

Wellesley College Wellesley College Digital Scholarship and Archive

Honors Thesis Collection

2016

Price trends for cloud computing services

Liang Zhang
lzhang3@wellesley.edu

Follow this and additional works at: <https://repository.wellesley.edu/thesiscollection>

Recommended Citation

Zhang, Liang, "Price trends for cloud computing services" (2016). *Honors Thesis Collection*. 386.
<https://repository.wellesley.edu/thesiscollection/386>

This Dissertation/Thesis is brought to you for free and open access by Wellesley College Digital Scholarship and Archive. It has been accepted for inclusion in Honors Thesis Collection by an authorized administrator of Wellesley College Digital Scholarship and Archive. For more information, please contact ir@wellesley.edu.

PRICE TRENDS FOR CLOUD COMPUTING SERVICES

LIANG ZHANG

ADVISOR: PROFESSOR DANIEL SICHEL

Submitted in Partial Fulfillment of the Prerequisite for Honors in Economics

May 2016

© 2016 LIANG ZHANG

PRICE TRENDS FOR CLOUD COMPUTING SERVICES

SENIOR THESIS

AUTHOR:
LIANG ZHANG'16
WELLESLEY COLLEGE

THESIS ADVISOR:
DANIEL E. SICHEL
WELLESLEY COLLEGE

APRIL 21, 2016

Abstract

The landscape in the information technology (IT) sector is changing rapidly, and studying price trends for a newly developed IT service provides a great gateway to gauge these changes. Cloud computing—the on-demand delivery of computing resources and applications via the Internet—is rapidly expanding yet relatively little is known about trends in prices of cloud computing. For my senior economics thesis, I constructed quality-adjusted price indexes to quantify the rate of price change of Cloud Computing Services. I employ two hedonic regression methods as well as the matched model method and the results suggest that Cloud Computing Services experienced a sharp price decline from 2009-2015. My results also indicate that the price reduction trend of Cloud Computing Services is slightly sharper than that of Computing Products in general (hardware and software).

Acknowledgements

I would like to thank my advisor Professor Daniel Sichel massively for his generous guidance and encouragement during this process. I am extremely grateful for the opportunity to work with him on this topic. My thesis would not have become reality without him.

I would also like to thank Professor McKnight for organizing the Economics Research Seminar and her thoughtful suggestions on this thesis.

I would also like to thank my fellow thesising students Hui Li'16, Sharon Shiao'16, Amy Wickett'16, Jackie Li'16, Hannah Ruebeck'16, Hero Ashman'16, Taylor Cranor'16, Suzanne Barth'16, and Evelyn Taylor-McGregor'16 for their helpful insights.

All errors are solely mine.

Contents

1. Introduction
 2. Background Information on Cloud Computing
 3. Preliminaries on Price Index
 - 3.1. Bias in Matched Model
 - 3.2. Overview of the Hedonic Method
 - 3.2.1. Dummy Variable Methods
 - 3.2.2. Adjacent Period Method
 - 3.2.3. Bias Adjustment in Hedonic Model
 4. Data Collection
 - 4.1. Selection and Collection of Dataset
 - 4.2. Data Construction and Imputation
 5. Empirical Results
 - 5.1. Matched Model Method
 - 5.2. Hedonic DV Method
 - 5.2.1. Comparison of Matched Model Index and Hedonic DV Index
 - 5.3. Hedonic Adjacent Period Method
 - 5.4. Explanation of the Drastic Change of AWS Products in 2014
 6. Comparison of Price Trends between Computing Products and Cloud Computing Services
 7. Possible Issues
 8. Conclusion
- References

1. Introduction

The landscape in the information technology (IT) sector is changing rapidly, and studying price trends for a newly developed IT service provides a great gateway to gauge these changes. Cloud computing—the on-demand delivery of computing resources and applications via the Internet—is rapidly expanding yet relatively little is known about trends in prices of cloud computing.

Cloud computing represents the innovative idea of sharing high-performance large Warehouse Scale Computers (WSC) as a commodity instead of depending on local data centers. Without the burden to invest or maintain any hardware at all, users can simply rent computing resources according to their specific needs. Since 2009, Cloud Computing Services have been available to the public and growing rapidly. The major benefits of using cloud computing include energy efficiency, ease in scaling, and savings in administrative costs. The major providers are Amazon, Google, and Microsoft. As the leader in the industry, Amazon Web Services (AWS) has a leading impact on the general price trend.

A desirable way to understand the price trend of a certain group of products is with a price index. The most well known example of a price index is the Consumer Price Index (CPI), which is constructed by the Bureau of Labor Statistics (BLS) to measure changes in the cost of living. For Computing Products in general, the Bureau of Economic Analysis (BEA) constructs price indexes for investment in hardware (computers and peripheral equipment) and software as depicted in Figure 1. Figure 2 plots these indexes on a natural log scale, and we can observe a consistent price reduction trend in both indexes.

Table 5.3.4. Price Indexes for Private Fixed Investment by Type

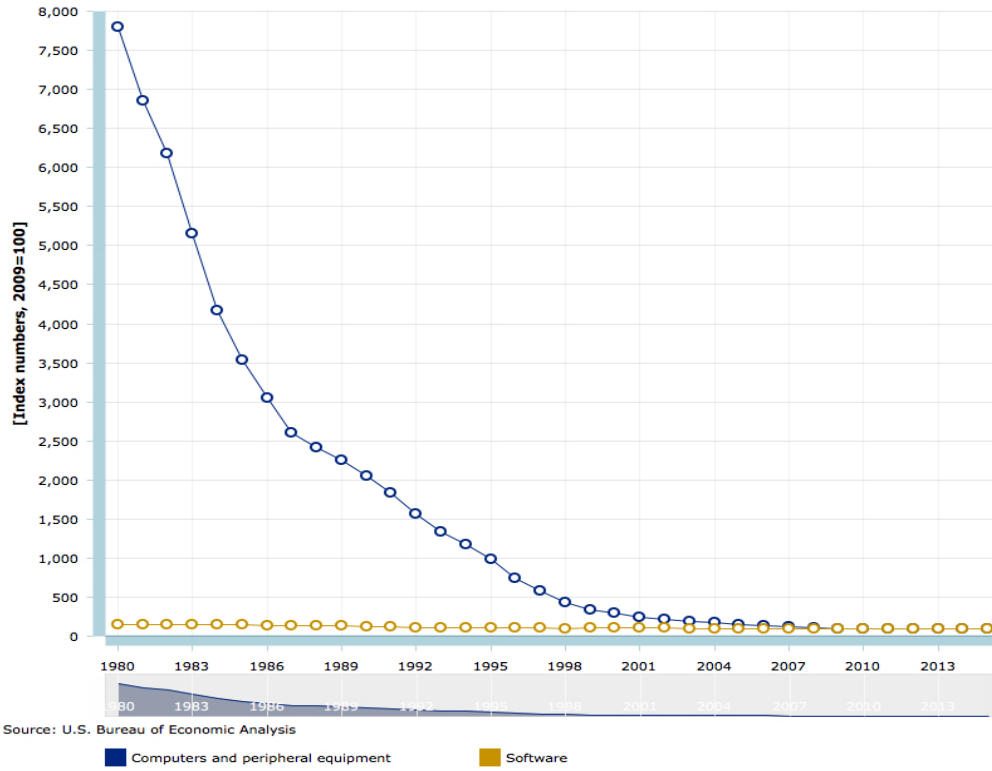


Figure 1

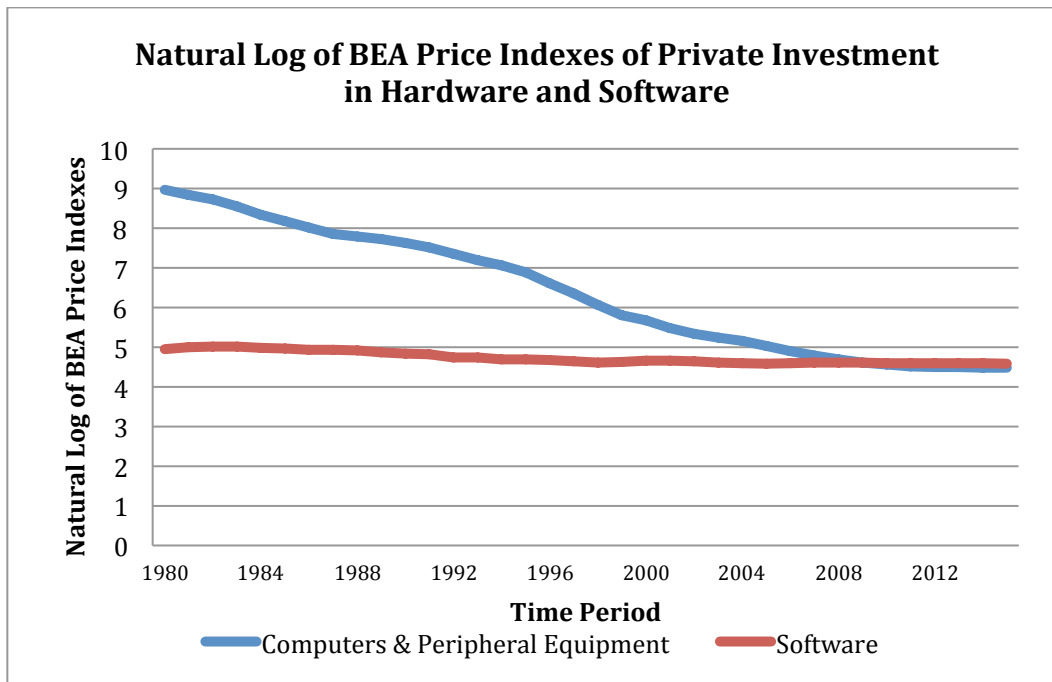


Figure 2

Two common ways to construct price indexes are the matched model method and the hedonic method. Both methods are used by BEA and BLS to construct various price indexes. The matched model method tracks a fixed pool of products and takes a weighted average of price changes. Matched model indexes control for quality change by only considering price changes for matched or homogeneous products. Hedonic indexes directly control for quality change with measures of product characteristics that reflect product quality. Compared to the hedonic regression method, the matched model method can exhibit quality bias and new goods bias. A matched model is unable to capture the implicit price change caused by a change in the quality of the product; the resulting bias is defined as quality bias. Matched model indexes only track a matched bundle of products, and the bias caused by its inability of capturing new products is defined as new goods bias.

In this thesis, I consider these important issues in price indexes in the context of cloud computing. Private sector technology trend reports have indicated that the price of Cloud Computing Services has been declining rapidly. However, no formal analysis has been made on this topic before, and the observations made by private sector reports are mostly anecdotal. There are many questions we can ask about Cloud Computing Services. What is the price trend of this new computing service? Is the price trend for cloud services in line with long-standing historical price trends for Computing Products in general? In recent years, when prices for computing hardware have fallen very slowly, do prices for Cloud Computing Services exhibit a similar sluggish rate of decline? To answer these questions, we need to construct quality-adjusted price index for Cloud Computing Services. With public online information from AWS, I was able to construct a

dataset of Cloud Computing Services prices and quality metrics. My thesis quantifies the rate of the price decline and analyzes the price trend of Cloud Computing Services by constructing quality-adjusted price indexes for AWS Elastic Cloud Computing (EC2) Services from 2009-2015.

From the hedonic index I constructed, I find that AWS EC2 services have experienced a sharp price decline from 2009 -2015. The price decreased gradually and modestly during 2009-2013 with small fluctuations. A visible price increase occurred in late 2013 due to Amazon's large scale restructuring of its product system – a large number of new products were introduced while almost all old products were eliminated. The price plummeted in early 2014 with a 38% (at annual rate 62%) drop from previous period and has remained unchanged through the end of 2015. The matched model gives a similar result from 2009-2013, but was unable to extend beyond 2014, for Amazon eliminated all the products in the fixed bundle and replaced them with new products. Compared to the price trend of overall traditional Computing Products (hardware and software), Cloud Computing Services have experienced a slightly sharper price decline since 2009.

My thesis is organized as follows. Section 2 provides background information on Cloud Computing Services. Section 3 gives an overview of different methods for constructing price indexes and summarizes available public information about Cloud Computing Services. Section 4 describes the selection, construction and imputation of data. Section 5 discusses the empirical results and compares traditional matched model results with hedonic regression results. Section 6 compares the price trend between

overall Computing Products and Cloud Computing Services. Section 7 addresses possible issues. Section 8 concludes.

2. Background Information on Cloud Computing

This section provides some background information on Cloud Computing Services.

When customers purchase computing power from AWS, they are renting an EC2 instance. I define an EC2 instance as a virtual machine (computer) that can be rented by the hour. Customers can rent different instances based on their specific needs. Different types of instances can be seen as different specifications of the virtual machine in terms of the amount of memory, number of EC2 Compute Units (ECU), storage, storage type, Input/output performance level and platform. I define these quality metrics as the characteristics of an instance. In the next paragraph, I will provide further explanation for each characteristic.

Each AWS EC2 instance has following characteristics:

- Memory: Memory is the random access memory (RAM) inside a computing machine. Computers use memory to store actively running programs. A specific amount of memory is attributed to each AWS EC2 instance. Memory is measured by Gigabyte (GB).
- EC2 Computing Units (ECU): ECU is the measure of central processing unit (CPU) for an instance. The amount of CPU that is attributed to a specific instance is expressed in terms of ECU. We can think of the instance as a virtual computer, and ECU is the measure of CPU power for this virtual machine. ECU is an essential characteristic, as it directly measures the computing power obtained by customers.

- Storage: Storage refers to permanent storage that is used to hold programs and data until purposely changed or removed by the user. Storage is measured by GB.
- Storage Type: Instance Storage and Solid State Drive-backed instance storage are the two types of storage options offered by AWS. Solid State Drive (SSD) is utilized in instance types that are designed to provide high input/output performance. In other words, SSD is preferred for heavy, database workloads.
- Input/output (I/O) Performance: I/O Performance measures how fast and efficient an instance can get the data into (and out of) the storage. I/O performance is essential for storage management and network capability.
- Platform: Platform describes the type of processor and is measured by bit. 64-bit platforms are widely used for computers.

It is notable that the characteristics of an instance are independent of regions, operating systems and payment options. The price of the same instance might differ based on these factors (region, operating systems and payment options), but the characteristics of the instance do not vary. Customers can choose to purchase an instance running on servers based in different locations, such as Virginia, California, or Oregon. Globally, AWS EC2 instances are also available in Europe and Asia. Customers can also choose the operating system based on their own needs. Linux and Window are the most popular operating system choices. There are three different payment options, and customers make their choice based on their business demands and usage:

- On-Demand: The on-demand option lets you rent the instance by the hour without long-term commitments or upfront payment. Customers pay for what they use.

- Reserved: The reserved option provides customers with a sizeable discount compared to the on-demand option, but with an upfront payment and commitment to long-term usage.
- Spot: The spot option charges customers an hourly rate that is slightly lower than the on-demand rate, but the price fluctuates based on supply and demand. Customers can stipulate a maximum hourly price they are willing to pay for an instance under the spot option, but their service will be shut down once the market price exceeds this maximum.¹

The information introduced above is crucial for data selection process, and I will return to this discussion in Section 4.

No formal academic analysis has been made of price trend of Cloud Computing Services. Some industry newsletters have reported on price changes for AWS services based on anecdotal evidence, however, no comprehensive data or analysis was provided.

¹ <https://aws.amazon.com/ec2/purchasing-options/>

² Differences in privacy laws across countries have led AWS to base servers in a wide

3. Preliminaries on Price Indexes

3.1. Bias in Matched Model

The matched model and the hedonic methods are widely used by government agencies like BLS and BEA to construct price indexes. Aizcorbe (2014) describes a typical matched model index using the Laspeyres formula:

$$I_{0,1}^L = \frac{\sum_{m=1}^M (P_{m,1} Q_{m,0})}{\sum_{m=1}^M (P_{m,0} Q_{m,0})} \quad (1)$$

where M goods are sold in period 0 and 1; $P_{m,t}$ represents the price of good m in period t ; $Q_{m,t}$ represents the quantity sold of good m in period t . The Laspeyres index measures the price change in period 1 by comparing the cost of purchasing the same amount of each good m in period 1 and period 0.

A matched model index controls for quality change by considering price changes only for products observed in both periods; accordingly, it cannot compare the quality of goods newly entering the market to those in the market or those exiting the market. For example, a new iPhone introduces new features that increase the “utility per dollar” and implicitly decreases the price, but this quality change could be missed by a matched model index. Similarly, the introduction of a totally new product, such as electronic reading devices, may have a substantial impact on the cost of living, but that improvement would be missed by a matched model index.

The IT industry changes rapidly with frequent entry and exit of new products as well as frequent changes in quality. Thus, the quality bias and new goods bias associated with a matched model is troublesome for a Cloud Computing price index. A matched model approach will overstate the rate of price change in an industry with such frequent technological improvements.

3.2. Overview of the Hedonic Method

Hedonic regression is another method to construct quality-adjusted price indexes. In contrast to the matched model method, hedonic regressions directly control for the characteristics of the goods. Thus, implicit price changes due to the increase in “utility per dollar” can be captured and measured, thereby mitigating quality change bias. Hedonic methods also mitigate the new goods bias in matched model indexes. As pointed out in Pakes (2003), hedonic regressions allow us to compare the value consumers attached to the characteristics of goods, and thus we can compare the prices of old goods to those of new goods directly instead of only considering a fixed bundle of goods over time.

The underlying assumption for hedonic regression is that there is a relationship between prices and products’ characteristics, and consumers attach a certain value to specific characteristics. As discussed in Pakes (2003), hedonic functions capture this relationship between price and characteristics. Different assumptions about market equilibrium provide different forms of hedonic functions. Here, I will use the example discussed in Pakes (2003) to demonstrate this relationship. We can think of a group of products whose price change we would like to measure over time. We can expect intuitively that for every single product with a normal demand curve, the price depends on marginal cost and a “mark-up” for some of its characteristics. For AWS EC2 services, the characteristics of an instance are: memory, EC2 Compute Unit (ECU), storage, storage type, platform, and I/O performance. The details about each characteristic can be found in Section 2.

As demonstrated in Pakes (2003), we denote the value of characteristics and the price of good i as (x_i, p_i) . Denote the characteristics and prices of other goods marketed as (x_{-i}, p_{-i}) . The demand function for good i is $D_i(\cdot) = D(x_i, p_i, x_{-i}, p_{-i}; A)$. The parameter A is the distribution of consumer attributes that determine consumer's preferences over characteristics. Assume all firms are single product firms and have marginal cost $mc(\cdot)$. We can express the price p_i as $p_i = mc(\cdot) + \frac{D_i(\cdot)}{|\partial D_i(\cdot)/\partial p|}$, where the second term is the “mark-up” that has an inverse relationship with demand elasticity. We can condition this equality on characteristics x_i . Therefore, we have

$$h(x_i) = E[p_i|x_i] = E(mc(\cdot)|x_i) + E\left(\frac{D_i(\cdot)}{|\partial D_i(\cdot)/\partial p|} |x_i\right)$$

Thus the hedonic function $h(x_i)$ here is the expectation of marginal cost and the mark-up conditioned on characteristics. Denote the hedonic function in period t as $h^t(x_i)$.

In the next two subsections, I will introduce two methods to construct a hedonic price index.

3.2.1. Dummy Variable Method

As discussed in Aizcorbe (2014), the Dummy Variable (DV) Method regresses prices on characteristics and includes dummy variables for each time period (except for the first/base period, which is dropped for reference). The DV method controls for quality by including characteristics in the regression, so the time dummy variables will capture all other changes in prices after controlling for quality. As a result, the difference in the coefficients of these time dummies measures the constant-quality price change over time.

As illustrated in Aizcorbe (2014), the hedonic function for this particular method is:

$$h_{DV}(x_i): \ln P_{i,t} = \alpha + \sum_k \beta_k X_{k,i,t} + \sum_t \delta_t D_{i,t} + \varepsilon_{i,t} \quad (2)$$

In this hedonic function, the logged price of each product i sold at time period t is expressed as a function of the quantities of each characteristic k in each product i at time t , and time dummies $D_{i,t}$, which equals 1 if a price for product i is observed at time t , and 0 otherwise (Aizcorbe 2014). $\varepsilon_{i,t}$ is the error term. For each product i , there are k characteristics that can affect the price and the term $\sum_k \beta_k X_{k,i,t}$ controls for the quality. The coefficient β_k tells us the effect of each characteristic on the log of price. The time dummy coefficients δ_t capture all the other price effects and the difference between δ_t s gives a measure of constant quality price change between time periods.

3.2.2. Adjacent Period Method

The Adjacent Period Method takes a very similar approach to the DV method. As discussed in Byrne et al. (2015), the adjacent period method runs the DV regression separately for each pair of adjacent periods. For $t \in \{t_j, t_{j+1}\}$

$$h_{adj}^{\{t_j, t_{j+1}\}}(x_{i,t}): \ln(P_{i,t}) = \alpha + \sum_k \beta_k \ln(x_{k,i,t}) + \delta_t D_{i,t(j+1)} + \varepsilon_{i,t} \quad (3)$$

The adjacent period method allows the β coefficients to change over time. It is also a helpful way to observe changes in characteristics' effect on prices over the time. For a specific characteristic k , an increasing coefficient β_k over time periods implies that characteristic k 's effect on price has become bigger.

The adjacent period method is particularly helpful in interpreting dramatic changes in price index, as the change of characteristics coefficients β provides insights

about the changing importance and relevance of characteristics over time. An increase or decrease in important characteristics can be a crucial factor for understanding the drastic changes in price. I will demonstrate my application of this method in section 5 to help explain the drastic AWS Cloud Computing Services price drop that happened in 2014.

3.2.3. Bias Adjustment in Hedonic Method

As Aizcorbe (2014) discussed, bias adjustment is essential in Hedonic regressions with dependent variable that takes form of natural log. As my ultimate goal is to develop an index for the level of prices instead of the natural log of price, an adjustment factor of $\exp(0.5\text{Var}(\varepsilon_{m,t}))$ will be applied to the time dummy coefficients in the hedonic regressions.

4. Data Collection

4.1. Selection and Construction of Dataset

To construct a price index, I need data on prices and characteristics from 2009 to the present. AWS first started to post EC2 instances price data online in 2009. Initially, the services were provided from servers based in Virginia and California, though the services could be accessed from anywhere in the world. In the middle of 2011, Amazon also started to provide EC2 from servers based in Oregon. Outside the U.S., Amazon EC2 services are available from servers based in Asia, Europe and South America.² For the purpose of this paper, I focus on the North America region.

AWS web archives are available from the Wayback Machine Internet Archive (<https://archive.org/web/>), which stores the same pricing webpage at different points of time. This allows me to construct a dataset, which pools price and characteristics of EC2 instances from 2009 to 2015. From the data collection process, I found that price changes usually occur once a year. Therefore, I collected data with a semi-annual frequency from 2009-2015 (every June and December). All the data are collected manually as it is hard to predict entries of new product and using code to extract these data might miss important insights. Furthermore, technical issues related to the web archive service (Way Back Machine) prevented me from using any sort of automated process and required maximum caution to avoid incorporating bad data.

As discussed previously, EC2 instances are available in from servers based in various regions and various operating systems (Linux, Windows etc.), and the dimensionality of the data grows larger when new regions and operating systems are

² Differences in privacy laws across countries have led AWS to base servers in a wide range of locations.

introduced. If we track the entire pool of EC2 instances (including all regions and all systems) over multiple time periods, we would need to collect thousands of observations. To avoid an unmanageable data collection process and to increase the efficiency of this project, I chose to focus on a subgroup of products that can represent the overall price trend. I narrowed my focus to AWS EC2 On-Demand Linux services from servers based in Virginia from 2009-2015. I will explain my reason behind choosing this particular subgroup step by step.

- On-Demand: As discussed in Section 2, there are three payment options for renting an EC2 instance—on-demand, reserved, and spot. All three options are available from servers based in various regions and running different operating systems. For first time users, the on-demand option is preferable, as it lets you rent the compute capacity by the hour without long-term commitments or upfront payment. Reserved Instances provide customers with a sizeable discount compared to the on-demand option, but with an upfront payment and a commitment to long-term usage. The spot option charges customer an hourly rate that is slightly lower than the on-demand rate, but the price fluctuates based on supply and demand for instances. Customers can stipulate a maximum hourly price they are willing to pay for an instance, but their service will be shut down once the market price exceeds this maximum. As we can see, the reserved option and the spot option are priced based on the on-demand option, with some discount or flexibility features. Therefore, it is reasonable to emphasize the on-demand pricing option and collect price data accordingly.

- Linux: EC2 instances were available for Linux and Windows operating systems when first introduced into the market. In the following years, more operating system options were added. As of 2016, all instances are available for seven different operating systems.³ Customers can choose the operating system that is compatible with their business needs. From the data collection process and spot checking across a wide range of data, I observed that although prices for instances differ across operating systems, the price trends of instances running different operating systems appear to be very similar. Linux is a very popular operating system for which instances have been available constantly throughout the time periods from 2009 to the present. Therefore, I chose instances running Linux as representative of the instances across other operating systems.
- Virginia: I chose instances running on servers based in Virginia for similar reasons that I chose to focus on the Linux operating system. Amazon first started to make EC2 instances available on servers running in Virginia, and I observed that when new instances enter the market, the first region where these new instances become available was Virginia. Although prices for the same instance differ across different regions, the price trends for different regions appear to be very similar. Therefore, I narrowed my focus to Virginia for its superior ability to quickly capture new products and to cover the maximum possible time period.

³ The seven systems are: Linux, RHEL, SLES, Windows, Windows with SQL Standard, Windows with SQL Web and Windows with SQL Enterprise.

While my data selection provides a sample of AWS EC2 prices, I believe that my sample is representative of the broader universe of prices. In the next section, I will discuss the process of organizing and cleaning these data.

4.2. Data Construction and Imputation

As previously discussed, I collected prices of on-demand Linux instances from servers based in Virginia from 2009-2015 semiannually, along with data on the characteristics of the instances. I will now use an example to clarify my data construction.

Suppose we want to identify the price of an On Demand Standard Small (API Name: m1.small) instance in Dec 2010. The price of this instance can be found on the web (from an archived version of AWS web pages) as shown in Figure 3. The characteristics associated with this instance can be found on a separate webpage that describes the instance (Figure 4).

On-Demand Instances

On-Demand Instances let you pay for compute capacity by the hour with no long-term commitments. This frees you from the costs and complexities of planning, purchasing, and maintaining hardware and transforms what are commonly large fixed costs into much smaller variable costs.

The pricing below includes the cost to run private and public AMIs on the specified operating system ("Windows Usage" prices apply to both Windows Server® 2003 and 2008). Amazon also provides you with additional instances for Amazon EC2 running Microsoft, Amazon EC2 running SUSE Linux Enterprise Server and Amazon EC2 running IBM that are priced differently.

US – N. Virginia	US – N. California	EU – Ireland	APAC – Singapore
Standard On-Demand Instances		Linux/UNIX Usage	Windows Usage
Small (Default)		\$0.085 per hour	\$0.12 per hour
Large		\$0.34 per hour	\$0.48 per hour
Extra Large		\$0.68 per hour	\$0.96 per hour

Figure 3

Available Instance Types

Standard Instances

Instances of this family are well suited for most applications.

Small Instance – default*

1.7 GB memory
1 EC2 Compute Unit (1 virtual core with 1 EC2 Compute Unit)
160 GB instance storage
32-bit or 64-bit platform
I/O Performance: Moderate
API name: m1.small

Figure 4

For On-Demand Small instances, I observe a price quote at \$0.085 per hour on 12/02/2010 from AWS's webpage archive. The instance's characteristics values are: 1.7GB memory, 1 EC2 Compute Unit, 160GB instance storage, 32 or 64 bit platform, and a moderate I/O performance. This information comprises one observation in my data set.

Another notable observation is Amazon's tendency to separate storage from instance in recent years. Starting from late 2013, some AWS EC2 instances are introduced as Amazon Elastic Block Store (EBS) instances. An Amazon EBS volume is a durable, block-level storage device that customers can create (at his or her own choice) and attach to a single EC2 instance. An EBS volume relies on SSD storage. According to Amazon, EBS offers customers flexibility in storage volumes, as customers can create EBS volumes up to 16TiB⁴ in size.⁵ Customers have the flexibility to choose between different size EBS volumes and pay for storage separately. Different size EBS volumes

⁴ 16 TiB = 17592 GB

⁵ <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html>

have different prices. Therefore, for instances that adopt EBS, customers pay for the instance that comes with no storage at an hourly price, and then pay for the EBS volume they choose to attach to the instance separately. Since these instances do not come with any storage, I set their storage values to zero.

5. Empirical Results

5.1. Matched Model Method

To observe a preliminary price trend of AWS Cloud Computing Services, I constructed a Laspeyres matched-model index. Since information on the quantity purchased of different AWS services is not available to the public, I assume an equal weight for each product in the sample. The Laspeyres index is the equal weight average of price changes. The equation

$$I_{0,1}^L = \sum_{m=1}^M (P_{m,1} Q_{m,0}) / \sum_{m=1}^M (P_{m,0} Q_{m,0}) \quad (1)$$

can be rewritten as:

$$I_{0,1}^L = \sum_{m=1}^M (w_{m,0} P_{m,1} / P_{m,0}) \quad (4)$$

where

$$w_{m,0} = (P_{m,0} Q_{m,0}) / \sum_{m=1}^M (P_{m,0} Q_{m,0}). \quad (5)$$

The fixed bundle consists of the seven products that have been available since December 2009. The Price index and the percent change from previous period are presented in Table 1.

Time Period	200912	201006	201012	201106	201112	201206	201212	201306	AAGR
% Change from previous period (at annual rate)	Base period	0.00%	0.00%	-9.30%	-12.05%	0.00%	0.00%	-30.64%	-7.43%
Laspeyres Index	1	1	1	0.952	0.893	0.893	0.893	0.744	--

Table 1

In figure 5, the Laspeyres index shows a clear and gradual reduction in prices from 2009 to 2013. On average, the price drop occurs every two periods (1 year). The average annual growth rates (AAGR) is -7.43%. Figure 6 displays the price trend for all

seven products within the bundle; the reduction patterns are very similar. However, this price index may understate the actual price declines because of new goods bias. During 2009-2013, 11 new instances became available but the Laspeyres index failed to capture these new goods. In order to confirm this result, I will compare the hedonic regression index and Laspeyres index in next chapters.

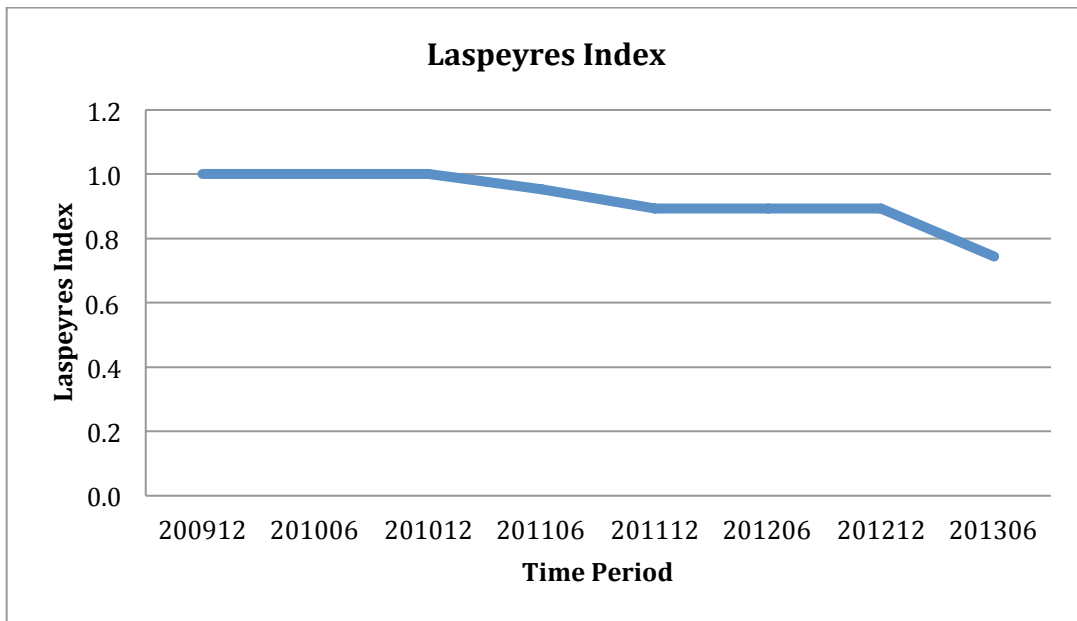


Figure 5

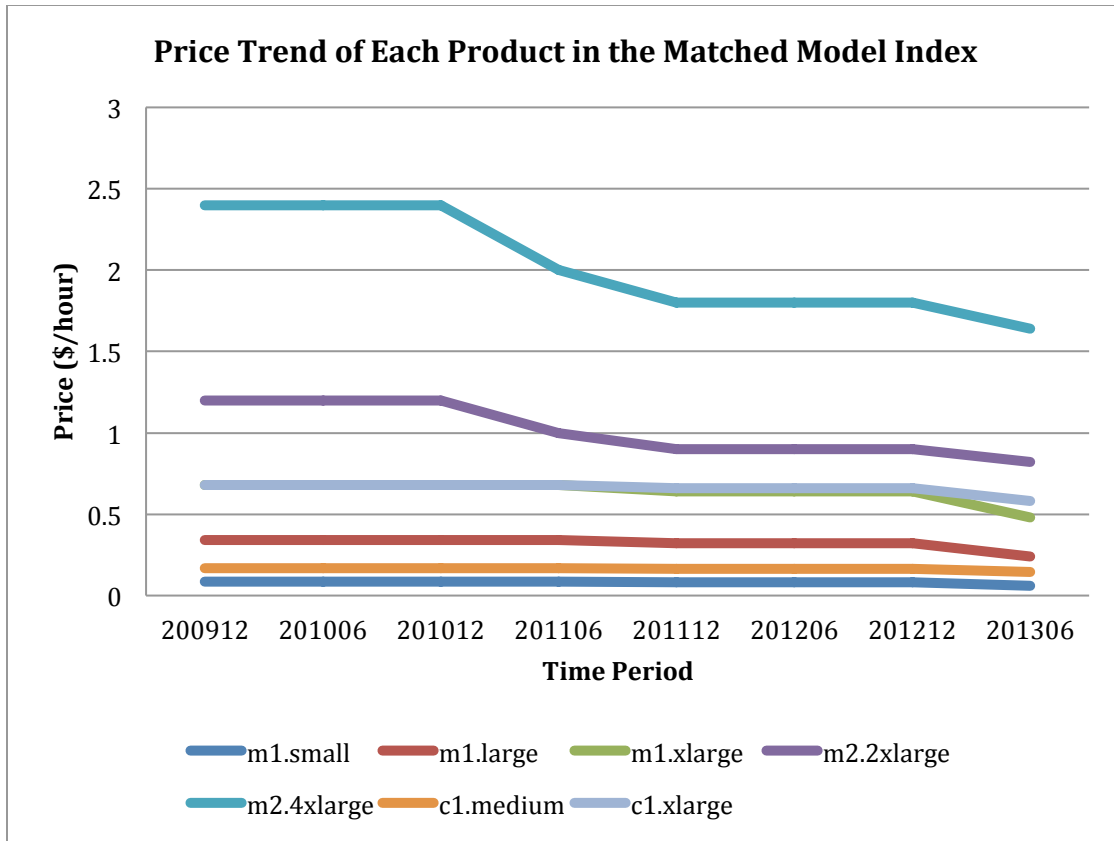


Figure 6

However, this Laspeyres index cannot be extended beyond June 2013. In this case, the flaws in a fixed-based index become troublesome. In late 2013, Amazon started to restructure their product types on a large scale. As a result, almost all the old types of instances are eliminated and a large number of new types of instances are introduced. As the old products that are used to construct the Laspeyres index are no longer available in 2013, the Laspeyres index stops at 2013.

5.2. Hedonic DV Method

As discussed previously, the hedonic regression used in the DV method takes the following form:

$$h_{DV}(x_i): \ln P_{i,t} = \alpha + \sum_k \beta_k X_{k,i,t} + \sum_t \delta_t D_{i,t} + \varepsilon_{i,t} \quad (2)$$

where $\sum_k \beta_k X_{k,i,t}$ capture product characteristics and $\sum_t \delta_t D_{i,t}$ are time dummies for each period. Differences in the coefficients of the time dummies give constant-quality price changes. I estimate equation 2 with my price and characteristic data for EC2 instances, using memory, number of EC2 Compute Units (ECU), storage, storage type, Input/output performance level and platform. The bias-adjusted and unadjusted Hedonic DV Method Price Index and the percent change from previous periods at an annual rate are presented in table 2 and figure 7. Table 3 shows the hedonic regression result.

	200912	201006	201012	201106	201112	201206	201212	201306	201312	201406	201412	201506	201512	AAGR
% Change from previous period (at annual rate)	Base period	-5.50%	-1.43%	-6.51%	-20.63%	3.90%	0.00%	-5.79%	1.43%	-60.71%	0.00%	0.00%	0.00%	-7.94%
Hedonic DV Method Price Index	1	0.972	0.965	0.933	0.831	0.847	0.847	0.822	0.828	0.519	0.519	0.519	0.519	--
% Change from previous period (at annual rate) - Adjusted	Base period	-5.63%	-1.47%	-6.66%	-21.09%	3.99%	0.00%	-5.93%	1.46%	-61.82%	0.00%	0.00%	0.00%	-8.10%
Hedonic DV Method Price Index - Adjusted	1.000	0.971	0.964	0.932	0.828	0.844	0.844	0.819	0.824	0.509	0.509	0.509	0.509	--

Table 2

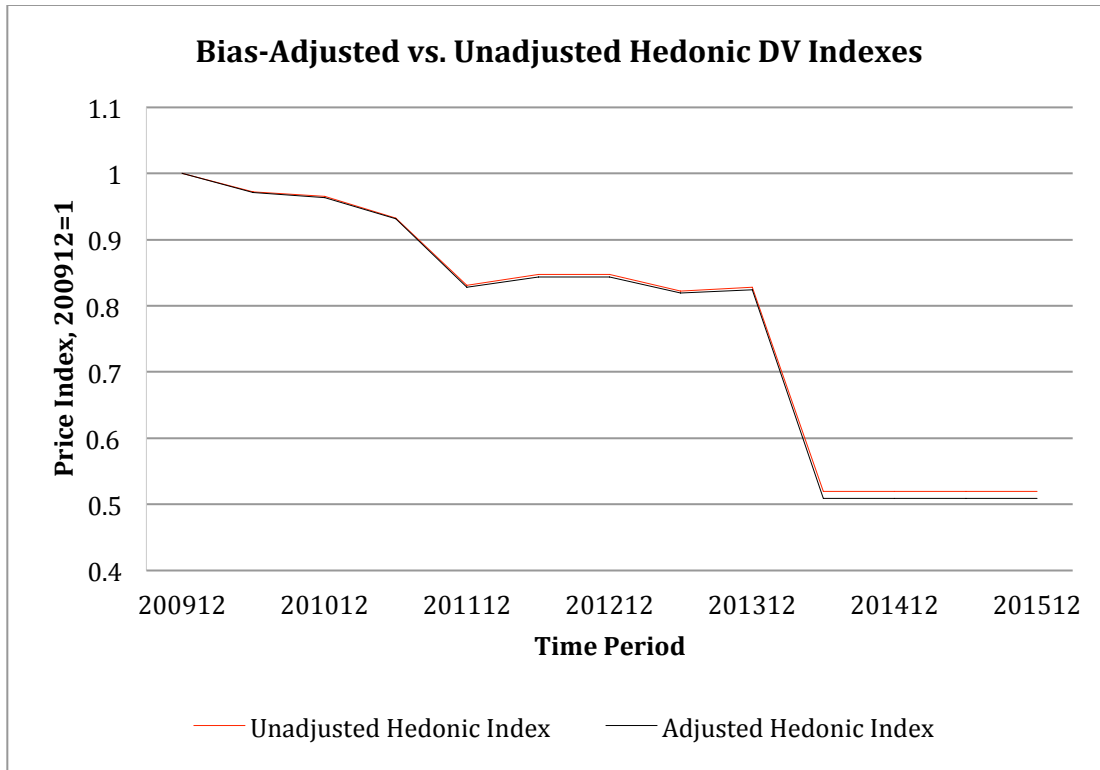


Figure 7

Variables	Hedonic Model	t-value
Memory	0.459***	(20.48)
ECU	0.441***	(16.36)
Storage	0.0725***	(10.25)
Platform	-0.194**	(-2.98)
I/O performance	0.148***	(5.21)
Storage Type	-0.113*	(-2.12)
Constant	-3.652***	(-25.88)
Observations	277	
Adjusted R ²	0.9792	

Notes: Natural logs are used for the Price, Memory, ECU, and Storage variables. The coefficients of year dummies are not presented here.

t statistics in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3

As presented in Table 3, all characteristics in the specification are statistically significant. With the exceptions of Storage Type and Platform, all the characteristics have positive coefficients. The adjusted R^2 of the model is 97.9%, which confirms that 97.9% of the variation in price can be explained by these characteristics and year dummy variables. Among all characteristics variables, Memory and ECU appear to be the most impactful, as Memory and ECU have the largest impact on R^2 when excluded from the model.

Overall, AWS Cloud Computing Services have experienced a sharp price decline from 2009 -2015. The price decreased gradually and modestly during 2009-2013 with small fluctuations. A visible price increase occurred in late 2013 due to Amazon's large scale restructuring of its product system – a large number of new products were introduced while almost all old products were eliminated. The price plummeted in early 2014 with a 38% drop from previous period (at annual rate 62%) and has remained unchanged ever since. The 38% drop is a price decline that cannot be observed in a Laspeyres model and can only be quantified in the context of a hedonic model. Overall, the bias-adjusted AAGR is -8.10%, while the unadjusted AAGR is -7.94%.

With the above price trend result of AWS Cloud Computing Services, we can proceed and ask more questions – Did some of the characteristics' effect on price change become bigger over the years? What is the explanation for the drastic 38% (at annual rate 62%) price drop that occurred in early 2014? Section 5.3 and 5.4 will provide more insights for these questions.

5.2.1. Comparison between Laspeyres Index and Adjusted Hedonic DV Index

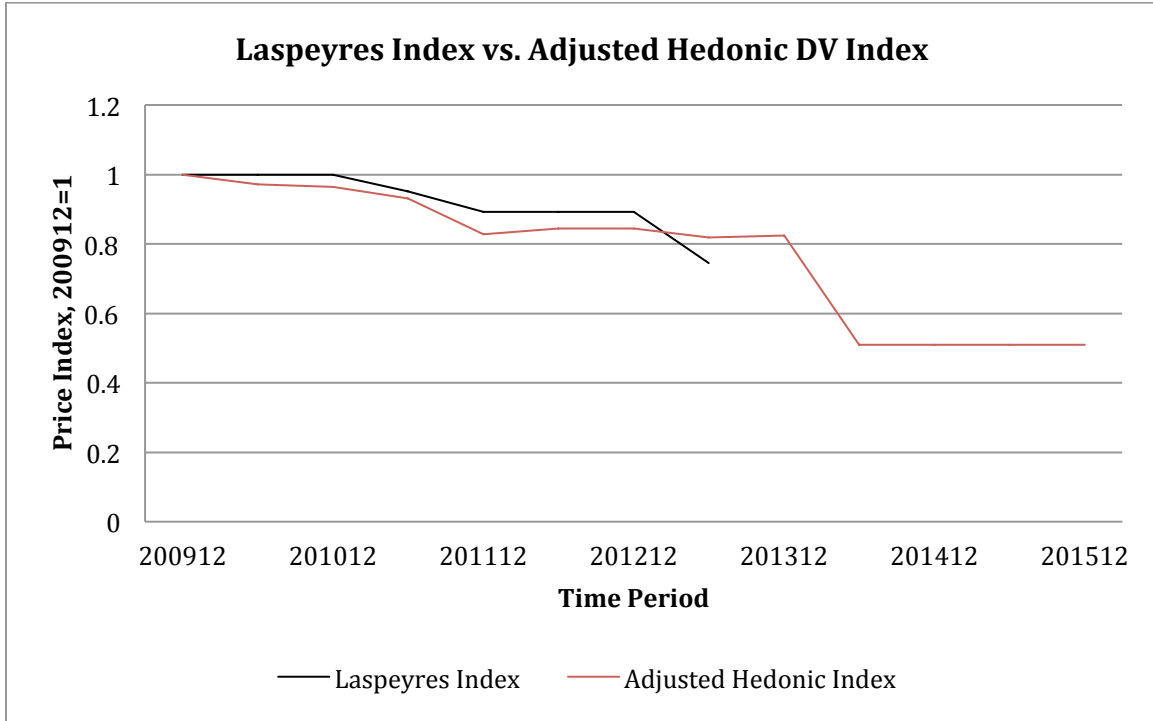


Figure 8

As shown in Figure 8, the obvious advantage of using a Hedonic index is its ability to more fully control for quality change and to correctly capture the quality of products entering and exiting the market. For example, although both indexes indicate that from 2009 to 2013, the price has declined gradually, for periods like June 2010, December 2010, June 2012 and December 2012, the Laspeyres Index fails to show any price fluctuation, while the Hedonic DV Index captures the price decline. The empirical results also are consistent with the frequent observation that Laspeyres indexes underestimate the price decline, because Laspeyres does not allow any substitution between products when relative prices change. The comparison in figure 8 also suggests that for the IT industry in general, fixed-based indexes are problematic and distorting.

5.3. Adjacent Period Method

The Adjacent Period Method employs a similar approach with Hedonic DV Method. As discussed in Section 3, the adjacent period method is running the DV regression separately for each pair of adjacent periods. For $t \in \{t_j, t_{j+1}\}$

$$h_{adj}^{\{t_j, t_{j+1}\}}(x_{i,t}): \ln(P_{i,t}) = \alpha + \sum_k \beta_k \ln(x_{k,i,t}) + \delta_t D_{i,t(j+1)} + \varepsilon_{i,t} \quad (3)$$

For n periods, the Adjacent Period Method runs $(n-1)$ DV regressions. Since the time periods for each regression are adjacent, we only have one time dummy, the coefficient of which directly captures the constant-quality price change from the previous period. From time dummies of all regressions we can construct an alternative hedonic price index. By letting each pair of adjacent period have a unique hedonic function, the Adjacent Year Method captures the unique relationship between characteristics and price for each pair of adjacent periods, thus is more accurate than the DV method. The average of adjusted R^2 value of Adjacent Period regressions is 98.43%, which is higher than the adjusted R^2 (97.92%) in DV Method. The results for bias-adjusted Hedonic Adjacent Period Method Price Index are listed in Table 4.

As we can observe, the AAGR of Hedonic Adjacent Period Method Price Index is -7.33%. Figure 9 displays the comparison between the (adjusted) Hedonic DV Index and the (adjusted) Hedonic Adjacent Period Index. Both indexes show very similar price trends, however, the DV index tends to underestimate the price level.

	200912	201006	201012	201106	201112	201206	201212	201306	201312	201406	201412	201506	201512	AAGR
Adjusted % Change from previous period (at annual rate)	Base period	-2.32%	4.28%	-6.55%	-19.63%	0.03%	0.00%	-15.46%	12.70%	-60.98%	0.00%	0.00%	0.00%	-7.33%
Adjusted Hedonic Adjacent Period Index	1	0.988	1.009	0.976	0.875	0.875	0.875	0.804	0.854	0.533	0.533	0.533	0.533	--

Table 4

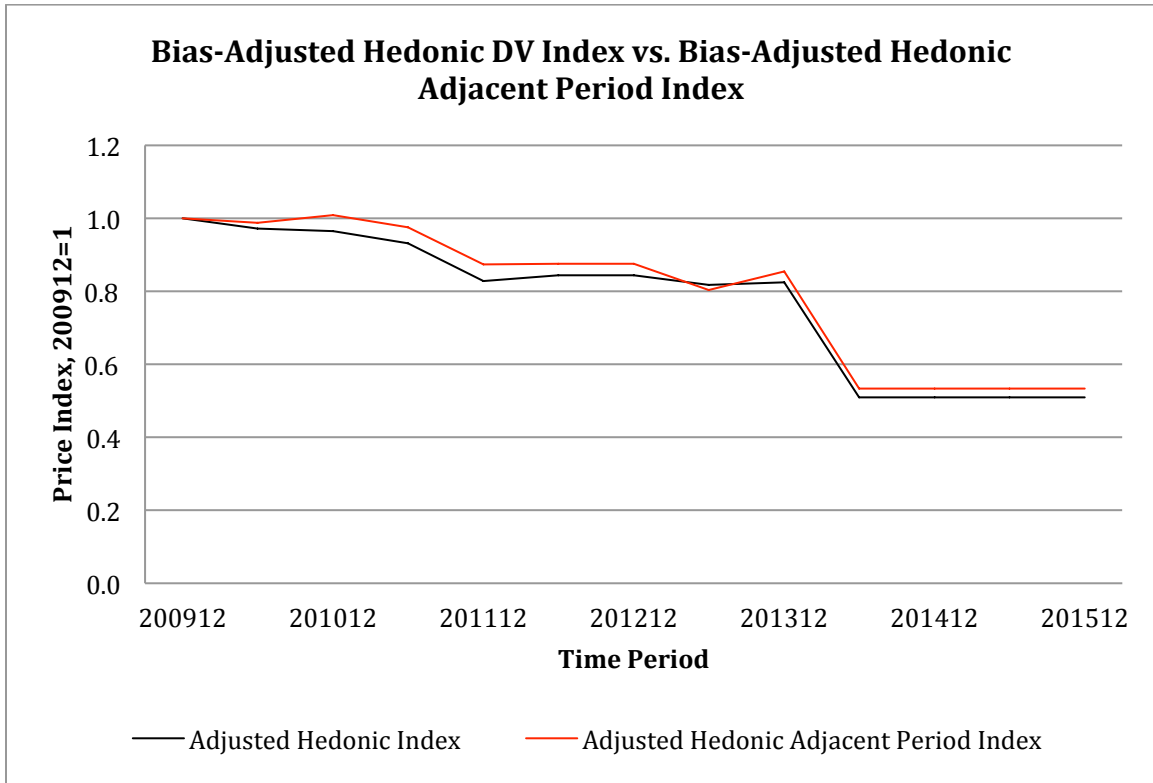


Figure 9

Another interesting observation from the Adjacent Period Regressions is the trend of characteristics coefficients β 's. The DV Method only provides only one β for each characteristic as DV Method includes all time periods in one regression, but the Adjacent Period Regressions provide a series of β 's for each characteristic over time periods. A drastic price drop can be caused by a sharp increase in "utility per dollar" obtained by customers. Under this circumstance, customers pay the same price, but get more quality per dollar spent. To understand what kind of quality increases can have a substantial

effect on price changes, I need to track the coefficients on characteristics over time.

Without the Adjacent Period Regressions, this analysis could not be done. Figure 10 plots all six characteristics coefficients from 2010-2015.

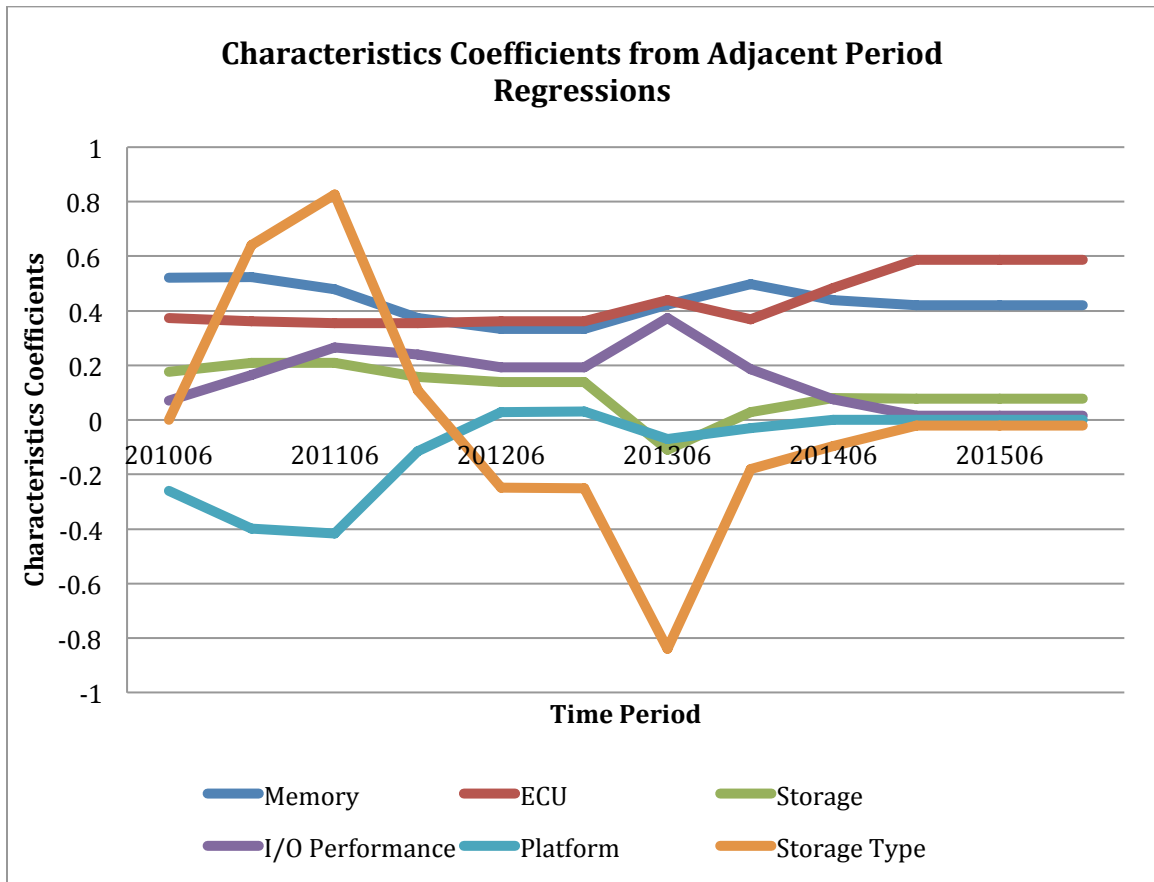


Figure 10

From this figure, we can find that among all characteristics, Memory and ECU's coefficients have relatively large magnitude and display consistency over the years. Thus, a sizeable increase in Memory or ECU can provide an explanation for drastic price change. On the other hand, the coefficients of Storage, Platform, I/O Performance and Storage Type converge to zero after June 2013. By the end of the 2015, all four coefficients approach zero. This result indicates that after 2014, an increase in Storage,

Platform, I/O Performance or Storage Type would barely have any effect on price. The increase in these characteristics cannot explain the huge drop that occurred in June 2014. The declining magnitude of storage coefficient also aligns with Amazon's effort in recent years to separate storage from instances and shift towards EBS storage volumes.

5.4. Explanation of the Drastic Change of AWS Products in 2014

As discussed previously, Amazon restructured their product system on a large scale from late 2013 to early 2014. All the old types of instances were eliminated and numerous new types of instances were introduced. This change caused a drastic 38% (at annual rate 62%) price decline. There are two possible explanations for this decline:

- The overall posted price level for EC2 Instances decreased sharply, but customers got similar amount of quality per dollar spent.
- The overall posted price level of EC2 Instances did not decrease sharply; the price decline was implicit as customers got more quality for each dollar they spent.

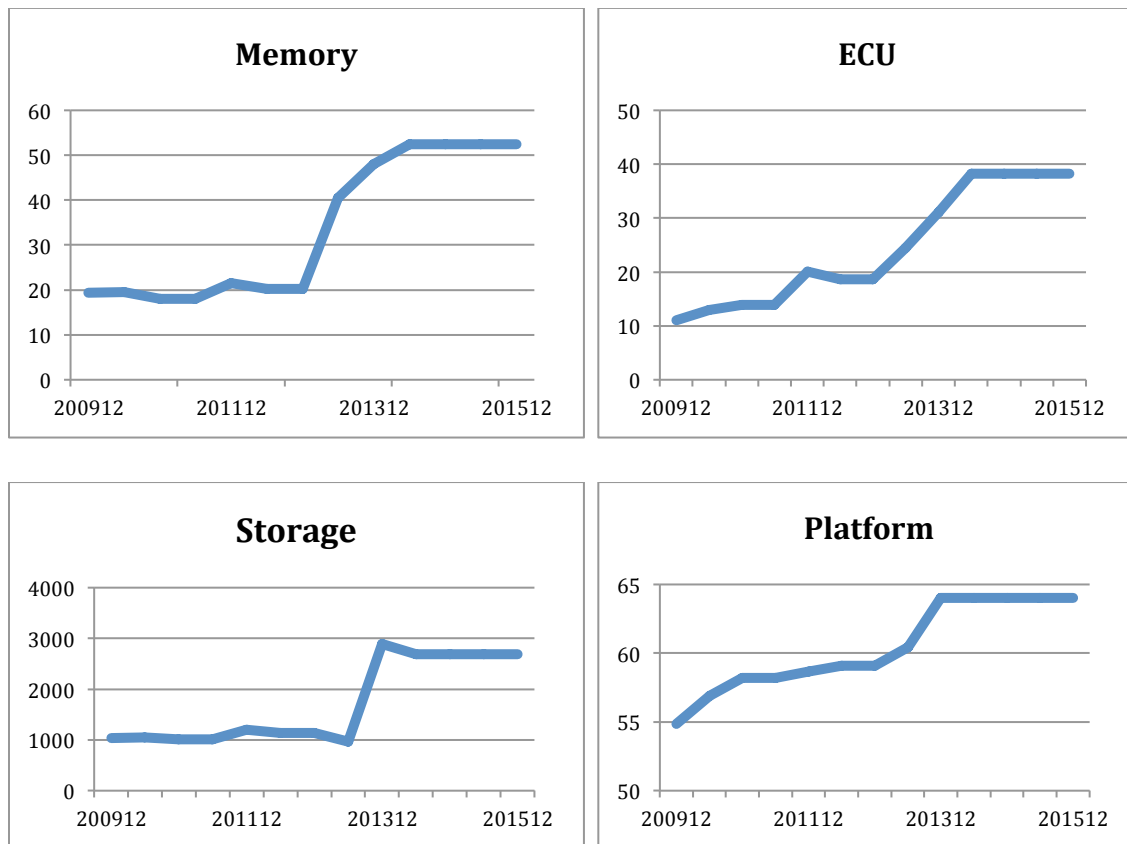
To examine the first possibility, I plotted the average price of all instances in each period. As shown in Figure 12, compared to the time periods before 2013, the average price for all instances is generally higher after 2013. On average, customers pay more for each instance after the restructuring of product system. Although the average price of all instances is not a perfect measure, it tells us that the drastic change in 2014 is unlikely a result of an explicit drop in posted prices.

Thus, I turn to examine the second option – that customers get more quality for each dollar they spent after 2013. In Section 5.4 we arrived at the conclusion that increase in Memory and ECU will have a sizeable impact on price, while increases or decreases in

other characteristics barely have any effect on prices. In figure 11, I plot the average characteristics value of all instances over the years (for each characteristic separately).

As shown in Figure 11, the average values of both Memory and ECU experience a sharp increase in 2013. On average, customers get more Memory and ECU with each instance after 2013. This sharp increase in quality provides an explanation for the drastic price drop in 2014 – customers get more quality for each dollar they spent.

It is noticeable that after the huge price drop in 2014, Amazon has formed a relatively well-developed system of Cloud Computing services, and prices for the instances in my price index have remained the same from 2014 to present (2016).



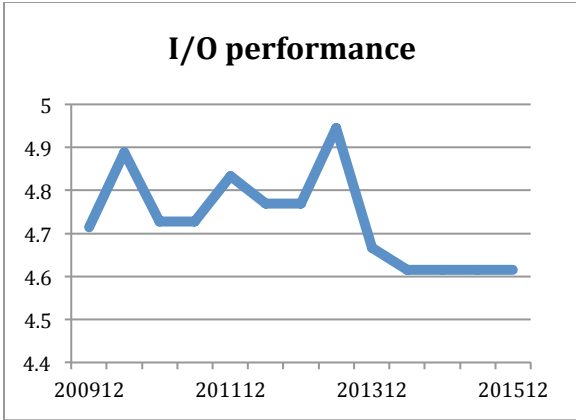


Figure 11

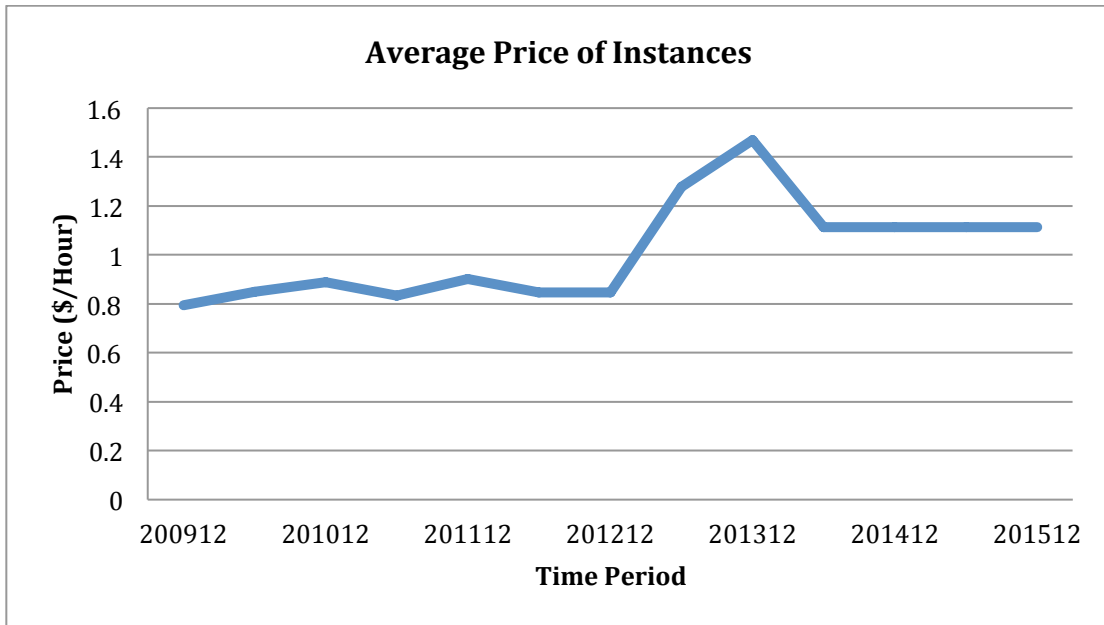


Figure 12

6. Comparison of Price Trend between Computing Products and Cloud Computing Services

It is also interesting to see if Cloud Computing Services share a parallel price trend with Computing Products (hardware and software) in general. Amazon needs to purchase Computing Products in order to construct server farms and provide Cloud Computing Services, and that suggests that the price of Computing Products should be highly correlated with the cost of Cloud Computing Services. Therefore, it is reasonable to expect that the two indexes would share a parallel price trend. Based on BEA's information of nominal shares and price indexes of private investment in hardware and software, I was able to construct a weighted average price index of hardware and software from 1980 to present. Figure 13 plots the weighted average price index along with my adjusted Hedonic DV Price Index. Figure 14 plots both indexes in logged forms in order to observe the price reduction trend. As observed in Figure 14, Cloud Computing Services and Computing Products share a similar price reduction trend, with Cloud Computing Services experiencing a sharper decline.

It is reasonable to see the price reduction trend of Cloud Computing Services from 2009-2015 as a continuation of Computing Products price reduction. As shown in Figure 14, Cloud Computing Services Index is an extension of the Computing Products Index, assuming the reduction pace is same before and after 2009. However, the price of Computing Products has been declining at a diminishing rate for about a decade. This pattern suggests two possibilities. First, as an innovative new IT products, Cloud Computing Services have been taking advantage of the price reduction trend of Computing Products in previous periods, but will eventually get caught up by the

diminishing rate of cost reduction. In other words, the drastic price drop of Cloud Computing Services in recent years is unlikely to be observed again in the future.

An alternative explanation, mentioned in Byrne, Oliner and Sichel (2015), is that the price trend for Computer Products is mismeasured and that those prices have been falling faster than shown by the blue line.

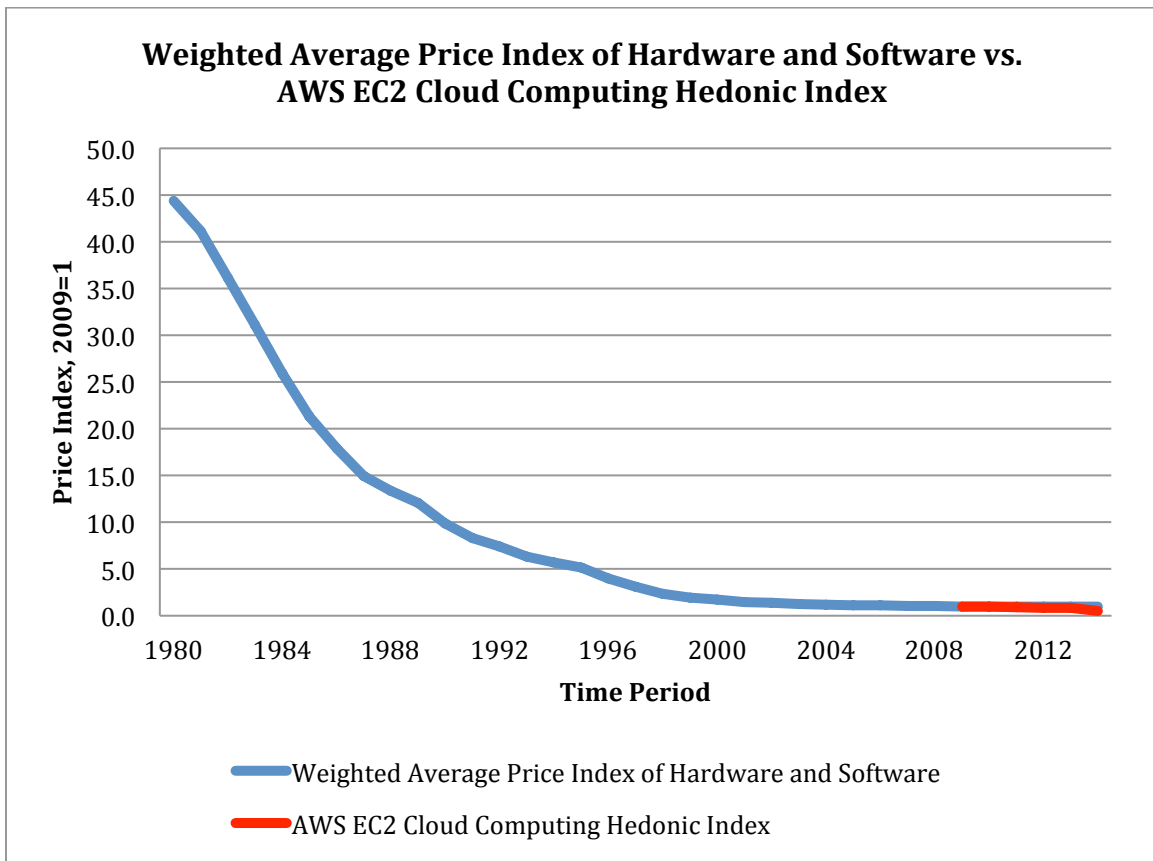


Figure 13

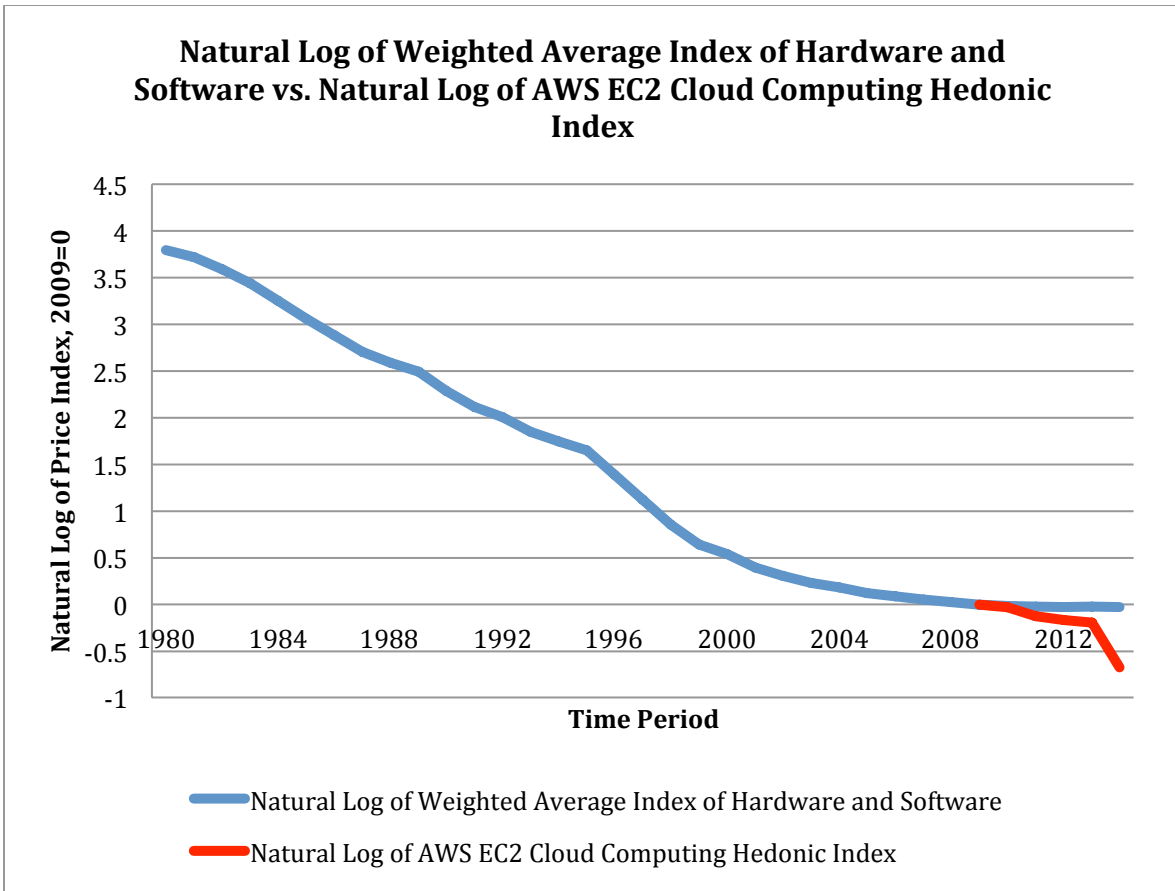


Figure 14

7. Possible Issues

Issues may arise from the methodology I adopted in this paper. First, my hedonic price indexes are constructed based on a small sample of data available. Because of the time constraint of this project, I chose to collect data of a group of instances that I considered to be a logical representation of the population. However, selection bias may occur and my results could be affected if more dimensions are incorporated.

Another issue is related to the possible flaws in hedonic method. The underlying assumption of my hedonic model is that the characteristics I defined precisely capture all the quality features of an instance. However, the set of characteristics might not be comprehensive at all. Producers may have other measures of quality features that factor into the pricing of instances, but these “hidden” characteristics are not available to the public. However, I would argue that compared with matched model method, hedonic methods yield more accurate results.

8. Conclusion

In this paper, I presented estimates for quality-adjusted price indexes of AWS Cloud Computing Services using the matched model method and the hedonic methods. For the hedonic methods, both DV Method and Adjacent Period Method suggest that AWS EC2 service has experienced a sharp price decline from 2009 -2015. The Adjusted DV Method Index shows an AAGR of -8.10% , and the Adjusted Adjacent Period Method Index shows an AAGR of -7.33% .

The price decreased gradually and modestly during 2009-2013 with small fluctuations. A slight price increase occurred in late 2013 due to Amazon's large scale restructuring of its product system – a large number of new products were introduced while almost all old products were eliminated. The price plummeted in early 2014 with a 38% (at annual rate 62%) drop from previous period and has remained unchanged through the end of 2015. The matched model yields a similar result from 2009-2013, but was unable to extend beyond 2014, for Amazon eliminated all the products in the fixed bundle and replaced them with new products. Compared to the price trend of overall Computing Products (hardware and software), Cloud Computing Services have experienced a slightly sharper price reduction trend over the years.

The results above offer helpful insights to future studies on price trends of innovative IT products and the relationship between these new products and the overall development of IT industry.

References

Aizcorbe, Ana. *A Practical Guide to Price Index and Hedonic Techniques*. N.p.: n.p., n.d. Print.

Byrne, David M., Stephen D. Oliner and Daniel E. Sichel, "How fast are semiconductor prices falling?", *NBER Working Paper*, 2015. <http://www.nber.org/papers/w21074.pdf>

Byrne, David M., Stephen D. Oliner, and Daniel E. Sichel, "Is the Information Technology Revolution Over?" *International Productivity Monitor*, 25, 20–36, 2013.

Getsova, Mariya, "A Quality Adjusted Price Index for Desktop Personal Computers", Wellesley College senior thesis, 2015.

Hennessy, John L., and David A. Patterson. *Computer architecture: a quantitative approach*. Elsevier, 2011.

Pakes, Ariel. "A Reconsideration of Hedonic Price Indexes with an Application to PC's." *American Economic Review* 93.5 (2003): 1578-596. Web.

Sun, Sophie, "What are We Paying For? A Quality-adjusted Price Index for Laptop Microprocessors," Wellesley College senior thesis, 2014.