Energy Research & Social Science 65 (2020) 101474

Contents lists available at ScienceDirect

Energy Research & Social Science

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

Perspective

Reducing the web's carbon footprint: Does improved electrical efficiency reduce webserver electricity use?

Ronald B. Mitchell[a], Richard York[b,*]

[a] Department of Political Science and Program in Environmental Studies, 1284 University of Oregon, Eugene OR 97403-1284 USA
[b] Department of Sociology and Program in Environmental Studies, 1291 University of Oregon, Eugene OR 97403-1291 USA

ARTICLE INFO

Keywords:
Jevons
rebound
efficiency
energy conservation
climate

ABSTRACT

Webservers are major consumers of electricity and, therefore, offer important opportunities for energy conservation. Server electrical efficiency has increased dramatically in recent years, suggesting that technological innovation can curtail electricity consumption. However, over a century ago, Jevons noted reasons to expect that technologies that increase the *efficiency* of resource use, paradoxically, can increase *consumption* of those resources. Here, we investigate the extent to which recent gains in server efficiency have translated into lower electricity use. We use the Standard Performance Evaluation Corporation's dataset on the electrical consumption and efficiency of over 600 server models tested between 2007 and 2019 to identify the extent to which improvements in electrical efficiency reduce server electricity use (watts) or increase server performance (operations per second). Our analysis estimates that server design innovation tends to favor the latter over the former. Electricity reductions typically equal one-quarter to one-third of a given improvement in electrical efficiency, suggesting a conservation-offsetting "rebound" but not one large enough to constitute a Jevons paradox in which efficiency actually increases resource use.

1. Do electrical efficiency innovations lead to webserver designs that reduce electricity use?

Global webservers and data centers consume substantial quantities of electricity, contributing significantly to carbon emissions if the electricity is generated from fossil fuels [1,2]. Technological innovations that increase energy efficiency seem to promise emission reductions [3]. We investigate how much such innovations deliver on this promise, examining webservers (servers) as a sector in which economic conditions favor investment, the technology context poses few barriers to innovation, and electricity costs encourage conservation. Indeed, although information technology (IT) companies have strong financial incentives to increase demand for web services, they also seek to minimize associated costs. Since electricity constitutes a major server operating cost and IT companies replace servers regularly, these considerations suggest that electricity-conserving innovation should face low financial and technological barriers while offering prompt returns on investment [4]. Indeed, server electrical efficiency increased ten-fold between 2007 and 2018 [4–6]. The debate over the environmental impact of improvements in resource efficiency range from optimistic claims that they can "restore the environment" [7] or at least reduce resource use proportionally to more skeptical claims that resource use typically declines less than proportionally ("rebounds") and may increase ("backfires" or "Jevons paradoxes") [8–12].

We focus on improvements in server electrical efficiency because of their substantive importance (data centers now account for 2 to 3% of global electricity use) and because they shed light on broader theoretical questions. Efforts to promote technological innovation dominate emission mitigation policy, not least because they are more politically attractive than limiting population, affluence, or consumption [13]. If constraining population and affluence growth (the non-technology components of the IPAT and Kaya identities) remain off the table [3,14], then it seems valuable to investigate whether technological innovations can deliver the 4% to 5% reductions in global emissions intensity (CO_2/\$GDP) that would be needed to avert the worst predicted climate impacts [15].

Since Jevons [8], scholars have argued that improving resource efficiency does not always generate corresponding reductions in resource use and, under some conditions, may increase it [16]. Automobiles illustrate the operative logic: improving fuel efficiency may fail to reduce fuel use because individuals may drive more (since cost per mile is cheaper) or spend freed-up income on other fuel-burning activities and, at national and international scales, because of indirect effects on global prices and industrial growth [10]. Pushing back, some

---

[*] Corresponding author.
E-mail addresses: rmitchel@uoregon.edu (R.B. Mitchell), rfyork@uoregon.edu (R. York).

https://doi.org/10.1016/j.erss.2020.101474
Received 10 November 2019; Received in revised form 9 February 2020; Accepted 13 February 2020

scholars contend that such effects are overstated and that energy-conserving innovations typically deliver 40% to 95% of their potential to reduce energy use [10]. Accurate estimates of the payoff of technological innovation matter, particularly for companies that operate data centers. Those companies increasingly seek to reduce energy use by improving the efficiency of individual servers, by optimizing the number of servers in data centers to most efficiently meet demand, and by reducing the indirect energy use of lighting, cooling, and the like [1].

Assessing the aggregate impact of improved server electrical efficiency on overall electricity use would require accounting for a broad array of direct and indirect effects, including global data on servers purchased, average load per server, replacement period, and the like for each server model. We skirt these complexities by focusing on the narrower question of how efforts to improve the electrical *efficiency* of servers have reduced their electricity *use*. In particular, we examine the tradeoffs that companies that design, develop, and sell servers make between server performance (measured as operations per second) and server electricity use.

To clarify, we are not analyzing how the choices by end-users of resource-conserving technologies influence aggregate resource consumption but, instead, how the design choices of engineering firms reflect the competing goals of increasing a technology's performance and reducing its resource use. In markets as fast-growing as data services, we should expect that economic incentives to meet growing consumer demand for data services may be so strong that firms design servers so that most increases in electrical efficiency show up as increases in server performance rather than decreases in server electricity use. This may mean that those buying servers find a market dominated by servers that are increasingly powerful rather than increasingly energy-conserving. Based on standardized benchmark data since 2007, we analyze trends in the relationship of server electricity efficiency and electricity use for each new model of server to infer whether server innovation favors performance or electricity conservation. Examining server performance and electricity use sheds light on the extent to which server designers channel improved electricity efficiency toward reducing electricity use or improving performance. Since we examine this relationship in a market where there are strong incentives for electricity conservation (because electricity dominates server operating costs), our estimates of the relationship of electrical efficiency to electricity use are likely to differ from that in economic sectors with weaker incentives for electricity conservation.

2. Methods: estimating the association of server performance, power, and efficiency

We use the Standard Performance Evaluation Corporation (SPEC) specpower_ssj2008 dataset of benchmark tests for 618 commercial web servers tested between November 2007 and April 2019 [6]. SPEC produces standardized, independent, lab-condition estimates of the power, performance, and electrical efficiency of unique server models. Their benchmark tests measure each model's performance (server-side Java operations or ssjops) and power use (watts) in 10% intervals from idle (0%) to maximum load (100%), as performance and power use vary by load. Electrical efficiency is the ratio of performance to power use (ssjops/watt). SPEC benchmarks of over 250 IBM and over 300 Oracle servers (as well as Sun and BEA servers) document increasing performance, power use, and electrical efficiency over time. Table 1 reports performance (ssjops), power use (watts), and electrical efficiency averaged across all IBM and Oracle servers tested in the first and last years of available data for each company. In aggregate, IBM and Oracle servers became far more electrically efficient (7-fold and 10-fold increases, respectively), but both companies' also increased power (3-fold and 9-fold) to deliver dramatic increases in performance (21-fold and 88-fold).

Table 1
Change in performance, power use, and electrical efficiency of IBM (2008 to 2015) and Oracle (2008 to 2019) servers.

	IBM		Change	Oracle		Change
Year	2008	2015		2008	2019	
Performance (ssjops)	205	4303	1999%	298	26,300	8725%
Power use (watts)	140	390	178%	240	2113	780%
Electrical efficiency (ssjops/watt)	1490	11,048	641%	1294	12,978	903%

Note: SPEC did not report benchmarks for any IBM servers after 2015.

To estimate the relationship between efficiency and total energy consumption analytically, using server-models as the unit of analysis, we regress electricity use on electrical efficiency using the natural logarithms of each while controlling for other factors. Using natural logarithms allows interpretation of coefficients as elasticities, in our case estimating the percentage change in electricity use for a one percent change in electrical efficiency. Our goal is to identify whether the co-variation of electricity use with electrical efficiency is more consistent with a proportionate reduction, rebound, or backfire hypothesis. To account for variation in electrical efficiency, electricity use, and their relationship due to server load, we conduct two sets of analyses of this relationship.

First, we estimate power use (watts) as a function of efficiency (ssjops per watt) with both averaged across all 10% load increments from 0% to 100%, effectively modeling servers as operating equal amounts of time at each load level (including when servers are idling at 0% load, using power but not performing any ssjops).

Second, we estimate this same relationship separately at each 10% load increment from 10% to 100% (but excluding 0% load as performance at that level is 0). We analyze each load level, since servers operate at quite varied load levels from cloud servers typically operating close to 60% load to “on-premises” servers usually operating between 10% and 20% loads, with the choice of any “average” load creating a risk of missing any load-dependent relationship [1,17,18].

Our units of analysis are individual server-models at average load in our first models and all specified loads in our second models. We regress the natural logarithm of power use (watts) on the natural logarithm of electricity efficiency (ssjops per watt), using dummy variables to control for chip manufacturer, vendor, and year of production. We used Stata 16 to estimate both ordinary least squares (reg) and robust (rreg) regression models. Our robust models exclude highly influential observations (those with Cook's distances over 1), and down-weight less influential cases based on iterations using first Huber weights until convergence, and iterate from that result using biweights until convergence. This specification sacrifices only 5% of an OLS model's efficiency when applied to data with normally distributed errors.

3. Results: server designs use improved electrical efficiency to increase performance more than to reduce electricity use

Table 2 reports our first analyses, in which we estimate both OLS and Robust models of power use averaged across load levels. Both models show higher electrical efficiency associated with lower electricity use, with the relationship being inelastic. Our OLS model estimates a coefficient for efficiency of -0.226, indicating a statistically significant (based on a two-tailed test with a 0.05 alpha level) association of a 1% increase in efficiency with a modest decrease of 0.226% in electricity use (95% confidence interval of -0.424 to -0.027). The corresponding coefficient in our Robust model is -0.330 (95% confidence interval of -0.485 to -0.176).

Linking our results back to our research question, recall that “rebounds” are the percentage difference between improvements in a technology's resource use *efficiency* and that technology's resource *use*. Server rebound estimates, then, are the complements of our electrical efficiency coefficients. Thus, when controlling for other factors and

Table 2
OLS and Robust (iterative Huber-biweight) elasticity models of total server electricity use. Total electricity use (ln(watts)) and efficiency (ln(ssjops per watt)) are averages of data across load levels from 0% to 100% in 10% increments.

	OLS Model	Robust Model
Independent variable	Coefficient (S.E.)	Coefficient (S.E.)
Server-side Java operations per watt	−0.226 (0.101)*	−0.330 (0.079)***
CPU (Intel omitted)		
AMD	0.201 (0.124)	0.326 (0.097)**
Vendor (Sun omitted)		
BEA	0.133 (0.481)	0.156 (0.374)
IBM	0.131 (0.448)	0.149 (0.348)
Oracle	−0.052 (0.432)	−0.053 (0.336)
Year (2019 omitted)		
2007	−2.247 (0.494)***	−2.739 (0.384)***
2008	−2.189 (0.342)***	−2.623 (0.266)***
2009	−1.884 (0.283)***	−2.382 (0.220)***
2010	−1.852 (0.259)***	−2.276 (0.201)***
2011	−1.578 (0.265)***	−2.300 (0.207)***
2012	−1.570 (0.196)***	−2.014 (0.153)***
2013	−1.881 (0.277)***	−2.380 (0.215)***
2014	−1.694 (0.395)***	−2.002 (0.307)***
2015	−2.022 (0.282)***	−2.157 (0.219)***
2016	−0.986 (0.259)***	−1.527 (0.201)***
2017	−0.999 (0.206)***	−0.988 (0.160)***
2018	−0.616 (0.184)**	−1.295 (0.143)***
y-intercept	8.727 (0.900)***	9.826 (0.700)***
R2	0.189	n/a
N	618	618

* $p<.05$ ** $p<.01$ *** $p<.001$ (two-tailed tests).

averaging across loads, we estimate improvements in electrical efficiency as having rebounds of 77.4% ((1 - 0.226) × 100) using OLS estimates and 67.0% ((1 - 0.330) × 100) using Robust estimates. Significant improvements in server electrical efficiency are not leading to corresponding reductions in energy use, as most of the improvements in efficiency are used to generate dramatic increases in performance (ssjops).

It is also important to note that over time, independent of other factors, server energy consumption has tended to increase, as reflected in the typically higher coefficient for the year-dummies closer to the present time. This is also reflected in the fact, presented in Table 1, that newer servers consume much more energy than earlier ones, even though being more efficient.

Table 3 reports our second set of analyses in which we regress identical models to those just described at each load level between 10% and 100%. To report multiple analyses concisely – and since our control variables had substantially similar coefficients to the “across loads” models of Table 2 – we present only our electrical efficiency coefficients for each load-specific model. Our load-specific estimates of reductions in electricity use (all else equal) range from 0.205% to 0.289% for each 1% improvement in electrical efficiency when estimated using OLS techniques (rebounds between 71.1% and 79.5%) and 0.300% to 0.374% when estimated using Robust techniques (rebounds between 62.6% and 70.0%). Both OLS and Robust techniques show that operating servers at higher loads sacrifices more of the potential benefits of efficiency innovations. This evidence suggests that, during server design, designers prioritize improving performance (increasing maximum server operations) over electricity conservation (reducing electricity use). That prioritization is not so strong as to eliminate reductions in electricity use (as Jevons might predict), but it does mean that resulting reductions prove to be only a fraction of those that efficiency improvements might lead one to expect. Once engineers improve electrical efficiency, server designers must choose to hold electricity use constant and increase performance, hold performance constant and reduce electricity use, or increase electricity use and dramatically increase performance. To illustrate, consider that IBM's 7.4-fold improvement in server efficiency (see Table 1) could have increased performance 7.4-fold without increasing electricity use, or could have held performance constant while using 86% less electricity, or, as it did, could increase performance 21-fold by increasing electricity 2.8-fold. Market incentives, not surprisingly, favor the last of these.

4. Conclusion: innovations in energy efficiency may fail to deliver on their energy-reducing promise

Our analyses suggest that server design choices prioritize targeting improved server electrical efficiency toward increasing performance rather than conserving electricity, with only about a quarter to a third of electrical efficiency improvements being realized as reductions in electricity consumption. Our findings stand in contrast to claims that most innovations deliver on their promised energy conservation [10]. Technology firms invest heavily in server innovation, seeking profits in a market characterized by dramatic growth in demand for web services and by high electricity costs. Server development firms must decide whether to prioritize server performance, electricity conservation, or a balance of the two. Our analysis suggests that, in practice, firms tend to design and market servers that dramatically increase server operations while reducing electricity consumption far less. As data service companies buy more and/or newer servers to meet increasing demand, our analysis suggests that electrical efficiency statistics are a bad proxy of electricity conservation. Indeed, servers available on the market are likely to offer far superior performance but deliver only around 30% of their nominal electrical efficiency improvements as reduced electricity costs and carbon emissions, independent of how efficiently data service companies actually operate those servers. Unfortunately, therefore, improvements in server electrical efficiency seem unlikely to alter the dramatic growth in electricity use that the data services sector has seen in the recent past.

Table 3
Association between electricity consumption and efficiency estimated across server load-levels from OLS and Robust Models.

	OLS Model	Robust Model
Load (% of maximum)	Coefficient (95% C.I.)	Coefficient (95% C.I.)
100	−0.205 (−0.422 – +0.012)	−0.300 (−0.467 – −0.132)
90	−0.217 (−0.429 – −0.004)	−0.315 (−0.482 – −0.149)
80	−0.208 (−0.415 – −0.001)	−0.319 (−0.481 – −0.157)
70	−0.219 (−0.422 – −0.017)	−0.332 (−0.490 – −0.175)
60	−0.241 (−0.439 – −0.042)	−0.353 (−0.508 – −0.199)
50	−0.256 (−0.451 – −0.060)	−0.363 (−0.515 – −0.211)
40	−0.260 (−0.453 – −0.066)	−0.358 (−0.510 – −0.207)
30	−0.266 (−0.456 – −0.076)	−0.362 (−0.510 – −0.214)
20	−0.270 (−0.456 – −0.084)	−0.363 (−0.505 – −0.220)
10	−0.289 (−0.469 – −0.109)	−0.374 (−0.510 – −0.238)
Combined	−0.226 (−0.424 – −0.027)	−0.330 (−0.485 – −0.176)

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare no conflicts of interest associated with this research.

Acknowledgments

The authors thank the Standard Performance Evaluation Corporation for making their dataset of SPECpower_ssj2008 benchmark results available for the analysis conducted in this article.

References

[1] A. Shehabi, S.J. Smith, D.A. Sartor, R.E. Brown, M. Herrlin, J.G. Koomey, E.R. Masanet, N. Horner, I.L. Azevedo, W. Lintner, United States data Center Energy Usage Report, Lawrence Berkeley National Laboratory, Berkeley, 2016http://eta-publications.lbl.gov/sites/default/files/lbnl-1005775_v2.pdf.

[2] N. Jones, How to stop data centres from gobbling up the world's electricity, Nature 561 (2018) 163–166 https://www.nature.com/articles/d41586-018-06610-y.

[3] Y. Kaya, K. Yokobori, Environment, energy, and economy: Strategies for Sustainable Development, United Nations University Press, Tokyo, 1997.

[4] J. Koomey, S. Berard, M. Sanchez, H. Wong, Implications of historical trends in the electrical efficiency of computing, IEEE Ann. Hist. Comput 33 (2011) 46–54 https://doi.org/10.1109/MAHC.2010.28.

[5] J. Witham, Achieving Data Center Energy Efficiency, Data Center Knowledge: Industry Perspectives, 2018, https://www.datacenterknowledge.com/industry-perspectives/achieving-data-center-energy-efficiency (Accessed 2 February 2020).

[6] SPECpower Committee, Standard Performance Evaluation Corporation, http://www.spec.org/power_ssj2008/results/power_ssj2008.html, (2019) (Accessed 3 July 2019).

[7] J.H. Ausubel, Can technology spare the earth? Am. Sci 84 (1996) 166–178 https://www.jstor.org/stable/29775636.

[8] W.S. Jevons, On the variation of prices and the value of the currency since, J. Stat. Soc. London 28 (1865) (1782) 294–320 https://www.jstor.org/stable/pdf/2338419.pdf.

[9] L.A. Greening, D.L. Greene, C. Difiglio, Energy efficiency and consumption - the rebound effect - a survey, Energy Policy 28 (2000) 389–401 https://doi.org/10.1016/S0301-4215(00)00021-5.

[10] K. Gillingham, M.J. Kotchen, D.S. Rapson, G. Wagner, The rebound effect is overplayed, Nature 493 (2013) 475 https://doi.org/10.1038/493475a.

[11] R. York, J.A. McGee, Understanding the Jevons paradox, Environ. Sociol 2 (2016) 77–87 https://doi.org/10.1080/23251042.2015.1106060.

[12] S. Sorrell, B. Gatersleben, A. Druckman, The limits of energy sufficiency: a review of the evidence for rebound effects and negative spillovers from behavioural change, Energy Res. Soc. Sci 64 (2020) Article 101439 https://doi.org/10.1016/j.erss.2020.101439.

[13] R.B. Mitchell, Technology is not enough: climate change, population, affluence, and consumption, J Environ. Dev 21 (2012) 33–36 https://doi.org/10.1177/1070496511435670.

[14] P.R. Ehrlich, J.P. Holdren, Impact of population growth, Science 171 (1971) 1212–1217 https://doi.org/10.1126/science.171.3977.1212.

[15] Intergovernmental Panel on Climate Change, Climate Change 2014: Mitigation of Climate Change (WG3 Contribution to the Fifth Assessment Report of the Intergovernmental Panel On Climate Change), New York, NY: Cambridge University Press, 2014https://www.ipcc.ch/site/assets/uploads/2018/02/ipcc_wg3_ar5_full.pdf (Accessed 2 February 2020).

[16] J.M. Polimeni, K. Mayumi, M. Giampietro, B. Alcott, The Jevons paradox and the Myth of Resource Efficiency Improvements, Earthscan, London, 2008.

[17] J. Barr, Cloud computing, Server utilization, and the Environment, AWS News Blog, 2015, https://aws.amazon.com/blogs/aws/cloud-computing-server-utilization-the-environment/ (Accessed 1 June 2019).

[18] A. Benik, The Sorry State of Server Utilization and the Impending Post-Hypervisor Era, GigaOm, 2013, https://gigaom.com/2013/11/30/the-sorry-state-of-server-utilization-and-the-impending-post-hypervisor-era/ (Accessed 1 June 2019).