

Accepted refereed manuscript of:

Comerford DA, Lange I & Moro M (2018) Proof of concept that requiring energy labels for dwellings can induce retrofitting, *Energy Economics*, 69, pp. 204-212.

DOI: [10.1016/j.eneco.2017.11.013](https://doi.org/10.1016/j.eneco.2017.11.013)

© 2017, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Proof of Concept that Requiring Energy Labels for Dwellings Can Induce Retrofitting

David A. Comerford*, Ian Lange[†], Mirko Moro[‡]

Abstract: How to induce households to install energy efficient technology remains a puzzle. Could an energy labelling requirement for residential real estate help? We propose that the salient color-letter grades on the English Energy Performance Certificate (EPC) served as targets, motivating vendors to invest in energy efficiency. To test our hypothesis we look to a random sample of over 16,000 homes in England. In the post-EPC data we find a cluster of homes with energy efficiency scores just above the D-grade threshold. This cluster was not present prior to the requirement, replicates in an independently-drawn random sample and is significantly larger amongst properties that can be identified as treated by the EPC requirement. We conclude that the EPC requirement induced investment, and hence that energy efficiency labels have potential to green the housing stock. We infer from our analysis how the design of the EPC could be altered to motivate greater investment in energy efficiency.

* corresponding author: david.comerford@stir.ac.uk; Division of Economics, University of Stirling, Stirling, FK9 4LA, UK, david.comerford@stir.ac.uk, 01786467317;

[†] ilange@mines.edu; Division of Economics and Business, Colorado School of Mines, Golden, CO, 80401, USA;

[‡] mirko.moro@stir.ac.uk; Division of Economics, University of Stirling, Stirling, FK9 4LA, UK. The paper benefited from comments from Sebastien Houde, Jim Sallee and Harrison Fell as well as from participants at the 2017 Royal Economics Society conference, 2016 Association of Environmental and Resource Economics meeting and the AERE session of the 2015 Midwest Economics Association Conference, at seminars at USA Environmental Protection Agency's National Center for Environmental Economics, University of Manchester, University of Stirling, Heriot Watt University, and Colorado School of Mines.

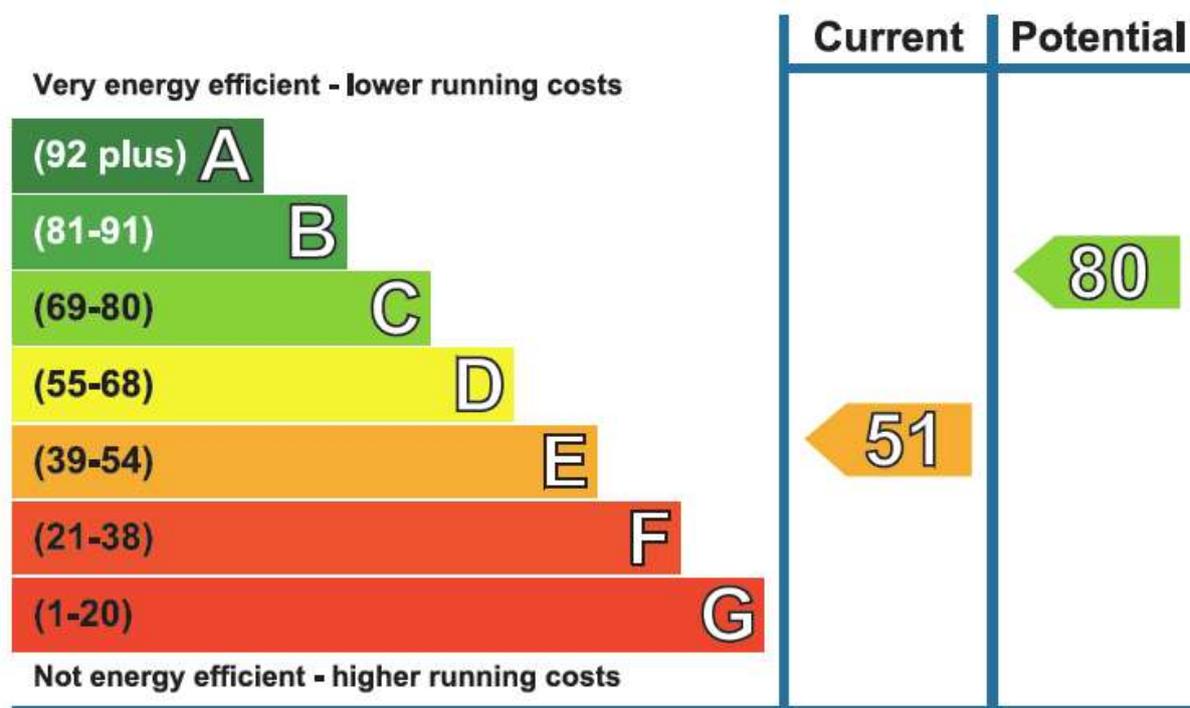
Keywords: Energy labels; Retrofitting; Selective Attention; Residential Energy Use; Notches; Energy Efficiency Gap

Introduction

There are two pathways to reducing household energy consumption: curtailment and retrofitting. While several prominent journal articles have pointed to curtailment in energy use induced by giving households feedback on their energy consumption (Allcott and Mullainathan, 2010; Allcott and Rogers, 2014, Costa and Kahn, 2013; Dolan and Metcalfe, 2013), retrofitting is the more effective, according to engineering models (Dietz et al, 2009; Gardner and Stern, 2008) and randomized control trials (RCT)¹. But retrofitting is rare: it is difficult to induce households to modify their property. Even when the monetary and non-monetary costs of installing energy efficient technologies are heavily reduced, few households retrofit (Allcott and Rogers, 2014; Fowlie, Greenstone and Wolfram, 2015a). Here we test whether retrofitting was induced by the requirement that homes offered for sale or rent in England display an Energy Performance Certificate (henceforth, EPC requirement). The contribution of this research is that it demonstrates a causal effect of energy labeling on investment in energy-saving technologies. Moreover, the observed effect cannot be accounted for by monetary incentives alone, and so our research sheds light on some non-monetary costs and benefits that influence the retrofitting decision. The explanation that best fits the observed effect is that investment is motivated by the goal to boost a home's energy rating to an arbitrary, though salient, reference point: the next color-coded letter grade on the visual display of the EPC. A sample EPC is presented in Figure 1. Our explanation is consistent with results showing that consumers selectively attend to more salient attributes at the expense of more diagnostic attributes (e.g. Gabaix and Laibson, 2006; Lacetera, Pope and Sydnor, 2012; Englmeier et al., 2017).

¹ The Weatherization Assistance Program, which retrofitted the homes of low income households, generated reductions of 7-8% in all forms of energy consumption (Fowlie et al, 2015b; Graff Zivin and Novan, 2015). The curtailment induced by the Opower study, which treated households with feedback on their electricity consumption relative to neighbors, generated 2% reductions in household electricity consumption only (Allcott and Rogers, 2014).

Figure 1: An English Energy Performance Certificate, showing SAP score (51) and letter grade (E)



Our identification strategy in this study is to look to a sample for which a marginal investment in energy efficiency will have particularly large positive impacts on the *appearance* of the resultant EPC. The English EPC reports home energy efficiency as both a 0-100 Standard Assessment Procedure (SAP) score and a concomitant color-coded letter grade (A green – G red). Previous research on selling prices in the English housing market has found that buildings with higher color-letter grades on the EPC enjoy a price premium (Fuerst et al, 2015). At certain initial SAP scores, a small investment in energy efficient technologies would shift a property into a higher color-coded letter grade on the EPC. We predict that some vendors will have invested in just enough energy efficiency so as to reach the next letter-grade. The testable implication is that, after the EPC requirement comes into effect, we will see a cluster of homes at the lowest point in a color-coded letter-grade.

To test for these clusters, we use data from the English House Condition Survey, which from 2002 onwards recorded SAP scores. We find a significant cluster at the lowest point in the D color-letter grade (55 SAP points) in the post-EPC data, but no cluster at this point in the distribution in the pre-EPC data. We rule out that this cluster was induced by planning requirements or that it was driven by new-build homes. We also replicate it in an independent sample. Additionally, when we restrict the sample to homes that had been treated by the EPC requirement – homes which had been traded since the EPC requirement – we find that this cluster is three-times as large and that there is a concomitant deficit of homes at the highest point in the E color-letter grade.

In the next section we present the background to our analysis. Section 3 reports our methods. Section 4 presents results. Section 5 discusses these results, paying particular attention to the question whether gaming could explain the observed clusters. We conclude that the EPC requirement prompted home vendors to make marginal investments so as to shift their homes to 55 SAP points. A necessarily crude but conservative back-of-the-envelope analysis of the cost savings induced by these marginal investments calculates them at £11.4m per annum.

1. Background

2.1. Prior research on inducing retrofitting

Despite its high expected benefits to both the private household and to society at large, retrofitting is a behavior that has proven difficult to induce. Fowlie et. al. (2015a) report the results of a resource-intensive RCT in which 7000 households, which were eligible for the Weatherization Assistance Program, were treated to an in-person visit by a field worker from their own community who explained the benefits of a retrofit of their heating and cooling

infrastructure, explained that the retrofit would be free of charge and offered to help the household complete the paperwork to apply for the Weatherization Assistance Program. The campaign did increase uptake of Weatherization relative to a control group, but only to 6 percent of eligible households, and at a cost of over \$1000 per weatherized home.

Qualitatively similar results come from the Opower RCT. Though the Opower RCT is justifiably cited as a success story in reducing energy consumption (e.g. Allcott and Mullainathan, 2010; Allcott and Rogers, 2014, Costa and Kahn, 2013), it was not successful at inducing retrofitting. Opower offered all their customers rebates on certain energy efficient purchases, e.g. \$50-\$75 for a washing machine; up to \$5000 for home insulation.

Additionally, households in the treatment group of the Opower RCT were sent bills that delivered feedback on their energy use and tailored recommendations on how to reduce energy consumption. For example, a household that consumes electricity heavily in summer would be shown the potential cost savings afforded by a new air conditioning unit. We know that this treatment group was motivated to reduce energy consumption because the results of the RCT show that they did reduce energy consumption through curtailment. Despite their motivation, the tailored information, and the offer of rebates, only 4.8 percent of them (compared to 4.4 percent in the control group) claimed rebates on energy efficient purchases (Allcott and Rogers, 2014).

Additional evidence of the stubbornness of retrofitting comes from the failure of the *Green Deal* in the UK. The *Green Deal* set aside £540m in public money as loans to private households toward retrofitting (Palmer, 2015). It was designed in consultation with the Behavioural Insights Team (BIT) and was intended to overcome loss aversion by allowing participants repay the loan out of savings on their energy bills. Notwithstanding its behaviourally-informed design, the scheme was scrapped in 2015 due to lack of interest. In

short, there is uncertainty regarding how to reduce the non-monetary costs that inhibit investment in energy efficient technology.

2.2. The EPC Requirement

A 2002 directive from the European Union (EU directive 2002/91/EC) required member states to ensure (1) certification of a building's energy performance, (2) that EPCs are made available when a building is constructed, sold, or rented, and (3) that these EPCs are comparable across member states. In the UK, this directive was passed into law through the Housing Act of 2004. Specifically, the law requires that, prior to a property being placed on the market, it is audited for energy efficiency. The vendor or lessor is obliged to show the resultant EPC to potential buyers or renters. In practice, EPC's tend to be included in advertising material for the property.

The EU Energy Label, codified in EU Directive 92/75/EC, is the presentation format that was adopted for the EPCs (for an example, see Figure 1). Crucial to this analysis, the UK grafted the 7 color-letter grades of the EU Energy Label on top of the UK's pre-existing measure of energy efficiency, the SAP score. Each letter-grade on the EPC spans a dozen or so SAP scores, so for example SAP scores from 39 – 54 are E grade (orange), and those from 55 – 68 are D grade (yellow).

The SAP score was developed in 1992 by the Building Research Establishment (BRE - at that time a government-funded research laboratory) to help the UK government monitor progress in residential energy efficiency. The SAP score measures “how much energy a dwelling will consume, when delivering a defined level of comfort and service provision” in terms of space heating, water-heating and lighting (Department of Energy and Climate Change, 2013). The SAP score is calculated based on a standardized audit of the building's fabric and fittings; the auditor records data that are subsequently entered into an algorithm and the algorithm

delivers the final SAP score. The SAP scores used in our analysis are drawn from audits conducted for the English House Condition Survey. These lasted on average one hour and were conducted by specially trained auditors who took notes and photographs, which were later cross-checked (Department of Community and Local Government. 2010). The SAP is what Stavins et al (2013) term an “energy integrity score,” meaning the score is entirely derived from building attributes. The SAP score is not confounded by the occupants’ use of the building or their behavior.

Though SAP scores existed prior to the EPC requirement, assessments were rare and their results were not meaningful to the public (is an SAP of 62 good or bad?). By contrast, an EPC is available for all properties built or placed on the market since December 2007, making it easy for the layperson to compare the energy efficiency of any two buildings.

2.3. Proposed Mechanism

The EPC audit includes recommendations on cost-effective means to improve energy efficiency (Figure 2). We propose that the news contained in that audit is particularly likely to be acted on for reasons of low marginal (non-monetary) costs and high marginal benefits.

Figure 2: Retrofitting Recommendations as Presented in an Energy Performance Certificate

Recommended measures	Indicative cost	Typical saving per year	Rating after improvement	
			Energy	Environment
1 Replace boiler with new condensing boiler	£2,200 - £3,000	£57	C 77	C 76
2 Solar water heating	£4,000 - £6,000	£32	C 78	C 78
3 Solar photovoltaic panels, 2.5 kWp	£5,000 - £8,000	£251	B 88	B 87

Fowlie et al (2015a) conclude from the limited success of their intensive intervention to increase uptake of the Weatherization Assistance Program that there must be high non-

monetary costs that inhibit retrofitting. Hassle is a sizable non-monetary cost of insulating a loft for example, according to a BIT investigation (BIT, 2011). When one's home is tidy and one's time is routinized, it is off-putting to schedule a work crew; empty the loft of the stuff that has accumulated; and tidy the stuff away while the work crew lays the insulation.

Because non-monetary costs are less fungible than monetary costs, there are moments when non-monetary costs are lower than usual. The moment at which property owners are given the results of the energy efficiency audit is when they are already preparing their homes for presentation on the market. When putting one's home on the market, one has to empty the loft anyway. Generally, the marginal costs of retrofitting in terms of hassle and disruption are lower when moving house.

Additionally, the benefits of investing in energy efficient technology can appear unusually high when preparing a home for the market. Around the time of selling a home, it is common for households to make changes to a property so as to make it more attractive to potential buyers or renters. Since the EPC must by law be shown to prospective buyers and renters, the EPC matters for presentation of the property. Whereas investment in home energy efficiency is typically presented as paying for itself through savings on energy bills over the long-term, the investment recommendations in the EPC audit can pay for themselves almost immediately through a higher selling price.

It is this marginal benefit mechanism that leads to the testable hypothesis of the current research. By construction, all along the distribution of SAP scores, a one point increase has the same expected value in terms of cost savings. In terms of salience, however, the expected value of a one point increase in SAP score is arbitrarily higher at certain points in the distribution. Lacetera, Pope and Sydnor (2012) show that selective attention leads to discontinuities in the relationship between mileage of a second-hand car and the price it achieves at auction. Specifically, an extra mile driven inflicts a more sizable reduction in

auction price when it moves an odometer from 9999 to 10000 than when it moves an odometer from 9986 to 9987. Similarly, an extra SAP point will be especially salient when it moves from 54 – an orange E grade – to 55 – a yellow D grade on the EPC.

Consumer Focus (2011), a British non-profit consumer advocacy group, ran a number of focus groups to learn which aspects of the EPCs were most and least helpful. Many respondents reported that the colors (green is good, red is bad) and the letters (A is good and F is bad) were how they utilized the information given. This finding is supported by a choice experiment conducted in the USA. Newell and Siikamaki (2014) compared people's willingness to purchase household appliances depending on the manner in which energy efficiency ratings were conveyed. They included one condition that used a color-letter grade very similar to that on the EPC. They found that the letter grades had a “powerful effect” on the likelihood of choosing an energy efficient product: “anyone who has been to school wants to get an A rather than a C or F” (Newell and Siikamaki, 2014).

EPC letter grades are also reflected in the selling prices of homes in England. Fuerst et al (2015) apply a hedonic model and find significant selling price premia for homes at higher letter grades and discounts for those at lower letter grades. In short, there is evidence to suggest that consumers are selectively attentive to the color-letter grades, and might be less sensitive to the SAP score, even though the SAP score is more informative. Selective attention of this sort has been documented in purchasing decisions of other costly consumer durables – second-hand cars (Lacetera et al, 2012) and household appliances (Houde, 2014a). The latter finds that consumers pay a premium over and above anticipated cost-savings to buy a refrigerator with *Energy Star* certification. We anticipate a similar salience effect in the UK property market. The fact that advertisements on market-leading property websites in England such as Rightmove.co.uk and Zoopla.co.uk routinely report only the letter grade of

the EPC, and not the SAP score, will likely compound this salience effect.² Relative to previous research, we take the implications of selective attention one step further by looking for a supply-side response. Our hypothesis is that vendors will be more likely to retrofit if doing so will boost their home's SAP score into the next letter grade. If our hypothesis is supported, it demonstrates that vendors of homes are acting as though potential purchasers are myopic (Gabaix and Laibson, 2006). Vendors invest in anticipation that consumers are selectively attentive to the color-code of the EPC and insufficiently attentive to the SAP score that underpins it.

Table 1: The increase in SAP points associated with various investments

Improvement	Rating can be improved by	Estimated Savings	Estimated costs for a four bed house, including installation	Cost per SAP point gained
Condensing Boiler	47 SAP points	£225+ per year	£2 - £6k (source: mumsnet)	£42-128
Cavity Insulation	13 SAP points	£100-£125 per year	£720 (source: Which?.co.uk)	£55
Roof Insulation	10 SAP points	£100-£125 per year	£395 (source: Which?.co.uk)	£39.50
Cylinder thermostat & Insulation	8 SAP points	£100-125 per year	£150 (source: energysavinghomes.org)	£18.75
Double Glazing	4 SAP points	£10-£15 per year	£4600 - £6000 (source: mybuilder.com)	£1150-£1500
Low Energy Lighting	2 SAP points	£10-£15 per year	£40 (source: Ebay)	£20

Source: Energykey.co.uk

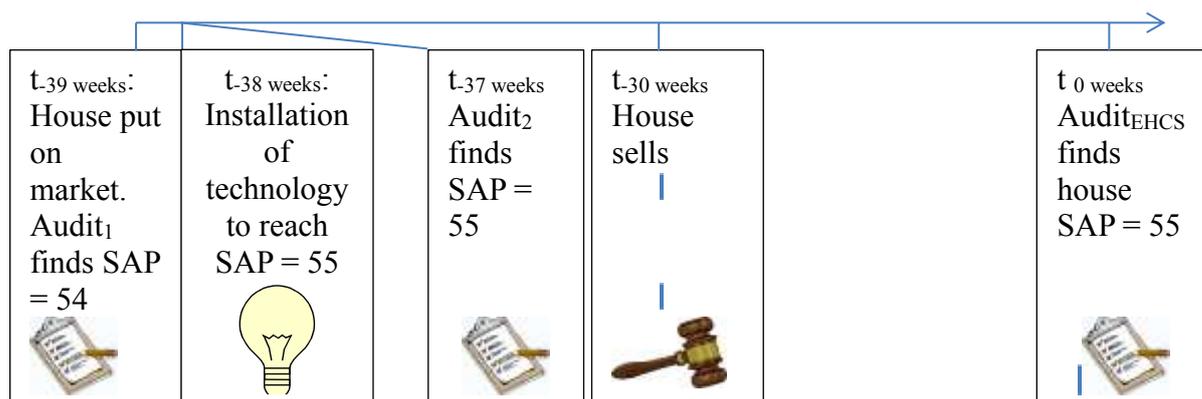
Put yourself in the position of a vendor. In line with the EPC requirement, you have had the energy efficiency of your property audited. You are now tidying the house to put it on the market. Would you also take the opportunity to replace the incandescent lightbulbs with more

²See, e.g., these advertisements: <http://web.archive.org/web/20160701103149/http://www.rightmove.co.uk/property-for-sale/property-41785911.html> and http://web.archive.org/web/20160701101959/http://www.zoopla.co.uk/for-sale/details/40897298?search_identifier=b9eec12f5ce7b4495ced4f97aedbdc08.

efficient LED lightbulbs? We predict that you become more likely to make such an investment in energy efficient technology if the audit has revealed that doing so would bump your home into the next color-letter grade on the EPC. The EPC audit returns a list of recommendations to improve energy efficiency. Replacement of incandescent lightbulbs with LED lightbulbs is estimated to increase a home’s energy efficiency by two SAP points (www.energykey.co.uk, see table 1). Purchasing twenty LED lightbulbs could shift a home from an E to a D color-letter grade at a cost of £40. Fuerst et al (2015) estimate from average market prices in England and Wales in 2014 that the price premium that D grade homes enjoy over E grade homes amounts to £1245. We predict that some vendors will target the next letter grade and make investments in order to reach that target.

The concrete steps through which we propose the EPC requirement influences our dependent variable are depicted in Figure 3. If consumers attend only to a home’s color-letter grade on the EPC, changing lightbulbs and commissioning a second EPC audit would be expected to yield a surplus of £1171.

Figure 3: Timeline



Notes: A timeline showing response to an initial audit of a property at 54 SAP points and how that response is predicted to lead to an overrepresentation of properties at 55 SAP points in the English House Condition Survey at t₀.

If our prediction is correct, we would expect to see a clustering in the English House

Condition Survey of properties just above a letter-grade threshold. For reasons of incentives,

we would expect this cluster to be largest in the middle of the distribution. Our logic is that properties at higher SAP scores are necessarily more energy efficient and so at the time of the initial audit ($t_{.39}$) are likely to have already exploited low hanging fruit i.e. they are more likely to be already equipped with low cost energy saving technologies or at the most common energy saving measures. Assuming this is true³, marginal investments in these energy efficient properties are likely to be less cost-effective than is the case lower in the energy efficiency distribution. At the bottom of the distribution, a different set of incentives operates. Over the period of our study, the English government was offering loans and subsidies to promote retrofitting (e.g. the *Green Deal*). Because retrofitting would be subsidized by the government, a vendor of a very energy inefficient property who was motivated to boost her SAP score could achieve a substantial effect at relatively low private cost by making a large-scale investment (e.g. replacing a boiler, replacing windows). Additionally, such investments would also make the home more attractive to would-be purchasers for reasons beyond its effect on the EPC label e.g. through increased reliability. Crucially, we cannot identify such investments because they would not be expected to cause clustering at the lowest point in an EPC color-letter grade.

Our prediction is lent support by recent empirical work in other contexts. Houde (2014b) finds a cluster of new household appliances released on to the market that are just energy efficient enough to garner Energy Star certification. In the specific context of energy labels for buildings, several studies have investigated response to LEED (Leadership in Energy and Environmental Design) certification. LEED certificates are awarded in Standard, Silver, Gold and Platinum varieties on the basis of points. Sallee (2012), Sallee and Slemrod (2012), Matisoff et al (2014) and De Lisle et al (2014) find an over-representation of buildings just

³ The English House Condition Survey does not detail the specific fittings that give rise to the SAP score and so we cannot test this assumption with the current data.

above silver, gold, and platinum threshold points, and a dearth of buildings just below these thresholds. Matisoff et al (2014) explain this response in terms of signaling: they hypothesize that actors invest in LEED certification as a means to signal their green credentials to stakeholders. Supporting this hypothesis, they find that non-profits build greener buildings than for-profit firms.

Our study makes a novel contribution relative to these previous papers for two important reasons. First, we look at the behavior not of firms, but of individual households. The main distinction of our research question, however, concerns its scope. Whereas these previous studies test for an effect of energy labeling on new buildings, our data tests for retrofitting i.e. altering existing buildings. When it comes to greening the building stock, a supply-side effect that permeates the entire housing stock is more impactful than one that affects only new buildings.

3. Methods

3.1. Data

We use data from the English House Condition Survey⁴, a national survey in repeated cross-section commissioned by the Department for Communities and Local Government (DCLG). It collects information about the condition and energy efficiency of housing in England on a stratified random sample of English homes⁵. In the course of this survey, a home energy audit is conducted that results in an SAP score. The home energy audit was conducted by Miller Mitchell Burley Lane, who “employ and manage a large field force of professional surveyors who work in close co-operation with the ONS interviewers to maximise response rates and deliver high quality data” (Department for Community and Local Government, 2014, p. 5).

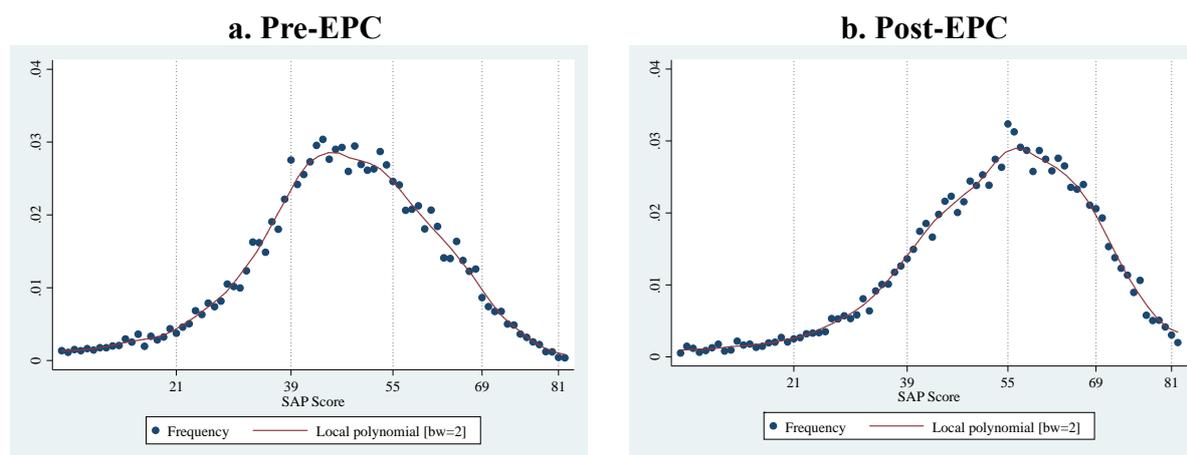
⁴ In 2008 the English House Condition Survey, which surveys properties, amalgamated with the Survey of English Housing, which surveys the residents of those same properties, to form the English Housing Survey. Our Post-EPC data is the 2009 wave of the English Housing Survey.

⁵In our analysis we account for stratification using sampling weights from the English House Condition Survey.

The EPC requirement was signed into law in November 2004. The 2003 wave of the English House Condition Survey (data collected April 2002 - March 2004; $n = 16648$, henceforth Pre-EPC) serves as our control group. The EPC requirement was implemented starting from August 2007, when properties of 4 or more bedrooms for sale or rent were required to display an EPC. The requirement applied to 3 bedroom homes from September 2007, and to all homes from December 2007 (*The Independent*, 2007). The EPC requirement came into effect during data collection for the 2007 wave (data collected April 2006 - March 2008; $n = 16217$, henceforth Introductory Phase). The EPC requirement was in effect throughout data collection for the 2009 wave (data collected April 2008 – March 2010, $n = 16150$, henceforth Post-EPC).

3.2. Analysis

Our core hypothesis is that a cluster occurs at a threshold in the centre of the distribution of Post-EPC data and that cluster is absent from the Pre-EPC data. Figure 4 depicts the frequency distributions of homes by SAP score for the Pre-EPC and Post-EPC samples. Additionally, Figure 4 shows a kernel-weighted local polynomial density estimation with a bandwidth of two, a non-parametric prediction of the distribution.

Figure 4. Frequency Distributions of Residential Properties in England By SAP Score

Notes: Frequency of houses by SAP score derived from the 2003 wave of the English Housing Survey (Pre-EPC, left) and from the 2009 wave of the English Housing Survey (Post-EPC, right). The first score in a new color-letter grade is represented by a vertical line. A kernel-weighted local polynomial smoothing with a bandwidth of two is shown. Frequencies are weighted using sampling dwelling weights. Summary statistics for these two distributions can be found in Table S1.

Figure 4b shows a cluster of homes at 55 SAP points, the lowest point in the D letter grade.

Our first test is whether this deviation is statistically significant. Our second set of analyses test whether this deviation can be explained by mechanisms other than the EPC requirement or whether the deviations occur in a manner consistent with the impact of the EPC requirement.

In order to test whether the deviation is statistically significant, we compared the observed density of homes at each SAP point to the density that would be predicted by a polynomial model of the smooth distribution. This approach is common in the literature on clustering because the polynomial model rules out as a source of spurious statistical significance the curvature of the distribution (e.g. Chetty, Friedman, Olsen, & Pistaferri. 2011; Lacetera, Pope, and Sydnor. 2012; Sallee and Slemrod, 2012). To derive the optimal polynomial, we run regressions of the form:

$$Freq_{SAP_i} = \alpha + g(i) + e_i \quad (1)$$

where $FreqSAP_i$ is the sample-weighted proportion of homes observed in the dataset with a given SAP score i (where i ranges from SAP score 3 to 82); the function g is a higher polynomial order that varies from 2nd to 22nd order; and e_i is an error term. Figure S1 graphs some of the resultant curves for the Post-EPC dataset. We select a polynomial order for which both the AIC and BIC statistics take the smallest value, which we denote as g^* . On the basis of these criteria, polynomial order $g = 19$ was selected for both the Post-EPC distribution and the Pre-EPC distribution.

Having selected g^* , the polynomial order that best fits the distribution, our next step is to test whether the observed data deviates significantly from it. We do this using the following model:

$$FreqSAP_i = \alpha + \beta(1=[i = \hat{i}]) + g^*(i) + e_i, \quad (2)$$

This model is very similar to that used to generate the smooth distribution. It attempts to predict the proportion of homes observed at each SAP point, i , using the optimal polynomial, $g^*(i)$. The independent variable of interest is the indicator variable $1=[i = \hat{i}]$, which takes the value of 1 for a single SAP score, \hat{i} , and a value of zero for the remaining 79 SAP scores. We run 80 regressions and in each a distinct SAP score, \hat{i} , is indicated. In other words, in our first regression \hat{i} indicates SAP = 3; in the second, \hat{i} indicates SAP = 4, and so on up to the 80th, where \hat{i} indicates SAP = 82. If the observed frequency of homes at a given SAP score is in line with the polynomial, then the indicator variable ($1=[i = \hat{i}]$) will add no explanatory power and its coefficient β will not differ significantly from zero.

4. Results

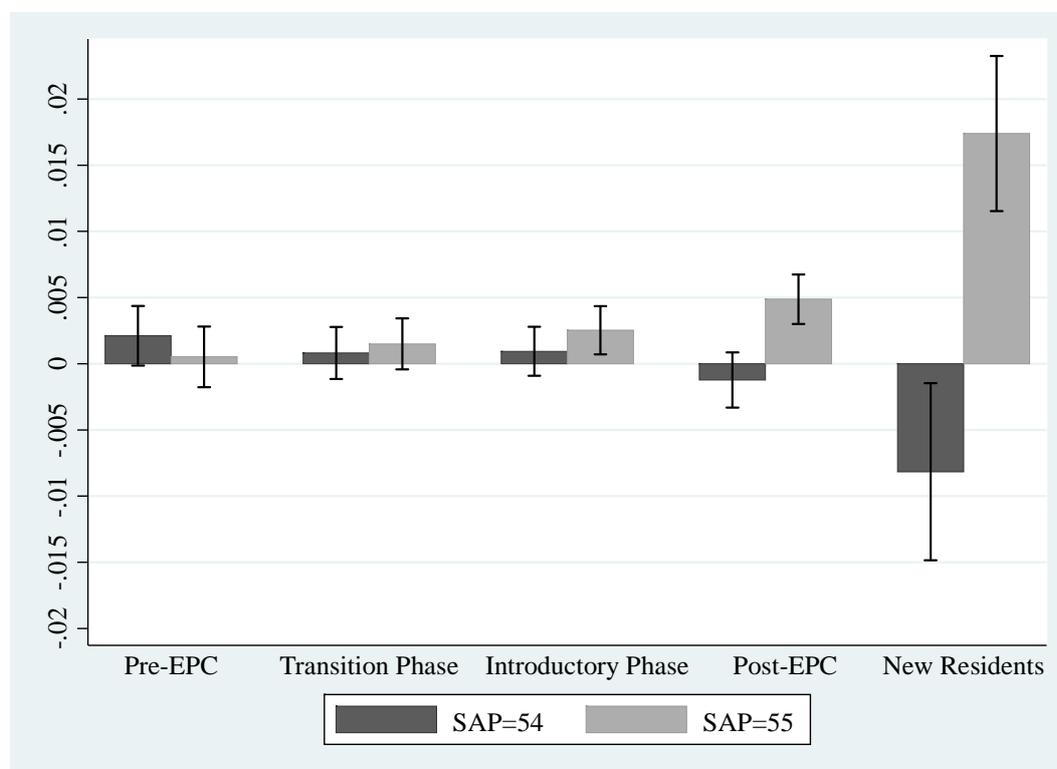
For the vast majority of SAP points, deviations are not significant (Figures S2 and S3 and Table S2 presents a list of significant deviations for each dataset used analysed). In the Post-

EPC data, there are more homes than would be expected by chance at SAP = 55 [N = 508; B = 0.005, S.E. = 0.001, $t(60) = 4.35$, $p < .001$], whereas there was no cluster at this same point in the distribution in the Pre-EPC data [N = 412; B = 0.0003, S.E. = 0.001, $t(60) = 0.036$, $p > .5$]. Our estimates show that the cluster comprises around 0.005% of all English homes – corresponding to an overrepresentation at SAP = 55 of 113,470 properties in 2009 (Department of Community and Local Government, 2012). Following Sallee and Ito (2017), we show that our results are similar when comparing against a smooth distribution in which the excess homes at SAP = 55 are spread uniformly across SAPs 50-54 (Figure S4). The purpose of this test is to rule out misspecification of the smooth distribution as a source of statistical significance. Specifically it rules out the concern that the loss in density from other SAP points leads to exaggeration of the cluster at SAP = 55. The observed cluster at SAP = 55 is also robust to variation in choice of polynomial order (Figure S5).

Two independent dose-response relationships implicate the EPC requirement as the causal mechanism for the cluster. The first of these considers the introductory phase of the data, which comes from an independent sample of homes relative to the Post-EPC data analyzed above. As would be expected from the staggered introduction of the EPC requirement, the Introductory Phase cluster is roughly half the size of the Post-EPC cluster [B = 0.003, S.E. = 0.001, $t(60) = 2.31$, $p > .024$, see Figure 5 and Figure S8].⁶ As well as supporting our dose-response prediction, this result also replicates our core finding in an independently-drawn sample of English homes.

⁶ Similarly, the transitional phase has a Post-EPC cluster at 55 that becomes statistically different than zero, see Figure 5 and Figure S9.

Figure 5: A timeline of deviations from the smooth distribution in the number of homes observed at 54 and 55 SAP points



Notes: 90 percent confidence intervals. The sample size is 80 for all analyses. *Pre-EPC* is the 2003 wave of the English House Condition Survey; *Transition Phase* is the pooled 2005 and 2006 waves ($n = 32939$); *Introductory Phase* is the 2007 wave; *Post-EPC* is 2009 wave; *New Residents* is the 2009 wave restricted to homes that have been occupied by their current residents for less than twelve months. A score of zero on the vertical axis means that there is the same percentage of homes at a given SAP score as predicted by the best-fit polynomial modeling of the smooth distribution. A positive and significant result indicates that there are in fact more homes at that SAP score than predicted by the smooth distribution; a negative and significant result indicates fewer homes than expected at that SAP score.

The second dose-response analysis looks to the subsample of homes that were actually impacted by the EPC requirement – homes that were placed on the rental or sale market.

When we restrict the 2009 wave to houses with new residents in the past 12 months (a proxy measure for homes being recently sold or rented, $n = 1526$), the size of the cluster at SAP = 55 trebles in size relative to the cluster observed for the full 2009 sample [$B = 0.017$, S.E. = 0.0035, $t(60) = 4.95$, $p < .001$, Figure 5, New Residents Sample and Figure S10).

Additionally, within this sample there are fewer homes than would be expected by chance at SAP = 54 ($B = -0.0082$, S.E. = 0.004, $t(60) = 2.04$, $p < .05$). Taken together, these two results

provide empirical support for the mechanism set out in Figure 3: home vendors respond to the news that their homes are at SAP = 54 by investing in energy efficiency.

We know that there was no national level policy in England that required or incentivized homes to attain a specific EPC letter grade. It is local councils, however, that are responsible for managing state-owned housing and for granting planning permission for new private construction. It is possible that some local councils enforced policies of which we are unaware. We rule out this alternative mechanism by dropping from the sample those properties for which local councils could conceivably influence the level of energy efficiency: state-owned property and properties built since the EPC requirement. Restricting our sample to privately-owned homes ($n = 11,195$), we continue to find significant clustering at SAP = 55 ($B = 0.005$, S.E. = 0.0013, $t(60) = 3.86$, $p > .001$, Figure S6). Our result is also robust to dropping from the sample homes that were built since the EPC requirement. The youngest age band reported in the English House Condition Survey are those built since 1990. Restricting the sample to homes built prior to 1990 ($n = 14,228$) there remains a significant cluster of homes at SAP = 55 ($B = -0.005$, S.E. = 0.0011, $t(60) = 4.84$, $p < .001$, Figure S7).

5. Discussion:

5.1. Are these clusters formed by real investments in energy efficiency?

The question that motivates this study is whether energy labeling can induce retrofitting of residential homes. Let's review the facts. Prior to the EPC requirement there were just as many homes at the lowest point in the D letter grade as predicted by a smooth distribution. Since the EPC requirement, there are 5 homes per thousand clustering at this SAP score that we would not expect given a smooth distribution. An analysis of homes with new residents demonstrates that this cluster was disproportionately achieved at or about the time when

homes were traded – precisely when EPCs are required. Out of every 1000 homes with new residents, 17 more than expected have SAP scores of 55, and 8 fewer than expected have SAP scores of 54.

We are confident that this pattern of results reflects real investment in energy efficiency for four reasons. First, the results are inconsistent with deliberate gaming and misreporting. Our theory predicted clustering on the basis of marginal costs. The marginal costs of misreporting are independent of a property's initial SAP score and so if misreporting occurs, we would expect to observe it either where the returns are greatest or else at every threshold, as in Hyland, Alberini and Lyons (2016). Our data shows a cluster at only one threshold and it is where returns to transitioning are 0.7 percent, less than a quarter of what is estimated to be returned by transitioning from C to B (Fuerst et al, 2015). Marginal benefits alone cannot account for the observed cluster. By contrast, marginal costs can explain the cluster. The data only shows a cluster where the marginal cost of increasing an SAP score is relatively low.

Second, the SAP data we analyze were collected specifically for an administrative survey, which reduces incentives for gaming. The assessor has a strong disincentive to deliberately misreport, since the firm (Miller Mitchell Burley Lane) is named in the summary report and would lose credibility and repeat business from the Department of Local and Community Government if it were shown to have falsified administrative data. More importantly, we can think of only one plausible story through which the homeowner has an incentive to influence an SAP score that will appear in an anonymized dataset. For completeness, we now review that story and show how it cannot account for our results. Assuming the English Housing Survey gifts to its respondents the EPC that derives from the audit, then homeowners would have an incentive to influence the resultant SAP score if they intend renting out or selling

their property in the near future⁷. Of course, we cannot observe in the data which homeowners were thinking of putting their properties on the market in the coming twelve months. Recall however that the observed cluster is driven by homes with residents who have moved in within the past twelve months. In order for reporting bias to explain our cluster it would have to be the case that the intention to sell or rent a home is substantially greater among new residents than amongst the rest of the sample. Is it plausible that motivation to trade the home they live in is three times as large amongst people who have just moved in than amongst other homeowners? We do not believe so.

Deliberate misreporting is one form of gaming. Gaming of a different sort is proposed to explain clustering at the minimum standard for LEED certification (Sallee, 2012). In this view, the rational response to noisy energy labels is to tweak blueprints so as to deliver a new-build that barely satisfies the criterion required for certification. Crucially, this can be achieved by removing energy-saving technologies from an initial blueprint. In this view, clusters can reflect losses in energy efficiency relative to the no-label counterfactual. There are two reasons to be confident that the cluster observed in the current study reflects positive changes to buildings in terms of SAP. First, properties that are missing at 54 SAP points appear at 55 SAP points. Second, homes built prior to 1990 were just as likely to cluster at 55 SAP points than were newer homes. The changes to energy efficiency that induced clustering in these older homes cannot be achieved merely by tweaking blueprints but must be achieved by altering the physical fabric of existing buildings.

5.2. Back of the envelope calculation of cost-savings

There are additional questions that a policy maker or curious reader will want answers to: How large are the savings induced by these effects? Is it welfare enhancing to stimulate

⁷ An EPC is valid just for one year.

retrofitting through the use of energy labeling? We do not have sufficient data to generate precise answers to these questions, but we do recognize that these questions warrant discussion. Our sense is that this discussion will profit from reference to relevant data, even if that data is imperfect, and so in what follows we sketch out some back-of-the-envelope implications of our results.

The first question concerns the scale of the effect of the EPC requirement. There were 22,694,000 dwellings in England in 2009 (Department of Community and Local Government, 2012). Extrapolating from our stratified random sample of English homes to the population level, there were 113,470 additional homes at 55 SAP points relative to what would have been the case in the absence of the EPC requirement. We do not know the initial level of energy efficiency of these homes, however. To estimate the change in energy efficiency induced by the EPC requirement, we look to the subsample of homes that were directly impacted by the EPC requirement: those with new residents. Amongst these, there were 17 additional homes per thousand at 55 SAP points relative to the no EPC counterfactual. Almost half of this cluster at 55 SAP points appears to comprise homes that had been at the highest point in the E colour-letter grade, as evidenced by the deficit of homes at 54 SAP points. We cannot observe whence the remainder of the cluster is drawn, and so we conservatively assume that it reflects an increase in energy efficiency of just two SAP points. We calculate, again conservatively, that this is equivalent to a reduction in energy use of 295 kWh per year per household.⁸ At benchmark electricity prices of £0.10 per kWh (derived from price comparison website UK Power.co.uk), this would save £29.50 per year per household. It would also save an estimated 0.02 tons of carbon emission per household

⁸ For the reason highlighted at the close of section 5.1, there can be no direct conversion from 1 SAP point to kWh. We derive our estimate in the following way: A gain of 2 SAP points is achieved by replacing 60 watt incandescent lightbulbs with 6 watt LED lightbulbs (table 1 above). That would save 393kWh per year, assuming 20 lightbulbs across a home, each illuminated for on average one hour a day. A gain of 1.5 SAP points is therefore equivalent to 295kWh per year.

(The Communities and Local Government Committee, 2008), which scales up to an annual saving of 2,269 tons of carbon. In 2008 the UK had successfully met its Kyoto protocol target of reducing carbon emissions by 12.5% of 1990 levels (The Guardian, 2010). We estimate that the labeling effect of the EPC identified here can account for 0.03 percent of that success.

We do not have data on which technologies were installed to boost SAP scores, and hence we cannot determine whether the estimated savings were worth their costs. One interesting reference point for interpreting these savings is to compare them with those that resulted from the Opower social norms manipulation. That intervention resulted in savings estimated at \$36 (Allcott, 2011, £27.90 at current market prices, XE.com), which is of similar magnitude to the effect found here. The Opower result relies on human behavior, which is less stable across individuals and over time than the effects of retrofitting. Allcott (2011) finds that people show signs of fatigue in response to the Opower intervention and so the initial curtailment effects tend to weaken. The energy savings delivered by fixed technology are likely to remain constant over time, although future research might investigate whether they are offset by changes in behavior e.g. the Jevons paradox (Sorrell, 2009).

5.3. Limitations

Based on the foregoing, we are confident that energy labels did prompt retrofitting. Did retrofitting lead to energy savings? Probably, but we cannot be certain. Recall that the SAP score is an index of how much it costs to maintain a home at a certain temperature. Kelly, Crawford-Brown and Pollitt (2012) observe that a home's SAP score would improve by replacing a bio-diesel stove with a coal-burning stove, a counterintuitive result which is explained by the fact that coal delivers more heat-per-£ than does bio-diesel. While we can think of no reason why the marginal investment observed in our data would be especially directed towards such spurious investments, it adds a caveat to our conclusion: energy labels

will only green the housing stock if those energy labels are derived from appropriate measures of energy efficiency.

It is important to note that this analysis should not be interpreted as an attempt to measure the overall effect of the EPC requirement on the energy efficiency of the housing stock. We tested for one very specific and localized consequence of the EPC requirement: the *additional* change in energy efficiency that is induced by targeting a letter-grade on the EPC, over and above other effects of the EPC requirement. In the language of policy evaluation, the result we found is a LATET – local average treatment effect on the treated. A cost benefit analysis would require knowledge of all of the effects of the EPC requirement on retrofitting. It is our sense that these cannot be observed with current data.

We can think of two other likely effects that seem plausible given the current result. First, the EPC audit gives the owner concrete recommendations on how to most efficiently improve their home's energy efficiency. Second, the EPC indicates not only a building's actual SAP score, but also a potential score, which is more useful as a target for investment than the arbitrary colour-coded letter grades. We could imagine that people who have just moved into a new house might retrofit in response to these recommendations and targets, however we must leave it to future research to investigate these additional effects of the EPC requirement.

With regard to the welfare implications of the EPC requirement, they look promising. Allcott (2016, p.4) points out that as a means to green the economy, "information and marketing is cheap compared to the fiscal cost of subsidies or the production costs of energy efficient goods". Certainly, other notable recent attempts to induce retrofitting have been costly – the Green Deal set aside £540m of public money; the Weatherization Assistance Program delivered social returns on investment of -9.5% (Fowlie et al, 2015b). The cost of an EPC audit is advertised as being £34 and is borne privately by the vendor or lessor of the property.

Further research is required before we conclude that energy efficiency labels will always and everywhere induce retrofitting. First, the EPC requirement comprises several policies and we cannot determine which of these are necessary or sufficient to induce investment.

Specifically, we cannot formally distinguish between the effect on investment that is driven by the requirement that EPCs be published, and the effect that is driven by the information and recommendations contained in the audit. There is a straightforward story to explain why the requirement that EPCs be published would induce targeting of a colour-grade: vendors and lessors want potential purchasers to see their property, and hence its EPC, looking as attractive as possible. If the EPC were not published, it is difficult to see why homeowners would care how their home measures up on its arbitrary scale.

Relatedly, the design of that arbitrary scale appears to play an important role in stimulating investment. It is interesting that we identify clustering only at the E to D letter grade threshold. We hypothesized that, for reasons of marginal cost, the effect would occur in the low to middle range of the energy efficiency distribution. A further factor that may have led the effect to be largest at the D grade threshold is the connotations of the colour-letter grades. In the UK school system, to move from an F to an E is to move from one fail grade to another, whereas to move from an E to a D is to move from a fail to a pass. Paraphrasing Newall and Siikamaki's (2014) conclusions from their choice experiment: Anyone who has been to school would rather get a pass than a fail. Other evidence that connotations affect choice comes from Ubel et al (2015), who found that people were attracted to health insurance plans labelled gold regardless of the substantive content of those health plans. It may be that the EPC requirement would have had less effect at stimulating investment if its layout were different e.g. if it only published SAP scores. Even though the colour-coded letter-grades on the EPC are redundant in terms of the information they add, they clearly play a role in stimulating retrofitting.

The audit that generates the EPC could stimulate investment in its own right. It serves a valuable function in providing the owner of a building with information on their home's current energy use and its potential for improvement. Our results may be partly explained by this provision of private information, and that portion of the effect could be achieved without the requirement that energy labels be made publicly available. Previous research suggests that the compulsion element of the policy is important; although homeowners should commission audits spontaneously if they expect private net-benefits from doing so, uptake on audits is low (Palmer and Walls, 2015).

Another limitation of this study is one that it shares with any natural experiment: it took place in a particular place at a particular time. We will have to leave it to future research to determine the environmental, institutional and cultural features that make investment receptive to energy labels (though see Kok et al, 2011). That said, our results did reveal one indication that energy labels have pervasive effects. It is not merely investment in new-build homes that responded to the EPC requirement; homes built prior to 1990 also demonstrated response. This result suggests that energy labels could green the housing stock even in environments where the housing stock is mature and new building is rare.

5.4. Conclusion

In spite of the questions that this research leaves open, it makes one key contribution. It demonstrates that an energy efficiency label stimulated retrofitting.

Notwithstanding its success at inducing investment, we can see scope to improve the design of the EPC as a means to induce retrofitting. It is possible that a home could be boosted a lettergrade by merely replacing lightbulbs (although our data we does not allow us determine what investments actually led homes to end up at 55 SAP points). Moreover, the targeting mechanism we identified as inducing retrofitting would only be expected to impact those

homes that are close to the next color-letter grade, which is a small proportion of the overall housing stock. Relatedly, once a higher color-letter grade is achieved, the current EPC does nothing to motivate further investment in energy efficient technologies. How can we design labels that motivate continued retrofitting? One approach would be to make the color-letter grades indicators of *relative* energy efficiency. Instead of fixing the seven color-letter grades at arbitrary absolute levels of energy efficiency (e.g. yellow D means 55-68 SAP points), the color-letter grades could indicate septiles of energy efficiency. Thus, as the distribution of home energy efficiency shifts upwards, so too would color-letter grades. If thresholds shift over time, then a higher proportion of homes relative to the current case will at some point in time be found either close to moving up or close to dropping down a color-letter grade. Additionally, the relative approach would be expected to provide stronger incentives than does the current EPC. Because people are loss averse and concerned with rank position (Vendrik and Woltjer, 2007; Brown, Gardner, Oswald and Qian, 2008), we would expect the prospect of dropping into a lower color-letter grade to be especially motivating.

In short, the current research provides evidence that energy labeling of residential homes merits a place in the toolkit for reducing household energy consumption. Retrofitting has proven a stubborn behavior to stimulate, but our results suggest that interventions aimed at people trading homes could be especially fruitful. We conclude that a well-designed energy labeling policy has potential to induce retrofitting and green the housing stock.

References:

Allcott, Hunt. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95(9): 1082-1095.

Allcott, Hunt. 2016. "Paternalism and Energy Efficiency: An Overview." *Annual Review of Economics* 8(1): 145-176.

Allcott, Hunt, and Sendhil Mullainathan. 2010. "Behavioral science and energy policy." *Science* 327(5970): 1204-1205.

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review* 104(10): 3003-3037.

MyBuilder.com (2011) How much does it cost to doubleglaze a 3 bedroom house with 12 standard upvc windows? <https://www.mybuilder.com/questions/v/459/how-much-does-it-cost-to-doubleglaze-a-3-bedroom-house-with-12-standard-upvc-windows>

B.I.T. (2011). Behaviour change and energy use. *London: Cabinet Office.*

BRE (2012) Energy Performance Certificates for Existing Dwellings. *Building Research Establishment Ltd.*

Brown, Gordon, Jonathan Gardner, and Andrew Oswald. 2008. "Does Wage Rank Affect Employees' Well-being?" *Industrial Relations: A Journal of Economy and Society*, 47(3), pp.355-389

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records." *The Quarterly Journal of Economics*, 126(2): 749-804

Consumer Focus. 2011. *As easy as EPC? Consumer views on the Content and Format of the Energy Performance Certificates*

Costa, Dora L. and Matthew E. Kahn. 2013. "Energy Conservation 'nudges' and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment" *Journal of the European Economic Association*, 11(3): 680-702

DeLisle, James, Grissom, Terry and Högberg, Lovisa. 2013. "Sustainable real estate: An empirical study of the behavioural response of developers and investors to the LEED rating system". *Journal of Property Investment & Finance*, 31(1), pp.10-40

Department of Community and Local Government. 2010. English House Condition Survey 2007: Technical Report. Retrieved at: <http://webarchive.nationalarchives.gov.uk/20120919132719/www.communities.gov.uk/publications/housing/ehcstechnicalreport2007>

Department of Community and Local Government. 2012. Live Tables on Dwelling Stock. <https://www.gov.uk/government/statistical-data-sets/live-tables-on-dwelling-stock-including-vacants>

Department of Community and Local Government. 2014. *Improving the energy efficiency of our buildings* Accessed on 12-5-16

Department of Energy and Climate Change. 2013. *Standard Assessment Procedure* http://web.archive.org/save/_embed/https://www.gov.uk/guidance/standard-assessment-procedure Accessed on 7-4-16

Dietz, Thomas, Gerald T. Gardner, Jonathan Gilligan, Paul C. Stern, and Michael P. Vandenbergh. 2009 "Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions." *Proceedings of the National Academy of Sciences* 106(44): 18452-18456.

Dolan, Paul, and Robert Metcalfe. 2013 "Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation." CEP Discussion Paper No 1222

EnergyKey.co.uk
<http://web.archive.org/web/20160412111200/http://www.energykey.co.uk/epc.html>. Accessed on 12-4-16

EnergySavingHomes.org (undated) Hot Water Cylinder Insulation Can Save Energy and Money http://www.energysavinghomes.org.uk/improvements/hot-water-cylinder-insulation_4

Englmaier, Florian, Arno Schmöller, and Till Stowasser. (forthcoming) "Price discontinuities in an online market for used cars." *Management Science*.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2015a. "Are the non-monetary costs of energy efficiency investments large? Understanding low take-up of a free energy efficiency program." *American Economic Review* 105 (5): 201-204.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2015b. "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program" NBER Working paper 21331

Fuerst, Franz, Patrick McAllister, Anupam Nanda, and Peter Wyatt. 2015. "Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England." *Energy Economics* 48:145-156.

Gabaix, Xavier, and Laibson, David. 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, 121(2), 505-540.

Gardner, Gerald T., and Paul C. Stern. 2008. "The short list: The most effective actions US households can take to curb climate change." *Environment: science and policy for sustainable development* 50(5): 12-25.

Graff Zivin, Joshua and Kevin Novan. 2015. Upgrading Efficiency and Behavior: Electricity Savings from Residential Weatherization Programs. Mimeo.

Houde, Sébastien. 2014a. "How Consumers Respond to Environmental Certification and the Value of Energy Information." NBER Working Paper w20019.

Houde, Sébastien. 2014b. "Bunching With the Stars: How Firms Respond to Environmental Certification", Working Paper.

Hyland, Marie. , Anna Alberini, and Sean Lyons. 2016. “The Effect of Energy Efficiency Labeling: Bunching and Prices in the Irish Residential Property Market” (No. tep0516). Trinity College Dublin, Department of Economics.

Independent. 2007. Everything you need to know about the Home Information Pack, 10 April 2007.

Ito, Kirocho. and James Sallee. 2017. “The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards.” *Review of Economics and Statistics* forthcoming.

Kelly, Scott, Douglas Crawford-Brown, and Michael Pollitt, 2012. Building performance evaluation and certification in the UK: Is SAP fit for purpose?. *Renewable and Sustainable Energy Reviews*, 16(9), 6861-6878.

Kok, Nils, Marquise McGraw, and John M. Quigley. 2011. “The diffusion of energy efficiency in building.” *The American Economic Review* 101(3): 77-82.

Lacetera, Nicola, Devin Pope, and Justin Sydnor. 2012. “Heuristic Thinking and Limited Attention in the Car Market”, *American Economic Review*, 102(5): 2206–36.

Matisoff, Daniel, Noonan, Douglas, and Mazzolini, Anne., 2014. “Performance or marketing benefits? The case of LEED certification”. *Environmental Science & Technology*, 48(3), pp.2001-2007.

Mumsnet. 2008. How much does it cost to replace a boiler
https://www.mumsnet.com/Talk/legal_money_matters/452064-how-much-does-it-cost-to-replace-a-boiler

Newell, Richard and Juha Siikamaki. 2014. “Nudging Energy Efficiency Behavior: The Role of Information Labels”, *Journal of the Association of Environmental and Resource Economists*, 1(4): 555-598.

Palmer, Kate 2015. Green Deal Funding Axed *The Telegraph*. 23/7/2016.

Palmer, Karen, and Margaret Walls. 2015. “Limited Attention and the Residential Energy Efficiency Gap”, *American Economic Review*, 105(5): 192-95.

Sallee, James. 2012. “Rational Inattention and Energy Efficiency”, *RFF Symposium Paper*

Sallee, James, and Joel Slemrod. 2012. “Car notches: Strategic automaker responses to fuel economy policy.” *Journal of Public Economics*, 96(11): 981-999.

Sorrell, Steven. 2009. Jevons’ Paradox revisited: The evidence for backfire from improved energy efficiency. *Energy Policy*, 37(4), 1456-1469.

Stavins, Robert, Todd Schatzki, and Jonathan Brock. 2013. *An Economic Perspective on Building Labelling Policies*, Analysis Group.

The Guardian. 2010. *UK carbon emissions fell by 2% in 2008, figures show* accessed at: <https://www.theguardian.com/environment/2010/feb/02/carbon-emissions-kyoto>

UKpower.co.uk, accessed on June 20 2016:

https://web.archive.org/web/20160620092650/https://www.ukpower.co.uk/home_energy/tariffs-per-unit-kwh

Vendrick, Marteen. and Woltjer, Geert., 2007. Happiness and loss aversion: Is utility concave or convex in relative income? *Journal of Public Economics*, 91(7), pp.1423-1448

Which? (undated) Cavity Wall Insulation Cost and Savings

<http://www.which.co.uk/reviews/insulation/article/cavity-wall-insulation-costs-and-savings>

Which? (undated) How to buy loft insulation

<http://www.which.co.uk/reviews/insulation/article/how-to-buy-loft-insulation/loft-insulation-costs-and-savings>

XE.com accessed 06/05/17.

Supplementary Material

Table S1: Data used for analysis

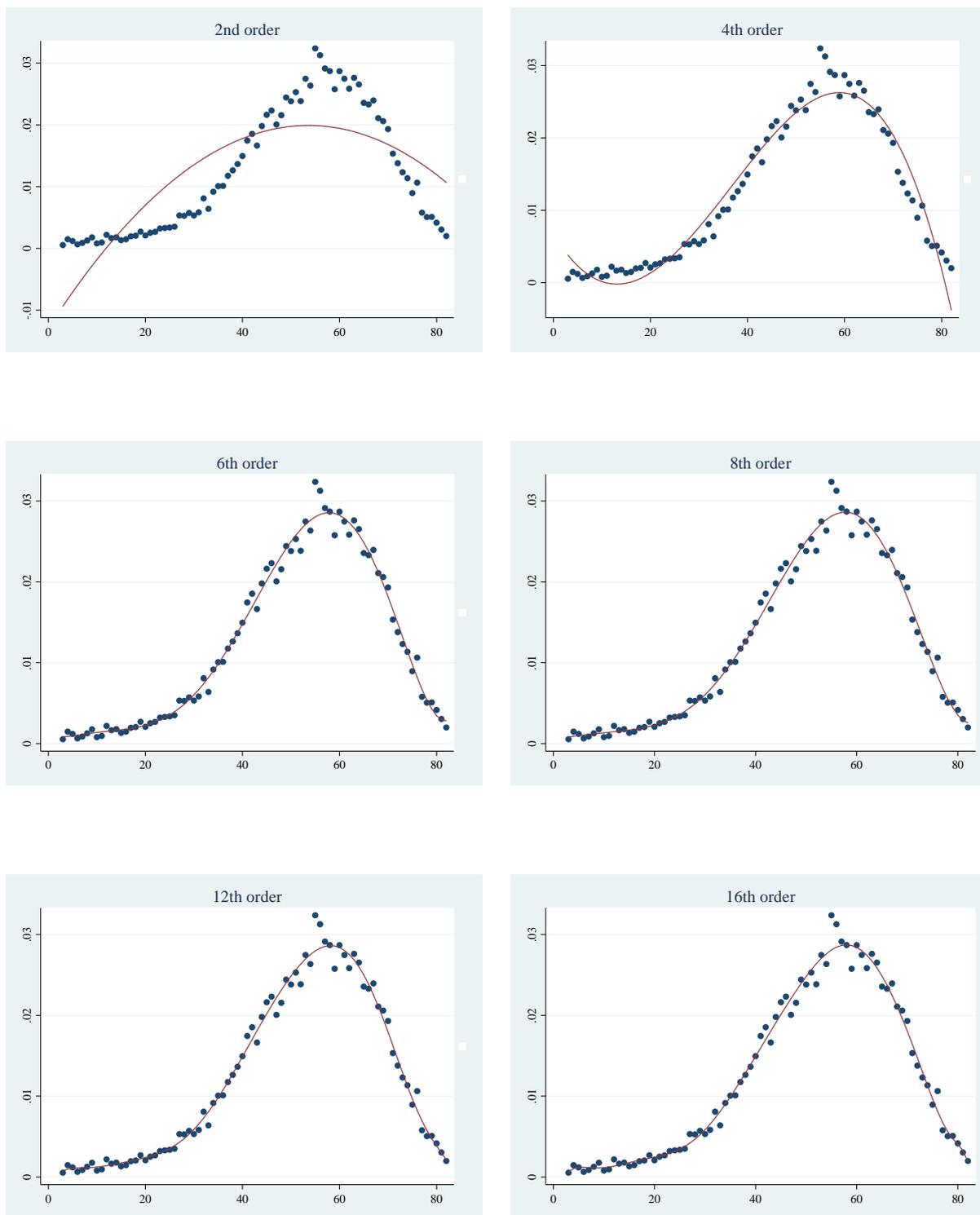
Sample	Survey wave	Survey period	Sample size	Mean SAP	Standard Deviation	Skewness	Kurtosis
Pre-EPC	EHCS 2003	April 02- March 04	16648	46.64	14.89	-0.55	3.52
Post-EPC	EHS 2009	April 08- March 10	16150	53.06	14.98	-0.81	3.90

Notes: Statistics are weighted using frequency weights reported in English House Condition Survey 2003 and English Housing Survey 2009. The Housing Act of 2004, which was granted royal assent in November 2004, stipulated that Energy Performance Certificates would be required for properties on the market.

Table S2: SAP scores that differ significantly ($p < .05$) from the best-fit distribution

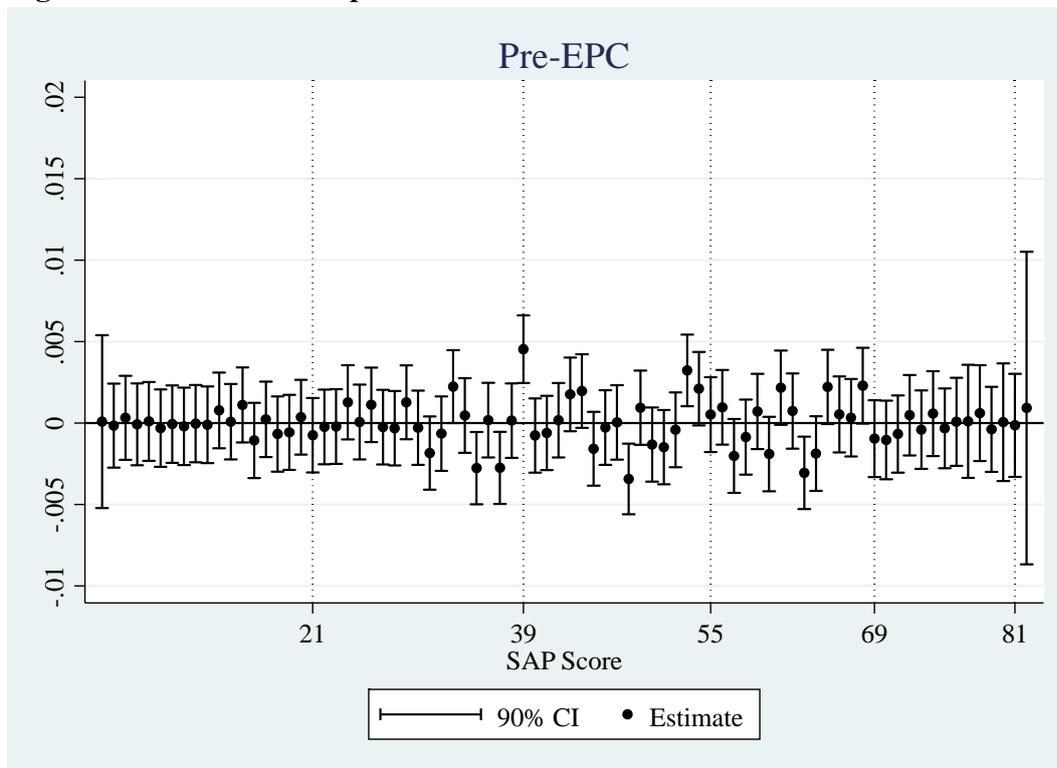
	Post-EPC	Pre-EPC	Introductory Phase
More houses than predicted	55, 56, 77	39, 48, 53, 65, 68	39, 50, 55, 60, 63, 72
Fewer houses than predicted	47, 52	35, 37, 63	42, 57

Figure S1: Polynomial fit graphs modelling SAP scores for the Post-EPC data



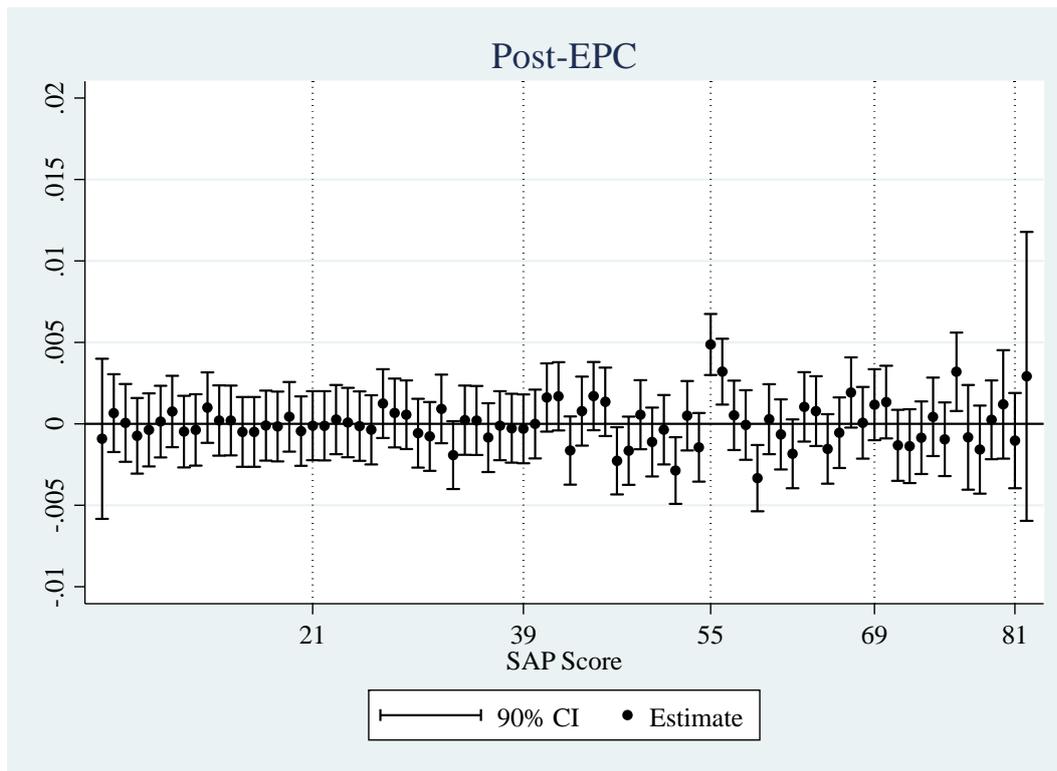
Notes: These figures show how well different polynomial orders match the distribution of our data in the Post-EPC data by displaying the actual data points from the 2009 English Housing Survey (represented by dots) and the fitted line from a given polynomial order (represented by a smooth line). The fit is similar for polynomial orders from 6 and above.

Figure S2: Pre-EPC Sample: Deviations from the smooth distribution



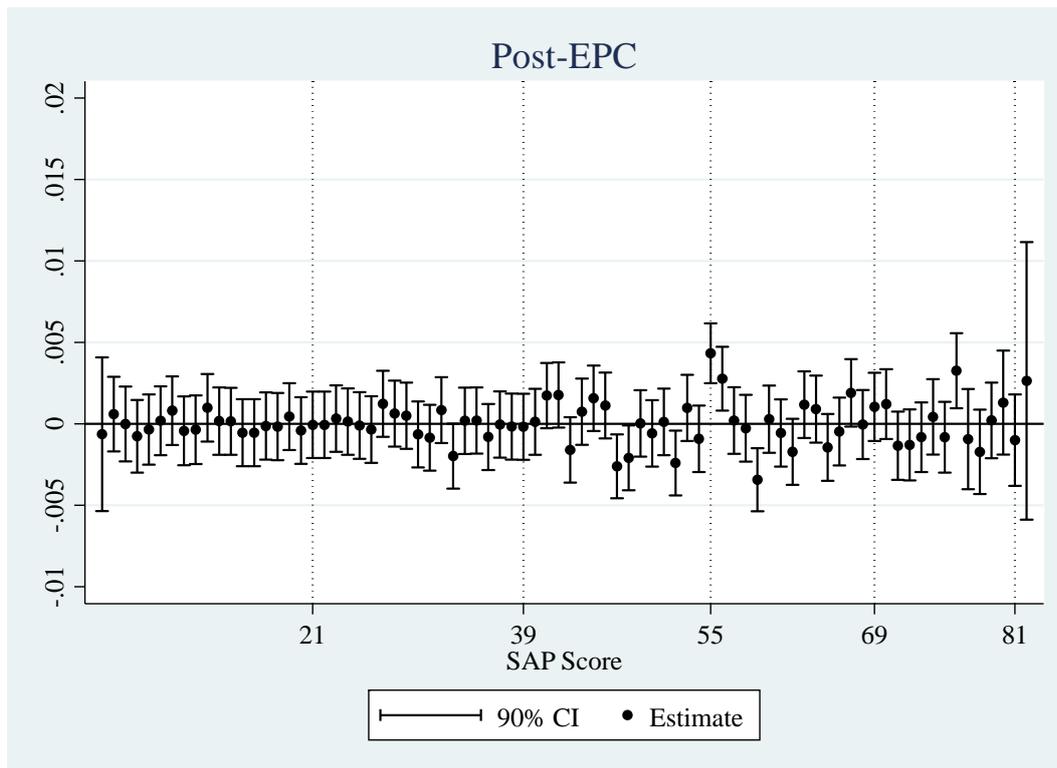
Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (19th order in this case) for the 2003 wave of the English House Condition Survey. The dependent variable is the frequency of SAP score. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Figure S3: Post-EPC Sample: Deviations from the smooth distribution



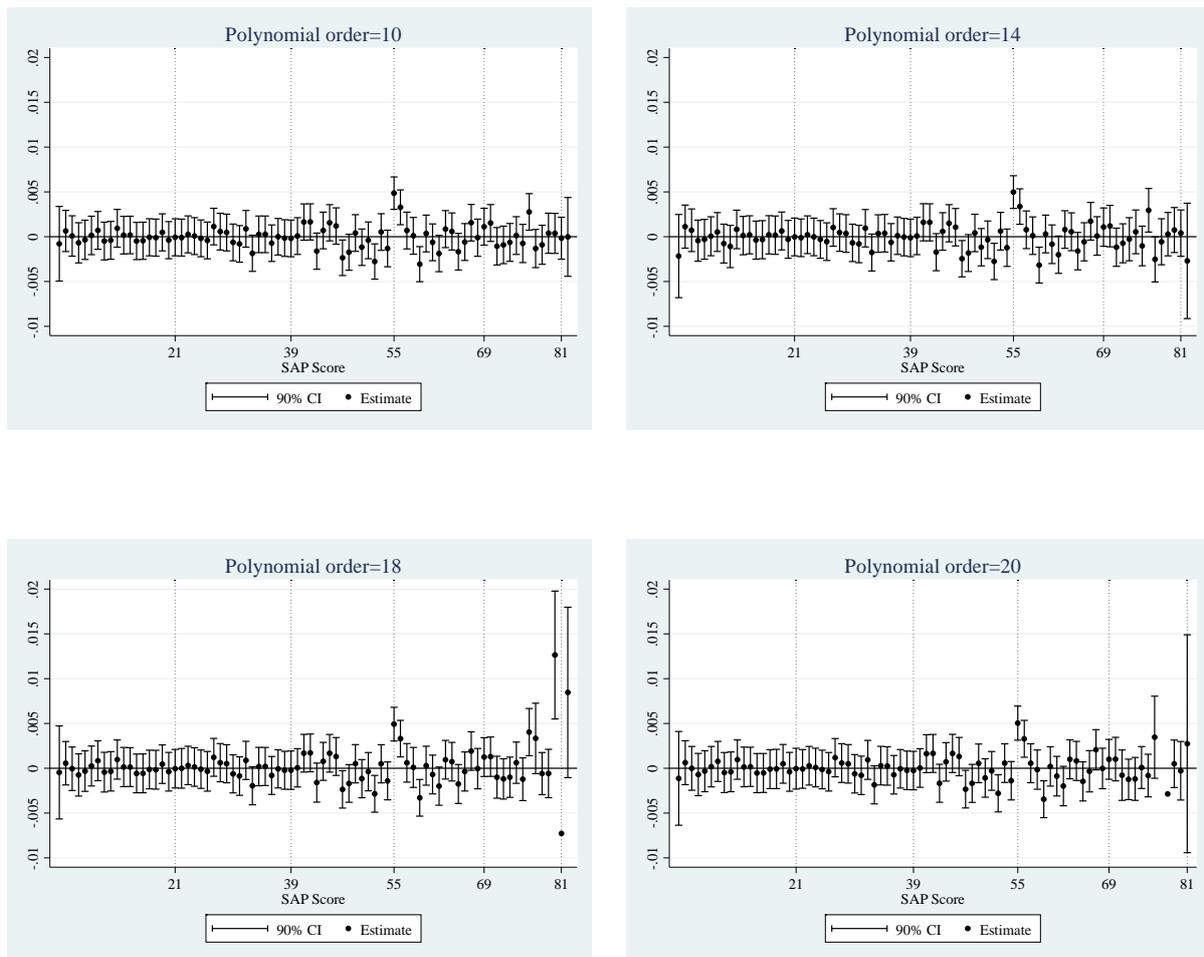
Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (19th order in this case) for the 2009 wave of the English House Condition Survey. The dependent variable is the frequency of SAP score. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Figure S4: Post-EPC Sample: With Excess Density Redistributed



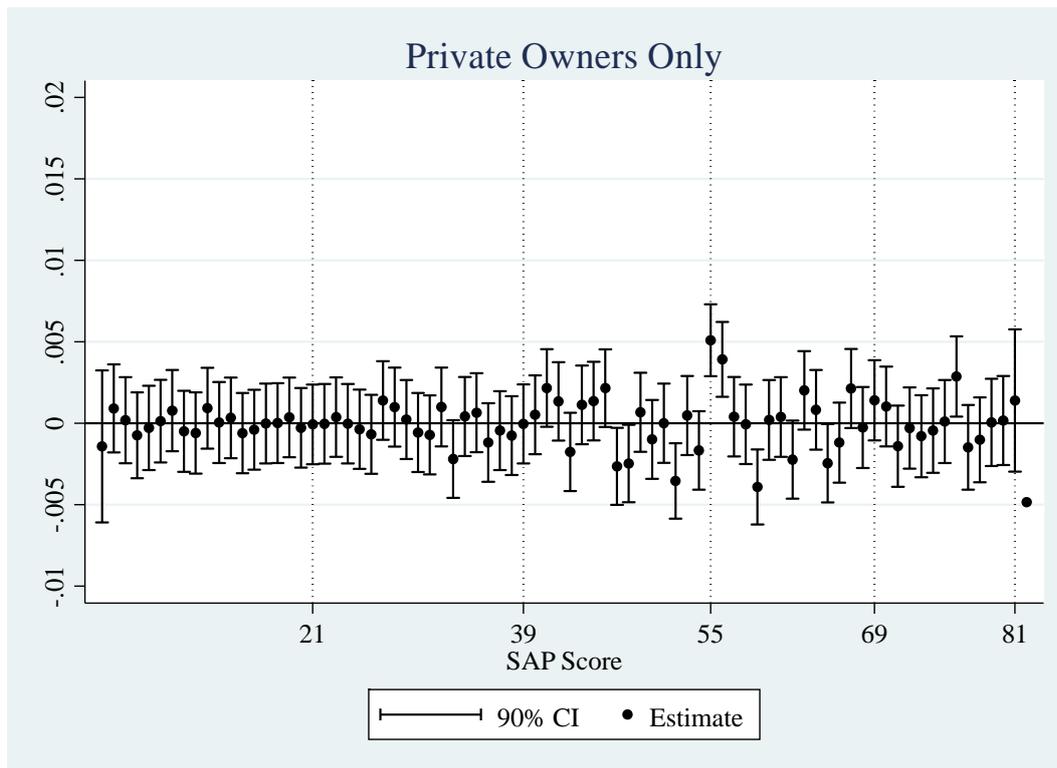
Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (19th order in this case) for the 2009 wave of the English Housing Survey. To control for the potential bias that would occur due to the actual distribution having excess density over the predicted distribution, the excess density at the 55 SAP score has been put back into the distribution uniformly over the SAP scores 50-54. This assumes that the excess homes in the 55 SAP bunch came from the 50-54 SAP scores. The dependent variable is the frequency of SAP score. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Figure S5: Post-EPC: Estimates of Deviation using Different Polynomial Orders



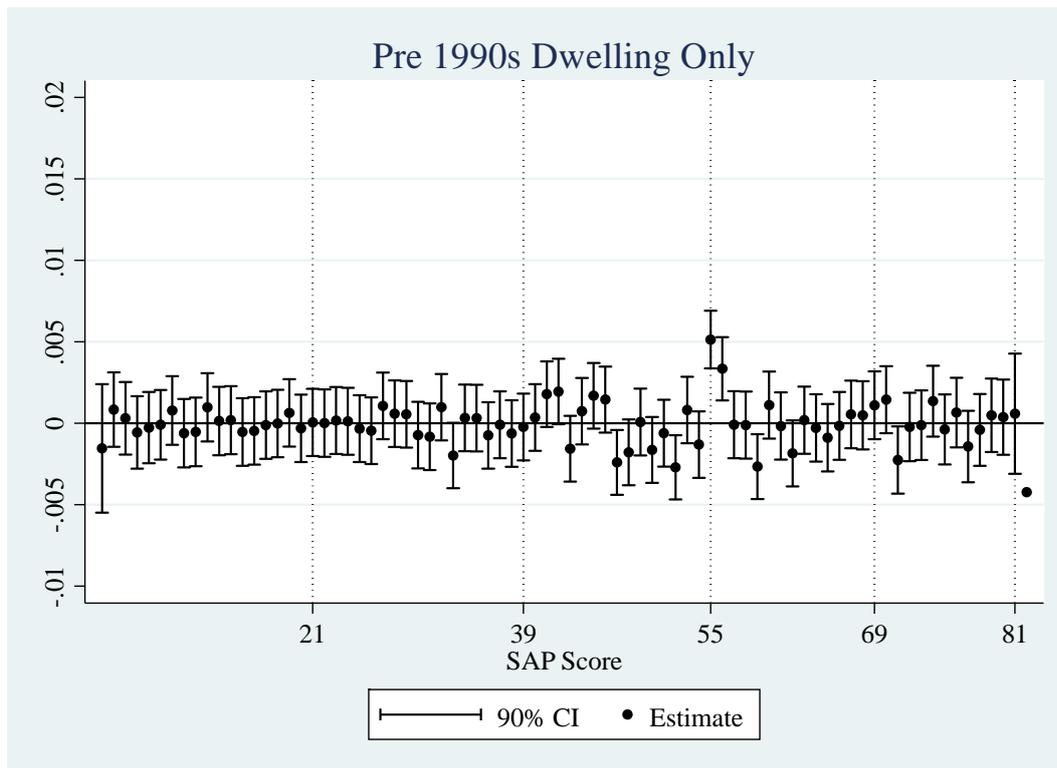
Note: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from several polynomial orders for the 2009 wave of the English Housing Survey. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa. Confidence intervals are not reported when larger than the vertical axis scale.

Figure S6: Post-EPC: Privately-Owned Homes Only



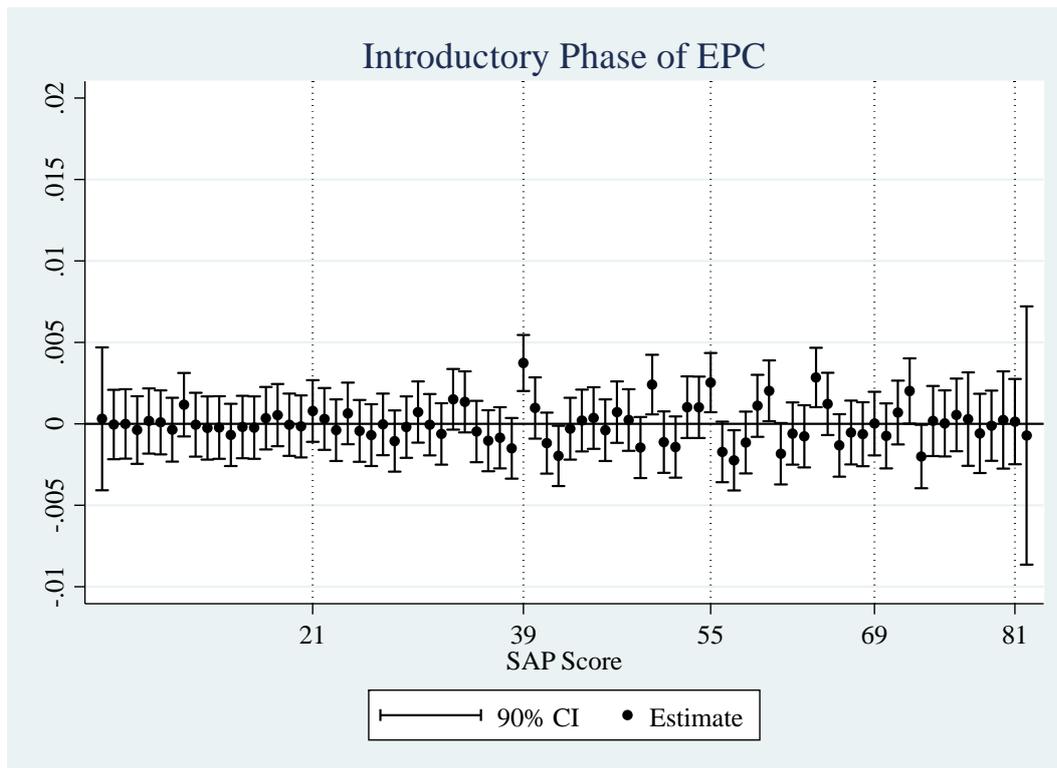
Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for the 2009 Wave of the English Housing Survey, with the sample restricted to privately owned homes. The dependent variable is the frequency of SAP score in the post-EPC years. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa. Confidence intervals are not reported when larger than the vertical axis scale.

Figure S7: Post-EPC: Pre 1990s Dwelling Only



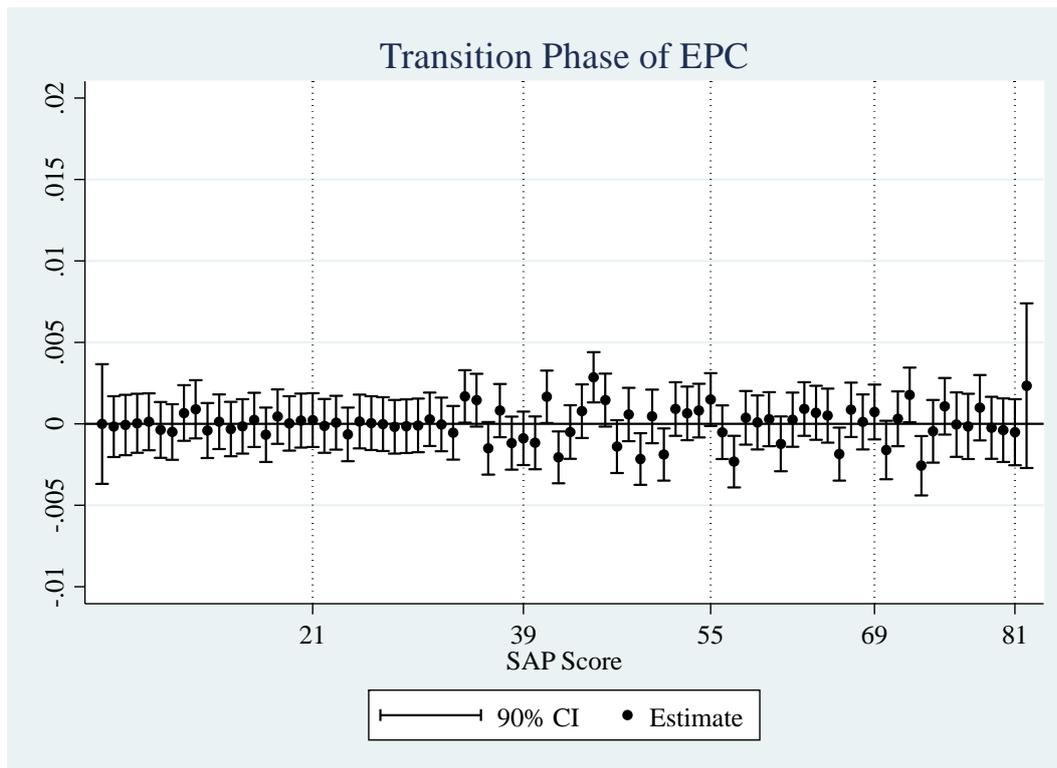
Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for the 2009 Wave of the English Housing Survey, with the sample restricted to homes built prior to 1990. The dependent variable is the frequency of SAP score in the post-EPC years. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa. Confidence intervals are not reported when larger than the vertical axis scale.

Figure S8: Introductory Phase – 2007 Wave



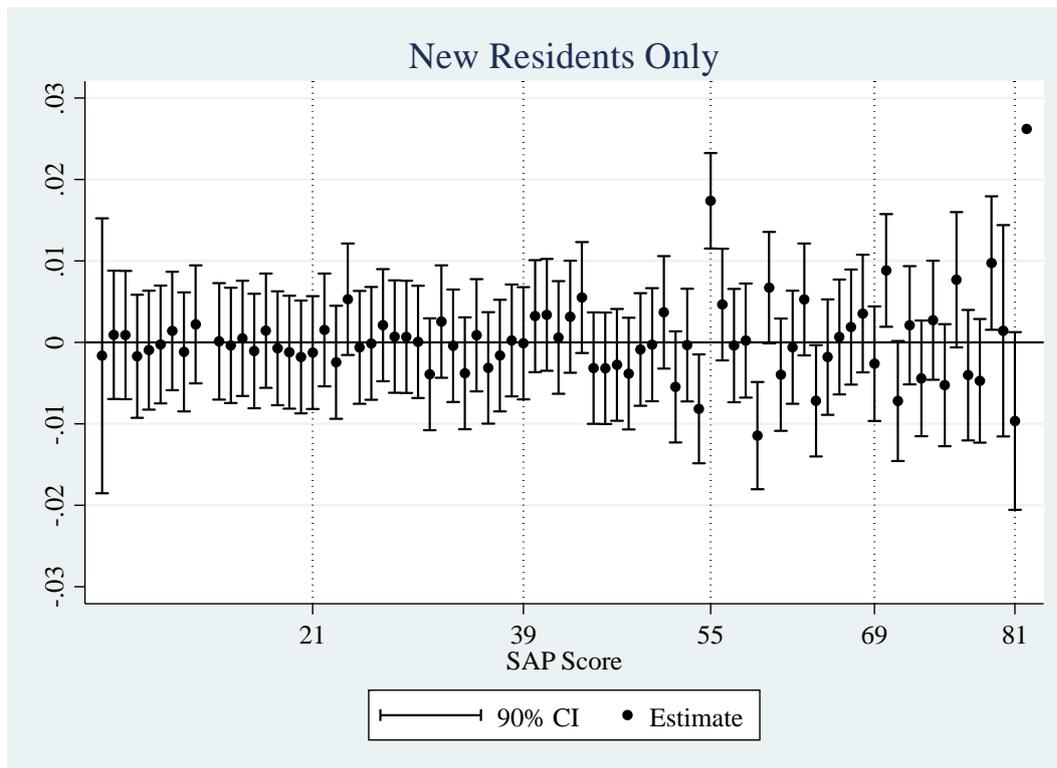
Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (19th order in this case) for the 2007 Wave of the English Housing Survey. The EPC requirement came into effect incrementally from August 2007. The dependent variable is the frequency of SAP score. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa. Confidence intervals are not reported when larger than the vertical axis scale.

Figure S9: Transition Phase – Pooled 2005 and 2006 Waves



Notes: The figure shows the estimated coefficient and 90th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (14th order in this case) for the pooled 2005 and 2006 waves of the English House Condition Survey. The dependent variable is the frequency of SAP score in the post-EPC years. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa. Confidence intervals are not reported when larger than the vertical axis scale.

Figure S10: New Residents Only - Estimates for Deviation from Optimal Polynomial



Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (18th order in this case) with the 2009 wave of the English Housing Survey restricted to dwellings occupied by new residents (i.e., which length of residents is less or equal to 1 year). A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa. Confidence intervals are not reported when larger than the vertical axis scale.