Division of Economics and Business
Working Paper Series

The Supply-side Effects of Energy Efficiency Labels

David Comerford
Ian Lange
Mirko Moro

Working Paper 2016-01
http://econbus.mines.edu/working-papers/wp201601.pdf

Colorado School of Mines
Division of Economics and Business
1500 Illinois Street
Golden, CO 80401

January 2016

Colorado School of Mines
Division of Economics and Business
Working Paper No. **2016-01**
January 2016

Title:
The Supply-side Effects of Energy Efficiency Labels*

Author(s):
David Comerford
Division of Economics
University of Stirling
Stirling, FK9 4LA
UK
david.comerford@stir.ac.uk

Ian Lange
Division of Economics and Business
Colorado School of Mines
Golden, CO 80401-1887
ilange@mines.edu

Mirko MoroDivision of Economics
University of Stirling
Stirling, FK9 4LA
UK
mirko.moro@stir.ac.uk

ABSTRACT

We build on research documenting demand-side consequences of energy-efficiency labels for buildings by testing for a supply-side response. We exploit a natural experiment to test whether the introduction of mandatory energy labels for residential homes influenced investment in home energy efficiency. From 2008, vendors and lettors in the UK were required to publish a property's energy performance certificate (EPC). The EPC evaluates home energy efficiency overlaying a color-coded letter grade (from a green A to red G, respectively) on a pre-existent 0-100 point scale, the Standard Assessment Procedure (SAP) score. We hypothesize that the salient color letter grades will serve as targets when home owners are deciding the scale of investment to make in home energy efficiency. Consistent with this hypothesis, we find fewer homes just below, and more homes just above, the D grade threshold in the treatment years relative to the control years. This clustering is higher for homes that were traded after the EPC requirement was in effect. We conclude that there is a supply-side response to energy-efficiency labels.

JEL **classifications: Q48, L15, Q58, H23**

Keywords: Energy efficiency, bunching, labels, thresholds

*The paper benefited from comments from Jim Sallee, Sebastien Houde, and Harrison Fell as well as from participants at the Association of Environmental and Resource Economics session of the 2015 Midwest Economics Association Conference at seminars at USA Environmental Protection Agencys National Center for Environmental Economics, University of Manchester, University of Stirling, Heriot Watt University, and Colorado School of Mines.

Introduction

There is debate regarding the efficacy of energy-efficiency labels as a means to "green" the economy. Previous research has tested for, and found, an effect of energy labeling on demand for energy efficient buildings (Eichholtz et al, 2010; Brounen and Kok, 2011). However, increased demand for energy efficient homes is not sufficient to insure a greener housing stock. "If building labeling affects property purchase and rental decisions, there will be no immediate impact on the energy efficiency of the building stock; labels may alter the choice among properties, but does not affect the energy efficiency of the overall building stock because it does not directly affect energy use and investment decisions." (Stavins et al, 2013, p. ES2-3) Demand for energy efficient homes will only translate into a greener housing stock if the supply side – builders and existing home owners - respond by investing in energy efficient construction. It is that supply-side response that we test for.

Our specific research question is whether a law requiring the publication of an Energy Performance Certificate for homes for sale or rent in the UK has had a causal effect on the level of energy efficiency of UK residential homes. This research contributes to the literature on the consequences of energy-efficiency labeling in several important respects. Firstly, we test for an effect of energy labeling on the outcome that is of direct interest for greening the economy: the energy efficiency of the housing stock. Secondly, we exploit a natural experiment to answer the policy-relevant question: does the introduction of an energy label have a causal effect on market behavior? Previous research has found a "green premium" in terms of rents and purchase prices paid for buildings that are labeled energy efficient (Eichholtz, et al, 2010; Brounen and Kok, 2011). Because the "green premium" is estimated from correlations and energy-efficiency labels are not randomly assigned, it remains possible that the "green premium" is, in part, a "nice building" premium. The "green premium" could derive from some desirable feature of a building that is not controlled for but that correlates with green certification. For instance, large windows maximize natural light, which is good for energy efficiency, but they also maximize views, which provide a hedonic benefit. In this example, a good view might be the attribute that is valued by the consumer. Moreover, even if the green premium is a real phenomenon, its existence does not prove that energy labeling impacts market behavior. The green premia reported in Eichholtz et al (2010) and Brounen and Kok (2011) are consistent with energy labels having no effect whatsoever. Consumers might attend directly to the green characteristics of a building (e.g. its construction, its heating bills etc.), and attach no decision weight to the energy label that proxies for these characteristics. We can draw a direct causal link between energy labeling and the behavior of market participants because we test whether the level of home energy-efficiency is predicted by a specific arbitrary detail in the design of the EPC energy label.

To determine the existence of a causal effect of the EPC requirement on home energy investment decisions, we exploit an idiosyncrasy in the UK's implementation of the requirement. The UK already had an energy efficiency measure in place prior to the announcement of the EPC requirement - the Standard Assessment Procedure (SAP) which measured efficiency on an unlabeled continuous scale from 0 to 100. The SAP score is a linear representation of home energy costs based on the attributes of the building, assuming constant preference for temperature among residents. A transition from 67 to 68 SAP should lead to the same reduction in energy costs as a transition from 68 to 69 SAP. In order to implement the EPC requirement, the UK coopted the SAP and overlaid its 0-100 scale with the EU energy label, a color-coded chart of letter grades (from A-G, where A is green and G is red). Crucial to our identification strategy, there are discontinuities in the mapping from SAP points to EPC letter grades. For instance, at 68 SAP points an extra point moves a home from a yellow "D" grade to a green "C" grade on the EPC, whereas at 67 SAP points an extra point would have a barely perceptible impact on the EPC – the certificate remains a yellow "D". In other words, depending on a home's initial SAP score, a small investment in home energy-efficiency can either have a very salient effect on the EPC – a change of color and letter - or virtually no effect. If the EPC requirement has induced investment in home energy efficiency, then we would expect it to occur where investment has most impact on the result of the published EPC: where investment pushes a home from a lower letter-grade into a higher one.

Our hypothesis then is that there will be an overrepresentation of homes just above a letter grade threshold in the years since the EPC requirement was announced. We test this hypothesis using data from the English Housing Survey on a representative sample of homes in England spanning the treatment period after the announcement of EPC requirement (2007-2009). Our econometric analyses find significant clustering just above the threshold in the middle of the distribution, grade D. There is no evidence of clustering around letter grades for homes that are exceptionally energy efficient, or for homes that are exceptionally inefficient. Relative to the distribution of homes in the control period before the introduction of EPCs, our regressions estimate that around 0.004% of homes – corresponding to about 100,000 homes -- moved to just above a letter grade threshold. Such a clustering could only be explained by chance or by the fact that homeowners responded to the EPC requirement by investing in the energy-efficiency of their homes. To rule out chance as an explanation, we test whether clustering occurred in the control years before the EPC requirement was announced.

To further test our supply-side story, we check whether homes which were traded in the treatment period have a stronger treatment effect than houses that were occupied by the same household throughout the study period, as one is likely to receive an EPC only when the home is for sale or rent. Regression results find that the amount of clustering is twice as large for residents who moved in after the treatment period than it is for residents who moved in before. We conclude that the EPC requirement has induced investment in home energy-efficiency.

Background and theory

In 1992, the Building Research Establishment (BRE - at that time a government-funded research laboratory) developed the SAP score to help the UK government monitor progress in residential energy efficiency. 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, and its concomitant color-coded letter grade, is not confounded by the occupants' use of the building, or their energy-saving behavior. As such, the EPC adds new information relative to what a potential purchaser could glean from other sources available in the market, e.g. utility bills.

SAP scores are derived from a professional audit of the building's structural properties. Importantly, the auditor does not decide on the final score. Instead, the attributes of the building, such as presence of loft insulation, boiler fuel and type, etc., are entered into an algorithm, which derives the score. This fact is useful for our analysis because it rules out the potential confound that auditors round up to the next letter grade or are biased to use scores that end in a 0 or a 5.

The EPC requirement is officially known as EU directive 2002/91/EC. It requires 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. The UK government made this directive law through the Housing Act of 2004. The EU Energy Label, codified in EU Directive 92/75/EC, is the presentation format that was adopted for the EPCs. Crucial to this analysis, the UK grafted the 7 color-coded letter grades of the EU Energy Label on top of the 0-100 SAP score. Each letter-grade 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).

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, the EPC is available for all properties for sale or rent since October 2008, making it easy for the layman to compare the energy efficiency of any two buildings. In what follows, we refer to this set of laws as the EPC requirement.

Theory: Salience, selective attention and transaction costs

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. Just as an extra mile driven is more salient when it moves an odometer from 9999 to 10000 than when it moves an odometer from 9986 to 9987, an extra SAP point is more salient when it moves from 68 – a yellow D grade – to 69 – a green C grade on the EPC. Consumer Focus (2011), a British non-profit consumer advocacy group, ran a number of focus groups to learn about 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-coded 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). In short, there is evidence that consumers are selectively attentive to the color-coded 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 – household appliances and second-hand cars. Houde (2014a) finds that consumers pay a premium over and above anticipated costs savings to buy a refrigerator with *Energy Star* certification. Lacetera et al (2012) demonstrate a price premium for cars just below a salient mileage threshold e.g. 100,000 miles, relative to those at other points in the distribution. If the value of mile driven can change as a result of its salience, it is plausible that the value of an SAP score might be similarly sensitive to salience. The fact that advertisements on market-leading property websites in UK such as Rightmove.co.uk and Zoopla routinely report only the letter grade of the EPC, and not the SAP score, will likely compound this salience effect.[1]

The salience of the letter-grades suggests that, at the margin, suppliers of homes might strategically invest in home energy efficiency to insure that their homes just meet the threshold to get into a letter grade. If this is so, we would expect to see a clustering of homes just above each letter-grade threshold. Our prediction is lent support by recent empirical work. 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, suggesting that suppliers are motivated to gain Energy Star certification for their products. Sallee (2012) investigates how building construction and design responds 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. He finds an over-representation of buildings just above silver, gold, and platinum threshold points, and a dearth of buildings just below these thresholds. He explains this pattern in terms of strategic response to LEED certification.

A complementary but distinct mechanism by which clustering above a letter-grade threshold might occur is through a reduction in transaction costs. Transaction costs fall at the home acquisition stage because the EPC requirement reduces the costs to the potential buyer of informing herself about the energy efficiency of homes in her choice set. Importantly, transaction costs also fall at the stage of home improvement. The EPC requirement reduces transaction costs by creating a single, easy-to-interpret, standardized metric through which a non-expert client can articulate to an architect, builder, or other contractor her preferred level of home energy efficiency. Let's take the case of a client who commissions a loft conversion. In the absence of a clear and precise metric, demand for home energy efficiency has to be expressed piecemeal. The client must demand energy-efficiency in each of the various products and

[1]See, e.g., these webpages: http://web.archive.org/web/20150619122901/http://www.rightmove.co.uk/new-homes-for-sale/property-48232456.html and http://web.archive.org/web/20150619122740/http://www.zoopla.co.uk/for-sale/details/36772562?search_identifier=cde062d8d1b9d2f207dfd087daad8681.

materials that go into the loft conversion. It is very costly for the non-expert client to research each product and material so as to choose the insulation, boiler, carpet, windows etc. that best trade off upfront costs against energy efficiency. Far more efficient to give the contractor a single target of energy efficiency for the project as a whole, and let the expert determine how to most efficiently hit that target. A standardized metric of energy efficiency for the building as a whole gives a shorthand by which the client can both express the specific level of energy-efficiency she demands, and also monitor whether her demand has been met. Reducing the costs of articulating demand should increase effective demand. Reducing monitoring costs should deter contractors from cheating (e.g. installing cheaper, less energy-efficient technologies to reduce costs) thereby increasing the realized level of energy efficiency. Reduced monitoring costs will also increase clients' confidence that they will get the product they demanded – increasing demand. In short, the EPC creates a clear and precise vocabulary which converts notional demand for energy efficient construction into effective demand.

Both the transaction costs story and the selective attention story suggest that investments in home energy efficiency will target the lowest SAP score in an EPC letter-grade. That yields the hypothesis that there will be an over-representation of homes at the lowest level of each letter grade. If this hypothesis is supported it will provide evidence that energy labeling affects the energy-efficiency of the housing stock.

In section 3 we set out our methodology in greater depth. In section 4, we present our results and discuss the implications. In section 5, we conclude.

Data, Hypotheses and Methodology

Our data are the 2002-2004 waves of the English Housing Conditions Survey and 2007-2009 waves of the (renamed) English Housing Survey. In the course of this survey, homes are audited and rated, resulting in an SAP score. SAP scores were measured on the same 1-100 scale in all waves of the survey utilized here.[2] It is only with the introduction of the EPC through the Housing Act of 2004 that SAP scores became a published metric of home energy efficiency. By the year 2008, EPCs became required for both sales and rentals of property. The analysis here will test for clustering just above the letter grade thresholds in the 2007-2009 data. The 2002-2004 data serves as a control group in which we do not expect to see clustering. Data from 2005 and 2006 are not included here as these years were the transition period between passage of the Housing Act of 2004 and its implementation.

A further hypothesis concerns heterogeneous treatment effects. We predict that our effect will be strongest in the middle of the energy-efficiency distribution. One reason for this is a sample size

[2]Unfortunately, from the 2010 English Housing Survey SAP scores were calculated using a different metric than had been previously utilized. BRE occasionally update the way SAP scores are calculated and it is not clear how to compare later versions of the score to earlier ones. All of the SAP scores used in this analysis are calculated according to SAP 2005.

consideration – the number of homes in the tails of the distribution is relatively small and so our test may lack sufficient power to detect an effect.

A second reason is that homes in the upper tail of the distribution have already exploited the most cost effective energy-saving technologies, and so the marginal investment required to move these homes across a letter-grade threshold is more costly than is the case in the middle of the distribution. Also, the exceptionally high energy-efficiency level of these homes suggest that their owners are motivated to invest in energy-efficiency by concerns other than crossing a letter-grade, such as environmental preferences or long-term cost-savings.

Motivation to cross to the lowest point in a letter grade is also likely to be less at the bottom of the distribution than it is in the middle. By definition, a home in the F letter grade will not be equipped with many energy-saving technologies. If its owners consider energy-efficiency at all, they will come to recognize that there are many energy-saving technologies available that would save money. Taking this into account, and also the fact that the UK government offered subsidies for the installation of energy-saving technologies in this period, it might be less costly for a home starting in the G letter grade to add 20 SAP points than it would be to add 2 SAP points.

A further factor that suggests heterogeneous treatment effects across the distribution is the connotations of the letter grades. In the UK school system, to move from an F to an E is to move from one fail grade to another. Paraphrasing Newall and Siikamaki's conclusions from their choice experiment: Anyone who has been to school would rather get a pass than a fail. Evidence that connotations such as this affect choice comes from Ubel et al (2015), which found that people were attracted to health insurance plans labelled gold regardless of the substantive content of those health plans. For all of the above reasons – sample size, cost and returns to investment in energy-efficiency technology, homeowner tastes for energy efficiency, and labelling connotations - we consider it most likely that we will observe our effect in the middle of the distribution e.g. E to D.

The SAP scale ranges from 0 to 100, but in practice we observe SAP scores ranging from 3 to 83 in the control and treatment years. Figure 2 shows the frequency distribution of homes by SAP score for the control and treatment years. The red vertical lines represent the letter-grade thresholds, though the letter grading system was not in existence during the control time period. Additionally, a kernel density estimation, a non-parametric prediction of the distribution, is shown.[3]

Comparing the two panels in the figure, it is clear that the distribution in the control years is smoother and more like a bell-curve. This visual impression is confirmed in the statistics given in

[3]Both kernels use a bandwidth equal to one to approximate each single SAP point.

Table 1. The distribution of the treatment years has a larger kurtosis and is left-skewed. [4] Of specific interest is the fact that the treatment distribution shows a large jump in frequency when moving from a SAP score of 54 (just below the D grade) to a score of 55 (the first SAP score in the D grade). The D grade is the first grade that is not a shade of red or associated with a fail grade in the British educational system. The shape of the distribution here is clearly different than the kernel estimation would predict, with the actual frequency at 54 smaller than the kernel prediction and the actual frequency at 55 much larger than the kernel prediction. The control distribution does not show any changes at this notch point.

Empirical analysis

While a visual inspection of the distributions is informative, a more rigorous examination of the data will reveal the extent to which the letter-grade thresholds impact the frequency of homes at each SAP score. Previous literature estimating bunching or clustering of a distribution has focused on kinks in income tax rates (Chetty et al, 2011; Saez, 2010) or automobile fuel efficiency notches (Sallee and Slemrod, 2012; Sallee and Ito, 2014). Our analyses test for systematic bunching at each point in the distribution and around letter-grade thresholds by comparing the actual density to the one predicted by the best fit polynomial. In our setting, the econometric model is given by:

$$FreqSAP_i = \alpha + \beta_1 (Dum\_SAP) + \beta_2 (Dum\_Threshold) + \beta_3 (polyns) + t_i + e_i \,, \tag{1}$$

where $FreqSAP_i$ is the frequency of homes with a given SAP score, *Dum_SAP* is a dummy for a given SAP score, *Dum_Threshold* is a dummy variable equal to one if the SAP is the lowest score in a colour-coded letter grade, *polyns* is the polynomial order used to fit the frequency distribution, t_i are year of survey dummies, and e_i is an error term. *Dum_Threshold* is added to ensure that potential jumps in the frequency induced by EPC do not affect other SAP score coefficient estimations.[5] We adopt the strategy of fitting a curve to the observed data so as to take account of the distribution and include optimal polynomial order, *polyns*. In this approach we follow the literature on bunching such as Lacetera et al (2012) and Sallee and Slemrod (2012) so that our results are not biased by the shape of the distribution.

[4]Formal tests of equality of distributions, Kolmogorov-Smirnov and Epps-Singleton reject the null hypothesis of equality of distributions at 1% significance level (both p-values=0.000).

[5]Equation 1 was estimated without the *Dum_Threshold* variables and there was no change in sign and statistical significance of the *Dum_SAP* coefficients.

Identifying the optimal polynomial order *–polyns* – is crucial. We run different polynomial regressions using from 3rd order to 15th order and identify the 12th as the best fit.[6] We select the model for which both the AIC and BIC statistics take the smallest value. Table 2 reports the AIC values while Figure 3 illustrates the scatter plot of the distribution and the polynomial fit for several polynomial models. The visual representation serves to show clearly that only higher order polynomials are justifiable in this case.

Previous literature, such as Sallee and Ito (2014), dummy out the thresholds when specifying the optimal polynomial order so that the order chosen is not influenced by the potential bunching. Our estimation of the polynomial order does not include these threshold dummies as Figure 3 shows a very good fit across a number of polynomial orders. Additionally, polynomial orders above 6 do a better job at replicating the distribution especially around the SAP score 55. One may be concerned that the distribution under the optimal polynomial order does not take into account the peak at SAP score 55 in an appropriate way. In order to alleviate this concern, Figure A1 shows the results of Equation 1 where the excess density (the difference between actual and predicted) at 55 is put back assuming a uniform distribution through SAP score 50-54. This assumes that the excess homes in the 55 SAP bunch came uniformly from the 50-54 SAP scores. It is identical in sign and statistical significance to the results in Figure 4.

Model (1) is run for each SAP score with a non-zero frequency where every iteration has a different *Dum_SAP* turned to one such that each SAP score has its dummy turned on once. In other words, this specification tests for deviations from the smooth distribution at each point in the distribution. Our prediction is that the coefficient estimate on *Dum_SAP*, β_1, will be positive and statistically significant in the treatment years at the lowest SAP scores in the letter grades at the centre of the distribution, i.e. SAP score 55, the transition from E to D, and SAP score 69, the transition D to C. We predict β_1, will not be statistically significant in the control years.

Results from Equation 1 demonstrates a statistically significant overrepresentation of homes at SAP score 55, the lowest point in the D letter grade, relative to the counterfactual of a smooth distribution. That result presents prima facie support for our theory, but it does not compare the post-EPC distribution against the most appropriate counterfactual available. There is no reason to believe that prior to the EPC requirement homes were smoothly distributed across the range of SAP scores. One advantage of our dataset is that it surveyed a representative sample of homes in repeated cross-section, which means that we can observe the distribution of homes prior to the EPC and compare the post-EPC distribution against that baseline. Accordingly, equation 2 tests for systematic bunching at each letter-grade threshold in the treatment years relative to the control group. In our setting, the econometric model is given by:

[6] Regressing on higher polynomial orders yields no results due to high collinearity.

$$FreqSAP_{treatment\,i} - FreqSAP_{control\,i} = \alpha + \beta_1 (Dum\_SAP) + \beta_2 (below\ threshold_i) + \beta_3 (above\ threshold_i) + \beta_4 (polyns) + t_i + e_i, \quad (2)$$

the dependent variable is now the difference in frequency between treatment and control group at each SAP point. Model (2) tests, whether there are more homes just above the threshold in the treatment years relative to the control years. If so, we would expect coefficient β_3 to be positive. The inclusion of polynomials in model (2) insures that we are not picking up significant results simply as a result of the shape of the distribution. Similarly to what we have done previously, Model (2) is run for each SAP score with a non-zero frequency where each iteration has a different *Dum_SAP* turned to one such that each SAP score has its dummy turned on once. The selection of the optimal polynomial order is obtained using AIC and BIC statistics as before. The optimal polynomial order is 19th. However such a high number of polynomials results in many coefficients being dropped because of collinearity. Because results with respect to salient SAP scores are identical independently of polynomial order, in what follows we present estimates from the 9th polynomial order, which retains the properties of having high AIC and BIC statistics together with parsimony.

Model (2) also tests a possible mechanism by which the result comes about. Our selective attention story suggests that homes that are at the highest SAP scores in a letter grade have the highest expected returns to marginal investments in home energy efficiency, since little investment would bump these homes into a higher letter grade. That story implies that the cluster of homes found in the treatment years will be disproportionately drawn from the highest SAP point of the immediately lower grade. If that is the case, we would expect β_2 to be negative and significant.

Results

Figures 4-5 contain the results from the estimations of Equation 1 for the treatment and control period separately. The coefficient and 95% Confidence Interval (CI) are shown for each SAP score.[7] Figure 4 shows the results for the treatment period. The 55 SAP score is one of two SAP scores statistically different than zero with a 95% CI. As would be expected, almost every coefficient is not statistically different than zero given that they are essentially tests for deviations from the optimal polynomial order. The estimate of the amount of the bunch is 0.003% of the frequency. According to the Census 2011, there are around 25 million houses in the UK. Assuming that the data are a representative sample of homes in the UK, that would imply about 75,000 houses bunched at this SAP score. Figure 5 shows the results for the control period. The 55 SAP score is not statistically different than zero for this sample, implying that the

[7]Confidence intervals were computed from robust standard errors.

number of homes at a 55 SAP score were initially in line with the rest of the distribution and then spiked after EPCs were required.

So far there is a systematic higher density just above grades at the middle-right of the distribution (letters, E to D). Figure 6 reports results from an estimation of Model (2), which compares the treatment and control distributions. If households made marginal investments to move from just below to just above each notch, we would expect a negative sign on β_2 and a positive sign on β_3. Our results clearly show a positive sign on β_3 around thresholds D and C (i.e., SAP scores 55 and 69), indicating again that there is clustering of homes at a threshold. The coefficient on SAP score 55 is 0.004, which translates to around 100,000 homes bunched at this SAP score using the 2011 Census figures. The coefficient on β_2 is not significant, indicating that there are no fewer homes just below a threshold in the treatment years than in the control years. This result tells us that the homes that cluster on the threshold are not drawn disproportionately from just below the threshold, indicating that our results are not driven by small marginal investments in home energy efficiency.

Figures 7 and 8 presents a further test of the mechanism which leads houses to bunch at the 55 SAP score. Given that EPCs are required when one rents or sells their home, it is expected that homes that were traded since the EPC requirement was announced are more likely than others to manifest marginal investment. Figure 7 shows the results only for homes with new residents. The 55 SAP score is again positive and statistically different than zero. The 55 SAP score coefficient is larger for new residents than when looking at all homes in the treatment period (Figure 4), which is consistent with the mechanism that the information on the EPC leads to changes in the homes SAP score. Figure 8 shows the results for homes which the residents have not moved during the treatment period. The 55 SAP score is not statistically different than zero, though it is statistically different at the 90% CI (not shown).

Conclusion

We tested whether the requirement that an EPC be published for every home on the market affected the level of investment in home energy efficiency. We take advantage of a feature of the energy efficiency labeling of houses whereby a color-letter grade was overlaid on to an existing 0-100 scale. As a result, there are discontinuities in the mapping from numbers to color-letter grades. For instance, improving the efficiency of a house just below this discontinuity leads to a change in the letter-color grade, whereas improving the efficiency in some other parts of the distribution does not change the letter-color grade.

Our empirical strategy is to test for bunches by comparing the distribution of homes before and after the initiation of the color-letter grades. We find that the EPC lead to a bunch at the first number of the letter "D" grade. Using Census statistics, this would imply about 100,000 houses bunched at this SAP score, while no evidence of bunching is found when looking at the distribution before the introduction of the letter-color labeling.

There are a number of limitations to our study and further research is required before we conclude that energy efficiency labels will always and everywhere have an effect on investment in home energy efficiency. The first limitation is one that it shares with any natural experiment: it took place at a specific time and place and so cannot be exactly replicated. We will have to leave it to future research to determine the institutional and cultural features that make a population receptive to EPC labels.

An important point to stress regarding these results is that they cannot determine whether the overall impact of the EPC requirement was to increase energy efficiency relative to the amount of investment that would have taken place in the absence of the EPC requirement, or whether the EPC requirement reduced investment relative to the counterfactual. We found that relative to the control years, there was an overrepresentation of homes at the C and D thresholds in the treatment years, but our prediction was not supported that those homes arrived at those points due to small marginal investments: We tested for an underrepresentation of homes just below the thresholds in the treatment years relative to the control years, but did not find it. Our transaction costs story is consistent with this pattern of results, and also with a positive effect of the EPC requirement on investment. The notional demand that the EPC requirement converted into effective demand for energy efficiency is no more likely to occur the upper end of an SAP letter grade than at any other point, so it follows that we would not see a deficit of homes at the top of a letter grade with the introduction of the EPC requirement. On the other hand, our selective attention mechanism suggests a story by which the EPC requirement could have reduced investment. Imagine a property developer recognized that, once a letter grade was achieved, consumers pay little attention to further gains in energy-efficiency. A strategic property developer might revise their blueprints so as to reduce the resultant SAP score from the middle of a letter-grade to its bottom, and thereby reduce construction costs. That disinvestment would manifest itself as one more home clustering at the lowest SAP score in the letter grade. Our data cannot tell us that that home would otherwise have been at a higher SAP score in the middle of the letter grade. The overall impact of the EPC requirement on home energy efficiency is a question that should be explored in future research as it has important implications for the environmental effects of energy labels.

A related caveat to these results is that we tested for one very specific and localised consequence of the EPC requirement. Our parameter estimate does not approximate the average change in energy efficiency caused by the EPC requirement. Instead, the parameter estimates the *additional* change in energy efficiency that is induced by targeting a letter-grade on the EPC, over and above any effect that the publication of EPCs has at other points in the distribution.

In spite of the questions that this research leaves open, it contributes to the policy debate around energy efficiency labels in three important respects. Firstly, it demonstrates that energy labels have a causal effect on the behavior of agents in the residential housing market. Secondly, it demonstrates an effect on the level of investment in energy efficiency technologies, thus providing evidence of a supply-side response to energy labels. Thirdly, it demonstrates that energy labels affect an outcome that is of direct environmental relevance: the energy efficiency of the housing stock. Though these results cannot determine whether the environmental impact of the EPC requirement has been positive or negative overall, we see these results as coming down on the positive side in the debate on energy labeling. They are the first evidence of any supply-side response in the residential housing market to energy labels, and so suggest that a well-designed energy label policy could green the housing stock.

References:

Allcott, Hunt and Michael Greenstone (2012) Is there an Energy Efficiency Gap?, Journal of Economic Perspectives, 26(1): 3-28.

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky (2014) Energy Policy with Externalities and Internalities, Journal of Public Economics, 112: 72-88.

Brounen, Dirk and Nils Kok (2011) On the Economics of Energy Labels in the Housing Market, Journal of Environmental Economics and Management, 62(2): 166-179.

Busse, Megan, Nicola Lacetera, Devin Pope, Jorge Silva-Risso and Justin Sydnor (2013) Estimating the Effect of Salience in the Wholesale and Retail Car Markets, American Economic Review Papers and Proceedings, 103(3): 575-579.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 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

DellaVigna, Stefano (2009) Psychology and Economics: Evidence from the Field, Journal of Economic Literature, 47 (2): 315–72.

Eichholtz, Piet, Nils Kok, and John Quigley (2010) Doing well by doing good: Green office buildings, American Economic Review, 100 (5): 2492–2509

Gillingham, Kenneth and Karen Palmer. Bridging the Energy Efficiency Gap: Policy Insights from Economic Theory and Empirical Analysis. Review of Environmental Economics & Policy (2014), 8(1): 18-38

Houde, Sebastien (2014a) How Consumers Respond to Environmental Certification and the Value of Energy Information, Working Paper

Houde, Sebastien (2014b) Bunching With the Stars: How Firms Respond to Environmental Certification, Working Paper

Ito, Koichiro, and James M. Sallee. (2014) The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards. No. w20500. National Bureau of Economic Research

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

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

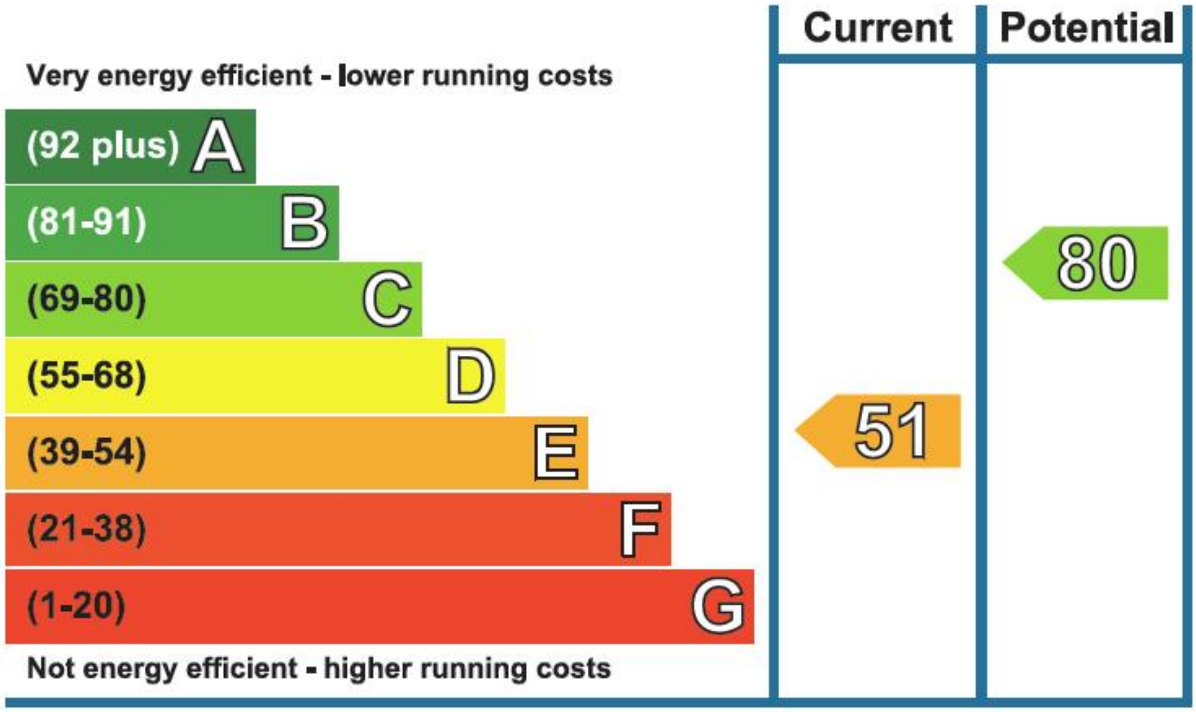
Saez, Emmanuel. (2010) Do taxpayers bunch at kink points?, American Economic Journal: Economic Policy 2(3): 180-212.

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.

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

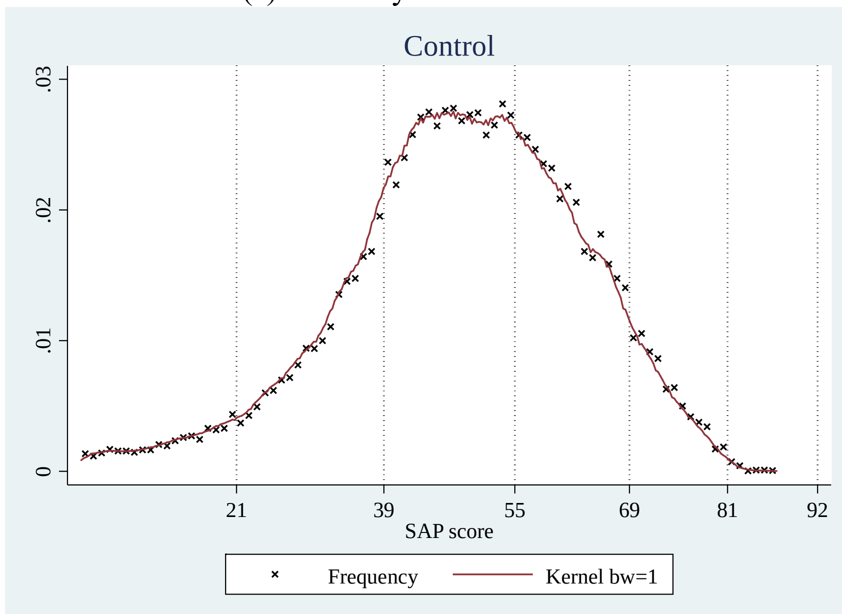
Figure 1: An Energy Performance Certificate, showing SAP scores (0-100) and letter grades



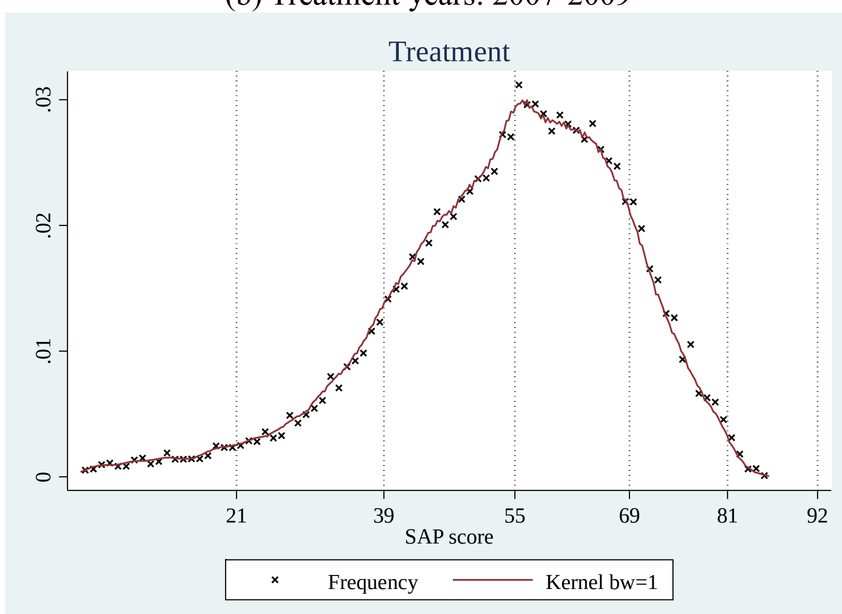
Notes: An example Energy Performance Certificate showing the SAP numerical score and the letter-color grade.

Figure 2: Frequency of houses by SAP score for control and treatment years

(a) Control years: 2002-2004

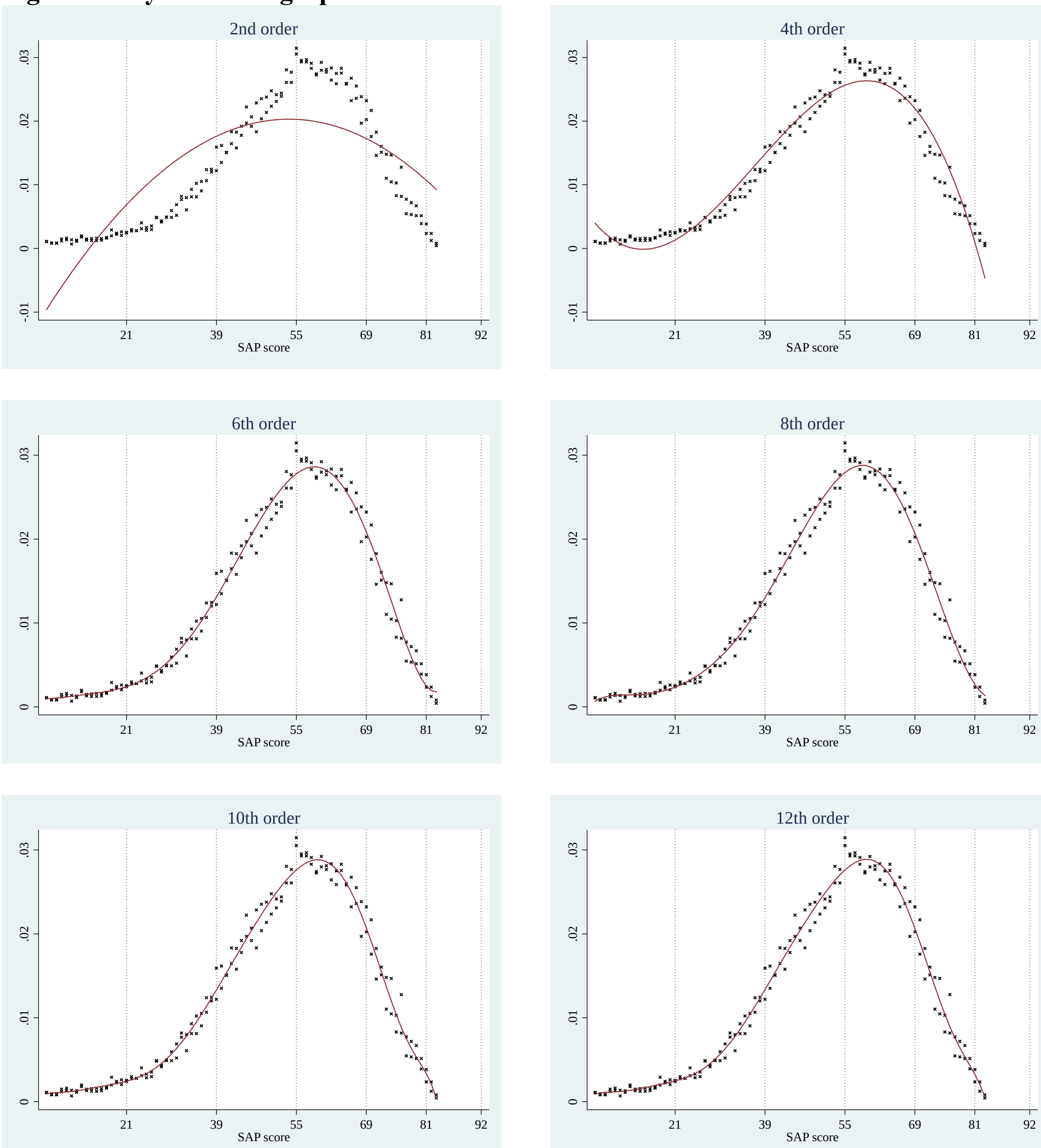


(b) Treatment years: 2007-2009



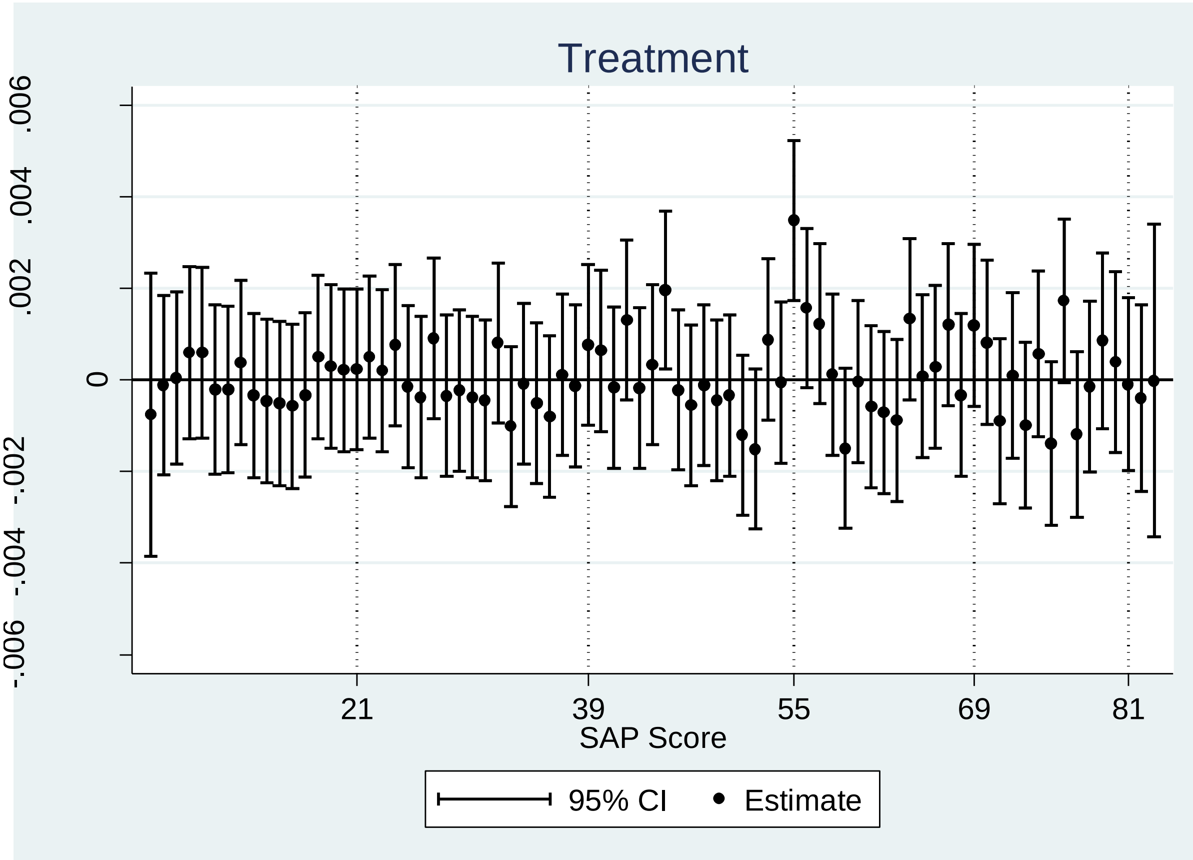
Notes: Frequency of houses by SAP score for the years 2002-2004 in panel (a) and for the years 2007-2009 in panel (b) with the first score in a new color-coded letter grade represented with the vertical lines. A kernel density estimation with a bandwidth of one is shown.

Figure 3 Polynomial fit graphs



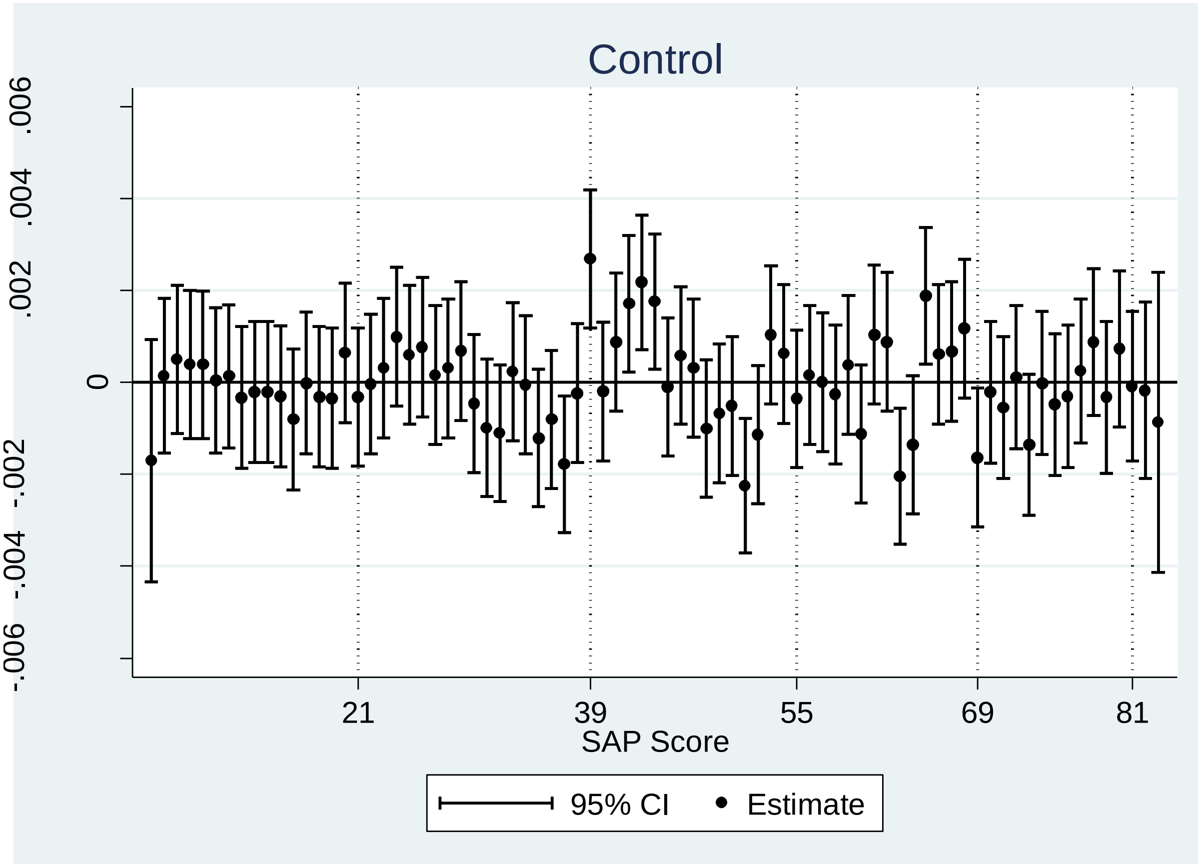
Notes: These figures show how well different polynomial orders match the distribution of our data in the treatment years by displaying the actual data points and the fitted line from a given polynomial order. The fit is similar for polynomial orders from 6 and above. See Table 2 for statistical tests of goodness-of-fit.

Figure 4: Estimates for Deviation from Optimal Polynomial-Treatment Sample



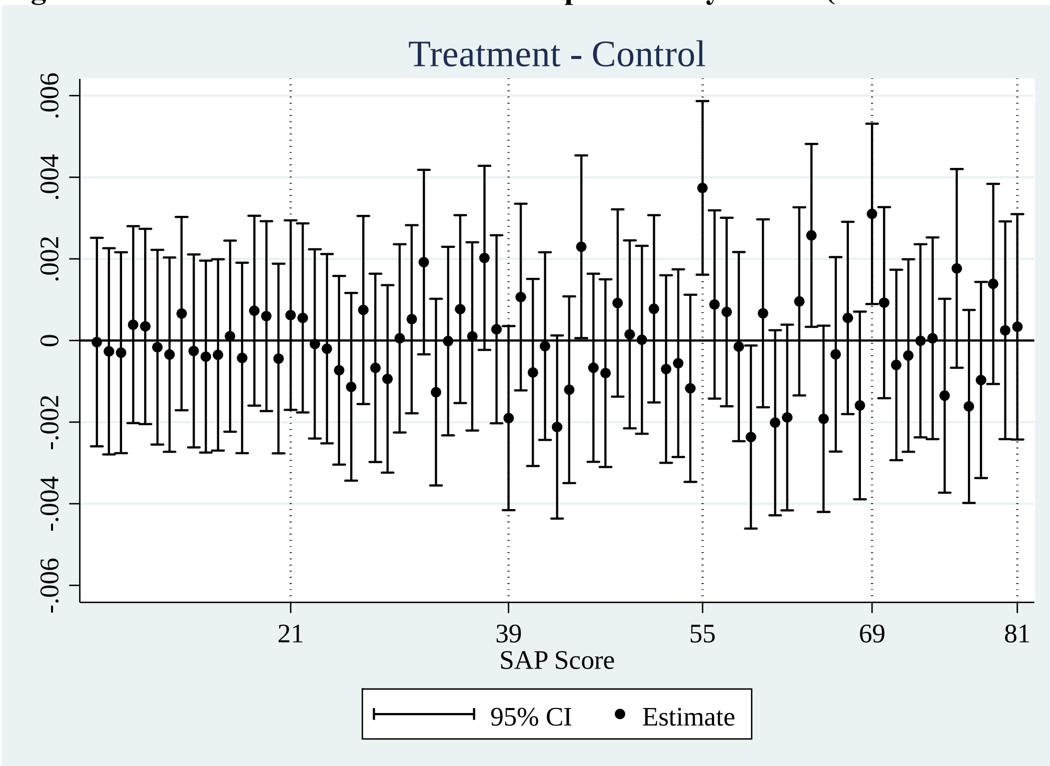
Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for the years in which Energy Performance Certificates were required. The dependent variable is the frequency of SAP score in the years 2007-2009. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Figure 5: Estimates for Deviation from Optimal Polynomial-Control Sample



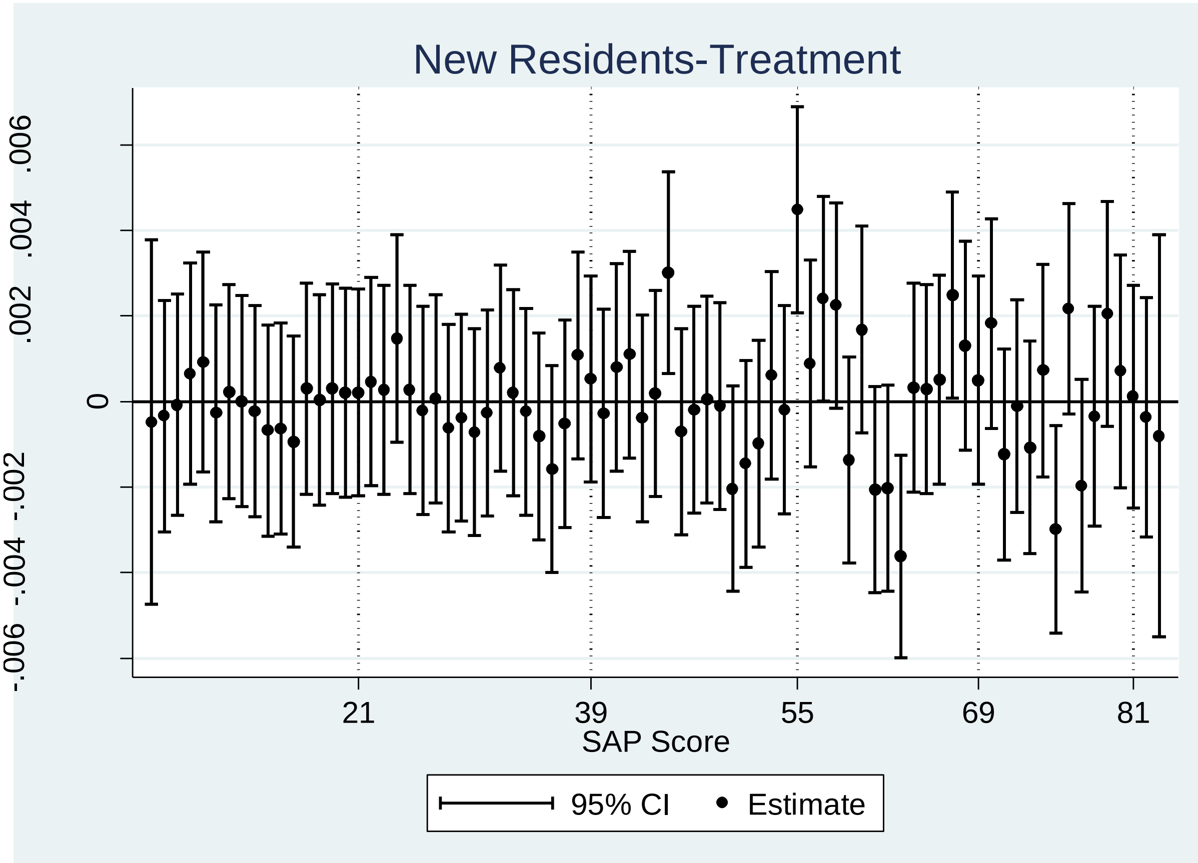
Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for the years in which Energy Performance Certificates were *not* required. The dependent variable is the frequency of SAP score in the years 2002-2004. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Figure 6: Estimates for Deviation from Optimal Polynomial (Treatment – Control)



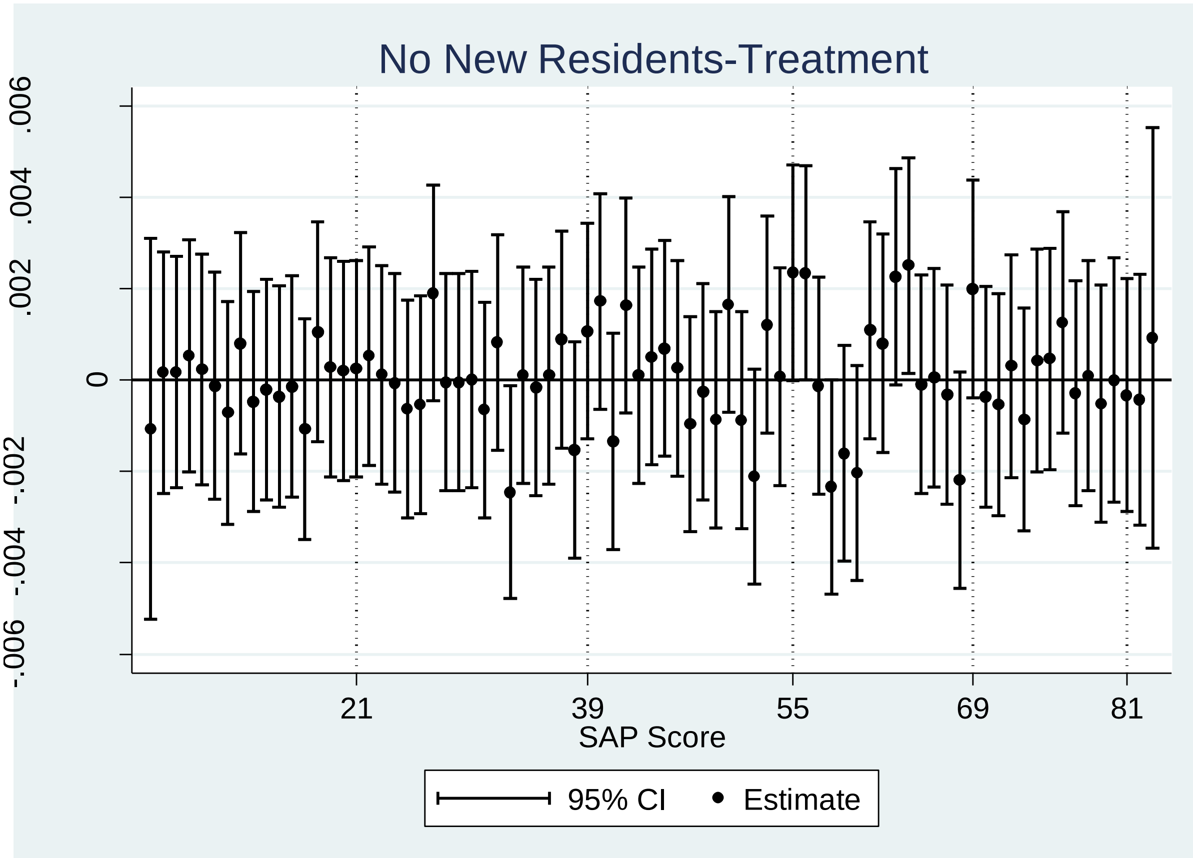
Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation difference of each SAP point from the 9th its optimal polynomial order. The dependent variable is for the treatment period minus the difference in frequency of SAP score in the treatment years – control years, similar to a difference-in-difference estimate. A positive number implies more homes in the treatment years versus control years at a SAP point than predicted by the optimal polynomial order in the treatment period relative to the control period and vice versa.

Figure 7: Estimates for Deviation from Optimal Polynomial-New Residents



Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for the for residents who moved in after an Energy Performance Certificates was required. The dependent variable is the frequency of SAP score for homes in the years 2007-2009 with residents who moved in since 2007. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Figure 8: Estimates for Deviation from Optimal Polynomial-No New Residents



Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for homes surveyed after the Energy Performance Certificates were required but the residents who moved in before Energy Performance Certificates were required. The dependent variable is the frequency of SAP score for homes in the years 2007-2009 with residents that moved in before 2007. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.

Table 1: Characteristics of the Distributions

Status	Mean	St dev	Skewness	Kurtosis
Control	48.43	14.49	-0.42	3.14
Treatment	53.94	14.36	-0.68	3.49

Notes: This table provides summary statistics of the distribution of homes along the SAP score for the years 2002-2004 (control) and the years 2007-2009 (treatment). Energy Performance Certificates were required from 2008 on after being included in the Housing Act of 2004.

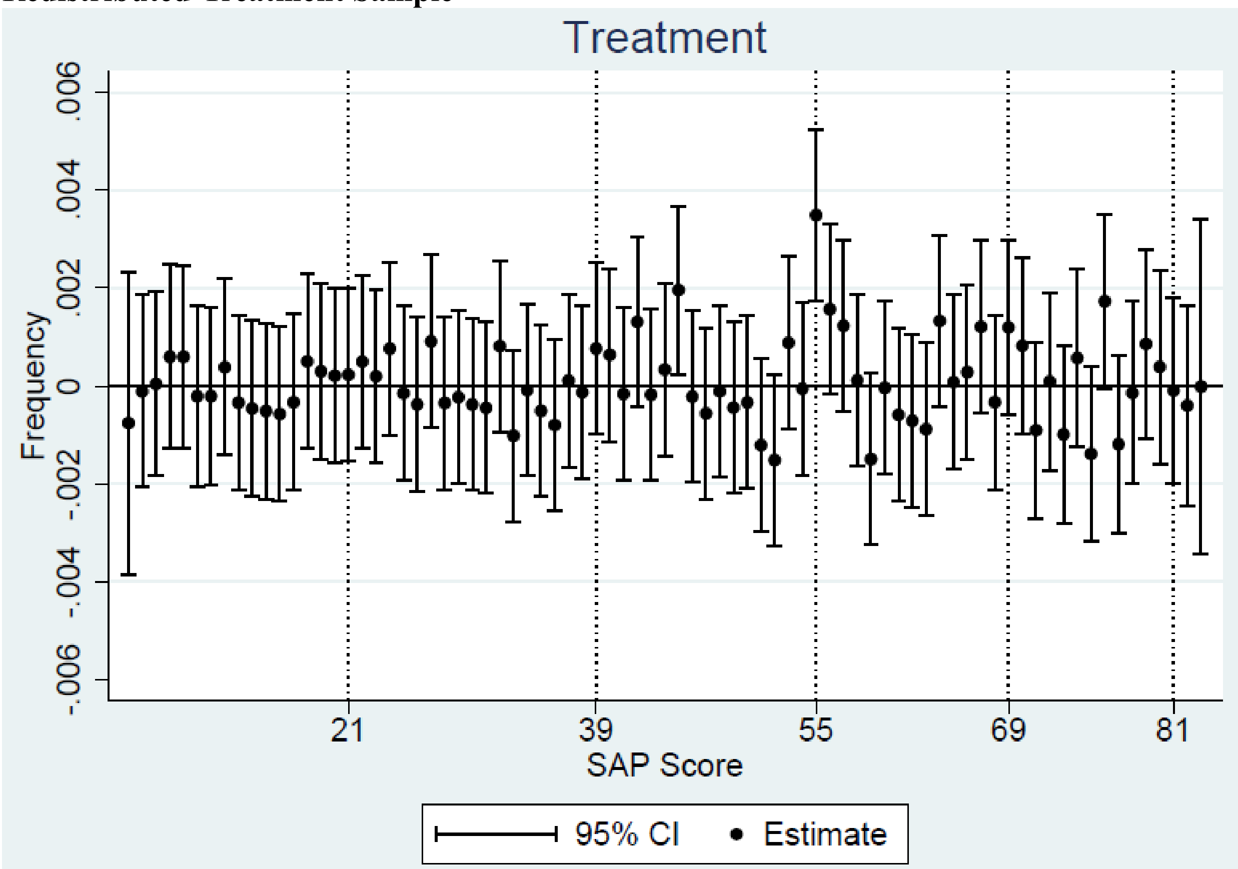
Table 2: Selection of optimal polynomial order for treatment group

Polynomial order	**2**	**3**	**4**	**5**	**6**	**7**	**8**
N	158	158	158	158	158	158	158
AIC	-1161.89	-1477.89	-1478.02	-1625.06	-1658.63	-1658.71	-1660.65
BIC	-1152.7	-1465.64	-1465.77	-1612.81	-1643.32	-1643.39	-1645.34
Polynomial order	**9**	**10**	**11**	**12**	**13**	**14**	**15**
N	158	158	158	158	158	158	158
AIC	-1667.11	-1668.16	-1668.39	-1669.65	-1668.09	-1657.99	-1669.21
BIC	-1651.8	-1649.78	-1650.02	-1654.33	-1649.72	-1642.67	-1650.83

Note: This table reports AIC and BIC statistics for polynomial regressions with different degrees of polynomial from 2 to 15 for data in the treatment years.

Appendix A

Figure A1: Estimates for Deviation from Optimal Polynomial with Excess Density Redistributed-Treatment Sample



Notes: The figure shows the estimated coefficient and 95th percentile confidence intervals for the deviation of each SAP point from the optimal polynomial order (12th order in this case) for the years in which Energy Performance Certificates were required. 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 in the years 2007-2009. A positive number implies more homes at a SAP point than predicted by the optimal polynomial order and vice versa.