Working Paper No. 489

August 2014

Estimating Building Energy Ratings for the Residential Building Stock: Location and Occupancy

Subsequently published in "Using census and administrative records to identify the location and occupancy type of energy inefficient residential properties", *Sustainable Cities and Society*, Vol. 18, 2015, pp. 56-65, http://dx.doi.org/10.1016/j.scs.2015.06.001

John Curtis, Niamh Devitt, Adele Whelan

Abstract: The common EU framework for assessing the energy performance of residential buildings and awarding Energy Performance Certificates (EPCs) is an important resource in the context of informing effective policy measures to improve energy efficiency. However, properties that have been assessed to-date are not likely to be fully representative of the entire housing stock and therefore provide a faulty baseline from which to devise policy actions. The paper presents a methodology to estimate the energy performance of all residential properties and, combined with census data, identifies what distinguishes the most energy inefficient properties, whether it is location, ownership, age or other characteristic. Data from the Irish EPC database suggest that 25% of the Irish residential housing stock is in the most energy inefficient categories, whereas the methodology developed here suggests that it is substantially higher at 35%. The results also indicate that there is a substantially greater likelihood that the elderly, or families living in rental properties live in the most energy inefficient properties.

Corresponding Author: John.Curtis@esri.ie

Estimating Building Energy Ratings for the Residential Building Stock: Location and Occupancy

1. Introduction

The European Union, through the Energy Efficiency Directive, has established a common framework of measures to promote energy efficiency and achieve its headline target of a 20 per cent improvement in energy efficiency by 2020 (CEC (2010); EP and CEC (2012)). Within energy efficiency the greatest energy saving potential lies in buildings. Nearly 40% of final energy consumption occurs in buildings, and specifically within residential buildings, two-thirds of energy use is for space heating (CEC (2011b)). It is widely recognised that the energy savings potential within the residential building stock is large but given the heterogeneity of buildings both between and within countries there is very little precise information on exactly where the savings will be realised. Without such information it is difficult to track progress toward the policy target or identify types of residential buildings where energy efficiency savings can be achieved most easily.

Under the 2002 EU directive on energy performance of buildings (EP and CEC (2002)), which established a methodological framework for calculating energy performance, EU member states have been developing certification systems for rating the energy performance of buildings. These certification systems, which differ between countries, provide a standardised framework to provide an indication of the energy performance of residential buildings. One benefit of building energy ratings is they can act as a signal to property market participants who are willing to pay extra for more energy-efficient properties with lower running costs. Several studies have confirmed that properties with high energy rating certification can command a price premium both in residential (Australian Bureau of Statistics (2008); Brounen and Kok (2011); Hyland *et al.* (2013); Cajias and Piazolo (2013)) and commercial buildings (Kok and Jennen (2012); Eichholtz *et al.* (2010); Reichardt *et al.* (2012)). Another benefit of building energy ratings is the information they provide relating to the energy efficiency of the building stock, which is of considerable interest in the context of energy efficiency policy targets. For instance, knowledge of energy performance across the entire building stock can provide a strong underlying basis for plans and measures to improve energy performance. Profiling the building energy rating by occupants provides an indicative assessment on the extent to which policy measures will be financed by occupants or whether grant schemes are likely to be more effective.

The EU's target of a 20 per cent improvement in energy efficiency by 2020, of which building energy efficiency is a major component, is a Europe-wide target. Cost effectively achieving that target would entail investing in energy efficiency where it is cheapest to do so, irrespective of geography. But there is neither sufficient information nor policy mechanisms to consider a broad whole-Europe policy response to the target. Instead the policy response is at individual country level, where countries are required to prepare national energy efficiency action plans (NEEAPs) and revise them on a three-yearly basis. Among the outputs envisaged within the Energy Efficiency Directive (EP and CEC (2012)) and to be reported upon in NEEAPs are evidence-based strategies for mobilising investment in the renovation of the national stock of residential and other buildings. A good understanding of the status of the national building stock is an important basis for that work, as is knowledge of the means of households living within different building archetypes.

National databases on the energy performance of buildings offer a means to improve knowledge about national residential building stocks, their energy performance and where potential energy savings exist. These databases contain the information underlying building Energy Performance Certificates (EPC), which are usually required as part of mandatory data disclosure associated with selling or renting property. Exploiting national EPC databases will enable more accurate assessment of potential energy efficiency gains and the associated investment cost of retrofitting residential buildings. However, as the underlying legislation is relatively recent, EPC databases may not be representative samples of the building stock. Relatively new properties as well as properties that have been rented or sold in recent years are likely to be over-represented. The implication is that EPC databases may misrepresent the energy efficiency of the residential building stock, possibly indicating a higher level of energy efficiency performance than can be realized in practice. EPC databases contain information on a large proportion of residential properties but in many instances less than half of all residential properties have been assessed; 50% in England and Wales, 40% in Scotland, and 30% in the Republic of Ireland[1]. In Ireland the EPC database is over-represented with relatively recently built properties, with 17% of properties built since 2006 (or 54% since 1991). This compares with census data (CSO (2012a)) which suggest that about 10% of properties were built since 2006 (or 41% since 1991). Nonetheless, EPC databases do contain sufficient information to enable estimation of the building energy ratings for the entire national residential building stock, as opposed to actual EPC ratings for an unrepresentative sample. With EPC estimates for the entire building stock, policy makers will be better informed as they design measures to improve the residential sector's performance with respect to both energy efficiency and climate policy targets.

This paper presents a methodology to estimate EPCs for a national residential building stock using limited information on building characteristics without the need to undertake costly, on-site EPC assessments. The methodology also provides insight into the profile of the households living within properties across EPC ratings. We use data from Ireland to

[1] For England and Wales see www.epcregister.com/lodgementStats.html; For Scotland see www.scotland.gov.uk/Topics/Built-Environment/Building/Building-standards/enerperfor/epcstats; and for the Republic of Ireland see www.seai.ie/Your_Building/BER/BER_FAQ/FAQ_BER/General/BER_Statistics.html. Accessed 24 March 2014.

demonstrate the methodology but the approach should be easily replicable in other EU member states, where the same policy context exists in terms of developing evidence-based strategies for mobilising investment in building energy efficiency. Developing a methodology to estimate the energy performance of the national building stock serves a number of policy-related purposes. In the context of improving energy efficiency in buildings it helps identify the type, number and geographical location of buildings where the most significant energy efficiency gains can be achieved. It also provides detailed information about households living in energy inefficient properties, including tenure of occupancy, household disability status, and the age profile of occupants. Knowledge of such information can assist in the design of targeted incentive schemes for investment in energy efficiency upgrades. It will also help in estimating investment costs associated with energy efficiency improvements to achieve policy targets, as well as estimating the energy efficiency potential from remediation works on the residential building stock.

The paper proceeds with further discussion of the policy drivers for improving energy efficiency and how enhancing the knowledge base surrounding the residential housing stock will facilitate that work. Section 3 outlines the methodology employed to improve the estimates of residential energy efficiency. The data used to illustrate the methodology are described in Section 4 and the estimation results are presented in Section 5. Section 6 contains a discussion of the results drawing policy conclusions about where public policy should prioritise measures to improve residential energy efficiency. Conclusions are summarised in Section 7.

2. Policy Context

The Energy Efficiency Directive (EP and CEC (2012)) is the EU's main legislative mechanism for implementing energy efficiency policy supported by a number of other directives such as the energy performance of buildings directives (EP and CEC (2002); EP and CEC (2010)). Energy efficiency is central to the EU's Europe 2020 Strategy for smart, sustainable and inclusive growth and for the transition to a resource efficient economy (CEC (2010)). The strategy recognises that energy efficiency is the most cost effective way to reduce emissions, improve energy security, increase competitiveness and employment, as well as make energy more affordable. Energy efficiency is the flagship initiative under the Europe 2020 Strategy (CEC (2011a)) and within the Commission's energy efficiency action plan buildings are identified as the area where the greatest energy saving potential lies (CEC (2011b)). The action plan focuses on developing instruments to trigger the renovation process in buildings and to improve building energy performance. A key element of any plans to trigger building renovation is a better understanding of the quality of the building stock and identifying where building improvements will achieve the greatest gains. Understanding the extent to which the greatest gains in energy efficiency coincide with occupants/owners (un)willingness or (in)ability to finance building refurbishment should affect National Energy Efficiency Action Plans.

Across Europe the scale of energy savings potential is immense. Between 1996 and 2007 energy efficiency savings totalled approximately 160 million tonnes of oil equivalent (ADEME (2009)) but even with such progress the EU is on course to achieve only half of its ambitious 20% target for energy efficiency improvement by 2020 (CEC (2011b)). Similarly impressive energy efficiency savings have been achieved in Ireland with 26% of the national savings target for 2020 achieved by 2012, representing savings of €470 million in energy costs and 2 million tonnes of carbon dioxide equivalent emissions (DCENR (2012)). In Ireland much of the savings to date are attributable to new building regulations but future savings are most likely to be realised within existing buildings. Similarly across Europe the extensive refurbishment of the existing housing stock will be necessary to achieve energy efficiency targets. Good knowledge of the building stock and its energy performance will be necessary to develop effective policy measures to incentivise the refurbishment of existing dwellings. While there are instances where there is detailed information to inform policy (e.g. DCLG (2010)) in many cases the knowledge base is quite limited. On-site energy assessments to develop representative sample databases are feasible but prohibitively costly. An alternative is to combine existing administrative datasets as a cost effective means to inform policy action.

NEEAPs are a reporting requirement of the Energy Efficiency Directive, with revised NEEAPs required on a three-yearly basis. Improved understanding of the housing stock will better inform the development of policy measures, including the revision of NEEAPs. Some of the efficiency savings will be easily achieved through general measures and policies, for example, subsidy schemes for insulation retrofitting. Increasingly more detailed information will be required to understand where potential efficiency savings are located and what trigger mechanisms are needed to encourage investment to secure those benefits. Detailed information is also necessary to effectively target public funds to alleviate households either experiencing fuel poverty or who are unable to afford investment in energy efficiency upgrades. Estimates of the energy performance of residential building stocks help fill such information deficits.

3. Methodology

We use data from Ireland to demonstrate the methodology but the approach is easily replicable in other countries. The first stage uses data from an EPC database to estimate a model using limited information on building characteristics (e.g. age, building type, heating type) to categorise buildings with an estimated EPC rating. We confine the explanatory variables to variables that are also contained in the buildings database used in the second stage, which calculates EPCs for each property in the residential buildings database based on the parameter estimates in stage one. In Ireland we use individual household records from the 2011 census, as Ireland does not have a buildings database. An advantage of census records is that they provide details both of property characteristics (e.g. age, building type, heating type) and the property's occupants. With the latter information we can, for example, identify if vulnerable families are more likely to live in energy inefficient properties.

A shortcoming of census data is that it relates to occupied properties only and excludes some 290,000 vacant properties in 2011 many of which are a legacy of the property bubble.

Actual EPC assessments entail much more technical information about the property than used as explanatory variables in the model, consequently the EPC equation is miss-specified. The miss-specification due to omitted variables will lead to bias in the parameter estimates (Greene (2002)). While the parameter estimates may be biased we are not specifically interested in individual parameters or their magnitude. Instead we want to use the estimated equation to map the limited information available about properties into an EPC rating for those properties.

The EPC classification model is specified with EPC as a linear function of the household characteristic variables (X_i), where $i = 1\ to\ N$ households.

$$\log\left(\mathrm{EPC}_i\right) = \beta X_i \qquad (1)$$

The vector β contains the model parameters to be estimated. EPC rating is a continuous variable, denoted in kWh/m^2/year, of which the log transformation is taken such that the dependent variable more closely satisfies the assumptions for the classical linear model, that it is distributed normally. The log transformation also reduces extrema in the data, though extreme outliers with EPC assessments in excess of 2000 kWh/m^2/year, which is over four times the threshold for the poorest category rating, were excluded. The explanatory variables used in estimation are the year the property was built, the type of dwelling (e.g. apartment), and fuel used for space heating.

4. Data

The Sustainable Energy Authority of Ireland (SEAI) maintains a public register of completed EPCs, termed Building Energy Ratings (BERs) in Ireland.[2] BERs are calculated by a standard assessment procedure, which models energy consumption under standard occupancy and normal climatic conditions, in line with European legislation.[3] The ratings are based on the technical specifications of the property including building type, year of construction, heating and cooling equipment, fuels used and are calculated using a bespoke software. A BER rating is reported as the total primary energy used expressed in kilowatt hours per unit area per annum (kWh/m^2/year), which is subsequently classified into a 15-point alpha-numeric scale. Table 1 shows the breakdown of the 15-point BER scale and the proportion of BER assessed properties within each BER sub-category. A low kWh/m^2/year indicates a good energy performance with the alpha-numeric rating beginning with A1 for the most efficient down to G, the lowest energy performance. The median property in the BER database has a

[2] The database of BERs is available to download at http://www.seai.ie/Your_Building/BER/National_BER_Research_Tool/

[3] Details of the assessment procedure and software are available at http://www.seai.ie/Your_Building/BER/BER_Assessors/Technical/DEAP/

C3 BER rating, which is between 200-225 kWh/m^2/year. Roughly 25-30% of residential properties have completed BER assessments, as BER assessments are only mandatory when selling or renting property.

The BER database contains a myriad of variables relating to technical information about individual properties. For example, measures of thermal resistance of windows, efficiency factors of heating systems, and building fabric. While such detailed information is important in determining a property's specific BER rating, the analysis here utilises only general household characteristic information in a BER classification regression. Specifically, only variables about property characteristics that are also available in the census dataset are included in the analysis. Descriptive statistics for these variables are presented in Table 2. Both the BER database and the census contain information on property size but there was not a clear relationship between the variables used in the two datasets and consequently property size was not used as an explanatory variable in the classification regression.

5. Estimation

The OLS estimates of equation (1) are reported in Table 3 and as mentioned earlier, the parameter estimates suffer from omitted variable bias. However, based on the relative magnitude of the parameter estimates we can learn about relative weightings within the EPC assessment process. Ranging from the oldest to the newest properties estimated parameters are increasingly negative in magnitude, meaning that newer properties require less energy for lighting and heating and generally have better BER ratings (i.e. lower kWh/m^2/year). This may reflect improvements in building practices and technologies, as well as improvements in building regulations, which have evolved over time with many of the developments leading to improvements in building energy efficiency. The negative coefficients for apartments and terrace houses indicate that such properties are considered more energy efficient in the BER rating scheme compared to semi-detached and detached houses (reference category in regression). This relative weighting will reflect the fact that such properties have less external wall area, which is vulnerable to heating loss. Across central heating types the lower estimated coefficient values for gas (mains and LPG) and wood compared to electricity, coal and peat also show that such heating systems are considered more energy efficient in the BER rating assessment.

Though the model uses just a sub-set of the technical data utilised in making an actual BER assessment it explains 57% of the total variation of outcomes based on the coefficient of determination. To assess the performance of the model we compared model predictions with actual BER assessments. Figure 1 plots the difference in estimated and actual BER values sorted in increasing value. The average error across all properties is just 12 kWh/m^2/year though there are also quite substantial errors for some properties. Overall, 83% of the errors are within 100 kWh/m^2/year of the actual value. Some of the largest prediction errors occur in properties that have an actual BER rating considerably in excess of

450 kWh/m^2/year, the threshold for the G BER classification. The model is poor at predicting properties with high energy efficiency but across the entire BER dataset just 6% of properties were classified in the top 5 BER categories (A1 to B2). The model is better suited to predicting properties in the bottom ten BER categories. The most frequent BER classifications for Irish residential property are classes C1 to D2, accounting for 62% of properties assessed with a further 24% in classes E1 to G.

We also compare predicted versus actual BER assessments by letter category, A to G, in Table 4. Of the 410,336 dwellings used in estimation 54,508 had a B rating (i.e. B1, B2, or B3) and the model correctly predicted a B rating for 33,821 of those dwellings. This is equivalent to 62% of all B rated dwellings. A further 25.9% of dwellings with an assessed B rating had a model prediction of a C rating. The model does not predict any A rated dwellings, which is not surprising given that just 0.5% of dwellings were A rated within the BER database. Instead the model predicts a B rating for 75% of actual A rated properties. In total the model correctly predicts the correct letter category for 47% of all dwellings. A prediction accuracy of roughly 1 in 2 is not impressive but in the context of the relatively simple model it is not particularly poor.

From a policy perspective more rather than less information is preferred but if the focus is on identifying the most energy inefficient properties the analytical focus would not be on which letter rating a property is but whether it is below a specific threshold. In such circumstances the focus might be on identifying dwellings in the E, F or G categories. Model predictions using such a threshold are reported in Table 5.[4] In the BER database there are 100,960 dwellings rated E, F or G (EFG). The model predicts that 102,033 dwellings are rated EFG but that includes some 37,265 false positive predictions. In total the model correctly predicts the BER rating for 82% of households. With relatively basic and limited information about residential properties the model is reasonably strong at classifying properties' energy efficiency. It therefore becomes a useful tool to predict energy inefficient properties within the entire residential building stock for the purposes of informing public policy. The model could undoubtedly be improved adding additional explanatory variables but that would thwart the second stage of the analysis, which combines the estimated model with the census dataset to investigate whether energy inefficient properties are more prevalent with particular socio-economic cohorts or geographical locations.

The second stage of the analysis is calculating BER ratings for all residential properties in the census dataset. Taking the model parameter estimates from Table 3, denoted as $\hat{\beta}$, a BER rating measured in kWh/m^2/year is calculated by

[4] An alternative modelling approach in this instance could be a logit or probit type model with a discrete dependent variable indicating whether a dwelling is above the threshold or not. With such a model the overall within sample prediction accuracy is similar to the OLS model results presented here.

$$e^{\widehat{\beta} X_{Census}} \qquad (2)$$

Where X_{Census} represents the census dataset variables for building age, property type and central heating fuel type. Each property in the census dataset is classified as being either energy efficient (ABCD) or energy inefficient (EFG).[5]

The ratio of energy efficient and inefficient properties is plotted in Figure 2 both those estimated from census data and for comparison the ratios from the BER database. The BER database suggests that 25% of the residential housing stock is energy inefficient (EFG), whereas the estimate based on census data suggests that it is substantially higher at 35%. Relying on the BER database as being representative of the housing stock leads to a large overestimate of the energy efficiency of the housing stock.

Table 6 reports some statistics comparing the distribution of estimated and actual BER ratings, as measured in kWh/m^2/year. There are some notable differences in the distributions. The means are similar (261 versus 276) but the shapes of the distributions are quite different resulting in a notable difference in the median rating (225 versus 255). Both distributions have positive skewness, indicating that the right side tail is longer or fatter than the left, but this is greater in the case of the BER database. A high kurtosis value indicates a sharper peak and longer, fatter tails, which is the case for the BER database compared to the estimated BER ratings from the census data. In summary, the BER database suggests a sharp peak or concentration around the mean, whereas the estimated BER distribution from the census data suggests that the distribution has a more rounded peak and the shape of the distribution is more similar to the normal distribution. There is no reason to expect that the distribution of BERs should follow a normal distribution, in fact, right skewness reflects the abundance of energy inefficient properties. Equally, it is difficult to understand why the true distribution of BERs (as a measure of energy efficiency in the residential housing stock) would be a high kurtosis distribution. Relying solely on the BER database to profile the energy efficiency of the residential building stock is likely to result in under-estimating the energy load in residential properties. We approximate that underestimate to be at least of the order of 5%.[6]

6. Who Lives in Energy Inefficient Properties?

An issue of potentially greater interest to policymakers than the number or distribution of energy inefficient properties is an appreciation of what distinguishes such properties, whether it is location, occupancy, ownership or other characteristic. Combining the BER

[5] While categorisation by 7 letter categories, or 15 alpha-numeric categories is feasible, it implies a level of model precision not merited, as indicated in the discussion of Table 4. Analysis of whether properties are energy inefficient (EFG) is sufficient to inform policy decisions.

[6] This estimate is based on using the BER database for average floor size by letter rating; the mid-point of the BER letter ratings in kWh/m^2 /year; the total number of private residential housing units from the census; the share of properties by letter rating both estimated here and from the BER database.

estimates with other information from the census of population enables such analysis to be undertaken. While policy is focused at improving energy efficiency across the housing stock a particular interest of policymakers is identifying whether there are particular barriers to upgrading the least energy efficient properties. The census data enables us to determine whether a particular cohort of the population is more likely to reside in such properties. The analysis will initially proceed with a cross-tabulation of variables against energy efficiency and will be followed with multivariate regression analysis.

Tenure

From the census of population we focus on the 1.6 million private residential dwellings enumerated in the 2011 census (CSO (2012a)).[7] Roughly 560,000 of these properties are estimated to have an EFG BER rating representing 35% of the residential stock (see Table 7). In terms of occupancy the likelihood of a property being EFG is slightly higher when tenure is rental versus owner occupier, 38% versus 33%. Within the rental market a higher proportion of private rental properties are EFG compared to local authority housing, 40% versus 36%. Overall, differences in the likelihood of properties being energy inefficient by tenure type are relatively small and consequently energy inefficiency does not appear to be a problem particularly associated with type of tenure.

Rental Values

There may be some preconception that low value rental properties are coincident with poor energy efficiency with a justifying logic that properties with low rents may not generate sufficient income to justify investment in energy efficiency. An issue for policymakers is whether EFG properties are more prevalent in the lower value rental market and whether that is a target group for policy intervention. A cross tabulation of rental values versus energy efficiency finds no strong evidence that that is the case but some evidence of the converse. Roughly 40% of low rental value properties are EFG regardless of whether rental rates are relatively low at €50/week or much higher at €500/week, as shown in Table 8. The highest incidence of EFG properties in the rental market occurs in properties that the weekly rental rate exceeds €500/week. When the multivariate analysis is presented later we find that low value rental properties are much more likely to be EFG rated compared to owner-occupied property.

Disability

Another area of policy focus is whether vulnerable or disadvantaged households are more likely to find themselves living in energy inefficient properties. Depending on the circumstance, energy inefficient properties may aggravate the vulnerability or disadvantage. The census recorded 535,000 people in 432,000 households with a self-declared disability, which included vision and hearing impairment, intellectual disability, mobility disability,

---

[7] It should be noted that the census identified 2 million permanent dwellings or housing units but this figure includes both vacant properties and holiday homes.

psychological conditions or chronic illnesses. Approximately 41% of people with disabilities (217,000 persons) live in EFG properties compared to 32% of the total population who live in EFG properties.

Age

The very young and old are among potentially vulnerable cohorts in the population. In 2011 there were 1.2 million children living in Ireland with 25% estimated to live in EFG rated properties. As a group children are less likely to live in the least energy inefficient properties. The same cannot be said for the elderly. About 57% of people aged 75 and over live in EFG rated properties, as reported in Table 9, with similar statistics when elderly is defined as above aged 60 or 65. Depending on these three definitions between 17% and 41% of EFG properties are occupied by elderly people. In the context of policy targets to improve the energy efficiency of the residential housing stock, a considerable proportion of the most energy inefficient properties are occupied by residents that are likely to be hesitant to engage in significant investment, irrespective of the issue of affordability.

Employment

Ability to pay for building improvements is likely to be an important consideration in policy measures to encourage households to upgrade their homes' energy efficiency. The census does not contain information on incomes but does provide information on households that have gainfully employed occupants. The latter is not a proxy for income and not gainfully employed can include a range of circumstances, such as occupants that are retired, unemployed, students, permanently unable to work due to disability or sickness, or caring for family. Table 10 reports the proportion of households with and without any occupants gainfully employed by the type of energy efficient residence. The disparity is quite significant. Of households with at least one person gainfully employed, 30% live in EFG rated properties, whereas the rate rises to 44% where there is no occupant gainfully employed. Income is potentially one reason for this disparity, which is examined next, but other factors could be at play. Regardless of the reason, the absence of gainfully employment within a household is an indicator that that there is a higher likelihood that the property is energy inefficient.

Income

To examine whether occupancy in a property with poor energy efficiency varies by income we apply the same methodology as applied earlier but use the CSO's Household Budget Survey (HBS) dataset instead of the census. The HBS is a representative random sample of all private households in the State collecting household income and expenditure data. The survey was most recently undertaken between August 2009 and September 2010 with 5,891 household participants (CSO (2012b)). Applying the BER estimation methodology yields similar results in terms of the distribution of energy efficiency, as shown in Figure 3. Table 11 reports estimated BER ratings by household disposable income. Families in the lowest disposable income categories have a higher likelihood of living in EFG rated properties. The likelihood that families in the two lowest income categories in Table 11 live in EFG rated properties is between 41-52% compared to 20-30% for all higher income categories.

Location

The physical concentration of energy inefficient properties has potential relevance both for policy makers and energy efficiency retrofit contractors. If inefficient properties are geographically concentrated rather than dispersed devising methods to identify and target relevant properties may differ, or when undertaking investment there may be economies of scale if inefficient properties are highly concentrated. The census data includes a number of spatial classifications. The first is agglomeration size. Across all town and village sizes there is a 68:32 ratio of ABCD:EFG properties. For rural areas this ratio is 58:42, so the likelihood of a rural property being more energy inefficient is higher. The ratio rises to 70:30 and above in towns ranging in size from 2,000 to 50,000 population though is somewhat less in larger cities, which may be a reflection of a greater proportion of much older properties. The most useful geographical classification in the census data is the Small Areas (SAs) classification, which are areas with between 50 and 200 dwellings, of which there are over 18,000 in the country. Examining the estimated BER ratings at this spatial resolution enables identification of locations with a high concentration of energy inefficient properties. Figure 4 maps the proportion of EFG properties by SA and shows that the midlands and north-west are the areas with high proportions of energy inefficient properties, i.e. rated EFG. To identify the locations with the highest concentration of energy inefficient properties the top 1000 Small Areas the greatest number of EFG rated properties were identified. These 1000 SAs contain roughly 110,000 properties, of which 90,000 are EFG rated. This represents 16% of all EFG properties and are concentrated within 6% of the census Small Areas. These areas are plotted in Figure 5 for the entire country and the urban areas of Dublin, Cork, Galway and Limerick are plotted in the subsequent figures. Looking at the country as a whole there is a somewhat unexpected concentration of EFG properties in the midlands. This reflects a high reliance on coal and peat as the primary fuel for heating within this region.

Multivariate Regression

The analysis to this point relied on cross-tabulations and while illustrative assumes a simplistic relationship between each attribute and building energy efficiency. We extend this analysis with multivariate regression to investigate how several variables are simultaneously associated with building energy efficiency. It is not a model of causality. We estimated a logistic model of building energy efficiency as a function of the household characteristics discussed above. The results are presented in Table 12 as odds ratios. The reference category for the regression is a household with a mortgage; a property that was built since 1991; occupants aged 19-74 that are not disabled and at least one occupant that is gainfully employed. We control in the regression for when a property was built, as we know from earlier that this has a major influence on the property's BER rating. Initially looking at the tenure categories, we find that compared to mortgaged properties that are owner occupied all other types of tenure are more likely to be EFG rated. Properties that are owned outright versus mortgaged are 37% more likely to be EFG rated. Low value rental properties, whether from private or local authority landlords, are more than twice as likely to be EFG rated compared to an owner-occupied mortgaged property. The likelihood declines somewhat for higher rental values.

Similar to the analysis earlier we find that older people are more likely to live in EFG rated properties. The likelihood that a property is EFG rated is 7% higher when an occupant is aged 75 or more. There is a similar likelihood when the head of household is aged 75 or over. However, when the logistic model was estimated using an age threshold of 65 rather than 75 the results were dramatically different. The likelihood that a property is EFG rated is 24% less when an occupant (or head of household) is aged 65 or more. We cannot provide a definitive explanation for this result but it is possible that the properties in which the older generation live are substantially different in an energy efficiency sense compared to the properties of the more recently retired.

Other distinguishing characteristics of the occupants of EFG properties are that the disabled are 6% more likely to live in such properties. Households where children are present are 9% less likely to live in EFG properties. And households where no occupant is in gainful employment are 10% more likely to live in EFG rated properties.

7. Discussion

As might be expected the energy performance of residential buildings is not randomly distributed across the population. Certain household characteristics such as age and tenure type are more highly associated with the most energy inefficient properties. The initial analysis suggested that there were some differences in the incidence of energy inefficient properties by tenure type and also to a limited extent with low rental value properties. The regression analysis suggests that the relationship is much stronger. Low rental value properties are twice as likely to be EFG rated as mortgaged, owner-occupied properties. There are some small differences in likelihoods depending on whether the landlords are private or local authorities but the overall message is the same; rental properties particularly in the low end of the market are substantially more likely to be EFG rated. In the lowest rental bracket, €0-150/week, properties are 120-140% more likely to be EFG rated compared to the reference category of mortgaged, owner-occupied properties. In higher rental brackets, which are predominantly owned by private landlords, the incidence is somewhat less but still 100% higher than the reference category. A clear policy implication is that any policy measures tackling residential building energy inefficiency should specifically target, if not prioritise, the rental sector, especially as it accounts for over 25% of the housing stock.

About 71% of the housing stock is owner-occupied, almost in equal proportions between owned outright and mortgaged financed. Another clear result of the analysis is that the second priority area for policy measures is properties that are owned outright. These properties are 37% more likely to be EFG rated compared to mortgage financed properties.

The earlier analysis indicated a number of vulnerable groups are more likely to live in energy inefficient properties. These results were confirmed in the multivariate analysis: people aged 75 and above, the disabled, and household units without anybody gainfully employed are to varying degrees more likely to live in EFG rated properties. While these results are not unexpected they pose practical difficulties for policy makers. These same groups may be less inclined to engage with policy measures and incentives to improve the energy efficiency of their homes. The earlier evidence also suggested that income is likely to be a relevant factor in terms of who is more likely to live in the most energy inefficient properties. Households with limited incomes are less likely to be in a position to engage in support schemes, such as subsidies, as invariably such schemes require co-funding from households. Energy efficiency policy measures or support schemes may need to actively target these vulnerable groups, otherwise a significant component of the EFG rated housing stock may not be accessible for upgrade. For instance, roughly 17% of the EFG rated properties are occupied by people aged 75 and over.

8. Summary

The EU directive on energy performance of buildings (EP and CEC (2002)) established a methodological framework for calculating the energy performance of buildings. Across Europe residential properties are being assessed for energy efficiency and the associated EPC databases contain a wealth of information that is potentially useful in developing policy measures to improve energy efficiency in the residential sector, as well as estimating the energy efficiency performance of national building stocks. Because EPC databases are relatively recent and as the properties assessed to date are not necessarily representative of the housing stock, the EPC database may misrepresent the status of energy efficiency across the residential housing stock. The analysis here suggests that the EPC database in the Republic of Ireland, with less than 30% of residential properties assessed, is over-represented with more energy efficient properties. Relying solely on the EPC database to inform policy would suggest a higher level of energy efficiency in the housing stock than reality and thereby provide faulty baseline from which to devise policy actions.

The paper presents a methodology to better estimate the energy efficiency performance of the housing stock using EPC and census data that should be easily replicable in other EU member states, where the same policy context exists in terms of developing evidence-based strategies for mobilising investment in building energy efficiency. Utilising census data adds a depth of policy relevant information that is not available in EPC data, at least in Ireland and most likely in other countries too. From the EPC databases we can compare EPC ratings against a range of engineering characteristics of properties but by combining that information with census data enables to assess the circumstances of people living in energy inefficient properties.

The data provide evidence to support the view that specific household characteristics, such as age and tenure type, are more highly associated with the most energy inefficient properties. Rental properties, and particularly low value rentals, have the highest likelihood of being the least energy efficient compared to other tenure categories. Over 17% of the total residential housing stock comprises low value rental property[8] in roughly equal proportions split between private and public housing. Therefore, a significant proportion of EFG rated properties is likely to be public sector ownership. Identifying EFG properties in public sector ownership may be burdensome but not difficult. The difficulty in improving the energy efficiency of such properties will be in securing the investment finance.

Combining EPC with census data facilitates spatial analysis.[9] The mapping of EFG properties revealed that the problem of energy inefficient properties is highly concentrated spatially. The maps highlighted concentrations of the most energy inefficient properties within relatively small areas. In many instances these locations are likely to coincide with housing estates. Within housing estates properties are likely to be very similar in terms of build materials and quality. It seems plausible to assume that there are cost efficiencies feasible if investments in adjacent properties occurred around the same time. For instance, because adjacent properties are likely to be similar, knowledge gained by contractors between properties will help improve their work efficiency and deliver economies of scale. With multiple property owners the practical difficulties of realising such scale economies may be difficult, though public sector housing may be one area where such economies could be more easily realised.

Acknowledgements

We acknowledge the Central Statistics Office (CSO) for access to an anonymous micro data file of the 2011 Census of Population data and we thank Gerry Walker for assistance with the data and clearance. We thank Seán Lyons, Alan Barrett and participants at seminars in ESRI for helpful comments and suggestions. We gratefully acknowledge financial support from the ESRI Energy Policy Research Centre.

Table 1: BER classification for residential buildings in Ireland

Rating	kWh/m2/yr	Frequency	%
A1	<= 25	9	0.0

[8] Just over 25% of the housing stock is in the rental sector.
[9] To date Irish EPC data has not been spatially coded.

A2	> 25	129	0.0
A3	> 50	2,209	0.5
B1	> 75	6,844	1.7
B2	> 100	15,182	3.7
B3	> 125	32,482	7.9
C1	> 150	44,938	11.0
C2	> 175	50,898	12.4
C3	> 200	53,350	13.0
D1	> 225	54,541	13.3
D2	> 260	48,794	11.9
E1	> 300	27,887	6.8
E2	> 340	21,836	5.3
F	> 380	21,693	5.3
G	> 450	29,544	7.2

Source: SEAI, February 2014

Table 2: Descriptive Statistics, Building Energy Rating Assessments

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
BER kWh/m2/year	410348	261.34	143.9	3.9	1997
log(BER)	410348	5.458	0.445	1.358	7.6
Year property built					
Pre 1919	410348	0.061	0.239	0	1
1919-1945	410348	0.059	0.235	0	1
1946-1960	410348	0.062	0.242	0	1
1961-1970	410348	0.055	0.227	0	1
1971-1980	410348	0.113	0.316	0	1
1981-1990	410348	0.105	0.306	0	1
1991-2000	410348	0.185	0.388	0	1
2001-2005	410348	0.189	0.392	0	1
2006-	410348	0.172	0.377	0	1
Property type					
House	410348	0.350	0.477	0	1
Semi-detached house	410348	0.246	0.431	0	1
Terrace house	410348	0.195	0.396	0	1
Apartment	410348	0.200	0.400	0	1
Other - maisonette etc.	410348	0.009	0.097	0	1
Main fuel for space heating					
Oil	410344	0.408	0.492	0	1
Mains gas	410344	0.376	0.484	0	1
Electricity	410344	0.141	0.348	0	1
Coal	410344	0.009	0.094	0	1
Peat	410344	0.002	0.049	0	1
Liquid Petroleum Gas (LPG)	410344	0.016	0.127	0	1

Wood	410344	0.003	0.058	0	1
Other	410344	0.031	0.173	0	1

Source: SEAI, February 2014

Table 3: Regression estimates

Variable	Coefficient	Standard Error	
Pre 1919	Ref		
1919-1945	-0.079	0.003	***
1946-1960	-0.132	0.003	***
1961-1970	-0.250	0.003	***
1971-1980	-0.336	0.003	***
1981-1990	-0.426	0.003	***
1991-2000	-0.523	0.003	***
2001-2005	-0.638	0.003	***
2006-	-0.900	0.003	***
Detached house	Ref		
Semi-detached house	0.030	0.001	***
Terrace house	-0.039	0.002	***
Apartment	-0.049	0.002	***
Other, maisonette etc.	-0.044	0.005	***
No central heating	Ref		
Oil	0.120	0.006	***
Mains gas	0.092	0.005	***
Electricity	0.563	0.006	***
Coal	0.683	0.008	***
Peat	0.657	0.013	***
LPG	0.084	0.006	***
Wood	0.115	0.014	***
Other	0.662	0.006	***
Constant	5.757	0.006	***
Dependent variable: Log(BER)			
R^2 = 0.556			
Number of observations = 410,344			

Ref indicates reference category, *** p<0.01

Table 4: Within-sample BER predictions

		Model Predictions, dwellings						
	BER ratings	B	C	D	E	F	G	Total
Actual BER Assessments	A	1,760	137	426	12	3	9	2,347
	B	33,821	14,099	4,433	2,047	39	69	54,508
	C	22,982	79,983	34,702	11,174	174	171	149,186
	D	1,207	34,678	43,883	21,036	1,460	1,071	103,335
	E	85	4,527	20,608	20,429	2,170	1,904	49,723
	F	26	609	6,550	11,443	1,374	1,691	21,693
	G	28	222	3,537	11,286	2,311	12,160	29,544
	Total	59,909	134,255	114,139	77,427	7,531	17,075	410,336
				Row Percentages				
	A	75.0	5.8	18.2	0.5	0.1	0.4	100.0

	B	62.0	25.9	8.1	3.8	0.1	0.1	100.0
	C	15.4	53.6	23.3	7.5	0.1	0.1	100.0
	D	1.2	33.6	42.5	20.4	1.4	1.0	100.0
	E	0.2	9.1	41.4	41.1	4.4	3.8	100.0
	F	0.1	2.8	30.2	52.7	6.3	7.8	100.0
	G	0.1	0.8	12.0	38.2	7.8	41.2	100.0

Table 5: E, F, G threshold, within-sample BER predictions

			Model Predictions		
			A-D	E-G	Total
Actual BER Assessments	A, B, C or D (ABCD)	No.	272,111	37,265	309,376
		%	66.3	9.1	75.4
	E, F or G (EFG)	No.	36,192	64,768	100,960
		%	8.8	15.8	24.6
		Total	308,303	102,033	410,336
		%	75.1	24.9	100.0

Table 6: Descriptive statistics of the BER ratings distribution

	Assessed BERs	Estimated BERs
Data	BER database	Census data
Observations	0.410 million	1.65 million
Mean (kWh/m^2/year)	261	276
Median (kWh/m^2/year)	225	255
Skewness	3.06	1.29
Kurtosis	19.24	4.8

Table 7: BER ratings by residential tenure type

Estimated BER rating	ABCD	EFG
Residential Housing Stock	65%	35%
Owner Occupier	67%	33%
Rental	62%	38%
Rent free	41%	59%
Rental Accommodation	61%	39%
Private Landlords	60%	40%
Local Authority Landlords	64%	36%
Voluntary/Co-op	66%	34%

Table 8: BER ratings by weekly rent

BER rating	<€50	<€100	<€150	<€200	<€250	<€300	<€500	€500+
EFG	39%	42%	39%	35%	34%	33%	38%	45%

Table 9: Estimates of the elderly living in energy inefficient properties

	Age 60+			Age 65+			Age 75+		
	Properties	Persons	% 60+ persons	Properties	Persons	% 65+ persons	Properties	Persons	% 75+ persons
EFG	232,812	318,897	45%	180,787	236,273	48%	95,761	112,807	57%

Table 10: Proportion of households with gainfully employed, by BER rating

Estimated BER Rating	Gainfully employed occupant(s)	No gainfully employed occupants
ABCD	70	56
EFG	30	44

Table 11: Weekly Household Disposable Income by estimated BER rating

Estimated BER Rating	<€250	€250-500	€500-750	€750-1000	€1000-1250	€1250-1500	€1500-2000	>€2000
ABCD	48%	59%	71%	75%	77%	79%	79%	73%
EFG	52%	41%	29%	25%	23%	21%	21%	27%
Households	606	1,106	1,246	909	707	496	492	329

Source: Author calculations and HBS.

Table 12: Logistic Regression, Energy Inefficient housing

Number of observations = 1558386 Log likelihood = -553609.12	LR chi2(14) = 843168.03 Prob > chi2 = 0.0000 Pseudo R2 = 0.4323		
Explanatory Variables	Odds Ratio	Standard Error	
Own outright	1.372	0.009	***
Private landlord & rent €0-150	2.231	0.022	***
Private landlord & rent €150-250	2.003	0.021	***
Private landlord & rent €250+	1.639	0.028	***

Local Authority landlord & rent €0-150	2.438	0.023	***
Local Authority landlord & rent €150+	1.779	0.087	***
Other tenure	2.186	0.031	***
-1945	669.985	11.355	***
1946-1970	8.540	0.056	***
1971-1990	1.999	0.013	***
Age 75+	1.072	0.009	***
Disabled person in house	1.062	0.006	***
Child(ren) in house	0.914	0.005	***
No gainful employment	1.095	0.006	***
Constant	0.069	0.000	***

Reference category: Own with mortgage, built 1991+, occupants aged 19-74, an occupant gainfully employed*** p<0.01

Figure 1: BER Prediction Errors, sorted in increasing value

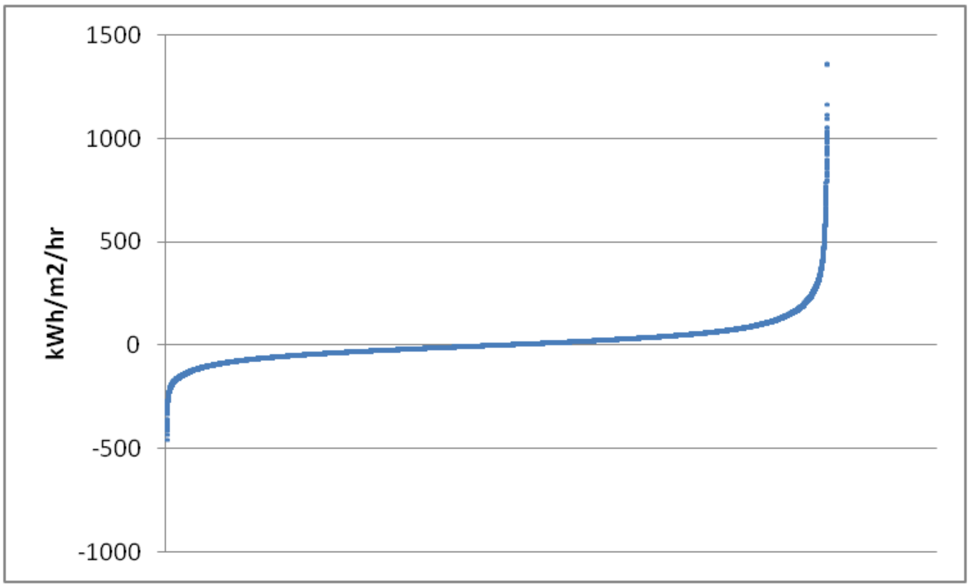


Figure 2: Distribution of estimated and assessed BER ratings

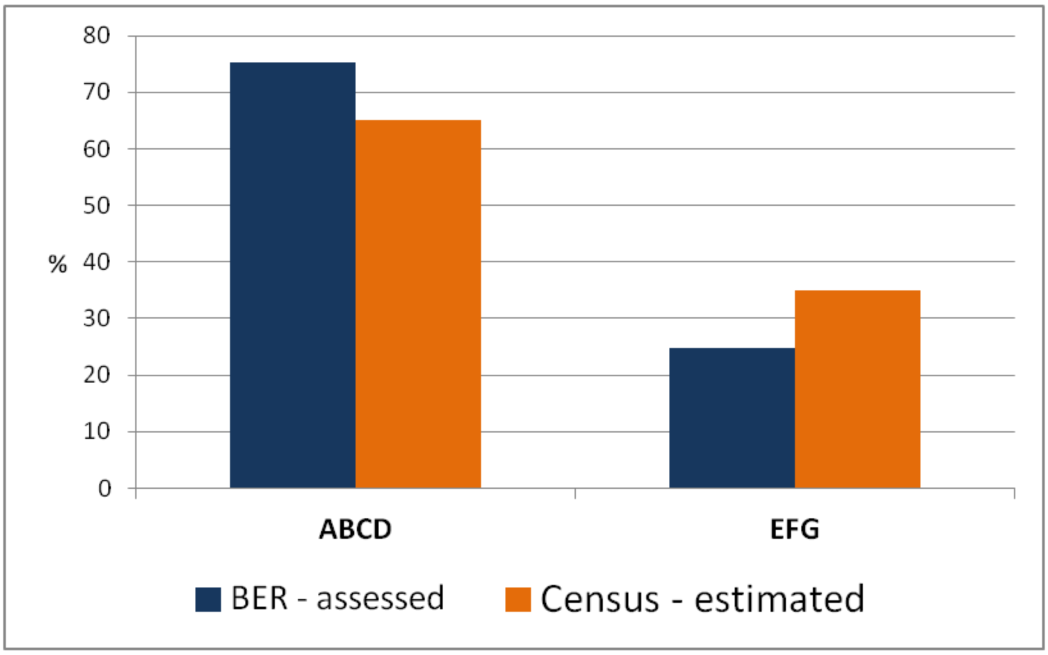


Figure 3: Distribution of estimated BER ratings, Census and Household Budget Survey

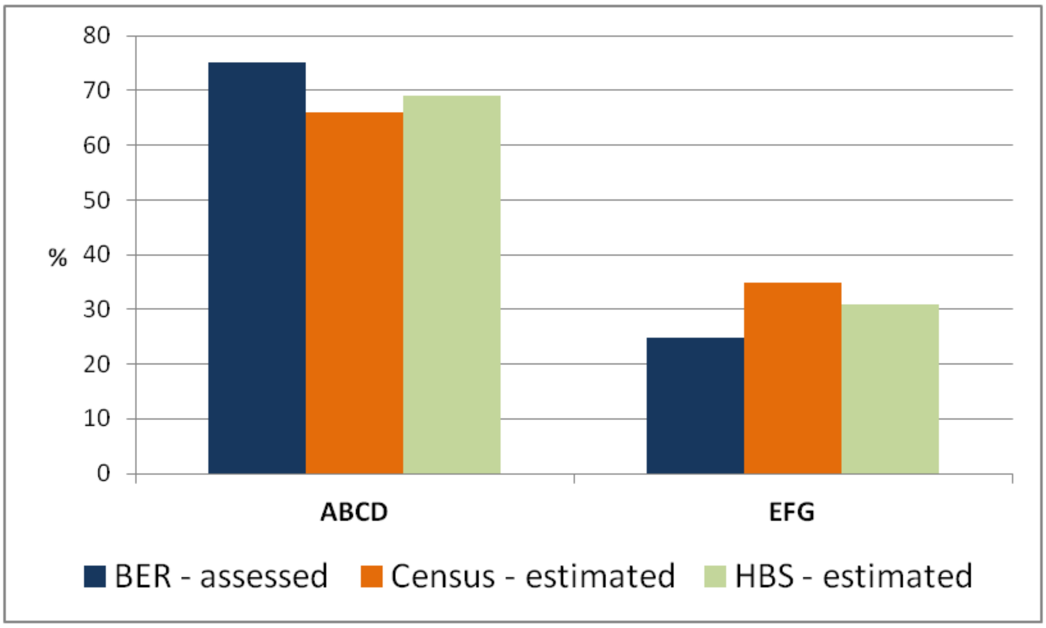


Figure 4: Proportion of properties with estimated EFG ratings by Small Area

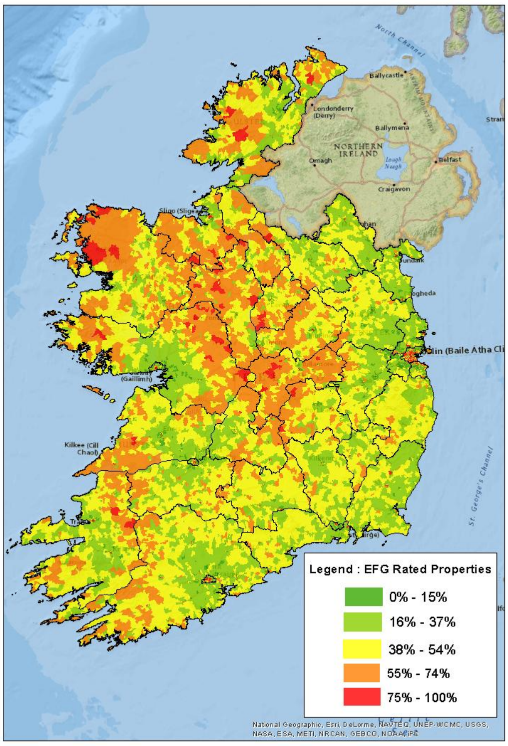


Figure 5: Small Areas with highest number of EFG rated properties - Ireland

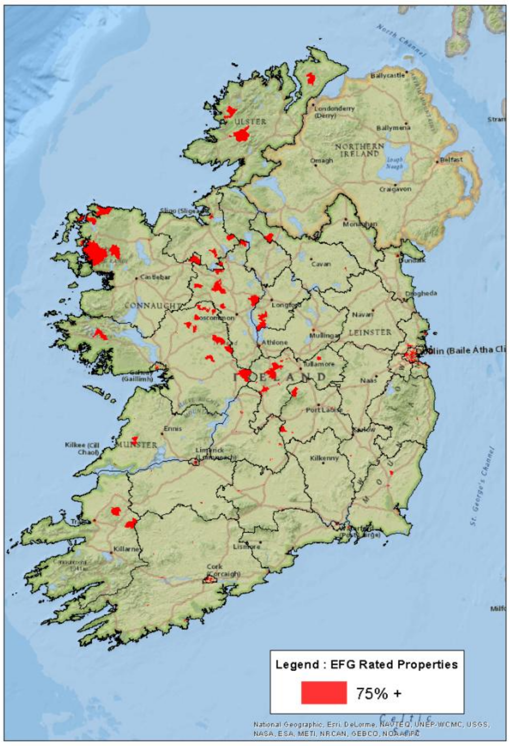


Figure 6: Small Areas with highest number of EFG rated properties - Dublin

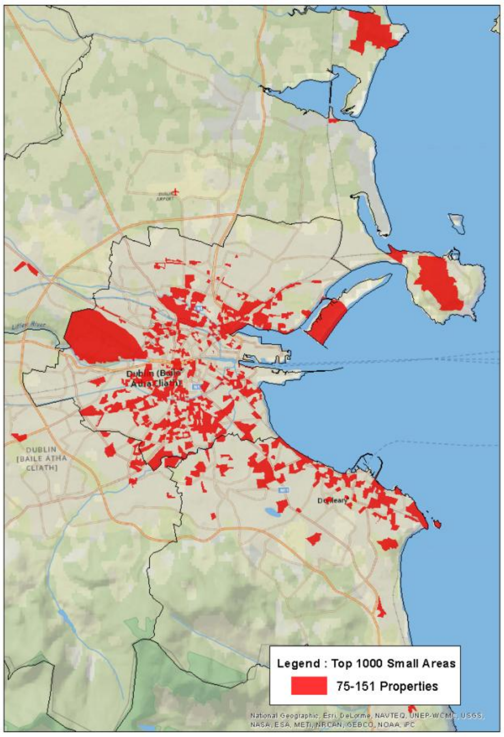


Figure 7: Small Areas with highest number of EFG rated properties - Cork

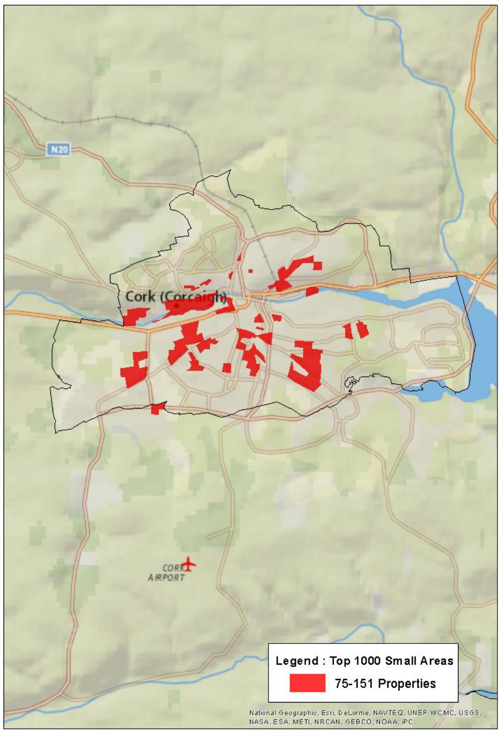


Figure 8: Small Areas with highest number of EFG rated properties - Galway

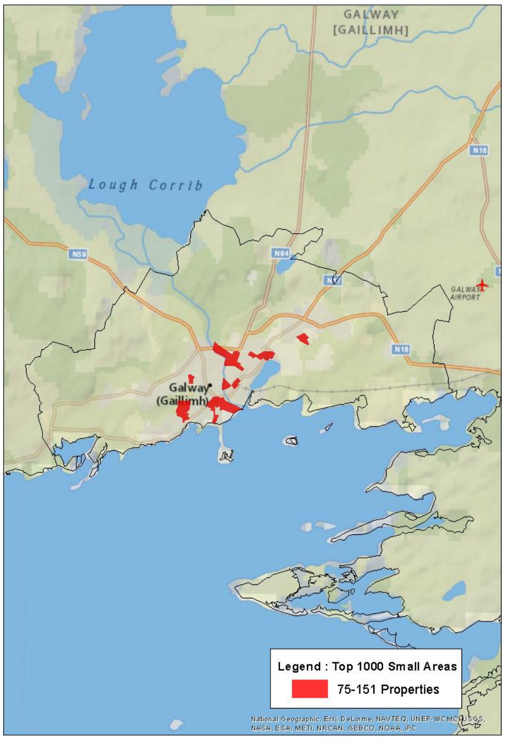
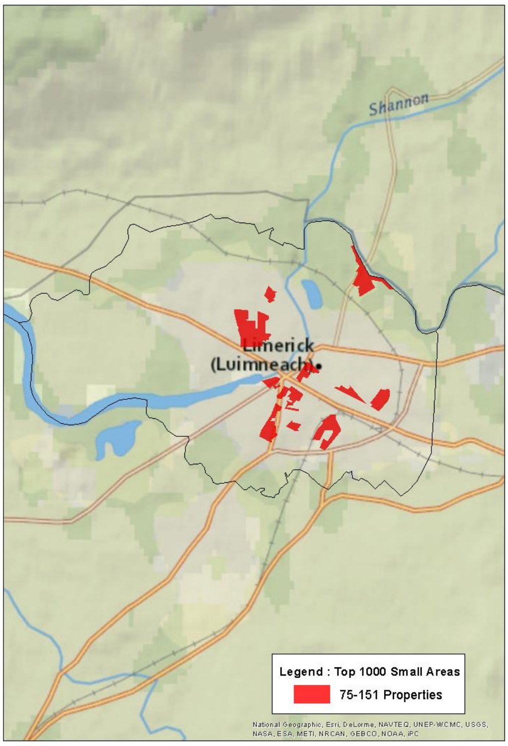


Figure 9: Small Areas with highest number of EFG rated properties - Limerick



References

ADEME. 2009. Energy Efficiency Trends and Policies in the EU 27, ADEME, Paris.

Australian Bureau of Statistics. 2008. Energy efficiency rating and house price in the ACT, Department of the Environment, Water, Heritage and the Arts, Canberra.

Brounen, D. and Kok, N. 2011. "On the economics of energy labels in the housing market", Journal of Environmental Economics and Management, vol. 62, no. 2, pp. 166-179.

Cajias, M. and Piazolo, D. 2013. "Green performs better: energy efficiency and financial return on buildings", Journal of Corporate Real Estate, vol. 15, no. 1, pp. 53-72.

CEC. 2010. Europe 2020 - A strategy for Smart, Sustainable and Inclusive Growth, European Commission, Brussels, COM(2010) 2020.

CEC. 2011a. A resource-efficient Europe – Flagship initiative under the Europe 2020 Strategy, European Commission, Brussels, COM(2011) 21.

CEC. 2011b. Energy Efficiency Plan 2011, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2011) 109.

CSO. 2012a. Census 2011 Profile 4 - The roof over our heads, Stationery Office, Dublin, Ireland.

CSO. 2012b. Household Budget Survey 2009-2010, Central Statistics Office, Dublin.

DCENR. 2012. Ireland's second National Energy Efficiency Action Plan to 2020, Department of Communications, Energy and Natural Resources (DCENR), Dublin.

DCLG. 2010. English Housing Survey: Housing stock report 2008, Department for Communities and Local Government (DCLG), London.

Eichholtz, P., Kok, N., and Quigley, J. M. 2010. "Doing Well by Doing Good? Green Office Buildings", American Economic Review, vol. 100, no. 5, 2492-2509.

EP and CEC. 2002. Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings.

EP and CEC, 2010. "Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast)",

EP and CEC. 2012. Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC.

Greene, W. H. 2002, Econometric Analysis. 5th edn, Prentice Hall, New Jersey.

Huber, P. J. 1967. The behavior of maximum likelihood estimates under nonstandard conditions, Vol. 1 of Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, pp. 221-233.

Hyland, M., Lyons, R. C., and Lyons, S. 2013. "The value of domestic building energy efficiency - evidence from Ireland", Energy Economics, vol. 40, pp. 943-952.

Kok, N. and Jennen, M. 2012. "The impact of energy labels and accessibility on office rents", Energy Policy, vol. 46, no. 0, pp. 489-497.

Reichardt, A., Fuerst, F., Rottke, N. B., and Zietz, J. 2012. "Sustainable Building Certification and the Rent Premium: A Panel Data Approach", Journal of Real Estate Research, vol. 34, no. 1, 99-126.

White, H. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity", Econometrica, vol. 48, pp. 817-838.

Year	Number	Title/Author(s) ESRI Authors/Co-authors *Italicised*
2014		
	488	Gaming in the Irish Single Electricity Market and Potential Effects on Wholesale Prices *Darragh Walsh* and *Laura Malaguzzi Valeri*
	487	Household Formation and Tenure Choice: Did the great Irish housing bust alter consumer behaviour? *David Byrne, David Duffy* and *John FitzGerald*
	486	Changing Time: Possible Effects on Peak Electricity Generation Sara Crowley (UCC), *John FitzGerald* and *Laura Malaguzzi Valeri*
	485	Changes in Labour Market Transitions in Ireland over the Great Recession *Adele Bergin, Elish Kelly, Seamus McGuinness*
	484	Examining the Relationships between Labour Market Mismatches, Earnings and Job Satisfaction among Immigrant Graduates in Europe *Seamus McGuinness* and Delma Byrne
	483	Worth a Try: A Statistical Analysis of Brian O'Driscoll's Contribution to the Irish Rugby Team *Pete Lunn* and *David Duffy*
	482	Who Should We Ask? Employer and Employee Perceptions of Skill Gaps within Firms *Seamus McGuinness* and Luis Ortizd
	481	Trends in Socio-Economic Inequalities in Mortality by Sex in Ireland from the 1980s to the 2000s *Richard Layte, Joanne Banks*
	480	Innovation and Productivity in Services: Evidence from Germany, Ireland and the United Kingdom Bettina Peters, Rebecca Riley, *Iulia Siedschlag,* Priit Vahter and *John McQuinn*
	479	What Can I Get For It? The Relationship Between the Choice Equivalent, Willingness to Accept and Willingness to Pay *Pete Lunn* and Mary Lunn
	478	Changes in Household Fuel Expenditure Associated with Improvements in Building Energy Efficiency *John Curtis* and Anne Pentecost

For earlier *Working Papers* see

http://www.esri.ie/publications/search_for_a_working_pape/search_results/index.xm