This supplement contains additional material about the paper

SungJin Cho, "An Empirical Model of Mainframe Computer Investment" Journal of Applied Econometrics.

The data used in this paper are very confidential data and the property of one of the largest telecommunications companies in the world. The company handles over 60 percent of all telephone services in its particular market. It also offers several other telecommunication services,including cellular PCS (Personal Communications Service),internet, cable, and satellite communication services. The computer systems in this company can be divided into two categories according to their use: (i) research, and (ii) management and/or delivering services. The following table illustrates the different groups of major tasks and number of systems in terms of the two CPU standards. All mainframe computer systems are associated with specific tasks.

Computers included in the sample in terms of CPU standards

CPU standard	MIPS	TPC
Number*	48	57
Tasks	Billing-Development	Business Info-Management
	Billing-Management	Customer Development
	General Management	Total Document
	New Customer Info-system	Pre-Billing
	Super High Speed Printer	Line-Management
		Material information

*: number of computers in the sample

Since computers used for research use are purchased and replaced on a project basis, their maintenance activities do not reflect technological depreciation. In this paper, I consider only those computer systems used for management or delivering services. I also do not include the replacement of PCs in my model, since in PC replacement there is no upgrade activity, but rather only block purchases or replacements. The time frame of the dataset starts in 1989 and ends in 1999. The data prior to 1989 are incomplete, though some computer systems have a history starting as early as 1977. For the 1989-1999 period, I have the full history of upgrades and replacements for 105 of the company's computer systems. The data consists of purchase dates, purchase prices, specifications, upgrade and replacement dates, upgrade and replacement prices, and other details on each replacement or upgrade such as system specifications. Monthly data on the firm's number of customers is also available.

The estimation procedure by NLS-NFXP is a two step procedure. First, outside of the system, the parameters θ_0 for state variables are estimated separately from the structural parameters. Next, inside of the system, the structural parameters θ_1 are estimated by the nested fixed point algorithm. In other words, inside of the maximum likelihood estimation, the nonlinear least squares estimation (NLS), which is in fact, parametric estimation is performed, and fixed points EV_θ are calculated. Based on the fixed points, the maximum likelihood estimation is performed.

The general method to solve the fixed point problem is a discretization of observed state variables. When the observed state variable is continuous, the required fixed point is in fact an infinite dimensional object.

Therefore, in order to solve the fixed point problem, it is necessary to discretize the state space so that the state variable takes on only finitely many values. But there are limits regarding this method: (i) "curse of dimensionality"; (ii) the limits it places on our ability to solve high-dimensional DP problems. Despite these limits, this method have been used in many literature.

The discretization method may not be appropriate to computer replacement research to solve the fixed point problem, because of the aforementioned problems. The details of shortcomings of discretization method and its alternative methods are in the following

0.1 An attempt of discretization of the state variables

The most conservative dimension of a possible combination of state variables resulting from discretization in the proposed model is 540,000. Discrete variables, capacity and age are discretized as follows. First, I discretize the age variable, g_t, into bimonthly cycle, even though I have monthly data. Thus, age 1 represents a new computer (literally 2 months old.), and an absorbing state 30 means 5 years of age (for estimation purpose, I discretize the age variables into months instead of bi-monthly cycle).

Second, regarding the capacity level, the current dataset of the capacity consists of the three elements of CPU, hard drive and memory size. In order to concretize and transform the capacities into actual numbers which can represent the capacity of each computer system, I take a weighted average of these three elements. Since CPU is the most important factor in the capacity of computer systems, I give it a weight of 0.5. On the other hand, I give equal weights to Hard Drive and Memory size, namely 0.25. At this time, I do not separate the capacity into the two standards of CPU benchmark, TPC and MIPS. Even though the weights were confirmed with the system administrators in the firm, their appropriateness will be verified in further research. With transformed capacities of computer systems, I discretize the capacity from 1 to 40. The last state 40 is the absorbing state. Difference between each step is 30. Therefore, 1 represents (1,...,30), and 2 represents (31,...,60), and 40 represents the range, (1171,...,$+\infty$). These two discrete variables should be discretized regardless of the parametric approximation.

The continuous variables, demand and cost per capacity, can be discretized as follows. First, I discretize demand from 1 to 30. Like the actual capacity, the last state 30 is the absorbing state. Demand 1 represents 100,000 to 105,000 users and the absorbing state 30 is from 245,001 to ∞ users.

Second, I discretize the cost per capacity into 15 possible costs such as $\{15, 14, ..., 1\}$. Difference between subsequent prices is a 20% price drop. I restrict maximum price drops in one period to just 2 steps. These assumptions are based on several research data, computer industry databooks, and Moore's Law. Therefore, the resulting dimension from the discretization is $540,000 = 30 \times 40 \times 30 \times 15$.

The transition probability matrices, $p(d_{t+1}|d_t)$ and $p(c_{t+1}|c_t)$ are in the appendix.

0.2 Computational Burden

First, solving the fixed point problem requires calculation of the expected value function. That is, $EV_\theta = \int\limits_y \int\limits_\eta V_\theta(y, \eta) p(dy, d\eta | x_t, \varepsilon_t, a, \theta_0)$. Even though the Markov transition probability from discretization is a sparse matrix, it still requires extensive time to calculate expectation of value function. Second, the polyalgorithm method by Rust (1987) takes advantage of the complimentary behavior of the two iterations, which are a combination of contraction iteration and policy iteration. (Newton Kantorovich method) This algorithm enjoys a substantial reduction in time calculating the fixed point. However, it is not applicable to solving a dynamic programing model. The reason is as follows: One must have a Frechet derivative $(I - T_\theta^{'})$ in order to use policy iteration method.(the idea of the policy iteration method, i.e., the Newton Kantorovich iteration is to find a zero solution of the nonlinear operator $F = (I - T_\theta)$ instead of finding a fixed point $EV_\theta = T_\theta(EV_\theta)$. With invertibility of $(I - T_\theta)$ and existence of a Frechet derivative $(I - T_\theta^{'})$, one can do a following Taylor expansion: $0 = [I - T_\theta](EV_l)\tilde{}[I - T_\theta](EV_{l-1}) + [I - T_\theta'](EV_l - EV_{l-1}) \Longrightarrow EV_l = EV_{l-1} - [I - T_\theta']^{-1}(I - T_\theta)(EV_{l-1})$.) But, the dimensionality problem makes it impossible to get the derivatives of T_θ. Thus, the algorithm for the DP problem consists solely of a backward induction, which is simple but takes more time to solve. Therefore, the extended time caused by the two aforementioned reasons seriously affects the calculation time of a nested fixed point algorithm, because the nested fixed point algorithm uses the fixed point algorithm outside of the maximum likelihood estimation.

0.3 3. Parametric Estimation

The detailed parametric estimation in this paper is as follows;

To begin with, one needs functional forms for the three value functions, keep, upgrade, and replacement.

$$V(a = 0, x) = H(x, \lambda_0) + \psi_0$$
$$V(a = U, x) = H(x, \lambda_U) + \psi_U$$
$$V(a = K_r, x) = H(x, \lambda_{K_r}) + \psi_{K_r}$$

where $H(x,\lambda_0)$, $H(x,\lambda_U)$ and $H(x,\lambda_{K_r})$ are flexible functions and linear in λ. ψ_0, ψ_U, and ψ_{K_r} are assumed to be distributed as $N(0,1)$

First, I choose the best functional forms for each value function according to the criteria, $\overline{R}^2$. After extended search for the appropriate functional forms of the three value functions, I have the following results.$V(\widehat{a=0},x)$ has 12 parameters ($=\lambda_0$) with 0.983 of $\overline{R}^2$, $V(\widehat{a=U},x)$ has 15 parameters ($=\lambda_U$) with 0.962 of $\overline{R}^2$ and $V(a=\widehat{(K_1...K_n)},x)$ has 18 parameters ($=\lambda_{K_r}$) with 0.962 of $\overline{R}^2$. Therefore, we have $H(x,\lambda_0)\approxeq\sum_{i=1}^{12}\lambda_0^i\vartheta_{0,i}(x)$, $H(x,\lambda_U)\approxeq\sum_{i=1}^{15}\lambda_U^i\vartheta_{U,i}(x)$, and $H(x,\lambda_{K_r})\approxeq\sum_{i=1}^{18}\lambda_{Kr}^i\vartheta_{K_r,i}(x(K_r))$.

Second, with the approximated functional forms of the three value functions, I estimate all 45 parameters $(\lambda_{0,}\lambda_U,\lambda_{K_r})$ with nonlinear least square estimation, such as

$$\min_{\lambda_0,\lambda_U,\lambda_{K_r}}\sum_j\sum_a[V_a(x_j)-U_a]^2$$

where

$U_1=[(\{u(x_t,a_t=0,\theta_1)$
$+\beta\int_y\sigma\log\left(\sum_{a'\in A(y)}\exp[V_{\alpha'}(y)/\sigma]\right)p(dy|x_t,a_t=0,\theta_0)\})]$

and

$U_2=[(\{u(x_t,a=U,\theta_1)$
$+\beta\int_y\sigma\log\left(\sum_{a'\in A(y)}\exp[V_{\alpha'}(y)/\sigma]\right)p(dy|x_t,a_t=U,\theta_0)\})]$

and

$U_3=[(\{u(x_t,a\ =K_r))$
$+\beta\int_y\sigma\log\left(\sum_{a'\in A(y)}\exp[V_{\alpha'}(y)/\sigma]\right)p(dy|x_t,a_t=K_r,\theta_0)\})]$.[1]

Solving the above minimization problem enables us to estimate all parameters $\hat\lambda_0,\hat\lambda_U$, and $\hat\lambda_{K_r}$.In fact, a parametric approximation procedure converts the contraction fixed-point problem into a nonlinear least squares problem.

SungJin Cho
sungcho [AT] snu.ac.kr

[1]The above three expectations are calculated by a quadrature method.