Non-convexities in dynamic programming problems

Giovanni Gallipoli
University of British Columbia

Lars Nesheim
CEMMAP
UCL and IFS

September 2013

Abstract

Models where agents choose over non-convex budget sets are commonly used in the analysis of economic problems with "extensive margin" decisions and fixed costs. Their solutions have interesting and distinctive features that are especially relevant in quantitative applications. We describe how problems with non-convex choice sets differ from standard problems and investigate under which circumstances the inclusion of random shocks makes their solution identical to the solution of standard problems. A simple framework is provided for the analysis of these problems and a numerical example is illustrated.

1 Introduction

Economic models often entail fixed costs and/or extensive margin binary choices. The associated budget sets are non-convex, removing one of the standard assumptions which guarantee uniqueness and continuity of optimal policies. Examples of choice problems characterized by non-convex budget sets abound in the literature on education, retirement, labor supply and investment, among others. These problems become very relevant whenever a numerical implementation is required because standard solution methods are often unsuitable when uniqueness and continuity is lost. Researchers often resort to introducing randomness in the shape of price or preference shocks which serves the purpose of smoothing the problem and make it more similar to the standard case.[1] This paper provides a framework to analyze problems with non-convex choice sets and discusses under which circumstances the introduction of random shocks is effective in reducing the non-convex problem to one resembling the convex problem. The main results of the paper are illustrated through a simple labor supply problem. We also describe an algorithm which is effective in solving for the numerical solution of problems with non convex choice sets and report some examples.

2 A Lifecycle Consumption-Leisure Problem

Consider the problem of an individual who supplies labor in a competitive market and chooses savings and consumption. Life starts at age 1 and lasts at most T periods. The (gross) rate of return on risk-free assets is r_t for $t \leq T+1$. Over the life cycle the agent chooses a consumption/labor supply path that maximizes expected lifetime utility.

We use $a_t \in A$ to denote individual asset holdings at age t. The market wage at age t is denoted as $w_t \in W$. Nonlabour income at t is $y_t \in Y$.

[1] For example, Gomes, Greenwood, and Rebelo (2001) study an infinite horizon search-theoretic model of equilibrium unemployment in which the employment decision is a binary variable. They find that the outer envelope of the value of being employed or searching is not concave due to the depression where these two conditional values intersect over the assets' domain. This kind of "butterfly" value functions are a common event in the presence of binary choices. Since concavity of value functions is a highly desirable thing, Gomes, Greenwood, and Rebelo (2001) concavize the value function by including a normally distributed shock that "fills up" the convex hull generated by the "butterfly" crossing.

Assumption 1 $A = R_+$, $W = \left[w_t^L, w_t^H\right]$, *and* $Y = \left[y_t^L, y_t^H\right]$ *with* $0 \leq w_t^L < w_t^H \leq \infty$ *and* $0 \leq y_t^L < y_t^H \leq \infty$.

Individual labor supply is denoted as $n \in [0,1]$ and leisure is $l = 1 - n$. Wages and nonlabour income are stochastic while asset returns are deterministic. Individual consumption at age t is denoted as $c_t \in \mathbf{R}_+$. The intertemporal discount factor is $\beta > 0$. The period utility $u : \mathbf{R}_+ \times [0,1] \longrightarrow \mathbf{R}$ depends on consumption c and leisure $l = 1 - n$. The terminal value function $v_{T+1} : \mathbf{R} \longrightarrow \mathbf{R}$ depends on final assets a. We make the following assumption on $u(c,l)$ and $\overline{v}(a)$:

Assumption 2 *u and $\overline{v}$ are twice continuously differentiable on their domains, strictly increasing and strictly concave in all their arguments. Moreovoer, u satisfies the Inada conditions (e.g.* $\lim_{c \longrightarrow 0} u_c = \infty$, $\lim_{c \longrightarrow \infty} u_c = 0$, *and* $\lim_{l \longrightarrow 0} u_l = \infty$*).*

These assumptions imply that consumption and leisure are never zero and savings never equal available resources.

Assumption 3 $a_t \geq a_t^L$ *for every* $t \in \{1, ...T+1\}$ *with* $a_{T+1}^L = 0$.

The above inequality is a standard borrowing constraint imposing a lower bound on asset holding. In what follows we consider the case of $a_t^L = 0$ for all t. We also assume that $r_t = r$ for all t.

For computational convenience we define an identical transformation for labour and nonlabour income:

$$y_t = \left(\frac{y_t^L + \exp\left(z_t\left(1\right)\right)}{1 + \frac{\exp\left(z_t(1)\right)}{y_t^H}} \right) \tag{1}$$

$$w_t = \left(\frac{w_t^L + \exp\left(z_t\left(2\right)\right)}{1 + \frac{\exp\left(z_t(2)\right)}{w_t^H}} \right) \tag{2}$$

This implies that:

$$\begin{array}{rcll} z_t\left(1\right) & = & \ln\left(\frac{y_t - y_t^L}{1 - \frac{y_t}{y_t^H}}\right) & \text{if } y_t^H < \infty \\ & & \ln\left(y_t - y_t^L\right) & \text{if } y_t^H = \infty \\ z_t\left(2\right) & = & \ln\left(\frac{w_t - w_t^L}{1 - \frac{w_t}{w_t^H}}\right) & \text{if } w_t^H < \infty \\ & & \ln\left(w_t - w_t^L\right) & \text{if } w_t^H = \infty \end{array}$$

We assume that $z_t \in R_2$ is an AR(1) process. That is

$$z_t = K_t + \Gamma z_{t-1} + \varepsilon_t$$

where K_t is a deterministic trend, Γ is a transition matrix and where $\varepsilon_t \sim N\left(\mu_\varepsilon, \Sigma_\varepsilon\right)$.

Assumption 4 $(A, \mathcal{A})$ *is a measurable space for assets where $\mathcal{A}$ is the Borel $\sigma-$algebra on A. Given* $X : A \times [0,1]$, $(X, \digamma(X))$ *is a measurable space.for assets and labor supplies such that* $(a, n) \in X$.

Assumption 5 *Let* $Z = Y \times W$ *and let* $\digamma(Z) = \digamma(Y) \times \digamma(W)$ *be the Borel $\sigma-$algebra on* $Z = Y \times W$. $(Z, \digamma(Z))$ *is a measurable space on* $\mathbf{R}_+$. *Also,* $g\left(z_t \mid Z^{t-1}\right)$ *is a transition function on* $(Z, \digamma(Z))$, *with* $z_1 = \bar{z}_1$ *with* $g\left(z_t \left| Z^{t-1}\right.\right) > 0$.

In the final period the household solves

$$v(a_T, z_T) = \max_{(a_{T+1}, n_t)} \left\{ u(c,l) + \alpha\beta \int \overline{v}(r_{T+1}a_{T+1} + y_{T+1})\, g(y_{T+1} | z_T)\, dy_{T+1} \right\}$$

subject to

$$c_T + a_{T+1} = a_T + w_T n_T + y_T$$
$$c_T \geq 0$$
$$0 \leq n_t \leq 1$$
$$l_t = 1 - n_t$$
$$a_{T+1} \geq a_T^L$$

where α is an altruism parameter and $\overline{v}$ is the utility the agent gets from assets upon death. Some properties of the final period problem are:

1. Feasible set is nonempty if $a_T + y_T + w_T - a_T^L \geq 0$;
2. Given $\left(a_T, w_T, y_T, a_T^L\right)$, the feasible set is compact valued and convex valued;
3. $\overline{v}$ is concave;
4. if the integral is finite (that is, if $\overline{v}$ is integrable w.r.t to g), the problem has a unique solution, v is concave in a and differentiable in a. The optimal policies $\left(a_{T+1}^*, n_T^*\right)$ are single valued functions.

In summary, the properties required for a solution (possibly represented as a value function) which is both finite and continuous are:

- for finiteness, the integrand must have finite moments when integrated against $g(\cdot)$. This comes naturally from the finite life-cycle and the compactness of choice sets. If choice sets were not compact, then "enough" concavity of the objective function would still guarantee finiteness of solution.
- for continuity, it must be the case that the order of the limits can be reversed. This condition can be intuitively compared to a dominated convergence theorem.

Let π denote a plan: this is a sequence of functions $\{\pi_t\}_{t=0}^T$ where $\pi_t : Z^t \longrightarrow P$,with Z^t the t-product of Z and P the set of possible actions that are available to the agent. In each period t the agent chooses an action from a subset of feasible alternatives in P. The constraints on these choices are described by a correspondence $\Gamma : A \times Z \to P$; in other words, $\Gamma(a, z)$ is the set of feasible actions in the current state $(a, z) \in A \times Z$. A plan π is feasible from (a_1, z_1) if it satisfies the following conditions:

1. $\pi_1 \in \Gamma(a_1, z_1)$;
2. $\pi_t(z^t) \in \Gamma[a_t^\pi(z^t), z_t]$, all $z^t \in Z^t$, t=1,2,..,T;

   The functions $a_t^\pi(z^t) : Z^t \longrightarrow X$, t=1,2,...,T are defined recursively by

   $$a_t^\pi\left(z^t\right) = \phi\left(a_{t-1}^\pi\left(z^{t-1}\right), \pi_{t-1}\left(z^{t-1}\right), z_t\right) \quad \text{for all } z^t \in Z^t, t = 2, 3, .., T$$

   Finally, we define the objective function. Let $(a_{t+1}, n_t) : A \times Z^t \to R_+ \times [0,1]$ be measureable decision functions for all t. Given some initial conditions (a_1, y_1, w_1), the age t agent's (sequential) utility over savings and work plans, $a = \{a_2, ..., a_{T+1}\}$ and $n = \{n_1, ..., n_T\}$, is denoted as

$U(a_1, z_1, a, n)$ and can be written as the expected discounted sum of period utilities

$$U(a_1, z_1, a, n) = \sum_{s=t}^{T} \beta^{s-t} \int u(c_s, 1 - n_s)\, g\left(z_s \mid Z^{s-1}\right) dz_s + \tag{3}$$

$$+\alpha\beta^{T-t} \int \overline{v}\left(r_{T+1}a_{T+1} + y_{T+1}\right) g\left(z_{T+1} \mid Z^{s-1}\right)$$

subject to

$c_t + a_{t+1} = r_t a_t + y_t + n_t w_t$ for all t

$a_{t+1} \geq a_t^L$ for all t

$c_t \geq 0$ for all t

$0 \leq n_t \leq 1$ for all t

2.1 Household's Problem

We assume that initial asset levels are finite and non-negative; i.e. $0 \leq a_1 < \infty$. Since a_{t+1} cannot exceed available resources we have that for all t

$$a_{t+1} \leq r_t a_t + y_t^H + w_t^H. \tag{4}$$

Therefore

$$a_{t+1} \in \left[a_t^L, r_t a_t + y_t^H + w_t^H\right]. \tag{5}$$

The set of feasible policies Γ is nonempty if $a_t \geq \frac{a_t^L - y_t^L - w_t^L}{r_t}$ and is compact valued if $y_t^H, w_t^H \leq \upsilon < \infty$. The Γ correspondence is convex-valued (as the feasible policies take values in the the space of two convex intervals).

In what follows we assume that the altruism parameter α is equal to 0 which means that assets have no value after the last period of life. For each period, define

$$v(a_t, z_t, t, \pi_t) = u\left(r_t a_t + y_t + w_t n_t^* - a_{t+1}^*, 1 - n_t^*\right) + \beta \int v\left(a_{t+1}^*, z_{t+1}, t+1\right) g\left(z_{t+1} \left| Z^t\right.\right) dz_{t+1}.$$

The period-1 expected utility function $\overline{U}$, given initial conditions (a_1, z_1), and the policy functions sequence $\pi = (\pi_1, \pi_2, .., \pi_T)$ where

$$\pi_t(a_t, z_t) = \left(a_{t+1}^*, n_t^*\right)$$

is defined as

$$\overline{U}(a_1, z_1, 1, \pi) = E\left[\left.\sum_{t=1}^{T} \beta^{t-1} v(a_t, z_t, t, \pi_t)\right| |z_1\right] \tag{6}$$

The household (sequential) problem can finally be stated as

$$\overline{U}^*(a_1, z_1, 1) = \sup_{\pi \in B(a_1, z_1)} \overline{U}(a_1, z_1, 1, \pi) \tag{7}$$

where $B(a_1, z_1)$ identifies the set of feasible policy sequences

$$B(a_1, z_1) \equiv \{\Gamma(a_t, z_t)\}_{t=1}^{T}. \tag{8}$$

2.2 Optimality and Value Functions

In this section we claim that an optimal plan exists and discuss the value function representation of this problem .

Lemma 1 *The set $B(a_t, z_t)$ is compact in the $T-$product topology. In fact at each t, a_{t+1} is bounded as shown in equation (4). Then we can define the set*

$$B(a_1, z_1) \equiv \prod_{t=1}^{T} \left[a_t^L, a_{t+1}\right] \times [0,1]$$

and the feasible set $B(a_1, z_1)$ is the finite product of compact sets and it follows that $B(a_1, z_1)$ is compact in the T product topology. This result is also known as Tychonoff theorem.

By backward induction, we argue that the solution of problem (7) exists and has several desirable properties.

Proposition 1 *Given compactness of the feasible set and continuity of the objective function we know that the final period problem (time T) has a maximum. Monotonicity of the objective function also guarantees that such maximum is unique.*

In the final period we have an optimal policy π_T^* and an indirect utility function $v(\pi_T^*, a_T, z_T)$. As a result, the period $T-1$ sequential problem has a solution and an optimal policy sequence $\left(\pi_{T-1}^*, \pi_T^*\right)$ exists. By induction we have that π_t^* exists for every t and so does $v(\pi_t^*, a_t, z_t)$.

Corollary 1 *There exists an optimal plan $\pi^* = \left\{\pi_2^*, \pi_3^*, ..., \pi_{T+1}^*\right\}$ such that $\breve{U}(a_1, z_1, \pi)$ as defined in equation (6) is equal to the supremum as defined in equation (7).*

We use functional equations to characterize the optimal path[2]. Next we show that a functional equation is an equivalent and unique approach to the household's sequence problem (7). The functional equation for this problem is

$$J_t(a_t, z_t) = \sup_{\pi \in B(a_t, z_t)} v(\pi_t, a_t, z_t) + \beta \int_Z J_{t+1}(a_{t+1}, z_{t+1})\, g\left(z_{t+1} \mid Z^t\right) dz_{t+1} \tag{9}$$

Using recursive substitution T times it can be shown that

$$J_t(a_1, z_1) = \sup_{\pi \in \{\Gamma(a_t, z_t)\}_{t=1}^T} E_Z \sum_{t=1}^{T} \beta^{t-1} v(\pi_t, a_t, z_t) \tag{10}$$

which coincides with the sequence problem in (7). Next we state a proposition on the equivalence between the sequential problem and the functional equation representation.

Proposition 2 *The value of the functional equation $J(a_t, z_t)$ defined in (9) achieves the value of the sequence problem $U^*(a_1, z_1)$ defined in (7). It follows that, for given (a_1, z_1), the value of $J(a_1, z_1)$ is unique and equivalent to $U^*(a_1, z_1)$.*[3]

It can then be shown that a value function does exist. In the next section we proceed by considering a model which departs from the standard case only for the presence of a fixed cost.

[2] In this section we use an hyphen "′" to identify next period unknown values and often omit the age/time subscripts for notational simplicity.

[3] The value of $\overline{U}^*(a_1, z_1)$ is the supremum of $\breve{U}(a_1, z_1, \pi)$ in (6) over the set of policies P, and the supremum of any function is unique. $J(a_1, z_1)$ achieves $\overline{U}^*(a_1, z_1)$, and therefore $J(a_1, z_1)$ is also unique and equivalent to $\overline{U}^*(a_1, z_1)$.

3 A Departure from the Standard Case: General Model with Fixed Costs

We depart from the model described in the previous section by introducing the possibility of a lump-sum payment when an agent chooses to work positive hours. The objective function is the same as in (3); however the budget constraint changes because of the fixed cost of participating in the job market, F. We denote by d_t the variable indicating participation, which takes value 1 when an agent chooses to pay the fixed cost and enter the labor market. It is important to notice that in this case the choice sets identified by the correspondence $\Gamma(a, z)$ are generally not convex-valued because of the fixed cost F. However the choice sets are still convex when conditioning on a value of d_t.

3.1 Value Functions and Optimal Policies

Denote savings carried over to next period as s_t. The consumer's problem in period t is then

$$v(a_t, y_t, w_t, t) = \max_{\{d_t, s_t, n_t\}} \left\{ u(c_t, 1 - n_t) + \beta \int v(a_{t+1}, y_{t+1}, w_{t+1}, t+1)\, g(z_{t+1} | z_t)\, dz_{t+1} \right\} \quad (11)$$

subject to

$$c_t + s_t = a_t + y_t + w_t n_t - F d_t$$
$$a_{t+1} = r_{t+1} s_t$$
$$(1 - d_t)\, n_t = 0$$
$$c_t \geq 0$$
$$0 \leq n_t \leq 1$$
$$a_{t+1} \geq a_t^L$$

It is useful to consider the value functions conditional on entering or not entering the market. We define the value function conditional on $d_t = 0$ as

$$v_0(a_t, y_t, w_t, t) = \max_{\{s_t\}} \left\{ u(c_t, 1) + \beta \int v(a_{t+1}, y_{t+1}, w_{t+1}, t+1)\, g(z_{t+1} | z_t)\, dz_{t+1} \right\} \quad (12)$$

subject to

$$c_t + s_t = a_t + y_t$$
$$a_{t+1} = r_{t+1} s_t$$
$$c_t \geq 0$$
$$a_{t+1} \geq a_t^L.$$

When $d_t = 0$ the labor supply is zero and leisure in the utility function is set to one. Therefore the maximum depends only on the savings' decision.

Similarly, define the value function conditional on $d_t = 1$ as

$$v_1(a_t, y_t, w_t, t) = \max_{\{s_t, n_t\}} \left\{ u(c_t, 1 - n_t) + \beta \int v(a_{t+1}, y_{t+1}, w_{t+1}, t+1)\, g(z_{t+1} | z_t)\, dz_{t+1} \right\} \quad (13)$$

subject to

$$c_t + s_t = a_t + y_t + w_t n_t - F$$
$$a_{t+1} = r_{t+1} s_t$$
$$c_t \geq 0$$
$$0 \leq n_t \leq 1$$
$$a_{t+1} \geq a_t^L$$

In this case the agent can choose positive labor supply values and the maximization depends on choosing an optimal level of both savings and leisure.

3.1.1 Zero fixed costs: $F = 0$

To illustrate the implications of introducing discrete choices and a fixed cost we characterize some properties of the value functions and optimal policies under the assumption that the fixed cost F is equal to zero.

Assumption 6 *Assume u is C^2, strictly increasing in c, strictly increasing in leisure and strictly concave in (c, l).*

Assumption 7 *Assume $v(a, y, w, t+1)$ is C^2, increasing and concave in a.*

Assumption 8 *Let $(s_t^*, n_t^*) = (\pi_s, \pi_n)$ be the solution to (11). Assume that, when $F = 0$, there exists $a_R(0)$ such that $n^* > 0$ for all $a < a_R(0)$ and $n^* = 0$ for all $a \geq a_R(0)$.*

Lemma 2 *Then*

$$\begin{aligned} v(a, y, w, t) &= v_1(a, y, w, t) = v_0(a, y, w, t) \text{ for all } a \geq a_R(0) \\ v(a, y, w, t) &= v_1(a, y, w, t) \geq v_0(a, y, w, t) \text{ for all } a < a_R(0) \end{aligned}$$

The labour supply policy looks like

$$n = \pi_n(a, y, w, t) = \begin{cases} \pi_{n1}(a, y, w, t) > \pi_{n0}(a, y, w, t) = 0 & a < a_R(0) \\ \pi_{n1}(a, y, w, t) = \pi_{n0}(a, y, w, t) = 0 & a \geq a_R(0) \end{cases}.$$

The optimal savings policy is

$$s_t = \pi_s(a, y, w, t) = \begin{cases} \pi_{s1}(a, y, w, t) & a < a_R(0) \\ \pi_{s1}(a, y, w, t) = \pi_{s0}(a, y, w, t) & a \geq a_R(0) \end{cases}.$$

Lemma 3 *Given assumptions 6, 7 and 8, then the period t policies are continous functions of (a, y, w) and differentiable almost everywhere in a. All three period t value functions are continuous in (a, y, w). In addition, all three period t value functions are concave and differentiable in a.*

Under further restrictions we could say more about (π_n, π_s). For example, under certain conditions, (π_n, π_s) are weakly monotonic in a.

3.1.2 Positive fixed costs: $F > 0$

If the fixed cost F is strictly positive, things change. Let $v(a, y, w, t, F)$, $v_0(a, y, w, t, F)$ and $v_1(a, y, w, t, F)$ be the unconditional and conditional value functions given fixed cost F. First we notice that introducing fixed costs F corresponds to shifting down the value of work (with no fixed costs) over the assets' space.

Lemma 4 *The following equalities hold:*

$$\begin{aligned} v_0(a, y, w, t, F) &= v_0(a, y, w, t, 0) \\ v_1(a, y, w, t, F) &= v_1(a - F, y, w, t, 0) \\ v(a, y, w, t, F) &= \max\{v_0(a, y, w, t, 0), v_1(a - F, y, w, t, 0)\} \end{aligned}$$

The optimal participation decision can be characterized in terms of an asset threshold. Bounds for this thereshold can be identified.

Lemma 5 *Under assumptions 6, 7 and 8,*

$$\begin{aligned} v(a, y, w, t, F) &= v_1(a - F, y, w, t, 0) > v_0(a, y, w, t, 0) \text{ for all } a < a_R(F) \\ v(a, y, w, t, F) &= v_1(a - F, y, w, t, 0) = v_0(a, y, w, t, 0) \text{ for } a = a_R(F) \\ v(a, y, w, t, F) &= v_0(a, y, w, t, 0) > v_1(a - F, y, w, t, 0) \text{ for all } a > a_R(F) \end{aligned}$$

where

$$a_R(0) - F < a_R(F) < a_R(0).$$

Proof. The inequality $a_R(0) - F < a_R(F)$ follows from the fact that

$$v_0(a_R(0) - F, y, w, t) \leq v_1(a_R(0) - F, y, w, t)$$

as stated in 2. The inequality $a_R(F) < a_R(0)$ follows obviously from the fact that $v_1(a, y, w, t, F) = v_1(a - F, y, w, t, 0)$. ■

In the presence of postitive fixed costs the value functions and associated optimal policies have different characteristics.

Lemma 6 *$v_0(a, y, w, t, F)$ and $v_1(a, y, w, t, F)$ are continuous in (a, y, w) and concave and differentiable in a.*

Lemma 7 *$v(a, y, w, t, F)$ is continuous in (a, y, w). It is differentiable with respect to a almost everywhere and piecewise concave in a.*

Lemma 8 *The conditional policy functions (π_{n1}, π_{s1}) and (π_{n0}, π_{s0}) are continuous functions and differentiable almost everywhere in a.*

Lemma 9 *The conditional policy functions satisfy*

$$\begin{aligned}\pi_{n1}(a, y, w, t, F) &= \pi_{n1}(a - F, y, w, t, 0) \\ \pi_{s1}(a, y, w, t, F) &= \pi_{s1}(a - F, y, w, t, 0)\end{aligned}$$

Lemma 10 *The unconditional labour supply policy correspondence is*

$$\pi_n(a, y, w, t, F) = \left\{ \begin{array}{cc} \pi_{n1}(a - F, y, w, t, 0) & \text{for } a \leq a_R(F) \\ 0 & \text{for } a \geq a_R(F) \end{array} \right\}$$

and

$$\pi_{n1}(a, y, w, t, F) = \pi_{n1}(a - F, y, w, t, 0) > 0 \text{ for all } a < a_R(0) + F$$

In particular,

$$\pi_{n1}(a_R(F), y, w, t, F) = \pi_{n1}(a_R(F) - F, y, w, t, 0) > 0 = \pi_{n0}(a_R(F), y, w, t, F)$$

The optimal policy $\pi_n(a, y, w, t, F)$ is a function differentiable almost everywhere in a.

3.2 Characterizing the Indifference Set

Conditional on (a, y, w, t), the worker is indifferent between working and not working when

$$v_0(a, y, w, t) = v_1(a - F, y, w, t, 0) \tag{14}$$

Assumption 9 *Suppose $v_1(a, y, w, t, F) - v_0(a, y, w, t)$ is continuously differentiable in (a, y, w), that there exists a point (a_0, y_0, w_0) such that*

$$v_1(a_0 - F, y_0, w_0, t, 0) - v_0(a_0, y_0, w_0, t) = 0 \tag{15}$$

and

$$\frac{\partial v_1(a_0 - F, y_0, w_0, t, 0)}{\partial a} - \frac{\partial v_0(a_0, y_0, w_0, t)}{\partial a} \neq 0. \tag{16}$$

Lemma 11 *Then the conditions of the local implicit function theorem are satisfied and there exists a function defined on a neighborhood of (y_0, w_0) such that*

$$\begin{aligned}a &= b(y_0, w_0) \\ \frac{\partial b}{\partial y} &= \frac{\frac{\partial v_0(a_0, y_0, w_0, t)}{\partial y} - \frac{\partial v_1(a_0 - F, y_0, w_0, t, 0)}{\partial y}}{\frac{\partial v_0(a_0, y_0, w_0, t)}{\partial a} - \frac{\partial v_1(a_0 - F, y_0, w_0, t, 0)}{\partial a}} \\ \frac{\partial b}{\partial w} &= \frac{\frac{\partial v_0(a_0, y_0, w_0, t)}{\partial w} - \frac{\partial v_1(a_0 - F, y_0, w_0, t, 0)}{\partial w}}{\frac{\partial v_0(a_0, y_0, w_0, t)}{\partial a} - \frac{\partial v_1(a_0 - F, y_0, w_0, t, 0)}{\partial a}}\end{aligned}$$

The above result is useful to characterize the relationship between the marginal asset level a and the other state variables y and w. This relationship is local and idenitifies the change in assets which is required to keep an agent indifferent between participating or not as y or w changes. Notice that, given a current asset level a_t, it is possible to define two distinct regions in the (y, w) space: we denote as $A_0(a_t)$ the region where the optimal $d_t = 0$, and as $A_1(a_t)$ the region where the optimal $d_t = 1$.

3.3 The Shape of the Value Function

We now turn our attention to the shape of the value function: in particular, we study the slope (first derivative) and curvature (second derivative), in order to assess whether the value function is concave when we introduce fixed costs. For the sake of simplicity we assume no persistence in (y_t, w_t). In what follows we also denote current assets as s_{t-1}.

Let $A_0(s_{t-1})$ be the region where $d_t = 0$. This is the region

$$A_0(s_{t-1}) = \{(y_t, w_t) : v_0(r_t s_{t-1} + y_t, t) \geq v_1(r_t s_{t-1} + y_t - F, w_t, t)\} \tag{17}$$

Let $A_1(s_{t-1})$ be the region where $d_t = 1$. This is the region

$$A_1(s_{t-1}) = \{(y_t, w_t) : v_0(r_t s_{t-1} + y_t, t) \leq v_1(r_t s_{t-1} + y_t - F, w_t, t)\} \tag{18}$$

We assume that the indifference set

$$A_{eq}(s_{t-1}) = \{(y_t, w_t) : v_0(r_t s_{t-1} + y_t, t) = v_1(r_t s_{t-1} + y_t - F, w_t, t)\} \tag{19}$$

has measure zero. The expected value function, before y and w are realized, can be written as

$$\begin{aligned} Ev(s_{t-1}) &= \int\limits_{A_0(s_{t-1})} v_0(r_t s_{t-1} + y_t, t)\, g(y_t, w_t)\, dy_t dw_t \\ &\quad + \int\limits_{A_1(s_{t-1})} v_1(r_t s_{t-1} + y_t - F, w_t, t)\, g(y_t, w_t)\, dy_t dw_t \end{aligned} \tag{20}$$

The first derivative of the unconditional value (20) with respect to current assets can be written as

$$\begin{aligned} \frac{\partial Ev(s_{t-1})}{\partial s_{t-1}} &= \int\limits_{A_0(s_{t-1})} r_t \frac{\partial v_0(r_t s_{t-1} + y_t, t)}{\partial s_{t-1}} g(y_t, w_t)\, dy_t dw_t \\ &\quad - \frac{\partial A_0(s_{t-1})}{\partial s_{t-1}} \int\limits_{\partial A_0(s_{t-1})} v_0(r_t s_{t-1} + y_t, t)\, g(y_t, w_t)\, dy_t dw_t \\ &\quad + \int\limits_{A_1(s_{t-1})} r_t \frac{\partial v_1(r_t s_{t-1} + y_t - F, w_t, t)}{\partial s_{t-1}} g(y_t, w_t)\, dy_t dw_t \\ &\quad + \frac{\partial A_0(s_{t-1})}{\partial s_{t-1}} \int\limits_{\partial A_0(s_{t-1})} v_1(r_t s_{t-1} + y_t - F, w_t, t)\, g(y_t, w_t)\, dy_t dw_t \end{aligned}$$

The notation is meant to indicate that when s_{t-1} changes, the boundary changes and we evaluate the impact on the contribution of v_0 and of v_1.In the first derivative these "boundary effects"cancel out because by definition $v_0 = v_1$ on the boundary. Then we can write..

$$\begin{aligned} \frac{\partial Ev(s_{t-1})}{\partial s_{t-1}} &= \int\limits_{A_0(s_{t-1})} r_t \frac{\partial v_0(r_t s_{t-1} + y_t, t)}{\partial s_{t-1}} g(y_t, w_t)\, dy_t dw_t \\ &\quad + \int\limits_{A_1(s_{t-1})} r_t \frac{\partial v_1(r_t s_{t-1} + y_t - F, w_t, t)}{\partial s_{t-1}} g(y_t, w_t)\, dy_t dw_t. \end{aligned} \tag{21}$$

The second derivative of the expected value is

$$\begin{aligned}\frac{\partial^2 Ev(s_{t-1})}{\partial s_{t-1}^2} &= \int\limits_{A_0(s_{t-1})} r_t^2 \frac{\partial^2 v_0(r_t s_{t-1}+y_t,t)}{\partial s_{t-1}^2} g(y_t,w_t)\,dy_t dw_t \\ &+ \int\limits_{A_1(s_{t-1})} r_t^2 \frac{\partial^2 v_1(r_t s_{t-1}+y_t-F,w_t,t)}{\partial s_{t-1}^2} g(y_t,w_t)\,dy_t dw_t \\ &+ \frac{\partial A_0(s_{t-1})}{\partial s_{t-1}} \int\limits_{\partial A_0(s_{t-1})} r_t \frac{\partial v_0(r_t s_{t-1}+y_t,t)}{\partial s_{t-1}} g(y_t,w_t)\,dy_t dw_t \\ &- \frac{\partial A_0(s_{t-1})}{\partial s_{t-1}} \int\limits_{\partial A_0(s_{t-1})} r_t \frac{\partial v_1(r_t s_{t-1}+y_t-F,w_t,t)}{\partial s_{t-1}} g(y_t,w_t)\,dy_t dw_t.\end{aligned}$$

The above expression can be rearranged to give

$$\begin{aligned}\frac{\partial^2 Ev(s_{t-1})}{\partial s_{t-1}^2} &= \int\limits_{A_0(s_{t-1})} r_t^2 \frac{\partial^2 v_0(r_t s_{t-1}+y_t,t)}{\partial s_{t-1}^2} g(y_t,w_t)\,dy_t dw_t \\ &+ \int\limits_{A_1(s_{t-1})} r_t^2 \frac{\partial^2 v_1(r_t s_{t-1}+y_t-F,w_t,t)}{\partial s_{t-1}^2} g(y_t,w_t)\,dy_t dw_t \\ &+ \frac{\partial A_0(s_{t-1})}{\partial s_{t-1}} \int\limits_{\partial A_0(s_{t-1})} r_t \left(\frac{\partial v_0(r_t s_{t-1}+y_t,t)}{\partial s_{t-1}} - \frac{\partial v_1(r_t s_{t-1}+y_t-F,w_t,t)}{\partial s_{t-1}}\right) g(y_t,w_t)\,dy_t dw_t\end{aligned} \tag{22}$$

3.4 Relevant Parameters for Concavity of the Unconditional Value Function

The second derivative in equation (22) has 3 components: the first two of them are clearly negative, because of the concavity of the conditional values. However the third component is

$$\frac{\partial A_0(s_{t-1})}{\partial s_{t-1}} \int\limits_{\partial A_0(s_{t-1})} r_t \left(\frac{\partial v_0(r_t s_{t-1}+y_t,t)}{\partial s_{t-1}} - \frac{\partial v_1(r_t s_{t-1}+y_t-F,w_t,t)}{\partial s_{t-1}}\right) g(y_t,w_t)\,dy_t dw_t$$

and it can assume positive values, depending on the sign of $\frac{\partial A_0(s_{t-1})}{\partial s_{t-1}}$ and $\frac{\partial v_0(r_t s_{t-1}+y_t,t)}{\partial s_{t-1}} - \frac{\partial v_1(r_t s_{t-1}+y_t-F,w_t,t)}{\partial s_{t-1}}$. In the labor market participation problem we are studying we can assume that the third term is positive. The absolute size of this term will increase in the difference of the slopes of v_0 and v_1, the size of $\frac{\partial A_0(s_{t-1})}{\partial s_{t-1}}$, the size of the interest rate r, the (local) density weight $g(y_t,w_t)$ and the size of fixed cost F.

We can then define a vector of parameters $\theta = (g, F, v_0, v_1, r)$. The expected utility is

$$\begin{aligned}Ev(s_{t-1},\theta) &= \int\limits_{A_0(s_{t-1})} v_0(r_t s_{t-1}+y_t,t)\, g(y_t,w_t)\,dy_t dw_t \\ &+ \int\limits_{A_1(s_{t-1})} v_1(r_t s_{t-1}+y_t-F,w_t,t)\, g(y_t,w_t)\,dy_t dw_t.\end{aligned}$$

Given θ, we can graph this function as a a function of s_{t-1} for all $s_{t-1} \in [s_L, s_H]$. Then we can vary θ and show how concavity of Ev depends on θ. For example, when $F = 0$, Ev is always concave. On the other hand when F is large (which is associated to large differences in the derivatives of the conditional value functions), the (local) density $g(y_t,w_t)$ must put relatively little weight on the boundary set for concavity to hold: we argue that this is more likely to happen when the probability distribution over the (y,w) space has relatively large variance.

4 Quantitative Relevance of Non-Convexities

Our discussion so far has shown how the presence of discrete choices and fixed costs in a standard model introduces the possibility that value functions are non-concave. In particular, we have stressed that introducing random shocks does not necessarily achieve the objective of buying back concavity and we have discussed which parameters of the problem play an important role in determining the final shape of the expected value functions.

In this section we ask whether the non-concavity problem can arise under a 'reasonable' parametrization of our simple model. Using numerical simulations we also investigate the quantitative importance of different parameters in inducing non-concavity.

4.1 Choosing Parameters for the Standard Lifecycle Model

To assess the quantitative importance of different parameters for the non-concavity of the expected value function, we first need to assign reasonable 'benchmark' values to them. For simplicity we use a two-period numerical counterpart of the life-cycle model with no fixed costs.

The parameters to assign are the following:

- Period utility. We assume a standard CRRA power specification over non-separable leisure and consumption

$$u(c, 1-n) = \frac{\left(c^{\alpha}(1-n)^{1-\alpha}\right)^{1-\sigma} - 1}{1-\sigma}$$

  where n denotes hours worked and $(1-n)$ is leisure; consumption is denoted by c. The elasticity of intertemporal substitution (IES) is defined as $\frac{1}{1-\alpha(1-\sigma)}$. Of course there are several (α, σ) pairs that are consistent with a given IES: we choose $\alpha = 0.6$, which delivers reasonable labor supply (see Ríos-Rull (1993)). The value of σ is set to 1.6. The implied IES is around 0.73 which is close to estimates by Blundell, Browning, and Meghir (1994) and Attanasio and Weber (1993).

- Discount factor. We set $\beta = 0.99$.

- Asset space. The lower bound for assets (a_t^L) is set to zero and the upper bound is set to 1.

- Wage process. We model the wage rate in a given period to be a stochastic (normal) variable with mean μ_w and variance σ_w. We set $\mu_w = 0.025$. The main question is: what is a realistic value for σ_w? We set this variance to be equal to estimates (see Gallipoli, Meghir, and Violante (2008)) of the variance of the innovation of the AR(1) process which is often used to describe persistent wage shocks in the literature. This variance (in logs) is about0 0.020, which is roughly equivalent to 2% of the wage rate (in our case the mean wage). The wage process we use does not incorporate most deterministic components (efficiency, age, etc) which all contribute to an agent's mean wage. However we strip down relevant uncertainty about wages to the perturbation of the AR(1) process: this we interpret as the genuine wage uncertainty faced in each period by a worker.

- Non-labour income. Non labour income has two components in the model. The first is the risk free return (r_t) on asset savings. The second (denoted as y_t in the model) is a stochastic component of non-labour income: in the current draft we choose to set y_t to zero, so that all non-labour income is risk-free. To the best of our knowledge there is no available estimate of the random process followed by non-labour income, which makes the parameterization of y_t difficult. In the future we plan to experiment with alternative assumptions about this process.

- Risk-free (gross) interest r_t. We set this value to 1.04.

- Fixed cost of participating in the labour market. In the benchmark model we set $F = 0$.

Using these parameters we simulate a two period version of the model and we find, as expected, that:

- $v_0(a_t) < v_1(a_t)$ for every level of current assets at which it is optimal to supply labor. This should obviously be the case when $F = 0$ and v_0 is simply a constrained (no labour supply) version of v_1.
- The slope of $v_0(a_t)$ is much larger than the slope of $v_1(a_t)$ for low assets, but it gets closer as assets grow and becomes exactly the same when no labor supply is chosen at the optimum.
- The optimal consumption associated to $v_1(a_t)$ is larger than that associated to $v_0(a_t)$ for all current asset levels at which it is optimal to work.
- the expected utility is a smooth concave function of future assets (savings) and its gradient is a smooth, decreasing, convex function.

4.2 Introducing a Positive Labor-Participation Cost

After verifying that the benchmark model, under a reasonable parametrization, generates sensible results, we introduce a positive fixed cost of participating in the labor market.

In the following table we report the size of the fixed cost we introduce, also relative to mean and variance of wages, and we indicate whether the expected value function, as defined in (20), preserves concavity and whether its gradient preserves convexity.

F				
value	$\frac{F}{\mu_w}$	$\frac{F}{\sigma_w}$	$Ev(s_{t-1})$ concave	$\frac{\partial Ev(s_{t-1})}{\partial s_{t-1}}$ convex
0.0025	0.1	5	yes	no
0.0050	0.2	10	yes	no
0.0075	0.3	15	no	no
0.0100	0.4	20	no	no

Table 1: Implications of different fixed costs of participation

These results confirm that the introduction of fixed costs of participation in the labor market does not necessarily imply non-concavity of the expected value function (although its gradient does lose convexity).

However we find that a relatively small $F = 0.3 \times \mu_w$ is already associated with a non-concave $Ev(s_{t-1})$. Given this fixed cost (as a share of μ_w) we also find that concavity can be obtained only by roughly doubling the variance σ_w, which brings back the $\frac{F}{\sigma_w}$ ratio to 10. In fact, there seems to be a relationship between the size of $\frac{F}{\sigma_w}$ and the shape of $Ev(s_{t-1})$.

We also find that increasing the interest rate r implies occurrence of the non-concavity at even lower levels of F, as predicted by our analytical results.

5 Conclusion

In this paper we study under what circumstances the introduction of discrete choices with fixed costs, in an otherwise standard life-cycle model, can lead to non-concavity of the expected value functions. Our findings suggest that relatively small fixed costs can generate non-concavity of the expected continuation value with an associated loss of uniqueness of the optimal policies. Introducing random shocks helps to gain back concavity of the expected continuation value but only if the variance of the shocks is sufficiently large.

References

ATTANASIO, O., AND G. WEBER (1993): "Consumption Growth, the Interest Rate and Aggregation," *Review of Economic Studies*, 60(3), 631–649.

BLUNDELL, R., M. BROWNING, AND C. MEGHIR (1994): "Consumer Demand and the Life-Cycle Allocation of Household Expenditure," *Review of Economic Studies*, 61(1), 57–80.

GALLIPOLI, G., C. MEGHIR, AND G. VIOLANTE (2008): "Equilibrium Effects of Education Policies: a Quantitative Evaluation," Mimeo.

GOMES, J., J. GREENWOOD, AND S. REBELO (2001): "Equilibrium Unemployment," *Journal of Monetary Economics*, 48, 109–152.

RÍOS-RULL, J.-V. (1993): "Working in the Market, Working at Home and the Acquisition of Skills: A General Equilibrium Approach," *American Economic Review*, 83(4), 893–907.