Information Technology Investment and Carbon Intensity in the Era of Cloud Computing: A Cross-National Study

Jason Dedrick
Syracuse University, jdedrick@syr.edu

Jooho Kim
Baruch College, City University of New York, jooho.kim@baruch.cuny.edu

Jiyong Park
University of North Carolina at Greensboro, jiyong.park@uncg.edu

Information Technology Investment and Carbon Intensity in the Era of Cloud Computing: A Cross-National Study

Emergent Research Forum (ERF)

Jason Dedrick
Syracuse University
jdedrick@syr.edu

Jooho Kim
Baruch college, CUNY
jooho.kim@baruch.cuny.edu

Jiyong Park
University of North Carolina at Greensboro
jiyong.park@uncg.edu

Abstract

To tackle climate change in the digital economy, there has been increasing attention to the role of information technology (IT) investment in decoupling economic growth from greenhouse gas emissions, or reducing carbon intensity. This research examines the impact of cloud computing on carbon intensity and further scrutinizes how the advent of cloud computing has altered the relationship between IT capital and carbon intensity. We combine data on IT capital stock for 51 countries during 1995-2014 with a natural experiment involving the staggered launches of cloud data centers across countries. Our preliminary findings suggest that cloud on-ramps availability is positively associated with carbon intensity, whereas it negatively moderates the impact of IT capital on carbon intensity. Taken together, our preliminary evidence implies that the environmental impact of cloud computing may not be as adverse as conjectured if we factor in its indirect effect on making overall IT capital greener.

Keywords

Cloud computing, IT capital, sustainable development, sustainability, green IT, green IS.

Introduction

The Intergovernmental Panel on Climate Change's Sixth Assessment Report (IPCC 2021) states that human activity has already raised global temperatures and is partly responsible for extreme weather events occurring increasingly, which urges that drastic reductions in greenhouse gas emissions must be achieved to tackle climate change. For sustainable development—defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (Brundtland 1987, p. 8), reducing levels of carbon dioxide (CO2) emissions associated with economic output (what we refer to as carbon intensity of economic production) has been emphasized as a substantial and realistic pathway to addressing climate change while sustaining historic rates of economic growth.

By virtue of its central role as a growth engine in today's economy (Dedrick et al. 2013; Dewan and Kraemer 2000), information technology (IT) investment has been brought to the forefront of the sustainability debate, and many information systems (IS) scholars have paid attention to the role IT plays in environmental sustainability (Melville 2010; Watson et al. 2010). As a large portion of IT investment has shifted rapidly to cloud computing (Society for Information Management 2020), cloud computing has been considered as a double-edged sword for IT in pursuing environmental sustainability. On one hand, cloud computing could worsen the environment and elevate carbon intensity in the overall economy in that data centers that support and offer cloud services consume a massive amount of electricity (Shehabi et al. 2016). On the other hand, it could allow organizations to optimize IT resource utilization and make operations more energy-efficient (Park et al. 2017). Despite the increasing significance of cloud computing, there has been little cross-country evidence on the role of cloud computing in environmental sustainability and carbon intensity in particular. To bridge this important gap in the extant literature, this research aims to examine the impact of cloud computing on carbon intensity. In order to reveal its indirect role in making IT greener, we further scrutinize how the advent of cloud computing has altered the relationship between IT capital and carbon intensity.

For empirical analysis, we construct a unique panel dataset of national IT capital stock and CO2 emissions for 51 countries, covering developing to developed countries, during 1995-2014. To identify the impact of cloud computing, we exploit a natural experimental setting in which cloud on-ramps data centers that provide direct connectivity to a cloud provider have launched in different countries in different times. In particular, we consider cloud data centers around the world for global big 4 cloud providers (Amazon Web Services, Microsoft Azure, Google Cloud, and Alibaba Cloud). By incorporating the natural experiment involving the staggered cloud data center launches into a production function framework, we develop a model of cloud computing and carbon intensity and estimate the impact of cloud on-ramps availability on carbon intensity as well as its moderating effect on the relationship between IT capital and carbon intensity.

Our preliminary findings suggest that cloud on-ramps availability appears to be positively associated with carbon intensity, possibly owing to the direct effect from data centers. In contrast, we also find that cloud computing negatively moderates the impact of IT capital on carbon intensity, possibly reflecting its indirect effects for better utilization of IT resources and improved efficiency. Taken together, our preliminary evidence implies that the environmental impact of cloud computing may not be as adverse as conjectured if we factor in its indirect effect on making overall IT capital greener. We expect that this study can contribute to the literature on green IT by presenting empirical evidence that cloud computing can be a viable IT investment that drives economic growth while maintaining the historical level of carbon intensity.

Data and Methods

Data

Using multiple data sources, we consolidate a country-level panel dataset of cloud on-ramps availability, IT capital investment, and CO2 emissions per GDP for 51 countries from 1995 to 2014. To proxy the cloud service availability in a country, we use the data on the locations of cloud on-ramps of data centers (cloud data centers, in short) that provide a direct connectivity to the cloud services provided by one of global big 4 cloud providers, obtained from Telegeography. As the data on private cloud data centers are not available, note that this dataset covers only public cloud data centers. To determine when each data center began to provide cloud services, we also collect the information about the launch date of each provider's cloud service regions across countries since 2006 from their own websites and news media. Taking advantage of the staggered launches of cloud data centers, we define a cloud service availability indicator if there is at least one cloud data center available in a country in a given year. We also consider the number of cloud data centers to measure the intensity of cloud service penetration.

For nationwide IT capital investment, we obtain IT spending data on computer hardware from International Data Corporation, which has been used in cross-country studies in the literature (Dedrick et al. 2013). Specifically, the hardware spending mainly includes server/storage, enterprise networks and infrastructure, personal computing and mobile devices, and peripherals. As our main dependent variable, carbon intensity is measured as the ratio of CO2 emissions to GDP using the data from World Bank. Total capital stock and total labor hours are adopted from Penn World Table, version 9.1 (Feenstra et al. 2015). For control variables, we also consider manufacturing share of GDP and gasoline price from World Bank, percentage of households with electricity and density of road network from Euromonitor International, and percentage of electricity generated from renewable energy from International Energy Statistics.

The monetary values of capital stock and GDP are measured in international dollars which is based on Purchasing Power Parity (PPP) conversion. Following the prior work (Dedrick et al. 2013; Dewan and Kraemer 2000), we convert the quality-adjusted *flow* of IT spending from current dollars to constant dollars by using the computers and peripherals price index from the U.S. Bureau of Economic Analysis. To construct IT capital *stock*, IT spending is accumulated with 10-year rolling windows based on the depreciation rate following Dedrick et al. (2013). Non-IT capital stock is calculated by subtracting the values of IT stock from the values of total capital stock. Total labor hours are calculated by multiplying the number of persons engaged by the average annual working hours. Gasoline price data is measured every other year.

Variable	Mean	Std. Dev.	Description
Carbon Intensity	0.30	0.15	CO2 emitted per unit of GDP (Kg per 2011 PPP $)
Cloud Service Availability	0.09	0.28	Cloud service availability indicator, encoded as 1 if a cloud on-ramps data center is available or 0 otherwise
Cloud Data Centers	0.38	2.31	Cumulative number of cloud data centers available
IT Capital Stock	26,543	83,791	IT capital stock in 2011 PPP adjusted international dollars (in millions)
Non-IT Capital Stock	3,970,546	7,822,781	Total capital stock less than IT capital stock in 2011 PPP adjusted international dollars (in millions)
Total Labor Hours	47,517.42	11,949.25	Total labor hours, measured by multiplying employment by average working hours (in millions)
Summary statistics for all other control variables are omitted to meet page length requirements.			

Table 1. Summary Statistics

A Model of Cloud Computing and Carbon Intensity

We draw upon a production function framework to examine the relationship between IT capital and carbon intensity. In country i in year t, Equations (1) and (2) represent the production functions for dual outputs, GDP and CO2 emissions, where IT capital stock (IT_{it}), non-IT capital stock (K_{it}), and total labor hours (L_{it}) are considered as factor inputs:

$$GDP_{it} = f(IT_{it}, K_{it}, L_{it}),\ CO2_{it} = f(IT_{it}, K_{it}, L_{it}) \tag{1), (2}$$

Specifically, we choose the Cobb-Douglas (CD) production function which has been widely used in the IT productivity literature (Brynjolfsson and Hitt 1996; Dedrick et al. 2013; Dewan and Kraemer 2000) as well as in IT and sustainability studies (Añón Higón et al. 2017; Park et al. 2017). By dividing Equation (2) by Equation (1) in the CD form, we can obtain the following equation for carbon intensity ($CI_{it} \equiv \frac{CO2_{it}}{GDP_{it}}$):

$$CI_{it} = \alpha IT_{it}^{\beta_1} K_{it}^{\beta_2} L_{it}^{\beta_3} \tag{3}$$

By taking the logarithm to Equation (3), we re-write the equation in a linear form:

$$ln(CI)_{it} = A + \beta_1 ln(IT)_{it} + \beta_2 ln(K)_{it} + \beta_3 ln(L)_{it} \tag{4}$$

where the coefficients correspond to the elasticity of each input in terms of carbon intensity, and A captures the portion of carbon intensity not explained by the amount of inputs in production.

In order to incorporate the role cloud computing plays in affecting carbon intensity independently and in concert with IT capital, we extend the model in two ways. First, we decompose the portion of carbon intensity not explained by the amount of inputs in production (A) into cloud-related and cloud-nonrelated in an additive form (i.e., $A = \theta Cloud_{it} + A'$). Thus, the coefficient θ is expected to reflect the direct impact of cloud service availability on carbon intensity, irrespective of other production factors including IT capital. Second, following a varying coefficient model that has been used to estimate the indirect effects of one input on augmenting other inputs (Gong et al. 2016; Mittal and Nault 2009), we model the elasticity of IT in terms of carbon intensity (β_1) as a function of cloud computing in an additive form (i.e., $\beta_1 = \gamma Cloud_{it} + \beta'_1$). The coefficient γ is expected to capture the indirect role of cloud computing in moderating IT's impact on carbon intensity. It is noteworthy that cloud data center availability ($Cloud_{it}$) constitutes a natural experimental design, where the treatment group includes countries that have at least one cloud data center available in a given year. Thus, the extended model allows us to identify the causal effect of cloud on-ramps availability on carbon intensity and its moderating effect on the relationship between IT and carbon intensity. Our final model of cloud computing and carbon intensity is presented in Equation (5):

$$ln(CI)_{it} = A' + \theta Cloud_{it} + \beta'_1 ln(IT)_{it} + \gamma (Cloud)_{it} \times ln(IT)_{it} + \beta_2 ln(K)_{it} + \beta_3 ln(L)_{it} + \beta_4 ln(X)_{it} \tag{5}$$

Model Specification

The estimation model for carbon intensity includes not only production factors and cloud data centers, but also control variables (X_{it}), as shown in Equation (6). Given the longitudinal nature of our data, we also account for country fixed effects (σ_i) and year fixed effects (τ_t) to control for time-invariant county heterogeneities as well as global sustainability trends (e.g., Paris Agreement).

$$ln(CI)_{it} = \theta Cloud_{it} + \beta'_1 ln(IT)_{it} + \gamma(Cloud)_{it} \times ln(IT)_{it} + \beta_2 ln(K)_{it} + \beta_3 ln(L)_{it} + \beta_4 ln(X)_{it} + \sigma_i + \tau_t + \varepsilon_{it} \quad (6)$$

Our set of control variables include manufacturing share of GDP, electricity from renewable energy, household electricity penetration, gasoline price, and road density. We include the manufacturing share to control for economic structures, as the manufacturing sector is expected to have higher carbon emissions than primarily service sectors. With electricity generation being a major source of carbon emissions, the extent to which electricity is generated from renewable energy is also considered. In addition, we proxy residential electricity usage with the percentage of households with electricity. Given that transportation is one of the major contributors to carbon emissions, we also control for gasoline price and road density.

Preliminary Results

The results are reported in Table 2. The first and second columns represent estimation results using cloud service availability indicator and the number of cloud data centers as the cloud computing measure, respectively. For production inputs, we confirm that IT capital stock can reduce the reduction in carbon emissions per unit of GDP, whereas other inputs do not have significant effects on carbon intensity. This implies that the sustainability benefits of IT capital investment seem to outweigh it environmental costs.

More importantly, our findings suggest that the cloud service availability is positively associated with carbon intensity, indicating that the presence of cloud on-ramps data centers in a country increases the amount of carbon emissions per unit of economic outputs. On the other hand, the interaction term between IT capital and cloud computing is found to be negative and significant, indicating that cloud data centers negatively moderate the relationship between IT capital and carbon intensity. In other words, we interpret that cloud computing can make IT greener by allowing it to reduce further carbon emissions per GDP.

Dependent Variable: ln(CO2 emissions / GDP)	(1)	(2)
Cloud Service Availability	0.077***	
Cloud Service Availability * ln(IT)	-0.006***	
ln(Cloud Data Centers)		0.052***
ln(Cloud Data Centers) * ln(IT)		-0.004***
ln(IT)	-0.022***	-0.023***
ln(Non-IT)	-0.006	-0.007
ln(Labor)	-0.017	-0.016
Observations	976	976
R-squared	0.947	0.947
Country & Year Fixed Effects	YES	YES
Robust standard errors in parentheses and results for all other control variables are omitted to meet page length requirements; *** p<0.01, ** p<0.05, * p<0.1		

Table 2. Estimation Results

Conclusion and Future Work

IT has been emphasized as a primary means to establishing a win-win relationship between economic growth and environmental sustainability (Gholami et al. 2016). Given that more and more IT resources have moved on to cloud-based systems and infrastructure, cloud computing has become the new "electricity," which is a backbone of the digital economy (Brynjolfsson et al. 2010). To advance the body of knowledge on the role of IT in promoting environmental sustainability, this study conducted a cross-national analysis to reveal the impacts of cloud computing on carbon intensity across countries. To this end, we incorporated the natural experiment involving the staggered cloud data center launches into a production function framework, proposing a model of cloud computing and carbon intensity where cloud computing plays a role in affecting carbon intensity independently and in tandem with IT capital.

The preliminary evidence implies that the environmental impacts of cloud computing are multi-faceted: While the infrastructure that supports cloud services, such as cloud data centers, could deteriorate the environment, it can have an indirect impact on facilitating IT reducing carbon intensity in a more environment-friendly way. This highlights the significant role of cloud computing as a catalyst for driving productivity growth while maintaining the historical level of carbon intensity in the overall economy, to the extent that we can leverage cloud computing in ways that make overall IT capital more energy-efficient and sustainable.

To continue developing our study further, we plan to address a potential endogeneity and implement additional analysis to make our initial findings robust. First, we consider leveraging a matching scheme to compare the countries with available cloud services and the similar countries with no services in terms of country-specific characteristics. In this way, we are able to address a possible selection bias which may also affect the launch of cloud services. Second, it would be interesting to examine the impacts of cloud computing on energy consumption as the use of cloud computing consumes a portion of energy but at the same, it also enables operations more energy efficient. We expect that this study can provide important implications for both research and practice for environmental sustainability in the era of cloud computing.

References

Añón Higón, D., Gholami, R., and Shirazi, F. 2017. "ICT and Environmental Sustainability: A Global Perspective," *Telematics and Informatics* (34:4), Elsevier Ltd, pp. 85–95.

Brundtland, G. H. 1987. *Our Common Future*, Oxford: Oxford University Press.

Brynjolfsson, E., and Hitt, L. 1996. "Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending," *Management Science* (42:4), pp. 541–558.

Brynjolfsson, E., Hofmann, P., and Jordan, J. 2010. "Cloud Computing and Electricity: Beyond the Utility Model," *Communications of the ACM* (53:5), pp. 32–34.

Dedrick, J., Kraemer, K. L., and Shih, E. 2013. "Information Technology and Productivity in Developed and Developing Countries," *Journal of Management Information Systems* (30:1), pp. 97–122.

Dewan, S., and Kraemer, K. L. 2000. "Information Technology and Productivity: Evidence from Country-Level Data," *Management Science* (46:4), pp. 548–562.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. 2015. "The Next Generation of the Penn World Table," *American Economic Review* (105:10), American Economic Association, pp. 3150–3182.

Gholami, R., Watson, R. T., Hasan, H., Molla, A., and Bjørn-andersen, N. 2016. "Information Systems Solutions for Environmental Sustainability: How Can We Do More?," *Journal of the Association for Information Systems* (17:8), pp. 521–536.

Gong, F., Nault, B. R., and Rahman, M. S. 2016. "An Internet-Enabled Move to the Market in Logistics," *Information Systems Research* (27:2), pp. 440–452.

IPCC. 2021. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press.

Melville, N. P. 2010. "Information Systems Innovation for Environmental Sustainability," *MIS Quarterly* (34:1), pp. 1–21.

Mittal, N., and Nault, B. R. 2009. "Investments in Information Technology: Indirect Effects and Information Technology Intensity," *Information Systems Research* (20:1), pp. 140–154.

Park, J., Han, K., and Lee, B. 2017. "Energy Efficiency in the Cloud: An Empirical Analysis of Information Technology Outsourcing, Cloud Computing, and Energy," *Proceedings of International Conference on Information Systems 2017*.

Shehabi, A., Smith, S. J., Sartor, D. A., Brown, R. E., Herrlin, M., Koomey, J. G., Masanet, E. R., Horner, N., Azevedo, I. L., and Lintner, W. 2016. *United States Data Center Energy Usage Report*, Berkeley, California: Lawrence Berkeley National Laboratory.

Society for Information Management. 2020. "2020 SIM IT Trends," Mount Laurel, NJ.

Watson, R. T., Boudreau, M.-C., and Chen, A. J. 2010. "Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community," *MIS Quarterly* (34:1), pp. 23–38.