

Hedonic Pricing of Cloud Computing Services

Caesar Wu, Adel Nadjaran Toosi, Rajkumar Buyya, and Kotagiri Ramamohanarao

Abstract— Cloud service providers (CSP) and cloud consumers often need to forecast the cloud price to optimize their business strategy. However, pricing of cloud services is a challenging task due to its services complexity and dynamic nature of the ever-changing environment. Moreover, the cloud pricing based on consumers' willingness to pay (W2P) becomes even more challenging due to the subjectiveness of consumers' experiences and implicit values of some non-marketable features, such as burstable CPU, dedicated server, and cloud data center global footprints. Unfortunately, many existing pricing models often cannot support value-based pricing. In this paper, we propose a novel solution based on value-based pricing, which does not only consider how much does the service cost (or intrinsic values) to a CSP but also how much a customer is willing to pay (or extrinsic values) for the service. We demonstrate that the cloud extrinsic values would not only become one of the competitive advantages for CSPs to lead the cloud market but also increase the profit margin. Our approach is often referred to as a hedonic pricing model. We show that our model can capture the value of non-marketable features. This value is about 43.4% on average above the baseline, which is often ignored by many traditional cloud pricing models. We also show that Average Annual Growth Rate (AAGR) of Amazon Web Services' (AWS) is about -20.0% per annum between 2008 and 2017, ceteris paribus. In comparison with Moore's law (-50% per annum), it is at a far slower pace. We argue this value is Moore's law equivalent in the cloud. The primary goal of this research is to provide a less biased pricing model for cloud decision makers to develop their optimizing investment strategy.

Index Terms—Cloud Characteristics, Time Dummy, Extrinsic, Intrinsic Variables, Hedonic Pricing,

1 INTRODUCTION

PRICING cloud computing has always been a big challenge not only to many Cloud Service Providers (CSPs) but also to many cloud consumers because of the exponential growth of new service features or characteristics appear almost daily. Although pricing of cloud service delivery has often been drawn an analogy as a new public utility service [1], the underlying structure of cloud pricing is much more complicated than the traditional public utility services due to the rapid development of cloud technologies and multiple layers of service delivery models (or Anything as a Service: XaaS).

As Weinman [2] had noticed, the utility pricing or Pay-As-Your-Go (PAYG) is not the only possible model for the cloud. Some firms have begun to explore their marketing strategy to support "pay-what-you-like". He indicated one of the important lessons that CSPs should learn from other industries is that relying on innovative cloud services and technologies is not enough. CSP has to also come up with new pricing models for their services. This means that CSPs should "move beyond competition just on price to competition on pricing." The question of how to move beyond competition just on price leads to the idea of how to establish innovative pricing models for cloud services. The primary objective of cloud pricing model is to capture cloud service values along with its pricing variation as well as the dynamic nature of cloud technology development.

Our observation shows that the revenue growth of Amazon Web Services (AWS), one of the leading global CSPs, has a positive correlation with its cloud characteristics (See Fig. 1). This means various cloud service features, such as PAYG, burstable CPU, data center global footprint, GPU, one account for all location, etc. (Notice that the number of characteristics has been increased from just a few in 2006 to more than thousand in 2017 due to AWS' continuous cloud innovation [24]). The basic question is "Will the cloud characteristics impact its service price or customer willingness to pay (W2P)?" If so, what is the relationship between cloud characteristics and its service prices? Most importantly, how we can calculate or estimate the values of these characteristics. One of the solutions is a so-called hedonic model. The compelling reason to propose the hedonic model is that it can capture non-market values (extrinsic values) for the cloud ecosystem and evolutionary characteristics that either directly or indirectly impact on its service prices.

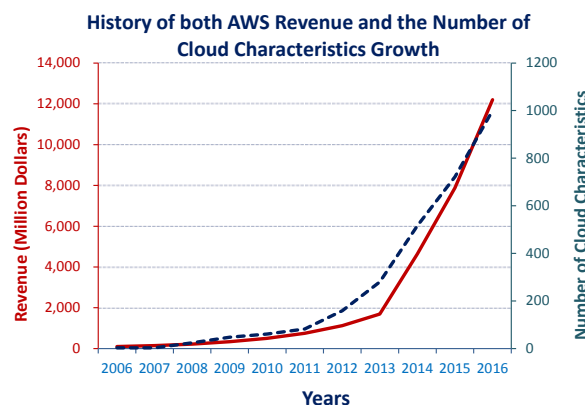


Fig. 1: AWS Revenue Expansion and Characteristics

Caesar Wu, Rajkumar Buyya, and Kotagiri Ramamohanarao are with Cloud Computing and Distributed Systems (CLOUDS) Lab, School of Computing and Information Systems, The University of Melbourne, Victoria 3010 Australia. A. N. Toosi is now with Faculty of Information Systems, Monash University, Australia
caesar.wu@computer.org, adel.n.toosi@monash.edu,
rbyyya@unimelb.edu.au, rkotagiri@unimelb.edu.au

Empirically, the basic premise or assumption of the hedonic function is that the product price difference is closely aligned with its characteristics (or features) variation. This means that if we can successfully establish a relationship between cloud service price differences with various cloud service characteristics, we will be able to estimate the price of cloud services accurately.

Another advantage to consider the hedonic approach is that the cloud price can be modeled by the regression analysis for the cloud service features along with its price variation over a period. In comparison with other methods, such as survey-based or contingent valuation [3] or Delphi [4] method, hedonic regression approach is quick and cost-effective if the chosen dataset is sufficiently large for the regression analysis. Moreover, it can be easily updated. It is a great fit for the cloud environment because of its ever-changing market conditions and rapid technological innovations.

Historically, the hedonic model has two different objectives. One is to predict the future price of goods or services that customers are willing to pay. This purpose of hedonic prediction is to help decision makers to make an optimized strategic decision. The other is a hedonic index, which is to establish a price ratio by comparing it with a price in a base period. The goal of the hedonic index is to monitor the price of either inflationary or deflationary, which is to verify what has happened in the past.

In this paper, we mainly focus on hedonic prediction or estimation. In order to achieve a better estimation, we introduce the concept of both intrinsic and extrinsic variables for hedonic function model, which is inspired by G.E. Moore [9] as a solution for cloud pricing problem. The intrinsic variables of cloud instances are defined as cloud resources, such as memory, CPU, storage, and network performance. They often appear as numerical variables. In contrast, the extrinsic variables can be anything from Burstable CPU, OpenStack compatible API, the global footprint of Cloud Data Center (DC), Mobile Application, vertical scaling without a reboot, to even one account for all locations. They are binary or categorical variables. In this paper, we propose a pricing model based on hedonic principles to capture the values of both intrinsic and extrinsic variables. This can help both CSPs and cloud consumers to estimate cloud prices more accurately. In addition, it explains the reasons why some market leaders of CSP do not only compete based on the price of intrinsic value but also on the price of an extrinsic one. Our proposed model will help many cloud decision makers to understand the price differentiation. We believe it will become a practical tool in a price modeling toolbox for many CSPs and it will also provide a pricing technique for many cloud consumers to select the right CSP for their application need. In summary, we have made the following contributions:

1) We articulate that cloud prices are dependent on both intrinsic and extrinsic variables according to the utility theory. We have also demonstrated how to compute these extrinsic values practically.

2) We construct a novel form of hedonic function for

cloud pricing, which consists of three explanatory variables: intrinsic, extrinsic and time dummy.

3) To the best of our knowledge, this is the first attempt to use the time dummy variable to correctly calculate Average Annual Growth Rate (AAGR) for cloud service. If we use AWS as a benchmark, it is about -20.0% per annum. This rate basically captures Moore's law behaviors. It is also the first time to comprehensively describe the context regarding the hedonic model for the cloud pricing. Moreover, it attempts towards pricing the cloud services with both panel and cross-sectional datasets.

4) We show that cloud price is declining but at a slower pace than what Moore's Law predicts for computing hardware [38]. We argue this slow pace is due to the non-marketable pricing values (by alone these features have no value) namely, extrinsic variables or characteristics.

5) We exhibit that our novelty pricing model can provide a good and simple solution to predict cloud price. We also show that a customer is paying more than a typical baseline service price (a standard configuration of cloud instance) on average for their business needs.

This study uses AWS data in 2014 to generate a simple hedonic regression model. Based on this model, we estimate a cloud price (by an average configuration of cloud instance) in 2017 and then compare with the real price in 2017. Our results show that the model can predict with an average accuracy of 87%. We use AWS 10 years unbalanced panel (longitudinal) data to construct a hedonic model with time dummy variables. According to this model, we can calculate the value of AAGR. By using AAGR, we can revise our estimation of cloud price. However, this price estimation does not take into a consideration of the extrinsic variables. In order to capture the extrinsic values, we develop a comprehensive hedonic model to calculate the value of each extrinsic characteristic based on the cross-sectional data of five CSPs. Finally, we update the estimated cloud price to achieve the much accurate results based on the particular type of workload.

The rest of paper is organized as follows: Section 2 provides the background information. Section 3 reviews related works and introduces the hedonic concept. It consists of three parts: the empirical work of hedonic analysis, the hedonic pricing model for computer prices and the hedonic model for cloud. Section 4 defines the hedonic function for cloud pricing. Section 5 provides a performance evaluation. Section 6 analyses the results with detailed discussion. The final section draws conclusions and indicates the future directions.

2 BACKGROUND

To set the background, we consider a scenario where a Chief Information Officer (CIO) of a firm needs to make a strategic investment decision whether to build their own private cloud (on-premises) or just migrate IT workloads to the cloud provider (off-premises, either private or public cloud infrastructure). Assume that the firm has its own on-premises IT infrastructure that still supports its existing business applications and the book value of IT

assets that cannot be written off for the next 12 ~ 36 months.

In this discussion, we ignore other issues such as types of IT workload, migration cost and system lifecycle management (SLCM) cost. The fundamental issue can then be boiled down to “how can we estimate the future market price of cloud services for the next 12~36 months?”. The logic behind this line of reasoning is if we can successfully predict or estimate the cloud price along with its service features (or cloud characteristics) that the business requires, we can select either building or buying or a hybrid solution for IT infrastructure. This means that if we can use the pricing model to predict the future price of cloud services accurately, we can help the CIO to develop a better IT investment strategy. However, cloud pricing modeling is much more complicated due to hedonic nature of many of its characteristics or features.

The term “hedonic” or “hedonism” was derived from a Cyrenaic parable in ancient Greek. It literally means “The Choice of Pleasure” [5] in contrast to “pain”. Economically, the connotation of hedonic is the meaning of gain, which is opposite to lose. From a cloud consumer perspective, hedonic values can be interpreted as some implicit benefits that are derived from specific cloud characteristics offered by a particular cloud service. Often, these service values are not only dependent on its intrinsic variables but also many extrinsic variables.

Traditionally, the price of any given cloud service (typically IaaS) is often determined by its cost components or required resources. It is referred to as cost-based pricing. With cost-based pricing, one of the disadvantages is that it cannot capture many cloud service characteristics. The other conventional approach to pricing is based on supply and demand, which is dependent on the market competition or the existing market conditions. We often call it market-based pricing. Unfortunately, many innovative services and cutting-edge technologies do not have an existing market to decide the price of goods. In contrast, the hedonic pricing model can overcome these issues to some extent because it can capture both intrinsic values (resource costs) and extrinsic values (service characteristics) and can estimate the missing or future price based on the existing market [46] [31]. This way, we can present a hedonic based pricing model to CIOs to estimate the future cloud price accurately. (Table 1 lists all the acronyms used in this paper.)

TABLE 1 Acronyms Used in this Paper

Acronym	Definition	Acronym	Definition
AAGR	Average Annual Growth Rate	I/O	Input / Output
API	Application Programming Interface	IaaS	Infrastructure as a Service
AWS	Amazon Web Services	OLS	Ordinary Least Square
CAGR	Compound Average Growth Rate	PAYG	Pay As You Go
CIO	Chief Information Officer	RAM	Random Access Memory
CSP	Cloud Service Providers	SLCM	System Lifecycle Management
EB3	Enterprise block Store	SSD	Solid State Drive
EC2	Elastic Compute Cloud	vCPU	Virtual Central Processing Unit
ECU	Elastic Compute Unit	VM	Virtual Machine
GCP	Google Cloud Platform	W2P	Willingness to Pay
GPU	Graphic Processing Unit	XaaS	Anything as a Service
HDD	Hard Disk Drive	YoY	Year on Year

3 RELATED WORK

The modern hedonic theory can be traced back to the founder of modern utilitarianism, Jeremy Bentham [6]. In Bentham’s view, the hedonic value is the sensational pleasure. He identifies seven main variables (IDCNPFEE) to calculate hedonic values. We show these values and their relevance to cloud computing values in Table 2.

TABLE 2
Bentham’s Seven Hedonic Variables Relevant to Cloud

Bentham’s Hedonic Variables (IDCNPFEE)	Bentham’s Definition	Value Range	Hedonic Values Relevant to Cloud
Intensity (I)	the amount of quality for pleasure or pain	0-10	Quality of Services
Duration (D)	how long the pleasure or pain will last	From minutes to weeks	Usage Time
Certainty (C)	the probability of the pleasure or pain will occur	0 – 100%	Certainty of price discount
Nearness (N)	how far off in the future	Now – Years	When discount price starts & ends
Purity (P)	how the decency of pleasure	0-100%	Dependent conditions to obtain cloud service
Fecundity (F)	the probability of reproducing the pleasure or other pleasures	0-100%	Probability of having discount price & more cloud service features continuously in future
Extent (E)	the number of people will be impacted by the pleasure	One or Many	Number of people can share the Cloud services

In contrast to Bentham’s view, John Stuart Mill [7] emphasized a higher level of intellectual happiness, which differs from Bentham’s pure hedonic value. He stated, “It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied.” In today’s a cloud pricing term, Mill’s hedonic value means to pursue a good result for business applications while Bentham’s hedonic value emphasizes to maximize the number of cloud service characteristics for the maximizing number of cloud customers.

G.E. Moore [8] divided the hedonic values into two kinds: intrinsic (or non-instrumental) and extrinsic (or instrumental) [9]. This is Moore’s significant innovation to hedonic theory. The concept of intrinsic value means that something is good or valuable on its own and the value is independent of others. For example, RAM, CPU, and storage can be considered as intrinsic values. In contrast, the extrinsic value is determined by the relationship to others, such as PAYG, burstable CPU and 24x7 supports, which are dependent on RAM and CPU. G.E. Moore’s concept of intrinsic and extrinsic values underpins our hedonic model.

3.1 The Empirical Hedonic Analysis

The empirical hedonic analysis had been adopted as early as the 1920’s. Zvi Griliches [10] generalized the hedonic regression model along with a semi-logarithmic form for the vehicles’ application in the 1960s. Griliches noticed many practical issues of the hedonic model analysis [31]. One of them was “How should the regression framework be expanded, what variables should be added to it, so as to keep the resulting estimates stable in facing of changing circumstances?”. He emphasized the essence of hedonic analysis is to estimate the “missing” prices or values due to quality or characteristics change, which influences our hedonic models for the cloud pricing.

3.2 Hedonic Model for Computer Price

In addition to the property and automobile applications, another popular application of the hedonic model is computer hardware, such as a mainframe, workstation, and personal computer. Since later 1970s, there have been countless publications regarding of hedonic price index of workstation and Personal Computer (PC). One of the earlier works was contributed by R. Michaels' [11]. He demonstrated how to establish a hedonic function with CPU performance, memory size, the speed of I/O, storage capacity and high-speed storage characteristics plus brand name and time dummy variables. Based on the regression analysis, the paper indicated that brand name had an insufficient impact on implicit prices and the deviation of quality-adjusted prices is smaller for the high-end computer equipment. The main conclusion of the paper was "observed price variations to be consistent with the economic theory" (value for money).

For the same topic, Cole et al. [12] presented and compared different PC hedonic price indexes with matched-model index and demonstrated that the traditional matched-model index is inadequate for PC product because the index excluded many new replacement PC models due to rapid technology improvement in PC industry. However, the authors did not give an explanation for why was the reason for PC price deflation.

Ernst R Berndt and Zvi Griliches [13] separated the price-decreasing problem into two issues: one is a price index and the other is the ratio of performance against price. They provided a variety of price indexes to serve the purpose of the deflation explanation for the microcomputer. The indexes were a kind of benchmark to measure "a technological frontier in the PC market" based on an unbalanced panel data. The paper reported testing results with various hedonic regression models, especially leveraging many dummy variables, such as year, vintage, process bit-length and age of PC. One of the apparent results was the PC price was decreasing although the quality of the PC was improving. Moreover, the authors noticed the issue of the parameters of the regression model has high variances and is unstable. The decision to select a set of variables from a pool of characteristics was arbitrary.

In contrast to many indexes oriented hedonic analysis, Rao et al. [14] mainly addressed the issue how to economically analyze information system (IS), which is how to acquire workstation hardware in the 1990s for many large organizations. The authors presented a hedonic function in the Box-Cox [15] transformation form (Equation 1) in order to extract a pattern between prices and the hardware characteristics.

$$\frac{y^\lambda - 1}{\lambda} = \beta_0 + \beta_1 \frac{x_1^\lambda - 1}{\lambda} + \beta_2 \frac{x_2^\lambda - 1}{\lambda} + \dots + \beta_n \frac{x_n^\lambda - 1}{\lambda} \quad (1)$$

where y is the workstation price, x_i^λ is the workstation i th characteristic. $\beta_1 \dots \beta_n$ are the coefficients, β_0 is the intercept value. λ is the transformation power parameter. The authors had noticed there were many difficulties to construct a hedonic function form, some of which still exist for determining cloud service pricing. These issues

include:

- 1) How to aggregate the characteristics of a good or service at a box level.
- 2) How to specify the characteristics in detail.
- 3) How to select each characteristic that can reflect both customers' values and resource costs.
- 4) How to handle the evolutionary characteristics.
- 5) How to trace and measure these characteristics at the box level.
- 6) How to apply the hedonic model or appropriate hedonic function at the box level.

In comparison with Rao's hedonic model, Pakes' paper [16] demonstrated a relatively easy way to construct a hedonic model from an index perspective. Parkes' empirical results show that PC's hedonic price had a sharp decline while the traditional matched model exhibited the near-zero values. According to Hulten [17], Pakes made three major contributions to the hedonic analysis:

- 1) The coefficients of hedonic function are not always fixed over time. Moreover, the sign of the coefficient is not necessary to be positive. In other words, some product's characteristics may have a negative impact on the overall hedonic values.
- 2) Two hedonic functions of the same product could be different from each other.
- 3) Each hedonic function is sufficient to make a quality judgment.

In addition, Pakes' theory of hedonic function is much easier to be grasped in comparison to other forms that have too many "restrictive assumptions." It can be directly derived from the theory of microeconomics [18], which the hedonic price reflects the price elasticity. Let (x_i, p_i) denote the characteristics and the price of the product "i" and Q_i is the quantity of demand of the product. Note Q_i is dependent on the price p_i and x_i . We can graphically show the product's price in Fig. 2.

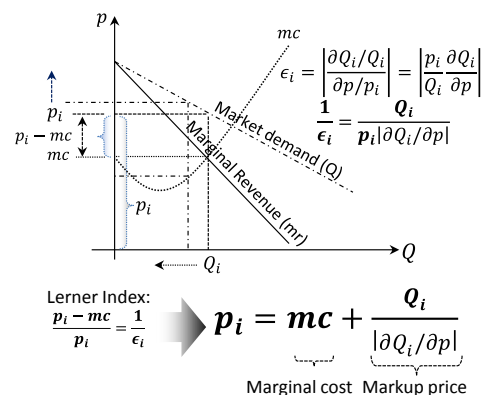


Fig.2: Theoretical Interpretation of Hedonic Price

From the Lerner index, we should have the Equation (2) [19]:

$$L_i = \frac{p_i - mc}{p_i} = \frac{1}{\epsilon_i} \quad (2)$$

where " p_i " is a market price of the product, and " mc " is the marginal cost for the product and ϵ_i is the elasticity. From microeconomics theory [18], the elasticity can also

be represented using equation (3):

$$\epsilon_i = \left| \frac{p_i}{Q_i} \frac{\partial Q_i}{\partial p} \right| \quad (3)$$

both Rosen and Halvorsen et al. [25] indicated that “The appropriate functional form for the hedonic equation cannot in general be specified on theoretical grounds”. This means that a practical solution to select a particular

TABLE 3 Common Regression Function Forms for Hedonic Analysis

Function form	Regression Equations	Lerner Index (Inverse of elasticity)	Hedonic Price
Linear	$p = \beta_0 + \sum_{i=1}^k \beta_i x_i + \epsilon_i$	$\frac{1}{\epsilon} = \beta_i \frac{x_i}{p}$	$\frac{\partial p}{\partial x_i} = \beta_i$
Quadratic	$p = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{i+1} x_i^2 + \epsilon_i$	$\frac{1}{\epsilon} = (\beta_i + 2\beta_{i+1}x_i) \frac{x_i}{p}$	$\frac{\partial p}{\partial x_i} = \beta_i + 2\beta_{i+1}x_i$
Cubic	$p = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{i+1} x_i^2 + \sum_{i=1}^k \beta_{i+2} x_i^3 + \epsilon_i$	$\frac{1}{\epsilon} = (\beta_i + 2\beta_{i+1}x_i + 3\beta_{i+2}x_i^2) \frac{x_i}{p}$	$\frac{\partial p}{\partial x_i} = \beta_i + 2\beta_{i+1}x_i + 3\beta_{i+2}x_i^2$
Semi-log Intrinsic & Extrinsic & Time Dummy	$\ln[p(X)] = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{j=1}^l \xi_j z_j + \sum_{t=1}^T \delta_t d_t + \epsilon$	$\frac{1}{\epsilon} = \left[\frac{\partial p}{\partial x_i} x_i, \frac{\partial p}{\partial z_j} z_j, \frac{\partial p}{\partial d_t} d_t \right]^T$	$\nabla p = \left[\frac{\partial p}{\partial x_i}, \frac{\partial p}{\partial z_j}, \frac{\partial p}{\partial d_t} \right]^T p$
Exponential or Semi-log	$p = \beta_0 \prod_{i=1}^k e^{\beta_i x_i + \epsilon_i}$, or $\ln(p) = \beta_0 + \sum_{i=1}^k \beta_i (x_i) + \epsilon_i$	$\frac{1}{\epsilon} = \beta_i x_i$	$\frac{\partial p}{\partial x_i} = \beta_i p$
Semi-log + Dummy Variable	$\ln[p(X)] = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{t=1}^T \delta_t d_t + \epsilon$	$\frac{1}{\epsilon} = \left[\frac{\partial p}{\partial x_i} x_i, \frac{\partial p}{\partial d_t} d_t \right]^T$	$\nabla p = \left[\frac{\partial p}{\partial x_i}, \frac{\partial p}{\partial d_t} \right]^T p$
Power or Double log (or Cobb-Douglas model)	$p = \beta_0 \prod_{i=1}^k x_i^{\beta_i} \times \epsilon_i$ or $\ln(p) = \ln(\beta_0) + \sum_{i=1}^k \beta_i \ln(x_i) + \epsilon_i$	$1/\epsilon = \beta_i$	$\frac{\partial p_i}{\partial x_i} = \frac{\beta_i}{x_i} p$
Logarithmic	$p = \beta_0 + \sum_{i=1}^k \beta_i \ln(x_i) + \epsilon_i$	$\frac{1}{\epsilon} = \frac{\beta_i}{p}$	$\frac{\partial p}{\partial x_i} = \frac{\beta_i}{x_i}$

From (2) and (3), we have the equation (4):

$$p_i = mc + \frac{Q_i}{|\partial Q_i / \partial p|} \quad (4)$$

Subsequently, the hedonic function can be written as:

$$h(x_i) \equiv E[p_i | x_i] = E[mc | x_i] + E \left[\frac{Q_i}{|\partial Q_i / \partial p|} \middle| x_i \right] \quad (5)$$

This equation consists of both marginal cost (first term) and markup price (second term). The first term is also dependent on customers' demand. The challenging question is how the first and the second terms interact with each other and how to calculate the market price.

Fortunately, we can use the regression analysis as an empirical tool to estimate the relationship between the response variable (cloud price) and explanatory variables (cloud characteristics). This is the basic idea of the hedonic approach. The idea of predicting hedonic price has been consolidated by Haas, Court, and Waugh and theorized by Lancaster [20] and Rosen [21]. According to Brachinger [22], the functional relationship of hedonic prices can be defined as:

$$MWTP = \frac{\partial p}{\partial x_i}(x) = \frac{\partial h}{\partial x_i}(x), (i = 1 \cdots k) \quad (6)$$

where “MWTP” is the marginal willingness to pay, “p(x)” is the price function, “h(x)” is the hedonic function. x_i is the characteristic of a product. Practically, there are four common types of hedonic forms (linear, semi-log, log-log or Cobb-Douglas and logarithmic, see Table 3). But, as

function form is really dependent on a dataset in hand, which is to examine which function form to be goodness-of-fit with a collected dataset. Halvorsen proposed a statistical procedure to select a functional form with a Box-Cox methodology that is basically to use likelihood ratio to examine the appropriateness of the alternative functional forms. However, Cassel et al. [26] argued that Box-Cox transformation is inadequate for the purpose of predicting hedonic prices because:

- 1) It is not necessary to increase the accuracy of price prediction. In fact, it could lead to a poorly estimated result, which it had been demonstrated by Rao [14].
- 2) The collected data may contain some negative values, but the traditional Box-Cos function does not allow any negative values because any negative number raised to non-integer real power would become imaginary.
- 3) Because the mean predicted the value of the untransformed dependent variable is not necessary to be equal to the estimated mean that has been transformed. As a result, the nonlinear transformation will introduce a bias for the untransformed variable.

Overall, the nonlinear transformation results would be challenging to be explained.

3.3 Hedonic Model for Cloud Price

To the best our knowledge, only limited studies of hedonic analysis had been conducted for cloud pricing, alt-

though the hedonic model has been widely applied in other industries, such as real estate, automobile, hotel, airline, and recreation. El Kihal et al. [27] were among the first presented a simple hedonic analysis regarding Infrastructure as a Service (IaaS) clouds. The result of the hedonic analysis is not compelling because the adjusted R-squared was 43% (IBM). Nevertheless, they initiated the hedonic model for further study of cloud prices. Mitropoulou et al. [41] [45] provide a hedonic price index for cloud price comparison purposes among 23 CSPs.

In summary, previous studies left a large gap of hedonic modeling for the cloud pricing in term of exploring different alternative hedonic forms, reducing regression errors, increasing R-squared values and adding practical values for cloud decision makers. In this paper, we show how to overcome many of these issues.

4 HEDONIC FUNCTION FOR CLOUD PRICING

4.1 Hedonic Function

By the extension of previous research for cloud prices, we first define the simplest hedonic function form of linear regression using OLS (Ordinary Least Square) method for our initial test. It can be directly interpreted as the mean coefficient values multiplied by independent variables (a vector of cloud characteristics) plus an error term:

$$p(X) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \varepsilon \quad (7)$$

where $X = (x_1, x_2, \dots, x_k)$, x_i are independent variables and also a vector to represent different cloud characteristics, such as RAM, CPU core, virtual CPU, storage size and network bandwidth. The “ k ” is the number of cloud characteristics. “ p ” is a dependent variable to represent cloud instance price, which can be observed from CSP’s web price catalog. Both independent and dependent variables are numerical values. β_i is the linear coefficient and β_0 is the interception point of the linear equation and ε is the error term or noise. The issues of the linear model are:

- 1) It may create substantial errors because of underfitting. The previous analytic results [27] demonstrated the R-squared value could be as lower as 46%.
- 2) This model cannot capture the price change due to time variation for the unbalanced panel data. In other words, it is impossible to measure the price change along with the temporal domain.
- 3) This model also ignored extrinsic features.
- 4) Moreover, some of the cloud characteristics provided by each individual CSP, such as dedicated server, burstable CPU, and OpenStack API cannot be captured due to the binary nature of these features. Therefore, it could lead to inaccurate price estimation.

In order to overcome these issues, we have to develop much sophisticated hedonic function forms to minimize the regression errors based on the collected datasets.

4.2 New Hedonic Function Form

One of the solutions to minimize regression error due to time dependency is to add another independent varia-

ble for the unbalanced panel data, namely time dummy or indicator variables to the OLS equation. This variable can capture the chronological influence of the cloud prices. As a result, the equation (7) would become the equation (8) as:

$$p(X) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{t=1}^T \delta_t d_t + \varepsilon \quad (8)$$

$$X = \langle x_1 \dots x_k, d_1 \dots d_T \rangle, d_t \in \{0,1\}, \sum_{t=0}^T d_t = 1$$

Here, d_t is the time dummy variable. Often, the unit of T is the number of years. δ_t is the coefficient value. ε is the error term that generates by both numerical and binary variables.

Furthermore, in order to capture the categorical variable of cloud service characteristics we separate all cloud characteristics into two categories, namely intrinsic and extrinsic characteristics. The intrinsic characteristics are closely associated with cloud infrastructure cost. They often appear to be the continuous variables. In contrast, the extrinsic characteristics are the binary variable. It means that CSPs can either support or not for a particular cloud instance. These service features will only add values to the customers when some intrinsic cloud characteristics are enabled. Let alone they often have no instrumental values to customers. Subsequently, we can develop further the equation (9) to be as following:

$$p(X) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{j=1}^l \xi_j z_j + \sum_{t=0}^T \delta_t d_t + \varepsilon, \quad (9)$$

$$X = \langle x_1 \dots x_k, z_1 \dots z_l, d_1 \dots d_T \rangle$$

$$z_j \in \{0,1\}, d_t \in \{0,1\}, \sum_{t=0}^T d_t = 1$$

where z_j is the binary variable (In general, z_j can be a categorical variable.) that represents extrinsic cloud characteristics j and “ l ” is the number of the extrinsic characteristics. ξ_j is the coefficient of the binary variable. ε is the term of combination errors for both intrinsic and extrinsic characteristics plus time dummy variable. If we take the derivative of equation 9, we should have a vector of derivatives.

$$\nabla p(X) = [\beta_1 \beta_2 \dots \beta_k \ \xi_1 \xi_2 \dots \xi_l \ d_1 d_2 \dots d_T]^T \quad (10)$$

Intuitively, the extrinsic cloud characteristics are similar to spatial fixed effects in the property data application. As Kuminoff et al. [34] suggested adopting a combination of spatial fixed effects, quasi-experimental identification and temporal controls would provide an unbiased result because of many unobserved characteristics. If all characteristics are explicit, Cropper et al. [35] suggested that linear and quadratic Box-Cox forms would produce the best results.

However, Triplett [28],[29] Griliches [30] and Gordon [32] indicated that the semi-log form has frequently emerged as “best” in hedonic function form tests. As a result, we can rewrite the equation (9) as a semi-log form. It can handle the substantial prices variation of cloud in-

stances for a long-time period.

$$\ln[p(X)] = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{j=1}^l \xi_j z_j + \sum_{t=1}^T \delta_t d_t + \varepsilon, \quad (11)$$

$$X = \langle x_1 \cdots x_k, z_1 \cdots z_l, d_1 \cdots d_T \rangle$$

$$z_j \in \{0,1\}, d_t \in \{0,1\}, \sum_{t=1}^T d_t = 1$$

Transformations will make sense if the dataset has the following features [33]:

- 1) The variance of the errors is unequal or heteroscedasticity.
- 2) The ratio between max and min is greater than 5.
- 3) The scatterplot of dependent and independent variables is curved.
- 4) The data points are skewed, which the data has a long right tail.
- 5) All values are positive.

Generally, the transformation will consider the response variable (cloud instance price) first and then both explanatory and response variables. Another solution to reduce the regression errors is to develop a polynomial regression formula, which is to add multiple high order terms for the independent variables if the collected dataset shows that the relationship between dependent variable (cloud price) and independent variables (cloud service characteristics) is not linear.

We considered a variety of hedonic function forms, as shown in Table 3, to minimize estimated errors.

5 PERFORMANCE EVALUATION

5.1 Datasets and Assumptions

5.1.1 AWS Panel Data

The AWS panel data comes from two sources: 1) internet archive [23], 2) Amazon annual reports [24]. The data was recorded or sorted based on the time sequence that AWS released a new service catalog every time.

Although Amazon started its AWS business as early as in 2006, AWS had a limited number of characteristics for its cloud services. Most of them belonged to intrinsic characteristics. In fact, AWS did not offer the cloud services to the general public until 2007. Consequently, the cut off time for the panel data test began in 2008. In the beginning, AWS offered only four instances to the public. Later, AWS gradually added more types of cloud instances to its service catalog. Each instance has a particular configuration, Application Programming Interface (API) name, and its price tag. After 2013, AWS superseded some previous generation of Elastic Compute Cloud (EC2) and replaced with a current generation of instances.

AWS pricing catalog is evolving from time to time due to the innovation of cloud technologies and pricing models. Some intrinsic variables are mixed with numerical and categorical values. Moreover, AWS sometime changes its CPU measurement in response to the cloud market competition [36]. Therefore, we have made the following assumptions in order to simplify the AWS panel dataset:

1) For optimized instances, AWS uses HDD for d-serial instance and Non-Volatile Memory (NVMe) SSD for i3-serial instance. The rest of the instances are either SSD or EBS only. In order to simplify the calculation, we assume these different characteristics of instance storage to be the same as HDD, in term of unit cost.

However, the prices of SSD, NVMe SSD and HDD are different in the market. So, this assumption will contribute the certain price variations in our analysis.

2) The networking performance in the AWS catalog is mixed with numerical and categorical variables. As a result, we unified all variables with the same numerical unit, which the category of very low is equal to "1", "low" is equal to "2", "low to moderate" is equal to "3", "Moderate" is equal to "4", "high" is equal to "5" and "Up to 10 GBits" is equal to "6". This assumption might also create some errors because "1" might not be necessarily equivalent to 0.1 GBits link.

3) AWS has two different types of instance prices for two operation systems: Linux and Windows. For this paper, we only use Linux price on-demand. The price ratio of Linux and Windows is ranging between 1.00 and 2.05. It is dependent on the size or capacity of the instance. AWS provides customer long-term subscription discount if cloud customers have a long-term commitment, which is so-called "reserved price". This is another aspect of the problem that will be dealt with separately in other research.

5.1.2 Computer Hardware Data

In order to make a price comparison between cloud service (IaaS) and general computer hardware with the influence of Moore's law, we include the general computer hardware market data of CPU, GPU, SSD, flash memory, storage, Hard Disk Drive (HDD) [47]. There have been some other works [39] [40] [44] for cloud price comparison, but they only focused on the cloud compute or storage resources in isolation. Our study takes into account all dependent variables.

5.1.3 Cross-Sectional Data

The cloud characteristics are released by different CSPs almost daily. Capturing all cloud characteristics is impossible. Due to the limitation of the dataset, we only have the total of 55 extrinsic cloud characteristics. Among them, 48 are considered to be the typical cloud characteristics, such as Pay-As-You-Go, Web interface, API, and Free-Transfer-In, which nearly all CSPs provide these common characteristics for their service. As a result, they have become the baseline of extrinsic cloud characteristics. In this study, we limit the number of extrinsic characteristics for our analysis because some of the extrinsic characteristics are insignificant (p-value > 0.05) such as vertical scaling without a reboot, OpenStack-compatible API and backup snapshot due to a limited number of data points. Furthermore, each CSP started the cloud business at the different time. Some of them just launched the cloud business recently.

5.1.4 All Cloud Instances of Five Leading CSPs

According to the latest Gartner's magic quadrant market report for the public cloud of IaaS [37] AWS, Microsoft and Google are the market leaders and Rackspace is one of the challenges and closely follows these three (see Table 4). Linode is one of the leading competitors in US IaaS market.

Note that some of the extrinsic characteristics add extra costs for the cloud services, for example, 10-node Hadoop

TABLE 4
Five Leading Public Cloud Service Providers

Name of CSP	No of Instances prices	No of baseline Characteristics	No of Host Domains (30-Jan-17)	No of Host Domains (30-Mar-17)
AWS	76	48	948,207	1,015,002
Microsoft Azure	69	48	142,854	149,175
Google Cloud Platform	21	48	599,846	630,117
Rackspace	19	48	504,624	487,827
Linode	14	48	210,106	220,717

cluster would have the extra cost of 0.15/per hour. In order to make a fair and horizontal comparison among different CSPs, we only track some extrinsic cloud characteristics across the board, which have no extra charge for an instance price. We assume CSPs do not charge an extra price for their baseline service configuration in their price catalog. These extrinsic characteristics of cloud service often have the binary values, which are either 0 or 1.

5.2 Test Design, Roadmap and Results

We start with the 1st test that is designed to analyze the cloud instance price. We adopt AWS cloud catalog dataset for the 2014 year (see Section 5.2.1). It is a simple OLS test. The purpose of this test is to examine the relationship between cloud instance prices (on-demand price for Linux OS) and its intrinsic characteristics. According to AWS, ECU (virtual server) resource is equivalent to CPU capacity of one 1.0-1.2GHz 2007 Opteron or 2007 Xeon processor. However, AWS has quietly adopted the unit of vCPU measurement in 2014. Each vCPU would correspond to a hyperthread of Intel Xeon core (clock speed) except t-serial instances. The purpose of a hyperthread technology is to increase CPU performance by sharing the computational workload among multiple cores. The value of ECU usually is higher than vCPU except for t-serial instances.

The second test consists of time dummy variables based on AWS unbalanced panel dataset between 2008 and 2017 (as discussed in Section 5.2.2). This test is an extension of the OLS. However, we add the second and third order polynomial terms into the linear equation in order to increase R-squared and reduce p-values.

The last test is to compare cloud instance prices among five different CSPs based on the cross-sectional dataset in 2017. This test is designed to add the extrinsic variables into the hedonic function form. It is to analyze the impact of cloud extrinsic characteristics on the price of baseline instance configuration (as discussed in Section 5.2.3). A roadmap of these three tests is illustrated in Fig 3, which

illustrates how we demonstrate the cloud extrinsic characteristics. In doing so, we report some performance and decision parameters of the preliminary models and then the full cross-sectional data for the final model.

We used R and R Studio to implement both panel data

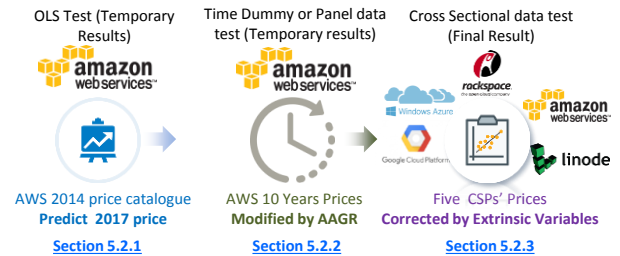


Fig.3: Simple Roadmap of Three Tests

and cross-sectional data regression analysis.

5.2.1 AWS Instance Price Test

According to our test design, we construct a simple linear regression model between Linux on-demand price and six explanatory intrinsic variables and then have a normality test and residual plots with instance price for the dataset.

Both R-squared and adjusted R-squared values are about 0.82-0.83 so that OLS only explains 82% of data points. Based on both the normality test and residual plots, we can see two outlier data points. These outlier points may cause regression errors. If we excluded these two points, the R-squared values could be increased.

We also notice that the coefficient of both vCPU and Bit (Architecture 32-bit or 64-bit) is negative. This may also be triggered by the regression errors. By excluding the new large GPU instances or outlier data points, we can improve the residual values of this OLS dramatically. The R-squared values are lifted to about 93% (Table 5).

Note: “**” means a significant code of p-value, “***” = p<0.001, “***” = p<0.01, “*” = p < 0.05

Furthermore, the p-values of ECU, CPU, and Bit be-

TABLE 5
The Linear Form of Hedonic Function for 2014

Coefficients	Estimated β	Std. Error	t-value	Pr(> t)
Intercept	-0.3377	1.06E-01	-3.176	0.00186 **
RAM	0.0049	3.98E-04	12.326	< 2e-16 ***
VCPU	0.0181	5.94E-03	3.044	0.00283 **
Storage	0.00005	8.42E-06	5.897	3.01e-08 ***
network performance	0.1755	2.67E-02	6.586	1.02e-09 ***

Residual standard error: 0.5949 on 130 degrees of freedom
Multiple R-squared: 0.9273, Adjusted R-squared: 0.9251
F-statistic: 414.8 on 4 and 130 DF, p-value: < 2.2e-16

come insignificant. The test has proved the Gartner's claim [36], which AWS quietly shifted from ECU to vCPU. Therefore, we can safely exclude ECU and CPU as independent variables with limited impact on R-squared and adjusted R-squared values. By extracting hidden values from the intercept (or beta zero), we can transform it to a semi-log form and add polynomial higher order terms into the OLS equation (as shown in Table 6).

TABLE 6
The Semi-log Form of Hedonic Function for 2014

Coefficients	Estimated β	Std. Error	t-value	Pr(> t)	EXP(β_i)
Intercept	-5.04E+00	2.14E-01	-23.483	< 2e-16 ***	0.01
RAM	8.71E-03	1.64E-03	5.307	4.82e-07 ***	1.01
RAM^2	-1.02E-05	2.32E-06	-4.402	2.25e-05 ***	1.00
RAM^3	3.81E-09	8.07E-10	4.726	5.98e-06 ***	1.00
VCPU	7.87E-02	1.23E-02	6.414	2.55e-09 ***	1.08
VCPU^2	-6.83E-04	2.14E-04	-3.194	0.00177 **	1.00
storage	1.99E-05	7.13E-06	2.792	0.00605 **	1.00
network performance	1.28E+00	9.45E-02	13.571	< 2e-16 ***	3.60
network performance^2	-9.40E-02	7.29E-03	-12.882	< 2e-16 ***	0.91
EBS.O	-1.22E-04	4.18E-05	-2.913	0.00423 **	1.00

Residual standard error: 0.4793 on 125 degrees of freedom
Multiple R-squared: 0.9118, Adjusted R-squared: 0.9055
F-statistic: 143.6 on 9 and 125 Df, p-value: < 2.2e-16

One issue with the linear form is that the absolute value of the intercept β_0 (-0.338) is the highest in comparison with other β_i (or hedonic) values. A practical interpretation of this negative β_0 is that AWS would pay customers upfront for on-demand instance which is not the case. One of the reasons for the higher absolute β_0 value is there are other hidden variables within β_0 . With the semi-log form, the β_0 value is down from 0.338 to 0.0064. Although both R-squared values slightly decline by about 2%, the β_0 value is reduced nearly 53 folds. On the other hand, the result of the semi-log form is difficult to be interpreted because of the higher order polynomial terms with the negative β values. The model becomes quite sensitive for the large instance configuration, especially for the characteristics of RAM and network. One of the reasons is that AWS may insert a volume discount mechanism for the large instances. The other possible reason is AWS does not give the resource-level permission to reboot, start, delete, detach EBS volume etc. for cloud customers to specify a resource in every instance action, in order to maintain control of its cloud infrastructure resource pool.

As noted in AWS 2017 catalog, AWS offers a wide variety of configurations for its computing instances such as cc1.4xlarge (cluster compute quadruple extra-large VM), cg1.4xlarge (GPU VM), and m1.small (general purpose small resource VM). To predict a cloud price of an average configuration resource in AWS 2017 catalog, we used the m4.10xlarge instance, which is one of the general purpose instances and provides a balance of computing memory and network resources. It is designed to support different computing environments such as web applications or line of business or LoB (The letter “m” stands for “general purpose,” “4.10” means the size of computing and network resources, “xlarge” stands for extra-large.) The detail configuration of this instance is RAM=160, ECU=124.5, vCPU=40, CPU=3.112, storage=0, Network Performance=10, EBS.O= 4000.

Based on this configuration, we can predict the price of the m4.10xlarge instance as \$2.925 (linear form) or \$2.961 (semi-log form). The real price for m4.10xlarge instance is \$2.155 (see Table 7). Although this prediction value is within 95% of the confidence interval, the predicted fitted value is about 36% higher than the real price value, and the price range between low and upper bound is high,

TABLE 7
Predicting Price of a Cloud Instance with m4.10xlarge Instance Configuration

	With 95% confidence interval	Fitted Value	Real price	Price difference Δp	Accuracy	Lower	Upper
Predicted Value by Linear form		\$2.925	2.155	\$0.77	64.3%	1.716	4.134
Predicted value by semi-log form		\$2.961	2.155	\$0.806	62.9%	1.110	7.898

but the linear form is slightly better than the semi-log. This might be due to many factors, such as different function forms, sample size, and skew dataset. Moreover, we have not taken consideration of time impact. Based on Moore’s law, the price of computer resource should decrease about -50% per annum. This issue leads to our next topic of analysis, namely time dummy variable.

5.2.2 AWS Panel Data Test with Time Dummy Variables

If we consider the time variables, the total number of data points (instances) of an unbalanced panel dataset is 837 between 2008 and 2017. The number of explanatory or intrinsic variables is almost identical either using vCPU or ECU. However, ECU is AWS long-term measurement for CPU resource. The time dummy variables are 9 (10 years, T-1 time dummy variable, Table 8). The linear Q-Q plot shows that it is highly skewed, but after a semi-log transformation of instance prices, the Q-Q plot appears to be much better (see Fig. 4).

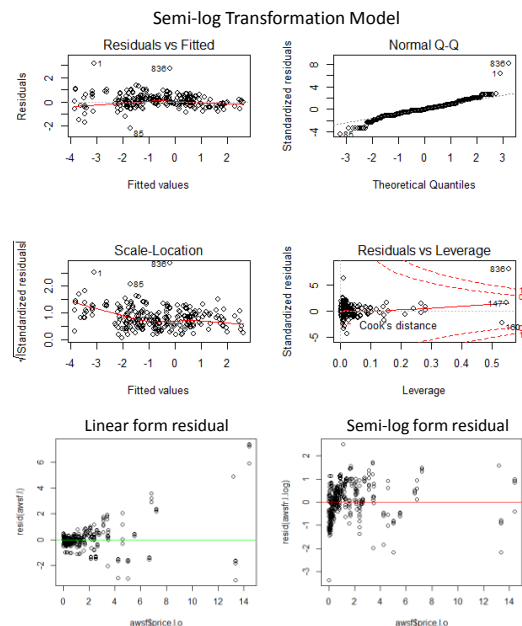


Fig.4: Log Transformation model and Residual Errors Plots Comparison between Linear and Semi-Log-Transformation 2008-2017

The main objective of the semi-log transformation is for time dummy variables inference. The initial linear model test only shows 7 variables including four-time dummy variables (2017, 2016, 2015 and 2014) are significant. It means that we can only inference four years. The R-squared and adjusted R-squared value are 0.8271 and

0.8235 respectively.

If we take the semi-log transformation, more time dummy variables become highly significant. The R-squared and adjusted R-squared values drop slightly to 0.8148 and 0.8109 respectively. If we add high order polynomial terms into the semi-log form, the test result is promised (see Table 8). There are two additional considerations to transfer hedonic function from linear to semi-log form:

TABLE 8
AWS Panel Data Regression Test with Time Dummy Variables (2008-2017)

Coefficients	Estimated β, δ	Std. Error	t-value	Pr(> t)	Annual Rate A_t	CAGR C_t
intercept	-3.87E+00	1.42E-01	-27.167	< 2e-16 ***		
RAM	2.46E-03	1.99E-04	12.315	< 2e-16 ***		
ECU	4.48E-02	1.93E-03	23.287	< 2e-16 ***		
ECU^2	-2.47E-04	1.59E-05	-15.578	< 2e-16 ***		
ECU^3	3.29E-07	3.23E-08	10.202	< 2e-16 ***		
Storage	2.26E-05	2.77E-06	8.159	1.27e-15 ***		
Net Perf	6.08E-01	2.77E-06	9.983	< 2e-16 ***		
Net Perf ^2	-6.46E-02	7.01E-03	-9.218	< 2e-16 ***		
Net Perf ^3	1.95E-03	2.27E-04	8.592	< 2e-16 ***		
bit	4.95E-02	2.17E-03	2.17E-03	< 2e-16 ***		
d17	-2.70E+00	1.62E-01	-16.709	< 2e-16 ***	-1.49%	
d16	-2.69E+00	1.54E-01	17.479	< 2e-16 ***	0.00%	
d15	-2.69E+00	1.54E-01	-17.479	< 2e-16 ***	-9.15%	
d14	-2.59E+00	1.55E-01	-16.769	< 2e-16 ***	-44.07%	
d13	-2.01E+00	1.64E-01	-12.261	< 2e-16 ***	-77.71%	
d12	-5.08E-01	1.21E-01	-4.196	3.02e-05 ***	-14.10%	
d11	-3.56E-01	1.24E-01	-2.859	0.00435 **	-6.69%	
d10	-2.87E-01	1.25E-01	-2.366	0.01822 *	-18.54%	
d9	-8.16E-02	1.25E-01	-0.65	0.51563	-7.83%	
d8	0.00E+00	baseline	baseline	baseline		
AAGR					-20.0%	
Compound Average Growth Rate						-25.9%
Residual standard error: 0.4913 on 817 degrees of freedom						
Multiple R-squared: 0.902, Adjusted R-squared: 0.8999						
F-statistic: 418 on 18 and 817 DF, p-value: < 2.2e-16						

- 1) Price of cloud infrastructure is closely associated with computer hardware. According to Moore's law, the hardware price depreciation rate is exponential in the time domain.
- 2) Previous experiences [42], [43] suggested the adoption of the semi-log model if a test is designed for a longer term comparison.

Based on the above test result with the time dummy, we can calculate Annual Growth Rate (A_t) and Average Annual Growth Rate (AAGR) by the following two equations:

$$A_t = \frac{(e^{\delta_t} - e^{\delta_{t-1}})}{e^{\delta_{t-1}}} \quad (12)$$

$$AAGR = \left(\left(\sqrt[T]{\prod_{t=2}^T (1 + A_t)} \right) - 1 \right) \times 100 \quad (13)$$

Note:

- 1) $\beta_0 = -0.021$, It is a combination of all explanatory variables. The value appears to be close to zero. It is a good indicator.
- 2) The coefficient values δ_1 is relative to 2008. It is emerged into β_0 value. Subsequently, "t" starts from 2.
- 3) There were no price changes between 2015 and 2016 after a significant discount in 2013 and 2014.
- 4) We use the geometric mean method to compute AAGR for

the years 2008 to 2017, the rate of depreciation is $-19.98\% \approx -20\%$

Overall, AWS AAGR or price reduction rate is far less than what Moore's law prediction [38], which is about 50% per annum in general. The gap between AWS AAGR and Moore's law prediction is $50\% - 20\% = 30\%$. To a certain extent, this price gap indicates why cloud customers are willing to pay more than what the benchmark price of computer hardware (see Fig. 5). We also see that AWS made a substantial price discount in 2013 and 2014. It may indicate a seven year's life cycle of computer asset if we consider AWS bought its cloud hardware assets in 2006. This is actually in align with Walker's [39] conclusion.

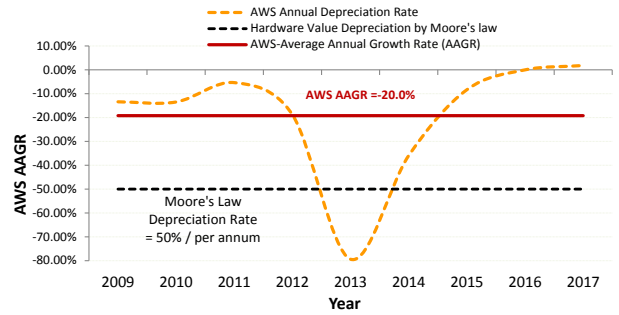


Fig.5: Comparison of AWS AAGR Price and Hardware Value Depreciation Rate via Moore's Law

The logic for this comparison can be justified by the following reasoning if we assume that the cloud instance configuration is *ceteris paribus*. Moreover, we assume that the independent time dummy variables do not interact with other independent variables, then, we should have the following two equations:

$$\ln((1+r)^t p) = \beta_0 + \sum_{i=1}^k \beta_i X_i + \delta_t d_t + \varepsilon \quad (14)$$

$$\ln((1+r)^{t-1} p) = \beta_0 + \sum_{i=1}^k \beta_i X_i + \delta_{t-1} d_{t-1} + \varepsilon \quad (15)$$

where, r = depreciation rate. Subtract (14) with (15) we should have the following equation (16).

$$r + 1 = e^{\delta_t - \delta_{t-1}} \quad (16)$$

$$r = \frac{e^{\delta_t} - e^{\delta_{t-1}}}{e^{\delta_{t-1}}}, \quad r = A_t$$

Based on the proof, we can derivate our conclusion that the AWS AAGR is around -20.0%/per annum in comparison with Moore's law.

By taking consideration of the impact of the time-dummy variable, the predicted price can be further updated. Alternatively, we can also use Compound Average Growth Rate (CAGR) to estimate the time impact, which is approximately close to AAGR. The CAGR formula is:

$$CAGR = C_t = \left(\frac{V_e}{V_s} \right)^{\frac{1}{T-1}} - 1 \quad (17)$$

where C_t is the compound average growth rate, V_e is the end value of the time period of "T" and V_s is the start val-

ue of the time period. Using the above prediction price in Table 7 as an example, we can correct the prediction result with CAGR in the following formula.

$$P_f = P_p \times (1 + A_t)^{(t_f - t_p)} \quad (18)$$

where P_f is for the future price and P_p is the present price, t_f is the future year value and t_p is the present year.

In order to predict future price accurately, we have to exclude the future year from our dataset when we calculate AAGR. In our case, it is 2017 data points. Subsequently, the value of $AAGR_{2008-2016} \approx -17\%$.

If we use this $AAGR_{2008-2016}$ to predict the instance price of m4.10xlarge in 2017 based on the 2014 price catalog, we should have the following result (see Table 9) and, the price difference between the real price and the predicted price (Δp) becomes negative.

TABLE 9
Estimate AWS Instance Price by Leveraging Time Dummy Variable

Within 95% confidence interval	Fitted Value	Real Price	Price difference Δp	Accuracy	Lower	Upper
Predicted value (semi-log)	\$1.693	2.155	-\$0.46	78.59%	0.635	4.516

Now, the question is why consumers should be willing to pay more than the predicted price. The possible answer is non-market characteristics of cloud services. From a CSP perspective, it is a part of CSP's marketing strategy to lead the cloud market. The common term is product or service differentiation. It leads to our next topic – cross-sectional dataset test, which is to examine the cloud instance price that is contributed by extrinsic variables.

5.2.3 Cross-Sectional Data Test

Based on the five CSPs' product catalogs, we constructed a dataset that consists of the entire 199 cloud instances. The initial linear model shows that R-squared and adjusted R-squared values are about 0.8077 0.7885 respectively and the Q-Q plot shows the data is highly skewed. According to the above five principles of transformation (discussed in Section 4.2), we can transfer it into semi-log form. Once the transformation is done, the Q-Q plot shows a better result (see Fig. 6) in comparison

Semi-log form

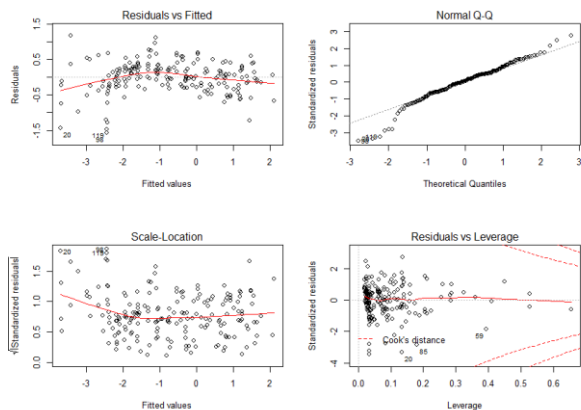


Fig.6: Semi-log transformation Form

with the linear form.

By a combination of semi-log transformation, adding a high order of polynomial terms and excluding just a few highly outlier points, the R-squared and adjusted R-squared values are increased more than 10% up to 0.913 and 0.904 respectively (see Table 10). A discussion of these elements is noted below:

TABLE 10
Cross-Section Data Analysis Results with the Semi-log Transformation

Coefficients	Estimated β, ξ	Std. Error	t-value	Pr(> t)	EXP(Z)	Above the baseline
Intercept : β_0	-2.68E+00	9.29E-02	-28.806	< 2e-16 ***		
RAM	2.50E-02	3.08E-03	8.122	8.14e-14 ***		
RAM^2	-1.14E-04	1.73E-05	-6.56	5.92e-10 ***		
RAM^3	1.47E-07	2.51E-08	5.875	2.10e-08 ***		
VCPU	1.98E-01	2.43E-02	8.154	6.72e-14 ***		
VCPU^2	-6.03E-03	9.76E-04	-6.18	4.41e-09 ***		
VCPU^3	5.55E-05	1.02E-05	5.458	1.64e-07 ***		
Storage	2.68E-04	6.57E-05	4.073	7.03e-05 ***		
Storage^2	-2.53E-08	7.13E-09	-3.551	0.000494 ***		
Storage^3	4.20E-13	1.22E-13	3.428	0.000758 ***		
Network Performance	1.77E-05	2.70E-04	2.479	0.014257 *		
Arch	-7.67E-02	1.13E-02	-6.786	1.73e-10 ***		
Arch^2	1.19E-03	1.75E-04	6.826	1.40e-10 ***		
Free Transfer to dedicated location ξ_2	5.49E-01	1.68E-01	3.275	0.001292 **	1.732	73.2%
GPU instance ξ_3	3.43E-01	1.63E-01	2.104	0.036771 *	1.409	40.9%
Burstable CPU ξ_4	5.33E-01	2.19E-01	2.44	0.015692 *	1.704	70.4%
Dedicated servers ξ_5	3.20E-01	1.29E-01	2.48	0.014097 *	1.377	37.7%
One account for all locations ξ_6	2.45E-03	3.33E-04	7.342	9.66e-12 ***	1.002	0.2%
Data Center Global Foot Print (AUS) ξ_7	2.82E-01	1.29E-01	2.19	0.029850 *	1.326	32.6%
Collocation ξ_8	4.00E-01	1.25E-01	3.194	0.001666 **	1.491	49.1%
48 Baseline Extrinsic Characteristics	0	Baseline	Baseline	Baseline	1	-
Average Extrinsic Price Value (AEPV)						43.4%

Residual standard error: 0.4607 on 173 degrees of freedom
Multiple R-squared: 0.9128, Adjusted R-squared: 0.9042
F-statistic: 106.5 on 17 and 173 DF, p-value: < 2.2e-16

- 1) Our analysis selected 5 intrinsic variables for cross-sectional data. Some intrinsic variables, such as a storage feature of Enterprise Block Store (EBS) optimized excluded from this test because it is insignificant for the regression analysis.
- 2) Based on the available dataset, we can make inference for 7 extrinsic variables (p-value is less than 0.05) with respect to a baseline characteristics of instance configuration (including, API, PAYG, Web interface, auto-scaling, resource usage monitoring, free transfer in, Free IP, load balancing, firewall, backup storage, credit card payment, volume discounts, free entry-level service and etc.).
- 3) The value of ζ_1 that represents the baseline characteristics have been emerged into the β_0 value. Different baseline configurations will result different β_0 values. It is dependent on the cross sectional dataset. Ideally, the β_0 value should be zero. However, it can only approach to zero in the reality.
- 4) Dedicated servers can be considered as extra resources.
- 5) Similar, free transfer to a dedicated location will give cloud customer mobility.
- 6) Burstable CPU can save the CPU price. If you do not use your specified capacity, CSP will give you credit so that you can withdraw when you need.
- 7) The price of GPU Instance is much higher than the baseline instance with the configuration of Intel CPU. AWS, GCP, and Azure provide the option of NVIDIA Tesla K80 GPU (launch price \$3,169/per unit in 2017.

In comparison with Intel Xeon E5-2673 V3 2.4-GHz chip, it costs \$700/ per unit Jul 2017).

- 8) The value ξ_6 of one account for allocation is minimal. It is basically submerged into the baseline characteristics, which all CSPs provide this feature without extra cost.
- 9) As Griliches indicated, the resulting of regression is sometimes unstable. It could be varied along with different circumstances. In the above case, if we change the configuration of the baseline extrinsic characteristics, the result will be totally different.

Now, we can answer the question that is raised before: “why cloud consumers are willing to pay nearly more than the predicted price.” If we use Table 9 to further revise our price prediction by taking consideration of cloud extrinsic variables, we can find the predicted price is very close to the real price (see Table 11).

TABLE 11
Predicted Price Including Extrinsic Values

With 95% confidence interval	Fitted Value	Real price	Price difference Δp	Accuracy	Lower	Upper
Predicted Price with Ave Extrinsic Value	\$2.428	2.155	\$0.273	87.32%	0.911	6.476
Predicted Price with particular cloud extrinsic characteristic	\$2.245	2.155	\$0.09	95.81%	0.842	5.988

5.2.4 Predict Cloud Prices for Different Instances

Notice that we can generalize the equation 18 as the equation 19 for future price prediction.

$$\hat{P}_f(X, Y) = (1 + AAGR)^{Y-Y_0} \times \left(\beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{j=1}^l \xi_j z_j \right) \pm 1.96 \sqrt{\frac{\hat{P}_f(X, Y)(1 - \hat{P}_f(X, Y))}{n}} \quad (19)$$

where $X = \langle x_1 \dots x_k, z_1 \dots z_l, Y \rangle$, Y = future year, Y_0 = current year or present year, n = size of population for a dataset. (We adopt 95% Wald confidence intervals or first approximation). Furthermore, if we take the semi-log form, the equation can be presented as following:

$$\ln \hat{P}_f(X, Y) = (Y - Y_0) \times \ln(1 + AAGR) + \left(\beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{j=1}^l \xi_j z_j \right) \pm 1.96 \sqrt{\frac{\hat{P}_f(X, Y)(1 - \hat{P}_f(X, Y))}{n}} \quad (20)$$

We use this equation to estimate the future price of different cloud instances. The comparison of different AWS

TABLE 12
Predicted AWS Cloud Prices with Different Instance

Instance types	API name	Fitted Value with 95% CI	Real price	Price difference Δp	Accuracy	Lower	Upper
Standard	m1.medium	0.0842	\$0.120	-0.036	70.2%	0.031	0.225
Standard	m4.10xlarge	\$1.693	\$2.155	-\$0.462	78.6%	0.635	4.516
3 rd Gen. Cluster	i3.8xlarge	\$1.69	\$2.496	-\$0.808	67.6%	0.553	4.85
4 th Gen. Cluster	c4.8xlarge	\$0.979	\$1.591	-\$0.612	61.6%	0.252	3.197

cloud instance produces the following prediction results (shown in Table 12).

We highlight three points for the prediction results:

- 1) The predicted prices usually are less than the real price. It means that AWS holds the price reduction pace due to its extrinsic values of a cloud instance.
- 2) For the standard instance, the predicted accuracy is approximately higher than 70% without consideration of extrinsic characteristics. (With one CSP, the extrinsic value cannot be compared)
- 3) For the latest generation cluster, the prediction accuracy is below 70%. It might be due to more extrinsic values that AWS has built into its price catalog.

Overall, once the predicted cloud price emerges, it can underpin the CIO to make a right strategic investment decision for IT infrastructure. Of course, he or she has to take consideration of other factors, such as business risks, workload growth, and volume discount and workload portability issues (or cloud vendor lock-in syndrome: “free to come and pay to leave”).

6 ANALYSIS AND DISCUSSION

We have illustrated how to use the hedonic analysis to predict the cloud instance price. From the unbalanced panel data, we can calculate the AWS’ AAGR is approximate -20.0% per annum. Statistically, the time dummy variable is the same as a fixed effect. The net effect is the hedonic function to be shifted downwards (see Fig. 7)

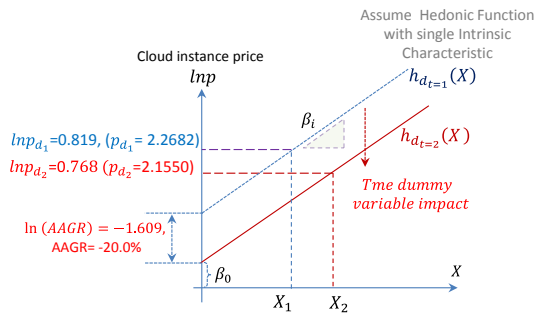


Fig. 7: Impact of Time Dummy Variable on AWS Cloud Instance Price

In comparison with Moore’s law prediction, the AWS price change rate (deflation) is at much less slow pace than what Moore’s law has predicted (-50% per annum). The reason that AWS can move beyond the competition just on price is its extrinsic characteristics that AWS can differentiate its cloud service from its competitors. AWS has developed more than 1,000 different cloud characteristics or features since 2006. Although we would not be able to analyze all extrinsic characteristics here, we can highlight some of the extrinsic characteristics among 5 leading CSPs (shown in Fig. 8). The characteristic of GPU instance is about 40.9% of cloud extrinsic value and data center global footprint (Australia) is 32.6% in comparison with the baseline configuration.

Ultimately, the impact of the extrinsic variable is similar to the time dummy variable (or fixed effect). It only shifts the hedonic function either up or down. This means

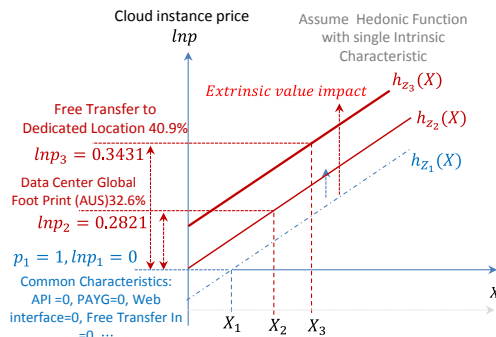


Fig.8: Impact of Extrinsic Variables on AWS Cloud Instance Price

in order to avoid an estimated bias we should include the required cloud characteristics not only intrinsic variables. If we just compare the cloud instance prices based on the intrinsic variables alone (for average configuration), AWS price is not the cheapest in comparison with top 30 global leading CSPs. Its price is just slightly above the median one (The market median price is \$146 marked as a notch. AWS instance price is \$149 is marked by a dashed line shown in a boxplot, Fig. 9). However, AWS can still maintain over 31% of IaaS global market share and keep double digits revenue growth year on year (YoY). This is mainly due to the contribution of AWS extrinsic values of its cloud services which cloud customers are willing to pay for.

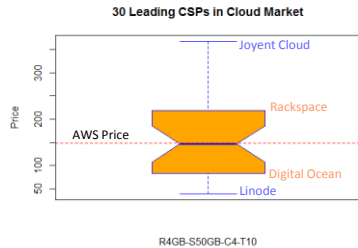


Fig.9: Box plot of 30 CSPs (an instance of 4GB-RAM, 50GB-storage, 4 -vCPU and 10 GB transfer out)

In this paper, we introduce the new concept of intrinsic and extrinsic variables that have been applied to the hedonic analysis of cloud pricing model. Moreover, we have mathematically proved that the time dummy or AAGR is equivalent to Moore's law impact if ceteris paribus. The AAGR plays a vital role in cloud price prediction.

In contrast to the previous studies that ignored the extrinsic variables impact on the cloud prices, we have clearly demonstrated that many extrinsic variables have significant values or fixed effect on the cloud price. The effective combination (or bundle) of intrinsic and extrinsic values does not only allow CSPs to slow the price reduction pace but also underpin the cloud market leadership.

Generally, the hedonic analysis is a practical or empirical approach to disclose the latent values of what customers are willing to pay for the quality changes. Ultimately, this research is to leverage the hedonic concept to discover homogeneous cloud pricing patterns that are closely associated with heterogeneous cloud service characteris-

tics, which are often hidden behind the complicated cloud pricing structure.

Our novel approach enables cloud customers to predict cloud service prices accurately based on their business application needs rather than purely on the cost of IaaS comparison. It means that cloud consumers can avoid many pricing estimation biases.

Another important implication is that it allows many CSPs to establish the correct performance benchmark based on the true value proposition of cloud services to compete with their market leader not only just on the price.

7 CONCLUSIONS AND FUTURE WORK

The conclusion of this work is the cloud instance price cannot be just examined by its intrinsic characteristics (mainly cost components, such as RAM, CPU, network performance, storage, etc.) alone. It will inevitably lead to the pricing estimation bias because the cloud price prediction is ultimately determined by three key factors or variables, namely, intrinsic, extrinsic and time dummy. Many traditional cloud pricing models cannot reflect cloud extrinsic values (such as burstable CPU, dedicated server, data center global footprint, etc.). However, it does not mean we can ignore these extrinsic characteristics. In fact, they have a heavy influence on the cloud service price. Throughout this paper, we have shown the process of how to calculate and predict the cloud price accurately and how to avoid the price estimation bias. The novelty of our work is that we present and prove that the value of AAGR is equivalent to Moore's law in cloud services.

We argue that the hedonic pricing model is a better approach to estimate the cloud price accurately if we can establish the adequate hedonic function form based on the available dataset in hand. Furthermore, we exhibit the AWS cloud price has been declining over the last 10 years but at a much slower pace in comparison with Moore's law prediction. One of the major influenced factors of this declining is due to the cloud of extrinsic values or characteristics. They have become AWS competitive advantages to lead in the cloud (IaaS) market.

We understand that some of our model assumptions can impact on the accuracy of cloud price prediction. However, if we can fully access many CSPs' datasets, we can improve the prediction results. In the future, we plan to refine our estimations with the availability of new data.

REFERENCES

- [1] R. Buyya, J. Broberg, and A. M. Goscinski, Cloud computing: Principles and paradigms, Vol. 87. John Wiley & Sons, 2010.
- [2] J. Weinman, Clouconomics: The business value of cloud computing, John Wiley & Sons, 2012.
- [3] L. Venkatachalam, "The contingent valuation method: a review," Environmental impact assessment review 24.1 2004, pp. 89-124.
- [4] H. A. Linstone, and T. Murray, The Delphi method: Techniques and applications, Vol. 29. Reading, MA: Addison-Wesley, 1975.
- [5] K. Lampe, The Birth of Hedonism: The Cyrenaic Philosophers and Pleasure as a Way of Life, Princeton University Press, 2014.
- [6] J. Bentham, The collected works of Jeremy Bentham: An introduction to

- the principles of morals and legislation, Clarendon Press, 1996.
- [7] J. S. Mill, *Utilitarianism*, Chapter II, "What Utilitarianism Is" Chapter II, "Of the Ultimate Sanction of the Principle of Utility, 1863.
 - [8] G. E. Moore, *Principia Ethica*, Chapter III, Hedonism, Cambridge at the University Press 1903, pp. 59-108
 - [9] T. Hurka, "Moore in the Middle," *Ethics* 113.3, 2003, pp. 599-628.
 - [10] Z. Griliches, "Hedonic price indexes for automobiles: An econometric of quality change," *The Price Statistics of the Federal Government*. NBER, 1961, pp. 173-196.
 - [11] R. Michaels, "Hedonic prices and the structure of the digital computer industry," *the Journal of Industrial Economics* 1979, pp. 263-275.
 - [12] R. Cole, Y.C. Chen, J. A. Barquin-Stoliman, and J. H. Hodge, "Quality-adjusted price indexes for computer processors and selected peripheral equipment," *Survey of Current Business* 66.1, 1986, pp. 41-50.
 - [13] E.R. Berndt, and Z. Griliches, "Price indexes for microcomputers: an exploratory study." Price measurements and their uses. University of Chicago Press, 1993, pp. 63-100.
 - [14] R. H. Raghav, and B. D. Lynch, "Hedonic price analysis of workstation attributes," *Communications of the ACM* 36.12 1993, pp. 95-102.
 - [15] R.M. Sakia, "The Box-Cox transformation technique: a review," *The statistician* 1992, pp. 169-178.
 - [16] A. Pakes, "A Reconsideration of Hedonic Price Indexes with an Application to PCs," *The American Economic Review* 93.5 2003, pp. 1578-1596.
 - [17] C. R. Hulten, "Price hedonics: a critical review," 2003.
 - [18] P. Davis, and G. Eliana, *Quantitative techniques for competition and antitrust analysis*, Princeton University Press, 2009, pp.286-287, pp.47-61
 - [19] P. Belleflamme, and P. Martin, *Industrial organization: markets and strategies*, Cambridge University Press, 2015.
 - [20] K.J. Lancaster, "A new approach to consumer theory," *Journal of political economy* 74.2, 1966, pp. 132-157.
 - [21] S. Rosen, "Hedonic prices and implicit markets: product differentiation in pure competition," *Journal of political economy* 82.1, 1974, pp. 34-55.
 - [22] H. W. Brachinger, "Statistical theory of hedonic price indices," *University of Fribourg*, 2002.
 - [23] Internet Archive, <https://archive.org/web/>
 - [24] AWS annual report, <http://phx.corporateir.net/phoenix.zhtml?c=97664&p=irol-reportsannual/>
 - [25] R. Halvorsen, and H. O. Pollakowski, "Choice of functional form for hedonic price equations," *Journal of urban economics* 10.1, 1981, pp. 37-49.
 - [26] E. Cassel, and R. Mendelsohn, "The choice of functional forms for hedonic price equations: comment," *Journal of Urban Economics* 18.2, 1985, pp. 135-142.
 - [27] S. El Kihal, C. Schlereth, and B. Skiera, "Price Comparison for Infrastructure-as-a-Service," *ECIS*, 2012.
 - [28] J. E. Triplett, "Hedonic methods in statistical agency environments: an intellectual biopsy." Fifty years of economic measurement: the jubilee of the conference on research in income and wealth. University of Chicago Press, 1991.
 - [29] J.E. Triplett, Draft copy Handbook on quality adjustment of price indexes for information and communication technology products, Paris: OECD, 2000.
 - [30] Z. Griliches, "Price indexes and quality change," *Harvard University Press*, 1971.
 - [31] Z. Griliches, "Hedonic price indexes and the measurement of capital and productivity: some historical reflections." Fifty years of economic measurement: The jubilee of the conference on research in income and wealth, University of Chicago Press, 1991.
 - [32] J. R. Gordon, *The measurement of durable goods prices*, University of Chicago Press, 1990, pp. 188-240.
 - [33] J. H. McDonald, *Handbook of biological statistics*, Vol. 2. Baltimore, MD: Sparky House Publishing, 2009.
 - [34] N.V. Kuminoff, C. F. Parmeter, and J. C. Pope, "Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?" *Journal of Environmental Economics and Management* 60.3 2010, pp. 145-160.
 - [35] M. L. Cropper, L. B. Deck, and K. E. McConnell, "On the choice of functional form for hedonic price functions," *The Review of Economics and Statistics* 1988, pp. 668-675.
 - [36] H. Kyle, "AWS moves from ECU to vCPU," <http://blogs.gartner.com/kylehilgendorf/2014/04/16/aws-moves-from-ecu-to-vcpu/>
 - [37] J. Read, What is an ECU? CPU Benchmarking in the Cloud, <http://blog.cloudharmony.com/20-10/05/what-is-ecu-cpu-benchmarking-in-cloud.html>
 - [38] C. Mack, "The Multiple Lives of Moore's Law," <http://spectrum.ieee.org/magazine/2015/April>
 - [39] E. Walker, W. Briskin, and J. Romney. "To lease or not to lease from storage clouds," *Computer* 43.4, 2010, pp. 44-50.
 - [40] E. Walker, "The Real Cost of a CPU Hour," *Computer*, (4), 2009, pp. 135
 - [41] P. Mitropoulou, E. Filiopoulou, C. Michalakelis, and M. Nikolaidou, "Pricing cloud IaaS services based on a hedonic price index," *Computing* 98(11), 2016, pp. 1075-1089.
 - [42] M. Silver, and S. Heravi, "The Difference Between Hedonic Imputation Indexes And Time Dummy Hedonic Indexes," Washington, D.C. International Monetary Fund, 2006.
 - [43] A. Aizcorbe, "The Stability Of Dummy Variable Price Measures Obtained From Hedonic Regressions," Mimeo, Federal Reserve Board, Washington DC, 2003
 - [44] M. Naldi, and L. Mastroeni, "Economic decision criteria for the migration to cloud storage," *European Journal of Information Systems*, 25(1): 2016, pp. 16-28
 - [45] P. Mitropoulou, E. Filiopoulou, M. Nikolaidou, and C. Michalakelis, "Pricing IaaS: A Hedonic Price Index Approach," *Lecture Notes In Computer Science* 2017, pp. 18-28
 - [46] A. Baranzini, J. Ramirez, C. Schaerer, and P. Thalmann, *Hedonic methods in housing markets: Pricing environmental amenities and segregation*, Springer Science & Business Media, 2008
 - [47] Storage, Full history 1957 - present, <https://hblock.net/blog/storage/>



Caesar Wu is a senior IEEE member. He is one of the authors of *Cloud Data Center and Cost Modeling*. He was a senior domain specialist in Telstra. He managed and operated many of Telstra's enterprises and IT data centers. He has over 30 years working, researching and academic experiences across various industries.



Adel Nadjaran Toosi was a research fellow in the CLOUDS Laboratory at the University of Melbourne. He is now a lecturer in Faculty of Information Technology, Monash University. He is currently working on resource management for software-defined clouds. For more details, please visit his homepage: <http://adelnadjarantoosi.info>.



Rajkumar Buyya is a Redmond Barry Distinguished Professor and Director of the CLOUDS Laboratory at the University of Melbourne. He has authored more than 600 publications. He is recognized as a "Web of Science Highly Cited Researcher" both in 2016 and 2017 by Thomson Reuters, a Fellow of IEEE, and Excellence in Innovative Research Award by Elsevier for his outstanding contributions to Cloud computing.



Kotagiri Ramamohanarao (Rao) received the Ph.D. degree from Monash University. He was awarded the Alexander von Humboldt Fellowship in 1983. He was Research Director for the Cooperative Research Center for Intelligent Decision Systems, the program co-chair for VLDB, PAKDD, DASFAA, and DOOD conferences. He was awarded Distinguished Contribution Award in 2009 by the Computing Research and Education Association of Australasia. He is a Fellow of Australian Academy of Science, a Fellow of Australian Academy of Technology and Engineering and a Fellow of Institution of Engineers Australia.