Revenue Forecasting in Technological Services: Evidence from Large Data Centers

Jyotirmoy Sarkar, Snehanhsu Saha, *Senior Member, IEEE,* Saibal Kar, Bidisha Goswami, *Member, IEEE* and Anand MN, *Member, IEEE*

Abstract—Data Center is a facility, which houses computer systems and associated components, such as telecommunications and storage systems. It generally includes power supply equipments, communication connections, and cooling equipments. A large data center can use as much electricity as a small town. In todays world due to the emergence of data center based compute services, it has become necessary to examine how the costs associated with data centers evolve over time, mainly in view of efficiency issues. In this paper, we present a general approach to optimizing the revenue/cost of data centers using CES production function. Next, we develop the optimization model for data center. We utilize the mathematical basis of CES functions and apply the same for cost optimization and revenue forecasting in data centers. We have applied factorial analysis in our model and tried to find out the factors, which are significantly contributing towards the response variables. For illustration and clarity, we offer a number of graphs generated over real-time data set, which calculates the precise cost elasticity of the optimization exercise and identifies the factors that maximize revenue significantly. The results are quite useful in view of production reorganization in large data centers around the world.

Index Terms—CES production function, data center, concavity, revenue function, profit function.

1 Introduction

A data center is a facility, which houses a large number of computing equipments like servers, routers, switches and firewalls. Along with servers, routers supporting components like air conditioning, backup equipment, and fire suppression facilities are required to run a data center.A data center can be complex if it requires a dedicated building or simple in case it is residing in a room with few servers. A data center can be shared by multiple organizations(shared data center) or it can be owned by the single organization(private data center).Due to the emerging fields of parallel computing and evolving data centers cloud computing services are becoming popular. Cloud computing can be defined as a pool of computer resources, which provide computing function as utility services. Microsoft, Google, Amazon, and IBM are few major leading IT companies in cloud computing field and they are increasing their footprint by investing significantly in data distribution and computational hosting [2].

In recent years, a lot of investments have been made in data centers to support cloud by large organizations. In a competitive world, IT organizations are under pressure to provide services at low cost and faster way. Cloud computing has become popular as it can provide a simultaneous way to lower costs, increase responsiveness and flexibility, and improve the quality of service. But over the last 10 years server cost, power cost and maintenance cost of a data center rose drastically. According to the Gartner group survey, energy consumptions incurs up to 10% of the current data center operational expenses(OPEX) and this cost may surge up to 50% in near future [3]. The power consumed by the computing systems dissipates as heat.A major slice of total heat approximately 70% is generated by the data center infrastructure[5]. To mitigate the heat, a cooling system is required. The cost of the cooling system may range between $2 to $5 million per year for a classical data center[4]. Failure to keep temperatures within operational limit may disrupt cloud services, which can result in SLA violation.In cloud data center, automation is mandatory and is the functional principle of design[2]. The ratio of stuff members to servers in a typical well established data center is 1:1000. Examining the different segments of data center cost may help us to answer a simple question "where does all the cost go in?". We will take help of James Hamilton's paper to quantify the cost structure. We are considering a data center which houses 50000 servers and built with superior quality, using state-of-the-art techniques and equipments. Table 1 provides major cost segments associated with the data center. Costs are amortized, i.e., one-time investment is allocated over reasonable time-frame, the opportunity cost of investment here to be 5%.[1]. managements of the IT organizations

Amortized Cost	Component	Sub-component
$\tilde{4}5\%$	Servers	CPU, memory, storage systems
$\tilde{2}5\%$	Infrastructures	Power distribution and cooling
15%	Power draw	Electrical utility cost
15 %	Network	Links, transit, equipment

TABLE 1
Cost segmentation in Datacenter

- *J. Sarkar is with GE Healthcare, Bangalore, 560066.India. E-mail: jyotirmoy.sarkar@ge.com*
- *S. Saha is with Center for Applied Mathematical Modeling and Simulation (CAMMS) & Department of Computer Science and Engineering, PESIT-BSC, Bangalore, 560100, India. E-mail: snehanshusaha@pes.edu*
- *B. Goswami is with the Department of Computer Science and Engineering, PESIT-BSC, Bangalore, 560100, India. E-mail: bidishagoswami@pes.edu*
- *Anand MN is with the Department of Computer Science and Engineering, BITS Pilani, Hyderabad, India. E-mail: anandmn@gmail.com*

Manuscript received April 19, 2005; revised August 26, 2015.

are looking for a solution, which will produce maximum revenue by minimizing the cost. CES production function is the backbone of our proposed model.The CES production function is a neoclassical production function that exhibits constant elasticity of substitution. In other words, the production technology has a constant percentage change in factor (e.g. labour and capital) proportions due to a percentage change in a marginal rate of technical substitution.In this paper, we try to optimize the revenue,essentially by minimizing the cost while using CES production function. We have considered server cost, power& cooling, and infrastructure cost as an input to the cost function while the physical inputs go into the production function. The flexibility of CES function is that it can be expanded with any number of input variables. We applied our model to real-time data set collected from various sources(we discuss this shortly) and we found the result quit relevant for the current scenario. We have demonstrated the elasticity range for which optimization can be achieved and we tried to prove the same by generating graphs based on the real-time data set. Factor analysis also has been applied to the factor input to find out if these contribute significantly towards revenue generation. Several statistical results and functions have been generated in support of factorial analysis.

The remainder of the paper has been organized as follows. Section II discusses the related work from the literature. In Section III, Revenue Optimization in Data Center is discussed. This section explains the mathematical foundation of the proposed model based on CES. Section IV offers the results and subsequent discussion in order to qualify the result and highlights the possible implementation of the proposed model for different datasets. Section V talks about the factors that provide an insight to revenue maximization in the CES model. Conclusion and future work have been explained in section VI and the appendix under section VII provides detailed derivations.

2 RELATED WORK

Revenue forecasting, whether for a country, a state or a firm, accommodates many allied considerations including substantial uncertainty regarding the respective fundamentals [18]. Notwithstanding, the importance of revenue forecasting for decision making within a given business is indisputable [19]. However, most such forecasting is associated with sales order recognition, which may be different from the operating revenue at a point in time. Indeed, the latter should be treated as a better predictor. Consequently, a number of approaches have developed such that the possible deviations associated with predictions and the observed values, ex post, is minimized. The use of asymmetric loss function to this end can isolate the forecast rationality and the costs associated with under-forecasting [20]. Alternatively, seasonal fractionally differenced ARIMA models with shorter lags can be quite useful as compared to long-run predictions and have been tried earlier. In this paper, since the purpose is to minimize cost in order to maximize revenue, we focus intensively on the inputs [21]. James Hamilton [6] has shown that, quite significantly, power is not the largest cost, if the amortization cost of power, cooling infrastructure for 15 years and new server amortization cost over 3 years are taken into consideration. He concluded that, cooling amortization and server amortization monthly payments have been computed using 5% per annum cost and server hardware costs are the largest. But power infrastructure cost rise and server hardware cost may fall, resulting in the domination of power cost over all other data center expenses in not so distant future.

Generally, a typical data center comprises 100 fully loaded racks with the current generation 1U servers needing $1.2 million for power and an additional $1.2 million for cooling infrastructure per annum. Moreover, $1.8 million annual cost is incurred due to maintenance, amortization of power and cooling equipment. Thus, power is the most significant cost of the data center while server hardware contributes to the biggest chunk of the total operating cost. The Constant Elasticity of Substitution (CES) production function is part of microeconomics and an early example appeared in Solow (1956) and Arrow et al (1961) rigorously studied the two input form of CES function. Hicks introduced a concept that, for a given output y (i.e., along an isoquant), with only two inputs the effect of a change in the factor price ratio on the input ratio can be characterized by a scalar. Blackorby and Russell (1989) argued that the main CES functions exhibit several drawback, instead the alternative the so-called Morishima Elasticity of Substitution (MES), which reduces to the original Hicksian concept in the two-dimensional case, is the natural one. Mahdi Ghamkhari in his paper investigates the trade-off between minimizing data centers energy expenditure and maximizing their revenue for various Internet and cloud computing services that they may offer.They also proposed a novel optimization-based profit maximization strategy for data centers for two different cases, without and with behind-the-meter renewable generators.Xiaobo Fan in his paper represent the characteristics of aggregate power usage of large collections of servers (up to 15 thousand) for different classes of applications over a period of approximately six months. They used modeling framework to estimate power management schemes to reduce peak power usage and the opportunities for power and energy savings are significant, but greater at the cluster-level (thousands of servers) than at the rack-level (tens). Gurumurthi et al. have tried to achieve an optimal tradeoff between energy efficiency and service performance over a set of distributed internet data centers with dynamic demand. Several effective techniques have been developed to limit the disk power consumption for laptops and workstations.But it may create new challenges to apply the same in data center.Dynamic rotations per minute approach to speed control in server disk arrays can be one of the approaches to save significant amount power consumption in I/O system without sacrificing the performance [7]. Since, energy consumption of cloud data centers is a key concern for owners ow- ing to rising energy costs (fuel), CO2 emissions related to this consumption have become relevant [8]. Therefore, saving money in the energy budget of a cloud data center, without sacrificing Service Level Agreements (SLA) is an excellent incentive for cloud data center owners, and would at the same time be a great success for environmental sustainability. The ICT resources, servers, storage devices and network equipment consume maximum power. Processors [9] are the main contributors to the server's power consumption whereas other components [10] like multiple level caches, RAM, I/O activities also contribute to the total power consumption of the server. The storage devices range from a single hard disk to SAN (Storage Area Network) devices, which consume a significant amount of power. The other significant contributors to power consumption are network equipment which include routers and switch fabrics.

Clearly, apart from the techniques associated with optimizing the performance of data centers, the problem of cost minimization remains an important issue. We are mindful of the cross-effects of reducing the use of one factor vis-a-vis another and apply the CES production function to see the impact of reducing the cost of all of the contributing factors on revenue maximization at the data centers.

3 REVENUE OPTIMIZATION AND DATA CENTERS

CES belongs to the family of neoclassical macroeconomics production functions. Cobb-Douglas is one of the form of constant elasticity substitution.The CES production for two input factors can be represented in the below form

$$Q(L,K) = (\alpha L^{\rho} + (1-\alpha)K^{\rho})^{1/\rho} \tag{1}$$

where
Q= Quantity of output
L, K =Input variables or factors
$\rho = \frac{s-1}{s}$
$s = \frac{1}{1-\rho}$, Elasticity of substitution
and α = Share parameter

The elasticity of substitution is constant for the CES production function. More specifically, the elasticity of substitution measures the percentage change in the factor ratio divided by the percentage change in the technical rate of substitution, whilst holding output fixed.

Case I: Linear production function:
When we set $\rho = 1$, the production function becomes: $y = K + L$, where the two inputs, capital and labor are perfect substitutes.

Case 2: Cobb Douglas production function:
When ρ tends to 0, i.e. $\lim_{\rho\to 0} y$, the isoquants of the CES production function look very much like those of the CobbDouglas production function. This can be shown a variety of different ways mathematically, but the easiest is to compute the technical rate of substitution. As such, the two inputs in this case are imperfect substitutes, where the production isoquants are downward sloping.

Case 3: Leontieff Production function:
When ρ tends to $-\infty$, i.e. $\lim_{\rho\to -\infty} y$, the production isoquants become Lshaped, which we associate with the perfect complements case for inputs.

3.1 Production Maximization

Consider an enterprise that has to choose its consumption bundle (S, I, P, N) where S, I, P and N are number of servers, investment in infrastructure, cost of power and networking cost respectively of a cloud data center. The enterprise wants to maximize its production, subjected to the constraint that the total cost of the bundle does not exceed a particular amount. The company has to keep the budget constraint in mind and keep total spending within this amount. The production maximization is done using Lagrangian Multiplier. The CES function is:

$$f(S,I,P,N) = (S^{\alpha} + I^{\alpha} + P^{\alpha} + N^{\alpha})^{\frac{1}{\alpha}} \tag{2}$$

Let m be the cost of the inputs that should not be exceeded.

$$w_1 S + w_2 I + w_3 P + w_4 N = m \tag{3}$$

w_1: Unit cost of servers
w_2: Unit cost of infrastructure
w_3: Unit cost of power
w_4: Unit cost of network

Optimization problem for production maximization
max f(S,I,P,N) subject to m

The following values of S, I, P and N thus obtained are the values for which the data center has maximum production of satisfying the given constraints on the total investment.

$$S = \frac{m w_1{}^{\frac{1}{\alpha-1}}}{w_1^{\frac{\alpha}{\alpha-1}} + w_2^{\frac{\alpha}{\alpha-1}} + w_3^{\frac{\alpha}{\alpha-1}} + w_4^{\frac{\alpha}{\alpha-1}}} \tag{4}$$

$$I = \frac{m w_2{}^{\frac{1}{\alpha-1}}}{w_1^{\frac{\alpha}{\alpha-1}} + w_2^{\frac{\alpha}{\alpha-1}} + w_3^{\frac{\alpha}{\alpha-1}} + w_4^{\frac{\alpha}{\alpha-1}}} \tag{5}$$

$$N = \frac{m w_3{}^{\frac{1}{\alpha-1}}}{w_1^{\frac{\alpha}{\alpha-1}} + w_2^{\frac{\alpha}{\alpha-1}} + w_3^{\frac{\alpha}{\alpha-1}} + w_4^{\frac{\alpha}{\alpha-1}}} \tag{6}$$

$$P = \frac{m w_4{}^{\frac{1}{\alpha-1}}}{w_1^{\frac{\alpha}{\alpha-1}} + w_2^{\frac{\alpha}{\alpha-1}} + w_3^{\frac{\alpha}{\alpha-1}} + w_4^{\frac{\alpha}{\alpha-1}}} \tag{7}$$

The above results are proved in Appendix 1.

3.2 Cost Minimization

Consider an enterprise that has a target level of output to achieve by investing a minimum amount.The CES function is of the form:

$$y_{tar} = f(S,I,P,N) = (S^{\alpha} + I^{\alpha} + P^{\alpha} + N^{\alpha})^{\frac{1}{\alpha}} \tag{8}$$

y_{tar} is the target output of the firm that needs to be achieved and w_1, w_2, w_3 and w_4 are unit prices of servers, infrastructure, power and network respectively. Cost minimization problem is:

$$\min_{S,I,P,N} w_1 S + w_2 I + w_3 P + w_4 N \; subject\; to\; y_{tar} \tag{9}$$

The cost for producing y_{tar} units in cheapest way is c, where

$$c = w_1 S + w_2 I + w_3 P + w_4 N \tag{10}$$

c can be written as:

$$c = \left(\frac{y_{tar}}{w_1^{\frac{\alpha}{\alpha-1}} + w_2^{\frac{\alpha}{\alpha-1}} + w_3^{\frac{\alpha}{\alpha-1}} + w_4^{\frac{\alpha}{\alpha-1}}}\right)^{\frac{1}{\alpha}-1} \tag{11}$$

The above results are proved in Appendix 2.

3.2.1 Global Minima for Cost minimization: A heuristic approach

Apart from the above calculation, Gradient Descent method has been used to retrieve the values of elasticities where cost minimization is ensured. For simplification of equations, let us consider two cost segments X and Y. w_1 and w_2 are unit prices of X and Y.Rewriting the cost function using the newly elected variables, we obtain

$$c = w_1 X + w_2 Y \tag{12}$$

The newly formed CES function is,

$$y_{tar} = (X^\alpha + Y^\alpha)^{\frac{1}{\alpha}}$$
$$y_{tar}^\alpha = X^\alpha + Y^\alpha$$
$$X^\alpha = y_{tar}^\alpha - Y^\alpha$$
$$X = (y_{tar}^\alpha - Y^\alpha)^{\frac{1}{\alpha}}$$

Putting the value of X in cost function 12, we obtain

$$c = W_1 \left(y_{tar}^\alpha - Y^\alpha\right)^{\frac{1}{\alpha}} + W_2 Y$$

$$\frac{\partial c}{\partial \alpha} = \frac{-w_1 \left(y_{tar}^\alpha - Y^\alpha\right)^{\frac{1}{\alpha}}}{\alpha^2} ln \left(y_{tar}^\alpha - Y^\alpha\right) \\ \left(y_{tar}^\alpha ln y_{tar} - Y^\alpha ln Y\right)$$

The above partial derivatives are used in gradient descent method for cost minimization.

Gradient Descent: Algorithm

1) **procedure** GRADIENTDESCENT()
2) $\frac{\partial c}{\partial \alpha} = \frac{-w_1 (y_{tar}^\alpha - Y^\alpha)^{\frac{1}{\alpha}}}{\alpha^2} ln \left(y_{tar}^\alpha - Y^\alpha\right)$ $(y_{tar}^\alpha ln y_{tar} - Y^\alpha ln Y)$
3) **repeat**
4) $\alpha_{n+1} \leftarrow \alpha_n - \delta \frac{\partial c}{\partial \alpha}$
5) $\alpha_n \leftarrow \alpha_{n+1}$
6) **until** $(\alpha_{n+1} > 0)$
7) **end procedure**

Using the above algorithm, the optimal values of α and cost have been computed(cf. Result & Discussion).

3.3 Profit Maximization

Considering an enterprise which wants to maximize the profit

$$Revenue = (S^\rho + I^\rho + P^\rho + N^\rho)^{\frac{1}{\rho}} \tag{13}$$

$$Cost = w_1 S + w_2 I + w_3 P + w_4 N \tag{14}$$

$$Profit = (S^\rho + I^\rho + P^\rho + N^\rho)^{\frac{1}{\rho}} - (w_1 S + w_2 I + w_3 P + w_4 N) \tag{15}$$

Say, the cost should not surpass the threshold c_{thresh} So we need to maximize $(S^\alpha + I^\alpha + P^\alpha + N^\alpha) - (w_1 S + w_2 I + w_3 P + w_4 N)$ subject to $w_1 S + w_2 I + w_3 P + w_4 N = c_{thresh}$. The following values of S, I, P and N thus obtained are the values for which the data center can attain maximum profit by staying within constraints.

$$S = \frac{c_{thresh} w_1^{\frac{1}{\rho-1}}}{w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} + w_3^{\frac{\rho}{\rho-1}} + w_4^{\frac{\rho}{\rho-1}}} \tag{16}$$

$$I = \frac{c_{thresh} w_2^{\frac{1}{\rho-1}}}{w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} + w_3^{\frac{\rho}{\rho-1}} + w_4^{\frac{\rho}{\rho-1}}} \tag{17}$$

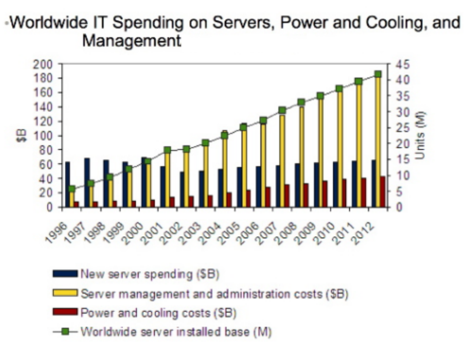
$$P = \frac{c_{thresh} w_3^{\frac{1}{\rho-1}}}{w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} + w_3^{\frac{\rho}{\rho-1}} + w_4^{\frac{\rho}{\rho-1}}} \tag{18}$$

$$N = \frac{c_{thresh} w_4^{\frac{1}{\rho-1}}}{w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} + w_3^{\frac{\rho}{\rho-1}} + w_4^{\frac{\rho}{\rho-1}}} \tag{19}$$

The above results are elaborated in Appendix 3.

4 Result and Discussion

As mentioned earlier, server and power/cooling costs form the biggest chunk of the total cost. These two inputs are considered for computing the values of the elasticities using 3D plots. However, the results obtained hold good for any number of inputs.It is also possible to aggregate the inputs into two broad categories,operational expenditure and capital expenditure and use these as the two inputs in the proposed cost model. Operational expenditures include the recurring costs like power/cooling, server management etc; whereas capital expenditure includes initial investment such as new server cost, infrastructure costs etc. The data associated with data center costs from various sources have been accumulated and CES function is applied to find the optimal solution for revenue of data center, and finally revenue maximization is demonstrated graphically. All simulation results have been generated by a computer system using Matlab.



The approximate data from the figure1 for two types of costs, namely server management/administrative cost and power/cooling cost are captured. The optimal elasticity of each input and maximum revenue for each year using Matlab code are obtained. The experiment has been conducted for three scenario.

1. In case $\rho < 1$
2. In case $\rho = 1$
3. In case $\rho > 1$

4.1 Case 1: $\rho < 1$

Applying the constraints $\rho < 1$
$\rho > 0$
to the function
$f = (x^\rho + y^\rho)^{\frac{1}{\rho}}$
Using fmincon function of matlab[Appendix 7], the values of elasticities for which revenue is maximized for each year are obtained.

TABLE 2
Simulation output

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	62	5	1.11	0.1000	19511.9
1997	65	5	1.11	0.1000	20038.22
1998	62	10	1.11	0.1000	26579.25
1999	60	10	1.11	0.1000	26108.41
2000	65	10	1.11	0.1000	27274.01
2001	55	15	1.11	0.1000	30038.95
2002	45	15	1.11	0.1000	27008.5
2003	47	15	1.11	0.1000	27635.78
2004	50	20	1.11	0.1000	32723.23
2005	52	20	1.11	0.1000	33401.89
2006	55	20	1.11	0.1000	34399.27
2007	56	30	1.11	0.1000	42176.36
2008	57	30	1.11	0.1000	42563.18
2009	58	30	1.11	0.1000	42947.06
2010	59	40	1.11	0.1000	49839.75
2011	60	40	1.11	0.1000	50268.74
2012	60	40	1.11	0.1000	50268.74

In the above table all units are in $B. The optimal revenue of all the years is obtained at ρ= 0.100 and s=1.11. 2D graphs have been generated using these results.

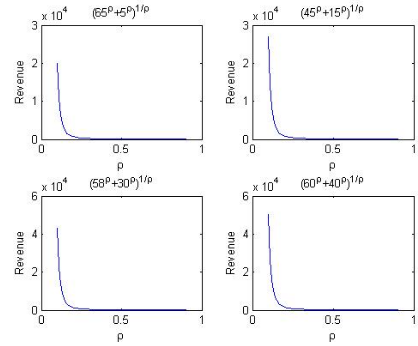


Fig. 1. Graph representing the Revenue against years when $\rho < 1$

In the above graph X axis represent ρ and Y axis represents Revenue. It is quit prominent from the graph that the maximum revenue is achieved where the value of ρ is 0.100. The graphs are displaying the same optimal values represented in the table 1. All the graphs are depicting a common pattern, where the revenue falls sharply after the $\rho > 0.100$ region. Another pattern, which is observable through out the 4 graphs is that there is no major fluctuation in revenue after $\rho > 0.100$ till 0.9.

4.2 Case 2: $\rho = 1$

In this section, we will discuss the effect of CES function on Data center revenue generation when $\rho = 1$. As there is no variations or range of ρ, which is visible in other cases, therefore the application of fmincon function is insignificant. We will investigate whether the revenue optimization can be attained and the behavior of revenue when $\rho = 1$.The scenario exhibits linear production function,where response revenue output becomes the sum of input factors. The observations after applying on worldwide IT spending data set have been shown in table 2. If we compare the revenues with the case $\rho > 1$, it is noticeable that the revenues are slightly higher than the later case. In 1996, the difference of revenue between these two cases is 1.5, whereas the difference rises to almost 6 in the year of 2012.

4.3 Case 3: $\rho > 1$

Applying the constraints $\rho < 2$
$\rho > 1$
to the function
$f = (x^\rho + y^\rho)^{\frac{1}{\rho}}$
Using fmincon function of matlab[Appendix 7], the values of elasticities for which revenue is maximized for each year are obtained.

TABLE 3
Simulation output

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	62	5	0	1	67
1997	65	5	0	1	70
1998	62	10	0	1	72
1999	60	10	0	1	70
2000	65	10	0	1	75
2001	55	15	0	1	70
2002	45	15	0	1	60
2003	47	15	0	1	62
2004	50	20	0	1	70
2005	52	20	0	1	72
2006	55	20	0	1	75
2007	56	30	0	1	86
2008	57	30	0	1	87
2009	58	30	0	1	88
2010	59	40	0	1	99
2011	60	40	0	1	100
2012	60	40	0	1	100

TABLE 4
Simulation output

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	62	5	-10	1.1000	65.5239
1997	65	5	-10	1.1000	68.5078
1998	62	10	-10	1.1000	69.5306
1999	60	10	-10	1.1000	67.5538
2000	65	10	-10	1.1000	72.4972
2001	55	15	-10	1.1000	66.8546
2002	45	15	-10	1.1000	57.0674
2003	47	15	-10	1.1000	59.0214
2004	50	20	-10	1.1000	66.3455
2005	52	20	-10	1.1000	68.2907
2006	55	20	-10	1.1000	71.2122
2007	56	30	-10	1.1000	81.1220
2008	57	30	-10	1.1000	82.0859
2009	58	30	-10	1.1000	83.0503
2010	59	40	-10	1.1000	93.1262
2011	60	40	-10	1.1000	94.0819
2012	60	40	-10	1.1000	94.0819

In Table 2, all units are in \$B. The optimal revenue for all years are obtained at ρ = 1.1 and s= -10 Using these results, 2D-simulations are created and the graphs are obtained.

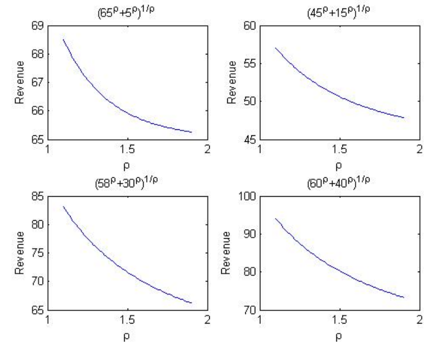


Fig. 2. Graph representing the Revenue against years, when $\rho > 1$

In the above figures X axis represents ρ and Y axis represents revenue. The maximum revenue is lying in the region, where the ρ is close to 1.1. Apart from first figure there is no immediate sudden fall of revenue after ρ =1.1.The graphs depicting in the above figure are predominantly concave decreasing.

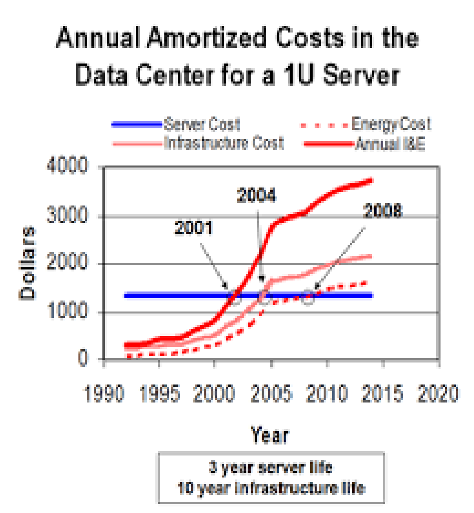


Fig. 3. Graph representing the annual amortized costs for a 1U size server

Figure3 is the graphical representation of Annual Amortization Costs in data center for 1U server. The graph has been sourced from ”In the data center, power and cooling costs more than the it equipment it supports” document [22]. All units are in \$. We have taken fairly accurate data from the graph and represented in a tabular format below on a yearly basis. Maximum revenue and optimal elasticity constants are demonstrated in the same table.

4.4 Case 1: $\rho < 1$

Applying the constraints $\rho < 1$
$\rho > 0$
to the function
$f = (x^{\rho} + y^{\rho})^{\frac{1}{\rho}}$

Fmincon function of matlab[Appendix 7] has been used to get optimal elasticity at which maximum revenue is attained. In the above table optimal revenues have been

TABLE 5
Simulation output

Year	Server Cost	Energy cost	Infrastructure	Annual I and E	Elasticity($\frac{\rho}{s}$)		Max. Revenue
1992	1400	50	200	220	1.1	0.1	265159191.08
1995	1400	75	250	300	1.1	0.1	329860151.91
2000	1400	200	500	1200	1.1	0.1	691685894.7
2005	1400	1000	1000	2800	1.1	0.1	1644247792.93
2010	1400	1600	1500	3400	1.1	0.1	2029519563.1

achieved at ρ=0.1 and s= 1.1. 2D graphs are generated using the same data set and displayed below.

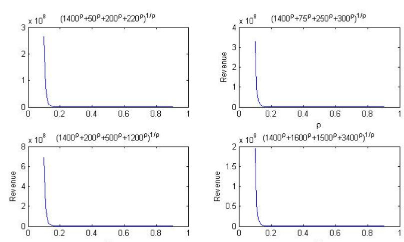


Fig. 4. Graph representing the Revenue against years, when $\rho < 1$

Cost data of the years 1992, 1995, 2000 and 2005 have been used in graph generation. The graphs are not displaying any concavity or convexity feature. The common pattern, which we already been observed in the previous dataset that the revenue falls sharply after ρ. Maximum revenue is attained in the the region which is closer to ρ= 0.1.

4.5 Case 2: $\rho = 1$

In this scenario, the CES function behaves as a linear production function, where the out put revenue is the sum of all the input factors. Single ρ value has been applied over the data set and revenues have been calculated.The revenues generated in this case are lesser than the previous case but revenues are higher than the case, where $\rho > 1$. The revenue rises almost 4 times from the year of 1992 to 2010. The rise can be explained by the surge of input cost factors.

4.6 Case 3: $\rho > 1$

Applying the constraints $\rho < 2$
$\rho > 1$
to the function
$f = (x^\rho + y^\rho)^{\frac{1}{\rho}}$
Fmincon function of matlab[Appendix 7] has been used to obtain optimal elasticity. In the table, it is noticeable that the revenue rises to 4 folds from 1992 to 2010. The maximum

TABLE 6
Simulation output

Year	Server Cost	Energy cost	Infrastructure	Annual I and E	Elasticity($\frac{\rho}{s}$)		Max. Revenue
1992	1400	50	200	220	0	1	1870
1995	1400	75	250	300	0	1	2025
2000	1400	200	500	1200	0	1	3300
2005	1400	1000	1000	2800	0	1	6200
2010	1400	1600	1500	3400	0	1	7900

TABLE 7
Simulation output

Year	Server Cost	Energy cost	Infrastructure	Annual I and E	Elasticity($\frac{\rho}{s}$)		Max. Revenue
1992	1400	50	200	220	-10	1.1	1619.02
1995	1400	75	250	300	-10	1.1	1733.20
2000	1400	200	500	1200	-10	1.1	2738.19
2005	1400	1000	1000	2800	-10	1.1	5066.52
2010	1400	1600	1500	3400	-10	1.1	6424.43

revenue is attained at ρ=1.1. Using the above dataset 2D simulation graphs are produced.
Using the cost data of 1992, 1995, 2000, 2005, the graphs are produced. The common feature of the above graphs is that all are concave decreasing and the maximum revenue attained at ρ=1.1. These features are observable across the datasets when ρ¿1. Gradient descent method has been applied on the same world wide IT spending dataset to find out optimal elasticities for cost minimization. The initial value of the ρ has been assumed as 1.2 whereas step size for each iteration has been considered as 0.001 (Table 5). In the gradient descent calculation, we assumed the target revenue as $240B and unit cost of new server installation as 0.6.

Now we are going to explore the behavior of profits, using our proposed model. We will consider two instances for better understanding the behavior of profit function.The worldwide IT spending dataset has been taken for profit analysis. The data set consists of two cost component New Server installation cost and Power & Cooling cost. In cost function, the weight of each cost segment need to incorporated. Weights of New Server Installation and Power & Cooling are decided as 0.3 and 0.4
1) $\rho < 1$
2) $\rho > 1$

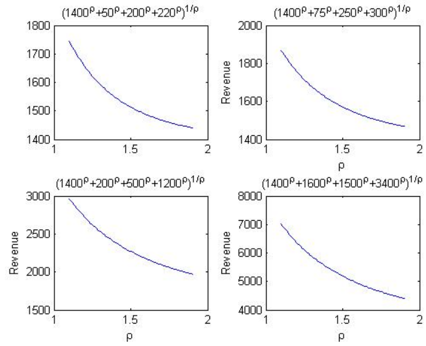


Fig. 5. Graph representing the Revenue against years, when $\rho < 1$

TABLE 8
Gradient Descent output for cost minimization

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Min. Cost
1996	62	5	1.007	0.0078	49.6
1997	65	5	1.21	0.1764	52.32
1998	62	10	1.007	0.0078	49.6
1999	60	10	1.03	0.0305	48
2000	65	10	1.21	0.1764	52.32
2001	55	15	1.09	0.0835	44
2002	45	15	.99	0.0076	36
2003	47	15	1.19	0.1600	37.61
2004	50	20	1.15	0.1323	40
2005	52	20	1.12	0.1132	41.6
2006	55	20	1.09	0.0835	44
2007	56	30	1.079	0.0733	44.8
2008	57	30	1.06	0.0629	45.6
2009	58	30	1.055	0.0523	46.4
2010	59	40	1.04	0.0415	47.2
2011	60	40	1.03	0.0305	48
2012	60	40	1.03	0.0305	48

4.7 Case 1: $\rho < 1$

Applying the constraints
$\rho < 1$
$\rho > 0$
to the function
$f = (x^\rho + y^\rho)^{\frac{1}{\rho}}$
Fmincon function of Matlab is used to find the optimal ρ value , at which maximum profit can be obtained.In table 9, New server cost and Power & Cooling cost of world wide it spending data set along with optimal profit have been displayed. It is observable from the Table 9 that like maximum revenue, maximum profit also obtained where ρ is 0.1. The same data set is used to generate 2D graphs.Cost data of the years 1999, 2002, 2009 and 2012 have been used in graph generation. X, Y axis represent Profit, ρ respectively.The graphs don not exhibit any concavity or convexity feature. The common pattern, which we already been observed in revenue graphs in the same instance that like revenue, profit falls sharply after ρ 0.1 . Maximum profit is attained in the the region which is closer toρ = 0.1. For giving more insight, Profit against the years graphs are demonstrated along with 'Profit Vs ρ' graphs. The profits from year 1999 to 2012 has been depicted in Figure 6. Profit from year 1999 to 2001 rises and it has decreased in the year 2001, a sudden dip in the graph is visible as both. After the year 2001 the profits rise

TABLE 9
Simulation output for profit maximization

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	62	5	1.1	0.1	19486
1997	65	5	1.1	0.1	20011
1998	62	10	1.1	0.1	26551
1999	60	10	1.1	0.1	26081
2000	65	10	1.1	0.1	27245
2001	55	15	1.1	0.1	30012
2002	45	15	1.1	0.1	26986
2003	47	15	1.1	0.1	27612
2004	50	20	1.1	0.1	32697
2005	52	20	1.1	0.1	33375
2006	55	20	1.1	0.1	34371
2007	56	30	1.1	0.1	42145
2008	57	30	1.1	0.1	42531
2009	58	30	1.1	0.1	42915
2010	59	40	1.1	0.1	49804
2011	60	40	1.1	0.1	50233
2012	60	40	1.1	0.1	50233

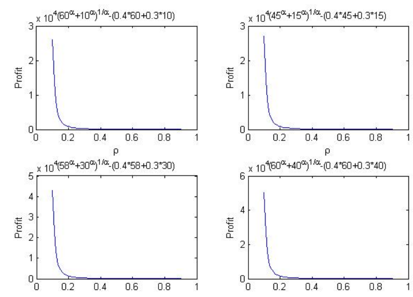


Fig. 6. Graph representing the profit of four different year against ρ, when $\rho < 1$

sharply till 2012 and the highest jump of profit is noticeable between 2006 to 2008. In this figure the X axis represents Years and Y axis represents Profit.

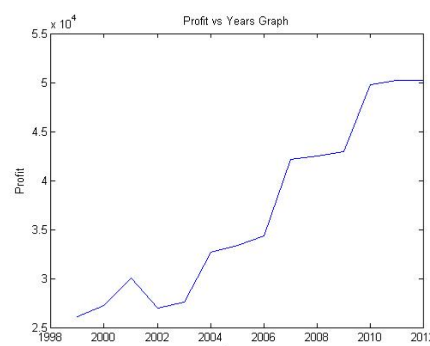


Fig. 7. Graph representing the profit against years when $\rho < 1$

4.8 Case 2: $\rho > 1$

Applying the constraints
$\rho > 1$
$\rho < 2$
to the function
$f = (x^\rho + y^\rho)^{\frac{1}{\rho}}$

Like previous case, Fmincon function of Matlab is utilized to get optimal ρ. In Table 10, the result after applying Fmincon function is represented which include maximum profit, optimal ρ and two cost segments. Using the cost

TABLE 10
Simulation output for profit maximization

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	62	5	-10	1.1	39.2239
1997	65	5	-10	1.1	41.0078
1998	62	10	-10	1.1	41.7306
1999	60	10	-10	1.1	40.5538
2000	65	10	-10	1.1	43.4972
2001	55	15	-10	1.1	40.3546
2002	45	15	-10	1.1	34.5674
2003	47	15	-10	1.1	35.7214
2004	50	20	-10	1.1	40.3455
2005	52	20	-10	1.1	41.4907
2006	55	20	-10	1.1	43.2122
2007	56	30	-10	1.1	49.7220
2008	57	30	-10	1.1	50.2859
2009	58	30	-10	1.1	50.8503
2010	59	40	-10	1.1	57.5262
2011	60	40	-10	1.1	58.0819
2012	60	40	-10	1.1	58.0819

data of 1999, 2002, 2009, 2012, the graphs are produced. The common feature of the above graphs is that all are concave decreasing and the maximum revenue attained at =1.1. X axis represents Profit, whereas Y axis represents ρ. The profits from year 1999 to 2012 has been depicted in

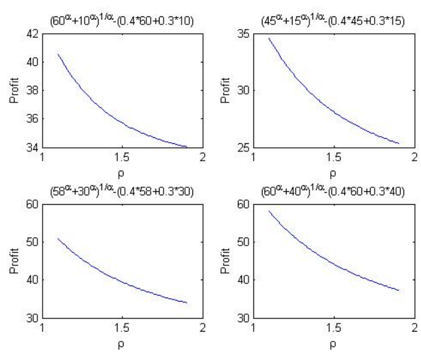


Fig. 8. Graph representing the profit of four different year against ρ, when $\rho > 1$

Figure 9. Profit from year 2000 declines significantly to reach the lowest profit in the year of 2002.The profit rises surges after the year 2002 without any sudden dip till it touches the highest profit in 2012. In this figure the X axis represents Years and Y axis represents Profit.

5 Concentration of firms in Data Center Field

As we have seen in previous section, with the increase of cost, profit is rising in all the cases. So it is require to investigate the level of concentration of firms in the field of Data center. For setting up a data center, a firm needs huge amount of initial investment hence big organizations are capable enough to have their own data centers. Big

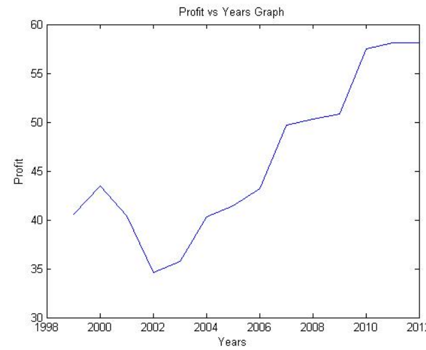


Fig. 9. Graph representing the profit against years, when $\rho > 1$

firms such Amazone, Google, Microsoft have constructed massive computing infrastructure to support their websites and services. These organizations have started to rent their infrastructure to developers and small firms, who dont have their own data center. Now we would like to know the stiff competition, these giants are facing. We will employ Hirschman-Herfindahl Index (HHI) to know the competition or concentration of firms.

The Herfindahl-Hirschman index (HHI) is widely used technique to measure the market concentration and it is calculated by squaring the market share of each firm competing in a market, and then summing the resulting numbers. The HHI number can vary from close to zero to 10,000. The HHI is expressed as:

$$HHI = s1^2 + s2^2 + s3^2 + ... + sn^2 (wheresn \\ isthemarketshareoftheithfirm). \tag{20}$$

If HHI is high, meaning a few firms control the business, then new cost outlay raises revenue and that raises profit. If many organization are playing in the business and there is tough competition, the HHI value will be less. First we will discuss about the market competition in asia pecific region. The region has generated just over USD 20 billion in data center infrastructure revenues for the worlds leading technology vendors and the market having grown by 23% from the previous year, according to data from Synergy Research Group [24].

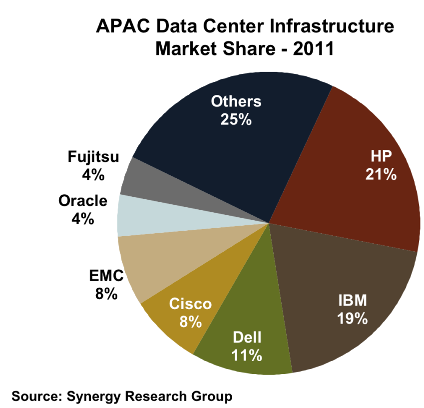


Fig. 10. Asia pacific region Data Center Market share in 2011

$$HHI = 21^2 + 19^2 + 11^2 + 8^2 + 8^2 + 4^2 + 4^2 + 25^2 = 1708$$

The U.S. Department of Justice considers a market with a result of less than 1,000 to be a competitive marketplace; a result of 1,000-1,800 to be a moderately concentrated marketplace; and a result of 1,800 or greater to be a highly concentrated marketplace [23]. In apac region the concentration is moderate and inclined towards highly concentrated market place. Next we will try to find out the firm concentration in Iaas market place. The Iaas market share data has been collected from a article of businessinsider website [25].
The HHI for Iaas market share is below

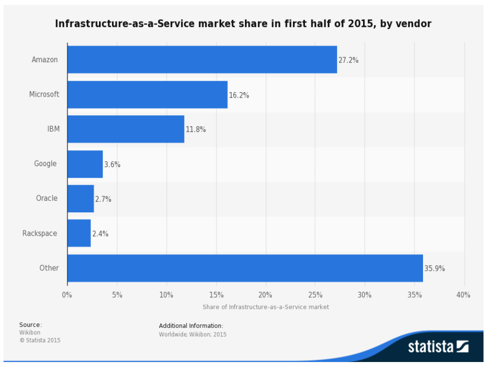


Fig. 11. Infrastructure-as-a-Service Market share in 2015 first half

$$HHI = 27.2^2 + 16.6^2 + 11.8^2 + 3.6^2 + 2.7^2 + 2.4^2 + 35.9^2 = 2456.34$$

If we exclude others from HHI calculation, it becomes 1167.53. Still the market can not be considered as competitive. Few firms are controlling the major share of the Infrastructure as a service market place.

6 Factor Analysis in CES Model

CES production can accommodate any number of factors as it can be expanded to n number of input variables. In our discussion, we have taken so far either 4 or 2 factors to calculate optimal revenue. In this section, we are going to discuss how these factors and their interactions are contributing towards revenue output. We may find out after factor analysis that few factors are insignificant and few factors are yielding more output in comparison to others. Factor analysis can uncover relationship patterns underlying hundreds of interacting phenomenon. Apart form individual factors, all possible interactions among the factors are also considered in factor analysis. So factor analysis will be helpful to understand how factors(input variables) are contributing to revenue output in CES function. To ease the calculation, we have considered 2 cost segments New Server Cost, Power & Cooling Cost as factors.We need to identify the high and low point of each factor.In result & discussion section we have considered two datasets, consisting of 2 and 4 factors but here the world wide IT spending data set is used to produce regression equation and relevant allocated variances will be demonstrated.

6.1 About Factor Analysis

Factor analysis is an important useful statistical tool to investigate hidden relationship for complex concepts such as socioeconomic status, dietary patterns, or psychological scales. The key concept behind factor analysis is that similar patterns are being followed by multiple observed variables as they are all associated with a latent (i.e. not directly measured) variable. In every factor analysis, there are the same number of factors as there are variables. Each factor is responsible for a certain amount of variance in the measured output and the factors generally listed in order how much they can detect the variation. Basically factor analysis takes thousands and potentially millions of measurements and their corresponding qualitative observations and resolves them into distinct patterns of occurrence [11].Factor analysis can be applied in a vast range of interest fields such as content area, structure a domain, map unknown concepts, classify or reduce data, illuminate causal nexuses, screen or transform data, define relationships, test hypotheses, formulate theories, control variables etc.

6.2 Factor Analysis for Proposed Model

The objective of the factorial analysis is to identify the significant factors , different cost associated with data center for which maximum revenue can be achieved. The contribution of each factor helps to determine how factors are impacting the revenue. This factor analysis identifies the percentage of contribution of each factor. These details can be used to understand and decide how the factor can be controlled to generate better revenue.

6.2.1 The 3^2 Design

The factors, which have been considered for 3 levels factorials design are listed below:

- New Server installation cost
- Power and Cooling

Instead of pointing a single value , we have considered a range for high, medium and low level for each factor.

TABLE 11
Factor 1 level

Level	Range
Low	5-15
Medium	16-29
High	30-40

TABLE 12
Factor 3 Level

Level	Range
Low	45-52
Medium	53-59
High	60-65

In both the tables, all units are in$B. Let us define two variables x_A and x_B as representation of New Server cost and Power & Cooling cost. The mapping of high, medium and low levels of each factors are demonstrated in below

TABLE 13
Factor Initialization

Factor	High	Medium	Low
x_A	2	1	0
x_B	2	1	0

table.

The Revenue y can now be formulated on x_A and x_B using a nonlinear regression model of the form:

$$y = q_0 + q_1 x_A + q_2 x_B + q_{12} x_{AB} + q_{11} x_A^2 + q_{22} x_B^2 \quad (21)$$

The effect of the factors are being calculated by the proportion of the total variation in the output, contributed by the factor. Sum of contribution (SC) can be represented by below equation:

$$SC = q_1 + q_2 + q_{11} + q_{12} + q_{22} \quad (22)$$

Where: $q_1 + q_{11}$ is the portion of SC that is contributed by New Server cost. $q_2 + q_{22}$ is the portion of SC that is contributed by Power & Cooling cost. q_{12} is the portion of SC that is contributed by the interactions of New Server cost and Power & Cooling cost.

$$SC = SCA + SCB + SCB \quad (23)$$

Fraction of variation contributed by $A = SCA/SC$
Fraction of variation contributed by $B = SCB/SC$
Fraction of variation contributed by $AB = SCAB/SC$
The objective behind applying this methodology to identify the major cost factors to accommodate in the proposed CES model.

6.3 Experimental Observations

6.3.1 Experiment 1 : With two factors$\rho < 1$

	Power and Cooling		
Server	Low	medium	High
Low	287322.14	33062.56	
Medium	30038.95	34399.27	44381.5875
High	23902.358		50268.74

TABLE 14
Experiment 1: With two factors$\rho < 1$

After solving all the equations, the regression equation is:

$$\begin{gathered} y = 27322.14 + 7143.511 x_A - 4462.551 x_B \\ -1380.1 x_A x_B - 4426.701 x_A^2 + 10202.971 x_B^2 \end{gathered} \quad (24)$$

Substituting the parameter in equation (14)
$SC = SCA + SCB + SCB$
$=7143.51 + 4462.551 + 1380.1 + 4426.701 + 10202.971$
$=27615.834$
After analysis the result, we can interpret the effect of each factor as follows: The contribution of New Server cost on revenue is 41.89%. The contribution of Power & Cooling cost on revenue is 53.10%. And the contribution by the interaction of New Server cost and Power & Cooling is 4.99%.

	Power and Cooling		
Server	Low	Medium	High
Low	61	71	
Medium	70	75	90
High	70.8		100

TABLE 15
Experiment 2: With two factors$\rho = 1$

6.3.2 Experiment 2 : With two factors$\rho = 1$

After solving all the equations, the regression equation is:

$$y = 61 + 13.1 x_A - 4.6 x_B - 5 x_A x_B - 4.6 x_A^2 + 14.6 x_B^2 \quad (25)$$

Substituting the parameter in equation (14)
$SC = q_1 + q_2 + q_{11} + q_{12} + q_{22}$
$=4.1 + 13.1 + 5 + 14.6 + 4.6$
$=41.4$

After analysis the result, we can interpret the effect of each factor as follows: The contribution of New Server cost on revenue is 41.54%. The contribution of Power & Cooling cost on revenue is 46.37%. And the contribution by the interaction of New Server cost and Power & Cooling cost is 12.07%.

6.3.3 Experiment 3 : With two factors$\rho > 1$

	Power and Cooling		
Server	Low	Medium	High
Low	58.0444	67.3181	
Medium	66.8546	71.2122	84.8461
High	68.72266		94.0819

TABLE 16
Experiment 3: With two factors$\rho > 1$

After solving all the equations, the regression equation is:

$$\begin{aligned} y = 58.0444 + 12.2814 x_A + 19.17105 x_B - 4.9161 x_A x_B \\ -3.4712 x_A^2 - 9.89735 x_B^2 \end{aligned} \quad (26)$$

Substituting the parameter in equation (14)
$SC = q_1 + q_2 + q_{11} + q_{12} + q_{22}$
$=12.2814 + 19.17105 + 4.9161 + 3.4712 + 9.89735$
$=49.7371$
After analysis the result, we can interpret the effect of each factor as follows: The contribution of New Server cost on revenue is 31.67%. The contribution of Power & Cooling cost on revenue is 58.44%. And the contribution by the interaction of New Server cost and Power & Cooling cost is 9.8841%.

7 Randomization of Data

The data collected from various sources are not sufficient enough to identify the effect of factors on the revenue. So we have to find the probability distribution of the original data set and random data set need to generate which will follow the same distribution of real data set. We have found through experiment that original data set follows normal distribution. Figure 12 depicts the normal distribution of server cost of original and random data. Figure 13 displays the normal distribution for power & cooling cost. The maximum revenue for random data has been calculated and presented in table 15 and 16. Shapiro-Wilk Original Test and Chi Square- Goodness have been conducted on the original and actual data set to identify the normal distribution behavior. The Null Hypotheses h_0: After adding noise to the original data set, the data follows normal Distribution. If $h_0 = 1$, the null hypothesis is rejected at 5% significance level. if h0=0, the null hypothesis is accepted at 5% significance level. After the experiment, we found that data set follows a normal distribution with 95% confidence level i.e. $h_0 = 0$. The details of Shapiro-Wilk Original Test and Chi Square-Goodness have been elaborated in Appendix 6.

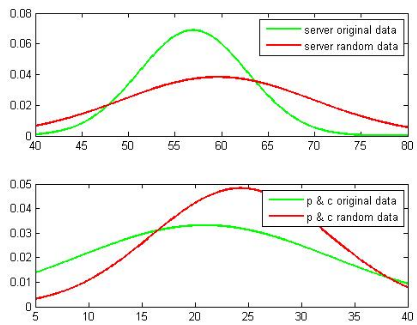


Fig. 12. The Original and Generated Server,Power & Cooling Data that follows Normal Distribution

TABLE 17
Simulation output for Random data, when $\rho < 1$

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	78	15	1.11	0.1000	36235.29
1997	60	35	1.11	0.1000	47096.27
1998	55	18	1.11	0.1000	32725.54
1999	44	20	1.11	0.1000	30613.65
2000	75	30	1.11	0.1000	49084.85
2001	62	15	1.11	0.1000	32023.18
2002	55	25	1.11	0.1000	38267.07
2003	49	18	1.11	0.1000	30794.72
2004	55	25	1.11	0.1000	38267.07
2005	57	24	1.11	0.1000	38229.9
2006	65	35	1.11	0.1000	49076.16
2007	46	38	1.11	0.1000	42832.04
2008	77	27	1.11	0.1000	47335.4
2009	68	32	1.11	0.1000	48107.5
2010	59	16	1.11	0.1000	32138.37
2011	48	10	1.11	0.1000	23134.73
2012	60	30	1.11	0.1000	43706.2

8 NON-PARAMETRIC ESTIMATION

Non-parametric statistic is not belong to the family of probability distributions.It can be both descriptive and statistical. No assumption about the probability distribution of the sample data, has been made in non-parametric estimation. The typical parameters mean, variance are accessed in the estimation. The typical parameters are the mean, variance, etc. Unlike parametric statistics, non-parametric statistics make no assumptions about the probability distributions of the variables being assessed.In parametric, there is no specific distinction between true models and fitted models. In contrast, non-parametric methods able to distinguish between the true and fitted models. The drawback of non-parametric tests are less powerful in comparison to parametric test. We have performed non-parametric estimation on the data set where we have ignored whether the data set part of probability distribution. Figure 13 and 14 showcase the result of non-parametric estimation on the on the original dataset. Both the cost segments, Server and power & cooling have been displayed in the figure. Figure 13 suggests interaction between the two cost component to generate revenue. Figure 14 reveals no interaction between the factors. The non-parametric estimation on generated data has been shown in figure 16 and 17. None off the figure demonstrate any interaction between factors.

TABLE 18
Simulation output for Random data, when $\rho > 1$

Year	New Server	Power and Cooling	Elasticity(s)	ρ	Max. Revenue
1996	78	15	-10	1.1000	85.55
1997	60	35	-10	1.1000	85.57
1998	55	18	-10	1.1000	66.57
1999	44	20	-10	1.1000	58.08
2000	75	30	-10	1.1000	95.04
2001	62	15	-10	1.1000	70.62
2002	55	25	-10	1.1000	72.4
2003	49	18	-10	1.1000	61.01
2004	55	25	-10	1.1000	72.4
2005	57	24	-10	1.1000	73.4
2006	65	35	-10	1.1000	90.12
2007	46	38	-10	1.1000	75.5
2008	77	27	-10	1.1000	94.38
2009	68	32	-10	1.1000	90.3
2010	59	16	-10	1.1000	68.6
2011	48	10	-10	1.1000	53.5
2012	60	30	-10	1.1000	81.2

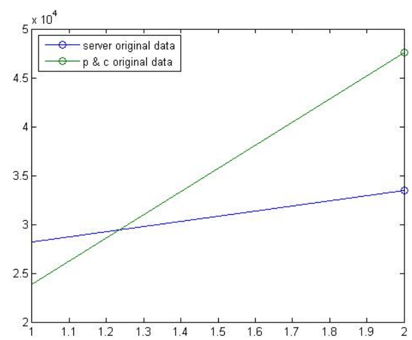


Fig. 13. Non parametric estimation on orignial data, when $\rho < 1$

9 EXPERIMENTS

Considering again the CES production function

$$y = (A^{\alpha} + B^{\alpha})^{\frac{1}{\alpha}}$$

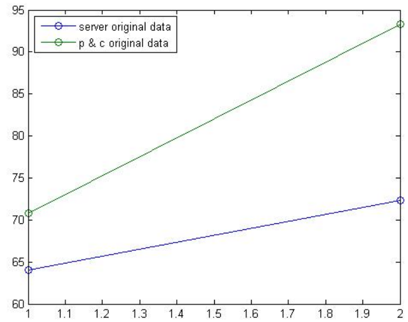


Fig. 14. Non parametric estimation on orignial data when $\rho = 1$

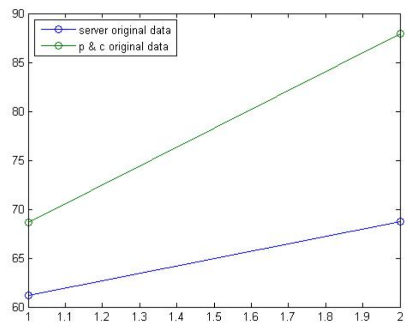


Fig. 15. Non parametric estimation on orignial data when $\rho > 1$

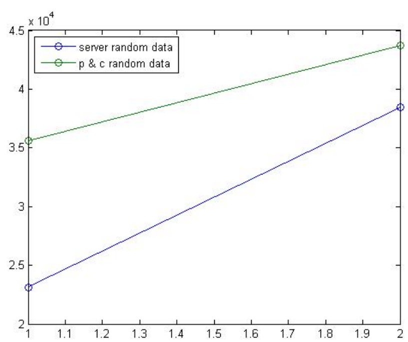


Fig. 16. Non parametric estimation on random data when $\rho < 1$

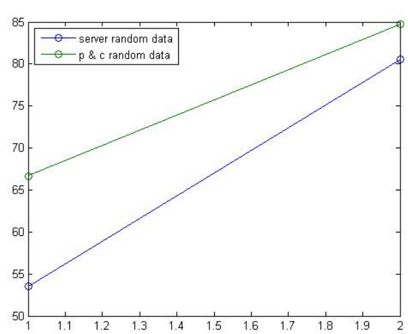


Fig. 17. Non parametric estimation on random data when $\rho > 1$

If we have N number of data points. The above equation can be rewritten as below.

$$y_1 = (A_1^\alpha + B_1^\alpha)^{\frac{1}{\alpha}}$$

$$y_N = (A_N^\alpha + B_N^\alpha)^{\frac{1}{\alpha}}$$

As the number of data point is greater than the number of unknown, it is an over-determined system. Two instances of least square approaches, we will discuss.

9.1 No Constraints

As the CES function is non-linear, we will explore non-linear least square method.Our objective is to fit a set of observation with our proposed model, which is non-linear. We assumed that there is no restriction or constraints on the parameters.The basis of the method is to approximate the model by a linear one and to refine the parameters by successive iterations. Considering m data points $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, and the model function as $y = f(x, \beta)$. The variable x is dependent on n parameters,$(\beta_1, \beta_2, \ldots \beta_n)$.So the expectation is to find the vector ofβ so that the curve best fit in the observed data set, means the sum of squares

$$S = \sum_{i=0}^{n} r_i^2$$

is minimized, where the residuals (errors) r_i are given by

$$r_i = y_i - f(x_i, \beta)$$
$$i = 1, 2 \ldots m$$

The minimum value will be attained where gradient is 0, that means there will be n gradient equations

$$\frac{\partial S}{\partial \beta_j} = 2 \sum_i r_i \frac{\partial r}{\partial \beta_j} (j = 1 \ldots n)$$

The gradient equations are the combinations of the independent variables and parameters in non-linear system. So these equations don't have any closed solution.Hence the values are obtained by successive approximation during each iteration.

$$\beta_j \approx \beta_j^{k+1} = \beta_j^k + \Delta\beta_j$$

K is the number of iteration and $\Delta\beta$ is the increment. Talyor series has been used to linearize the model at each iteration. We have used lsqnonlin function of matlab, which deals with non-linear least square approach. The details of this approach is available in Appendix 5. The results after applying non-linear least square has been shown in table 18.

	$\rho < 1$	$\rho > 1$
ρ	0.100	1.100

TABLE 19
Least Square Result without constraints

With constraints we have utilized the same matlab lsqnonlin function. The details has been elaborated in Appendix 5. Table 19 displays the result, we have obtained after applying least square method with constraints.

	$\rho < 1$	$\rho > 1$
ρ	0.100	1.100

TABLE 20
Least Square Result with constraints

10 REPLICATION

A replication experiment is performed to estimate the imprecision or random error of the analytical method. Methods of measurements are almost always subject to some random variation. Repeat measurements will usually reveal slightly different results, sometimes a little higher, sometimes a little lower. Determining the amount of random error is usually one of the first steps in a method validation study[16].To measure the variability associated with an experiment, repetition performs in various fields such as engineering, science and statistics.As we are going to elaborate 2^2 factorial design, atleast two replications need to be considered. The result of replication experiment has been displayed in Table 20. The difference with the previous 2^2 experimental design in terms of variations is visible. The interaction factor, which was the second most significant in case of $\rho < 1$, has reduced significantly and reaches to 2.25%. Still new server is the most major contributor and power and cooling contributing 9.27 %, which we can ignore. In case of $\rho > 1$, the power and cooling contribution has decreased towards revenue generation and interaction factor still is the least significant factor. Both the cases, it is observable that the experiment error is quite high.

In statistics, we perform a variety of intervals to understand the behavior of the produced result.Out of these, confidence interval is widely used. A confidence interval defines a range values, which has been derived from a sample dataset. The range may contain the value of unknown parameter. Due to the random characteristic, i is unlikely that two sample collected from a population yield an identical confidence interval.But by repeating sample many times, we may find a confidence interval, which holds the unknown parameter. Here we are going to apply the same concept to the each of the coefficient(q_A, q_B, q_AB). We are 90% confident that each coefficient value will fall in the ranges, shown in table 21.

	$\rho < 1$	$\rho > 1$
New Server	63.1	47.74
Power and Cooling	9.27	27.417
Interaction	2.25	4.10
Error	25.2	20.73

TABLE 21
Percentage variations

	$\rho < 1$ CI	$\rho > 1$ CI
q_0	33553.785-37008.065	68.204-74.516
q_A	4200.75-7649.03	8.1265-14.4385
q_B	2760.63-6214.91	1.171-7.4835
q_AB	9.395-9.395	-0.8985- 5.4135

TABLE 22
Percentage variations

11 PREDICTION AND FORECASTING

The linear regression models are restricted in three ways.First, only one predictor variable need to consider. Second the predictor variable should be quantitative and third response must be linear function of predictor.Multiple linear regression is a technique, where more than one predictor variables can be considered [17]. A multiple linear regression helps one to predict a response variable y as a function of k predictor variables $x_1, x_2, x_3 \cdots x_k$ using a linear model of the following form.

$$y = b_0 + b_1x_1 + b_2x_2 + \cdot + b_kx_k + e$$

Here,$b_0, b_1, \cdot, b_k$ are the $k + 1$ fixed parameters and e is the error terms. Given a sample dataset $(x_11, x_21, \cdot, x_k1, y_1), \cdot, (x_1n, x_2n, \cdot, x_kn, y_n)$ of n observations, the model consists of the below n equations:

$$\begin{aligned} y_1 &= b_0 + b_1x_11 + b_2x_21 + \cdot + b_kx_k1 + e_1 \\ y_2 &= b_0 + b_1x_12 + b_2x_22 + \cdot + b_kx_k2 + e_2 \\ &\vdots \\ y_n &= b_0 + b_1x_1n + b_2x_2n + \cdot + b_knx_kn + e_n \end{aligned}$$

So vector notation we can rewrite the above equations as

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{kn} \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_k \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

or in simplified notation $y = Xb + e$

- b = A column vector with k+1 elements are$b_0, b_1, ..., b_k$
- y= A column vector of n observed values of $y = y_1, ..., y_n$
- X= An n row by K+1 column matrix whose $(i, j + 1)^{th}$ element $X_{i,j+1}$ is 1 if j is 0 else x_{ij}
- e= a column vector of n error terms $e_1, e_2\cdot, e_n$

Parameter estimation:

$$b = (X^TX)^{-1}(X^Ty) \tag{27}$$

Allocation of variation:

$$SSY = \sum_{i=1}^{n} y_i^2 \tag{28}$$

$$SS0 = n\overline{y}^2 \tag{29}$$

$$SST = SSY - SS0 \tag{30}$$

$$SSE = y^Ty - b^TX^Ty \tag{31}$$

$$SSR = SST - SSE \tag{32}$$

Where
SSY=sum of squares of Y
SST=total sum of squares
SS0=sum of squares of y
SSE=sum of squared errrors

SSR= sum of squares given by regression

Coefficient of determination:

$$R^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST} \tag{33}$$

Coefficient of multiple correlation

$$R = \sqrt{\frac{SSR}{SST}} \tag{34}$$

The multiple regression has been applied on dataset,in both the cases where $\rho < 1$ and $\rho > 1$. The findings of different parameters and coefficient of determination have been demonstrated in table 22. The R squared implies that how well the linear regression model can explain the dataset. It is also known as coefficient of determination and the coefficient of multiple determination for multiple regression. The linear expressions produced by the datasets fitted well as the R squared are 99.08 % and 99.99 %.

$$y = 10073 + 113.85x_1 + 848.85x_2 \tag{35}$$

$$y = -1.5782 + 1.0072x_1 + 0.8782x_2 \tag{36}$$

	$\rho < 1$	$\rho > 1$
SSY	$2.1652 * 10^{10}$	$9.5578 * 10^4$
SSO	$1.9979 * 10^{10}$	$9.3381 * 10^4$
SST	$1.6738 * 10^9$	$2.1970 * 10^3$
SSR	$1.6584 * 10^9$	$2.1967 * 10^3$
SSE	$1.5397 * 10^7$	0.3271
R squared	0.9908	.9999

TABLE 23
Multiple Linear Regression Results

12 Conclusion

The data center has become an integral part of IT organizations.Due to more binding budget constraints, most firms typically take recourse to cost minimization to remain competitive in situations where newer technologies have led to significant fall in product prices. Though power is the fastest growing cost among all other cost, but it is not the largest cost if we consider the amortization cost of power, cooling infrastructure for 15 years and new server amortization cost over 3 years [12].Considering all these facts, we tried to identify statistically the major cost contributors towards optimal revenue generation.First we tried to establish our mathematical model based on CES production function to achieve maximum revenue staying within budgetary constraints. Next we used the proposed mathematical model to minimize cost by maintaining the target revenue.In the result and discussion section, we have applied our proposed model on real time dataset, collected from various sources and shown optimal elasticity, ρ value where maximum revenue obtained. 2D graphs are also demonstrated to support the simulation. HHI has been used to identify the concentration of firms and the level of competition in the data center market place. We found competition is low and concentration is moderate in Data center field.Factor analysis procedure has been explored to identify major factors, which are contributing significantly towards output.We have demonstrated two types of factor analysis. In the first case, we consider two cost factors and in the other case take three cost segments.Regression equations, statistical parameters and Pareto, plot graphs are presented in support of the factor analysis.In each of the cases, three scenarios have been explored to identify the major factors. In first case, the server installation cost is the most significant factor, whereas in the second case, We found that Annual I&E, Infrastructure cost etc are the significant factors. Interaction factors, which are not significant at 5% level, have been ignored. Nonparametric estimation has been performed over original and random dataset and graphical representations are displayed to demonstrate the behavior of interaction factors. Almost all the figures exhibit no interaction between the factors apart from scenario $\rho < 1$, where interaction between two factors are happening. A replication experiment is performed and contributions of each factors including effects of interaction factors to revenue generation have been calculated.Confidence interval has been computed for each of the effects for replication experiment. None of the result of confidence interval includes zero.All the effects are significantly different from zero at this confidence level. We are 90% confident that the values of each factor will fall in the range.Multiple linear technique has been applied on the worldwide IT spending data set and it will help to predict the output response, which is in our case revenue, if we know the different cost segments(predictor variables).The optimal ρ value obtained by Least square approach endorse the findings, which we have elaborated in Result & Discussion section.

Since CES function is more general and takes account of number of parameters, it is easy to estimate and that all types of returns can be covered by this production function.

Operation of a Data Center have considered four key cost components (space, power, cooling, operation) in total cost calculation. Individual cost component has been discussed in details based on their dependency on different parameters such as amortization cost, maintenance cost and the influence on the total cost. These apart,other cost factors(Licensing cost, Personnel cost, Operation cost) have also been incorporated in the cost model [14].In contrast, our proposed model are not dependent on the number of cost components included. Our model not only highlighted cost optimization but also had shown how to achieve maximization of target revenue. Using data from large data centers set we have established the efficiency of CES production functions in optimizing revenue. Various challenges related to data center have been discussed in the process and neural network has been implemented to build the mathematical framework. Energy efficiency is the prime objective of the optimization model and the performance of the model is limited by the quality and quantity of data inputs like any other machine learning applications [15].In contrast, our model is independent on the training set and methods of training the machine.

Acknowledgments

The authors would like to thank...

REFERENCES

[1] H. Kopka and P. W. Daly, *A Guide to LATEX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

[2] The Cost of a Cloud: Research Problems in Data Center Networks Albert Greenberg, James Hamilton, David A. Maltz, Parveen Patel Microsoft Research, Redmond, WA, USA

[3] B. P. Rimal, Eunmi Choi, I. Lumb, A Taxonomy and Survey of Cloud Computing Systems, the Fifth International Joint Conference on INC, IMS and IDC, pp. 44 51, 2009.

[4] Gartner Group, available at: http://www.gartner.com/

[5] J. Moore, J. Chase, P. Ranganathan and R. Sharma, Making Scheduling Cool: Temperature-Aware Workload Placement in Data Centers, USENIX Annual Technical Conference, 2005.

[6] N. Rasmussen, Calculating Total Cooling Requirements for Data Centers, white paper, APC Legendary Reliability, available at: http://www.ptsdcs.com/whitepapers/23.pdf

[7] Hamilton,J. Cost of Power in Large-Scale Data Centers [WWW document] http://perspectives.mvdirona.com/2008/11/cost-of-power-in-large-scale-data-centers/ (November 28, 2008).

[8] M. K. S. Gurumurthi, A. Sivasubramaniam and H. Franke, Reducing disk power consumption in servers with drpm, Computer, vol. 36, no. 12, pp. 5966, Dec. 2003.

[9] Robert Basmadjian, Hermann De Meer, Ricardo Lent and Giovanni Giuliani (2012), Cloud computing and its interest in saving energy: the use case of a private cloud, Journal of Cloud Computing: Advances, Systems and Applications 2012, 1:5 doi:10.1186/2192-113X-1-5.

[10] Fan X, Weber WD, Barroso LA (2007), Power provisioning for a warehouse-sized computer, ACM SIGARCH Computer Architecture News, 2007, 35(2): 13—23.

[11] Saravana M, Govidan S, Lefurgy C, Dholakia A (2009),Using online power modeling for server power capping, in Workshop on Energy-Ecient Design 2009, University of Texas and IBM.

[12] https://www.hawaii.edu/powerkills/UFA.HTM

[13] James Hamilton, Cost of Power in Large-Scale Data Centers

[14] http://www.yourarticlelibrary.com/economics/production-functions-4-most-important-production-functions-managerial-economics/28657/

[15] Chandrakant D. Patel et al, Cost Model for Planning, Development and Operation of a Data Center

[16] Lecture Notes on Nonparametrics Bruce E. Hansen University of Wisconsin

[17] https://www.westgard.com/lesson22.htm

[18] The art of computer systems performance analysis, Raj jain, wiley professional computing.

[19] Buettner, Thiess; Kauder, Bjoern;Revenue Forecasting Practices: Differences across Countries and Consequences for Forecasting Performance, Fiscal Studies, September 2010, v. 31, iss. 3, pp. 313-40.

[20] Whitfield, R. I.; Duffy, A. H. B.;Extended Revenue Forecasting within a Service Industry, International Journal of Production Economics, February 2013, v. 141, iss. 2, pp. 505-18.

[21] Krol, Robert;Evaluating State Revenue Forecasting under a Flexible Loss Function, International Journal of Forecasting, April-June 2013, v. 29, iss. 2, pp. 282-89 .

[22] Ray, Bonnie K.;Long-Range Forecasting of IBM Product Revenues Using a Seasonal Fractionally Differenced ARMA Model, International Journal of Forecasting, August 1993, v. 9, iss. 2, pp. 255-69 .

[23] Christian L. Belady, P.E.; In the data center, power and cooling costs more than the it equipment it supports, Feb 2007

[24] http://www.investopedia.com/terms/h/hhi.asp

[25] https://www.srgresearch.com/articles/hp-ibm-and-dell-lead-burgeoning-apac-data-center-infrastructure-market

[26] http://www.businessinsider.in/The-cloud-wars-explained-Why-nobody-can-catch-up-with-Amazon/articleshow/49706488.cms

Jyotirmoy Sarkar Jyotirmoy Sarkar is Senior Design Engineer at GE Healthcare, Bangalore. Jyotirmoy studied for a BTech in Computer Science & Engineering from Govt. College of Engg. & Leather Technology, Kolkata. He worked as a software engineer at Tech Mahindra. During this period, Jyotirmoy undertook MTech at Bits, Pilani in Software systems. Jyotirmoy's fields of interest include Mathematical modeling,cloud computing, machine learning.

PLACE PHOTO HERE

D r. Snehanshu Saha, Ph.D. was born in Calcutta, India, where he completed degrees in Mathematics and Computer Science. He went on to earn his Masters in Mathematical Sciences at Clemson University and his Ph.D. from the Department of Mathematics at University of Texas at Arlington in 2008. After working briefly at his Alma matter, Snehanshu moved to the University of Texas El Paso, where he currently teaches Differential equations, Calculus and Linear Algebra. He is a Professor of Computer Science and Engineering at PESIT South since 2011 and heads the Center for Basic Initiatives in Mathematical Modelling. He has published 30 peer-reviewed articles in International journals and conferences.

Saibal Kar Saibal Kar is in the faculty of economics at the Centre for Studies in Social Sciences, Calcutta, India. He has been a Humboldt Fellow during 2008-09 and held visiting positions in Amsterdam School of Economics, Santa Fe Institute, HWWI, UNU-WIDER and University of East Anglia, in recent years. He is also a Research Fellow of IZA, Bonn. His research interests are broadly in the areas of Labor Economics and International Economics. His books, The outsiders: Economic Reform and Informal Labour in a Developing Economy (with Sugata Marjit), International Trade and Economic Development (with Rajat Acharyya) have been published by the OUP, New Delhi and Clarendon, UK in 2011 and 2014, alongside several papers in internationally reputed journals.

PLACE PHOTO HERE

Bidisha Goswami

PLACE PHOTO HERE

Anand MN