Investment under Risk with Discrete and Continuous Assets: Solution and Estimation

Chris Elbers, Jan Willem Gunning and Melinda Vigh
VU University Amsterdam and Tinbergen Institute

May 11, 2010

Abstract

This paper investigates the savings and investment decisions of rural households who face risks but cannot borrow. Although indivisibilities loom large in such decisions, they are usually not modeled. This paper considers a general class of stochastic dynamic choice models with discrete and continuous decision variables. We propose a method based on value function iteration for solving and estimating this class of models. We use an example to show the importance of modeling indivisibilities: welfare effects of introducing insurance in the correct model with indivisibilities are very different from the effects estimated with a simplified, continuous representation.

Keywords: value function iteration, mixed continuous/discrete controls, stochastic dynamic choice model
JEL classifications: C61, C63, C51, E12, G11, Q12

1 Introduction

The vast majority of the world's poor are to be found in rural areas. This has inspired a large literature on the savings and investment decisions of rural households. In most models in this literature households face risks and cannot borrow. Also, households are usually assumed to have access to a safe asset. A well known example of such a model is Deaton (1991). In most models assets are treated as continuous. There is a long standing interest in relaxing one or both of these two key assumptions.

It is obviously important to investigate the consequences of relaxing the assumption that investment can be modeled as a continuous variable since many agricultural investments, for example in livestock, irrigation pumps or in cash crops with substantial entry costs, are characterized by indivisibilities. Rosenzweig and Wolpin (1993) modeled this for Indian households. Each of the three assets they distinguished (pumps, bullocks and calves) could take only a limited number of discrete values. Their model does not have a continuous asset. This severely (and clearly unrealistically) constrains intertemporal adjustment: households cannot save small amounts. In addition, since investment can take only a small number of values, such models are often numerically not well identified (Elbers *et al.*, 2009).[1]

Other papers have relaxed the safe asset assumption. Dercon (2002) used simulation experiments to show that the effectiveness of consumption smoothing (the only form of

[1] Rosenzweig and Wolpin therefore fixed the value of one of the parameters.

risk management available in the Deaton model) is reduced if the return to savings is itself subject to risk. Gunning (2008) used a two-period model to show that the effect of risk on savings can change sign if risk affects assets rather than (labor) income. Hence, while being exposed to risk necessarily induces more (precautionary) savings in the Deaton model, the effect may well be negative if there is no safe asset. Zimmerman and Carter (2003) find that exposure to risk can act as a poverty trap by inhibiting investment. Elbers *et al.* (2007) and Pan (2008) estimate Ramsey models with both asset and income risk for Zimbabwe and Ethiopia respectively. They find large negative *ex ante* effects of risk[2] on livestock accumulation. In all these papers investment is modeled as a continuous variable, as in Deaton (1991).

A few papers have attempted to allow for both asset risk and indivisibilities. Fafchamps and Pender (1997) specify a model with this conjunction and present the corresponding Bellman equations. However, they do not solve this model. Their estimating equation does not follow from the Bellman equations but is a simplified representation of the model. Dercon (1998) presents a model with livestock indivisibilities and asset risk but also does not solve the Bellman equation.[3] Vigh (2008) estimates a Ramsey model with both discrete (oxen) and continuous (sheep) assets. While Elbers *et al.* (2007) and Pan (2008) reported a strong negative *ex ante* effect of risk on investment using models with only continuous assets, Vigh shows that under indivisibility the effect can be positive. However, the policy functions she uses are only first-order approximations to the true policy functions.

In this paper we investigate how a model with indivisibilities can be solved and estimated and whether it is necessary to take such indivisibilities into account. In section 2 we define a class of models with both types of risk and both continuous and discrete assets. In section 3 we show how these models can be solved, and how they can be estimated using simulation. In section 4 we use a simple example to investigate the importance of indivisibilities. The results suggest that a continuous specification can be quite misleading: notably, for very poor households actuarially fair insurance has much larger effects on welfare in the correct model (with indivisibilities) than in a simplified, continuous representation. Section 5 concludes.

2 The model

We start by defining a class of economic models where the agent decides on investing in a lumpy asset under risk. This class can be written using a stochastic dynamic choice model with the following structural assumptions: (i) the agent maximizes expected discounted utility, (ii) he chooses the values d and x of two assets where d is discrete and x is continuous, (iii) the agent decides on his asset holdings knowing the realization of shocks in the current period, s_t, but before shocks $s_{t+1}$are realized, (iv) he knows the distribution of all future shocks. The problem of the agent in period τ can be written as

$$V_\tau(F(x_{\tau-1}, d_{\tau-1}, s_\tau)) = \max E_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau} u(c_t). \tag{1}$$

with

$$c_t = F(x_{t-1}, d_{t-1}, s_t) - g(x_t, d_t) \tag{2}$$

[2]The *ex ante* effect of risk measures how the propensity to invest (as a function of current wealth) is affected by risk.

[3]Dercon (1998), Appendix A.

and satisfying

$$\begin{array}{rcll} c_t & \geq & 0 & (3) \\ d_t & \in & \mathbb{D} = \{0, 1, ..., D\} & (4) \\ x_t & \in & \mathbb{X} \subseteq \mathbb{R} & (5) \\ d_t, x_t & & \text{are measurable w.r.t. the event space} & (6) \\ & & \text{generated by} \quad \{F(x_{t-i}, d_{t-i}, s_{t-i+1}) | i = 1, ..., t - \tau + 1\} & \end{array}$$

for $t = \tau, \tau + 1, ...,$ given $d_{\tau-1}, x_{\tau-1}$.

Here d_t denotes a discrete asset, x_t a continuous asset and s_t the shock variables. $\mathbb{X}$ specifies all possible values of the continuous asset.[4] Function $F(\cdot)$ determines wealth at hand. Function $g(\cdot)$ specifies the cost of investment in d and x in terms of the consumption good. We assume that the shocks are distributed i.i.d., so that the maximization problem is stationary. As a consequence, the value function $V_\tau(w)$ is time-independent, and in the following it is denoted by $V(w)$.

The maximization problem can be reformulated as a dynamic programming problem

$$\begin{array}{rcll} V(F(x_{t-1}, d_{t-1}, s_t)) & = & \max\limits_{x_t, d_t} u(F(x_{t-1}, d_{t-1}, s_t) - g(x_t, d_t)) + \beta E_t V(F(x_t, d_t, s_{t+1})) & \\ s.t. & & & (7) \\ & & F(x_{t-1}, d_{t-1}, s_t) \geq g(x_t, d_t) & (8) \\ & & d_t \in \{0, 1, ..., D\} & (9) \\ & & x_t \in \mathbb{X} \subseteq \mathbb{R} & (10) \end{array}$$

where $V(\cdot)$ denotes the value function in the Bellman equation (7). The state variable of the problem is wealth at hand, denoted by $w_t \equiv F(x_{t-1}, d_{t-1}, s_t)$. Note that the curse of dimensionality does not apply here: the dimension of the state space is one, irrespective of the number of controls (x_t, d_t) and shock (s_t) variables. The implicit assumption (reflected in the function $g(x_t, d_t)$) is that the two assets can be bought or sold at fixed prices in terms of consumption. This makes it possible to describe the household's wealth with a single state variable, w_t. Clearly, given the distribution of shocks the optimal values of d_t and x_t will depend only on wealth at hand: $d_t = \psi_d(w_t)$ and $x_t = \psi_x(w_t)$.

Before proceeding to the description of the solution algorithm, we summarize all the assumptions of the model together with the ones implied by the dynamic programming formulation of model (1)-(6):

Assumption 1: Properties of the utility function: $u'(\cdot) > 0$ and $u''(\cdot) < 0$.

Assumption 2: Properties of the wealth function: $F_x(\cdot) > 0$ and $F_s(\cdot) > 0$, where $F_v(\cdot)$ denotes the derivative of $F(\cdot)$ w.r.t. variable v. Further, $F(x, d_1, s) - F(x, d_0, s) > 0$ for all $d_1 > d_0 \geq 0$.

Assumption 3: The two assets x and d can be traded at given constant prices, identical for buying and selling.

Assumption 4: Shocks do not affect the cost of investment.

[4] $\mathbb{X}$ can be defined flexibly. For example, it can be applied to an entry decision model with continuous investment levels in x. In this model, d is the binary entry decision variable such that $d = 1$ if entry and $d = 0$ otherwise. Then, $x = 0$ if $d = 0$ and $x \geq X^{min}$ if $d = 1$, which implies $\mathbb{X} = \{0\} \bigcup \{x \in \mathbb{R} | x \geq X^{min}\}$ with $X^{min} > 0$.

Assumption 5: Shocks are i.i.d.[5]

Assumption 6: There is a finite number $D+1$ of possible values for the discrete asset.

Assumption 7: The continuous asset is bounded from below.

If a dynamic choice problem can be written in the form of (1)-(6) and Assumptions 1-7 are satisfied then the solution method of section 3 can be used to solve the problem. This algorithm works satisfactorily if there is at least one discrete control in the problem at hand. In case of problems with only continuous controls, Pan (2008) shows for the stochastic Ramsey model that a solution method using policy function iteration outperforms the method with value function iteration.

The formulation (1)-(6) is more general than that of Fafchamps and Pender (1997) because it allows a decision on the discrete asset holdings in every period, irrespective of the choices before. It is straightforward to extend the number of assets in x and d. However, an expansion of the decision space will increase the computational burden of approximating the policy function.

3 A solution algorithm using value function iteration

In this section, we describe a solution algorithm for problem (7)-(10). The algorithm is based on value function iteration and it approximates the solution of the dynamic programming problem in a non-parametric way.[6] Since the problem is recursive we suppress the time subscript and use the notation $x = x_t$, $d = d_t$, $s^+ = s_{t+1}$, $w = F(x_{t-1}, d_{t-1}, s_t)$ and $w^+ = F(x, d, s^+)$ The problem (7) can then be written as

$$V(w) = \max_{d,x} u(w - g(x,d)) + \beta EV(w^+). \tag{11}$$

3.1 Discretization of the state space and integration

To approximate the expectation in equation (11) we discretize the state variable, $w \in \mathbb{R}$, as $w \in \{w_0, w_1, ..., w_J\}$ for a sufficiently large J. We denote by $p(w_j^+|x,d)$ the probability that next period's wealth at hand will be w_j^+ given x and d. Then,

$$V(w_j) = \max_{d,x} u(w_j - g(x,d)) + \beta \sum_{k=0}^{K} p(w_k^+|x,d) V(w_k^+). \tag{12}$$

The evaluation of $V(w_j)$ involves the calculation of the conditional probabilities $p(w_k^+|x,d)$ for all w_k^+ during the maximization procedure w.r.t. d and x. Alternatively, the expectation of next period's value function can be approximated using a discretization of the shock variable, s. Take M discrete realizations of the income shock, s_m, and the corresponding probability weights, θ_m^s, for $m = 1, ..., M$. Then,

[5] In case of serially correlated shocks, the state space would have to be expanded by at least s_t, therefore we do not consider this case.

[6] This is in the spirit of Rust (1996). The main difference is that in the Bellman equation (11) we discretize the value function on the left hand side (by choosing grid points for wealth) but not the value function on the right hand side. In addition, the choice variable x is continuous in our method but discretized in Rust's approach.

$$V(w_j) = \max_{d,x} u(w_j - g(x,d)) + \beta \sum_{m=1}^{M} \theta_m^s V(F(x,d,s_m)). \quad (13)$$

This is the approach used in this paper.

3.2 Numerical and Monte Carlo integration methods

The shocks and corresponding weights in (13) can be drawn is a number of ways. The most popular of these are Monte Carlo integration and Gaussian quadrature methods. Gauss-Hermite quadrature can be used if the underlying density has a factor of $\exp(-s^2)$. Hence, Gauss-Hermite quadrature can be used in case of normally or log-normally distributed disturbances. However, the approximation will only be accurate if the function in the integral is close to a polynomial in the random variable, because the algorithm chooses nodes and probability weights in such a way that with M number of nodes the integral of a polynomial up to order $2M-1$ can be solved exactly.[7]

In case of Monte Carlo integration, a random sample from the true distribution is drawn with probability weights $1/M$. This algorithm performs well with multi-dimensional integrals (with dimension 3 or larger), however it is not as efficient as numerical integration in case of a single dimension.[8]

As an alternative, we propose a third method that we use fruitfully in the application of Section 4. We take equidistant nodes on the interval $(-4,4)$ for the standard normal disturbances, $(e_1, e_2, ..., e_M)$, and then use these to construct the log-normal shocks as $s_m = \exp(\alpha_0 + \alpha_1 e_m)$, where α_0 and α_1 are chosen such that $\sum_{m=1}^{M} \theta_m^s s_m = 1$ and $\sum_{m=1}^{M} \theta_m^s (\log s_m - \sum_m \theta_m^s \log s_m)^2 = \sigma^2$. Probability weights, θ_m^s, are based on the standard normal density of e_m normalized so as to ensure that $\sum \theta_m^s = 1$. Of the three alternative approaches the third worked best in our particular application.

3.3 Value function iteration algorithm

Before we can start the value function iteration algorithm, we need to first specify the state space, w_j, the shock realizations s_m and their probability weights θ_m^s, and initial guesses for the value and policy functions at each w_j. We initially choose equidistant wealth at hand realizations for w_j on a relevant interval for the problem. An initial approximation of the value and policy functions are necessary so that we can calculate the RHS values of $V(\cdot)$ in (13). The initial guesses of the policy functions are used as starting values in the value function maximization step.

In the iteration procedure we use the approximated value and policy functions from the previous iteration, $V^{r-1}(\cdot)$ and $\psi_x^{r-1}(\cdot), \psi_d^{r-1}(\cdot)$, and their linear interpolation or extrapolation to obtain the value of the value and policy functions at a given wealth level. Therefore, in each iteration we evaluate the value function at wealth levels $\{w_1, w_2, ..., w_J\}$. These values can serve to approximate next period's value function at states $\{w_1, w_2, ..., w_J\}$ in the next iteration. Thus, in each iteration we solve

$$V^r(w_j) = \max_{d_j,x_j} u(w_j - g(x_j,d_j)) + \beta \sum_{m=1}^{M} \theta_m^s \hat{V}^{r-1}(w_m) \quad (14)$$

[7] For more information on the quadrature methods see chapter 7 in Judd (1998) or Chapter 4 in Press *et al.* (1992).

[8] On Monte Carlo integration see e.g. chapter 8 in Judd (1998).

with $w_m = F(x_j, d_j, s_m)$ and $\hat{V}^{r-1}(\cdot)$ the interpolated or extrapolated values of the value function for $j = 0, 1, ..., J$.[9]

The iteration procedure can be summarized in the following steps:

Step 1. Initialization: take J values of w_j, such that $w_1 < w_2 < ... < w_J < \infty$. These are the nodes where the value function will be evaluated. Choose initial values $V(w_j)$ and $\psi_x^0(w_j), \psi_d^0(w_j)$ for each w_j, which give the initial approximation for the value and policy functions, using the equidistant integration method described in section 3.2. This gives the M shocks s_m and the corresponding probability weights θ_m^s.

Step 2. Iteration: at each iteration r a new approximation of $V^r(w_j)$ and $\psi_x^r(w_j), \psi_d^r(w_j)$ is calculated for each value of w_j. In the evaluation, the approximations at the previous iteration are linearly interpolated or extrapolated to account for the approximation of the policy functions at wealth levels not included in w_j. At each iteration the following steps are implemented to find a new value function and optimal investment rule for each w_j:

Step 2.1. For all feasible values of the discrete asset, $\bar{d}_k \in \{d \in \mathbb{D} | g(0, d) \leq w_j\}$, maximize

$$\begin{aligned} V_k^r(w_j) &= \max_{x_k} u(w_j - g(x_k, \bar{d}_k)) + \beta \sum_{m=1}^{M} \theta_m^s \hat{V}^{r-1}(F(x_k, \bar{d}_k, s_m)) \quad (15) \\ s.t. \quad & g(x_k, \bar{d}_k) \leq w_j \end{aligned}$$

and store the optimal values $(V_k^r(w_j), x_k, \bar{d}_k)$.

Step 2.2. Update the value function

$$V^r(w_j) = \max_k V_k^r(w_j)$$

and the policy functions

$$\{\psi_d^r(w_j), \psi_x^r(w_j)\} = \{\bar{d}_{k_j^*}, x_{k_j^*}\}$$

with $k^* = \arg\max_k V_k^r(w_j)$.

Step 3. Convergence: the approximation of the value and policy functions converges to their true values, because $V(w)$ satisfies Blackwell's sufficient conditions.[10] Hence, the contraction mapping theorem[11] applies to the value function iteration. In practice, we stop the iteration procedure when the difference between the values of $(V^{r-1}(w_j), \psi_x^{r-1}(w_j), \psi_d^{r-1}(w_j))$ and $(V^r(w_j), \psi_x^r(w_j), \psi_d^r(w_j))$ becomes very small for all j.

To demonstrate the use of the algorithm, in the next section we take a simple application and show how our solution method can be used to approximate the policy function.

[9] A hat over a function denotes that we use interpolation or extrapolation to evaluate the function at the given value.

[10] Blackwell's sufficient conditions require $V(w)$ to be monotonic and satisfy discounting. Further, $V(w)$ is bounded since $0 < w_j < w_J < \infty$ for all j. See Stokey and Lucas (1989).

[11] See Stokey and Lucas (1989).

4 An example

In this section, we discuss a simple example to demonstrate the solution of a dynamic choice model with both discrete and continuous controls. We describe how the structural parameters of such a model can be estimated, and present results on the distribution of the parameter estimates using Monte Carlo simulation. We use the model to investigate whether the policy implications would change if we replace the discrete asset by a continuous representation.

4.1 The model

Consider a farmer who earns his income from cultivating land. He has an expected yield of $y_0 = 0.5$ units, however its value is affected by weather shocks summarized in s. He can save by storing grain, x. Grain is a continuous, safe asset. In each period he also has the option to rent a pair of oxen, d, which he can use for ploughing. Oxen rental costs him 1 unit in each period; it increases his expected income to $y_1 = 2$ in the next period.[12] Hence, for reasonable discount rates if the farmer can afford to rent the oxen it is beneficial for him to do so, at least in expectation.

The farmer's problem can be formalized in the dynamic programming framework with the following Bellman equation:

$$
\begin{aligned}
V(w) \quad &= \quad \max_{x,d} u(w-x) + \beta EV(x - d + s(y_0 + d(y_1 - y_0))) && (16)\\
s.t. & \\
& w \geq x && (17)\\
& x \geq d && (18)\\
& d \in \{0, 1\} && (19)
\end{aligned}
$$

with $y_0 = 0.5$ and $y_1 = 2$. Let $\beta = 0.9$ and assume CRRA utility $u(c) = (c^{1-\gamma} - 1)/(1-\gamma)$ with $\gamma = 0.95$.[13] The distribution of the shock is $\log s \sim N(-\sigma^2/2, \sigma^2)$. Note that x represents the total holdings of assets, while d denotes the part of assets that is invested in the more productive technology (i.e. the pair of oxen).

For this problem we approximate the optimal savings and investment decision using numerical optimization. However, for the deterministic case ($s = 1$ with probability 1), we can derive the exact solution analytically. This is done in the next section, so that we can compare the results of the approximated value and policy functions with $\sigma = 0$ to the analytical solution.

4.2 Benchmark: analytical solution of the deterministic case

We first solve the deterministic case of (16)-(19) analytically. It is optimal for the agent to reach and stay at the steady state with $d = 1$ and $x = 1$. We can simplify the problem by noting that once the farmer saves 1 in x, he will invest this money in d. Thus, we can rewrite the problem without d as

[12] Alternatively, we can interpret the problem as follows: the farmer can buy either a pair of oxen or no oxen. His return on owning a pair of oxen is $s(y_1 - y_0) - 1$, hence the risk on the return of oxen is perfectly correlated to income risk.

[13] This is very close to log utility ($\gamma = 1$).

$$
\begin{aligned}
V(w) &= \max_x u(w-x) + \beta V(y_0 + I_{x\geq 1}(y_1 - y_0 - 1) + x) && (20)\\
& s.t. \\
& \quad w \geq x && (21)\\
& \quad x \geq 0. && (22)
\end{aligned}
$$

We have to find the wealth level, $\bar{w}_1$, at which the agent will decide to invest 1 unit in x (and hence in d). At this wealth level the agent will be indifferent between investing 1 unit now or postponing the investment one period.

$$
V^0(\bar{w}_1) \equiv u(\bar{w}_1 - 1) + \beta V(y_1) = \max_x u(\bar{w}_1 - x) + \beta u(y_0 + x - 1) + \beta^2 V(y_1) \equiv V^1(\bar{w}_1). \quad (23)
$$

Using the FOC to solve for x in equation (23) we can find the switching wealth level $\bar{w}_1$ and the savings level just below it, x_1^-. Equation (23) has two solutions but we are only interested in the lower solution, $\bar{w}_1 = 1.093$. The policy function is discontinuous at this point; the limit from below is $x^- = 0.780$ and the limit from above is $x^+ = d = 1$.

Next, we want to find the switch point, $\bar{w}_2$, where the agent is indifferent between investing in $d = 1$ next period and in two periods, and so on. A detailed description of the analytical solution can be found in Appendix A. Panel A of table 1 reports all the switching wealth levels, savings decisions and value function values at these points. Note that an agent with wealth of 0.5 needs to save for three periods before he can invest in d.[14]

The policy function of total savings, x, is shown in Figure 1. The graph is linear between the switching points because on these intervals the Euler equation of the problem is satisfied.[15] Thus, the optimal savings decision is

$$
x = \frac{\beta^{1/\gamma}}{1+\beta^{1/\gamma}} w + \frac{1}{1+\beta^{1/\gamma}} (x^*_{i+1} - y_0) \quad (24)
$$

for wealth levels $w < \bar{w}_1$. Then there is a range of wealth where x remains constant at 1. While for very high values of w

$$
x = \frac{\beta^{1/\gamma}}{1+\beta^{1/\gamma}} w + \frac{1}{1+\beta^{1/\gamma}} (x^*_{i+1} - y_1) \quad (25)
$$

where i stands for the number of periods that are needed to reach $d = 1$ and x^*_{i+1} denotes the optimal savings decision in the next period. Wealth levels above y_1 are higher than at the steady state, therefore the agent wants to consume the excess wealth. The first order condition implies that it is optimal to consume all excess wealth at once if the wealth at hand is not greater than 2.117, and spread out the consumption over more periods according to equation (25) otherwise. However, there is no switch-point at this wealth level, the transition is smooth.

[14] Since each agent receives at least 0.5 unit of income this implies that all agents will be able to use the higher technology, i.e. there is no poverty trap here.

[15] See Appendix A for further details.

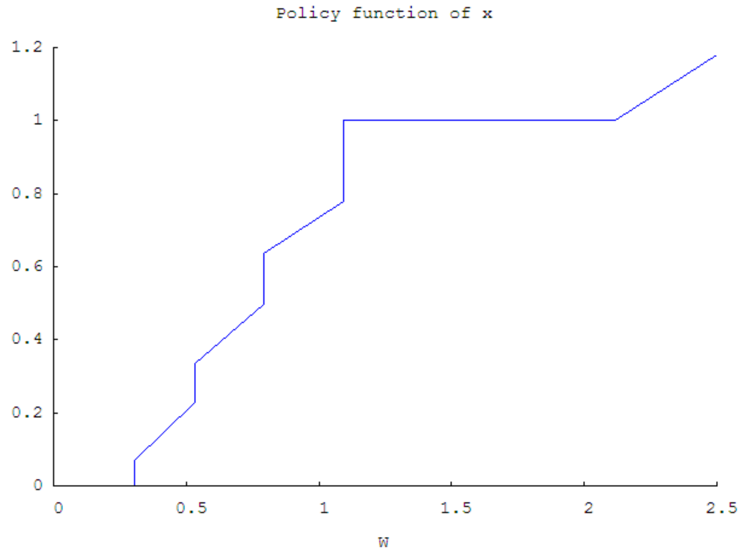


Figure 1: Policy function of the problem in equations (16)-(19).

4.3 Solution method using value function iteration

We solve (16)-(19) using value function iteration. For this model, Step 2.1 of the solution algorithm involves maximizing

$$V_0^r(w_i) = \max_{x_i^0 \geq 0} u(w_i - x_i^0) + \beta \sum_{m=1}^{M} \theta_m^s \bar{V}^{r-1}(x_i^0 + s_m y_0) \tag{26}$$

and

$$V_1^r(w_i) = \max_{x_i^1 \geq 1} u(w_i - x_i^1) + \beta \sum_{m=1}^{M} \theta_m^s \bar{V}^{r-1}(x_i^1 - 1 + s_m y_1) \tag{27}$$

for every wealth level, where it is feasible.

Further details on the implementation of the solution algorithm can be found in Appendix B.

4.4 Simulation results

First, we report the results of the value function iteration for the deterministic case ($\sigma = 0$). The results of the iteration procedure are plotted in Figure 2 and the estimates of the switch points are reported in Panel B of table 1. In the table w_{p^-} and w_{p^+} denote the grid-points where the discontinuities occur. The distance between w_{p^-} and w_{p^+} can be reduced by evaluating the policy function at more nodes. For wealth level w_{p^-}, x_{p^-} is the estimated stock of assets held by the household, while it is x_{p^+} at w_{p^+}. $V(w_{p^-})$ stands for the level of the value function at w_{p^-}. The shape of the obtained policy functions is very similar to the analytical solution. The values of $(w, x^-, x^+, V(w))$ at all switch-points are estimated with a high precision. The true switch-points, $\bar{w}$, always fall in the range found by the value function iteration algorithm if we use the initial-value-search routine described in Appendix B. Without this routine, the switch-points are estimated several grid-points away from the true value except at the wealth level where the investment in d

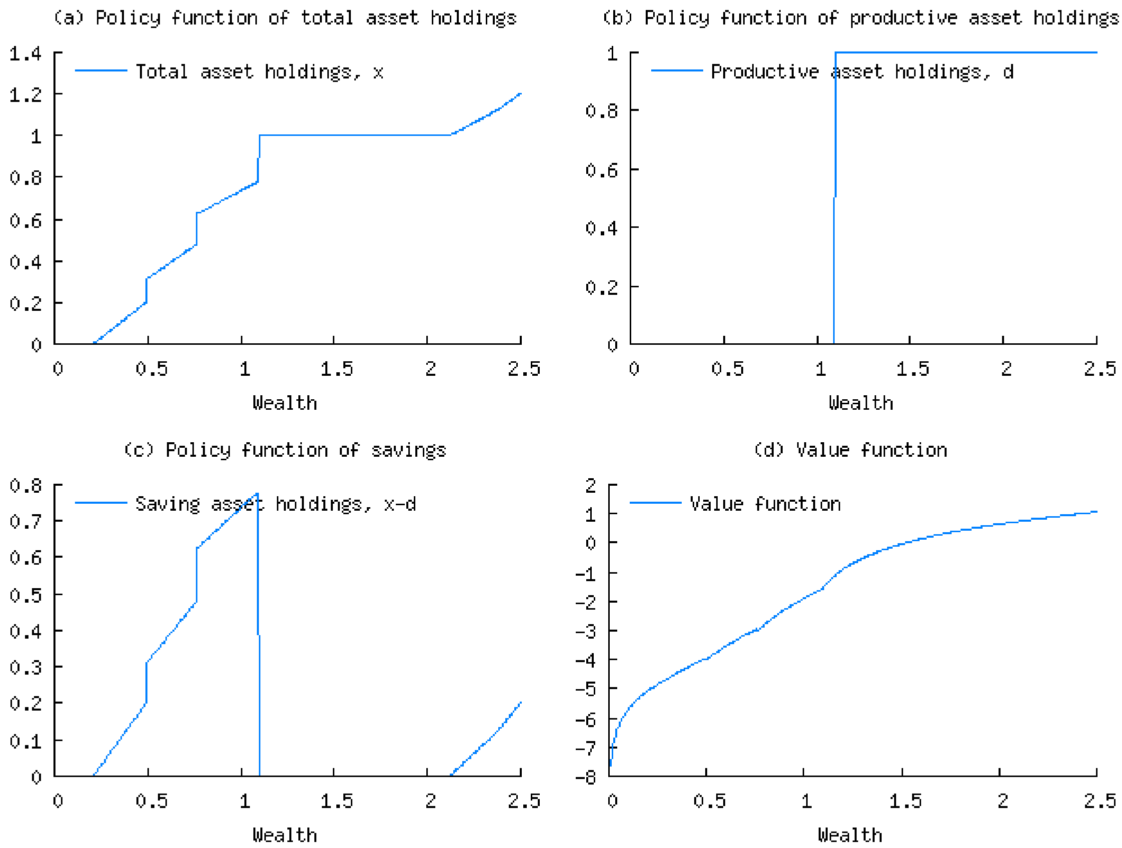


Figure 2: Value function iteration results

is made.[16] The values of x also have the right magnitudes. The values of x_{p^-} are slightly lower than the analytical solution $\bar{x}_p^-$ because the estimated values w_{p^-} are also a little bit smaller than $\bar{w}_p$. The same holds with reverse sign for x_{p^+} and $\bar{x}_p^+$. Notice, however, that at $w = 0.305$ the value function iteration algorithm does not find a switch point. The policy function continues to decline to zero, which it reaches at $w = 0.203$.[17]

Next, we simulated the model for various values of the parameters using as the benchmark case the set $(\beta, \gamma, \sigma, y_0, y_1) = (0.9, 0.95, 0, 0.5, 2)$.[18] Figure 3 shows that the policy function is quite insensitive to changes in γ.

Figure 4 plots the policy and value function for different values of y_0. We observe that household with a low basic income ($y_0 = 0.3$) save more than other households at low wealth levels, so that they build up enough wealth to invest in the higher return activity. For higher values of y_0 the difference between the two activities, $y_1 - y_0$ approaches the cost of switching between the activities. As a result switching becomes less attractive and the investment into the productive technology occurs at higher wealth levels. If the difference in the returns becomes less than 1, the households will (obviously) only save if their income is above y_0 and will never invest in the high return activity.

Figure 5 shows the policy function for increasing levels of risk. Note that as risk

[16] In this case the discontinuity occurs due to a change in d, hence the flatness of the value function w.r.t. x is not an issue here.

[17] We checked the optimality of this solution, and we found that this solution yields the same welfare (discounted utility) to the agent as the solution with the switch point at $w = 0.305$.

[18] Results are available upon request from the authors.

p	Panel A: Analytical				Panel B: Value function iteration				
	$\bar{w}_p$	$\bar{x}_p^-$	$\bar{x}_p^+$	$V(\bar{w}_n)$	w_{p^-}	w_{p^+}	x_{p^-}	x_{p^+}	$V(w_{p^-})$
	2.117^a	1.000	1.000	0.111	2.123^a	2.130^a	1.000	1.007	0.232
0	2.000^b	1.000	1.000	0.000	1.999^b	2.006^b	1.000	1.000	0.029
1	1.093	0.780	1.000	-2.240	1.089	1.096	0.777	1.000	-2.219
2	0.792	0.498	0.638	-3.483	0.789	0.795	0.496	0.639	-3.463
3	0.532	0.230	0.335	-4.481	0.527	0.534	0.227	0.335	-4.465
4	0.305	0.000	0.069	-5.275	0.300^b	0.307^b	0.064	0.071	-5.262
5					0.203^a	0.210^a	0.000	0.002	-5.412

Notes: a. Change in slope, no jump
b. No change in slope, no jump

Table 1: Comparing switch-points of the analytical solution to the those using value function iteration in the deterministic case with $\beta = 0.9$, $\gamma = 0.95$, $y_0 = 0.5$ and $y_1 = 2$

increases, the discrete step in the policy function disappear. At the end, we only have one discontinuity in the policy function when investing in the high return technology. Also, the households build up precautionary savings when their wealth level increases above y_1 due to a positive shock. In this case, savings are the only way to cope with income risk and prevent the household form dropping back to the basic technology.

4.5 Estimation

We now turn to the estimation of the structural parameters of the model. We develop an estimator for the following situation: the sample consists of N identical households; we observe their asset holdings x and d for two consecutive periods ($\bar{x}$ and $\bar{d}$ for the earlier period and x and d for the following period). The log-likelihood function can be written as

$$\begin{aligned} \ell(\mathbf{x}, \mathbf{d}|\theta) &= \sum_{i=1}^{N} \log p\left(x = x_i, d = d_i | \bar{x}_i, \bar{d}_i, \theta\right) \\ &= \sum_{i=1}^{N} \log p(x = x_i | \bar{x}_i, \bar{d}_i, \theta) + \log p(d = d_i | x_i, \bar{x}_i, \bar{d}_i, \theta) \end{aligned} \tag{28}$$

where $p(x = x_i, d = d_i|\bar{x}_i, \bar{d}_i, \theta)$ is the joint probability of observing $x = x_i$ and $d = d_i$, given previous observations $\bar{x}_i$ and $\bar{d}_i$ and parameters $\theta = (\beta, \gamma, \sigma, y_0, y_1)$. To simplify notation, we introduce $p_\theta(x = x_i, d = d_i) \equiv p(x = x_i, d = d_i|\bar{x}_i, \bar{d}_i, \theta)$. Then, $p_\theta(x = x_i)$ is the conditional probability of $x = x_i$ given the past observations and parameter values. Similarly, $p_\theta(d = 0)$ denotes the probability that $d_i = 0$ given the previous observations and parameter values. Note that d_i can only take values 0 and 1, therefore, $p_\theta(d = 1) = 1 - p_\theta(d = 0)$.

First, we derive $P_\theta(x < x_i) = \int_{-\infty}^{x_i} p_\theta(x = x_s)dx_s$. Recall that we denote the policy function for the total assets by $\psi_x(w)$. $\Phi(\cdot)$ denotes the standard normal cumulative distribution function $\phi(\cdot)$ is the corresponding density. Then,

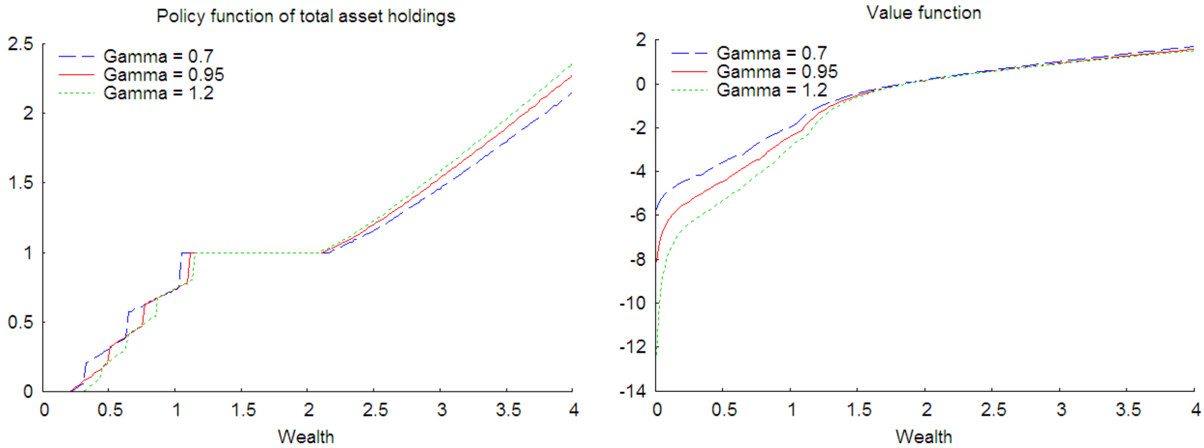


Figure 3: Policy and value functions for different values of γ

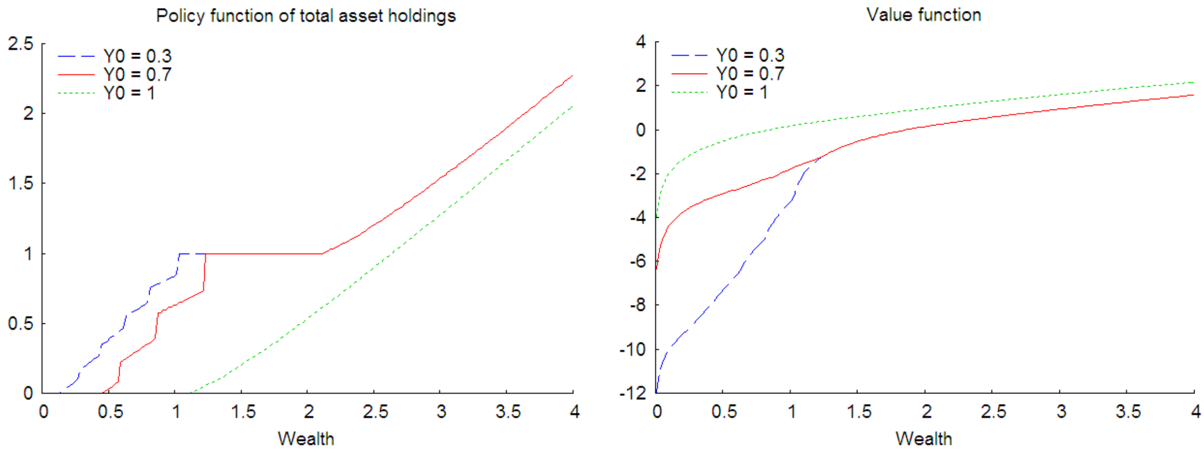


Figure 4: Policy and value functions for different values of y_0

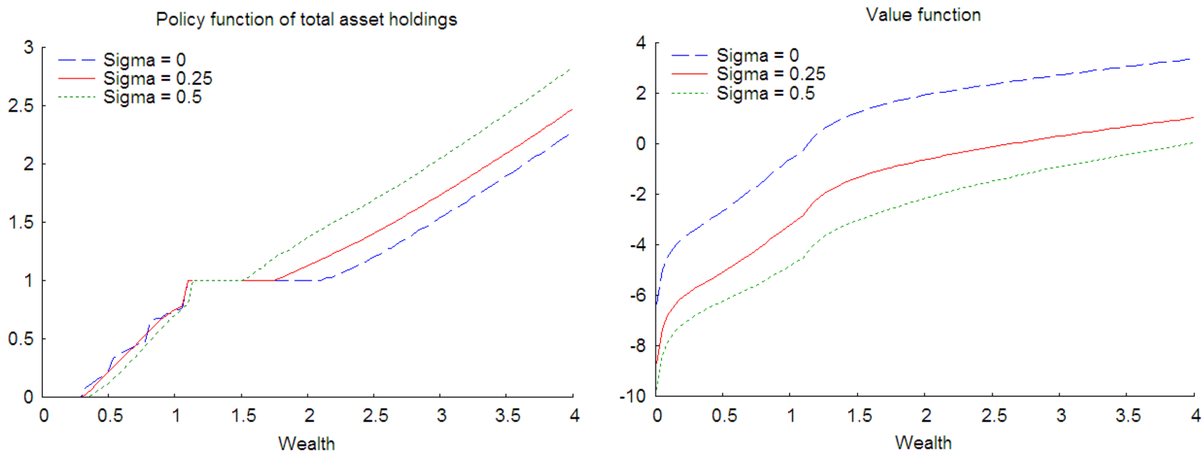


Figure 5: Policy and value functions for different values of σ

$$
\begin{aligned}
P_\theta(x < x_i) &= P_\theta\left(\psi_x(\bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0))) < x_i\right) && (29)\\
&= P_\theta\left(\bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0)) < \psi_x^{-1}(x_i)\right) && (30)\\
&= P_\theta\left(\log s < \log\left(\frac{\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i}{y_0 + \bar{d}_i(y_1 - y_0)}\right)\right) && (31)\\
&= \Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right) && (32)
\end{aligned}
$$

Now, the density of $x = x_i$ can be calculated for $x_i \notin \{0, 1\}$ as

$$
\begin{aligned}
p_\theta(x = x_i) &= \frac{d}{dx_i} P_\theta(x < x_i)\\
&= \frac{\left(\psi_x^{-1}\right)'(x_i)}{\sigma(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i)} \cdot\\
&\quad\ \phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right) && (33)
\end{aligned}
$$

The inverse of the policy function is not well-defined at $x_i = 0$ and $x_i = 1$. For these observations we have to use the appropriate probabilities:

$$
\begin{aligned}
P_\theta(x = 0) &= P_\theta\left(\bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0)) < \bar{\psi}_x^{-1}(0)\right) && (34)\\
&= P_\theta\left(\bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0)) < \bar{\psi}_x^{-1}(0)\right) && (35)\\
&= \Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\bar{\psi}_x^{-1}(0) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right) && (36)
\end{aligned}
$$

and

$$
\begin{aligned}
P_\theta(x = 1) &= P_\theta\left(\underline{\psi}_x^{-1}(1) \le \bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0)) < \bar{\psi}_x^{-1}(1)\right) && (37)\\
&= P_\theta\left(\bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0)) < \bar{\psi}_x^{-1}(1)\right)\\
&\quad - P_\theta\left(\bar{x}_i - \bar{d}_i + s(y_0 + \bar{d}_i(y_1 - y_0)) < \underline{\psi}_x^{-1}(1)\right) && (38)\\
&= \Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\bar{\psi}_x^{-1}(1) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\\
&\quad - \Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\underline{\psi}_x^{-1}(1) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right) && (39)
\end{aligned}
$$

where $\underline{\psi}_x^{-1}(\chi)$ stands for the lowest wealth level for which $x_i = \chi$ and $\bar{\psi}_x^{-1}(\chi)$ for the highest wealth level for which $x_i = \chi$. Note that if $\bar{\psi}_x^{-1}(0) < \bar{x}_i - \bar{d}_i$ then $P_\theta(x_i = 0) = 0$ and if $\underline{\psi}_x^{-1}(1) < \bar{x}_i - \bar{d}_i$ then the second term in (39) is 0.

As a last step, we need to calculate the conditional probability of observing $d = d_i$ given the value x_i and the past asset holdings. However, the value of d is determined by the observation of x_i given the policy function because $d = \psi_d(\psi_x^{-1}(x_i))$. Therefore, $p_\theta(d = d_i|x_i)$ is either 0 or 1. The log of 1 is 0, however, the log of 0 does not exist, therefore in the log- likelihood function we need to replace the conditional probability of d by a penalty (P_ε) if the predicted value of the discrete asset is different from the observed value.

Finally, we can write the log-likelihood function as

$$
\begin{aligned}
\ell(\mathbf{x},\mathbf{d}|\theta) \quad = \quad & \sum_{i=1}^{N} I\left(x_i = 1\right) \cdot \\
& \log\left[\Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\bar{\psi}_x^{-1}(1) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\right. \\
& \left. -\Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\underline{\psi}_x^{-1}(1) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\right] \\
+ \quad & \sum_{i=1}^{N} I(x_i = 0)\log\Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\bar{\psi}_x^{-1}(0) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right) \\
+ \quad & \sum_{i=1}^{N} I\left(x_i \notin \{0,1\}\right)\left[\log\left(\left(\psi_x^{-1}\right)'(x_i)\right) - \log(\sigma(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i))\right. \\
& \left. + \log\phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\right] \\
+ \quad & \sum_{i=1}^{N} P_\varepsilon I(d_i \neq \varphi_d(x_i))
\end{aligned} \tag{40}
$$

where $\varphi_d(x_i) \equiv \psi_d(\psi_x^{-1}(x_i))$ is the predicted value of d_i given x_i and $P_\varepsilon < 0$ is the penalty for a wrong prediction.

We use Simulated Annealing[19] in the estimation of the model. This method does not require differentiability of the objective function and it is a global optimization method. For the estimation it is important that the policy function changes monotonically for changes in the parameter values, otherwise the likelihood function will not be smooth. Our value function iteration algorithm satisfies this monotonicity condition.

The estimation results are shown in Table 2. The second column of the table shows the true parameter values used in data generation. The third and fourth columns report the Monte Carlo estimate (MC) of the expectation and standard deviation of $\hat{\theta}$, respectively, based on $R = 200$ different random samples each with $N = 1000$ observations.[20] The sixth column gives the MC standard deviation of the MC expectation of $\hat{\theta}$, which measures the accuracy of the MC estimates in the third column. By increasing R, this number can be made arbitrarily small. We use a small number of replications since the estimation of θ is very time consuming.[21] The fifth column reports the root mean squared error (RMSE) of the parameter estimates.

Note from Figure 6 that the parameter estimates are remarkably precise except for γ. Recall, however, that the policy function is very insensitive to changes in γ, so that this imprecision has no serious consequences.[22]

[19] For details see Goffe *et al.* (1994).

[20] We also calculated the standard error of $\hat{\theta}$ based on the curvature of the log-likelihood function, however it seriously underestimated the true standard error of the parameter estimates (not reported here).

[21] It takes around 4 days to estimate one set of parameters.

[22] The estimate for y_1 is also not very precise. Again, the policy function is not affected as shown by sensitivity analysis (not reported here).

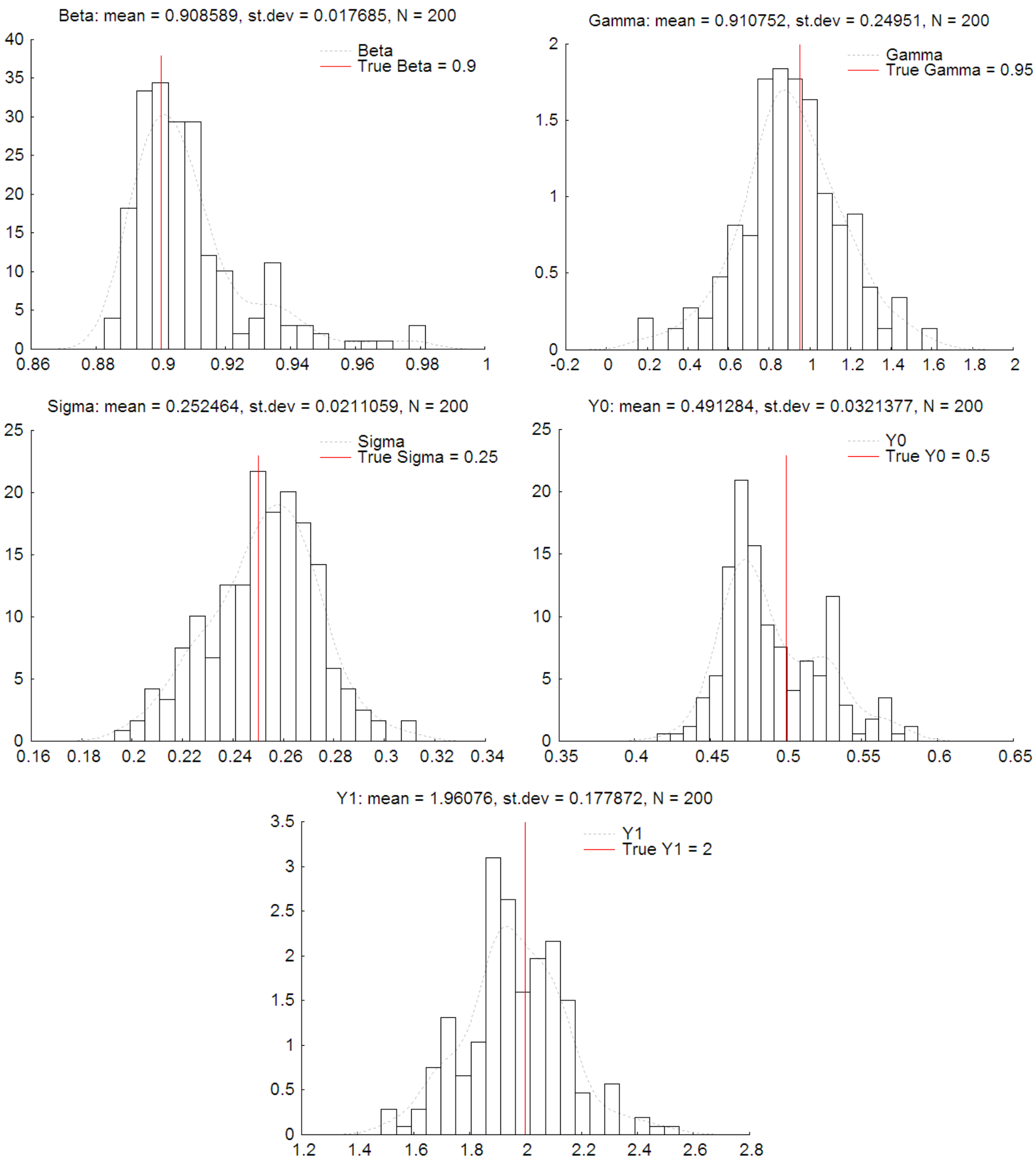


Figure 6: Distribution of Monte Carlo parameter estimates

Parameter	True	Mean	St.dev.	RMSE	MC st.dev.
β	0.90	0.9086	0.0177	0.0197	0.0013
γ	0.95	0.9108	0.2495	0.2526	0.0176
σ	0.25	0.2525	0.0211	0.0212	0.0015
y_0	0.50	0.4913	0.0321	0.0333	0.0023
y_1	2.00	1.9608	0.1779	0.1821	0.0126

Notes:
1. N = 1000, R = 200
2. Mean is calculated as the mean of the parameter estimates
3. St.dev. calculated as the standard deviation of the parameter estimates around their mean
4. RMSE is calculated as the root mean squared error from the true parameter values
5. MCst.dev. calculated as Std.dev/$\sqrt{R}$
6. Parameter estimates calculated using Simulated Annealing with T=100, NS=20, NT=10, RT=0.85

Table 2: Parameter estimates from Monte Carlo simulation

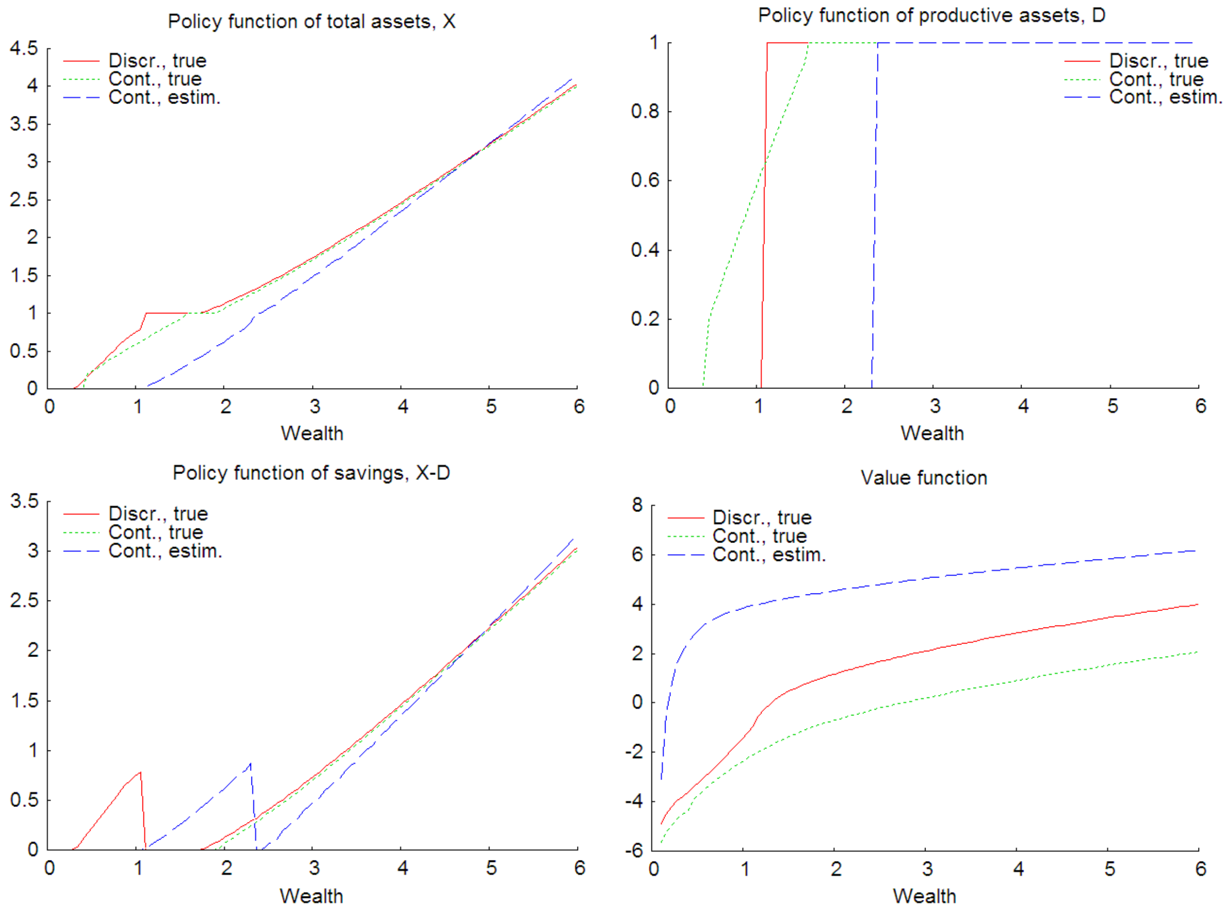


Figure 7: Comparing the policy and value functions of the true model and continuous representations

4.6 Continuous representations

In practice, researchers usually model investment behavior as continuous, even though they are aware of the importance of indivisibilities in rural societies. This raises the question whether policy conclusions are sensitive to the use of continuous or discrete investment models. We investigate this for our example by using the discrete model to generate the data, which are then used by the econometrician to estimate a continuous model.

The continuous model is specified as (16)-(18) and $d \in [0, 1]$, hence we only change the possible values of the productive asset, which are now assumed continuous between 0 and 1. This would be the case if it is possible for a group of households to share a rented pair of oxen.[23]

The estimation results of the incorrect continuous representation are shown in Table 3.[24] The standard deviations of the estimates have greatly increased relative to Table 2. The most notable changes are the very large increases in the mean estimates of γ and y_0. We already know that the policy function is fairly insensitive to changes in γ. It is, however, quite sensitive to changes in y_0. This is illustrated in Figure 7. The figures compare three cases: the true model, the continuous model but using the the correct parameter values

[23] Further details on the continuous model and its estimator can be found in Appendix C.

[24] In Appendix C. we show that the estimator is very precise when the data are generated with the continuous model.

Parameter	True	Mean	Std.dev.	RMSE	MC std.dev.
β	0.90	0.9458	0.0260	0.0527	0.0018
γ	0.95	1.8292	0.6753	1.1086	0.0477
σ	0.25	0.3334	0.1850	0.2029	0.0131
y_0	0.50	1.3061	0.4874	0.9420	0.0345
y_1	2.00	2.3435	0.4806	0.5907	0.0340

Notes:
1. N = 1000, R = 200
2. Mean is calculated as the mean of the parameter estimates
3. Std.dev. calculated as the standard deviation of the parameter estimates around their mean
4. RMSE is calculated as the root mean squared error from the true parameter values
5. MC std.dev. calculated as Std.dev/$\sqrt{R}$
6. Parameter estimates calculated using Simulated Annealing with T=100, NS=30, NT=20, RT=0.85

Table 3: Parameter estimates from Monte Carlo simulation in the misspecified model

(θ_0), and the continuous model using the estimated and therefore incorrect parameter values ($\hat{\theta}$).[25] In interpreting these graphs, one should recall that the econometrician is supposed to have data only on total and productive assets (x and d) but not on current wealth at hand (which includes the value of the current shock). As a result, the policy function estimated under the assumption that all assets are continuous has a shape which quite accurately follows the shape of the true policy function. However, it is displaced to the right: the estimated policy function implies that households invest at higher levels of wealth than in the true model. This reflects an overestimate of the basic income (y_0): the estimated value is more than twice the true value. As a result, households are estimated to be at unrealistically high levels of wealth.

4.7 Welfare effects of misspecification

Now, we investigate whether using the incorrect continuous representation might lead to different policy conclusions. We do this by estimating the effect on welfare (in the sense of discounted utility) of the removal of all risk, or equivalently the introduction of actuarially fair insurance.

Figure 8 shows welfare as a function of wealth at hand with the situation with risk and without risk. Obviously, actuarially fair insurance is welfare improving. The question is whether the gain would be correctly estimated if the continuous representation were used. This is the amount (paid in period 0) that would make a household indifferent between the situation with risk and the situation with insurance. Our results show that in this example households are willing to pay quite high premium. For example, in the discrete model a household with a wealth level of 0.5 (the basic income level) would be willing to give up 40% of wealth to acquire insurance. Figure 4.7 compares the risk premium for the two cases as a fraction of wealth: true values in the discrete model and risk premium derived from the continuous representation.[26] The results are strikingly different. First, risk premium for very poor households are extremely high relative to their wealth level in the former and extremely low in the latter case. Secondly, while in the discrete model the premium first falls and then rises (again as a fraction of wealth) with increasing wealth, estimating the continuous model would suggest a very different pattern: the risk premium first rises as a fraction of wealth and then remains fairly stable. Clearly, using the continuous representation can lead to very misleading conclusions as to what groups would

[25] Note that $\theta_0 = (0.90, 0.95, 0.25, 0.5, 2)$ and $\hat{\theta}$ is the mean in Table 3.

[26] The wobbly parts of relative risk premium graphs on Figure 4.7 are the consequence of slope changes in the value functions, as shown on Figure 8. They may also be the artifact of the algorithm.

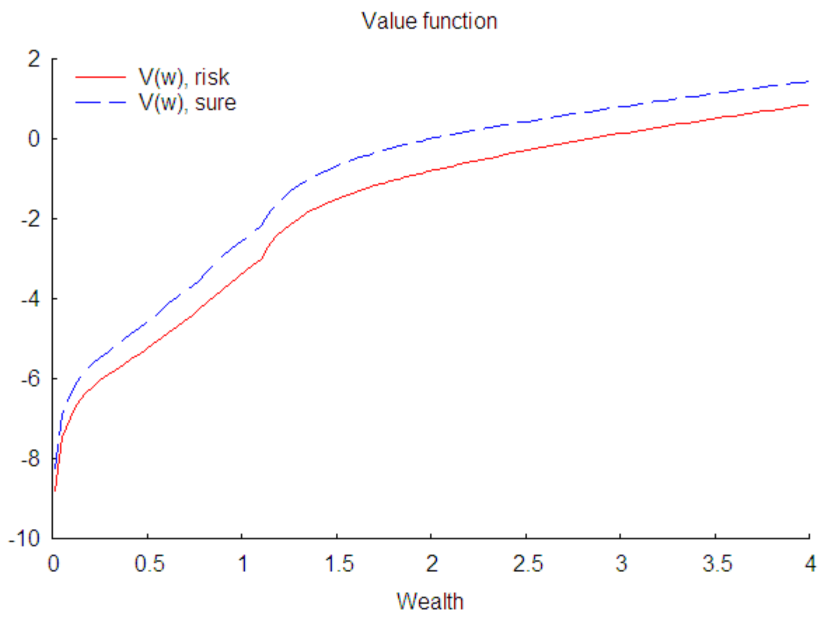


Figure 8: Value function of the discrete model under risk and no risk

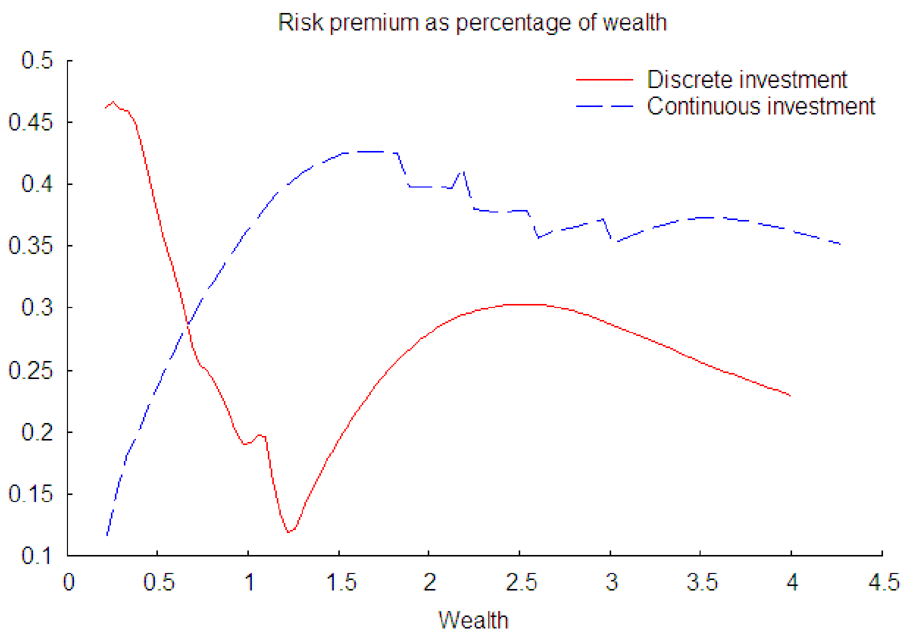


Figure 9: Risk premium as a function of wealth

benefit from the introduction of insurance and to what extent. Notably, the continuous representation could dramatically underestimate the benefits of insurance for very poor people.

5 Conclusion

In this paper we have defined a class of stochastic dynamic choice models with both discrete and continuous decision variables. The key characteristics of this class are that agents cannot borrow and that (given the distribution of the stochastic variables) the agent's wealth at hand is the only information required for investment and consumption decisions. This class contains most dynamic programming models that have been used to analyze intertemporal decisions of rural households under risk. For these models we propose a solution based on value function iteration.

This solution method is primarily developed for models that incorporate both discrete and continuous decision variables. We are particularly interested in settings where credit constrained rural households make savings decisions in terms of a continuous asset (cash or small livestock such as sheep) and investment decisions in terms of a discrete asset (such as cattle).

We illustrate the solution method with a simple example. The model in the example can be solved analytically in the deterministic case, which allows us to study the accuracy of our solution method. We derive a simulated maximum likelihood estimator for this model. The example is used to investigate the importance of modeling indivisibilities in investment. We show that the effects of introducing actuarially fair insurance are very different in the correct model (with indivisibilities) compared to a simplified, continuous representation. Notably, such a continuous representation could erroneously suggest that very poor people have little to gain from insurance. The example therefore suggests the importance of explicitly modeling indivisibilities.

References

[1] Deaton, A. (1991): Saving and Liquidity Constraints. *Econometrica 59(5)*, pp. 1221-1248

[2] Dercon, S. (1998): Wealth, Risk and Activity Choice: Cattle in Western Tanzania. *Journal of Development Economics 55(1)*, pp. 1-42.

[3] Dercon, S. (2002): Income Risk, Coping Strategies and Safety Nets. *The World Bank Research Observer 17(2)*, pp. 309-166.

[4] Elbers, C.; Gunning, J.W. and Kinsey, B. (2007): Growth and Risk: Methodology and Micro Evidence. *The World Bank Economic Review 21(1)*, pp. 1-20.

[5] Elbers, C.; Gunning, J.W. and Pan, L. (2009): Insurance and Rural Welfare: What Can Panel Data Tell Us? *Applied Economics 41(24)*, pp. 3093-3103.

[6] Fafchamps, M. and Pender, J. (1997): Precautionary Saving, Credit Constraints, and Irreversible Investment: Theory and Evidence From Semiarid India. *Journal of Business and Economic Statistics 15*, pp. 180-194.

[7] Goffe, W. L., Ferrier, G. D. and Rogers, J. (1994): Global Optimization of Statistical Functions with Simulated Annealing. *Journal of Econometrics 60(1-2)*, pp. 65-99.

[8] Gunning, J.W. (2008): Risk and Savings: a Taxonomy. Tinbergen Institute Discussion Paper 2008-071/2. Forthcoming in *Economics Letters*.

[9] Judd, K. (1998): *Numerical Methods in Economics*. The MIT Press. Cambridge, Massachusetts.

[10] Pakes, A. (1991): Dynamic Structural Models: Problems and Prospects. Mixed Continuous Discrete Controls and Market Interactions. *Cowles Foundation Discussion Paper No. 984*.

[11] Pan, L. (2008): Poverty, Risk and Insurance: Evidence from Ethiopia and Yemen. *PhD Thesis*, Tinbergen Institute, Amsterdam.

[12] Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992): *Numerical Recipes in C: the Art of Scientific Computing*. Second Edition, Cambridge University Press, World Wide Web sample. www.nr.com

[13] Rosenzweig, M.R. and Wolpin, K.I. (1993): Credit Market Constraints, Consumption Smoothing, and Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India. *Journal of Political Economy 101(21)*, pp. 223-244.

[14] Rust, J. (1996): Numerical Dynamic Programming in Economics in H. Amman, D. Kendrick and J. Rust (eds.), *Handbook of Computational Economics: Agent-Based Computational Economics*, vol. 2, Amsterdam: North-Holland.

[15] Stokey, N.L. and Lucas, R.E. jr. with Prescott, E.C. (1989): Recursive Methods in Economic Dynamics. Harvard University Press.

[16] Vigh, M. (2008): Risk Coping in Rural Ethiopia: the Liquidity and Productivity of Livestock. *MPhil Thesis*, Tinbergen Institute, Amsterdam.

[17] Zimmerman, F.J. and Carter, M.R. (2003): Asset Smoothing, Consumption Smoothing and the Reproduction of Inequality under Risk and Subsistence Constraints. *Journal of Development Economics 71*, pp. 233-260.

Appendix A. Analytical Solution of the Deterministic Model

This section describes the analytical solution of the model with Bellman equation (20)-(22), which is presented again for convenience

$$V(w) = \max_x u(w-x) + \beta V(y_0 + I_{x>=1}(y_1 - y_0 - 1) + x) \tag{41}$$

$$s.t.$$

$$w \geq x \tag{42}$$

$$x \geq 0 \tag{43}$$

In order to solve the problem, first we have to find the wealth level, $\bar{w}_1$, at which the agent will decide to invest 1 in d. At this will wealth level the agent will be indifferent between investing 1 in d today ($V^0(\bar{w}_1)$) and making the same investment tomorrow ($V^1(\bar{w}_1)$). Hence, we can write that

$$V^0(\bar{w}_1) \equiv u(\bar{w}_1 - 1) + \beta V(y_1) = \max_x u(\bar{w}_1 - x) + \beta u(y_0 + x - 1) + \beta^2 V(y_1) \equiv V^1(\bar{w}_1) \tag{44}$$

First we need to find x as a function of $\bar{w}_1$ to be able to solve for $\bar{w}_1$. The FOC for the RHS of equation (44) yield the Euler equation

$$u'(\bar{w}_1 - x) = \beta u'(y_0 + x - 1). \tag{45}$$

Now, using the Euler equation we can solve for x and get

$$x = \frac{\beta^{1/\gamma}}{1+\beta^{1/\gamma}} \bar{w}_1 + \frac{1}{1+\beta^{1/\gamma}}(1 - y_0) \tag{46}$$

Substituting for x we are now able to solve for $\bar{w}_1$ in (44). The equation has two solutions as can be seen on Figure 10. We are interested in the lower solution that gives us $\bar{w}_1 = 1.080$, which is the wealth level below which the agent waits one more period to reach $d = 1$ and saves $x = 0.720$ today, and above which the agent invests in $d = 1$ today. Hence, we observe that there is a discontinuity in the policy function for x at $\bar{w}_1$.

Next we want to find the switch point, $\bar{w}_2$, where the agent is indifferent between investing in $d = 1$ tomorrow and in two days. Hence, we want to solve

$$\begin{aligned} V^1(\bar{w}_2) &\equiv \max_x u(\bar{w}_2 - x) + \beta u(y_0 + x - 1) + \beta^2 V(y_1) & (47) \\ &= \max_{x_0, x_1} u(\bar{w}_2 - x_0) + \beta u(y_0 + x_0 - x_1) + \beta^2 u(y_0 + x_1 - 1) + \beta^3 V(y_1) \equiv V^2(\bar{w}_2) \end{aligned}$$

Again, we need to solve for the x's first through the FOC's, which yield

$$x = \frac{\beta^{1/\gamma}}{1+\beta^{1/\gamma}} \bar{w}_2 + \frac{1}{1+\beta^{1/\gamma}}(1 - y_0) \tag{48}$$

$$x_0 = \frac{\beta^{1/\gamma}}{1+\beta^{1/\gamma}} \bar{w}_2 + \frac{1}{1+\beta^{1/\gamma}}(x_1 - y_0) \tag{49}$$

$$x_1 = \frac{\beta^{1/\gamma}}{1+\beta^{1/\gamma}} (x_0 + y_0) + \frac{1}{1+\beta^{1/\gamma}}(1 - y_0) \tag{50}$$

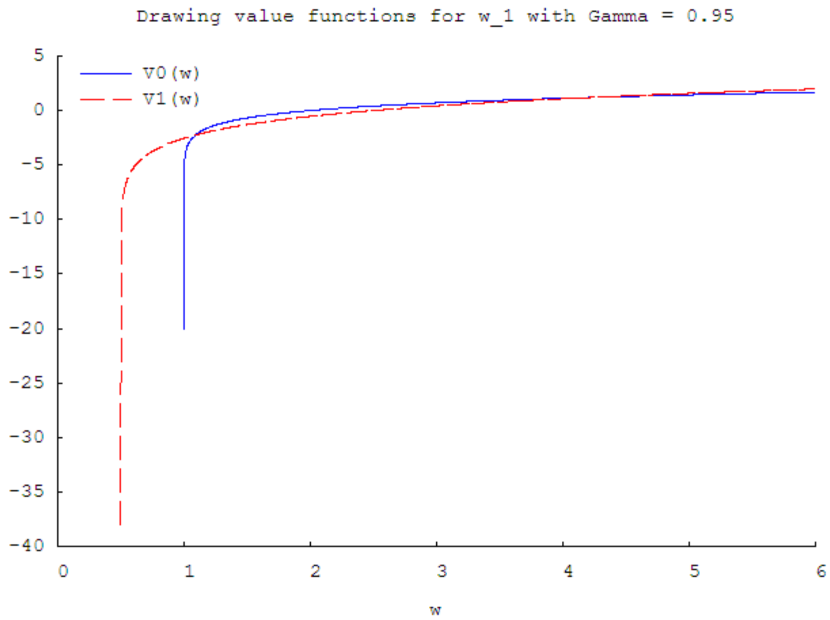


Figure 10: Solving for $\bar{w}_1$: value functions $V^0(\bar{w}_1)$ and $V^1(\bar{w}_1)$

Equation (47) has two solutions again, from which the lower one is of interest to us: $\bar{w}_2 = 0.677$.[27] At this wealth level the savings of the agent who invests in d tomorrow is $x = 0.530$, and the savings of the agent who invests in d in 2 days is $x_0 = 0.350$ today and $x_1 = 0.659$ tomorrow. Between $\bar{w}_1$ and $\bar{w}_2$ agents reach $d = 1$ in one period, and their savings decision follows the FOC of (46).

The next step is to find the switch point, $\bar{w}_3$, where the agent is indifferent between investing in $d = 1$ in two days and in three days. However, we leave it to the reader to derive the remaining switch points.

[27] Note that the upper solution of the equation is not applicable, because at that wealth level the agent has income y_1 instead of y_0.

Appendix B. Notes on programming

This section contains our comments on the implementation of the solution algorithm for the model of section **??**. The algorithm is programmed in Ox.[28]

We set $N = 300$ with the values of w evenly spaced on $[0.01, 2.5]$. The iteration procedure terminates if the largest relative change in the value function and the policy functions become very small. Attention should be paid to the initial values of the policy and value functions and the starting values of the maximization algorithm. Choosing good starting values is important in achieving convergence. When setting up the program, it is a good idea to plot the function approximations after each iteration. For wealth level w_i we use $\psi_x^0(w_i) = 0.6w_i$, $\psi_d^0(w_i) = 1$ if $\psi_x^0(w_i) \geq 1$ and 0 otherwise. $V^0(w_i) = u(\psi_x^0(w_i) + y_0 + \psi_d^0(w_i)(y_1 - y_0 - 1))/(1-\beta)$ for the initial function values. Before applying the initial-value-search routine, we observed that when initial values of x contain zeros, those grids will not move away from zero anymore. This can occur as a result of the flatness of the policy function at the switch points and the log transformation in the optimization, which make the output of the optimization algorithm sensitive to starting values. Therefore, to be on the safe side, it is a good idea to assume some savings for every wealth level in the initial values.

In the optimization problem of Step 2.1 we apply the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method[29] on the log-transformed variable $\tilde{x} = \log(x - d)$, such that the unconstrained maximization algorithm returns $x > d$. The starting value in the maximization is chosen as $\tilde{x}_k^r = \log(\max\{\psi_x^{r-1}(w_i) - \bar{d}_k, \varepsilon\})$ with ε a small positive increment. This value is, however, modified in the neighborhood of the discontinuities in the policy function according to the initial-value-search routine.

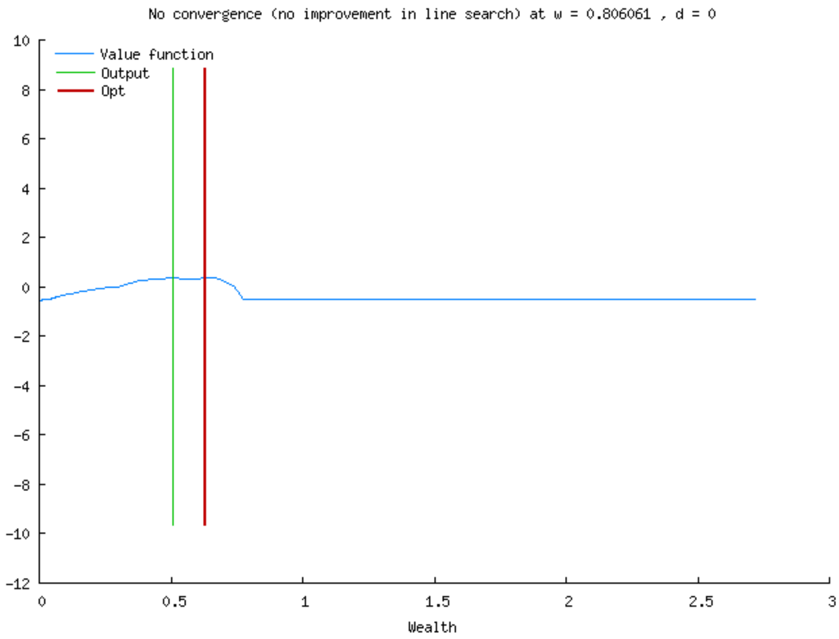
The *initial-value-search algorithm* does the following: if $|\psi_x^{r-1}(w_{i+1}) - \psi_x^{r-1}(w_i)| > w_{i+1} - w_i$ or $|\psi_x^{r-1}(w_i) - \psi_x^{r-1}(w_{i-1})| > w_i - w_{i-1}$, then in iteration r at w_i before executing the maximization routine, we evaluate value function $V(\tilde{x}_b|w_i, d_i)$ at B equidistant values of $\tilde{x}_b$ on range $\exp(\tilde{x}_b) \in (\min\{\exp(\tilde{x}) - 0.05, \varepsilon\}, \min\{\psi_x^{r-1}(w_{i+1}), w_i - \varepsilon\})$. From this we use $\tilde{x}_b$ with maximal value of $V(\tilde{x}_b|w_i, d_i)$ as the starting value in the maximization procedure. The larger we choose B, the closer the initial value of $\tilde{x}$ is going to be to the optimum.

The initial-value-search routine is an important part of the solution algorithm because at the switch points the value function has the same value for two different saving strategies (invest in the advanced technology after k or $k+1$ periods) that contain different optimal savings in x today. Figure 10 in Appendix A illustrates the value function for different values of wealth level assuming optimal decisions for x and d. From a different angle, Figure 11 shows the shape of $V(\tilde{x}_b|w_i, d_i)$ close to a switch point. This plot also highlights the importance of good starting values. Starting at around $x = 0.5$ the BFGS algorithm does not find a value of x that yield higher optimum, however thorough evaluation of the value function using the initial-value-search routine shows that the value function takes its optimum close to $x = 0.63$ instead of 0.5. Hence, without the initial-value-search routine we were not able to locate the discontinuity points in the policy function accurately.

We observe that many times the maximization algorithm reports weak or no convergence, however the value function iteration algorithm converges nonetheless. The problematic areas are (a) close to the discontinuities in the policy function, where the slope is

[28]The Ox code is available on request from Melinda Vigh (mvigh@feweb.vu.nl).

[29]See chapter 5 of Judd (1998) for more information on the BFGS method.



Note: Output is without using smart algorithm.

Figure 11: Optimization routine at a switch point

very flat (see Figure 11); (b) the areas where $x - d = 0$, because there the optimum of the log transformed problem is $\log(x - d) = -\infty$; and (c) sometimes it also occurs for the optimum with $d = 0$ when the optimal chose is $d = 1$. Applying the initial-value-search routine is able to reduce type (a) non-convergence messages.

In order to get a precise approximation of the critical points of the policy function (i.e. where the discontinuities occur), we change the grid of wealth levels after the policy function is close to convergence. More grid-points are added around the wealth levels where the slope of the policy function is changing, and less nodes are used at the intervals where the policy function is (close to) linear. Additional grid points are also added around the location where x becomes positive and where it becomes 1.

With a convergence criterion of 0.001 for the largest relative change in function values compared to the previous iteration, the policy functions converges in 10 iterations, while it takes 40 iterations for the value function to achieve convergence with the same tolerance level. In some cases it might occur that the algorithm diverges for a specific starting value. To avoid the breakdown of the program, we restart the value function iteration with new random starting values if the convergence condition becomes too large. This is useful when estimating the model parameters.

Appendix C. Estimation of the continuous model

This section describes the solution of the continuous model, its estimator and the accuracy of the estimator. For convenience, we present the Bellman equation of the continuous model again:

$$\begin{aligned} V(w) &= \max_{x,d} u(w-x) + \beta EV(x-d+s(y_0+d(y_1-y_0))) && (51)\\ s.t. & \\ & w \geq x && (52)\\ & x \geq d && (53)\\ & d \in [0,1] && (54) \end{aligned}$$

We solve (51)-(54) using value function iteration as described in section **??**. For this model, Step 2.1 of the solution algorithm involves maximizing

$$V^r(w_i) = \max_{l_i,d_i} u(w_i - l_i - d_i) + \beta \sum_{m=1}^{M} \theta_m^s \bar{V}^{r-1}(l_i + s_m(y_0 + d_i(y_1 - y_0)) \qquad (55)$$

where $l_i \equiv x_i - d_i$. To ensure that $l_i > 0$ and $0 < d_i < 1$, we use logarithmic and logistic transformation of variables, respectively. The starting values for the iteration are $\psi_l^0(w_i) = 0.1w_i$, $\psi_d^0(w_i) = \min\{0.4w_i, 1\}$ and $V^0(w_i) = u(\psi_x^0(w_i) + y_0 + \psi_d^0(w_i)(y_1 - y_0 - 1))/(1-\beta)$. Figure 12 plots the policy function for $\theta_0 = (0.9, 0.95, 0.25, 0.5, 2)$, which are the values used to generate data in the baseline estimation.

Now, we can turn to the estimation of the model. Again, we assume that we have data on the current and past asset holdings (x, d, $\bar{x}$, $\bar{d}$) but not on the current wealth at hand (w). The estimator we use is very similar to the one derived in section 4.5. The only difference is that here d can take any value between 0 and 1, not just 0 or 1. Hence, we need to replace the penalty function used in place of the conditional probability $p_\theta(d = d_i|x_i)$.

We use a scaled normal density as a measure of penalty with mean $\varphi_d(x_i)$ and variance σ_ε^2. The size of the penalty depends on σ_ε, with a larger penalty for a smaller σ_ε. This function ensures that $p_\theta(d = d_i|x_i) > 0$ for all $d_i \geq x_i$, and we scale it such that $p_\theta(\varphi_d(x_i) = d_i) = 1$. Hence, the penalty function can be written as

$$P_{\sigma_\varepsilon}(d_i, x_i) = \sqrt{2\pi}\sigma_\varepsilon \phi\left(\frac{d_i - \varphi_d(x_i)}{\sigma_\varepsilon}\right) \qquad (56)$$

and the log-likelihood is

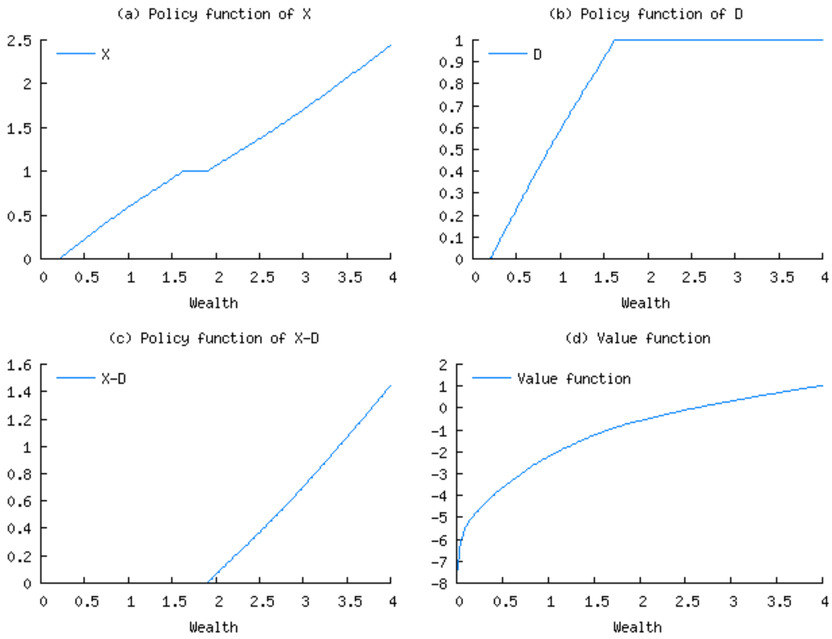


Figure 12: Policy and value functions in the continuous model using at θ_0

$$\begin{aligned}
\ell(\mathbf{x},\mathbf{d}|\theta) \quad = \quad & \sum_{i=1}^{N} I\left(x_i = 1\right) \cdot \\
& \log\left[\Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\bar{\psi}_x^{-1}(1) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\right. \\
& \left. -\Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\underline{\psi}_x^{-1}(1) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\right] \\
+ \quad & \sum_{i=1}^{N} I(x_i = 0)\log\Phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\bar{\psi}_x^{-1}(0) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right) \\
+ \quad & \sum_{i=1}^{N} I\left(x_i \notin \{0,1\}\right)\left[\log\left(\left(\psi_x^{-1}\right)'(x_i)\right) - \log(\sigma(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i))\right. \\
& \left. + \log\phi\left(\frac{\sigma}{2} + \frac{1}{\sigma}\log(\psi_x^{-1}(x_i) - \bar{x}_i + \bar{d}_i) - \frac{1}{\sigma}\log(y_0 + \bar{d}_i(y_1 - y_0))\right)\right] \\
+ \quad & \sum_{i=1}^{N} \sqrt{2\pi}\sigma_\varepsilon\phi\left(\frac{d_i - \varphi_d(x_i)}{\sigma_\varepsilon}\right)
\end{aligned} \tag{57}$$

Figure 13 and Table 4 show the accuracy of this estimator using a Monte Carlo simulation with R=200 replications and a sample size of $N = 1000$. We observe that the parameters are estimated very precisely with the exception of γ. Note that these results are very similar to those on the discrete model in section 4.5.

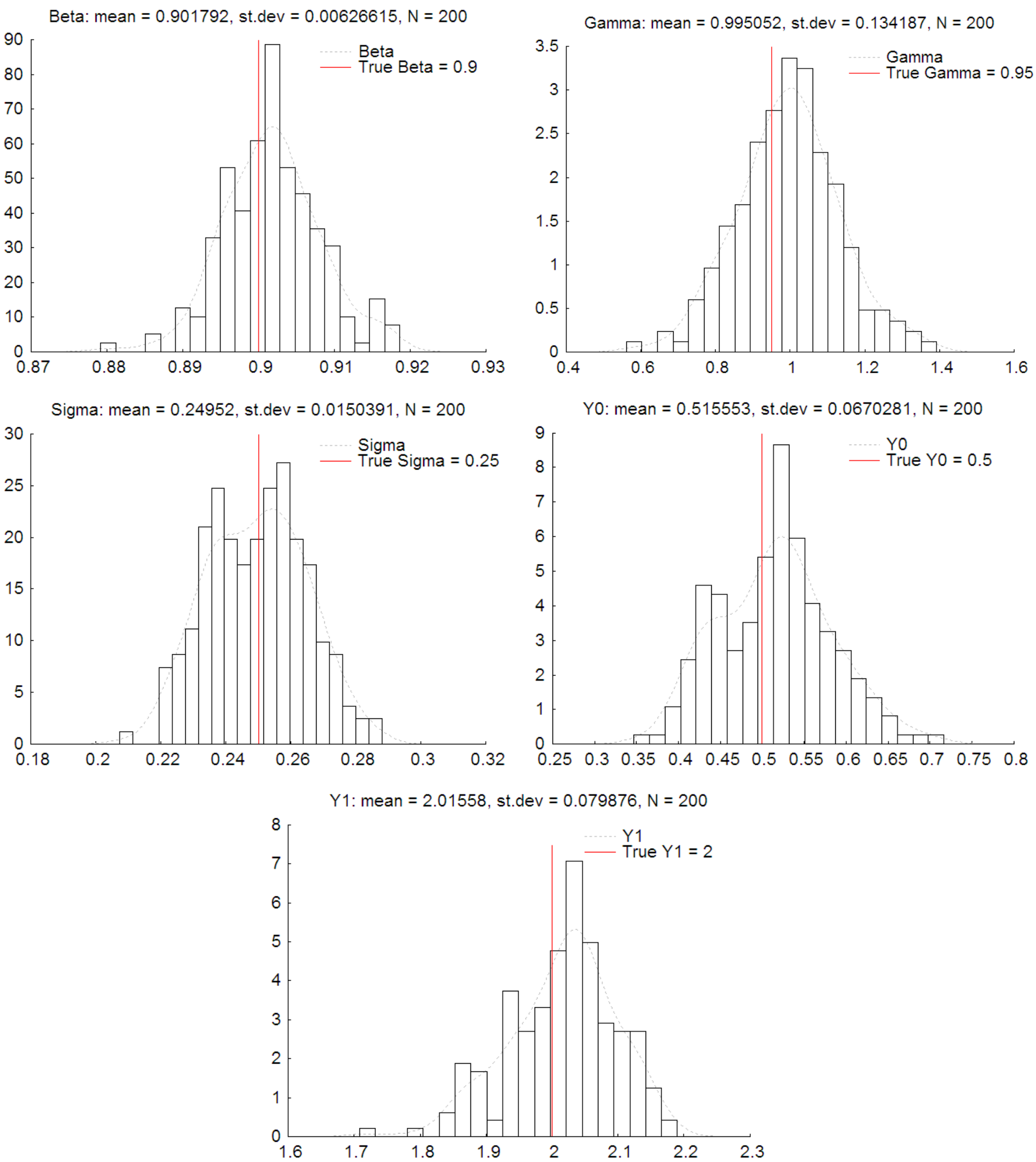


Figure 13: Distribution of Monte Carlo parameter estimates in the continuous investment model

Parameter	True	Mean	Std.dev.	RMSE	MC std.dev.
β	0.90	0.9018	0.0063	0.0065	0.0004
γ	0.95	0.9951	0.1342	0.1415	0.0095
σ	0.25	0.2495	0.0150	0.0150	0.0011
y_0	0.50	0.5156	0.0670	0.0688	0.0047
y_1	2.00	2.0156	0.0799	0.0814	0.0056

Notes:
1. N = 1000, R = 200
2. Mean is calculated as the mean of the parameter estimates
3. Std.dev. calculated as the standard deviation of the parameter estimates around their mean
4. RMSE is calculated as the root mean squared error from the true parameter values
5. MC std.dev. calculated as Std.dev/$\sqrt{R}$
6. Parameter estimates calculated using Simulated Annealing with T=100, NS=30, NT=20, RT=0.85

Table 4: Parameter estimates from Monte Carlo simulation in the continuous model