can be empirically tested. For this reason, what can be done is to test parallel pre-trends in which in the pre-intervention period, time trends in the outcome are the same in treated and control units.

Firstly, visual evidence may be helpful, that is plotting the mean outcomes over time for control and treated groups and then checking whether the two lines appear to be approximately parallel. In my case though the moment when the facilities become operational is clearly not the same for every observation because the 86 different facilities opened in different years. For this reason, I only plot certain facilities for illustration. I plot the following facilities: a CCGT plant, called Spalding, commissioned in 2004 (the red horizontal line), a bioenergy plant, Runcorn EfW, commissioned in 2014 and a solar park, Eveley, commissioned in 2016 are displayed in figures 4, 5 and 6. In such cases, the annual means are estimated and when year-to-year volatility is relatively low, it is simple to spot deviations from the common pre-trends assumption in a sufficiently long time series. However, it may be difficult to distinguish between statistical noise (especially when in the treated group there are few observations as in the case of Eveley solar park) and genuine deviations from the common pre-trends.

Figure 4: Spalding power plant

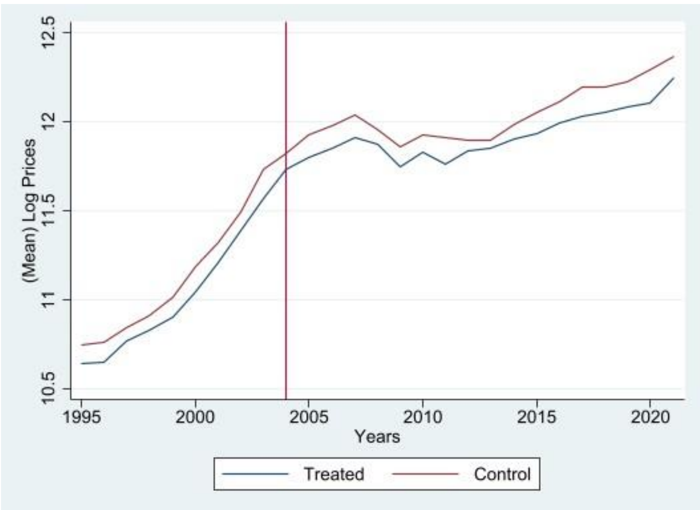


Figure 5: Runcorn EfW power plant

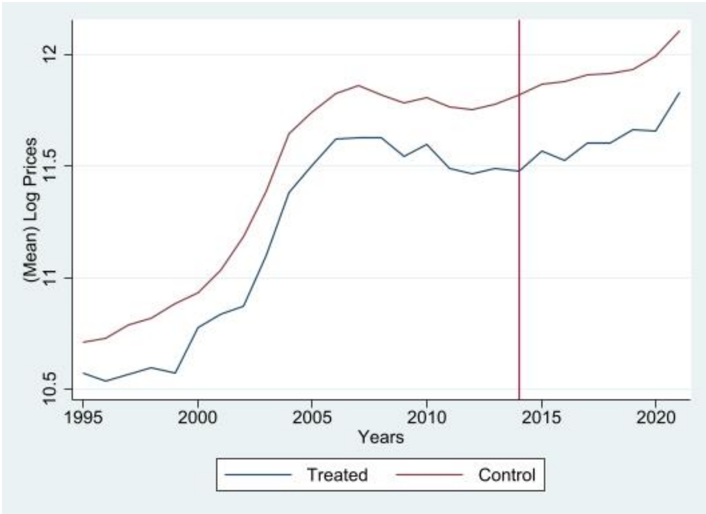
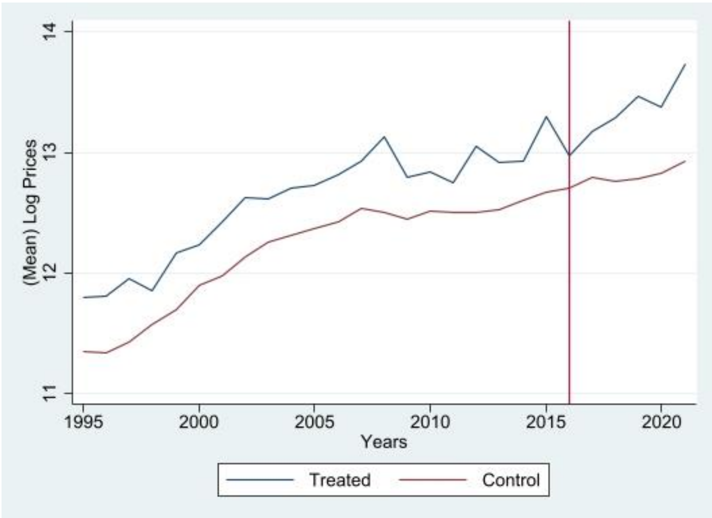


Figure 6: Eveley solar park



Other strategies for evaluating common pre-trends between treated and control units would be performing a Wald test to assess whether the linear trends are parallel prior to treatment. In order to do this, I would augment the diff-in-diff model and add another coefficient that captures the differences in slopes between treatment group and control group in pre-treatment periods. Another test is the so called Granger-type causality test, that can be useful to test whether future treatment exposures are anticipated by current outcomes. However, both tests require balanced panel data, while my data are unbalanced and therefore not applicable.

What can be done instead to empirically validate the DID framework and assumptions are some placebo tests. In particular, I will use “pseudo-outcomes”, i.e. some housing characteristics, which theoretically should not be affected by the treatment variable and set them as dependent variable and repeat the DID analysis. I would expect zero effect from this.

Finally, as already specified in the empirical strategy section, in all my regressions I will add covariates, also because they can account for factors that may capture differing trends across treated and untreated units over time. I will also estimate DID regressions by using different sub-samples and try computing DID with and without covariates. I would expect stability of estimated effects across such perturbations. More details can be found in the robustness checks section.

4.3 Empirical strategy

The research design is difference-in-differences (DID), which I implement using linear regression with fixed effects. The reason why I choose a regression framework rather than a simple differencing is mainly because I can control for other variables which may reduce the residual variance (and therefore smaller standard errors). Moreover, the fixed effects should control for all time-invariant characteristics of postcodes. In this sense, if houses in a given neighbourhood are by themselves more expensive compared to other neighbourhoods or postcodes (due to income, education, population, share of young adults and so on), this problem should be partially overcome by the postcode fixed effects.

Thus, the research strategy is to compare the average change in housing prices in areas where and when power plants become operational, with the average change in housing prices in some other comparator groups. It is a quasi-experimental design as I would like to establish a cause-and-effect relationship between the independent (the presence of facilities) and dependent variable (the housing prices). However, unlike a real experiment, a quasi-experiment does not rely on random assignment. Instead, subjects are assigned to groups based on non-random criteria. The unit of observation is indeed a transaction, which can be geolocated at the postcode level and I define the spatial treatment at the postcode level. As previously mentioned in the last sub-section, a generic transaction i will be treated if the property involved in the transaction is located within a treated postcode. A generic postcode p is defined as treated (T_p = 1) if the distance between the postcode and the nearest power plant is lower than (or equal to) 3 kilometres. A postcode will be instead in the control group (T_p = 0) if the median distance between the postcode and the nearest power plant is between 3 and

15 kilometres. It happens that a postcode p may be within 15 kilometres of more than one power plant, especially because some of them are located one next to the other. In this case the transaction is taken singularly.

The idea behind the choice to handle treatment and control group in this way is that there should be homes that are not affected by the opening of a power plant, but close enough to hold characteristics similar to the treatment group. In this sense, the threshold chosen (3 kilometres) has been set also by looking at similar works in this field: Davis (2011) which analyses traditional power plant set the threshold at 2 miles (3.2 km), while Gibbons (2015) and Lang et al. (2014) studying the effect of wind farms on prices, set the point at 2-4 km and 1 mile (2.25 km) respectively. In any case, I try different specifications to check whether the results are robust to this particular choice of the threshold.

I set January 1995–year of commissioning as the pre-treatment period (F_t = 0 where t is a year), and year of commissioning–April 2021 as the post-treatment period (F_t= 1). The main empirical strategy is a difference-in-differences specification,

$$\ln price_{pt} = \beta(T_p \times F_t) + \zeta F_t + \gamma \boldsymbol{X}_{pt} + \delta year\ dummy_t + v_p + \varepsilon_{pt} \qquad (2)$$

where β is the main coefficient of interest, while $\ln price_{it}$, the dependent variable, is the log mean housing transaction price in postcode p in day t. T_p is a dummy (1-0) treatment variable, indicating that postcode p is within at least one operational power plant 3 km distant. The variable $\boldsymbol{X}_{pt}$ is a set of control variables, including housing characteristics: whether the property is new; whether it is a flat, a terrace house, a semi-detached house, or a detached house; whether it is sold on a freehold or leasehold basis. Since data are aggregated at the postcode level, I take averages of these household-level characteristics across houses located in the same postcode. $year\ dummy_t$ represents the full set of year dummies (from 1995 to 2021) which controls for factors changing each year that are common to all houses (treated and control) for a given year, for example country-wide changes in housing prices over time. Finally, in this baseline specification I add postcode fixed effects, v_p, that captures all unobserved, time-constant factors that affect $\ln price_{it}$. The fact that v_p has no t subscript means that it does not change over time. The treated variable is not included here because its information is captured by the fixed effects themselves and it would be collinear with them, so the interaction term alone can still represent the intervention effect.

The estimates of β from the within-postcode fixed effects estimator should be interpreted as the average price change between the pre- and post-operation periods, given the time spanned by the housing transaction data. If there are price changes occurring prior to operation, or continuing after operation, this parameter will not coincide with price change occurring exactly around the time of operation. Because of the panel structure of data, it is preferred to estimate the difference-in-differences estimator using linear regression, instead of computing the double difference. More specifically, I use a fixed effects regression which is more appropriate for panel data, opposed to the case with repeated cross sections. In fact, it is unlikely that a step-change in prices happens coincidentally with plants operation because housing price changes evolve slowly. In addition, the panel is sparse and unbalanced, with missing periods where there are no transactions in some postcodes, so working with differences over specific time intervals within postcodes would result in a large reduction in sample size.

I also consider different alternative specifications to the baseline specification (2). I analyse alternative assignments of transactions to treatment and control groups, by using alternative bandwidths or considering buffer zones between treated and control areas. Finally, I estimate the treatment coefficient also by the size of facilities (in terms of capacity).

5. Results

5.1 Descriptive statistics

Table 2: Descriptive statistics in 2008

	All	$T_p = 1$	$T_p = 0$
	Total		
Observations	219,931	9,440	210,491
Average price	207,878	179,639	209,145
	Wind		
Observations	41,291	293	40,998
Average price	155,261	163,876	155,200
	CCGT		
Observations	112,225	6,044	106,181
Average price	201,155	168,083	203,037
	Solar		
Observations	41,723	1,660	40,063
Average price	227,816	238,997	227,353
	Bioenergy		
Observations	51,599	1,882	49,717
Average price	192,834	135,878	194,990
	Conv. Steam		
Observations	20,572	967	19,605
Average price	286,752	180,851	291,976
	OCGT		
Observations	1,344	51	1,344
Average price	202,275	162,692	203,777

Note: The unit of observation is a housing transaction. Transaction prices are expressed in pounds. T_p is equal to 1 for all transacted properties located in postcode p within 3 km of the energy facility, and to 0 for all transacted properties located between 3 and 15 km of an energy facility.

Table shows the statistics for transactions in total and in treated and control postcodes in an example year (2008). There are few differences between transactions of properties located in the two different locations. First of all, the number of observations between control and treated group is not commensurate: the observations in control group accounts for 95.7% of the total dataset (versus 4.3% of treated). The average prices in houses located closer to power stations are markedly lower, except for wind and solar farms where the average price in the treated group is slightly higher. For the latter case, this may be due to statistical noise and this is particularly true for wind farms, since the treated observations accounts for 0.71% of the total. On the other hand, homes near traditional power stations sold for less than homes farther away from the sites. This may occur due to the fact that typically plants are located in places where houses (if at all present) have lower value on average.

Table 3 displays the number of postcodes at different distances from power plants. Note that it refers to the average amount of postcodes per power plant since each type of facilities has a different amount of sites. It is evident that there are very few postcodes within 1 km, 2 km and partially also 3 km and this is particularly true for wind farms and solar parks, which are established in more remote locations and are less close to urban centres than the other types of power plants. This means that less properties can be potentially affected by the presence of wind farms and solar parks with respect to the other power plants. Taking into consideration all power plants, overall there are around 21 repeat observations for each postcode (=8,894,754/431,897 from the table).

Table 3: average amount of postcodes within *x* km per power plant and in parenthesis the total amount

	Total	**Wind**	**CCGT**	**Solar**	**Bioenergy**	**Steam**	**OCGT**
Postcodes within 1 km	12 (1,011)	1 (37)	32 (732)	5 (97)	26 (205)	49 (98)	2 (2)
Postcodes within 2 km	77 (6,651)	4 (143)	204 (4,684)	50.9 (967)	166 (1,330)	285 (569)	66 (66)
Postcodes within 3 km	201 (17,243)	24 (804)	485 (11,151)	159 (3,024)	442 (3,537)	787 (1,573)	171 (171)
Postcodes within 4 km	380 (32,708)	73 (2,479)	801 (18,415)	325 (6,168)	914 (7,313)	1,540 (3,080)	239 (239)
Postcodes within 8 km	1572 (135,193)	470 (15,972)	3,014 (69,321)	1,480 (28,125)	3,836 (30,691)	6,780 (13,560)	705 (705)
Postcodes within 15 km	5,022 (431,897)	2,569 (87,348)	9,460 (217,588)	4,600 (87,391)	12,064 (96,508)	18,979 (37,958)	3,397 (3,397)

Analysing the size of facilities, the biggest plants in term of capacity are manifestly those whose technology used is CCGT (see Table 4). The mean capacity for CCGT plants indeed is 897 MW, compared to solar parks, wind farms and bioenergy plants that have a mean capacity of 32 MW, 40 MW and 57 MW, respectively. That means that solar parks and wind farms have the least capacity. Related to that, the largest wind farms in the UK are off-shore. However, offshore sites are excluded from the research for obvious reasons, even though some of them are close to the British coasts and potentially visible from residential areas on shore but it would have complicated by a lot the analysis.

Table 4: Power plants capacity summary data 1995-2021 England and Wales

MW Capacity	**Number of plants**	**Mean**	**Standard Errors**	**Min**	**Max**
Overall	86	269.8	50.89	20	2,199
Wind	34	39.79	6.21	20	228
CCGT	23	897.43	113.27	56	2,199
Solar	19	31.78	3.11	21	72.2
Bioenergy	8	56.68	9.39	33.3	91
Conventional Steam	2	33.5	3.5	30	37
OCGT	1	140	-	140	140

5.2 Baseline results

OVERALL FINDINGS – I first assess the average price response after the construction of power plants, with the specification within 3 km and with no differentiations between different types. In Table 5, the estimates of specification (2) over the period January 1995 – April 2021 for the log price are showed. I have estimated four different variations of specification (2). Column 1 shows estimates of the "basic" difference-in-differences specification without postcode and time fixed effects and without additional controls other than the post construction time dummy variable, the treatment dummy variable and the interaction term between them, DID in the table. In columns 2 and 3, control variables and year fixed effects are added, respectively, whilst in column 4 fixed effects are further inserted. Hereafter, the R-squared presented is the within R-squared as it is more appropriate for the analysis. It represents how much of the variation in the (log) price within postcodes is captured by the model.

Table 5: Effect of power plants on (log) prices

	(1)	(2)	(3)	(4)
DID	-.0645***	-.0595***	-.0770***	-.0205***
	(.0058)	(.0051)	(.0047)	(.0022)
Post construction dummy	.6425***	.641***	.0148***	.0171***
	(.0014)	(.0013)	(.0019)	(.0007)
Treatment dummy	-.0912***	-.098***	-.065***	
	(.0058)	(.0049)	(.004)	
Newly built		.111***	.139***	.124***
		(.0029)	(.0022)	(.001)
Semi-detached		-.436***	-.433***	-.246***
		(.0018)	(.0017)	(.00077)
Terraced		-.639***	-.625***	-.344***
		(.0024)	(.0023)	(.00099)
Flat		-.241***	-.189***	-.581***
		(.0039)	(.0038)	(.0023)
Other type		-.322***	-.481***	-.804***
		(.011)	(.01)	(.0086)
Leasehold		-.291***	-.348***	-.202***
		(.0031)	(.003)	(.0021)
Controls	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes
Postcode fixed effects	No	No	No	Yes
Observations	8,894,754	8,894,754	8,894,754	8,894,754
R-squared	0.152	0.25	0.452	0.703

Note: Robust standard errors, clustered at the postcode level are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.
Each cell displays the result of a separate regression (specification 2). The unit of observation is a transaction. All specifications include post-construction time dummy and zone dummies. I exclude detached houses which therefore constitute the reference group for semi-detached, terraced, flat and other type variables.

The price decrease in (1) is around 6% after the construction of facilities in the proximity of them, compared to houses further away from the sites. The within R-squared is quite low in (1) and becomes higher as postcode and year fixed effects are introduced, as well as control variables. When adding control variables (2), the DID coefficient remains negative with a very similar decrease (-5.95%) to the first specification. Additionally, when year dummies and subsequently fixed effects are added, the sign continues to be negative (7% decrease and around 2% price decrease in the third and fourth columns, respectively). In all specifications the p-value is lower than 0.001 and therefore coefficients are all significant at 0.1% level. In terms of the post construction dummy, the magnitude decreases dramatically once the year dummies are included in the model. In fact, the omitted year dummies made the post construction dummies extremely high, because in the model factors changing (and specifically increasing) each year that are common to all houses were not initially accounted for. Finally, the treatment dummy is omitted in column 4, as it is collinear with fixed effects.

Looking at controls, the housing characteristics variables, I did not include the dummy variable that takes 1 if the transaction involves a detached house (and 0 otherwise) which therefore constitute the reference group. Therefore, taking into consideration column 4, given the negative signs of these variables, the interpretation of semi-detached, terraced, flat and other variables should be as follows: all the previous types of houses have been sold to a less price than detached houses. The cheapest type is the category "other" followed by flats, terraced houses and semi-detached houses. Moreover, the newly built houses have been priced on average 12% more than non-newly built ones. Finally, the price of houses sold on a leasehold basis houses are 20% lower than the price of houses sold on a freehold basis.

I also estimate the number of transactions post-power plants (Table 6). In this case, the dependent variable is the cumulative sum of sales in postcode p in a year and the regression provides a test for changes in the rate of transactions between the before and after operation periods. For this reason, the observations are markedly lower, because each postcode has been aggregated to the mean for each year, and not at a daily basis as before.

Table 6: Effect of power plants on number of transactions

	(1)	(2)	(3)	(4)
Number of transactions	-.0928***	-.079***	-.089***	-.0898***
	(.014)	(.012)	(.0127)	(.0107)
Control variables	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes
Postcode fixed effects	No	No	No	Yes
Observations	4,677,312	4,677,312	4,677,312	4,677,312
R-squared	0.0025	0.153	0.167	0.17

Note: Standard errors are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.
Each cell displays the result of a separate regression (specification 2). The dependent variable is the number of transactions within a postcode and a year. Each cell displays the result of specification 2, collapsed at the level of a postcode × year. Only the difference-in-differences coefficient is displayed, i.e., the coefficient before the spatial treatment interacted with a post-construction dummy. All specifications include post-construction time dummy and zone dummy.

In all regressions, the DID coefficients is significant at 0.1% level. The magnitude and values in all coefficients remain around -0.09 in all cases. The negative sign of the coefficient means that the presence of energy facilities has reduced the number of transactions in their proximity. In particular, considering postcode fixed effects and the set of controls, the power plants, when constructed and operational, determined the reduction of 0.1 housing transactions each year compared to the period before within 3 km.

ROBUSTNESS CHECKS – Following these first regressions, I also ran postcodes fixed effect regressions using different specifications (Table 7): in (1) the treated group is unchanged (postcodes ≤ 3 km), while the control group includes only those postcodes far from the facility more than or equal to 10 km. In (2), (3), (4) instead, a postcode is treated if its distance from the power plant is lower than 2, 4 and 8 km respectively.

The price impact is about -1% within 2 km and significant at 1% level, increasing to a price decrease of about -1.6% within 4 km, and -0.5% within 8 km. When the control group changes and consider only houses located 10 km or more from the facilities, the effect is about -2.3% and still statistically significant.

Table 7: Effects with different settings

	(1) ≤ 3 km and control group ≥ 10 km	(2) ≤ 2 km	(3) ≤ 4 km	(4) ≤ 8 km
DID	-.0235*** (.00075)	-.0093** (.0033)	-.0161*** (.0017)	-.00494*** (.0010)
Postcode fixed effects	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	4,963,723	8,894,754	8,894,754	8,894,754
R-squared	0.69	0.703	0.703	0.703

Note: Robust standard errors, clustered at the postcode level are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.

The different types of facilities are very heterogeneous and very diverse from one another. Traditional and fossil fuels based power plants produce local pollution and they usually have bigger sites, while renewable energy facilities do not create polluting substances and are typically evaluate for the noises or odours they produce and the visual and aesthetical impact they have on the landscape. For this reason, after these first set of results, I next analyse the different types by themselves.

FINDINGS BY TYPE OF POWER PLANT – I thus evaluate separately all type of power plants according to the classification based on technology made by DUKES (Digest of UK energy statistics): wind farms, solar parks, CCGT, bioenergy, conventional steam and OCGT plants. The results are reported in Table 8.

I find negative DID estimates for CCGT, bioenergy and conventional steam plants. The price reduction of houses in the proximity to the facilities after their construction is about 3.1%, 8.6% and around 12% for CCGT, bioenergy and conventional steam, respectively. On the other hand, for solar parks, the estimated coefficient is positive, meaning that the construction of photovoltaic plants has determined an increase in the housing transaction prices of about 2.1%. With regards to wind farms, the coefficient of interest is statistically insignificant (p-value = 0.4). Similarly, for the only OCGT plant, the DID coefficient is once again not significant at all, not even at 10% level (p-value = 0.193).

Table 8: Effect of power plants on (log) prices by type of facility

	Wind	**CCGT**	**Solar**	**Bio**	**Steam**	**OCGT**
DID	.0073	-.0312***	.0209***	-.0904***	-.1212***	.0636
	(.0086)	(.0028)	(.0046)	(.0056)	(.0075)	(.0488)
Post	.0261***	.0024***	.0328***	.04***	-.0044*	1.66***
construction	(.0014)	(.0008)	(.002)	(.0014)	(.0023)	(.02)
Newly built	.112***	.139***	.1046***	.136***	.1153***	.0584***
	(.0021)	(.0014)	(.0018)	(.0046)	(.0035)	(.0103)
Semi-	-.276***	-.24***	-.252***	-.258***	-.203***	-.228***
detached	(.0086)	(.0011)	(.0015)	(.0018)	(.0028)	(.0061)
Terraced	-.403***	-.338***	-.339***	-.364***	-.272***	-.324***
	(.0022)	(.0014)	(.002)	(.0022)	(.0033)	(.008)
Flat	-.668***	-.608***	-.419***	-.633***	-.359***	-.355***
	(.0058)	(.0032)	(.0052)	(.0046)	(.0067)	(.0276)
Other types	-.6643***	-.867***	-.543***	-.956***	-.948***	-.323***
	(.0174)	(.012)	(.0166)	(.02)	(.0032)	(.0514)
Leasehold	-.069***	-.19***	-.362***	-.151***	-.3926***	-.462***
	(.0038)	(.0028)	(.0056)	(.0038)	(.0067)	(.0328)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Postcode fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,615,464	4,547,625	1,783,619	2,013,942	830,111	60,633
R-squared	0.656	0.703	0.747	0.702	0.738	0.734

Note: Robust standard errors, clustered at the postcode level are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.

In terms of control variables, the coefficients are substantially similar in terms of magnitude across the different regressions and to the overall findings seen above and all significant at 0.1% level. For the OCGT plant only, there is one difference: with respect to detached houses, flats cost more than other type of houses. The post construction dummy variable, alike overall findings, is overall extremely weak in all regressions taken into consideration.

I also try specifications with and without controls for property characteristics (and also with and without year dummies) and it makes little difference to the results, suggesting that unobserved price trends or differences in types of housing being sold do not affect the results massively.

Once again, I try different specifications, as already done before with overall findings. The results of DID coefficient are shown in table 9 (controls and year and postcode fixed effects are always included in the regressions).

Table 9: Effect with different specifications by type of facility

	Wind	**CCGT**	**Solar**	**Bio**	**Steam**	**OCGT**
≤ 3 km	.0048	-.0367***	.0155***	-.103***	-.15***	.0546
contr. grp ≥ 10 km	(.0086)	(.0029)	(.0047)	(.0057)	(.0078)	(.0491)
≤ 2 km	.0466*	-.0183***	.0352***	-.1293***	-.108***	.125*
	(.024)	(.004)	(.0065)	(.0086)	(.0108)	(.054)
≤ 4km	.0092	-.0281***	.0122***	-.0676***	-.100***	-.0078
	(.0054)	(.0025)	(.0032)	(.0039)	(.0053)	(.039)
≤ 8 km	-.0017	-.0142***	-.0069**	-.043***	-.074***	-.034
	(.0025)	(.0016)	(.002)	(.0022)	(.0027)	(.0232)

Note: Robust standard errors, clustered at the postcode level are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.

Similar results in signs and magnitudes come out by changing the specifications for control and treated groups. For the only OCGT plant, the results are quite insignificant in all settings, except when setting the treated group ≤ 2 km presenting a stunning positive sign (almost 13% price increase). As far as CCGT, bioenergy and conventional steam plants, the impact of facilities is the smaller, the larger is the treated group considered. For CCGT plants, for example, the impact doubles when considering the houses within 4 km, with respect to 8 km (from 1.4% decrease within 8 km to 2.8% within 4 km and 3.1% within 3 km). For photovoltaic facilities instead the effect is the opposite: the larger the treated group, the smaller the increase in housing prices is (so much that the setting with ≤ 8 km, the coefficient becomes negative). Once again, wind farms do not have significant effects on home prices, not even by changing the treated group, except for houses within 2 km where the price increase, significant at 5% level, is about 5%.

5.3 Results on power plants size

The results so far have considered simply the power station development as a binary treatment, and have not taken into consideration the dimension in term of energy capacity of the power plant. Generally speaking, more capacity might result in bigger sites, more pollution and more noises perceived by close properties than smaller stations. Table 10 shows the DID coefficients in the diff-in-diff fixed effects regressions by type of technology and according to the size of facilities: power plants whose capacity is less than 50 MW; between 50 and 100 MW; between 400 and 1000 MW; and more than 1000 MW. Some remarks: there is only one wind farm with capacity greater than 100 MW (228 MW) and too few observations are recorded (in total 594 transactions), it is therefore excluded by the analysis. It is also worth to remind that 9 wind farms with different capacities stay within a range of 9 kilometres. Therefore, the results regarding wind farms are not straightforward to interpret. Finally, three CCGT plants have less than 100 MW capacity (but more than 50 MW).

Table 10: Effects by power plant capacity

Capacity	**Wind**	**CCGT**	**Solar**	**Bioenergy**
≤ 50 MW	.0128	-	-.0102**	-.0259**
	(.009)		(.0053)	(.0085)
$50 < x \leq 100$ MW	-.0123	-.0031	.1677***	-.1243***
	(.0178)	(.0062)	(.0088)	(.0068)
$100 < x \leq 400$ MW	-	-	-	-
$400 < x \leq 1000$ MW	-	-.0499***	-	-
		(.0032)		
> 1000 MW	-	-.0755	-	-
		(.008)		

Note: Robust standard errors, clustered at the postcode level are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.
Conventional Steam and OCGT plants are excluded from the table. The former type presents only one plant; the latter has two facilities with similar capacity.

As far as CCGT and bioenergy plants are concerned, the table shows that a dimension effect is in place: in fact, both for CCGT and bioenergy facilities, the bigger the plants are, the bigger and more negative is the effect on the local housing prices. In fact, the biggest CCGT facilities with 1000 MW capacity, despite not statistically significant, determined a decrease in the housing price within 3 kilometres of almost 8% compared to houses more than 3 km distant. The smallest CCGT plants on the other hand, does not seem to cause any effect on prices in its proximity. Wind farms and solar parks present different results. Considering wind farms, no significant effects has been found for sites with more or less than 50 MW capacity, with basically no differences. Solar parks with less than 50 MW capacity do seem to have a negative effect on housing values (1% decrease), whilst bigger sites determine a remarkable 17% increase.

Related to this, it is worth to mention that in first stages of the analysis I needed to remove the oldest and remodelled power plants that in the UK account for the biggest ones in term of energy production. This could have reduced the impact of results as even the biggest power plants have modest capacity compared to the dropped ones. Furthermore, all nuclear and hydroelectric power plants that typically have higher capacity and are bigger and more impactful facilities in term of visibility and detectability are excluded due to either their construction prior to 1995 or re-qualifications of older sites.

5.4 Robustness checks and further statistical checks

Firstly, in all my regressions I used robust standard errors clustered at the postcode level following Bertrand et al. (2004). For implementation of a difference in differences design, they suggest clustering the standard errors on the panel unit identifier in order to correct for both autocorrelation and heteroscedasticity. This approach allows one to relax the Gauss-Markov homoskedasticity assumption and specifies that the standard errors allow for intragroup correlation, relaxing the usual requirement that the observations should be independent, meaning that the observations are independent across groups (clusters) but not necessarily within groups. Moreover, clustered standard errors are often useful when treatment is assigned at the level of a cluster as in my case, instead of at the individual level. The postcode aggregates multiple houses and thus when estimating the standard error or confidence interval of the model, classical or even heteroscedasticity-robust standard errors would be inappropriate since housing prices within each postcode are not independently distributed. Instead, houses in postcodes located in richer neighbourhoods have extra high housing prices (regardless of whether they receive the experimental treatment) while houses in postcodes located in poor neighbourhoods have especially low housing prices. In addition, in many panel data settings, such as diff-in-diff model used here, clustering often offers an effective way to account for non-independence between periods within each unit (sometimes referred to as "autocorrelation in residuals").

As mentioned in the previous chapter, I also repeat the regression using as dependent variable the housing characteristics used before as control variables. More specifically, I run different panel data fixed effects logistic regressions with leasehold or freehold dummy variable, newly built houses dummy variable and type of house (detached, semi, terraced, flat) at the left-hand side. The aim is to investigate if there are within-postcode changes in the composition of the sample that coincide with the start of facilities operations. However, it is worth to say that these covariates may be affected by the construction of facilities: as seen above, in term of transactions, the DID effect determined less housing transactions and this could have concerned some types of houses more than others. The results are in Table 11.

Table 11: Tests for housing characteristics as dependent variables

	Detached	**Semi**	**Terraced**	**Flat**	**Lease**	**New**
Wind	.092	-.0143	.0194	-.353	.051	-.441**
	(.0637)	(.0569)	(.0647)	(.12)	(.12)	(.183)
CCGT	.0189	.058**	-.0029	-.109**	-.128***	-.449***
	(.0203)	(.017)	(.017)	(.033)	(.0245)	(.0404)
Solar	.0147	.005	.065*	-.014	-.0478	-.426***
	(.029)	(.028)	(.031)	(.049)	(.049)	(.102)
Bio	.03	.0356	.0534*	-.229***	-.248***	.255*
	(.043)	(.028)	(.029)	(.05)	(.04)	(.091)
Steam	-.34***	-.156**	-.0834*	.0983	.14**	-2.34***
	(.0704)	(.046)	(.039)	(.051)	(.051)	(.19)
OCGT	-.0348	-.128	.399	9.65	12.59	-.368
	(.329)	(.2917)	(.365)	(359.99)	(1132.9)	(.461)

Note: Standard errors are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05.
Other set of regressors are excluded.

As the table shows, coefficients for OCGT, wind and solar plants are not significant. Thus, overall there is no evidence that the finding of impacts from these power plants on prices arises from omitted variables or unobserved price trends. However, concerning CCGT, bioenergy and conventional steam plants, some covariates are significantly negatively correlated with DID coefficient, meaning that the plants have reduced the transactions involving specific types of house.

From these last set of regressions, it emerged that some explanatory variables are correlated with each other. In fact, the model can also be suffering from multicollinearity, whenever independent variables are highly correlated to one another. The variance inflation factor (VIF) can give an indication if this is the case. It is defined as

$$VIF_j = \frac{1}{1 - R_j^2}$$

where R_j^2 is the R^2 obtained by regressing the j^{th} predictor on the remaining predictors. The correlation matrix or covariance matrix for the group of variables can also help in this sense. I checked for these and found out that there are not major multicollinearity issues in the model, except for flat and leasehold variables, which are quite highly correlated to each other.

Finally, I check to see how stable coefficients are when different samples are used and different characterisations are in place. In particular, I want to make sure that the difference in differences estimator is unbiased and, in order to do that, the key assumption to check is if the treatment and control group would have followed the same trend in the absence of the commencement of the energy facilities. I try to corroborate this assumption in multiple ways.

In this sense, I first restrict the estimation sample to a more homogeneous set of postcodes for which differential time trends are less likely (in Table 12 column 2). In particular, I selected all the power stations (and consequently transactions) located only in the English region of Yorkshire and the Humber, finding no substantial differences compared to the baseline estimation (column 1). I also control for trends at county level (which is one of the administrative subdivisions of the UK) and verify whether the coefficients (in particular the DID coefficient) are fairly stable and the diff-in-diff model is credible (column 3). The estimates are, again, similar to baseline estimates. I subsequently include to the set of regressors a linear time trend interacted with the treatment dummy (column 4) to test directly for different trends between treatment and control postcodes. The diff-in-diff estimator is a little smaller than column 1 (-0.008 versus -.02) but still statistically significant. Moreover, the interaction term (treat × linear time trend in the table) is extremely small in magnitude. Thus, linear price trends do not seem to differ between houses close to energy sites and houses further away from such facilities.

Table 12: Robustness checks: sample restrictions, county and time linear trends and street level treatment.

	(1) Baseline estimate	**(2) Yorkshire sub-sample**	**(3) Counties trends**	**(4) Treatment × linear time trend**
DID	-.0205***	-.0284***	-.0240***	-.00817***
	(.0022)	(.0058)	(.0005)	(.0024)
Treat × linear time trend				-.0009***
				(.00017)
Observations	8,894,754	743,561	8,894,754	8,894,754
	(5) County linear time trend	**(6) Street level treatment**		
DID	-.0207***	-.0078**		
	(.0019)	(.0028)		
Observations	8,894,754	8,784,569		

Note: Robust standard errors, clustered at postcode or street level are reported in parenthesis. *** p<0.001, ** p<0.01, * p<0.05. All regressions include time and year dummies, postcode (or "street") fixed effects and property characteristics.

I ultimately include county specific linear time trends by adding 119 linear time trends at the county level to the set of regressors (column 5), to take into account possible different trajectories of each county. Once again, the DID coefficient is unchanged. I also run a specification similar to the first baseline, but I now define the treatment at the street level as identified in the houses' addresses and including "street" fixed effects to show that the aggregation of treatment at the postcode level is not an issue. The estimate (column 6) is slightly smaller than the baseline, but still statistically significant at 1% level. I performed this for all types of facilities separately but table 12 shows the results for the overall findings. Overall, these results confirm that there is no evidence that housing prices of the treatment and control group followed different trends before the energy facilities' commencement dates.

5.5 Comparison with the literature

As previously described, a huge literature on this topic is present. Referring to all the studies mentioned before in the literature review section, this study is in line with those researches that examined the negative effect of industrial facilities on the housing markets, starting from Blomquist (1974) analysing coal-burning power plants, and continuing with Kiel and McClain (1995a, 1995b) regarding waste incinerators and Davis (2011) on power plants opened during the 1990s. In fact, a decrease in housing prices has been found for traditional power plants and plants based on the combustion of biomass (including waste incinerator), with even similar degree of magnitude (around 3-7% decrease).

When it comes to wind farms, the results in the relevant literature are less straightforward. My findings are along the lines of those that found substantially no effect on housing market (Hoen et al. 2011; Sims et al. 2008 among others) in contrast with other positive and negative results. For example, Kahn (2013) attests that wind farm counties can yield benefits for their local communities because the revenues to land owners spill over to the community in general, through lower property tax rates and improved public expenditures. On the other hand, many papers report negative effect of wind farms on housing values (Gibbons 2015 among others).

With regards to solar parks, as previously said, very few papers analyse its relationship with housing prices. Al-Hamoodah et al. (2018) report that the majority of individuals believe that proximity to a solar installation has either no impact or a positive effect on property values, while estimates on property values show an opposite negative relationship. In addition, Carlisle et al. (2015) argue that 70 percent of US respondents in a survey developed by them, believe that solar installations will decrease property values. Instead, my results indicate a small increase in UK housing prices (around 2% increase).

However, stated preferences studies on multiple energy sources (Borchers et al, 2007; Cicia et al, 2012) are also in line with my findings: solar is in general the most preferred one (solar parks determine an increase in prices in my analysis), followed by wind (no effects found) and biomass (under the macro-category bioenergy, I found a strong negative effect). Moreover, Cicia et al. (2012) also report that individuals are willing to replace fossil power plant with renewables (CCGT and steam plants cause a decrease in housing prices in my results).

5.6 Discussion

Several improvements and concerns arose during the development of this dissertation. First, the baseline specification, by defining treatment at the postcode level, ignores the geography and topography of a postcode. For instance, the visibility especially for wind farms, where the turbines may have different heights and dimensions, as demonstrated by Gibbons (2015) plays an important role in determining the housing values, other than the physical distance from the actual facility.

In addition, determining manually the accuracy of power plants coordinates via satellite imagery was subject to human error. This might have affected the outcomes ending up with faulty results. Moreover, in the case of CCGT power stations, many of them have been built in the same exact locations where previous coal-fired plants or other types had been in place for many years and then eventually dismissed. This means that people may have eventually already absorbed the idea of a disamenity caused by the energy site and the further construction may not have affected their perception.

Perhaps the main concern though regards the possible changes in behaviour by agents before the official commencement date which I used to define the before and after the siting of the power plant. The date I used for the analysis coming from the official UK energy statistics (DUKES) is the “year of commission or year generation began”. However, households may begin adjusting their preferences when they first hear about a public announcement of a forthcoming facility in their local area. The valuation and predisposition over the facility may change over time, when more details come out and the externalities for the local communities are more obvious. Even when the facility is completely operational, the initial fear of the local community could have been overestimated (or underestimated) and preferences and consequently prices will rebound again. For these reasons, the siting life cycle may have been oversimplified by choosing only one point in time and may have ignored the total welfare. An example of a study of this kind is by Kiel and McClain (1995a, 1995b) who analysed the housing prices during different siting stages of a waste incinerator and concluded that the adjustment of prices is more prolonged and composite than thought. Prices response occurred also only when rumours of a facility in the local town came out and also when the plant is in fully operation and households could actually weigh its effect and make a decision to relocate or stay. Therefore when more information is available about siting life cycles (first announcement, actual construction, the early operation stage and when fully operational) of power plants further studies on how tastes evolve over time and how the housing values varies over various phases of power stations can be further carried out.

To further complement the analysis, it might be useful to estimate the regression across different quantiles of postcodes property price distribution to test for the presence of possible heterogeneous treatment effects. In particular, the opening of a power plant may affect differently the price distribution: housing prices may decline the stronger, the higher (or lower) is the quantile, namely the shift in the price would be more (or less) pronounced for high-value properties. In fact, a differential variation within postcodes may reflect differences across differentiated housing markets. Richer residents (or buyers of high-value properties) may have a higher willingness to pay for environmental quality and the presence of the power station thus may affect their valuation of a neighbourhood disproportionately with respect to low-value property buyers. These agents may also differ in their access to information, that is, in their capacity to process the new presence of facility, or they may differ in their mobility. This would be consistent with the idea that richer households value environmental amenities more, are better processing the shock, or are less subject to relocation frictions.

Due to the lack of data on income and households characteristics (education, ethnicity and so on) at a local or topography level, I estimated only the hedonic price function itself (that is, the equilibrium relationship between a house prices and its characteristics). This “first stage” in the hedonic model can be used, under certain conditions, to infer the marginal willingness to pay for individual housing characteristics, including environmental amenities, or disamenities as in my case (Freeman et al. 2014). While such information is valuable, the next step that can be done is to measure the change in total welfare from a local decrease of environmental quality. Moreover, also a change in the composition of neighbourhoods close to the facilities can be analysed and computed.

As a matter of fact, when analysing the coefficient of variables of interest, it is worth to keep in mind that even though the first thought might be that there should be a positive relationship between the distance from the power station and the property prices due to local disamenities caused by negative externalities, a negative effect may occur simultaneously capturing, for example, employees who wish to live near the workplace (and this applies more strongly for traditional and conventional thermal power plants) and, more generally, there may be people that have other preferences that does not fit in the model or the effect that is trying to be captured. In fact, traditional power plants are usually still important local employers. Renewable energy instead is mainly capital intensive and labour intensive only in the process of building the plants, but once the facilities are built (except in the case of biomass) they do not need much management and maintenance. Regarding wind and solar energy, the employment impact is mainly temporary and largely indirect (it concerns the components produced in the factory).

The opening of a power plant might also increase labour demand through spill over effects in the region. Furthermore, in regular intervals, large-scale renewal, maintenance work as well as routine inspections, installation of equipment, regular reporting, systems integrations and reviews take place at power sites. This work may require many external engineers and assembly operators from subcontracting firms to work at the site in addition to the regular work force. These external workers often stay in local hotels and accommodations and eat in local restaurants. The opening of a facility is therefore likely to partially benefit also local hotels, shops and the restaurant industry in terms of income and employment. For all these reasons, a further analysis on a positive local demand shock and spill over effect, that can capitalise in the value of housing and thus partially compensate for the environmental disamenities can be undertaken.

Finally, as discussed by Banzhaf (2021), when applying the difference-in-differences method to the hedonic econometric model, there is not a clear and straightforward interpretation of the estimate. The derivative of a price function with respect to the characteristic of interest is the marginal willingness to pay. However, it is not obvious how this marginal willingness to pay can be identified in a diff-in-diff hedonic regression model. In fact, the dependent variable (housing price) mixes information from two different equilibria, from treated and control groups. It is argued that when a small subset of houses is examined, for changes to the set of characteristics, the equilibrium hedonic price function can be considered as constant over a short period of time, so that the DID model can be interpreted within a single equilibrium (Palmquist, 1992). However in the more general case, a large change in the supply of an amenity will shift the hedonic price function. In addition to that, other changes in the economic environment (income, other amenities) occurring over the longer time periods, such as ten years, would also shift the price function. When the hedonic price function shifts for either reason, panel data studies comparing prices at two different equilibria acquire potentially confusing results.

Consequently, in my study the diff-in-diff model implies also that after the start of the power plant activity, a shift in the demand is likely to occur from treated to control locations. Banzhaf (2021) argues that what I found to be the difference in differences estimates are the lower bounds of the true demand shift, especially when the Stable Unit Treatment Value Assumption is violated. This assumption requires that the potential outcome for any unit should be unaffected by the particular assignment of treatments to the other units. The control zones that are close to the plant despite being in the control group can be negatively impacted by the new openings of power stations. These control zones may also be indirectly affected through the housing demand of residents from treated areas. In order to check whether the control group has been impacted by the plants openings and whether this spill over effect has occurred, I have considered a 'doughnut' specification with a sort of "buffer zone" between treated and control areas (postcodes within 3 km as treated while postcodes ≥ 10 km as control) and found no significant differences. Furthermore, control areas are markedly larger than treated areas (see table 2), and more densely populated on average, and this is especially true for wind farms and solar parks (see table 3).

6. Conclusions

This dissertation provided estimates of the effects of different types of power plants on housing prices in England and Wales. The analysis consisted of a difference in differences postcode fixed effects methodology. It used daily housing transactions aggregated at the postcode level spanning almost 26 years. It compared postcodes within 3 kilometres where power plants became operational, with postcodes more than 3 kilometres away (but still within 15 km). The results varies depending on the technology used: the price reduction is about 12% for steam plants, falling to around 9% for bioenergy sites and 3% for CCGT facilities. No effects are found for wind farms, while solar parks increase housing prices by around 2%. Moreover, the results show that bigger facilities have typically a bigger impact on housing prices. Different specifications for control and treated groups have also been undertaken as well as numerous robustness checks to verify the diff-in-diff parallel trends assumption holds and if the model is credible. The estimates are also in line with the relevant literature regarding fossil fuel and biomass power plants. For renewables, results are less in conformity with the literature (despite there are not many papers analysing the same topic), but the overall findings are totally consistent with outcomes from the stated preference literature, classifying fossil fuels as the least preferred, followed by biomass, wind and solar power, considered the most preferred, although the quantitative results are not directly comparable.

Rosen (1974) and the hedonic price literature (in spite of the critiques by Banzhaf 2021 explored in the discussion section) suggest that price differentials emerging between places closer to power stations and places further away from sites can be interpreted as household marginal willingness to pay to avoid the disamenity of having an energy facility. If this is the case and we take the estimates in this study as proxies of the mean willingness to pay to avoid power stations in locations exposed to them, the cost is quite considerable. For instance, if we assume an average house price of £180,000 (based on table 2) and a 5% interest rate, a household would be willing to pay around £270 per year to avoid having a CCGT plant within 3 km, around £810 to avoid a bioenergy plant and around £1,080 per year to avoid having a conventional steam plant.

These results shed light on some policy implications. Siting new energy facilities is becoming increasingly difficult over time, in large part because of the NIMBY issue, especially in places with large and growing populations. In this sense, policymakers have to balance several factors and often face difficult, politically contentious decisions about where to site plants. Local amenities are typically one of the most important components considered in this process. However, the lack of reliable empirical evidence about the magnitude of these costs has limited the use of cost-benefit analysis. Thus, these results can provide potential inputs for cost-benefit analysis related to the siting of energy facilities. It is quite clear that home buyers near energy sites value more negatively traditional power plants and plants based on the combustion of biomass than renewable energy facilities such as wind and solar farms. The findings could also suggest a first approximate estimate of a possible policy on monetary compensation for home owners for the loss of value in their homes near the power plant, such as tax relieves or tax deductions.

7. References

Al-Hamoodah, L., Koppa, K., Schieve, E., Reeves, D. K., Hoen, B., Seel, J., Rai, V. (2018). *An Exploration of Property-Value Impacts Near Utility-Scale Solar Installations*. Working Paper. Policy Research Project, LBJ School of Public Affairs, The University of Texas at Austin.

Bakker, R.H., Pedersen, E., van den Berg, G., Stewart, R. E., Lok, W., Bouma, J. (2012). “Effects of wind turbine sound on health and psychological distress”. *Science of the Total Environment*, 425: 42–51.

Banzhaf, H. S. (2021). “Difference-in-differences hedonics”. *Journal of Political Economy*, 129 (8): 2385 - 2414.

Bauer, T.K., Braun, S. and Kvasnicka, M. (2017). “Nuclear power plant closures and local housing values: Evidence from Fukushima and the German housing market”. *Journal of Urban Economics*, 99: 94–106.

Bertrand, M., Duflo, E. and Mullainathan, S. (2004). "How Much Should We Trust Differences-In-Differences Estimates?". *The Quarterly Journal of Economics*, 119 (1): 249–275.

Bielecki, A., Ernst, S., Skrodzka, W. (2020). “The externalities of energy production in the context of development of clean energy generation”. *Environmental Science and Pollution Research,* 27: 11506–11530.

Blomquist, G. (1974). “The Effect of Electric Utility Power Plant Location on Area Property Value,” *Land Economics,* 50 (1): 97-100.

Borchers, A., Duke, J., Parsons, G. (2007). “Does Willingness To Pay For Green Energy Differ by Source?”. *Energy Policy,* 35: 3327-3334.

Carlisle, J. E., Solan D., Kane S. and Joe J. (2016). “Utility-Scale Solar and Public Attitudes Toward Siting: A Critical Examination of Proximity.” *Land Use Policy,* 58: 491-501.

Cicia, G., Cembalo, L., Del Giudice, T., Paladino, A. (2012). “Fossil Energy versus Wind, Solar and Agricultural Biomass: Insights from an Italian National Survey”. *Energy Policy*, 42: 59-66.

Clark, D. E. and Allison, T. (1999). “Spent Nuclear Fuel and Residential Property Values: The Influence of Proximity, Visual Clues and Public Information”. *Papers in Regional Science,* 78: 403-421.

Clark, D. E., Michelbrink, L., Allison, T. and Metz, W. C. (1997). “Nuclear Power Plants and Residential Housing Prices”. *Growth and Change,* 28: 496–519.

Coulomb R. and Zylberberg Y. (2020). “Environmental risk and the anchoring role of mobility rigidities”. *Journal of the Association of Environmental and Resource Economists*, 8 (3): 509–542.

Cunningham, S. (2021). *Causal Inference: The Mixtape.* New Haven, Connecticut: Yale University Press.

Davis L. W. (2011). “The effect of power plants on local housing values and rents”. *The Review of Economics and Statistics,* 93 (4): 1391–1402.

Dröes, M. and Koster, H. (2020). "Wind turbines, solar farms, and house prices", CEPR Discussion Paper 15023.

European Environment Agency. (2008). *Air pollution from electricity-generating large combustion plants*. Tech. Rep. No. 4/2008.

Fink, A. and Stratmann, T. (2015). “U.S. housing prices and the Fukushima nuclear accident”. *Journal of Economic Behavior and Organization*, 117 (C): 309–326.

Freeman, A. M. III, Herriges, J., and Kling, C. (2014). *The Measurement of Environmental and Resource Values*, Third Edition. Washington D.C: Resources for the Future.

Gibbons, S. (2015). “Gone with the wind: Valuing the visual impacts of wind turbines through house prices”. *Journal of Environmental Economics and Management*, 72: 177–196.

Global Energy Observatory, Google, KTH Royal Institute of Technology in Stockholm, Enipedia, World Resources Institute. (2018). *Global Power Plant Database*. Published on Resource Watch and Google Earth Engine.

International Energy Agency. (2019a). *Energy Policies of IEA Countries: United Kingdom 2019 Review*. Paris: IEA Publications.

International Energy Agency. (2019b). *Does household use of solid biomass-based heating affect air quality?* iea.org, 18 March. https://www.iea.org/articles/does-household-use-of-solid-biomass-based-heating-affect-air-quality [last access: 12/08/2021].

International Energy Agency. (2021). *Global Energy Review 2021*. Paris: IEA Publications.

IRENA. (2021). *World Energy Transitions Outlook: 1.5°C Pathway.* International Renewable Energy Agency, Abu Dhabi.

Kahn, M. E. (2013). "Local non-market quality of life dynamics in new wind farms communities". *Energy Policy,* 59: 800–807.

Kiel, K. and MacClain, K. (1995a). "House Prices During Siting Decision Stages: The Case of an Incinerator from Rumor through Operation". *Journal of Environmental Economics and Management,* 28: 241-255.

Kiel, K. and MacClain, K. (1995b). "The Effect of an Incinerator Siting on Housing Appreciation Rates". *Journal of Urban Economics,* 37: 311-323.

Kroes, E. and Sheldon, R. (1988). "Stated Preference Methods: An Introduction". *Journal of Transport Economics and Policy,* 22 (1): 11-25.

Kuminoff, N., Parmeter, C. and Pope, J. (2010). "Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?" *Journal of Environmental Economics and Management*, 60 (3): 145–160.

Lang, C., Opaluch, J. and Sfinarolakis, G. (2014). "The Windy City: Property Value Impacts of Wind Turbines in an Urban Setting". *Energy Economics*, 44(C): 413-421.

Lipscomb, C. (2011). "Using Contingent Valuation to Measure Property Value Impacts". *Journal of Property Investment and Finance,* 29: 448-459.

Lovich, J. E. and Ennen, J. R. (2011). "Wildlife Conservation and Solar Energy Development in the Desert Southwest, United States". *Bioscience*, 61 (12): 982–992.

Mattman, M., Logar, I. and Brouwer, R. (2016). "Hydropower externalities: A meta-analysis". *Energy Economics*, 57: 66-77.

Muehlenbachs, L., Spiller, E. and Timmins, C. (2015). “The housing market impacts of shale gas development”. *American Economic Review,* 105 (12): 3633–3659.

Munday, M., Bristow, G. and Cowell, R. (2011). “Wind farms in rural areas: How far do community benefits from wind farms represent a local economic development opportunity?”. *Journal of Rural Studies,* 27: 1-12.

Mundy, B. and McLean, D. (1998). “The Addition of Contingent Valuation and Conjoint Analysis to the Required Body of Knowledge for the Estimation of Environmental Damage to Real Estate”. *The Journal of Real Estate Practice and Education*, 1 (1): 1-19.

Palmquist, R. B. (1992). “Valuing Localized Externalities”. *Journal of Urban Economics*, 31 (1): 59-68.

Roddewig, R. J., Frey, J. D. (2006). “Testing the Reliability of Contingent Valuation in the Real Estate Marketplace”. *The Appraisal Journal*, 74: 267-280.

Rosen, S. (1974). “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”. *The Journal of Political Economy*, 82 (1): 34-55.

Soland, M., Steimer, N. and Walter, G. (2013). “Local acceptance of existing biogas plants in Switzerland”. *Energy Policy,* 61: 802-810.

Welsch, H. (2016). “Electricity externalities, siting, and the energy mix: A survey”. Oldenburg Discussion Papers in Economics, No. V-394-16, University of Oldenburg, Department of Economics.

Welsch, H. and Ferreira, S. (2014). “Environment, Well-Being, and Experienced Preference”. Oldenburg Discussion Papers in Economics, No. V-367-14, University of Oldenburg, Department of Economics.

Whitehead, J., Pattanayak S., Van Houtven, G., Gelso, B. (2008). “Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: An Assessment of the State of the Science”. *Journal of Economic Surveys*, 22 (5): 872–908.

World Health Organization. (2019). *Health consequences of air pollution on populations*. Who.int, 15 Nov. https://www.who.int/news/item/15-11-2019-what-are-health-consequences-of-air-pollution-on-populations [last access: 12/08/2021].