Energy benchmarking of commercial buildings: a low-cost pathway toward urban sustainability

Matt Cox, Marilyn A Brown and Xiaojing Sun

School of Public Policy, Georgia Institute of Technology, D M Smith Building, 685 Cherry Street, Room 312, Atlanta, GA 30332-0345, USA

E-mail: marilyn.brown@pubpolicy.gatech.edu

Received 3 February 2013
Accepted for publication 14 June 2013
Published 16 July 2013
Online at stacks.iop.org/ERL/8/035018

Abstract

US cities are beginning to experiment with a regulatory approach to address information failures in the real estate market by mandating the energy benchmarking of commercial buildings. Understanding how a commercial building uses energy has many benefits; for example, it helps building owners and tenants identify poor-performing buildings and subsystems and it enables high-performing buildings to achieve greater occupancy rates, rents, and property values. This paper estimates the possible impacts of a national energy benchmarking mandate through analysis chiefly utilizing the Georgia Tech version of the National Energy Modeling System (GT-NEMS). Correcting input discount rates results in a 4.0% reduction in projected energy consumption for seven major classes of equipment relative to the reference case forecast in 2020, rising to 8.7% in 2035. Thus, the official US energy forecasts appear to overestimate future energy consumption by underestimating investments in energy-efficient equipment. Further discount rate reductions spurred by benchmarking policies yield another 1.3–1.4% in energy savings in 2020, increasing to 2.2–2.4% in 2035. Benchmarking would increase the purchase of energy-efficient equipment, reducing energy bills, CO$_2$ emissions, and conventional air pollution. Achieving comparable CO$_2$ savings would require more than tripling existing US solar capacity. Our analysis suggests that nearly 90% of the energy saved by a national benchmarking policy would benefit metropolitan areas, and the policy’s benefits would outweigh its costs, both to the private sector and society broadly.

Keywords: benchmarking, urban sustainability, energy efficiency, commercial buildings, information failures

Online supplementary data available from stacks.iop.org/ERL/8/035018/mmedia

1. Introduction

Commercial buildings accounted for nearly one-fifth of the energy consumed in the US in 2010, and their portion of the nation’s energy budget is expected to increase to 21% by 2035 (EIA 2011a). Commercial buildings dominate the urban landscape, and their energy requirements contribute to urban air quality and heat island effects. As a result, innovative policies that promote energy-efficient commercial buildings are critical to sustainable development. We focus here on the use of energy benchmarking to inform building owners and tenants about poor-performing buildings and subsystems and to enable high-performing buildings to achieve greater occupancy rates, rents, and property values. We estimate the possible impacts of a national policy mandating the energy benchmarking of US commercial buildings, emphasizing the benefits to sustainable urban development.

The commercial building sector suffers from three main information failures. First, there is the problem of information
asymmetry: building owners and managers know more about the energy performance and efficiency of their buildings than prospective buyers and tenants. Analogous to the case of ‘lemons’ in the used car market as described by Akerlof (1970), this can lead to inefficient transactions. Second, there are principal–agent problems in the sector, which occur when one party (the agent) makes decisions in a market and a different party (the principal) bears the consequences. This issue was found by Prindle (2007) to be significant and widespread in many end-use energy markets in both the US and other countries. In many commercial buildings, architects, engineers, and builders select equipment, duct systems, windows, and lighting for future building occupants. Similarly, landlords often purchase and maintain appliances and equipment for tenants who pay the energy bill, providing little incentive for the landlord to invest in efficient equipment (Brown 2001). Third, a decades-long research effort has identified discount rates related to equipment purchases that are far higher than theoretically anticipated, resulting in fewer purchases of high-efficiency equipment (Frederick et al 2002, Train 1985).

This analysis focuses on giving building owners in the country access to baseline information on their building’s energy consumption (‘benchmarking’), which is currently unavailable or underutilized in most parts of the US. This could be accomplished by requiring utilities to submit energy data in a standard format to a widely used database, such as Portfolio Manager1, which currently maintains information on hundreds of thousands of buildings in the US, provided by building owners and managers. Using existing software packages, meter data from utilities and building owners could be combined to provide a ‘virtual building meter’, allowing for building-wide assessments2. The data would then be available to the building owner and the utility and maintained by the Environmental Protection Agency (EPA).

According to a report sponsored by the US Green Building Council and others (Carbonell et al 2010), the EPA may have the authority to require utilities to submit building energy data under section 114 of the Clean Air Act. This utility data must be connected to individual buildings to be useful in providing building owners with baseline energy performance information. A uniform national building identification system, similar to the VIN system for cars, could facilitate this connection regardless of where a building is located, how it is used, or whether it has multiple street addresses—all currently issues for energy benchmarking. The benchmarking approach assessed here involves two features.

• Utilities are required to submit whole building aggregated energy consumption data for all tenants to the EPA Portfolio Manager.
• A national registry of commercial buildings, with each building receiving a unique building identification (BID) number is developed.

If implemented, better building energy data would become available to owners, tenants, and utilities. In turn, benchmarking efforts could be accelerated; demand side management programs could become more feasible; municipal governments would have a uniform system for building codes and mandated disclosure reporting; and the federal government would gain valuable data to inform the ENERGY STAR® building certification standards and the commercial building energy consumption survey. The real estate sector would be able to provide better information to clients, and energy performance could be better incorporated into property assessments.

2. Background

Benchmarking creates an energy consumption baseline for a specific building. If benchmarking is completed for a large set of buildings and stored in a shared database, comparisons become possible. Benchmarking also helps to set priorities for limited staff time and capital. EPA and the American Council for an Energy-Efficient Economy (ACEEE) both suggest that savings up to 10% can be made at little or no cost to building owners, savings which are frequently overlooked (Dunn 2011, Nadel 2011).

The federal government benchmarks its buildings as a result of Section 432 of the Energy Independence and Security Act of 2007. However, policy experience with benchmarking in the US is largely tied to mandated disclosure policies at the state and local level (figure 1). Most of these policies emphasize the residential sector or are under consideration, but six cities and two states (California and Washington) have adopted mandated disclosure, which necessitates benchmarking of commercial buildings. Every existing American program, including an international effort between the US and Canada, uses Portfolio Manager as the benchmarking tool (EPA 2011). As of 2012, Portfolio Manager includes data on the performance of more than 300 000 buildings in the US, providing normalized building scores that qualify buildings for ENERGY STAR certification and help achieve LEED certification.

The Institute for Market Transformation (IMT) summarized recent experiences of nine current US programs (Burr et al 2011). As a result of program reviews and in-depth stakeholder discussions, a series of best practices were recommended for benchmarking, the main one being to follow EPA guidelines surrounding the use of the Portfolio Manager. This recommendation enables jurisdictions to avoid debates over building use and building type classifications, but there are other benefits as well, including easy integration of building data into the Portfolio Manager format.

2.1. Results from implementing governments3

While Europe has used mandated disclosure and benchmarking programs for many years, the US is just beginning to implement these programs. Currently, the governments of

---

1 EPA Portfolio Manager www.energystar.gov/index.cfm?c=evaluate_performance.bus_portfoliomanager.

2 Certain utilities, like Consolidated Edison and Austin Energy, have developed meter aggregating tools to collect whole-building energy consumption data.

3 Program managers from New York City, Seattle, Austin, Washington, DC, and DOE’s Building Technology Program were interviewed.
Figure 1. Mandated disclosure and benchmarking efforts in the US. Reproduced with permission from Burr (2012).

Table 1. Main conclusions from program managers.

| Core lessons from program managers | Key areas for assistance |
|-----------------------------------|--------------------------|
| • Portfolio Manager and the SEED database are great tools | • Clarifying leadership roles at the federal level |
| • Programs are reducing information gaps | • Workforce certification programs |
| • Implementation is more difficult than anticipated | • Defining confidential data |
| • Building aggregation capacity is crucial | • Funding or rules for utilities to aggregate building data and facilitate release and access |

New York City, Seattle, Washington, DC, Philadelphia, and Austin, Texas are taking leadership roles (Keicher et al. 2012, Burr et al. 2011). Results have begun to be released, notably in New York City, where 2010 and 2011 data have been reported for over 2000 buildings. The 2010 report suggested that if the bottom half of large commercial buildings could be brought to just the median level of energy performance, energy consumption and greenhouse gases from this building class would fall by 18% and 20%, respectively (City of New York 2012a). The summary of the 2011 release suggests that such an effort would reduce total citywide greenhouse gases by 9% (City of New York 2012b). Interviews with key program managers from leading cities discovered consistent findings that can be informative for policymakers (table 1).

First, Portfolio Manager has found broad acceptance as the principal benchmarking tool. The time-series and cross-sectional comparison capabilities of the tool make it extremely attractive. The Standard Energy Efficiency Data Platform⁴ that DOE provides has also been well received because it helps the local governments share best practices. However, the dual-agency approach has led to confusion about federal roles, and some cities have suggested that clarifying leadership positions would be helpful.

Second, all program managers interviewed believe a large information gap related to building energy consumption existed in their jurisdictions prior to the benchmarking and mandated disclosure laws. While benchmarking efforts have assisted in reducing this gap by informing building owners about total building performance, the gap remains.

Third, tenant authorization is required for building owners to access energy consumption data in many jurisdictions. Rules and support for utilities to facilitate easy access and release of aggregated building data are particularly important. One manager stated that if he were starting over, aggregated building data rules would be the first thing instituted.

Fourth, every program experienced delays in implementation, largely due to the economic downturn in 2008. Frequently, timelines had to be amended after the program began in earnest. Lastly, a commonly noted issue is the lack of a qualified workforce—certification programs for contractors

---

⁴ DOE Standard Energy Efficiency Data Platform: www1.eere.energy.gov/buildings/commercial/seed_platform.html.
are strongly desired. Benchmarking and mandated disclosure efforts have the potential to create and expand markets for energy contractors, and some means to differentiate between contractors would reduce other information barriers for building owners.

2.2. Policy rationale

‘Policy actions…could, in principle, correct for the excessive present-mindedness of ordinary people’ (Solow 1991).

Benchmarking has the potential to reduce information asymmetries in the marketplace and to lower the discount rates used by consumers in the sector. A few scholars question the extent and evidence of such problems (Alcott and Gillingham 2012, Gillingham et al 2009, Jaffe and Stavins 1994). This skepticism seems to stem from the information assumptions of neoclassical economics. Policy tools based on such theory are unable to modify discount rates and provide no policy relevant advice for information-based gaps (Stern 1986). In contrast, empirical research has found that information can modify discount rates in use; providing information may address a barrier to the deployment of energy-efficient technologies that other approaches cannot.

Theoretically, one way discount rates are determined is by combining the market interest rate with a time preference premium and some level of uncertainty or risk; with efficient capital markets, discount rates should converge with interest rates (Fuchs 1982). Hausman theorized that rational actors would equate the potential stream of energy savings from more efficient technologies with the monetary savings from buying less-expensive equipment. His findings on implicit discount rates, however, did not match the theory; for efficient air conditioners, consumers used discount rates that were much higher than the market interest rate. ‘Other factors such as uncertainty and the possibility of technological change do not seem sufficient to explain the high discount rate which we found’ (Hausman 1979). Later research would find many instances where empirically observed discount rates deviated from theoretical predictions, finding that future gains receive higher discounting than future losses (Thaler 1981), that smaller anticipated results (either positive or negative) receive higher discount rates than larger anticipated results (Benzion et al 1989), and that consumers prefer improving sequences of outcomes (Frederick and Loewenstein 2002, Varey and Kahneman 1992). Furthermore, Sultan and Winer (1993) found no evidence of consumers using market-based discount rates across a number of appliances.

Research specific to equipment-purchasing decisions found numerous discount rates in use across the population. These discount rates vary over time and appliances (Train 1985, Koomey 1990). Frederick and Loewenstein (2002), in a review of the theoretical and empirical history of discount rates, found the following.

The implicit discount rate was 17–20% for air conditioners (Hausman 1979); 102% for gas water heaters, 138% for freezers, 243% for electric water heaters (Ruderman et al 1987); and from 45% to 300% for refrigerators, depending on assumptions made about the cost of electricity (Gately 1980).

Disparate findings in discount rates across the population pose theoretical difficulties, but open the door for different policy approaches and rationales. Tools in regulatory, financial, and information areas may help to address discount rate issues: for example, equipment standards and subsidies can result in choices that approximate the effect of lowering discount rates by eliminating low-efficiency choices and reducing first cost, respectively. However, information-based policies have the unique ability to modify the discount rate in use. Studies have found that providing information can reduce discount rates anywhere from 3% to 22% (Coller and Williams 1999, Goett 1983). Coller and Williams suggest that information about energy consumption will result in a 5% decline in discount rates for energy decisions made by the median population. Depending on the discount rate in use, an adjustment of this size could dramatically impact equipment decisions.

In the commercial sector, numerous studies (Christmas 2011, Campbell 2011, Miller et al 2008, Jackson 2009, Das et al 2011) show higher occupancy rates, higher rents, and higher property values for high-efficiency buildings. Benchmarking could increase the market demand for these buildings. Portfolio Manager itself has the potential to address some information gaps through its use of time-series data and cross-sectional comparisons. This may lead to more energy-efficient technology choices, reduced uncertainty in maintenance costs, lower fuel costs, and ease the attainment of building certifications like ENERGY STAR. The ties between Portfolio Manager and ENERGY STAR certification also reduce transaction costs for renters desiring high-performance space. This could reduce the size of the principal–agent problem by creating market and social pressure for building owners to consider energy in purchasing decisions, particularly when combined with mandated disclosure.

Recent studies show that benchmarking spurs energy efficiency investments (NMR Group, Inc. with Optimal Energy, Inc. 2012), one of the fastest, direct, and cost-effective means of reducing greenhouse gas emissions (Ciochetti and McGowan 2010). To the extent that benchmarking limits the emission of greenhouse gases, it helps to correct this negative externality and mitigates the threat posed by climate change. In essence, benchmarking has the potential to be a step toward better consumer choices without resorting to pricing instruments or regulation. In this way, it is a policy tool complementary to economically efficient approaches such as carbon pricing policies, guiding behavioral changes by energy managers and users. With the commercial sector representing 19% of US CO₂ emissions (over 3% of global emissions), and that percentage expected to grow (DOE 2012), managing the emissions of the sector is critical to ‘preventing dangerous interference with the climate system’ (UNFCCC 1992).

3. Methodology

Technological selection in many modeling efforts is based on a series of economic considerations such as first costs and discount rates. The impact on energy consumption is not insignificant, as the efficiency of technologies available
for meeting the same demand for energy services in the marketplace is quite varied. For modeling projections, technological forecasting, technological assessment and progress, as well as social factors that influence the adoption of technologies, are critical (Coates et al. 2001). Such forecasting and modeling efforts face long odds of accurately projecting future outcomes, but can be useful in providing estimates as well as informing policy debates (Silberglitt et al. 2003). Furthermore, many models also struggle with simulating macroeconomic spillover effects, where actions taken in one sector of the economy change conditions and thus alter decisions in other sectors of the economy.

Our analysis of benchmarking in the commercial sector utilizes the Georgia Tech version of the Energy Information Administration’s (EIA) 2011 National Energy Modeling System (GT-NEMS) (a detailed description of the NEMS commercial module, which was modified for this study, can be found in the Commercial Demand Module of the National Energy Modeling System: Model Documentation (EIA 2011b)). GT-NEMS contains a technology menu and forecast developed from manufacturer surveys of anticipated technological performance for several hundred types of equipment. GT-NEMS also has the capacity to simulate macroeconomic spillover effects. When selecting a technology in the commercial sector to meet a demand for energy services, GT-NEMS uses a combination of discount rates and the rate for US government ten-year Treasury notes to calculate consumer ‘hurdle rates’ used in evaluating equipment-purchasing decisions. While the macroeconomic module of GT-NEMS determines the rate for ten-year Treasury notes endogenously, the discount rates are exogenous. Modifying these inputs is the primary means of estimating the impact of benchmarking for the commercial sector in this analysis. This is done in two steps: first, by updating the discount rates to reflect a broader selection of the literature; and second, by adjusting the updated discount rates to account for the effects of a national benchmarking policy.

The GT-NEMS inputs for discount rates are separated by end use, including space heating, space cooling, ventilation, lighting, water heating, cooking, and refrigeration, divided into seven population segments for each end use. Each population segment is capable of using a different discount rate with regard to the end use in question each year. In the Annual Energy Outlook 2011 (EIA 2011a) reference case, these discount rates are quite high; for example, more than half of the consumer choices in lighting and space heating use discount rates greater than 100%, and less than 3% of the population uses discount rates under 15% (EIA 2011b).

Such high discount rates are not reflected by the bulk of the existing research. This problem has been recognized for some time in energy forecasting models (Decanio and Laitner 1997). An extensive literature review spanning four decades uncovered more than two-dozen studies estimating implicit discount rates for commercial consumers across the GT-NEMS series of appliances. The mean discount rates in this literature ranged from 17% (space heating and space cooling both) to 63% (refrigerators). The simulation and econometrics to analyze risk (SIMETAR) tool was used to develop continuous probability distribution functions for each end use. GRKS distributions were used for space cooling, lighting, cooking, and water heating. SIMETAR matched Weibull distributions as a better fit for space heating and refrigeration. Ventilation was the sole end use to have no specific studies; the space heating distribution was used to represent it (see supplementary material, available at stacks.iop.org/ERL/8/035018/mmedia, for full details of the discount rates and distributions used in the modeling).

Ideally, the continuous functions would be used to model the distribution of discount rates across the population. However, GT-NEMS is not suited for this due to the segmented approach previously described. Therefore, the probability density functions were divided into seven segments of equal area for each end use. The median value of these seven segments generates the updated discount rates scenario (UDR). To estimate the impact of benchmarking, two scenarios were modeled. In the Benchmarking 5% scenario, the findings of Coller and Williams (1999) were applied, meaning the median value declined by five percentage points. The quotient of this ‘benchmarked’ median discount rate and the updated median discount rate was calculated and used as an adjustment factor to the other six population segment medians. In this way, the findings of Coller and Williams are carried throughout the consumer population, since each population segment reduces by the same proportion as the median. Given the uncertainty in the estimates of information-based discount rate modifications and the wide range of reported implicit discount rates (Train 1985), we also produce the Benchmarking 10% scenario, which follows the same method but applies a 10% reduction to the median discount rate from the UDR scenario.

GT-NEMS adds the rate of ten-year Treasury notes to these values, which vary by year according to macroeconomic conditions. The reference case Treasury note rates were subtracted from the updated discount rates so that the final hurdle rate calculated by GT-NEMS are consistent with the values suggested by the literature. These modifications generate the main policy cases: ‘Benchmarking 5%’ and ‘Benchmarking 10%’. All policy scenarios begin implementation in 2015.

A number of sensitivities were also modeled, where benchmarking pushes R&D forward, bringing new, highly efficient technologies to the marketplace. This sensitivity (referred to as ‘Benchmarking +’) utilizes the EIA High Tech technology suite for the commercial sector where 40 new high-efficiency technologies are introduced (EIA 2011a), and is consistent with the ‘announcement effect’, which describes the phenomenon that firms and customers adjust their behavior in the interim between the announcement of a regulation and its implementation. However, it needs to be acknowledged that there is still a gap between the High Tech assumptions and state-of-the-art technologies that are currently available in the market place. This analysis would
be more thorough if GT-NEMS could account for the latest developments in technology innovation and deployment. Furthermore, the benchmarking policy combined with whole building design practices that consider the design, operations, and maintenance of buildings could lead to greater impacts. For example, GT-NEMS accounts for the operations and maintenance costs of equipment; however, the provision of energy information through a benchmarking policy could lead building owners to improve overarching operations and maintenance regimes (such as shell characteristics), which could be an unaccounted-for benefit of this policy approach.

To give a sense of the uncertainty in the analysis, two alternative scenarios are included that bound the potential benefits of modifying the discount rates of consumers. Energy expenditures are expected to severely impact the total benefits of the policy. The high benefits case reflects a future where EPA regulations make it more expensive to keep using coal for base-load power, resulting in higher electricity prices. As a result, consumers will demand more efficient technologies, and manufacturers will deliver them. To model this, we utilize the assumptions of the AEO 2011 High Coal Cost side case and the High Tech side case technology menu. The low benefits case reflects a case where there are diminished prices for coal, perhaps from reduced demand due to cheap natural gas. At the same time, we expand estimates of supply for shale gas, resulting in lower gas prices than in the GT-NEMS reference solution. This scenario uses the reference case technology suite. A description of the side cases can be found in AEO 2011 (EIA 2011a).

The macromodel of gas space heating, followed by ventilation. Electric space heating experiences an increase in consumption after 2025, following the adoption of more heat pumps in those years. Excluding ventilation, the average saving for an end use is 1% in 2020 and 1.2% in 2035 in the Benchmarking 10% scenario.

Both benchmarking scenarios result in a series of technology shifts across the major end uses. For space heating, GT-NEMS projects a fuel shift from natural gas to electric technologies. Benchmarking 5% mostly shows service demand shifting within each fuel type, but also a trade of 4 TBtu in service demand between natural gas space heating and electric space heating in 2020. By 2035, this service demand trading increases to 18 TBtus, accounted for by a shift toward electric heat pumps. Benchmarking 10% follows a similar trajectory; in 2020, the single most significant change in service demand is a move from typical natural gas furnaces to high-efficiency natural gas furnaces. However, by 2035, about 30 TBtus in service demand for natural gas heating are shifted to air-source heat pumps, representing a change in the fuels and technologies selected by consumers to meet space heating demand. This technology shift is partially responsible for greater natural gas savings than electricity. However, the energy price context is constantly evolving. GT-NEMS projects increasing natural gas prices, but a prolonged presence of cheap natural gas may drive the private sector to develop more and better natural gas end-use technologies, which could affect commercial sector technology choices.

### 4. Results

#### 4.1. Technology shifts

In both 2020 and 2035, the greatest savings are from natural gas space heating, followed by ventilation. Electric space heating experiences an increase in consumption after 2025, following the adoption of more heat pumps in those years. Excluding ventilation, the average saving for an end use is 1% in 2020 and 1.2% in 2035 in the Benchmarking 10% scenario.

Both benchmarking scenarios result in a series of technology shifts across the major end uses. For space heating, GT-NEMS projects a fuel shift from natural gas to electric technologies. Benchmarking 5% mostly shows service demand shifting within each fuel type, but also a trade of 4 TBtu in service demand between natural gas space heating and electric space heating in 2020. By 2035, this service demand trading increases to 18 TBtus, accounted for by a shift toward electric heat pumps. Benchmarking 10% follows a similar trajectory; in 2020, the single most significant change in service demand is a move from typical natural gas furnaces to high-efficiency natural gas furnaces. However, by 2035, about 30 TBtus in service demand for natural gas heating are shifted to air-source heat pumps, representing a change in the fuels and technologies selected by consumers to meet space heating demand. This technology shift is partially responsible for greater natural gas savings than electricity. However, the energy price context is constantly evolving. GT-NEMS projects increasing natural gas prices, but a prolonged presence of cheap natural gas may drive the private sector to develop more and better natural gas end-use technologies, which could affect commercial sector technology choices.

#### 4.2. Reduced energy consumption and expenditures

The benchmarking policy scenarios modeled in this study all target the seven major equipment classes that account for approximately 45% of the energy used by commercial buildings: space heating, space cooling, ventilation, lighting, water heating, cooking, and refrigeration. The impact can be seen in figure 2. The updated discount rate reduces projected energy consumption by 4.0% in 2020 and 8.7% in 2035, for these seven energy end uses. This finding suggests that the EIA reference case overestimates future US energy consumption by underestimating future investments in energy-efficient building equipment. Analysis of forecasts from the 1980s relative to actual US energy use indicates that this overestimation bias is long-standing (Laitner 2009).

Benchmarking 5% shows savings of 1.3% in 2020 (150 TBtu) and 2.2% (260 TBtu) in 2035 when compared to the UDR case. Benchmarking 10% projections are slightly larger at 1.4% (160 TBtu) in 2020 and 2.4% (270 TBtu) in 2035. In 2020, the high benefits case produces a 3.7% decline in energy consumption from the UDR case, rising to 10% in 2035.

Benchmarking reduces energy consumption without reducing the commercial sector’s growing spatial footprint. As a result, energy intensity, measured in Btu ft\(^{-2}\), declines, as does the nation’s energy intensity as a whole. In 2020, benchmarking results in a 1% improvement in energy intensity, relative to the UDR case.

Benchmarking 10% shows reduced energy demand over the modeling horizon; both natural gas and electricity consumption is down an average of 1.6% compared to the UDR case. The result is a reduction in the average price for natural gas of 0.3%. When Benchmarking 5% is compared to the UDR case, natural gas and electricity consumption decline by an average of 2.2% and 1.4%, respectively, with a corresponding 0.2% and 0.3% average reduction in price for each fuel. Rebound effects, where lower levels of energy consumption reduce prices and thus increase energy consumption, contribute to limiting energy savings in the modeled scenarios (Sorrell et al 2009).

Decreased demand combined with declining energy prices result in a reduction in energy expenditures by the
owners of commercial buildings. Compared with the UDR case, Benchmarking 10% expenditures drop by 0.7% in 2020, saving $1.2 billion; in 2035, expenditures drop 1.1%, saving $2.4 billion. On average, annual energy expenditures drop by 0.7%, valued at $1.4 billion. These savings cumulatively total $13 billion through 2035, and $16 billion over the lifetime of the installed equipment (at a 7% discount rate). In Benchmarking 5%, 2020 expenditures drop 0.8%, worth $1.5 billion; 2035 expenditures drop 0.9% and are worth $1.9 billion. Savings through 2035 have a net present value of $11 billion, increasing to $13 billion over the lifetime of the equipment (also at a 7% discount rate). While the savings appear modest compared to some other energy efficiency programs (Brown et al. 2013, Gillingham et al. 2006), we have shown that these differences change technology choices in the commercial sector, and we will show that they are still meaningful on the supply side and with respect to environmental benefits.

Several NEMS features may have restricted the energy-saving potential. First, NEMS models the discount rates used by commercial customers for only seven equipment classes. Office equipment and miscellaneous end uses are modeled through a different, simplified fashion with negligible efficiency improvements. In reality, a benchmarking policy could reduce energy consumption of all end uses, which means the consumption reduction presented in figure 2 is at the low-end of the saving potential. In addition, the technology choice decision rule used in NEMS presents a barrier to higher energy savings. According to the model, when it comes to end-use equipment retrofit or placement, consumers have a certain freedom in choosing technologies, but the majority of them are limited to use either a later version of the same technology or at least remain in the same fuel type. At a minimum, this restricts fuel switching in the sector and could potentially stop consumers from choosing economically feasible high-efficiency technology and dampen the energy savings. Previous studies of energy economic models (including NEMS) have found evidence that these models tend to reinforce the status quo, underestimate innovation (in both policy and technology), and miss market potential, issues that may explain some of the modest savings potential for benchmarking projected by GT-NEMS (Laitner et al. 2003, Laitner 2009).

4.3. Cost effectiveness

While the benchmarking policy option is modeled as ceasing in 2035, the benefits of the policy would extend into the future due to the lifetime of energy-saving technologies installed as a result of the policy. Energy-efficient technologies have varying lifetimes (for example, chillers and boilers last longer than natural gas water heaters)7. This analysis assumes that energy savings degrade at 5% annually (Brown et al. 1996). Therefore, technologies installed in 2035 provide the greatest savings in that year, with a linear decline in savings out to 2055, when energy savings are no longer expected. The same rationale is applied to emissions benefits.

Aside from the private sector benefits from reduced energy expenditures, there are social benefits from fewer pollutant emissions. Our analysis includes criteria pollutant (SO$_2$, NO$_x$, and PM$_{2.5}$ and PM$_{10}$) benefits and CO$_2$ benefits. Changing the regulatory framework for these pollutants and other changes (lower prices or innovations, for example) that result in dramatic departures from projected means of meeting energy demand would lead to different estimates of the costs and benefits associated with these pollutants.

Criteria pollutant benefits are calculated based on values from the National Research Council (2010), and take into account public health effects, damage to crops and timber, buildings, and recreation. Such damage tends to vary substantially depending on meteorological conditions, proximity of populations to emitters, and the sources and means of electricity generation (Fann and Wesson 2011). The National Research Council estimates exclude damage from mercury pollution, climate change, ecosystem impacts, and other areas where damage is difficult to monetize. Even

---

6 Examples of miscellaneous uses include service station equipment, automated teller machines, telecommunications equipment, and medical equipment.

7 See tables 5.3.9, 5.6.9, and 5.7.15 in the DOE Buildings Energy Data Book http://buildingsdatabook.eren.doe.gov/.
with this incompleteness, damage from coal power plants is estimated to exceed $62 billion annually, and recent analysis suggests that the damage from coal power plants exceeds the value added to the economy (Muller et al. 2011). The national average values provided for electricity generation and on-site use of energy sources are used to analyze the emissions benefits of benchmarking.

Carbon dioxide emissions are outputs of GT-NEMS and are the result of fuels used for energy on-site and in the electricity sector. The economic value of reductions in CO₂ is estimated by multiplying the annual decrement in emissions by the ‘social cost of carbon’ (SCC). In this analysis, the central values of the US Government Interagency Working Group on the Social Cost of Carbon (EPA 2010) are used, ranging from $25 per metric ton of CO₂ in 2015 to $47 per metric ton of CO₂ in 2050 (in 2009-$).

Benchmarking policies improve the ability of the market to operate effectively and take advantage of low-cost energy-saving opportunities. As a result, the damage from pollutants declines, representing significant public benefits over the duration of the policy timeline analyzed here. However, the two benchmarking scenarios strongly diverge on the emissions benefits. This is due to discontinuities in demand side choices that impact supply side decisions. For example, GT-NEMS projects a temporary increase in the use of coal for electric generation in the East North Central census division in the Benchmarking 5% scenario. This is largely due to national reductions in electricity consumption, reducing the price of coal and increasing the attractiveness of coal-fired electric generation in the decade post-policy initiation. The Benchmarking 10% scenario produces similar reductions in both natural gas and electricity, creating price effects that favor natural gas generation in the electric sector, particularly from combined heat and power. The result is that increased coal consumption and the related increase in emissions projected under Benchmarking 5% are avoided by Benchmarking 10%. The supply side is behaving very differently in these two cases. This non-linearity accentuates the value of a general equilibrium framework, because such uncertainties, tipping points, and supply–demand interplays would not otherwise be visible.

When compared to the UDR case, the cumulative value of avoided criteria pollutant emissions is estimated at $1.4–3.4 billion in 2020, growing to $3.0–8.2 billion in 2035 (ranges reported for Benchmarking 5% and Benchmarking 10% using a 3% discount rate). Including CO₂ benefits, the net present value of these emissions reductions would be $3.9–10.5 billion through 2035 with the potential to grow to more than $18 billion over the lifetime of equipment installed under the policy.

To put these emissions reductions in perspective, we estimated the scale of installed capacity required to achieve the same level of CO₂ mitigation from nuclear power and solar photovoltaics, as other carbon-free electricity sources. In 2020, the Benchmarking 5% and 10% projections show reductions of 4–22 million metric tons (MMT) CO₂ relative to the UDR case. Based on the average carbon intensity of grid-supplied electricity projected by GT-NEMS, achieving these carbon savings requires replacing 7.2–39.7 TWh of electricity generation in 2020 with carbon-free generation. Nuclear power typically has a much higher capacity factor than solar; for the purposes of this comparison, we assume an 89% capacity factor. In 2020, this would require 925 MW–5.1 GW of nuclear power. Assuming a capacity factor of 17% for solar photovoltaics, achieving the same carbon savings would require 4.8–26.6 GW of solar capacity.

While the reductions in CO₂ are significant, this result adds to a growing body of literature that is skeptical of technological innovation and behavioral change as sufficient to address the threat of climate change—regulatory and taxation approaches may well prove necessary (van den Bergh 2013). The US government has established goals for energy reductions in the commercial sector (the Better Buildings Initiative, aiming for a 20% reduction from projected 2020 energy consumption in commercial buildings) and greenhouse gas reductions for the entire economy (a 17% reduction in CO₂ from 2005 levels, established at the Copenhagen climate negotiations in 2009). Benchmarking policies can assist but many other initiatives will be necessary to achieve these goals.

Turning to costs associated with benchmarking, buildings with multiple tenants will require aggregation services in order to determine the energy footprint of an entire building. We call these compliance costs. These costs were determined using the 2003 Commercial Building Energy Consumption Survey (CBECS) data (EIA 2007), which provides the number of multi-tenant buildings with electric and natural gas services. The average square footage of a multi-tenant building from CBECS is used in conjunction with GT-NEMS projections of commercial floor space to produce estimates of the number of multi-tenant buildings that will exist between 2004 and 2035. Burr (2012) estimates that existing laws will require 60,000 buildings to undergo benchmarking regardless of this policy option, so these are subtracted from the total. It is assumed that compliance costs will be the same for each building, following the Consolidated Edison model in New York City, and is set at $102.50 (2011-$) for electricity and natural gas. Consolidated Edison provides both gas and electric data for $102.50, a value that was calculated based on anticipated labor costs to collect and aggregate data and approved by the New York State Public Service Commission (Consolidated Edison Company 2010). We model that a building needing aggregation for both fuels would incur costs of $205, based on the assumption that gas and electric services may not be provided by the same utility in every jurisdiction, making our estimate of cost higher than would likely be the case. The end result is an initial cost of $141 million (2009-$) in 2015. Costs for new buildings after 2015 are also included, ranging from $2.7 million in 2016 to $3.2 million in 2035 (2009-$).

These costs are modeled as public costs due to concerns about distributional impacts and policy viability. If these costs were directed to utilities, opposition would likely grow substantially. The costs of accounting upgrades and software development are probably minor, but the cost of benchmarking every building would not be. If costs were directed toward building owners, they would be incentivized...
to avoid complying. As the purpose of the policy is to identify and benchmark the energy consumption of as many buildings as possible in the US, that outcome would not complement the policy goals. Therefore, it is recommended that the federal government finance the compliance costs, distributing funding through special grants to utilities at levels determined by the proposed BID system. Such an approach would alleviate increased utility opposition and foster a cooperative environment.

Lastly, the costs associated with equipment are handled by GT-NEMS and determined from model outputs. Costs for installation, operations, maintenance, and eventual removal at the end of equipment life are estimated, showing reductions in cost under both benchmarking scenarios due to declining service demands. Investment costs for the major end uses are derived from the technology-specific details about energy service demand, technology cost per unit of energy service demand, and annual usage characteristics.

One shortcoming of this modeling approach is the weak linkage between demand for new equipment in one sector and the energy consumption related to its production in the industrial sector. The implication for our study is that the embedded emissions and energy consumption of producing equipment used in retrofits (i.e. replacing existing equipment before the end of equipment life with a more efficient unit) is not observable. The Commercial Building Consumption Energy Survey suggests that these types of retrofits affect less than 0.8% of buildings annually (EIA 2007), and a much smaller percentage of overall equipment purchases. Benchmarking is anticipated to increase the rate of retrofits. While GT-NEMS determines retrofit decisions, it does not separately report them, eliminating our ability to estimate “missing” energy consumption or emissions. However, for equipment purchased in new buildings or to replace expired equipment, the model is capable of more accurately assessing the associated industrial energy consumption and emissions.

Having tallied the benefits and costs of benchmarking to both the private and public sector, it is worthwhile to see how these compare from a societal perspective. In the first five years of the policy, compliance costs and the increases in criteria pollutant emissions are significant costs, but the commercial sector is showing net benefits of $11.6–13.6 billion compared to the UDR case. By 2035, cumulative energy savings, combined with the benefits of reduced emissions, exceed cumulative equipment and compliance costs by more than $40 billion. Once all new equipment has been retired, net benefits have grown to $61–64 billion using a 3% discount rate (table 2).

As is always the case with a benefit–cost analysis, there are costs and benefits that we are unable to capture, so it is crucial to recognize this effort as an incomplete best guess (Krutilla 1967). For example, the benefits of improved asset values for building owners and local governments, as well as numerous environmental benefits, are lacking from this analysis. A major benefit of benchmarking is the reduced transaction costs necessary to learn about building energy performance. Reducing these transaction costs is likely to be a large part of the policy rationale behind pursuing a policy like benchmarking, but methods to estimate the value of reduced transaction costs are currently lacking. Mandated disclosure laws would further reduce these costs.

Policymakers at all levels of government should consider that benchmarking could enable other policies, providing synergistic effects. For example, an analysis of the impact of a national commercial building code in the US projected 2035 energy savings of 0.94 Quads compared to the AEO 2011 reference case. The difference in commercial sector energy consumption between the reference case projection and the benchmarking scenarios for the same year is 1.33–1.35 Quads. If these policies were purely additive, then the expected result of having both policies in effect would be a reduction of 2.27–2.29 Quads. However, when modeled together, a slight synergy was shown in the model outputs, with reductions totaling 2.30–2.35 Quads. While these are small savings, benchmarking’s benefits could be expanded if matched wisely with other energy policies. Local policymakers have matched benchmarking with mandated disclosure laws and with financing programs, which may be synergistic pairings.

GT-NEMS shows that the benefits of this policy option are also sensitive to substitution effects and cross price

---

### Table 2. Benefit–cost analysis of commercial sector benchmarking

| Year | Energy expenditure savings | Value of avoided CO₂ | Value of avoided criteria pollutants | Lower equipment outlays | Total benefits | Compliance costs | Total costs | Net social benefits |
|------|-----------------------------|----------------------|-------------------------------------|-------------------------|---------------|-----------------|-------------|---------------------|
| 2020 | 6.3–2.8                     | -0.4–0.1             | 1.4–3.4                             | 6.4–5.4                 | 13.7–11.7     | 0.1             | 0.1         | 61.4–63.7          |
| 2035 | 28.3–22.0                   | 0.6–1.6              | 3.1–7.3                             | 18.0–21.7               | 50.0–52.6     | 0.1             | 0.1         | 61.4–63.7          |
| Total impact | 39.7–31.7                   | 0.9–2.3              | 3.0–8.2                             | 18.0–21.7               | 61.5–63.8     | 0.1             | 0.1         | 61.4–63.7          |

---

a Present value of costs and benefits were analyzed using a 3% discount rate. Values reported in table 2 are the Benchmarking 5% value followed by the Benchmarking 10% value.

b The total impact accounts for the energy savings and related benefits occurring throughout the lifetime of the commercial equipment, assuming an average lifetime of 20 years.

---

8 See section 4.7 and appendix A for more details: www.spp.gatech.edu/faculty/workingpapers/wp69.pdf.

9 www.spp.gatech.edu/faculty/workingpapers/wp71.pdf.
elastizities with respect to fuel choices in the utility sector. One sensitivity scenario we ran cut shale gas reserves by 50% from the reference case. The result was a dramatic expansion in the use of coal for electricity generation in the Midwestern region of the USA. In this scenario, energy prices and expenditures increased, and emissions benefits were largely eliminated. If the assumptions within the model about fuel price elastizities in the utility sector are correct, fuel switching may severely reduce the benefits of this policy option. The assumptions in this sensitivity case may be extreme, but they highlight an additional concern for policymakers—to ensure maximum benefits from benchmarking and other similar policy options, it may be necessary to take complementary actions that decrease the incentives to use coal. Such an action would prevent backsliding in emissions benefits.

Our review of the literature surrounding discount rates and the obvious impact of the updated discount rates scenario relative to the AEO 2011 reference case suggests that the Energy Information Administration should adjust the discount rates used in the NEMS model. Doing so would create a clearly measurable change in the technologies selected to meet energy demand in the commercial sector, impacting energy consumption, prices, and pollutant emissions projections from the model used to analyze national energy policy proposals.

Urban sustainability could be improved by this policy. Currently, the US lack data on the number and type of commercial buildings and their associated energy consumption in metropolitan statistical areas (MSAs)—hence our proposal of creating the BID system. However, GT-NEMS shows a 0.997 correlation between GDP and commercial floor space. 89% of 2010 US GDP occurred within US MSAs (US BEA 2010). Assuming GDP is a useful proxy for commercial sector benefits within MSAs, we would expect to see 89% of the benefits of this policy flow to MSAs. This would result in $55–57 billion in net societal benefits.

5. Summary

Efficient buildings contribute to sustainability problems in urban areas. Many improvements in commercial building energy efficiency could be driven by requiring utilities to submit building energy data to a uniform database accessible to building owners and tenants. Numerous other non-monetized advantages would also present themselves as a result of the proposed BID system.

With benchmarking, the market can see more clearly the advantage of superior energy performance, potentially spurring an end-user-driven shift. Building owners would have motivation to seek highly energy-efficient tenants, perhaps even incentivizing these tenants. Private organizations or government could grant recognition of quality energy management to specific tenants, further reducing transaction costs between tenants and building owners. This could enable market-based rewards for good energy practices by tenants, perhaps something similar to an ENERGY STAR program for tenants that allowed them to signal quality practices.

The impact of benchmarking shows a reduction in energy consumption of 160–180 TBtus over the UDR case in 2035, with nearly 90% of these savings benefitting metropolitan areas. It is estimated that the benefits of a benchmarking policy outweigh the costs, both to the private sector and society broadly. The net benefits of the policy are likely underestimated, due to the inability to fully monetize all potential benefit streams. For example, we do not incorporate the benefits of building envelope improvements such as low-emissivity windows and better insulation, and we do not monetize the full suite of environmental benefits from lower electricity consumption such as the health benefits of avoided mercury pollution and the ecosystem benefits of reduced coal mining. Overcoming some of the information barriers in the sector looks to be a worthy investment, mostly on the basis of the potential for energy savings. Opposition to benchmarking is likely to be grounded in concerns over principal–agent issues, tenant privacy, incurred costs (depending on policy design and implementation), and fear of the impact on the value of poor-performing buildings. Clarity from the federal government in policy design could substantially help cities address some of this opposition and improve the functionality of the marketplace.

Acknowledgments

Support for this research was provided by Oak Ridge National Laboratory and the Department of Energy’s Office of Policy and International Affairs. We are grateful for the input and guidance provided by the Energy Information Administration, in particular, Erin Boedecker. We would also like to thank Yeong Jae Kim and Paul Baer of the Climate and Energy Policy Laboratory at Georgia Tech for their inputs on the methodology. Lastly, we also wish to thank two anonymous reviewers from Environmental Research Letters for their constructive comments. Any remaining errors in this paper are the responsibility of the authors alone.

References

Akerlof G A 1970 The market for 'lemons': quality uncertainty and the market mechanism Q. J. Econ. 84 488–500
Alcott H and Greenstone M 2012 Is there an energy efficiency gap? J. Econ. Perspect. 26 3–28
Benzion U, Rapoport A and Yagil J 1989 Discount rates inferred from decisions: an experimental study Manag. Sci. 35 270–84
Brown M A 2001 Market failures and barriers as a basis for clean energy policies Energy Policy 29 1197–207
Brown M A, Baer P, Cox M and Kim Y J 2013 Evaluating the risks of alternative energy policies: a case study of industrial energy efficiency Energy Effic. at press (doi:10.1007/s12053-013-9196-8)
Brown M A, Wolfe A, Bordner R, Goett A, Kreitler V and Moe R 1996 Persistence of DSM Impacts: Methods, Applications, and Selected Findings EPRI TR-106193 (Oak Ridge, TN: Oak Ridge National Laboratory and Synergic Resources Corporation for the Department of Energy and Electric Power Research Institute)
Burr A C 2012 Benchmarking and Disclosure: State and Local Policy Design Guide and Sample Policy Language (Washington, DC: State and Local Energy Efficiency Action Network) (www1.eere.energy.gov/seeaction/pdfs/commercialbuildings_benchmarking_policy.pdf)
Sultan F and Winer R 1993 Time preferences for products and attributes and the adoption of technology-driven consumer durable innovations Environ. Psychol. 14 587–613

Thaler R 1981 Some empirical evidence on dynamic inconsistency Econ. Lett. 8 201–7

Train K 1985 Discount rates in consumers’ energy-related decisions: a review of the literature Energy I 1243–53

United Nations Framework Convention on Climate Change (UNFCCC) 1992 http://unfccc.int/resource/docs/convkp/conveng.pdf

US Bureau of Economic Analysis (BEA) 2010 Gross Domestic Product by Metro Area (www.bea.gov/iTable/iTable.cfm?ReqID=70&step=1&isuri=1&acrdn=2, retrieved 31 January 2013)

US Environmental Protection Agency (EPA) 2010 Technical Support Document: Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866 (www.epa.gov/otaq/climate/regulations/scc-tsd.pdf)

US Environmental Protection Agency (EPA) 2011 The Governments of United States and Canada Harmonize Approach to Save Energy in Commercial Buildings (http://yosemite.epa.gov/opa/admpress.nsf/d0cf6618525a9e9f85257359003f69/2d85476b30cc156785257943005aaf8?OpenDocument)

van Bergh J C J M 2013 Environmental and climate innovation: limitations, policies and prices Technol. Forecast. Soc. Change 80 11–23

Varey C and Kahneman D 1992 Experiences extended across time: evaluation of moments and episodes J. Behav. Decis. Mak. 5 169–85