Technical Appendix: The Optimum Quantity of Debt[†]

S. Rao Aiyagari and Ellen R. McGrattan

Federal Reserve Bank of Minneapolis

November 1995

[†] The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Technical Appendix: The Optimum Quantity of Debt

In this appendix, we describe the numerical methods used to compute an equilibrium in the economy with an inelastic labor supply and in the economy with an elastic labor supply (i.e., our benchmark economy). Although the economy with inelastically supplied labor is a special case of the benchmark economy, the equilibrium in the inelastic labor supply case is much easier to compute and is therefore treated separately. In each case, we start with the consumer's problem, assuming the consumer takes prices as given. We then show how the equilibrium prices are determined. To verify that the methods work well for our problem, we apply them to some related test problems that have known solutions.

1. The case with inelastically supplied labor

1.1. Computing the consumer's decision functions

The consumer chooses sequences of consumption and asset holdings to maximize expected utility; i.e.,[1]

$$\max_{\{\tilde{c}_t,\tilde{a}_{t+1}\}} E\big[\sum_{t=0}^{\infty}(\beta(1+g)^{1-\nu})^t\tilde{c}_t^{1-\nu}/(1-\nu)|\tilde{a}_0,e_0\big] \tag{1}$$

$$\text{subject to}\quad \tilde{c}_t+(1+g)\tilde{a}_{t+1}\leq(1+\bar{r})\tilde{a}_t+\bar{w}e_t+\chi, \tag{2}$$

$$\tilde{a}_t\geq 0, \tag{3}$$

with the after-tax interest rate $\bar{r}$, the after-tax wage rate $\bar{w}$, and a lump-sum transfer χ taken as given. The productivity level e_t is assumed to follow a Markov chain with $Ee_t=1$. The probability of transiting from state i to state j is given by $\pi_{i,j}$, $i,j=1,\ldots,m$.

To incorporate the inequality constraints in Eq. (3), we replace the objective in Eq. (1) with

$$E\left[\sum_{t=0}^{\infty}(\beta(1+g)^{1-\nu})^t\left\{\frac{\tilde{c}_t^{1-\nu}}{1-\nu}+\tfrac{1}{3}\zeta\min(\tilde{a}_t,0)^3\right\}|\tilde{a}_0,e_0\right], \tag{4}$$

[1] As in the paper, we normalize variables by dividing through by output, and we assume that $\tilde{a}_t\geq 0$ is sufficiently restrictive for the parameters that we consider.

where ζ is some positive parameter. Note that if $\tilde{a}_t$ is negative, then we subtract $-\zeta \tilde{a}_t^3$ from the consumer's value function. With this respecification, the optimization problem is as follows: given ζ, we choose sequences of consumption and assets that are optimal for Eq. (4) subject to Eq. (2). The parameter ζ is set so that Eq. (3) is approximately satisfied. In practice, the optimization is done either by iterating over a sequence 1, 10, 10^2, etc. for ζ, until the constraints are satisfied to within the tolerance, or by starting with a reasonably large value of ζ.[2]

Because the decision rules are stationary, we can compute the functions $c(x,i)$ and $\alpha(x,i)$ that satisfy the following first-order conditions for $i = 1, \ldots, m$ and $x \in [0, x_{max}]$:

$$(1+g)c(x,i)^{-\nu} = \beta(1+g)^{1-\nu}\{\sum_j \pi_{i,j}\,(1+\bar{r})c(x',j)^{-\nu} + \zeta \min(\alpha(x,i),0)^2\}, \tag{5}$$

$$c(x,i) + (1+g)\alpha(x,i) = (1+\bar{r})x + \bar{w}e(i) + \chi, \tag{6}$$

where $x' = \alpha(x,i)$ and $e(i)$ is the productivity level in state i. Note that these conditions assume an interior solution for $c(x,i)$. If we substitute the expression for $c(x,i)$ in (6) into Eq. (5), we have a functional equation in α. We denote this expression by $R(x,i;\alpha)$; i.e.,

$$\begin{aligned} R(x,i;\alpha) = {} & (1+g)\{(1+\bar{r})x + \bar{w}e(i) + \chi - (1+g)\alpha(x,i)\}^{-\nu} \\ & - \beta(1+g)^{1-\nu}\{\sum_j \pi_{i,j}\,\{(1+\bar{r})(1+\bar{r})\alpha(x,i) + \bar{w}e(j) \\ & + \chi - (1+g)\alpha\big(\alpha(x,i),j\big)\}^{-\nu} + \zeta \min(\alpha(x,i),0)^2\}. \end{aligned} \tag{7}$$

The computational task is, therefore, to find an approximation for $\alpha(x,i)$ – say $\alpha^h(x,i)$ – that implies $R(x,i;\alpha^h) \approx 0$ for all x and i. We accomplish this task by applying a finite element method. In particular, we do the following. We choose some discretization of the domain of our functions. Since only x is continuous, we need to specify some partition of $[0,x_{max}]$ where x_{max} is such that no $x > x_{max}$ would be chosen by the consumer. We refer to each subinterval of x as an element. On each element, we choose a set of basis functions for approximating α; that is, we assume α^h can be represented as a weighted sum of basis

[2] See R. Fletcher (1987), *Practical Methods of Optimization* (New York: Wiley) for a discussion of the relationship of ζ to the Lagrange multipliers of the constraints in Eq. (3).

functions, where the weights and basis functions may be different for each element. In our case, we choose linear basis functions for all elements; e.g.

$$\alpha^h(x,i) = \psi_e^i N_e(x) + \psi_{e+1}^i N_{e+1}(x), \qquad N_e(x) = \frac{x_{e+1}-x}{x_{e+1}-x_e},\ N_{e+1}(x) = \frac{x-x_e}{x_{e+1}-x_e},$$

on the element $[x_e, x_{e+1}]$. This choice is motivated by our test problem (see below). Notice that $\alpha^h(x_e,i) = \psi_e^i$ and $\alpha^h(x_{e+1},i) = \psi_{e+1}^i$. If we consider the approximation globally, we need to compute values for ψ_e^i for all nodes e (assume there are n) and for all levels of productivity i. We choose these values for ψ_e^i by setting the weighted residual to zero; i.e.,

$$\int R(x,i;\alpha^h)N_e(x)dx = 0, \qquad i = 1,\ldots,m,\ e = 1,\ldots n. \tag{8}$$

In effect, we solve a problem of the following form: find $\vec{\psi}$ such that $h(\vec{\psi}) = 0$, where $\vec{\psi}$ is the vector of coefficients that we are searching over and h is the system of equations in Eq. (8).

There are several practical points worth noting. The first point is that, for parameterizations in which the no-borrowing constraint binds, the penalty function only serves to get $\alpha(x,i) > -\epsilon$, where ϵ is small but positive. Thus, if the function is truly equal to zero for low values of x and i, then the algorithm will not yield a mass point at $x = 0$. To deal with this problem, we use a two-step procedure. At the first step, we apply the penalty function method and choose a sufficiently fine mesh to resolve the kink. The kink is defined as the grid point at which the second derivative is maximized. At the second step, we use the candidate solution to determine the boundary conditions $\alpha^h(x,i) = 0$, $x < x^*$, where x^* is the grid point at which the second derivative of $\alpha^h(x,i)$ is highest. These boundary constraints are imposed on the solution prior to the final run. The second point concerns the procedure for solving $h(\vec{\psi}) = 0$. We solve this system of equations by applying a Newton method. Therefore, we need to compute the derivative of $h(\vec{\psi})$ with respect to $\vec{\psi}$, and we need to invert it. We calculate analytical derivatives because it saves computing time. With respect to inverting the Jacobian, we can rely on the fact that it is very sparse. The sparseness is due to the fact that the approximation is done element by element; thus, the basis functions are nonzero on relatively small subdomains. In practice, however, we only use sparse matrix routines if the dimension of the matrix to be inverted is bigger than 2000×2000.

1.2. Computing the distribution of assets

To compute equilibrium prices, we must first construct the joint distribution over assets and productivities. This distribution can be derived from the decision rule for asset holdings; i.e.,

$$x_{t+1} = \alpha(x_t, e_t),$$

where x_t is asset holdings in t and e_t is the earnings shock in t. We want to compute the invariant distribution for asset holdings; namely, $H(x,i) = Pr(x_t < x \,|\, e_t = e(i))$. To compute the distribution, we solve the following functional equation:

$$H(x,i) = \sum_{j=1}^{m} \pi_{j,i} H(\alpha^{-1}(x,j), j) I(x \geq \alpha(0,j)), \tag{9}$$

where π is the transition matrix for the Markov chain governing earnings and I is an indicator function (i.e., $I(x > y)$ is equal to one if $x > y$ and is equal to zero otherwise). The form of the functional equation in (9) is motivated below.

Suppose X_1 is a random variable with density function f_1. Assume that X_2 is a second random variable that is given by $X_2 = h(X_1)$ for some function h. In this case, the density function for X_2 is

$$f_2(x_2) = \int_{-\infty}^{\infty} f_1(x_1)\delta(x_2 - h(x_1))dx_1, \tag{10}$$

where δ is the Dirac delta function. If we integrate the left-hand side of Eq. (10), we get

$$\begin{aligned} F_2(x_2) = \int_{-\infty}^{x_2} f_2(s)ds &= \int_{-\infty}^{x_2}\int_{-\infty}^{\infty} f_1(x_1)\delta(s - h(x_1))dx_1 ds \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{x_2} \delta(s - h(x_1))ds\right] f_1(x_1)dx_1 ds \\ &= \int_{-\infty}^{\infty} I(h(x_1) \leq x_2) f_1(x_1)dx_1 \\ &= \int_{-\infty}^{\infty} I(h(x_1) \leq x_2) dF_1(x_1). \end{aligned} \tag{11}$$

Now assume that h is strictly increasing and, hence, invertible. In this case, we can rewrite (11) as follows:

$$\begin{aligned}F_2(x_2) &= \int_{-\infty}^{\infty} I(h^{-1}(h(x_1)) \le h^{-1}(x_2))dF_1(x_1)\\ &= \int_{-\infty}^{\infty} I(x_1 \le h^{-1}(x_2))dF_1(x_1)\\ &= \int_{-\infty}^{h^{-1}(x_2)} dF_1(x_1)\\ &= F_1(h^{-1}(x_2)).\end{aligned}$$

In the example with assets, we have α increasing but not strictly. To motivate the functional equation in Eq. (9), we consider cases in which the borrowing constraints do not bind and cases in which they do and conclude that both cases imply the same functional equation for H. First, consider states of the world where the borrowing constraint does not bind (e.g., values of i such that $\alpha(x,i) > 0$ for all $x \ge 0$). Using (11), we know that the equation to be solved is

$$H_{n+1}(x',i) = \sum_{j=1}^{m} \pi_{j,i} \int_{-\infty}^{\infty} I(\alpha(x,j) \le x')dH_n(x,j), \tag{12}$$

where H_n is the cumulative distribution function for today's assets, H_{n+1} is the cumulative distribution function for tomorrow's assets, and both H_n and H_{n+1} are functions of the earnings state. If an invariant distribution exists, it is given by the fixed point of Eq. (12); i.e., $H = H_n = H_{n+1}$. In this case, $I(\alpha(x,j) \le x')$ is equal to one for $x' \ge \alpha(0,j)$ and is equal to zero for $x' < \alpha(0,j)$. Therefore,

$$\int_{-\infty}^{\infty} I(\alpha(x,j) \le x')dH_n(x,j) = \begin{cases} H_n(\alpha^{-1}(x',j),j) & \text{if } x' \ge \alpha(0,j)\\ 0 & \text{if } x' < \alpha(0,j),\end{cases}$$

and (12) can be rewritten as follows:

$$H_{n+1}(x',i) = \sum_{j=1}^{m} \pi_{j,i} H_n(\alpha^{-1}(x',j),j) I(x' \ge \alpha(0,j)). \tag{13}$$

Second, consider states of the world where the borrowing constraint binds (e.g., values of i such that $\alpha(x,i) = 0$ for $x \in [0,x^*]$ where $x^* > 0$). In these cases, if we assume that $\alpha^{-1}(0,j) = x^*$ where x^* is the point at which we see a kink, then we again have the mapping in (13). If we assume that $H_n = H_{n+1}$ at a fixed point, the functional equation for the invariant distribution is given by Eq. (9).

To compute H, we again apply the finite element method with linear basis functions. In this case, the residual is the difference between the right- and left-hand sides of Eq. (9). For this problem, we do not have to worry about inequality constraints directly, but we do have to deal with them indirectly. If inequality constraints bind in the consumer's problem, then the decision functions for low productivity levels will be set to zero for some interval $[0, x^*]$. Thus, there will be a mass point at $x = 0$ and throughout the distribution at points traversed prior to reaching the zero-asset position. The mass points in the distribution imply that the solution to (9) has discontinuities, possibly at a countably infinite number of points. Thus, we need to know if the finite element method, as we implement it, will yield a good approximation to the distribution H.

1.3. Computing an equilibrium

Assume that we have in place numerical algorithms for computing $\alpha(x, i)$ and $H(x, i)$. We now describe an algorithm for computing the equilibrium after-tax interest rate. All other prices and quantities can be determined from the interest rate.

We start with bounds on the interest rate r; that is, we assume that the equilibrium interest rate is in the interval $[r_l, r_u]$. Our inputs are the share of capital in production θ, the government consumption-to-GDP ratio γ, the debt-to-GDP ratio b, the discount factor β, the rate of capital depreciation δ, the utility parameters η and μ, the growth rate g, the initial bounds on r, the productivity levels $e(i)$, $i = 1, \ldots, m$, and the transition probability matrix π. Our initial guess for the equilibrium r is $\frac{1}{2}(r_l + r_u)$. If lump-sum taxes are assumed, then we set the taxes equal to $\gamma + \chi + (r - g)b$. If proportional taxes are assumed, then we back out the tax rate from the government budget constraint; i.e.,

$$\tau_y = \frac{\gamma + \chi + rb - gb}{1 + rb - \delta\theta/(r + \delta)}. \tag{14}$$

With τ_y and r, we have the after-tax interest rate $\bar{r} = (1 - \tau_y)r$. We can also compute the after-tax wage rate

$$\bar{w} = (1 - \tau_y)(1 - \theta). \tag{15}$$

We now have the inputs needed for computing finite element approximations of α and H. The next step is to calculate the mean asset holdings $E\tilde{a}_t$. Since our approximation for H is piecewise linear, an estimate of the mean asset holdings is simply

$$\frac{1}{2}\sum_j \sum_i (\hat{H}(i+1,j) - \hat{H}(i,j))(x(i+1) + x(i)),$$

where $x(i)$ is the ith grid point over asset holdings and $\hat{H}(i,j)$ is the finite element approximation of $H(x(i),j)$. Given this estimate of $E\tilde{a}_t$, we check to see if markets are cleared; i.e., if

$$E\tilde{a}_t = \frac{\theta}{r+\delta} + b. \tag{16}$$

If the right-hand side of (16) is larger than the left-hand side, then we set $r_l = r$ and repeat the above steps. Otherwise, we set $r_u = r$ and repeat. Iterations are made until Eq. (16) is satisfied to within some tolerance level.

2. The case with elastically supplied labor

2.1. Computing the consumer's decision functions

The consumer chooses sequences of consumption, asset holdings, and leisure to maximize expected utility; i.e.,

$$\max_{\{\tilde{c}_t,\tilde{a}_{t+1},\ell_t\}} E\Big[\sum_{t=0}^{\infty}(\beta(1+g)^{\eta(1-\mu)})^t(\tilde{c}_t^{\eta}\ell_t^{1-\eta})^{1-\mu}/(1-\mu)|\tilde{a}_0,e_0\Big] \tag{17}$$

$$\text{subject to}\quad \tilde{c}_t + (1+g)\tilde{a}_{t+1} \leq (1+\bar{r})\tilde{a}_t + \bar{w}e_t(1-\ell_t) + \chi, \tag{18}$$

$$\tilde{a}_t \geq 0, \tag{19}$$

$$\ell_t \leq 1, \tag{20}$$

with the after-tax interest rate $\bar{r}$, the after-tax wage rate $\bar{w}$, and a lump-sum transfer χ taken as given. In specifying the optimization problem, we have left out two inequality constraints; namely, $c_t \geq 0$ and $\ell_t \geq 0$. When we compute the decision functions, these inequality constraints are ignored, but we check the solution to make sure that they are satisfied. The productivity level e_t is assumed to follow a Markov chain with $Ee_t = 1$. The probability of transiting from state i to state j is given by $\pi_{i,j}$, $i, j = 1, \ldots, m$.

To incorporate the inequality constraints in equations (19) and (20), we replace (17) with

$$E\Bigg[\sum_{t=0}^{\infty}(\beta(1+g)^{\eta(1-\mu)})^t\Bigg\{\frac{(\tilde{c}_t^{\eta}\ell_t^{1-\eta})^{1-\mu}}{1-\mu}+\frac{1}{3}\zeta(\min(\tilde{a}_t,0)^3+\min(1-\ell_t,0)^3)\Bigg\}|\tilde{a}_0,e_0\Bigg], \quad (21)$$

where ζ is some positive parameter. Note that if $\tilde{a}_t$ is less than zero, then we subtract $-\zeta\tilde{a}_t^3$ from the consumer's value function. If ℓ_t is greater than one, then we subtract $-\zeta(1-\ell_t)^3$. With this respecification, the optimization problem is as follows: given ζ, we choose sequences of consumption, asset holdings, and leisure that are optimal for Eq. (21) subject to Eq. (18). The parameter ζ is set so that Eq. (19) is approximately satisfied. In practice, the optimization is done either by iterating over a sequence 1, 10, 10^2, etc. for ζ, until the constraints are satisfied to within the tolerance, or by starting with a reasonably large value of ζ.

Because the decision rules are stationary, we can compute the functions $c(x,i)$, $\alpha(x,i)$, and $\ell(x,i)$ that satisfy the following first-order conditions:

$$\eta(1+g)c(x,i)^{\eta(1-\mu)-1}\ell(x,i)^{(1-\eta)(1-\mu)} = \beta(1+g)^{\eta(1-\mu)}\{\sum_j \pi_{i,j}\,\eta(1+\bar{r}) \cdot c(x',j)^{\eta(1-\mu)-1}\ell(x',j)^{(1-\eta)(1-\mu)} + \zeta\min(\alpha(x,i),0)^2\}, \quad (22)$$

$$(1-\eta)c(x,i)^{\eta(1-\mu)}\ell(x,i)^{(1-\eta)(1-\mu)-1} = \eta\bar{w}e(i)c(x,i)^{\eta(1-\mu)-1}\ell(x,i)^{(1-\eta)(1-\mu)} + \zeta\min(1-\ell(x,i),0)^2, \quad (23)$$

$$c(x,i)+(1+g)\alpha(x,i) = (1+\bar{r})x+\bar{w}e(i)(1-\ell(x,i))+\chi, \quad (24)$$

$$x' = \alpha(x,i), \quad (25)$$

for $i = 1, \ldots, m$ and $x \in [0, x_{max}]$, where $e(i)$ is the productivity level in state i. Notice that $c(x,i)$ is a function of x, i, $e(i)$, the parameters, and the function $\alpha(\cdot)$. Therefore, with a candidate solution for α, we can back out c using (24). This is not the case for leisure, however, since (23) is an implicit function of $\ell(x,i)$ (if we assume that x, i, $e(i)$, the parameters, $\alpha(\cdot)$, and $c(\cdot)$ are known). However, we can construct a numerical solution for $\ell(x,i)$ by applying a Newton method with given values for η, μ, ζ, $\bar{r}$, $\bar{w}$, χ, g, x, i, $e(i)$, and $\alpha(x,i)$. We simply iterate on

$$\ell^{k+1} = \ell^k - f(\ell^k)/J(\ell^k), \qquad k = 0, 1, \ldots, \tag{26}$$

where

$$\begin{aligned} f(\ell) &= (1-\eta)c(\ell)^{\eta(1-\mu)}\ell^{(1-\eta)(1-\mu)-1} - \zeta\min(1-\ell,0)^2 \\ &\quad - \bar{w}e(i)\eta c(\ell)^{\eta(1-\mu)-1}\ell^{(1-\eta)(1-\mu)}, \\ c(\ell) &= (1+\bar{r})x + \bar{w}e(i)(1-\ell) + \chi - (1+g)\alpha(x,i), \end{aligned} \tag{27}$$

$$\begin{aligned} J(\ell) &= -2\bar{w}e(i)(1-\eta)\eta(1-\mu)c(\ell)^{\eta(1-\mu)-1}\ell^{(1-\eta)(1-\mu)-1} \\ &\quad + (1-\eta)((1-\eta)(1-\mu)-1)c(\ell)^{\eta(1-\mu)}\ell^{(1-\eta)(1-\mu)-2} \\ &\quad + \bar{w}^2 e(i)^2\eta(\eta(1-\mu)-1)c(\ell)^{\eta(1-\mu)-2}\ell^{(1-\eta)(1-\mu)} \\ &\quad + 2\zeta\min(1-\ell,0). \end{aligned} \tag{28}$$

The function f is the Euler equation in (23), the function c is consumption derived from the budget constraint, and the function J is the derivative of f with respect to ℓ. We start the iterations in Eq. (26) with an initial guess ℓ^0, and we stop when $|\ell^{k+1} - \ell^k|$ is less than some tolerance parameter.

Let $\ell^*(x,i;\alpha)$ be the solution to Eq. (26). Then we can write the first-order conditions in Eqs. (22)-(25) in terms of one residual; i.e.,

$$\begin{aligned} R(x,i;\alpha) &= \eta(1+g)c(\ell^*(x,i;\alpha))^{\eta(1-\mu)-1}\ell^*(x,i;\alpha)^{(1-\eta)(1-\mu)} \\ &\quad - \beta(1+g)^{\eta(1-\mu)}\{\sum_j \pi_{i,j}\,\eta(1+\bar{r})c(\ell^*(\alpha(x,i),j;\alpha))^{\eta(1-\mu)-1} \\ &\quad \cdot \ell^*(\alpha(x,i),j;\alpha)^{(1-\eta)(1-\mu)} + \zeta\min(\alpha(x,i),0)^2\}, \end{aligned}$$

where $c(\cdot)$ is defined in (27). If we ignore the constraint on leisure (i.e., $\ell_t \leq 1$), the residual is given by

$$R(x,i;\alpha) = \eta(1+g)\left\{\frac{1-\eta}{\eta\bar{w}e(i)}\right\}^{(1-\eta)(1-\mu)} \{(\bar{w}e(i) + (1+\bar{r})x - (1+g)\alpha(x,i) + \chi)\eta\}^{-\mu}$$

$$- \beta(1+g)^{\eta(1-\mu)} \Bigg\{ \sum_j \pi_{i,j}\, \eta(1+\bar{r}) \Bigg\{ \frac{1-\eta}{\eta \bar{w} e(j)} \Bigg\}^{(1-\eta)(1-\mu)} \{(\bar{w}e(j)$$

$$+ (1+\bar{r})\alpha(x,i) - (1+g)\alpha(\alpha(x,i),j) + \chi)\eta\}^{-\mu} + \zeta \min(\alpha(x,i),0)^2 \Bigg\}. \quad (29)$$

We show later that other calculations are also simplified if an interior solution for leisure is assumed. However, for some parameterizations, we need to impose the constraint on leisure. Therefore, we describe how to compute the equilibrium assuming the constraint is violated in some regions of the state space.

As in the inelastic labor case, the computational task is to find an approximation for $\alpha(x,i)$, say $\alpha^h(x,i)$, that implies $R(x,i;\alpha^h) \approx 0$ for all x and i. We follow the same procedure described for the inelastic labor case (i.e., we apply the finite element method).

2.2. *Computing the distribution of assets*

Computation of the cumulative distribution is the same for the elastic and inelastic labor cases. Therefore, the description in section 1.2 carries over to this case.

2.3. *Computing an equilibrium*

Assume that we have in place numerical algorithms for computing $\alpha(x,i)$ and $H(x,i)$. We now describe an algorithm for computing the equilibrium interest rate and aggregate hours. All other prices and quantities can be determined from these two.

We start with an initial guess for N. With N fixed, we apply a bisection method to calculate the interest rate that clears the asset market. Assume that the equilibrium r is in the interval $[r_l, r_u]$ and that our initial guess for the equilibrium r is $\frac{1}{2}(r_l+r_u)$. If lump-sum taxes are assumed, then we set the taxes equal to $\gamma + \chi + (r-g)b$. If proportional taxes are assumed, then we back out the tax rate from the government budget constraint. Given the tax rate and the guess for N, we can determine the after-tax interest rate $(1-\tau_y)r$ and the after-tax wage rate $(1-\tau_y)(1-\theta)/N$. These two rates are used to compute the finite element approximation of α. To check to see that the asset market has cleared, we also compute a finite element approximation of H. H is used to calculate the mean asset holdings $E\tilde{a}_t$. Given this estimate of $E\tilde{a}_t$, we check to see if markets are cleared; i.e., if

$$E\tilde{a}_t = \frac{\theta}{r+\delta} + b. \tag{30}$$

If the right-hand side of (30) is larger than the left-hand side, then we set $r_l = r$; otherwise, we set $r_u = r$. Iterations are made until Eq. (30) is satisfied to within some tolerance level. Once Eq. (30) is satisfied, we update our guess for N. The updating scheme is simply Newton-Raphson:

$$N^{k+1} = N^k - J(N^k)^{-1} f(N^k),$$

where

$$f(N) = \sum_i \int e(i)(1 - \ell^*(x, i; \alpha)) dH(x, i) - N, \tag{31}$$

and J is the derivative of f with respect to N. Note that the first term on the right-hand side of (31) is a function of N because the optimal decision rules depend on it.

3. Test problems

3.1. Computing the consumer's decision functions – a test case

Assume that the household solves

$$\max_{\{c_t, a_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$
$$\text{subject to} \quad c_t + a_{t+1} = (1+r)a_t + w.$$

This specification assumes that there is no uncertainty ($e_t = 1$); therefore, wages are constant.

The dynamic program for this example involves the following form for Bellman's equation:

$$v(x) = \max_{0 \le y \le Rx+w} \{u(Rx + w - y) + \beta v(y)\}, \tag{32}$$

where $y_t = x_{t+1}$ is the asset position next period and $R = 1 + r$ is the gross return. A conjectured solution is as follows:

$$\begin{cases} y = 0 & \text{if } x \in [0, m_1], \\ y = \frac{-m_1^2}{m_2 - m_1} + \frac{m_1}{m_2 - m_1}\ x & \text{if } x \in (m_1, m_2], \\ y = \frac{(m_3 m_1 - m_2^2)}{m_3 - m_2} + \frac{m_2 - m_1}{m_3 - m_2}\ x & \text{if } x \in (m_2, m_3], \\ \vdots \end{cases} \tag{33}$$

where m_j, $j = 1, 2, \ldots$, will be calculated below. Note that the solution assumes that if $x = m_{j+1}$, then $y = m_j$.

The Lagrangian for the maximization in the right-hand side of Eq. (32) is given by

$$L = u(xR + w - y) + \beta v(y) + p(Rx + w - y) + qy,$$

where p and q are multipliers. The first-order conditions for this problem are

$$\begin{aligned} & -u'(Rx + w - y) + \beta v'(y) - p + q = 0, \\ & p \geq 0, \quad Rx + w - y \geq 0, \quad p(Rx + w - y) = 0, \\ & q \geq 0, \quad y \geq 0, \quad qy = 0. \end{aligned} \tag{34}$$

If we assume that the conjecture above is correct, then when $x \in [0, m_1)$ we have the $y \geq 0$ constraint binding. Therefore, if we assume that $Rx + w > 0$, then it must be true that $y < Rx + w$, $p = 0$, and

$$v'(0) = \frac{1}{\beta} u'(Rx + w) - \frac{q}{\beta} \leq \frac{1}{\beta} u'(Rx + w).$$

Furthermore, from Bellman's equation, we get

$$v(x) = u(Rx + w) + \beta v(0), \quad \text{for } x \in (0, m_1) \ \text{ and } v(0) = \frac{\beta}{1 - \beta} u(w),$$

and taking derivatives, we get

$$v'(x) = Ru'(Rx + w) < \frac{1}{\beta} u'(Rx + w), \tag{35}$$

since $\beta R < 1$.

Consider next the interval $(m_1, m_2]$. The conjectured solution is such that in this interval, the constraint $y \geq 0$ is not binding. If we assume that $Rx + w > y$, then the first-order conditions imply

$$v'(y) = \frac{1}{\beta} u'(Rx + w - y). \tag{36}$$

If $y = 0$ at $x = m_1$, then

$$v'(0) = \frac{1}{\beta} u'(Rm_1 + w). \tag{37}$$

Using Eq. (35) evaluated at $x = 0$ and Eq. (37), we get

$$u'(Rm_1 + w) = \beta R u'(w),$$

which gives us an equation for m_1. For example, if $u(c) = c^{1-\mu}/(1-\mu)$, then

$$m_1 = \frac{w(1 - (\beta R)^{\frac{1}{\mu}})}{(\beta R)^{\frac{1}{\mu}} R},$$

and $y = 0$ in the interval $[0, m_1]$.

Now we want to compute the asset function for the next interval $(m_1, m_2]$. If the conjecture in Eq. (33) is correct, then Eq. (36) holds, as does

$$v'(x) = Ru'(Rx + w - g(x)), \tag{38}$$

which is the derivative of the value function once y is replaced by the optimal policy $y = g(x)$. The conjecture assumes that $y = m_1$ when $x = m_2$, and by Eq. (36) and Eq. (38), we have,

$$u'(Rm_2 + w - m_1) = \beta R u'(Rm_1 + w).$$

Note that this equation can be solved for m_2. If we follow the same logic for the remaining m's, we find that, in general,

$$u'(Rm_{j+1} + w - m_j) = \beta R u'(Rm_j + w - m_{j-1}), \quad j = 1, \ldots, \text{ and } m_0 = 0. \tag{39}$$

Thus, given m_1 and m_2, we can compute m_3 and so on.

What we have done is conjectured a solution and derived the functions analytically. It is easy to show that the solution is, in fact, piecewise linear and that the conjecture is correct.

Now we consider the finite element approximation. Let $\beta = 0.95$, $w = 1.0$, $r = 0.02$, $u(c) = c^{1-\mu}/(1-\mu)$, and $\mu = 3$. We can use the formula in Eq. (39) to derive the exact solution. In Figure 1a, we plot the true solution and the finite element approximation assuming the grid is given by m_j, $j = 0, \ldots, 14$. Notice that the two functions are in very good agreement. In Figure 1b, we look at exactly the same picture, but we focus on the corner, where the constraints bind. The finite element approximation in this case does not assume any boundary constraints are imposed. Notice that there is a slight deviation of the approximate and exact solutions because the penalty function ensures only approximate satisfaction of the inequality constraints. The approximate solution in Figure 1c imposes that the asset holdings are zero at the first two grid points. Notice that the fit is exact. In Figures 2a-2c, we show the results of a more realistic experiment. In this case, we start with a grid that is "stretched" exponentially: the intervals get exponentially larger the farther they are from the origin. To obtain the results of Figure 2a, we use 15 grid points ($n = 15$). Notice in Figure 2a that the grid points of the approximate solution do not line up exactly with the kinks in the exact solution. Notice also that boundary conditions are not imposed. In Figure 2b, we add two grid points around the value of x where the decision function becomes positive. The picture is similar to Figure 2a. In Figure 2c, we impose that asset holdings are zero for the first three grid points. Notice how close we get to the exact solution even though we do not use the best possible discretization. Finally, in Figure 3, we plot the decision functions for three different sized grids: $n = 5$, $n = 9$, and $n = 17$. In all three cases, we impose some boundary conditions. If we plot the solutions for all x, the picture looks like Figure 1a. In Figure 3, we focus on only the corner. Notice how close the approximate function is to the exact function, even with $n = 5$.

3.2. Computing the distribution of assets – a test case

Suppose the productivities take on two possible values and the decision functions are given by

$$\alpha(x, i) = \begin{cases} \max(0, -0.25 + x), & \text{if } i = 1 \\ 0.5 + 0.5x, & \text{if } i = 2, \end{cases}$$

with $\pi_{1,1} = \pi_{2,2} = 0.8$. In this case,

$$
\begin{aligned}
H(0,1) &= 0.8H(0.25,1), \\
H(0,2) &= 0.2H(0.25,1), \\
H(0.25,1) &= 0.8H(0.5,1), \\
H(0.25,2) &= 0.2H(0.5,1), \\
H(0.5,1) &= 0.8H(0.75,1) + 0.2H(0,2), \\
H(0.5,2) &= 0.2H(0.75,1) + 0.8H(0,2), \\
H(0.75,1) &= 0.8H(1.0,1) + 0.2H(0.5,2), \\
H(0.75,2) &= 0.2H(1.0,1) + 0.8H(0.5,2), \\
H(0.875,1) &= 0.8H(1.125,1) + 0.2H(0.75,2), \\
H(0.875,2) &= 0.2H(1.125,1) + 0.8H(0.75,2), \\
H(1.0,1) &= 0.5, \\
H(1.0,2) &= 0.5.
\end{aligned}
$$

If we assume that $H(x,i) = 0.5$ for $x > 1$, then the above expressions can easily be solved. We can first determine $H(0,j)$, $H(0.25,j)$, $H(0.5,j)$, and $H(0.75,j)$ for j =1,2 by solving $Ax = b$, where

$$
A = \begin{bmatrix}
1 & -.8 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -.8 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -.8 & -.2 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & -.2 & 0 \\
0 & -.2 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & -.2 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & -.2 & -.8 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -.8 & 1
\end{bmatrix}
\qquad
b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ .4 \\ 0 \\ 0 \\ 0 \\ .1 \end{bmatrix}.
$$

The solution is $H(0,1) = 0.225$, $H(0.25,1) = 0.282$, $H(0.5,1) = 0.352$, $H(0.75,1) = 0.426$, $H(0,2) = 0.056$, $H(0.25,2) = 0.070$, $H(0.5,2) = 0.130$, and $H(0.75,2) = 0.204$. Note that we can back out the other points from these solutions by applying the formula in Eq. (9). In Figure 4a, we plot the finite element approximation and the exact solution for a relatively coarse grid ($n = 13$) with evenly spaced partitions. Notice that a grid this coarse will not resolve the discontinuities. In Figure 4b, we refine the grid by doubling the number of partitions (i.e., $n = 25$); in this case, they are still evenly spaced. In Figures 4c and 4d, we recompute with an evenly spaced grid but with finer partitions (i.e., $n = 49$ and $n = 97$). In both of these cases, it is clear where the discontinuities occur. In Figure 4e, we plot the finite element approximation for a grid adapted to better resolve the discontinuities. Note that we get very good agreement with the above solution.

In Figures 4f and 4g, we focus on regions where jumps occur to illustrate how the method performs. Although the solution is relatively accurate, there are slight deviations at the discontinuities due to the fact that the method is essentially "centrally differencing." That is, to determine the solution at $x = .5$, information is used from above and below this point. The result is nonmonotonicity near discontinuities. To fix this nonmonotonicity, we simply adjust the final solution as follows. In regions just to the left of large gradients (i.e., discontinuities), if we see $\hat{H}(i,j) < \hat{H}(i-1,j)$, then we set $\hat{H}(i,j) = \hat{H}(i-1,j)$. In regions to the right of discontinuities, if we see $\hat{H}(i,j) < \hat{H}(i-1,j)$, then we look at points $k = i+1, i+2, \ldots$, until we find $\hat{H}(k,j) \geq \hat{H}(k-1,j)$. Once we find such a point, we set $\hat{H}(l,j) = \hat{H}(k-1,j)$, $l = i-1, \ldots, k-1$. The solution shown in Figure 4h applies this "trick." If we compare this solution to Figure 4f, we see that the nonmonotonicity is eliminated and that the solution is closer to the exact solution.

Finally, we should note that for this example, if values of x greater than one are not included in the grid, then the approximation is very inaccurate. This inaccuracy is due to the fact that values of $H(x,i)$, $x < 1$ are used to construct an estimate of $H(x,i)$, $x > 1$. For this example, this extrapolation will lead to the inaccurate approximation shown in Figure 5. Notice that the approximation in Figure 5 lies above the true solution and violates the terminal condition that $H(x,i) = 0.5$.

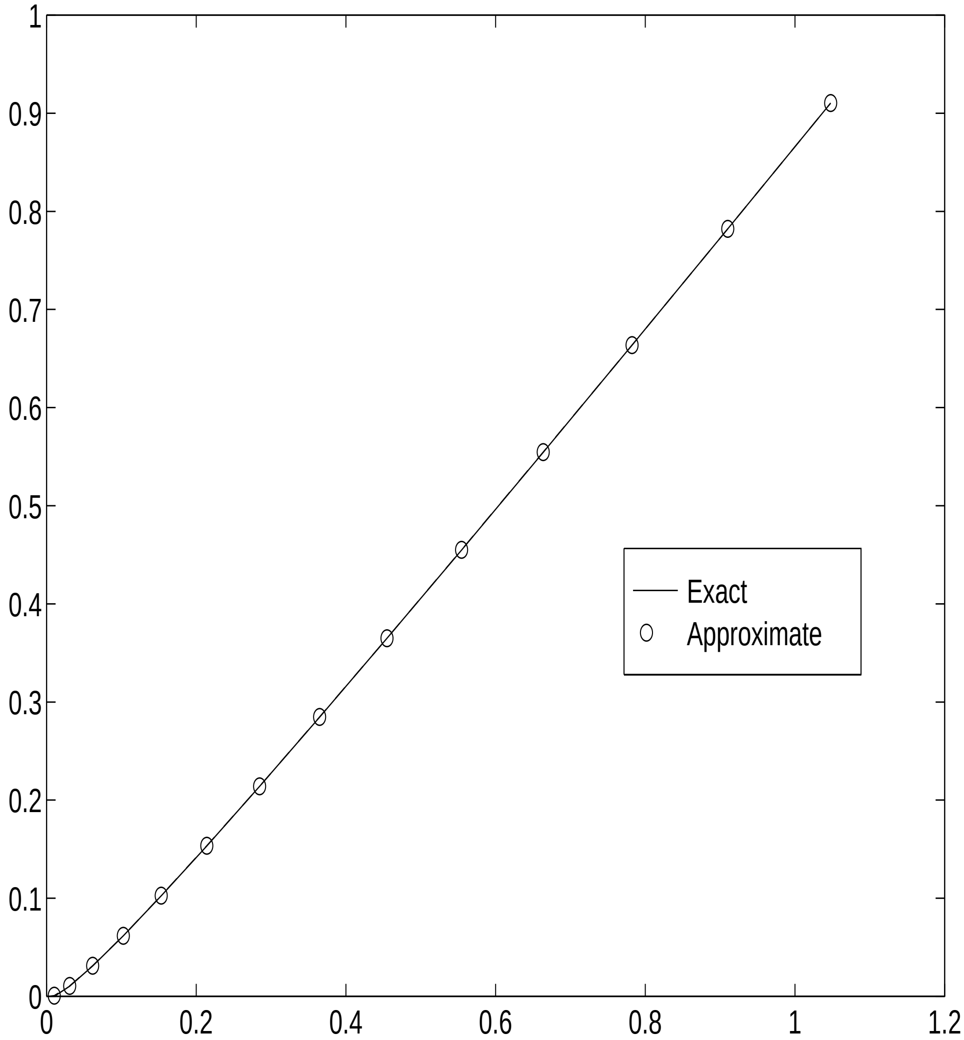

Figure 1a. Decision functions for the test case with correct discretization, no boundary conditions imposed, and $n = 15$.

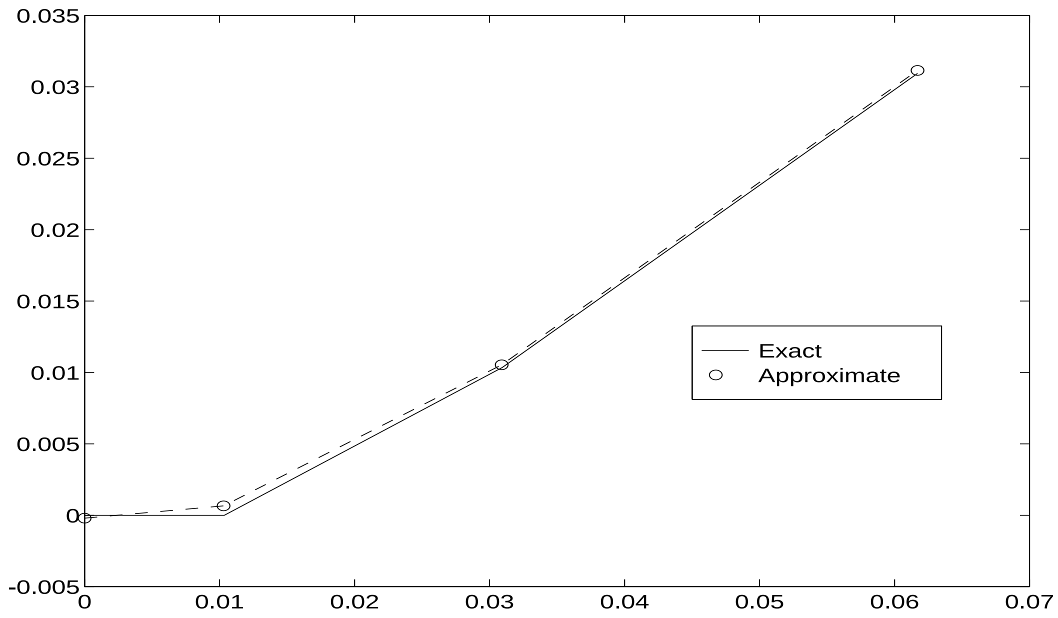

Figure 1b. Decision functions for the test case with correct discretization, no boundary conditions imposed, and $n = 15$. (Only $x \in [0, .07]$ shown.)

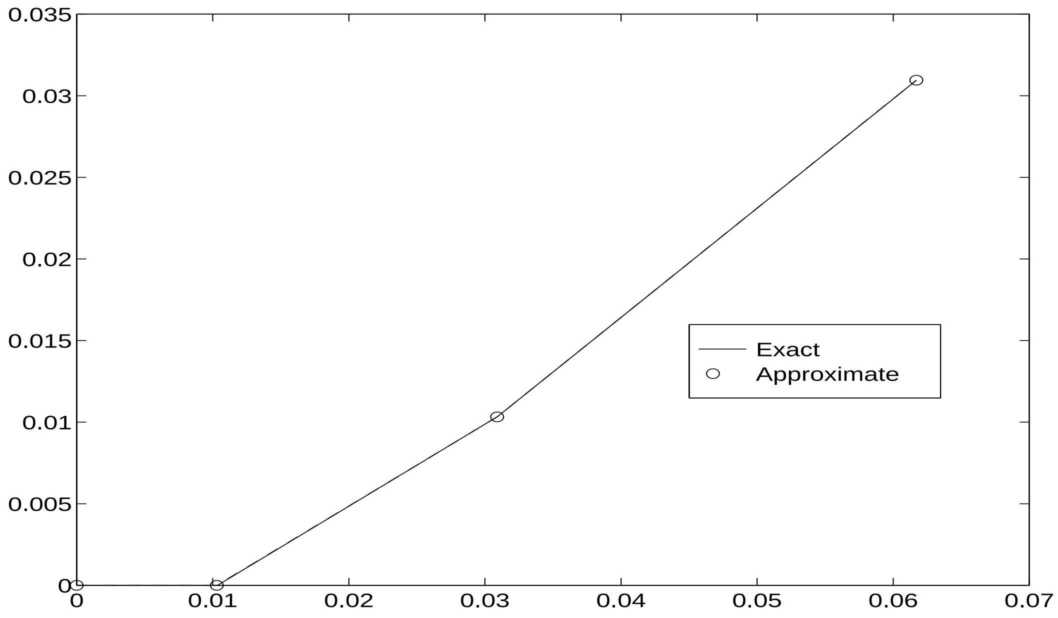

Figure 1c. Decision functions for the test case with correct discretization, boundary conditions imposed, and $n = 15$. (Only $x \in [0, .07]$ shown.)

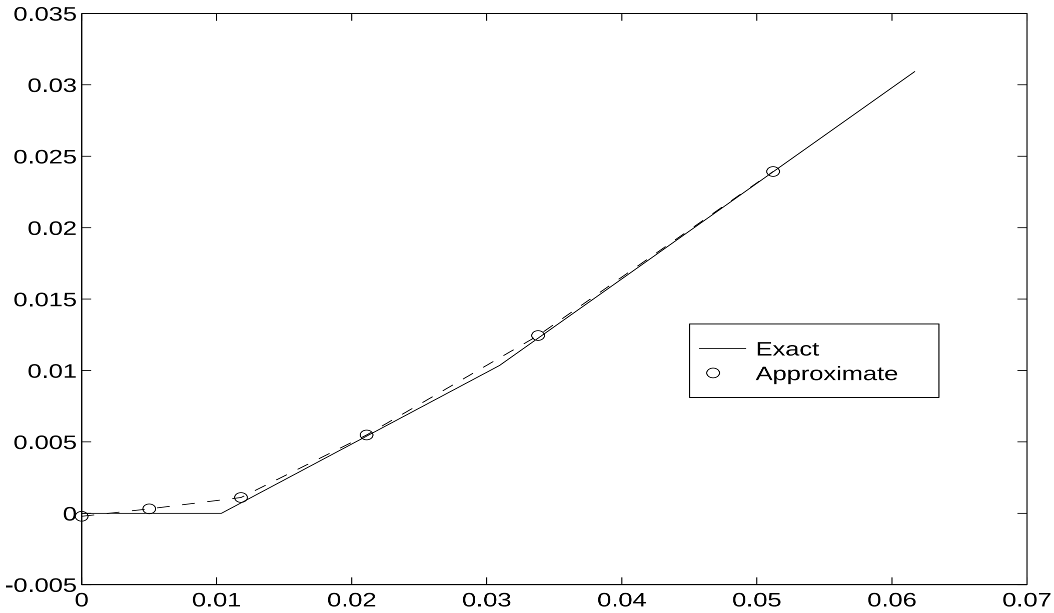

Figure 2a. Decision functions for the test case without correct discretization, without boundary conditions imposed, and $n = 15$. (Only $x \in [0, .07]$ shown.)

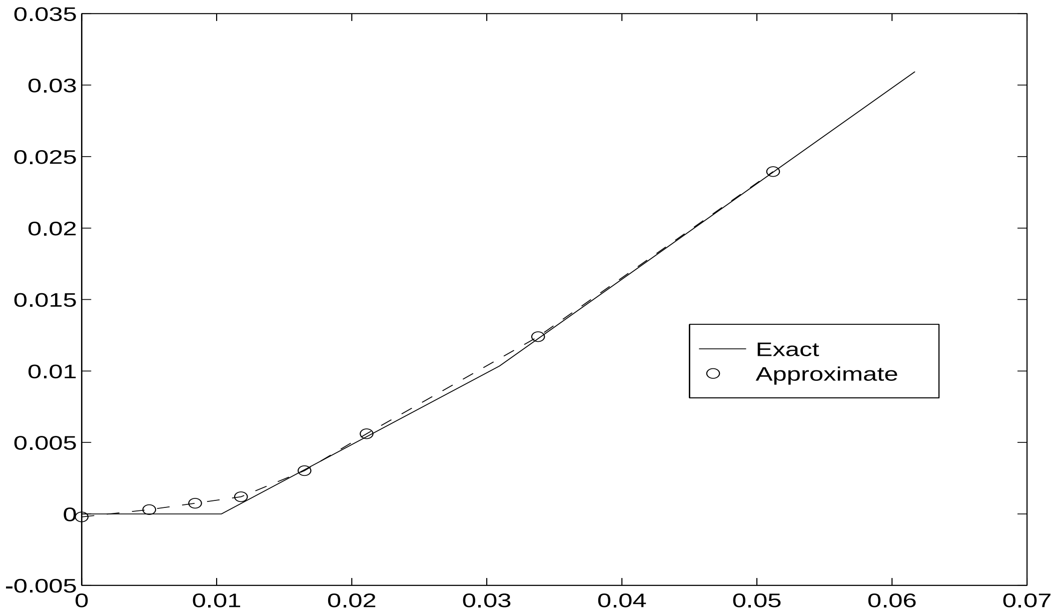

Figure 2b. Decision functions for the test case without correct discretization, without boundary conditions imposed, and $n = 17$. (Only $x \in [0, .07]$ shown.)

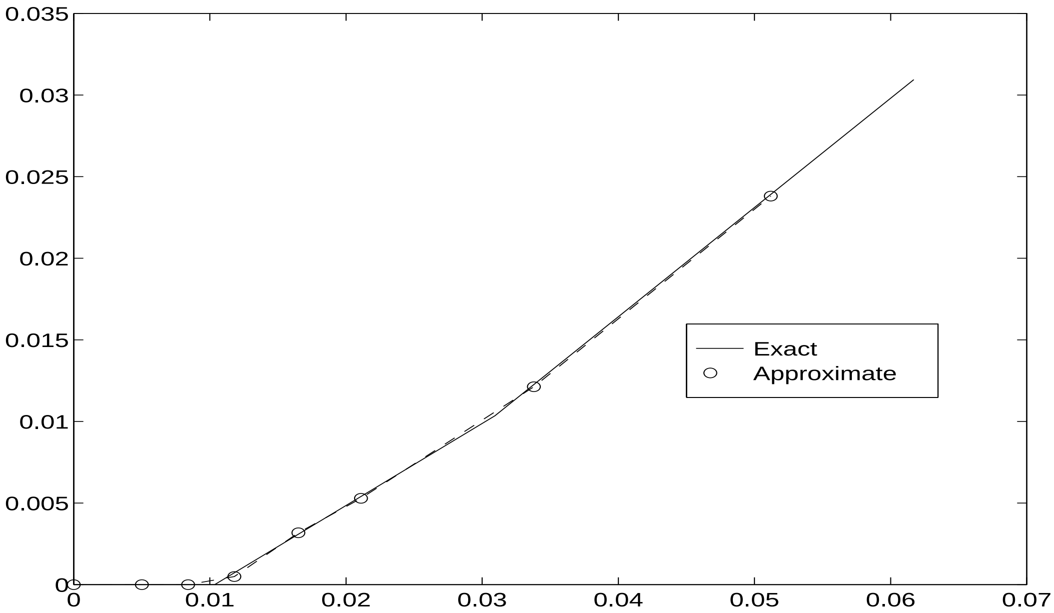

Figure 2c. Decision functions for the test case without correct discretization, with boundary conditions imposed, and $n = 17$. (Only $x \in [0, .07]$ shown.)

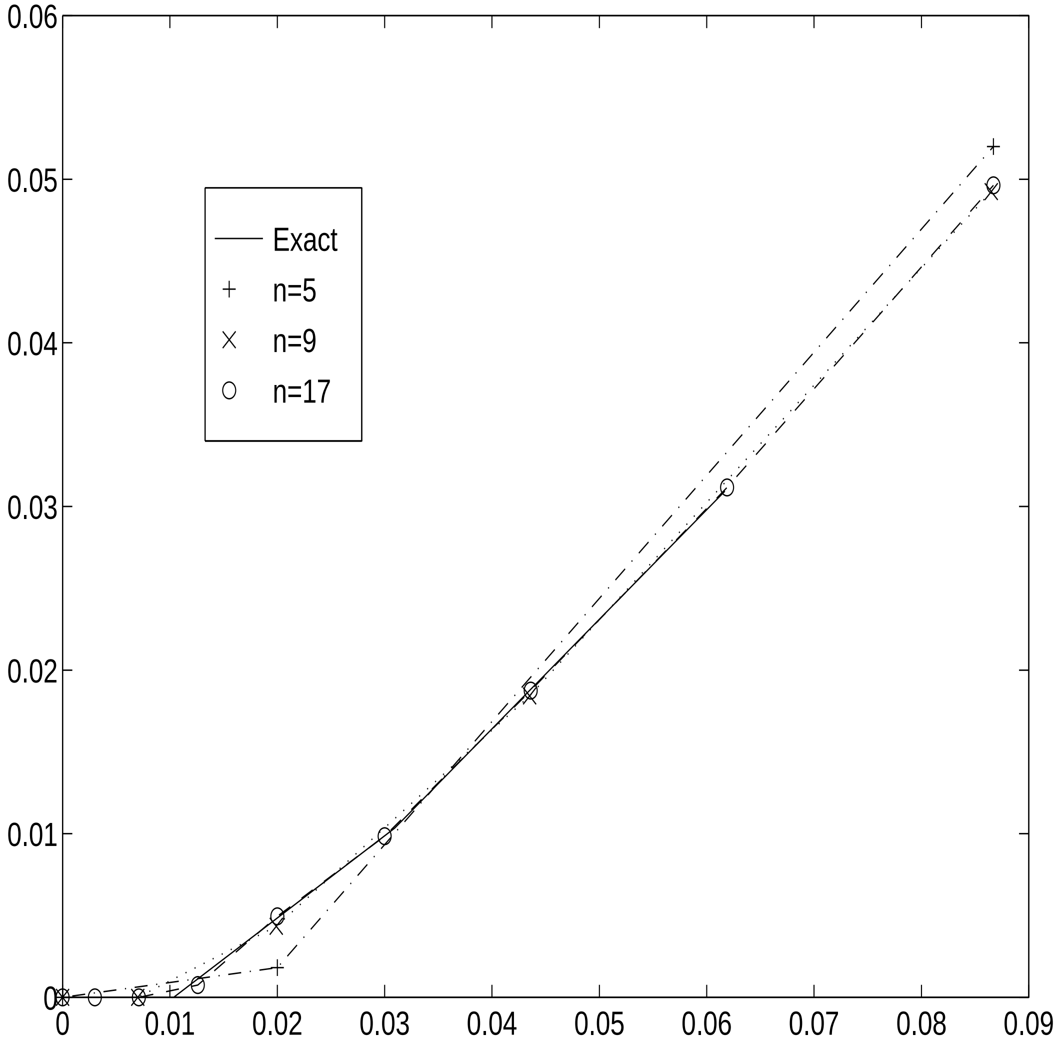

Figure 3. Decision functions for the test case without correct discretization, with boundary conditions imposed, and various n. (Only $x \in [0, .09]$ shown.)

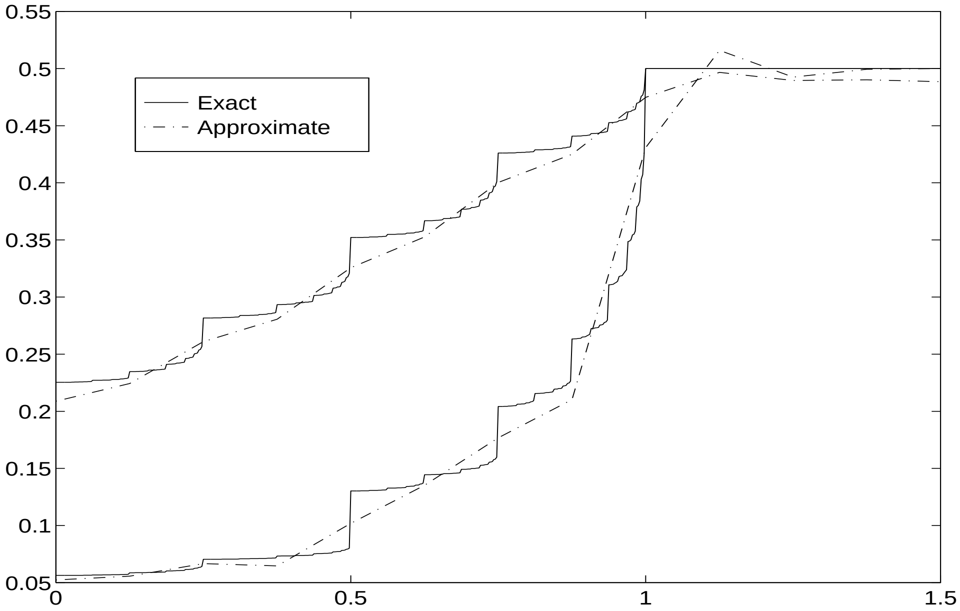

Figure 4a. Distribution for the test case with evenly spaced mesh and $n = 13$.

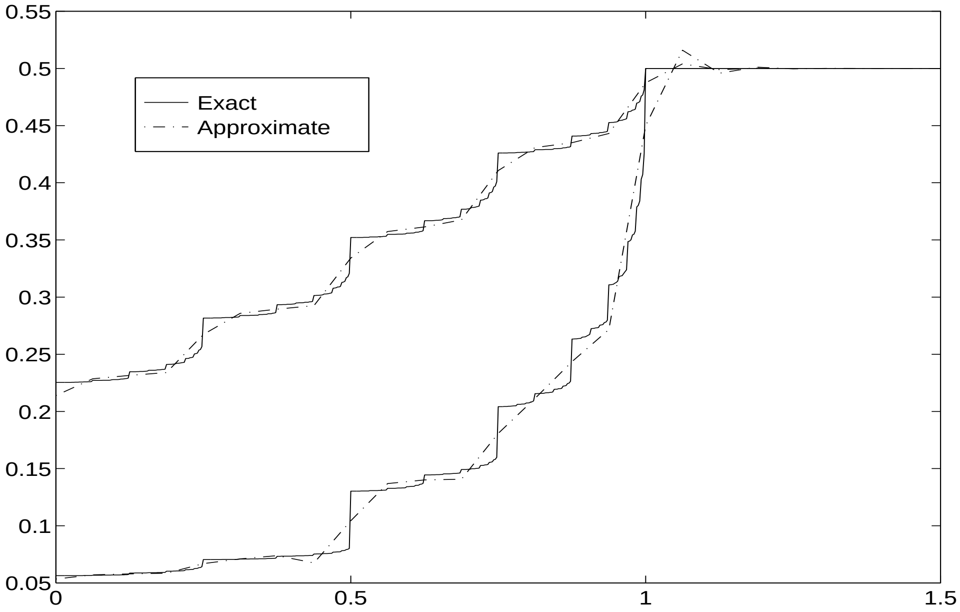

Figure 4b. Distribution for the test case with evenly spaced mesh and $n = 25$.

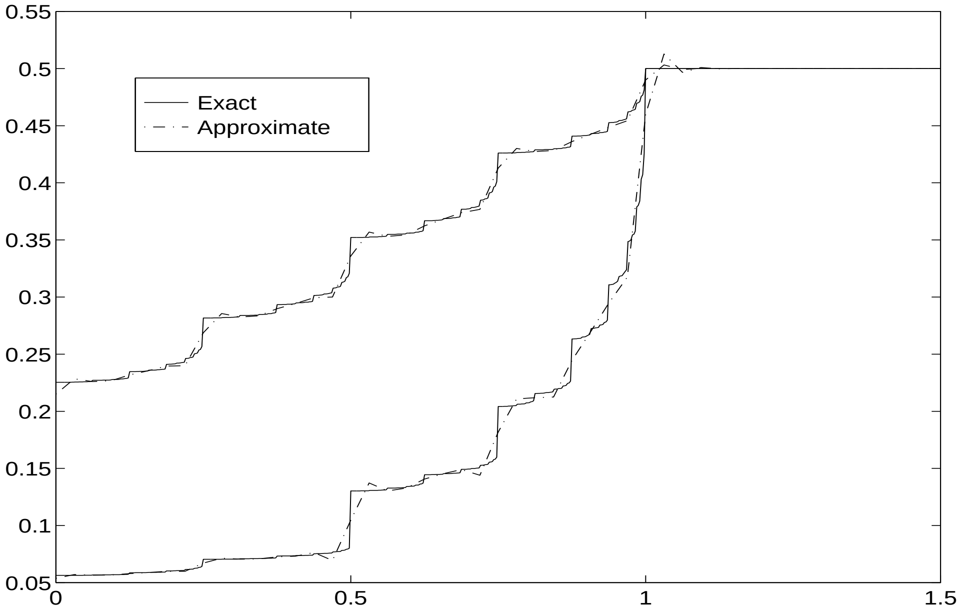

Figure 4c. Distribution for the test case with evenly spaced mesh and $n = 49$.

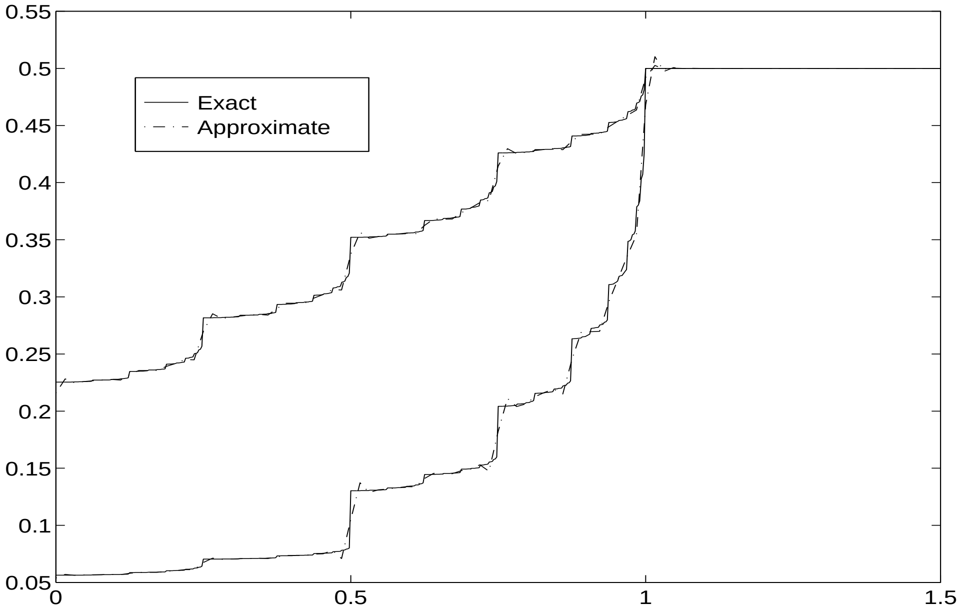

Figure 4d. Distribution for the test case with evenly spaced mesh and $n = 97$.

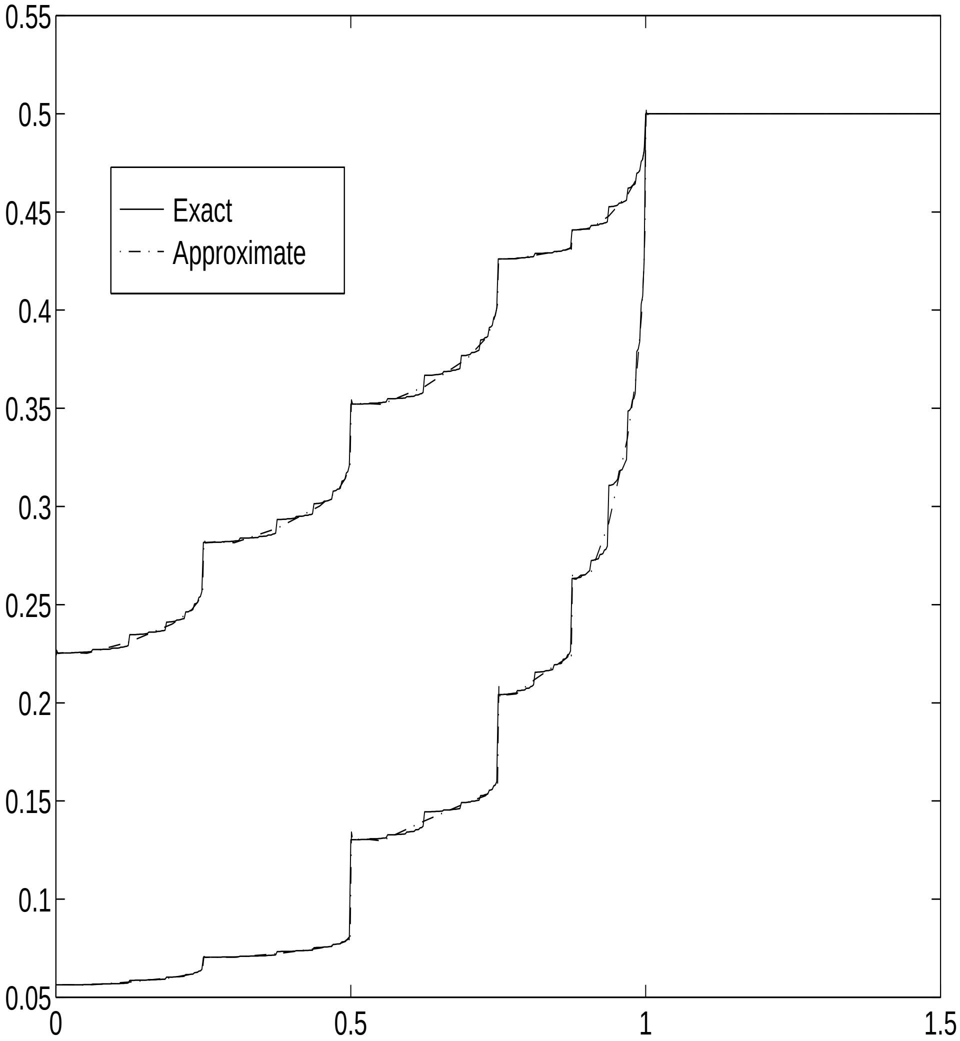

Figure 4e. Distribution for the test case with an adapted grid and $n = 73$.

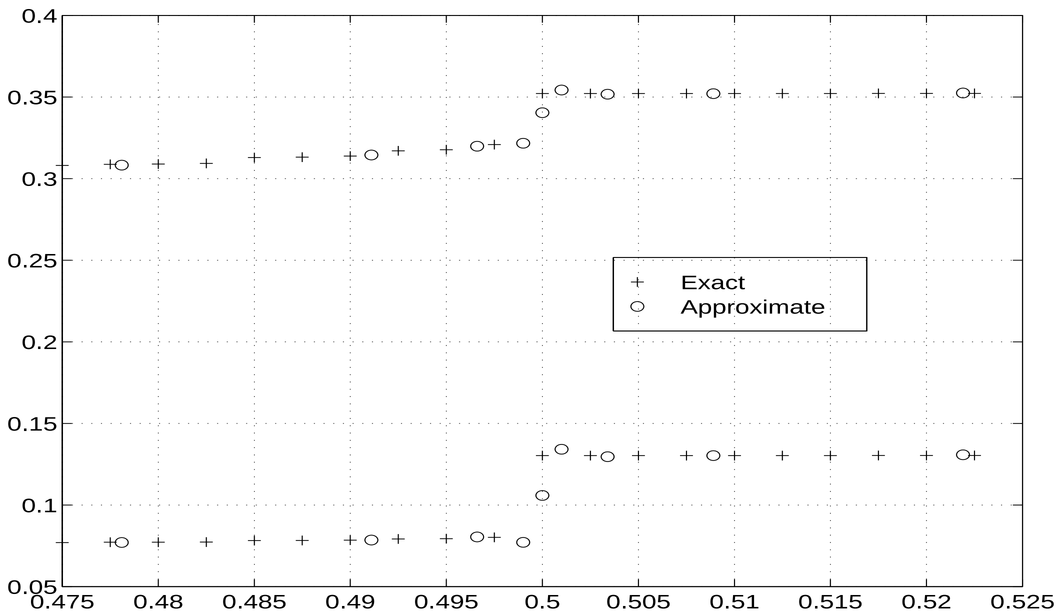

Figure 4f. Distribution for the test case with an adapted grid. (Only $x \in [0.475, 0.525]$ shown.)

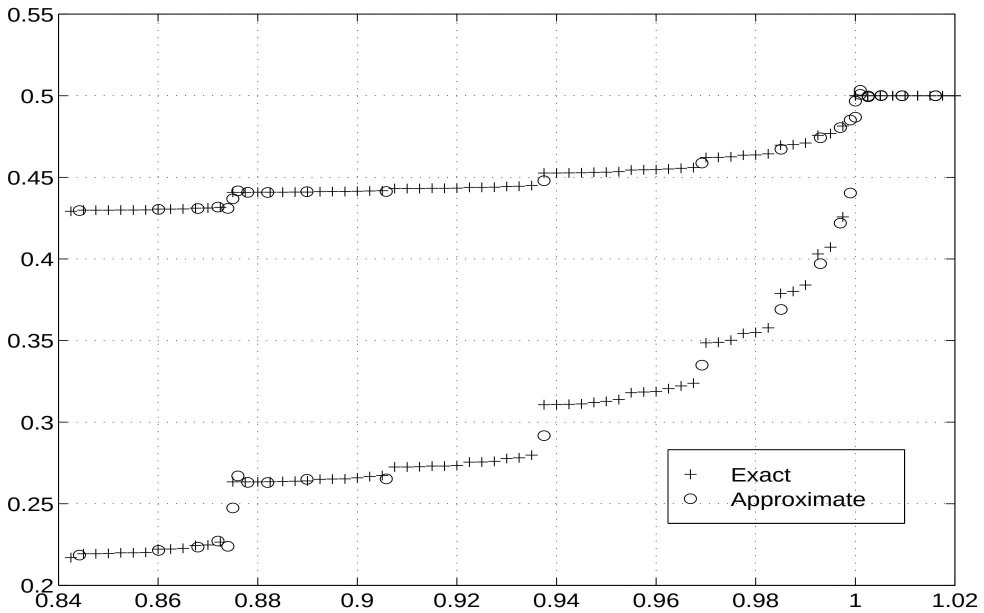

Figure 4g. Distribution for the test case with an adapted grid. (Only $x \in [0.84, 1.02]$ shown.)

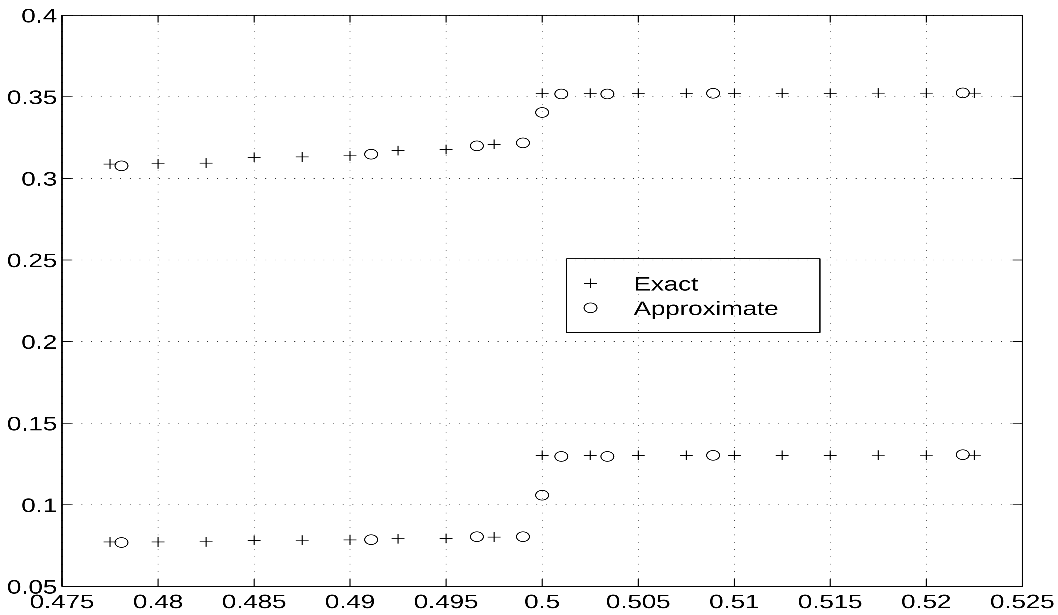

Figure 4h. Distribution for the test case with an adapted grid and nonmonotonicities eliminated. (Only $x \in [0.475, 0.525]$ shown.)

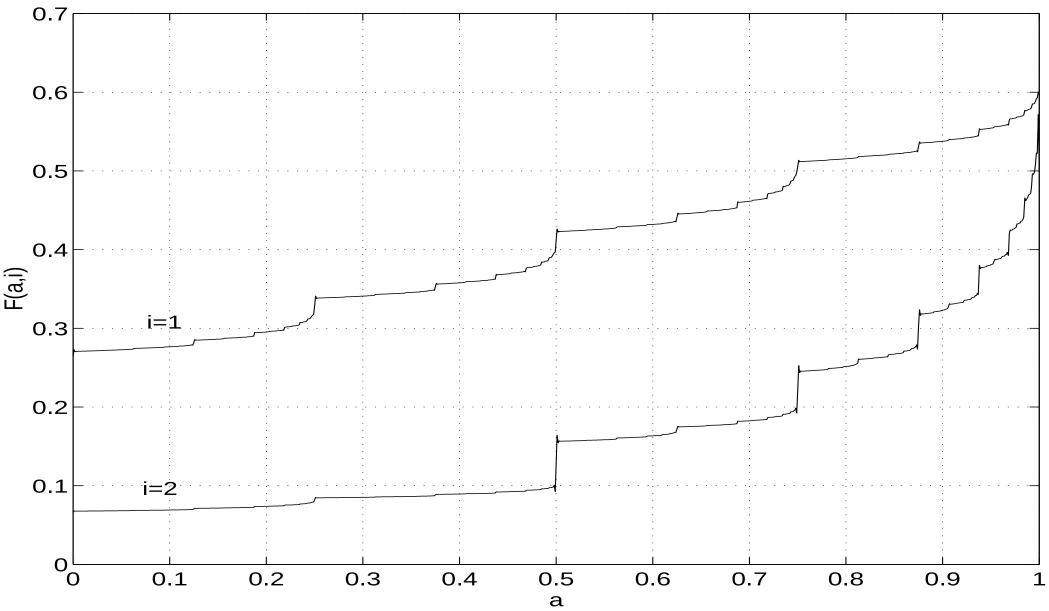

Figure 5. Inaccurate solution for the distribution.