Applied Energy 233–234 (2019) 816–826

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Modelling energy demand from higher education institutions: A case study of the UK

Zia Wadud[a,*], Sarah Royston[b], Jan Selby[b]

[a] Centre for Integrated Energy Research, University of Leeds, Leeds LS2 9JT, UK
[b] Department of International Relations, University of Sussex, Brighton BN1 9RH, UK

HIGHLIGHTS

- Energy consumption in higher education institutions (HEIs) modelled.
- For the first time, panel data used, with cross-sectional and time-series variations.
- Over time and across HEIs, energy use increases as income and size increase.
- Economies of scale as HEIs grow, but energy use still rising with sector growth.
- Research universities more energy intensive.

ARTICLE INFO

Keywords:
Energy demand
Econometric model
Panel data
Higher education
Building energy
Carbon emissions

ABSTRACT

Among the various sustainability goals of higher education institutions (HEIs), reducing energy use and carbon emissions are particularly important. However, not much is known about energy demand from the higher education sector – especially since there is a lack of robust models of energy demand in this sector. This paper, the first to utilize a panel dataset and advanced panel econometric techniques in order to model energy use in higher education, investigates variations in energy use between HEIs (cross-sectional analysis), and also changes in energy use over time (temporal analysis), using the UK as a case study. We argue that panel dataset and methods are more useful for understanding growth (and reduction) in energy use within the HE sector than the methods used within previous cross-sectional studies. Results show that, over time and also across the sector, energy consumption in the HEIs increases with increases in income and floor space, but at a slower rate. As HEIs grow overall (in terms of income, floor space, student and staff number) over time, they become more 'energy efficient' (using less energy per unit of area, population or income), indicating economies of scale in the temporal dimension. Results also show that after controlling for income and size, research intensive HEIs consume more energy. We also find a small but statistically significant effect of energy prices on energy consumption, as might be expected. Simulation using the model parameters for an example scenario suggests that energy consumption will continue to increase unless there is a significant change in the policies driving income growth and spatial expansion in the HE sector in the UK.

1. Introduction

Higher education institutions (HEIs) are not only places for knowledge creation and dissemination, but are also major employers, economic actors, and providers of cultural, recreational and infrastructure resources [1]. The higher education (HE) sector often presents itself as leading on sustainability challenges – be it through research and technology development, dissemination of knowledge, good citizenship and environmental awareness of communities, students and staff or through changing its own corporate behaviour toward more sustainable practices [2]. Reducing carbon emissions is the latest addition to the sustainability goals of the HE sector.

While past efforts to reduce energy use and carbon emissions from HEIs have generally focused on the energy performance of buildings, little is known about the underlying reasons for changes in energy demand, such as economic activity or population changes within the HE sector. This knowledge gap presents a substantial challenge in reducing energy use and carbon emissions, as gains from technical improvements

* Corresponding author.
E-mail addresses: Z.Wadud@leeds.ac.uk (Z. Wadud), S.Royston@sussex.ac.uk (S. Royston), J.Selby@sussex.ac.uk (J. Selby).

https://doi.org/10.1016/j.apenergy.2018.09.203
Received 21 April 2018; Received in revised form 18 September 2018; Accepted 24 September 2018
Available online 01 November 2018
0306-2619/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

to buildings and equipment can be negated by increases in energy demand that result from changes in "non-energy" policies and practices. For example, Royston [3] argues that non-energy strategies such as provision of luxurious student facilities, growth in numbers of students and staff, or increased research activity can lead to increases in demand that outweigh gains made through technical measures such as insulation or LED lighting. However, the impacts of these policies are rarely recognised, and are largely "invisible". Therefore it is important to understand the drivers of energy use in the HE sector, especially through quantitative modelling, which is an under-researched area. This paper addresses this gap in the international energy demand literature by developing an advanced econometric model to understand the effects of different factors on energy demand within HEIs.

The paper makes several methodological and empirical contributions to our current understanding of energy consumption from HEIs. Firstly, this is the first study to utilize temporal variations of energy use in a cross-section of HEIs, unlike previous studies ([4–6] in the USA and Taiwan), which use only cross-sectional data at one time point and have several important limitations (explained later) that make them less useful. Secondly, we apply advanced hybrid panel models [7], in addition to the more traditional panel data methods in modelling energy use from HEIs. Thirdly, our energy consumption model is also more comprehensive than previous ones with the inclusion of several new explanatory factors, including energy price, which is – somewhat surprisingly – missing from previous models. Finally, inclusion of temporal observations in the panel econometric models also allows forecasting of energy use in the future, which is important for energy and carbon planning and management purposes but is not possible from any of the previous cross-sectional models. Our energy modelling techniques and results thus have wider applications in the international HE sector.

The paper is laid out as follows: Section 2 briefly introduces the UK HE sector as a case study, including its energy reduction and carbon mitigation initiatives. Section 3 describes the existing international literature on energy use in the HE sector, identifying the key gaps and describing further the contributions of this paper. Section 4 describes the data and methods used in detail. Section 5 presents and discusses the modelling results, including simulation results for a future scenario. Section 6 draws conclusions and the broader implications of our results and models.

2. UK HE sector

2.1. State of the sector

As mentioned above, we use the UK as the case study country for modelling energy use from the HE sector. There were 162 publicly funded higher education providers in the UK during 2016/17, with 2.32 million students and around 420,000 staff [8]. Universities UK (a lobby group representing most UK universities) estimates that in 2014–15, UK universities generated £95 billion in gross output for the economy and contributed £21.5 billion to GDP, representing 1.2% of the UK's GDP [8]. If all the HEIs in the UK were concentrated in one city, that would be the fifth largest city-economy in the UK [9]. Official data published by the Higher Education Statistics Agency (HESA) shows that HEIs own approximately 28.8 million m^2 of gross internal space [10]. Higher education has also been a fast-growing export sector [11], and UK universities have substantially increased their international student intake in recent years: the proportion of international students studying at UK universities increased from 14% in 2006–07 to 19% in 2015–16 [12]. Universities are often likened to towns or cities rather than places of work or study, because of their large size, population and various complex activities taking place on campus, including residential facilities for students, teaching and research. These diverse activities all have implications for energy demand.

2.2. Energy reduction and carbon mitigation initiatives

Overall the education sector (including primary and secondary education) is responsible for 13% of energy use in the UK [13]. In keeping with UK's national carbon emission reduction target of 34% by 2020 from a 1990 baseline, in 2010 the regulator for English HEIs, Higher Education Funding Council for England (HEFCE), set a target of 43% reduction in carbon emissions (Scope 1 and 2)[1] from the English HEIs between 2005 and 2020 [14]. However, not every HEI has the same target, rather individual institutions were allowed to set their own targets. The collective impact of institutional targets is 38%, which is lower than the sectoral target [15]. In order to incentivise HEIs to reduce their emissions, the capital investment framework of HEFCE was also modified [14]; however, this later became largely irrelevant due to changes in the structure of HE funding in England. Meanwhile, all Scottish Universities as of 2016 signed the Universities and Colleges Climate Commitment for Scotland (a public declaration that the institution acknowledges the Scottish Government targets to reduce carbon emissions by 80% by 2050). Universities in Scotland must also report carbon emissions to the Scottish Government [16]. In Wales, the sector regulator Higher Education Funding Council for Wales requires universities to have a carbon management strategy and target (which is set by the university).

To date, there has been some limited progress in reducing emissions. The English HE sector's emissions peaked in 2009/10 at just over 2.1 million tonnes (CO_2 equivalent), and have since been on a downward trajectory (albeit with a slight rise in 2013/14). Sector emissions are now 1.7 million tonnes, which is 17% lower than the 2005 baseline used for HEFCE targets [17].

59% of English HEIs are currently behind schedule and will likely miss their targets [17]. 14 HEIs (11%) indeed increased their emissions between 2005/06 and 2015/16. Among the Russell Group of 20 research intensive HEIs, which are collectively responsible for half of all HEI emissions, emissions have only fallen by 11% from 2005 to 2015/16. If the current trend continues, the overall reduction from 2005 could be only 23% by 2020 [17]. Emissions from Scottish HEIs also increased by 6% between 2008/09 and 2014/15 [18]. While some improvement in carbon intensity of energy use has been achieved over time (through grid decarbonisation, use of renewables, installation of combined heat and power plants), increases in energy use – generally driven by growth –make it much harder to achieve absolute reductions in carbon emissions. As such, modelling energy demand from the HEIs is an important area of research not only from the cost and resource perspective, but also from the carbon management perspective.

3. Literature on modelling energy demand in HEIs

Although sustainability in the HE sector has attracted some academic interest, there are few studies investigating overall energy use. A large part of this literature focus on carbon emissions (or energy use) from individual institutions, specific buildings or specific interventions in an academic institution (e.g. [19–23]), which is outside our interest in whole sectoral studies. Studies investigating a group of schools (not HEIs) are not uncommon either (see e.g. [24–27]). Nearly all of these studies use energy audit or measurement data of a group of school buildings and use elementary descriptive statistics to explain consumption based on different characteristics. Pereira et al. [28] provide a useful survey of the studies on energy consumption in schools. Although simple correlation analysis has been employed in some studies

[1] 'Scope' defines the different categories of emissions: scope 1 refers to 'direct' emissions from sources owned by the HEI (e.g. fuel use in a boiler); scope 2 refers to 'upstream' emissions generated by purchased electricity by the HEI; Scope 3 refers to emissions resulting from activities of the HEI, but from sources not owned or controlled by the HEI (e.g. business travel by the academics).

to assess the strength of the relationship between various factors affecting energy use in schools, none of this literature attempts to use regression (econometric) methods to quantify the relationship.

In the UK, energy use in and carbon emissions from HEIs has received some attention, especially in the run-up to the Climate Change Act of 2008 and thereafter. Of these Fawcett [29] summarises estimates for emissions, including indirect emissions for business travel by students and staff. Ward et al. [11] use HEFCE data to review energy consumption, using primarily descriptive statistics of various energy indicators over various HEI groups. Distribution of fuel source has also been an area of investigation. Ward et al. [11] test pairwise correlation to understand the explanatory factors for energy consumption in HEIs, which is similar to Sinha et al.'s [30] research into GHG emissions in US universities. Robinson et al. [31] track GHG emissions from the Russell Group of research-intensive universities in the UK, and discuss their achievements against their own carbon targets. Altan [32] qualitatively investigates energy efficiency interventions in HEIs in the UK. Mazhar et al. [33] conduct semi-structured interviews of energy/environment-related managers to qualitatively discuss current management practices toward emissions reduction from the HEI sector. Royston [3] also conducts interviews with energy managers to understand current energy and carbon management practices in HEIs in the UK. Once again, quantitative modelling is missing from this body of work.

Only three studies – Fetcher [4], Klein-Banai and Theis [5] and Wang [6] – utilize econometric/regression modelling techniques to investigate energy use or carbon emissions from HEIs. The first two investigate greenhouse gas emissions and focus on US universities, while Wang [6] focuses on energy consumption from Taiwanese HEIs. Both the US studies use the same GHG inventory database – the American College & University Presidents' Climate Commitment (ACUPCC) reporting system – but from different years. Fetcher [4] uses HEI size and weather (mean temperature in summer and/or winter) to explain carbon emissions (Scope 1 and 2) from 238 US HEIs. He uses gross area and full time equivalent student population to represent size in different model specifications, but never uses both in the same model to avoid multicollinearity. His major finding is that larger HEIs are less carbon-efficient per unit area, especially if the HEI offers a Doctoral degree.

Klein-Banai and Theis [5] improve upon Fetcher's [4] model by dividing gross floor area into different use types (residential, laboratory, health care, etc.) and postulating that different types of use of building space have different carbon implications. They also replace mean temperature with heating degree days (HDD) and cooling degree days (CDD) to represent weather. Using regression models over a sample of 135 HEIs, Klein-Banai and Theis [5] also support Fetcher's [4] finding that as building space increases, Scope 1 and 2 emissions increase at a faster rate. This study also includes regressions for combined Scope 1, 2 and 3 emissions, for both number of occupants and floor space. Fetcher [4] also found that HEIs with medical schools emit more compared to those without, which was also observed by Larsen et al. [19].

Wang [6] also uses cross-sectional data to model energy consumption in 51 Taiwanese universities. He uses three different independent metrics to represent energy use: energy consumption, energy use per unit area per year and energy use per student. The energy consumption model used floor area, land area and building density as explanatory factors. As in the previous studies, floor area is strongly and positively linked with energy consumption. Aranda et al.'s [34] study on annual energy consumption in the Spanish banking sector is also relevant here because of their use of the multiple regression approach. Weather and floor space were found to be the key explanatory factors to predict energy consumption in this study.

While these studies are front-runners in modelling energy use and carbon emissions in the HEI sector, there are several key shortcomings. The statistical models are too simple and the choice of variables are often incomplete or appear entirely ad hoc. More importantly, all three studies on HEIs use cross-sectional observations, i.e. use data from different HEIs at one point in time. Although Klein-Banai and Theis [5] had access to multiple year data for the same HEI, they ignore temporal variations and instead focus on variations between the HEIs by selecting only one observation for each HEI. Such cross-sectional models can explain the variation in carbon emissions (or energy consumption) between the HEIs, but are not appropriate in understanding how emissions (or energy consumption) could increase or decrease within the HEIs in response to changes in the explanatory factors. However, it is the latter question which is more relevant for policymakers or stakeholders. In order to understand the evolution of energy use in the HEI sector, temporal observations are necessary, which is missing from all of these studies. Also, none of these studies include energy price as an explanatory factor, which is well known to affect demand for energy, or any normal good (see e.g. [35,36] for gasoline, [37] for diesel, [38] for natural gas).

4. Data, model and methods

4.1. Data source

The UK Higher Education Statistics Agency (HESA) collects various self-reported statistics from the HEIs in the UK. Although the primary motivation is to collect information on university finance, students and staff, energy consumption data have also been collected since 2001/02 as part of the Estates Management System dataset. The early years (between 2001/02 to 2007/08) of the energy dataset are limited to some extent: e.g. universities in Scotland, Wales or Northern Ireland are not covered and energy use from different fuel types is not available either. However, energy consumption data for non-residential (university business) and residential (student accommodation) use are separated. Fig. 1 presents the distribution of total energy consumption in the HEIs in 2014/15, which shows a large variation in energy consumption among the HEIs.

With a view to monitoring carbon emissions from HEIs, HESA's estates data collection was extended substantially from 2008/09: universities from all four UK nations were included, carbon emissions were also reported and different breakdowns of both energy and carbon emissions were reported. For example, carbon emissions were separated by Scope 1, Scope 2 and Scope 3, as per the World Resource Institute's [39] emissions reporting protocol. In this research we include only energy uses from Scope 1 and Scope 2 categories, emissions from which are under the HEFCE emissions reduction target, and which are less uncertain than Scope 3 emissions.

The number of universities reporting energy consumption in the

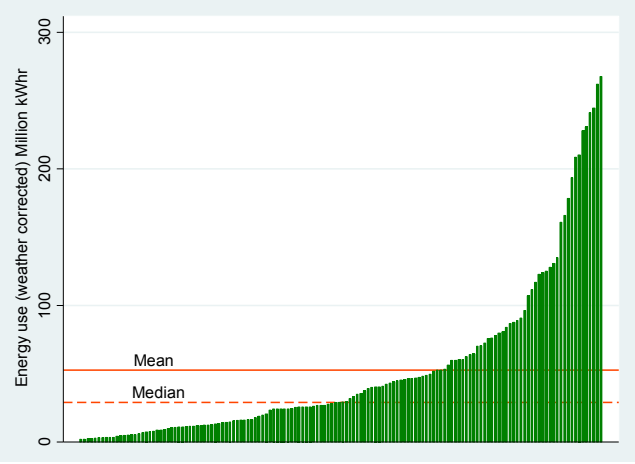

Fig. 1. Distribution of energy use among the HEIs in year 2014, in ascending order of energy use.

Fig. 2. Evolution of energy demand with respect to income for two types of universities.

HESA dataset increased from 112 during 2001/02 to 144 during 2014/15. However given the lack of data on full time equivalent students and staff before 2002/03, which is one of the explanatory factors in the energy model, our final dataset spans 2002/03 to 2014/15. Other relevant explanatory factors collected from the HESA dataset are building gross internal area (GIA) and total income of the HEIs. There are some data gaps in the continuous time series for individual HEIs because of non-reporting, potential misreporting and opening or closure of HEIs during the sample time period. After data cleaning, our final dataset contains 1530 observations from 140 HEIs: a minimum of 2 observations from each HEI, a maximum of 13, and on average 10.9 observations per HEI. To our knowledge this is the largest dataset assembled to understand energy use from HEIs.

Apart from some use in transport (which falls under Scope 3 and not part of this study), energy in HEIs is primarily used in buildings and a large share of building energy consumption is due to space heating during the winter. The energy required for heating a building to a target temperature depends – among other things – on the outside temperature. Since the weather pattern and thus temperature substantially varies temporally and spatially, it is important to control for these differences, which is generally done via the heating degree days (HDD) method. A heating degree day is a measure of how much (in degrees), and for how long (in days), the outside air temperature is below a certain level (www.degreedays.net) so that the building needs space heating. The assumption is that the greater the HDDs, the more artificial space heating will be used, and there is strong evidence that a building's energy consumption is directly linked to HDDs (or cooling degree days in warmer climate, see e.g. [5]). Spatially and temporally disaggregated monthly HDD data is collected from degreedays.net (2016) for eighteen regions within the UK. HDDs are then assigned to each HEI depending on its location and annual observation period, which is then used to normalize respective energy consumptions against weather effects, by dividing energy consumption by respective HDDs and multiplying by the sample average HDD. Fig. 2 presents the evolution of this normalized, weather-corrected energy use with respect to income for all the HEIs in the dataset. Each line represents an HEI, showing the trajectory of its energy use with respect to income over the period. Different colour codes are used for those which are research intensive and those which are not. We define research intensive universities as those with at least 13% of their income from research contracts and grants, based on recent HESA finance statistics.[2]

HEIs consume natural gas directly for space heating, while many HEIs own and operate Combined Heat and Power plants, which also use gas as feedstock; and nearly 30% of UK grid electricity is produced from natural gas, which is the largest share among all feedstocks [40]. As such natural gas prices are used to proxy for energy prices. Real gas price indices for the industrial sector are collected from the Department for Business, Energy and Industrial Strategy [41].[3] Nominal income is converted to real income using the Consumer Price Index (CPI) data from the Office for National Statistics [40].

Data on other characteristics of the HEIs, such as their membership of mission groups such as the Russell Group or the now-disbanded 1994 Group, are collected from relevant mission group websites. The arts/humanities/social science or science focus of a university is determined using the unit of assessments data of HEI submissions for the Research Assessment Framework (REF) in the UK.[4]

4.2. Explanatory factors

Energy consumption in buildings is clearly a function of building characteristics – especially the size of the building, materials used, façade types, insulation and the quality of construction [42]. HEI's energy use and gross internal area (GIA) – representing size – have the highest correlation. However, energy use patterns may vary substantially between different types of building use purposes, e.g. student dormitories, classrooms, laboratories, administrative offices, libraries, sports halls, etc. As such, it would have been preferable to have disaggregated GIA data, by these different uses. Unfortunately, such elaborate differentiation is not available in the HESA dataset. However, GIA is differentiated by residential and non-residential purposes. Therefore we include both residential and non-residential GIA as explanatory factors, instead of a single measure for total gross area.

In the UK, the technical energy efficiency of buildings (energy use per unit floor space) is expressed through energy labels such as the Energy Performance Certificates for residential buildings or Display Energy Certificates for public or commercial buildings.[5] While the most recent HESA dataset has some information on energy performance certificates, the data is incomplete (e.g. the total area under each energy efficiency groups do not equate to the total area of the HEI, and also not available for all years). Therefore it is not possible to include a building energy performance variable in our model. Age could have been a proxy, but different buildings in the same university are built in different eras, and also older buildings may or may not be refurbished to new standards, possibly making the correlation of age with energy use unclear.

Other explanatory factors included in the model are income and population (staff and student number, full time equivalent). The three variables of income, population and GIA are highly correlated and together represent the growth or contraction of HEIs. We also include income squared to account for the possibility of non-linear responses in energy consumption to increases in income.[6] All of these variables vary with time within one HEI, and obviously between HEIs. We also control for other time dependent external factors that could affect sectoral energy use through explicitly including time in the specification. The price of energy is included to account for the potential negative effect of

[2] 13% is the lowest ratio of research income to total income for the ex-G94 HEIs.

[3] Note that energy prices paid by the HEIs can often be negotiated with the utilities by the individual HEIs, and therefore could differ between them in practice. Such data, however is not available and our price data therefore varies only temporally, but not between the HEIs. We do not expect the within-effect estimates to be affected significantly, though.

[4] While HEIs can be quite strategic in their REF submissions and omit reporting weak departments, leading to some bias in this variable, this is the best data we have access to.

[5] It is important to note that the term "energy efficiency" is used in this paper to express a ratio of energy use to another variable (specifically: area, population or income). This is not intended to endorse prevalent discourses that promote "efficiency" as the sole or main route to reducing energy use or carbon emissions (see [53]).

[6] i.e. when the growth *rate* with respect to income depends on the income.

energy price on energy demand.[7]

In addition, we have information on the characteristics of HEIs which are time invariant, i.e. they are fixed over time for each HEI. It is quite possible that research focused, especially science and medical research focused HEIs, would consume more energy compared to a similar sized non-research intensive university and inclusion of such characteristics in the model improves the explanatory power further. Because of multicollinearity, however, we cannot include all of these variables at the same time (e.g. all Russell Group universities are research intensive). As such we test several combinations in our regression model. Energy use in HEIs thus has the following specification:

$$E = f(GIA_NR, GIA_R, INC, INCSQ, OCC, HDD, PRC, T) \quad (1)$$

where GIA_NR = Gross internal area – non-residential
GIA_R = Gross internal area – residential
INC = Income
$INCSQ$ = Income squared
OCC = Number of occupants (staff + students, full time equivalent)
HDD = Heating degree days (or, energy corrected for HDD)
PRC = Price of natural gas
T = Time (continuous or dummy)

All the continuous variables are converted to logarithms. This has two advantages: firstly, it reduces the potential heteroscedasticity problem (i.e. variance increases with larger values) and secondly, the parameter estimates directly represent the elasticities of energy use with respect to the explanatory factors. Table 1 presents the summary statistics for the key variables in the model.

4.3. Panel econometric method

Our dataset is known as panel data in the literature, whereby the variables are observed for cross-sectional units – the HEIs – over several time periods. While it is possible to pool all the data together and estimate the model parameters through Ordinary Least Squares (OLS) regression technique, such a 'pooled' OLS approach assumes that each observation is independent from each other, which is clearly not the case in a panel dataset. As such a simple OLS method can result in erroneous estimation of the parameters.

On the other hand, panel econometric techniques recognize the special structure of the data and can lead to more efficient estimation of the parameters. Especially, it is never possible to include all the explanatory variables in a typical regression model, and the absence of some variables could lead to an omitted variable bias. Panel regression can control for these unobservable factors and also recognizes that the HEIs are heterogeneous units that can differ from each other [43,44]. This is certainly a more plausible representation of reality than assuming all observations are similar, which is the implicit assumption of the pooled model.

The basic framework in a *hypothetical* one-way panel data model with one time varying and one time-invariant explanatory factor is:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \alpha_i + \varepsilon_{it} \quad (2)$$

where y_{it} = dependent variable
x_{it} = time variant explanatory factor
z_i = time invariant explanatory factor
α_i = intercept for cross sectional units
i = subscript for cross-sectional unit
t = subscript for time unit

[7] Another variable that could affect consumption is the length of academic year which may vary between 9 and 12 months. Unfortunately we do not have any data on that. Also, the effect will likely be minimal, since our main interest is in the temporal dimension and HEIs do not significantly alter their academic term durations between different years.

Table 1
Summary statistics for estimation sample.

Variable	Mean	Std. Dev.	Minimum	Maximum
Energy ('000 kW h)	54,000	52,200	1990	294,000
Energy-HDD normalized ('000 kW h)	55,300	53,000	2013	294,000
Nonresidential area (m^2)	136,781	115,975	3881	687,089
Residential area (m^2)	48,326	42,118	226	249,433
Nominal income ('000 GBP)	186,916	172,977	2412	1,429,389
Occupants	14,245	8662	504	45,922
Real gas price index	104.77	22.37	58.7	134.0
Number of Russell Group universities	23			
Number of ex-G94 universities	15			
Number of research intensive HEIs	62			
Number of medical research HEIs only	3			

There are several ways a panel regression model can be estimated. A 'between-effect' (BE) model is defined as one using the within group (within each cross-sectional unit) mean of the variables:

$$\bar{y}_i = \beta_0 + \beta_1 \bar{x}_i + \beta_2 z_i + \alpha_i + \bar{\varepsilon}_i \quad (3)$$

The BE parameters can explain the differences between the cross-sectional units, but lose the potentially rich information in the time dimension by reducing them to averages only. The parameter estimates can also be biased [45]. Still, the model can be useful in explaining whether there are any systematic differences in energy consumption between, say, research-intensive and teaching-intensive HEIs or high income and low income HEIs. The cross-sectional models used in [4–6] are all – in essence – BE models.

A fixed-effects (FE) model produces 'within-effect' estimates for parameters and is estimated by deducting Eq. (3) from Eq. (2) and running OLS on the transformed dataset:

$$(y_{it} - \bar{y}_i) = \beta_1 (x_{it} - \bar{x}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (4)$$

In practice, the FE or within-effect parameters represent how the dependent variable in the cross-sectional units – on average – responds to changes in the independent variables in the temporal dimension. This is also equivalent to estimating Eq. (2) by introducing dummy variables for each cross-sectional unit, and as such assuming that the α_is represent the effects of all unobservable characteristics of the cross-sectional units. The biggest advantage of the FE model is that it allows the fixed effects (α_is) to be correlated with the explanatory factors, which is often the case in practice and results in an unbiased estimate. The primary disadvantage is that the effects of time-invariant characteristics (e.g. research intensive vs. teaching intensive HEIs) cannot be determined as they are subsumed within the α_is. However, this is still the more popular method, given the researchers' interest is often in the effects of a 'change' which is time-dependent and the appropriate interpretation of the parameter estimates in this context. Since our primary interest lies in the effects of change in the explanatory factors, this is also the primary model we are interested in.

A random effect (RE) model, on the other hand, assumes that the α_is are uncorrelated with the explanatory factors and randomly distributed (with zero mean and constant variance). In essence, RE parameter estimates are weighted averages of between-effect and within-effect estimators and Generalized Least Squares (GLS) or Maximum Likelihood (ML) methods are employed for estimation. If the assumptions hold, RE model estimates are more efficient than FE ones. RE formulation also allows modelling the effects of time-invariant characteristics, but any misspecification has a more serious consequence compared to FE models [44].

There is a large literature on the choice between FE and RE models. FE models are often described as the standard default in panel data modelling, especially in the field of economics [46]. Subjectively, FE

models are used when the inference is conditional on the sample, while RE are used when inference about population from a sample of observations is required [47]. Interpretability of the parameters and robustness against misspecification (due to omitted variables, which is quite common) generally tilt the balance toward FE. Objective measures such as the Hausman test [48] are often used to guide the choice between FE or RE models.

4.4. Hybrid panel method

Mundlak [49] adds to the discussion by arguing that the choice is not between FE and RE models, but rather on how far the assumption of zero correlation between α_is and the explanatory factors can be relaxed and modelled within the RE framework. Mundlak [49] proposed a correlated random effect model, where he assumes that $\alpha_i = \pi\bar{x}_i + \nu_i$ and Eq. (2) becomes:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \pi\bar{x}_i + \nu_i + \varepsilon_{it} \quad (5)$$

In this formulation, the group mean $\bar{x}_i$ picks up the correlation between the variable and α_i, and ν_i becomes the error uncorrelated with the x_{it}s. The parameter estimate for x_{it} is the same as the within or FE estimator. Similar to Mundlak's [49] approach is the 'hybrid' model [7], which decomposes x_{it} into a between ($\bar{x}_i$) and a within effect ($x_{it} - \bar{x}_i$) in Eq. (2), again within the RE framework. The hybrid model thus takes the form:

$$y_{it} = \beta_0 + \beta_1(x_{it} - \bar{x}_i) + \beta_2 z_i + \beta_3\bar{x}_i + \mu_i + \varepsilon_{it} \quad (6)$$

As in a typical RE model, μ_i is randomly distributed and uncorrelated with the explanatory factors, any correlation being picked up by the inclusion of $\bar{x}_i$ in the model. In this formulation β_1 produces the same FE or within-effect as Eq. (5), while β_3 is the between-effect. The hybrid formulation allows us to test whether the within and between effects are the same, i.e. whether the effect of changes in one variable within a cross-sectional unit is the same as the effect of that variable on different cross-sectional units. More importantly it allows estimating the correct FE or within-effect parameter – which is the main interest – but also allows deciphering the effects of non-time varying characteristics (z_i) of the HEIs, although that can be measured with some bias. Hybrid models are becoming popular because of their flexibility and this will be the second model structure that we estimate. Bell and Jones [46] suggest that hybrid models should be preferred to FE models.

4.5. Econometric model specification

The final econometric specification of the model for HDD-corrected energy consumption from the HEIs is:

$$\begin{aligned} lnE_{it} = {} & \beta_0 + \beta_{GIANR} lnGIA_NR_{it} + \beta_{GIAR} lnGIA_R_{it} + \beta_{INC} lnINC_{it} \\ & + \beta_{INCSQ} lnINCSQ_{it} + \beta_{OCC} lnOCC_{it} + \beta_{PRC} lnPRC_{it} + \beta_R Res_i \\ & + \beta_{MR} MedRes_i + \beta_{Rus} Russell_i + \beta_{G94} ExG94_i + \sum \beta_t T_t + \alpha_i + \varepsilon_{it} \end{aligned} \quad (7)$$

Both the FE and hybrid models are estimated for this functional specification. For the FE models, the parameter estimates for time-invariant variables automatically vanishes. For the hybrid formulation, the continuous variables are decomposed into group means and deviations from group means.

5. Results and discussion

5.1. Choice of fixed-effects model

Table 2 presents the results of the fixed-effects (within-effect) model, our main interest, for a number of different specifications. The first step, however, is to check whether we need to apply panel econometric techniques or whether a pooled model is sufficient. This can be done by conducting an F-test on whether all α_is are equal to zero (which would mean there are no significant HEI specific effect). For our primary Model FE1, $F(139, 1373) = 27.84$ ($p < 0.001$), which rejects the null that the α_is are zero at 99% confidence, clearly indicating the superiority of FE model, and as such panel econometric techniques. Hausman test for RE vs. FE also indicates a preference for the FE model ($\chi^2(17) = 237.7$, $p < 0.001$).

Within the FE models, several model specifications are tested in order to select the best-possible one. Our first interest is the choice of dependent variable: energy normalized using HDD (Model FE1 and FE2), or uncorrected energy as a dependent variable with HDD included as an explanatory factor (Model FE3). Clearly Models FE1 and FE2, which utilize normalized energy have significantly better fit (within-R^2 0.413 vs. 0.313). As such all of the other models follow normalized energy as the dependent variable.[8]

Inclusion of energy prices requires some discussion. Energy prices are likely measured with some error (since we use one energy price for all HEIs, which varies only temporally) and inclusion of them in the model does not improve the model fit (Model FE1 vs. Model FE2). However, the parameter estimate for price becomes statistically significant, and also improves the significance of a few of the time fixed effects. Given the theoretical importance, we keep price in our model and also the time dummies to control for other time-dependent changes. As such Model FE1 is our preferred specification for the fixed effect model.[9] Fig. 3 presents the model predictions with actual values for model FE1, and shows good visual fit of the model.

5.2. Choice of hybrid model

Technically, the hybrid model is an RE model, and we have already determined the appropriateness of FE over RE using statistical tests. Nonetheless, we run the hybrid formulation of the RE model to understand the effects of time-invariant characteristics. The choice of appropriate time-invariant characteristics such as membership in mission groups (e.g. Russell Group), research intensiveness (yes/no), focus on arts and social sciences or presence of a medical school is made via model fit performance. Given Russell Group and ex-G94 universities are all research intensive, and nearly all HEIs offering medical degrees are also research intensive, there is a very high correlation among these variables. There is also a negative correlation between a focus on arts/social sciences and research intensity, albeit not a strong one. The preferred hybrid model HY1 includes indicator (dummy) variables for membership in the Russell Group, in the ex-G94 Group, and research intensiveness as explanatory factors for between-effect differences. Also included is an indicator variable for three HEIs which are purely for medical research. Model HY2 is not too different, where the variables for Russell and ex-94 Group memberships are dropped, and is marginally inferior to Model HY1. Other specifications with characteristics such as presence of a medical school, and primary focus on arts or humanities subjects, do not improve model fit at all, because of the multicollinearity between the HEI characteristics.

Table 3 presents the results of the hybrid model. As discussed earlier, the parameter estimates on the mean-differenced variables are the within-effects. We indeed find that these estimates are the same as the FE parameters (Model FE1) above. The means of respective variables are time invariant and these parameter estimates represent between-

[8] Note that Model FE3 shows that an increase in HDD increases energy consumption only at 85% confidence level (parameter estimate 0.246). The somewhat weak significance is because in a fixed effect model only the temporal changes in HDD come into play, which becomes correlated with individual time effects.

[9] We have also estimated a model without taking logarithms of the continuous variables, but that model performs poorly.

Table 2
Parameter estimates for fixed-effect models.

Model	FE1		FE2		FE3	
Dependent variable	Log of weather-corrected energy				Log of uncorrected energy	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
lnGIA_NR	0.322***	0.101	0.322***	0.101	0.320***	0.099
lnGIA_R	0.120***	0.021	0.120***	0.021	0.118***	0.020
lnINC	−0.861***	0.304	−0.861***	0.304	−0.834***	0.299
lnINCSQ	0.046***	0.016	0.046***	0.016	0.045***	0.015
lnOCC	−0.032	0.055	−0.032	0.055	−0.029	0.054
lnPRC	−0.208***	0.068	.	.	−0.209***	0.067
lnHDD	.	.	.	.	0.246	0.171
Time effects:						
2003/04	−0.016	0.016	−0.025	0.018	0.017	0.018
2004/05	0.080***	0.015	0.021	0.023	0.073***	0.015
2005/06	0.076***	0.023	−0.037	0.024	0.109***	0.023
2006/07	0.200***	0.017	0.091***	0.029	0.075**	0.034
2007/08	0.074***	0.018	−0.051	0.034	0.072***	0.017
2008/09	0.025	0.019	−0.118***	0.038	0.076***	0.022
2009/10	−0.031**	0.013	−0.149***	0.039	0.043*	0.022
2010/11	−0.049***	0.014	−0.180***	0.040	0.024	0.022
2011/12	0.074***	0.019	−0.082**	0.039	0.052***	0.019
2012/13	−0.041**	0.018	−0.213***	0.043	0.084***	0.033
2013/14	0.116***	0.014	−0.052	0.045	0.039*	0.023
2014/15	.	.	−0.144***	0.047	.	.
Constant	17.401***	2.161	16.553***	2.024	15.358***	2.338
No of HEIs	140		140		140	
No of observations	1530		1530		1530	
Within R^2	0.413		0.413		0.313	

Statistically significant at ***99%, **95%, *90% confidence level.

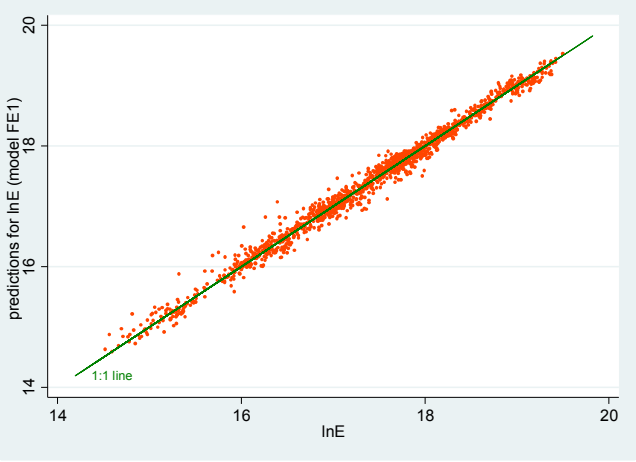

Fig. 3. Model prediction vs. actual values for model FE1.

effects of Eq. (3) earlier. Although they appear numerically different, sometimes substantially, from the within-effect parameters, pairwise statistical tests (e.g. *ΔlnGIA_NR* vs *Mean(lnGIA_NR)*) show the differences are not statistically significant for any of these, which is an alternate confirmation for the appropriateness of the FE models.

We also run a hybrid model for non-residential energy consumption (Table 3, Model HY3-NR), bearing in mind that this model directly provides the within-effect parameters, too. This model is the same as others, with residential gross area (GIA-R) missing from the explanatory factors. These results are presented in Table 3 too.

5.3. Effect estimates

The fixed effect model for temperature corrected energy use show that – within HEIs, over time – energy use increases with an increase in floor area, but the response is different with respect to residential and non-residential floor areas (Table 2, FE1). On average over all the HEIs, a 10% increase in non-residential floor space increases energy consumption by 3.2%, but a similar increase in residential floor area increases consumption by only 1.2%. This difference is expected given the multitude of energy-intensive uses of non-residential buildings (e.g. computing, heating of large spaces, research and teaching equipment, etc.). The less than proportional increase in energy consumption with respect to floor areas may indicate improved technical performance of the new buildings and/or the facilities therein (e.g. computers).

The effect of floor area is numerically slightly different when we consider within-HEI (temporal) and between-HEI effects. A 10% difference in mean non-residential floor area (i.e. the mean for each HEI over the years) is responsible for a 5.3% difference in energy use between different HEIs, while for mean residential space the difference is even less, at 1.8% (Table 3, HY1) – both of these effects are slightly larger than the respective within-HEI temporal effects (although not statistically significant). More importantly, both the between- and within-effect results point to the same qualitative conclusion that HEIs with larger floor space are more 'energy efficient' (use less energy per unit floor area) than those with smaller floor areas.

Note that the between-effect part of the hybrid model is conceptually similar to the cross-sectional models of carbon emissions from HEIs in the US by Fetcher [4] and Klein-Banai and Theis [5]. However both of them find that carbon emissions increase *more than proportionally* with respect to floor area. While recognising that carbon emissions and energy consumption are two different variables, this difference is worth considering. The potential discrepancy with our results arises most likely because the previous models did not include other correlated variables such as income or student and staff number. As such the floor space variables in those studies picked up the effects of not only floor area, but also income and number of occupants. As we explain in the next section, when we combine the different growth variables (floor area, income, occupants), we do not observe economies of scale in the cross-sectional analysis, supporting our explanation.

The effect of income on energy consumption requires further calculation because of the presence of the quadratic term in the model. The net effect of changes in income in the within-effect model is given by $\beta_{INC} + 2.\ \beta_{INCSQ}.\ lnINC$. Given the positive parameter estimate of

Table 3
Parameter estimates for hybrid models.

Model	HY1		HY2		HY3-NR	
Dependent variable	Log of Weather corrected energy, all				Log of Weather-corrected energy, non-residential	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
ΔlnGIA_NR	0.318***	0.103	0.319***	0.103	0.473***	0.128
ΔlnGIA_R	0.121***	0.021	0.121***	0.021		
ΔlnINC	−0.862***	0.306	−0.861***	0.306	−0.865***	0.313
ΔlnINCSQ	0.046***	0.016	0.046***	0.016	0.049***	0.015
ΔlnOCC	−0.034	0.056	−0.034	0.056	−0.017	0.084
ΔlnPRC	−0.211***	0.068	−0.211***	0.068	−0.166***	0.064
Mean(lnGIA_NR)	0.525***	0.079	0.510***	0.076	0.753***	0.000
Mean(lnGIA_R)	0.184***	0.028	0.187***	0.026		
Mean(lnINC)	−0.901*	0.556	−0.702*	0.438	−0.401	0.510
Mean(lnINCSQ)	0.050**	0.024	0.041**	0.018	0.030	0.023
Mean(lnOCC)	0.072	0.077	0.082	0.076	−0.020	0.076
Mean(lnPRC)	.	.	.	.	.	.
MedRes	1.026***	0.169	1.012***	0.168	0.464**	0.242
Univres	0.207***	0.068	0.233***	0.059	0.200***	0.061
Russell	−0.032	0.083	.	.	−0.009	0.085
ExG94	0.075	0.069	.	.	0.006	0.068
Time effects:						
2003/04	−0.016	0.016	−0.016	0.016	−0.003	0.017
2004/05	0.080***	0.015	0.080***	0.015	0.086***	0.016
2005/06	0.077***	0.023	0.077***	0.023	0.079***	0.023
2006/07	0.200***	0.017	0.200***	0.017	0.201***	0.019
2007/08	0.075***	0.018	0.075***	0.018	0.082***	0.018
2008/09	0.025	0.019	0.025	0.019	0.031	0.021
2009/10	−0.031**	0.013	−0.031**	0.013	−0.016	0.016
2010/11	−0.050***	0.015	−0.050***	0.015	−0.046***	0.018
2011/12	0.074***	0.019	0.074***	0.019	0.080***	0.021
2012/13	−0.041**	0.018	−0.041**	0.018	−0.031	0.020
2013/14	0.117***	0.014	0.117***	0.014	0.123***	0.015
2014/15	.	.	.	.	.	.
Constant	12.335***	3.162	11.309***	2.511	8.998***	2.950
No of HEIs	140		140		157	
No of observations	1530		1530		1789	
Within R^2	0.413		0.409		0.350	
Between R^2	0.972		0.971		0.973	
Overall R^2	0.962		0.962		0.963	

Statistically significant at ***99%, **95%, *90% confidence level

β_{INCSQ}, this indicates that the effect of income on energy consumption increases at larger incomes. At mean income of our estimation sample, a 1% change in income results in a 0.22% (std error 0.073) increase in energy consumption within the HEIs in the temporal dimension (Table 2, FE1). The parameter estimates for income in the between-effect part of the model are not too different from the within-effect part.

After controlling for income and floor space, the number of students and staff does not have a statistically significant within-effect on energy use in HEIs (Table 2, FE1). This is possibly not surprising as the student numbers rarely change substantially without concurrent changes in the physical size or income of the HEIs and the variables are highly correlated.[10] Among published studies, Fetcher [4] reports an increase in carbon emissions with an increase in the number of students, but again, this specification does not control for the fact that universities with larger student populations are also larger in physical size. Also, Sekki et al. [24] looked into individual building energy consumption data from schools and found that the connection between energy consumption and occupancy was *not* strong.

Our main interest from the between-effect parameters of the hybrid model is in the impacts of the time-invariant characteristics of the HEIs, which cannot be determined from the FE or within-effect parameters. The parameter indicating whether an HEI is research intensive or not is positive and statistically significant at 99% confidence. This indicates that after controlling for income and size (which are generally large for research intensive HEIs), research intensive HEIs still consume 23% more energy compared to other HEIs.[11] Clearly research activities – especially in the science, medicine and engineering domain – are often more energy intensive given the use of information technology services and other specialist energy-intensive equipment. After controlling for the research intensiveness, members of Russell or ex-G94 groups do not consume significantly more or less energy. The three medical research focused HEIs (*MedRes*) consume substantially more energy than their sizes would suggest. Most medical equipment is very energy intensive and building specific carbon footprint data shows that buildings housing medical departments were indeed the highest carbon emitters (and as such energy consumers) on a per unit area basis in a Norwegian university [19].[12]

For non-residential energy consumption (Model HY3-NR), the findings are similar: an increase in floor area and income increase non-residential energy consumption, increase in gas price reduces energy consumption. The effect of non-residential floor area on non-residential energy consumption is larger than the effect of non-residential floor area on overall energy consumption: this is expected given the variables

[10] The statistically insignificant parameter estimate for occupant number is likely a result of multicollinearity, which is a known issue in this type of energy modelling, e.g. see Huebner et al. [52] for domestic electricity consumption. Our model fits better with all three growth variables, compared to only floor area and income.

[11] $e^{0.207} - 1 = 0.23$

[12] We have also tested the effects of science-intensive universities, but given they are generally more research intensive, they return insignificant effect estimate.

now are more directly related.

5.4. Economies of scale

The parameter estimates above relate to a ceteris paribus condition, i.e. the parameters represent the effect of one variable keeping others constant. In reality, it is highly unlikely that one (e.g. income) will grow without an increase in another (e.g. the number of students). Given the recent growth and future growth strategy of most HEIs, it is important to understand the effect of future overall growth of HEIs on energy consumption. A relevant question is whether there are economies of scale in energy consumption, i.e. whether the HEIs become more 'energy efficient' as they grow over time. Given we have three variables that capture growth – namely floor area, income and number of occupants – we define economies of scale as the effect on energy consumption due to an increase in all of these factors simultaneously. It is important to consider the effect of these variables simultaneously because considering the effect of only one (as in parameters estimated in the previous section) would not represent overall growth correctly. This is equivalent to calculating $\beta_{GIANR} + \beta_{GIAR} + \beta_{INC} + 2.\ \beta_{INCSQ}.\ lnINC + \beta_{OCC}$. The combined effect of increasing these variables each by 1% is an increase in energy consumption of 0.63% (std error 0.094). This indicates that energy consumption grows less than proportionally as the HEIs grow over time from their current position, i.e. there is an economies of scale effect in energy consumption in the temporal dimension. This may be because newer buildings (and the facilities and equipment within them) tend to be more 'energy efficient' than older ones.

The economies of scale *within* the HEIs in the temporal dimension does not necessarily mean that larger and wealthier universities (in terms of floor area, income and population) at the same point in time are less energy intensive compared to smaller ones. The between-effect part of the hybrid model reveals that a simultaneous 1% difference in all of the three growth indicators above results in a 1.05% (std. error 0.031) difference in energy consumption between the HEIs, indicating the absence of an economies of scale effect between the HEIs at a specific point in time.[13]

The different findings in the temporal and cross-sectional dimensions show the usefulness of the hybrid models over the traditional cross-sectional models used before. The apparent contradiction between the economies of scale within HEIs as they grow over time and the lack thereof between HEIs at a specific time requires some consideration. From a mathematical point of view, it is driven by the larger estimates for each of the relevant between-effects compared to within-effects and opposing estimates for occupancy in Table 3. The underlying explanation may be linked with the temporalities of change in the sector, and the limitations of our dataset. As universities grow over time, they are likely to add energy-intensive new developments such as ancillary student facilities (e.g. sports halls, gymnasiums, swimming pools – which are relatively energy-intensive), but this may happen as a one-off event after the growth occurs, and not as a continuous process. Such 'structural changes' would have reduced the economies of scale, but might not be captured by the temporal dimension of our dataset, which only shows change that has occurred over the last decade – a relatively short and recent period of time. On the other hand, many universities in our dataset have existed for decades or even centuries and the cross-sectional analysis shows the present-day outcomes of their historical trajectories. The universities that are now relatively large and wealthy may have undergone this structural change process in the past, in response to growth; and as such we do not observe economies of scale when looking across the sector. In general, these large wealthy institutions are also likely to be older, with associated infrastructure that could be less 'energy efficient' to begin with. Note however, for practical forecasting purposes, it is the changes in the temporal dimension that are useful, not the between-effect estimates. And this is where our estimates are especially more reliable and useful, given all previous models reported only between-effect estimates.

5.5. Role of energy prices & HEFCE targets

The price of gas – which proxies for energy prices in general – is associated with a small but negative effect on energy consumption in the HEIs (Table 2, FE1). This is a typical expectation from economic theory. A 10% increase in industrial gas prices reduces energy consumption by 2%, indicating energy use is relatively inelastic with respect to prices. Even the small price response can have an important role in understanding the recent reduction in carbon emissions in HEIs in England. BriteGreen [17] reports a 17% reduction in carbon emissions between 2005/06 and 2015/16. Given that real energy prices almost doubled during that period, at least a part of that emission reduction could be associated with the demand response to energy price increases. Fig. 4 compares the predictions from model FE1 with actual prices during that period with predictions assuming prices were the same as in 2002 (only a few HEIs are shown for the purpose of clarity). Clearly, energy use would have been higher (solid lines) had energy prices not gone up during the observation period. As such BriteGreen's [17] projection that the English HEIs are set to reduce 23% of their emissions by 2020 (over 2005 emissions) may be optimistic if energy prices remain low in future. On the other hand, if prices go up in the future, emissions will likely be reduced further than projected.

Given the use of temporal fixed effects (the yearly dummy variables) in our model, it is not possible to ascertain whether the HEFCE carbon targets of 2010 had any role in reducing energy use.[14] However, it is possible for the HEIs to reduce their carbon emissions without taking any measures at all since the carbon intensity of grid electricity in the UK has been declining over time. Similarly, reduction in carbon emissions is possible despite an increase in energy use if the carbon intensity of energy is reduced through greater use of renewable sources. As such the observed reduction in carbon emissions by BriteGreen [17] is not at odds with our findings of an increase in energy use.

5.6. Projections into the future

The model parameters of HE1 can be used to simulate future energy use from the higher education sector in the UK (similar to [50] for supermarkets). However such simulation is only as good as the underlying forecasts of the explanatory factors, which can themselves be quite uncertain. Therefore, instead of making forecasts of future energy use, we simulate future energy use in a hypothetical scenario of 4% growth in income per year for every HEI (*average* growth was 5.2% for our observations, we assume a conservative growth rate to reflect Brexit related uncertainties in the higher education sector). Student and staff number and building floor area, however, do not grow as quickly as income. As such we use a similar ratio of growth for these explanatory factors (Table 4). Given the uncertainty of future gas prices, they are kept fixed at the same value as in 2014. Fig. 5 presents the results from the simulation, which runs until 2027 and shows that energy consumption from the HEIs will continue to increase in the assumed conservative growth scenario if there are no external policy changes.

6. Conclusions

This paper set out to understand how changes in the HE sector can affect universities' energy consumption by using the UK as a case study

[13] Note also that the larger and wealthier HEIs are nearly always research intensive – which also results in higher energy consumption.

[14] The dummy variable for post-HEFCE target period will be perfectly collinear with the time fixed effects during that period.

Fig. 4. Simulated effect of price on energy use in selected HEIs (within estimation sample).

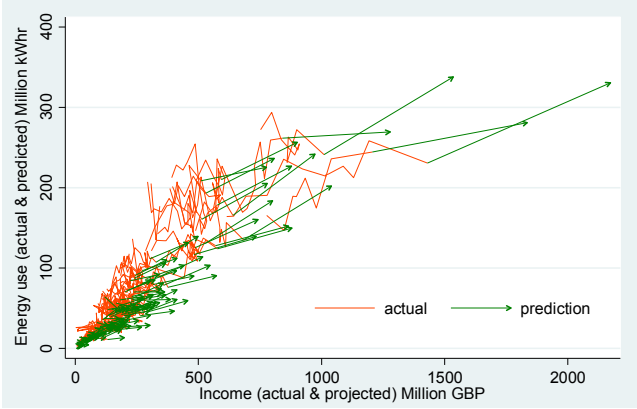

Fig. 5. Example simulation of future energy use from the HEIs.

country. We discussed the limitations of existing studies that use cross-sectional regression models and argued that temporal information and 'within-effect' panel regression techniques are more suited to understanding the factors affecting energy use from HEIs and to answering the relevant policy questions – especially since researchers and policymakers are often interested in the evolution of energy use in the temporal dimension. The traditional limitations of FE models are resolved by applying a hybrid modelling technique to determine both 'within' and 'between' effects from the same model. Results show that the physical size of HEIs (building floor area) and income affect energy use directly, although student and staff numbers do not have a statistically significant effect due to multicollinearity. The effects of changes in non-residential and residential floor area are also different, reflecting different energy consumption patterns in different types of buildings. Increasing energy prices reduce energy use, a finding that indicates that the relatively slow growth in energy use and reduction in carbon emissions in recent years could be – at least partially – a result of increases in energy prices during that period. A fall in energy prices in future could therefore jeopardize reductions in energy use or carbon emissions from the HEI sector.

We observe some economies of scale in energy consumption in the temporal dimension: as HEIs grow over time, they become more 'energy efficient' (use less energy per unit of area, population or income). While testing of economies of scale is itself novel in this sector, this finding also has important implications. Unlike this work, all of the previous studies used cross-sectional data and reported scale 'diseconomies', and any forecasting or prediction exercise using those earlier studies would substantially overestimate energy consumption (or carbon emissions, as the majority of those studies investigate). The hybrid model also shows that research-intensive HEIs generally consume more energy compared to teaching-focused HEIs even after controlling for differences in income, floor area and student and staff population. Although these results are UK specific, the general direction of the effects of different variables will likely hold for other countries, too, especially for those where the higher education sector is still growing. This is certainly the case in many emerging and developing countries.

We stop short of providing any forecast for potential future energy use from HEIs given the uncertainties in the forecasts of other input variables in future. However, we do show how our model parameters can be used to predict energy use under an imagined future scenario. By varying the explanatory variables, the model can be used to explore other scenarios too; for example, a rise in fuel prices, a shift from research to teaching, or a drop in income or size of the universities. Clearly, the growth path of the HEIs and prices of energy are both quite uncertain and that uncertainty will be carried into such simulated scenarios, but that is a necessary evil in all forecasting models using multiple regression, as in here. The modest growth scenario (in income, built area and student and staff numbers) tested here suggests that energy use will continue to grow in future unless there is significant change in the policies currently driving growth in the HE sector in the UK.

The panel econometric modelling brings a step change in understanding how energy use in the HE sector evolves with changes in the underlying explanatory factors. The underlying panel econometric method can also be applied to other countries where such panel data is available. Similarly the method can be applied to carbon emissions if the underlying emissions data are available. However, there is still room for improvement. Incorporation of reliable and comprehensive data on building energy efficiency would further improve the current model. Most of the explanatory factors are highly correlated and modelling each of the explanatory factors as a function of other explanatory factors (e.g. space as a function of students and staff number) – possibly using a structural equation or energy decomposition framework (e.g. [51]) – could improve the understanding of the pathways to changes in energy use. Stochastic frontier methods could also be applied to benchmark the HEIs against those performing best on energy consumption. Such benchmarking results would be beneficial to individual HEI energy or sustainability managers.

Finally, by showing that universities' income, size and research intensity are key influences on their energy consumption, this paper helps to build a picture of how non-energy policies (often invisible to policymakers and researchers) can contribute to escalating demand. The variations and changes in income, size and research intensity that are analysed here are not natural or inevitable, but are shaped and steered (among other things) by the policies of individual institutions and of wider cross-sectoral and national decision-makers. To date, these

Table 4
Changes in income, floor area, and population of HEIs and assumed future scenario.

	2003	2014	Annual growth rate, %	Assumed growth rate for future, %
Income ('000 GBP)	146,990	238,870	5.21	4.0
Nonresidential area (m^2)	132,005	153,795	1.38	1.25
Residential area (m^2)	47,281	50,862	0.63	0.6
Occupants	12,954	14,625	1.07	0.8
Energy-HDD normalized ('000 kW h)	58,900	60,000	0.156	–

policies have mostly been (often unintentionally) driving energy demand in an upward direction. By revealing the role of institutional changes in escalating energy demand, the analysis here also supports the argument of Royston [54] for a new research and policy agenda that takes seriously the impact of invisible energy policies on energy demand.

Acknowledgements

This research was partially funded by the Engineering and Physical Sciences Research Council (grant numbers EP/K011723/1) as part of the RCUK Energy Programme, and by EDF as part of the R&D ECLEER Programme. We thank Eduardo Henriques Dos Santos for processing the HESA dataset, to www.degreedays.net for kindly sharing the HDD data and to Phil Wheat for stimulating discussions. Thanks are also extended to the three anonymous reviewers for their constructive suggestions.

References

[1] Lambert R. Lambert Review of Business-University Collaboration. London: HM Treasury; 2003.
[2] HEFCE. Sustainable development in higher education: consultation on a framework for HEFCE; 2013.
[3] Royston S. Invisible energy policy in Higher Education. Presented at DEMAND conference, Lancaster, UK, 13–15 April 2016. 2016.
[4] Fetcher N. Effects of climate and institution size in greenhouse gas emissions from colleges and universities in the United States. Sustainability 2009;2(6):362–7.
[5] Klein-Banai C, Theis TL. Quantitative analysis of factor affecting greenhouse gas emissions at institutions of higher education. J Cleaner Prod 2013;48:29–38.
[6] Wang JC. A study on the energy performance of school buildings in Taiwan. Energy Build 2016;133:810–22.
[7] Allison PD. Fixed effects regression models. Thousand Oaks, CA: Sage; 2009.
[8] Universities UK. Higher education in numbers; 2018. Available: <http://www.universitiesuk.ac.uk/facts-and-stats/Pages/higher-education-data.aspx> [accessed: 13 April 2018].
[9] Universities UK. The impact of Universities on the UK economy. London: Universities UK; 2014.
[10] HESA. Higher education provider data; 2018. Available at: <https://www.hesa.ac.uk/data-and-analysis/providers/estates/table-1> [Accessed October].
[11] Ward I, Ogbonna A, Altan H. Sector review of UK higher education energy consumption. Energy Policy 2008;36:2939–49.
[12] Universities UK 2017. Patterns and Trends in UK Higher Education; 2017. Available: <http://www.universitiesuk.ac.uk/facts-and-stats/data-and-analysis/Pages/patterns-and-trends-2017.aspx> [accessed: 13 April 2018].
[13] Department for Business, Energy and industrial Strategy. Energy consumption in the UK; 2016. Available: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/573269/ECUK_November_2016.pdf> [accessed: April 2017].
[14] HEFCE. Carbon reduction target and strategy for higher education in England. Bristol: Higher Education Funding Council for England; 2010.
[15] HEFCE. Reducing carbon emissions; 2017. Available: <http://www.hefce.ac.uk/workprovide/carbon/> [accessed: March 2017].
[16] Environmental Association for Universities and Colleges, EAUC. Universities and colleges climate commitment for Scotland; 2018. Available: <http://www.eauc.org.uk/universities_and_colleges_climate_commitment_fo2> [accessed April 2018].
[17] BriteGreen. University of Sussex University Carbon Report for the Academic Year 2015/16; 2017.
[18] Carbon Forecast. Summary report on Scottish further and higher education carbon emissions 2014-2015; 2016. Available: <http://www.sustainabilityexchange.ac.uk/files/eauc_fhe_report_v2_-_key_findings_summary.pdf> [accessed: March 2017].
[19] Larsen HN, Pettersen J, Solli C, Hertwich EG. Investigating the carbon footprint of a university – the case of NTNU. J Cleaner Prod 2013;48:39–47.
[20] Escobedo A, Briceno S, Juarez H, Castillo D, Imaz M, Sheinbaum C. Energy consumption and GHG scenarios of a university campus in Mexico. Energy Sustain Develop 2014;18:49–57.
[21] Ozawa-Meida L, Brockway P, Letten K, Davies J, Fleming P. Measuring carbon performance in a áK university through a consumption-based carbon footprint: De Montfort University case study. J Cleaner Prod 2013;56:185–98.
[22] Sreedharan P, Farbes J, Cutter E, Woo CK, Wang J. Microgrid and renewable generation integration: University of California, San Diego. Appl Energy 2016;169:709–20.
[23] Wang Y, Zhao F-Y, Kuckelkorn J, Spliethoff H, Rank E. School building energy performance and classroom air environment with the heat recovery heat pump and displacement ventilation system. Appl Energy 2014;114:58–68.
[24] Sekki T, Airaksinen M, Saari A. Measures energy consumption of educational buildings in a Finnish city. Energy Build 2015;87:105–15.
[25] Gaitani N, Lehmann C, Santamouris M, Mihalakakou G, Patragias P. Using principal component and cluster analysis in the heating evaluation of the school building sector. Appl Energy 2010;87:2079–86.
[26] Desideri U, Proietti S. Analysis of energy consumption in the high schools of a province in central Italy. Energy Build 2002;34:1003–16.
[27] Raatikainen M, Skon J-P, Leiviska K, Kolehmainen. Intelligent analysis of energy consumption in school buildings. Appl Energy 2016;165:416–29.
[28] Pereira LD, Raimanodo D, Corgnati SP. Energy consumption in schools – a review paper. Renew Sustain Energy Rev 2014;40:911–22.
[29] Fawcett T. Energy use and carbon emissions from the higher education sector. UKERC Working Paper No. UKERC/WP/DR/2005/005. UK: Energy Research Centre; 2005.
[30] Sinha P, Schew WA, Sawant A, Kolwaite KJ, Strode SA. Greenhouse gas emissions from US institutions of higher education. J Air Waste Manage Assoc 2010;60:568–73.
[31] Robinson O, Kemp S, Williams I. Carbon management at universities: a reality check. J Cleaner Prod 2015;106:109–18.
[32] Altan H. Energy efficiency in UK higher education institutions. Energy Policy 2010;38:722–7731.
[33] Mazhar MU, Bull R, Lemon M, Mallaburn P. The current state of strategic carbon management within the UK higher education sector: leading the way forward? 9th international symposium on sustainable leadership, Salzburg, June 3–6. 2014.
[34] Aranda A, Ferreira G, Mainar-Toledo MD, Scarpellini S, Sastresa EL. Multiple regression models to predict the annual energy consumption in the Spanish banking sector. Energy Build 2012;49:380–7.
[35] Wadud Z, Graham DJ, Noland RB. Modeling fuel demand for different socio-economic groups. Appl Energy 2009;86:2740–9.
[36] Wadud Z, Noland RB, Graham DJ. A Semiparametric model of household gasoline demand. Energy Econ 2010;32(1):93–101.
[37] Wadud Z. Diesel demand in the road freight sector in the UK: estimates for different vehicle types. Appl Energy 2016;165:849–57.
[38] Wadud Z, Dey HS, Kabir MA, Khan SI. Modeling and forecasting natural gas demand in Bangladesh. Energy Policy 2011;39:7372–80.
[39] World Resources Institute. The Greenhouse Gas Protocol: a corporate accounting and reporting standard; 2001. Available at: <https://www.wri.org/sites/default/files/pdf/ghg_protocol_2001.pdf> [accessed: July 2017].
[40] Office for National Statistics. Electricity: United Kingdom Energy Statistics (DUKES); 2017 [chapter 5]. Available at: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/552059/Chapter_5_web.pdf> [accessed: March 2017].
[41] Department for Business, Energy and industrial Strategy. Fuel price indices for the Industrial sector; 2017. Available <https://www.gov.uk/government/statistical-data-sets/industrial-energy-price-indices> [accessed: February 2017].
[42] Huebner G, Shipworth D, Hamilton I, Chalabi Z, Oreszczyn T. Explaining domestic energy consumption: the comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes. Appl Energy 2015;177:692–702.
[43] Hsiao C. Analysis of panel data. Cambridge: Cambridge University Press; 2003.
[44] Baltagi BH. Econometric analysis of panel data. West Sussex: John Wiley & Sons Ltd; 2005.
[45] Schnuck R. Within and between estimates in random-effect models: advantages and drawbacks of correlated random effects and hybrid models. The Stata J 2013;13(1):65–76.
[46] Bell A, Jones K. Explaining fixed effects: random effects modelling of time-series-cross sectional and panel data. Polit Sci Res Methods 2015;3(1):133–53.
[47] Greene WH. Econometric analysis. New Delhi: Pearson Education Inc; 2005.
[48] Hausman JA. Specification tests in econometrics. Econometrica 1978;46:1251–371.
[49] Mundlak Y. On pooling of time series and cross section data. Econometrica 1978;46(1):69–85.
[50] Braun MR, Altan H, Beck SBM. Using regression analysis to predict the future energy consumption of a supermarket in the UK. Appl Energy 2014;130:305–13.
[51] Wadud Z. Decomposing the drivers of aviation fuel demand using simultaneous equation models. Energy 2015;83(C):551–9.
[52] Huebner G, Shipworth D, Hamilton I, Chalabi Z, Oreszczyn T. Understanding electricity consumption: a comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes. Appl Energy 2016;159:589–600.
[53] Shove E. What is wrong with energy efficiency? Build Res Informat 2017. https://doi.org/10.1080/09613218.2017.1361746.
[54] Royston S, Selby J, Shove E. Invisible energy policies: a new agenda for energy demand reduction. Energy Policy 2018;123:127–35.