Undiscounted dynamic discrete choice and regularized Koopmans problems

Nicola Rosaia*

June 9, 2019

EXTREMELY PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CIRCULATE

1 Model set up

In this section we review the basic dynamic discrete-choice setup. The state variable is $x \in X$, which we assume to take only a finite number of values. Agents choose actions $a \in A$ from a finite set A. A correspondence $(A(x))_{x \in X}$ describes the set of actions $a \in A(x)$ that are available to agents in each state x, hence $A = \cup_{x \in X} A(x)$. The instantaneous utility that an agent derives from choosing a is state x is

$$u_{a,x} + \epsilon_a$$

where $u \in \mathbb{R}^{A \times X}$ is the vetor of structural payoffs, and $\epsilon \in \mathbb{R}^A$ is a vector of utility shocks associated with each action. The vector (x, ϵ) evolves according to a first-order controlled Markov process. Conditional on the current state x and choice a, the future value of x, denoted by x', is drawn from a transition kernel $p_a(x'|x)$. The future value of ϵ, denoted by ϵ', is then drawn from a distribution $q(\cdot|x')$, which can depend on the realization of the state variable. In particular, following Rust (1987) and most of the literature, we rule out serially persistent forms of heterogeneity by imposing the so-called conditional independence assumption:

Assumption 1. *Conditionally on the current state x, the distribution of ϵ is independent on the previous history of play:*

$$Pr(x', \epsilon'|x, \epsilon, a) = Pr(\epsilon'|x')\, Pr(x'|x, a)$$

Although all our main results have an equivalent in the general case, for simplicity of exposition we assume that the distributions of utility shocks have full support. This allow us to neglect the tie-breaking in agents' optimality conditions and to obtain point-identification results.

Assumption 2. *Conditionally on each state x, the distribution $q(\cdot|x)$ has full support*

*Harvard University. nicolarosaia@g.harvard.edu

In order to simplify the exposition, we confine our analysis to the case in which the state space is "connected" under the transition probabilities in the following sense: every state can be reached with positive probability from every other state by following an appropriate sequence of choices.[1] We state this assumption formally:

Assumption 3. *For every set of states $Y \subsetneq X$, there exist $y \in Y$, $x \in X \setminus Y$ and $a \in A(y)$ such that $p_a(x|y) > 0$.*

1.1 Outcomes

As described in the next section, agents can condition their behavior on the privately observed value of the shock, which is not observed by the analyst. From the persepctive of the analyst, the outcome is a system of choice probabilities decribing the frequency according to which agents take different actions conditional on each state:

Definition 4. An *outcome* is a system of conditional choice probabilities σ, whose (x,a)th element is denoted by $\sigma(a|x)$.

For every state x and control $a \in A(x)$, $\sigma(a|x)$ denotes the probability of a in state x. In what follows we use the following shorthand notation for σ:

$$P_\sigma = \left[\sum_{a \in A(x)} p_a(y|x)\,\sigma(a|x) \right]_{x,y \in X}.$$

Hence the (x,y)th element of P_σ is the transition probability from x to y obtained by integrating the action-dependent transition probabilities with respect to σ. If an agent behave as described by σ from some initial state, then probability according to which she visits different states in the long-run is described by a stationary distribution of P^σ. Compounding this distribution with σ yelds a joint distribution over actions and states:

Definition 5. An *steady state outcome* is a probability distribution μ over $A \times X$ with the property that there exists an outcome σ such that the marginal of μ over X is a stationary distribution of P_σ and

$$\forall a, x: \quad \mu(a,x) = \sigma(a|x)\,\mu(x).$$

Actually steady state outcomes can be defined without reference to a particular system of conditional choice probabilities:

[1] In other wirds, define the relations "x' is directly reachable from x" if there exists $a \in A(x)$ such that $p_a(x'|x) > 0$, and "x' is reachable from x", written $x \to x'$, if there exists a sequence of states $x = x_1, x_2, \ldots, x_{n+1} = x'$ such that x_{k+1} is directly reachable from x_k for every $x = 1, \ldots, n$. Formally, we are assuming that X is irreducible according to relation $\to$. If this was not the case, our results would still apply separately to every class $Y \subseteq X$ which is recurrent according to this relation. States which are transient with respect to $\to$ are irrelevant for our analysis since they are never visited in the long run, no matter which choices the agents make.

Lemma 6. *$\mu \in \mathbb{R}_+^{A\times X}$ is a steady state outcome if and only if it satisfies*

$$\sum_{a,x} \mu(a,x) = 1$$

and

$$\forall x \in X: \quad \sum_{a\in A(x)} \mu(a,x) = \sum_{y} \sum_{a\in A(y)} p_a(x|y)\,\mu(a,y).$$

In what follows we denote by $\mathscr{M}$ the set of steady state outcomes, and by $\mathscr{M}_0$ the subset of its elements that correspond to "pure policies":

Definition 7. Say that $\mu \in \mathscr{M}$ is pure if, for every state x, there is at most one action $a \in A(x)$ such that $\mu(a,x) > 0$.

That is, μ is pure if and only if it can be associated with a policy that maps each state in a single action, and the associated system of choice probabilities σ can be chosen so that

$$\forall a,x: \quad \sigma(a|x) \in \{0,1\}.$$

This is in contrast with the general case, in which choice at each state is allowed to be "mixed", hence our terminology. Every outcome can be decomposed in terms of the weights it assigns to a finite set of pure outcomes. This follows from the fact that the set of outcomes is convex, hence it can be seen as the set of all convex combinations of its extreme points.[2]

Lemma 8. *$\mathscr{M}_0$ is the set of extreme points of $\mathscr{M}$.*

Since $\mathscr{M}_0$ might not be affinely independent, the representation of an outcome in terms of these extreme points in general is not unique. Instead, we can obtain a unique representation in terms of a barycentric coordinate system.[3]

Definition 9. A *barycentric coordinate system* is a maximal affinely independent subset of $\mathscr{M}_0$.

If $\{\mu_0, \mu_1, ..., \mu_N\}$ is a barycentric coordinate system, then for every $\mu \in \mathscr{M}$ there is a unique vector $\alpha \in \mathbb{R}^{N+1}$ such that $\sum_{n=0}^N \alpha_n = 1$ and $\sum_{n=0}^N \alpha_n \mu_n = \mu$.[4] In what follows we fix a coordinate system and represent each steady state outcoe in terms of its coordinates. The set of all coordinates corresponding to elements in $\mathscr{M}$ is denoted by $\mathscr{A}$, where it can be seen that

$$\mathscr{A} = \left\{ \alpha \in \mathbb{R}^{N+1}: \quad \sum_{n=0}^{N} \alpha_N = 1 \text{ and } \sum_{n=0}^{N} \alpha_n \mu_n \geq 0 \right\}.$$

[2] μ is an extreme point of $\mathscr{M}$ if it cannot be written as a convex combination of other elements in $\mathscr{M}$. It is a standard result in convex analysis that every element of a convex set can be written as a convex combination of its extreme points.

[3] $\{\mu_0, ..., \mu_N\}$ is affinely independent if there is no vector of weights $\alpha \neq 0$ such that $\sum_{n=0}^N \alpha_n \mu_n = 0$ and $\sum_{n=0}^N \alpha_n = 0$. It is a maximal affinely independent set if there is no $\mu_{N+1} \in \mathscr{M}_0$ such that $\{\mu_0, ..., \mu_N, \mu_{N+1}\}$ is affinely independent.

[4] We use the notation $\sum_{n=0}^N \alpha_n \mu_n$ for the vector defined by

$$\forall a,x: \quad \left(\sum_{n=0}^{N} \alpha_n \mu_n\right)(a,x) = \sum_{n=0}^{N} \alpha_n \mu_n(a,x).$$

2 The average stage payoff problem

In this section we describe the average stage payoff problem, which consists in chosing a stationary policy in order to maximize the long-run average of the expected future rewards. Agents can condition their behavior on the observation of state x and the realization of the preference shock ϵ, hence their behavior can be described by a function mapping these variables into a probability distribution over feasible controls. For convenience, we describe these mappings in terms of the joint distributions that they induce over actions and shocks:

Definition 10. A mixed stationary *policy* is a function π mapping each state x to a joint distribution over preference shocks and feasible controls $\pi(x) \in \Delta\left(A(x) \times \mathbb{R}^A\right)$ such that the marginal of $\pi(x)$ over $\mathbb{R}^A$ is $q(\cdot|x)$.

Conditional on x and a realization of ϵ, the probability according to which the agent chooses action a is given by $\pi(\epsilon, a|x) / q(\epsilon|x)$. Agents' objective is to maximize over all stationary policies the average payoff per stage

$$w_\pi(x_0, u) = \lim_{T\to\infty} \frac{1}{T+1} \mathrm{E}_\pi \sum_{t=0}^{T} \left(u_{a_t,x_t} + \epsilon_{t,a_t}\right) \tag{1}$$

for any given initial state x_0, where the expectation above is taken with respect to the distribution of future states, actions and shocks induced by π and p[5]. For each x, let $\sigma^\pi(x)$ denote the marginal of $\pi(x)$ over $A(x)$, and define

$$U_\pi = \left[\mathrm{E}_{\pi(x)}\left(u_{a,x} + \epsilon_a\right)\right]_{x\in X}.$$

Then σ^π is the system of conditional choice consistent with agents playing π, and since the (x, y)th element of $P^t_{\sigma^\pi}$ (P_{σ^π} to the tth power) is the t-step transition probability from x to y corresponding to σ^π, we can write 1 in matrix notation as:

$$w_\pi(u) = \lim_{T\to\infty} \left(\frac{1}{T+1} \sum_{t=0}^{T} P^t_{\sigma^\pi}\right) U_\pi.$$

Indeed, an important result regarding transition probabilities is that the limit in the preceding equation exists, and it is equal to the matrix of stationary distributions of P_{σ^π} conditional on each initial state (see references CIT). Hence it can be seen that the average payoff per stage 1 can be divided in a structural component and one component coming from unobserved heterogeneity:

$$w_\pi(x_0, u) = \sum_{x,a\in A(x)} \mu^{\pi,x_0}(a, x)\, u_{a,x} + \sum_x \mu^{\pi,x_0}(x)\, \mathrm{E}_{\pi(x)}(\epsilon_a) \tag{2}$$

where μ^{π,x_0} is the long-run outcome associated with σ^π when the system starts in state x_0.[6]

[5] Conditional on x_t, a_t, ϵ_t is distributed according to $\pi(x_t)$. Then, conditional on x_t, a_t, x_{t+1} is drawn according to $p_{a_t}(x_t)$.

[6] That is:

$$\forall a, x: \quad \mu^{\pi,x_0}(a, x) = \sigma^\pi(a|x) \left(\lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} P^t_{\sigma^\pi}\right)_{x_0,x}.$$

2.1 Representation

Our first goal is to write the problem in terms of long-run outcomes only, removing any explicit reference to policy functions, since the former are observed by the analyst while the latter are not. This will give a characterization of those outcomes that are consistent with a given vector u of structural stage payoffs. To do this, for every $\mu \in \mathcal{M}$ denote by σ^{μ} any system of chice probabilities that is consistent with it, which is uniquely defined at those states which are visited with positive probability under μ:

$$\forall x : \mu(x) > 0, \forall a \in A(x) \quad \sigma^{\mu}(a|x) = \frac{\mu(a,x)}{\mu(x)}.$$

At state x, consider the problem of optimal matching between shocks and actions compatibly with the choice probabilties at that state being given by $\sigma^{\mu}(x)$:

$$\begin{aligned} &\max_{\rho} \mathrm{E}_{\rho}(\epsilon_a) \\ &\text{s.t.}\rho \in \Pi_x\left(\sigma^{\mu}(x), q(\cdot|x)\right) \end{aligned} \tag{3}$$

where $\Pi_x\left(\sigma^{\mu}(x), q(\cdot|x)\right)$ denotes the set of probability distributions over $A(x) \times \mathbb{R}^A$ with marginals given by $\sigma^{\mu}(x)$ and $q(\cdot|x)$.

Our first result builds on the contribution of a recent literature (Galichon & coauthors CIT) which has shown that the value of this problem can be written in closed form in terms of $\sigma^{\mu}(x)$. Define the *social surplus function* at state x, $W_x : \mathbb{R}^{A(x)} \to \mathbb{R}$, by

$$\forall v \in \mathbb{R}^{A(x)} : \quad W_x(v) = \mathrm{E}_{q(\cdot|x)}\left[\max_{a \in A(x)} v_a + \epsilon_a\right].$$

That is, $W_x(v)$ is the expected indirect utility of a representative agent who faces a static discrete choice problem when the mean payoffs are given by v and the additive preference shocks are distributed according to $q(\cdot|x)$. The above-mentioned literature has shown that Problem 3 has value $-W_x^*(\sigma^{\mu}(x))$, where $W_x^* : \mathbb{R}^{A(x)} \to \mathbb{R}$ is the *convex conjugate* of W_x (see Galichon CIT).

Now note that problem of maximizing 2 can be divided in two nested problems. Conditionally on generating a long-run outcome μ, the optimal policy must solve the optimal matching Problem 3 at each state. The outer problem then selects the best long-run outcome among those compatible with the initial state, assuming an optimal solution of 3 for each μ.

Definition 11. Say that μ is rationalized by u if there exists an optimal policy π and an initial state x_0 such that $\mu = \mu^{\pi, x_0}$.

Theorem 12. *There exists a unique long-run outcome which is rationalized by u, given by*

$$\mu(u) = \arg\max_{\mu \in \mathcal{M}} \mu \cdot u - \sum_x \mu(x) W_x^*\left(\sigma^{\mu}(x)\right). \tag{4}$$

Uniqueness is driven by Assumption 2, which implies that the objective function is strictly concave. If this assumption were dropped there could be multiple long run outcomes rationalized by u, corresponding to the multiple solutions of 4. Assumption 3 instead guarantees that the problem is independent of the initial state. In particular, we can denote the optimal average payoff per stage by

$$w(u) = \max_{\mu \in \mathcal{M}} \mu \cdot u - \sum_x \mu(x) W_x^*\left(\sigma^{\mu}(x)\right).$$

2.2 Optimality conditions and relation with the discounted problem

Our second result introduces the analogue of Bellman's equation for the unidiscounted problem. Similarly to the well studied version of the problem without unobserved heterogeneity, all the solutions to this equation can be identified with the optimal average payoff per stage and a vector of differential valuations of each state (see Betrsekas CIT). The existence of a solution to these equation is guaranteed by the existence of a solution of Problem 4, while uniqess will be shown in Section CIT.

Theorem 13. *A scalar $w \in \mathbb{R}$, a vector $\psi \in \mathbb{R}^X$ and a long-run outcome $\mu \in \mathcal{M}$ satisfy*

$$\forall x \in X: \quad \psi(x) + w = E_{q(\cdot|x)}\left[\max_{a \in A(x)} u(a,x) + p_a(x)\psi + \epsilon_a\right] \tag{5}$$

and

$$\forall x \in X, a \in A(x): \quad \sigma^\mu(a|x) = q\left(a = \max_{a \in A(x)} u(a,x) + p_a(x)\psi + \epsilon_a | x\right) \tag{6}$$

if and only if (μ, w, ψ) is an optimal dual pair of Problem 4. Moreover, condition 5 implies that $w = w(u)$.

This result suggests one way of solving Problem 4 by finding a solution to 5 and then computing the optimal outcome from 6. In Section CIT we provide an algorithm to do so.

Intuitively, ψ in previous lemma can be thought as the vector of relative valuations of different states. To see this, notice that we can always normalize $\psi(x_0) = 0$ for an arbitrary reference state x_0 leaving Condition 5 satisfied. To gain further intuition, we relate (ψ, w) to the limit of the value function of the corresponding discounted problem. To do this, for each $\beta \in (0,1)$, let us consider the value V^β of a stationary policy π for the corresponding β-discounted problem. This is the unique solution of the Bellman equation

$$\forall x \in X: \quad V_x^\beta = \mathrm{E}_{q(\cdot|x)}\left[\max_{a \in A(x)} u_{a,x} + \epsilon_a + \beta p_a(x) V^\beta\right].$$

The optimal choice probabilities, given by

$$\forall x, a \in A(x): \quad \sigma^\beta(a|x) = q\left[a = \max_{a' \in A(x)} u_{a,x} + \epsilon_a + \beta p_a(x) V^\beta | x\right],$$

are associated with a unique optimal long-run outcome[7], which we denote $\mu^\beta(u)$.

Theorem 14. *Let ψ, w be as in Lemma 13. Then, as $\beta \to \infty$,*

$$\mu^\beta(u) \to \mu(u)$$

$$\forall x \in X: \quad (1-\beta) V_x^\beta \to w$$

$$\forall x, y \in X: \quad V_x^\beta - V_y^\beta \to \psi_x - \psi_y.$$

[7]Assumption 2 implies that $\sigma^\beta(a|x) > 0$ for every $x, a \in A(x)$. Hence Assumption 3 implies that P_{σ^β} is irreducible, so that it has a unique stationary distribution.

Hence the average payoff per stage problem approximates the discounted problem as the discount factor increases, and in particular, $w(u)$ can be interpreted as the limit of the average discounted payoff, which is independent of the initial state, and ψ is the limit of agents relative valuations for each state.

2.3 Computation

In this section we provide an algorithm for solving Problem 4 and prove its convergence. To our knowledge the results in this section are new, although they rely on the full support Assumption 2, and are not necessarily true when this assumption is dropped. Alternatively, relative value iteration can be used under more restrictive assumptions (See Bertsekas CIT).

The algorithm iterates on the vector of relative valuations ψ, normalized so that $\psi_{x_0} = 0$ for some arbitrary reference state x_0. Define the operator $T : \mathbb{R}^X \to \mathbb{R}^X$ by

$$\left(T\psi\right)_x = \mathrm{E}_{q(\cdot|x)}\left[\max_{a\in A(x)} u_{a,x} + \epsilon_a + p_a(x)\,\psi\right] - \mathrm{E}_{q(\cdot|x_0)}\left[\max_{a\in A(x_0)} u_{a,x_0} + \epsilon_a + p_a(x_0)\,\psi\right].$$

If ψ is a fixed point of T, defining

$$w = \mathrm{E}_{q(\cdot|x_0)}\left[\max_{a\in A(x_0)} u_{a,x_0} + \epsilon_a + p_a(x_0)\,\psi\right]$$

it can be seen that ψ, w are the optimal dual variables of Problem 4, hence they can be used to compute the optimal outcome as in Lemma 13.

If $\lambda \in (0,1)$ and K is an integer, define the iterates $\left(\psi^n(K,\lambda,x_0)\right)_{n=0}^{\infty}$ by:

$$\forall n \ge 0: \quad \psi^{n+1}(K,\lambda,x_0) = \lambda T^K \psi^n(K,\lambda,x_0) + (1-\lambda)\,\psi^n(K,\lambda,x_0).$$

where $\psi^0(K,\lambda,x_0)$ is initialized to an arbitrary vector in $\mathbb{R}^X$ with zero x_0-coordinate. We show that, for suitable values of K, this sequence converges to a fixed point of T as $n \to \infty$. To this end, define matrix $G \in \{0,1\}^{X\times X}$ by

$$G_{x,x'} = \begin{cases} 1 & \exists\, a \in A(x): \quad p_a\left(x'|x\right) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Hence $G_{x,x'} > 0$ if and only if x' is reachable with positive probability from x for an approriate choice of the control. Similarly it is easy to see that, for every $K \in \mathbb{N}$, $G^K_{x,x'} > 0$ if and only if x' is reachable with positive probability from x in exactly K steps. The following is an easy consequence of Assumption 3:

Lemma 15. *There exists an integer K such that*

$$\forall x \in X: \quad G^K_{x,x_0} > 0.$$

Theorem 16. *Fix an arbitrary state x_0. The operator T has a unique fixed point ψ^*. Moreover, let K be as inprevious Lemma. Then, for every $\lambda \in (0,1)$:*

$$\lim_{n\to\infty} \psi^n(K,\lambda,x_0) = \psi^*.$$

The proof consists in showing that T^K is non-expansive mapping with respect to a particular norm, and then applying a theorem due to Ishikawa CIT to show that this implies that $\psi^n(K,\lambda,x_0)$ converges to a fixed point ψ^* of T^K. We then prove that T^K has a unique fixed point, and that this implies that ψ^* is the unique fixed point of T.

This result gives some guideline on how to choose K and the reference state x_0. First, for given x_0, K should be chosen to be the smalles among the integers identified in Lemma 15. Second, one should pick x_0 that is associated with the smalles value of K. The following Corollary provides a special case of the result in a special case which is common in economic applications:

Corollary 17. *Suppose that there exists a state x_0 such that, for every $x \in X$, there exists $a \in A(x)$ such that $p_a(x_0|x) > 0$. Then*

$$\lim_{n\to\infty} \psi^*(1,\lambda,x_0)$$

for every $\lambda \in (0,1)$.

2.4 Identification and duality

In this section we characterize the set of structural stage payoff vectors that are identified from the observation of a long-run outcome, that is, the set of all us such that $\mu(u) = \mu^*$. In analogy with static discrete choice, we show that this set can be characterized in terms of the gradient of the convex conjugate of a long-run payoff function (see Galichon CIT). First of all, note that the solution to Problem 4 depends on u only through the set of long-run averages

$$\left\{ \sum_{a,x} \mu(a,x)\, u(a,x) : \quad \mu \in \mathcal{M} \right\}. \tag{7}$$

Hence payoff vectors producing the same set of long-run averages cannot be told apart from the observation of μ^*. Moreover, addying or subtracting a constant from all these averages does not affect the solution of Problem 4, hence we can only hope to identify the set 7 up to a constant. The first result of this section shows that this is indeed the empirical content of Problem 4. The result relies on the well known Arcididiacono-Miller CIT inversion, stating that the conditional choice probabilities at each state uniquely identify the differences in the dynamic payoffs associated to each action at that state (see also Chiong, Galichon and Shum CIT). Applied to our setting this yelds the following:

Lemma 18. *For each state x, there exists a unique vector $v_x \in \mathbb{R}^{A(x)}$ satisfying*

$$E_{q(\cdot|x)}\left[\max_{a\in A(x)} v_{a,x} + \epsilon_a\right] = 0 \text{ and } \forall a \in A(x): \quad \sigma^{\mu^*}(a|x) = q\left(a = \max_{a\in A(x)} v_{a,x} + \epsilon_a | x\right)$$

Theorem 19. *Let v be as in previous lemma. Then $\mu(u) = \mu^*$ if and only if*

$$\forall \mu \in \mathcal{M}: \quad \sum_{a,x} \mu(a,x)\, u_{a,x} = \sum_{a,x} \mu(a,x)\, v_{a,x} + w(u). \tag{8}$$

Proof. If $\mu^* = \mu(u)$, then taking ψ, w be as in Theorem 13 we must have

$$\forall x,\, a \in A(x): \quad v_{a,x} = u_{a,x} + p_a(x)\psi - \psi(x) - w.$$

Averaging both sides with respect to $\mu \in \mathscr{M}$ and exploiting the stationarity of μ yelds 8. Conversely, notice that $\mu(v) = \mu^*$ (to see this, take $\psi = 0$ and $w = 0$ in Theorem 13). Hence 8 implies that $\mu(u) = \mu^*$, since the solution to Problem 4 only depends on the differences between the average payoffs per stage associated to each long-run outcome. $\square$

In order to better understand the structure of the identified set, we exploit the representation of steady state outcomes in terms of their barycentric coordinate system identified in Section 1.1. Let $v \in \mathbb{R}^{N+1}$ denote a vector of average structural stage payoffs associated to each coordinate:

$$\forall n = 0, 1, ..., N: \quad v_n = \sum_{a,x} \mu_n(a,x)\, u_{a,x}.$$

Notice that the solution to Problem 4 depends on the structural payoffs only through the value of v. Indeed, its objective function can be written as

$$\begin{aligned}&\sum_{a,x} \mu(a,x)\, u_{a,x} - \sum_{x} \mu(x)\, W_x^*\left(\sigma^{\mu}(x)\right)\\ =&\sum_n \alpha_n v_n - w^*(\alpha)\end{aligned}$$

where α is the vector of coordinates of μ and

$$\forall \alpha \in \mathscr{A}: \quad w^*(\alpha) = \sum_x \sum_n \alpha_n \mu_n(x)\, W_x^*\left(\sigma^{\sum_n \alpha_n \mu_n}(x)\right).$$

Hence Problem 4 can be framed as the optimal choice of a vector of coordinates, and its value is a function of v, that is

$$w(v) = \max_{\alpha \in \mathscr{A}} \alpha \cdot v - w^*(\alpha).$$

and the solution of this problem, which we denote $\alpha(v)$, is the vector of coordinates of the optimal steady state outcome. In conclusion, the mathematical structure of the model is very similar to that of a static discrete choice: Problem 4 can be seen as an optimal assignment of weights to a finite set of coordinates, whose solution depends only on the finite-dimensional vector of payoffs associated to each coordinate. This observation leads us to the following result, whose equivalent in static discrete choice is known as the Williams–Daly–Zachary Theorem:

Theorem 20. *w and w^* are convex conjugates. In particular, for every $v \in \mathbb{R}^{N+1}$ and $\alpha \in \mathscr{A}$, the following are equivalent:*

1. $\alpha = \alpha(v)$
2. $\partial w(v) = \{\alpha\}$
3. $\partial w^*(\alpha) = \{v + k: \quad k \in \mathbb{R}\}$

Hence the identified set average stage payoffs is the gradient of the conjugate of the social welfare function evaluated at the obseved outcome. This observation will be useful in Section CIT, where we consider the problem of inversion in a model with persistent heterogeneity.

3 The logit model

In economics applications of dynamic discrete choice models, the most popular family of distributions of idiosyncratic preference shocks is the type-I extreme value, or logit. In this section we review this model in terms of the results of previous sections, and we characterize it in terms of the assumptions that it imposes on counterfactuals: we show that the average payoff per stage logit model is axiomatized by a dynamic analogoue of independence of irrelevant alternatives, an axiom characterizing the static logit model (Luce CIT).

If shocks are i.i.d. according to a standard logit [8], the social surplus function and its conjugate take the functional form[9]

$$W_x(v) = \log \sum_{a \in A(x)} \exp v_a \text{ and } W_x^*(\sigma(x)) = \sum_{a \in A(x)} \sigma(a|x) \log \sigma(a|x)$$

and the choice probablities have the simple expression

$$q\left(a = \max_{a \in A(x)} v_a + \epsilon_a | x\right) = \exp\left(v_a - W_x(v)\right).$$

Hence Problem 4 becomes

$$\max_{\mu \in \mathscr{M}} \sum_{x, a \in A(x)} \mu(a,x)\, u(a,x) - \sum_{a \in A(x)} \mu(a,x) \log \frac{\mu(a,x)}{\mu(x)}$$

Its optimality conditions 5 and 6 can be written as

$$\psi(x) + w = \log \sum_{a \in A(x)} \exp\left[u_{a,x} + p_a(x)\psi\right]$$
$$\log \sigma^{\mu^*}(a|x) = u_{a',x} + p_{a'}(x)\psi - w - \psi(x)$$

and the average payoffs are identified from

$$\forall \mu \in \mathscr{M}: \quad \sum_{x, a \in A(x)} \mu(a,x) \log \sigma^*(a|x) = \sum_{x, a \in A(x)} \mu(a,x)\, u_{a,x} - w(u). \tag{9}$$

This last expression is the one that we leverage for axiomatizing the model. The intuition is the same of the static case: removing certain alternatives from the choice set should not affect the relative likelyhood of those choices which are not removed. To make this observation precise, we need to clarify what does it mean to remove an outcome from the choice set. Once we do that, we will see that this property actually characterizes the dynamic logit model with no discounting.

[8] By standard logit we mean a mean zero, variance $\pi^2/6$ logit distribution. The choice of the mean is without loss of generality since this does not afect choice probabilities, while the choice of the variance corresponds to a normalization of the structural stage payoffs.

[9] See references CIT

3.1 Independence of irrelevant alternatives

The universe of all states and actions that can be chosen in each state is summarized by the pair $(X, A(\cdot))$. Once we limit the available choices at some particular states we obtain different sub-problems, corresponding to the maximal classes of states that are connected under the new feasible-action correspondence in the sense of Assumption 3:

Definition 21. A sub-problem is a pair $P = \left(X^P, A^P(\cdot)\right)$ given by a subset of states $X^P \subseteq X$ and a correspondence $A^P : X^P \to 2^A$ such that $\left(X^P, A^P(\cdot)\right)$ is "connected" in the sense of Assumption 3 and we have:

$$\forall x \in X^P : B(x) \subseteq A(x),$$
$$\forall x \in X^P, a \in A^P(x) : \quad p_a\left(y|x\right) > 0 \Rightarrow x \in X^P.$$

The collection of all sub-problems is denoted by $\mathscr{P}$. For each $P \in \mathscr{P}$, we can define the set long-run outcomes, $\mathscr{M}^P$, and of pure long-run outcomes, $\mathscr{M}_0^P$, as in Section CIT. Notice that we have

$$\mathscr{M}^P \subseteq \mathscr{M} \text{ and } \mathscr{M}_0^P \subseteq \mathscr{M}_0.$$

Suppose that we observe the outcome of every sub-problem: for each $P \subseteq \mathscr{P}$, μ^P denotes the long-run outcome observed under sub-problem P. Let $\boldsymbol{\mu} = \left(\mu^P\right)_{P \in \mathscr{P}}$.

Definition 22. $\boldsymbol{\mu}$ is rationalized by the logit model if there exists a vector $u \in \mathbb{R}^A$ such that, for all $P \in \mathscr{P}$:

$$\mu^P = \arg\max_{\mu \in \mathscr{M}^P} \sum_{x, a \in A^P(x)} \mu(a,x)\, u(a,x) - \sum_x \mu(x) \sum_{a \in A^P(x)} \sigma^\mu(a|x) \log \sigma^\mu(a|x).$$

For each $P \in \mathscr{P}$ and $\mu \in \mathscr{M}_0^P$, define the likelihood of μ^P given μ by

$$L\left(\mu^P|\mu\right) = \Pi_{x, a \in A(x)} \sigma^{\mu^P}(a|x)^{\mu(a,x)}.$$

Expression 9 applied to different sub-problems implies that the ratios of the likelyhoods of the observed outcome conditional on two different pure outcomes are independent of which other choices are available, similarly to Luce's model of static discrete choice. It can be shown that this property fully characterizes the dynamic logit model with no discounting:

Theorem 23. *$\boldsymbol{\mu}$ is rationalized by the logit model if and only if, for every $P, P' \in \mathscr{P}$ and every $\mu, \mu' \in \mathscr{M}_0^P \cap \mathscr{M}_0^{P'}$:*

$$\frac{L\left(\mu^P|\mu\right)}{L\left(\mu^P|\mu'\right)} = \frac{L\left(\mu^{P'}|\mu\right)}{L\left(\mu^{P'}|\mu'\right)}.$$

4 Mixing and inversion

Moving from the results in the end of previous section, we consider the possibility of mixing i.i.d. models in order to allow for richer substitution patterns. In the mixed model, agents differ in their individual taste parameters, and the observed steady state outcomes are derived as the aggregate outcome of agents decisions. Formally, consider an economy populated by a unit measure of agents, and suppose that the average stage payoffs associated to each coordinate are randomly distributed among the population. Write the vector of average stage payoffs for an arbitrary agent as the sum $\delta + \eta$ of an average component $\delta \in \mathbb{R}^{N+1}$ and a disturbance η which is distributed among the population according to some probability distribution $\Pr(\cdot|\theta)$ with mean zero, where θ is a parameter indicizing the distribution of the random component. The coordinates of the observed steady state outcome are obtained by aggregating them across agents:[10]

$$\alpha(\delta,\theta) \equiv \int \alpha(\delta+\eta) \Pr(\mathrm{d}\eta|\theta).$$

In this section we fix a parameter θ and consider the problem of inverting the function $\alpha(\cdot,\theta) : \mathbb{R}^{N+1} \to \mathscr{A}$ (where inversion is of course "up to a constant", since we have $\alpha(\delta,\theta) = \alpha(\delta+k,\theta)$ for each $k \in \mathbb{R}$). In the context of demand functions, invertibility is usually established by appealing to the gross substitutes property (see Berry, Gandhi and Halie CIT). Applied to the function $\alpha(\cdot,\theta)$ this property requires that, for each $n = 0,1,\dots,N$, $\alpha_n(\cdot,\theta)$ is increasing in δ_n and decreasing in δ_j for $j \neq n$. It is well known that when this property is satisfied some sort of tatonnement algorithm can usually be implemented to perform inversion. Unfortunately this is not our case, since it can be shown by example that $\alpha(\cdot|\theta)$ generally does not satisfy this property (see section CIT for an example). However we can still exploit the duality identified in Section 2.4. To this end, define the average surplus function

$$w(\delta,\theta) = \int w(\delta+\eta) \Pr(\mathrm{d}\eta|\theta).$$

By differentiating with respect to δ we obtain

$$\partial_\delta w(\delta,\theta) = \int \partial w(\delta+\eta) \Pr(\mathrm{d}\eta|\theta) = \int \alpha(\delta+\eta) \Pr(\mathrm{d}\eta|\theta) = \alpha(\delta,\theta)$$

implying that δ solves

$$\min_{\delta'} w(\delta',\theta) - \delta' \cdot \alpha(\delta,\theta). \tag{10}$$

Given this observation, the next Lemma is sufficient to establish invertibility:

Lemma 24. *If $\delta' \notin \{\delta + k : \quad k \in \mathbb{R}\}$ then, for every $\lambda \in (0,1)$, $w(\lambda\delta + (1-\lambda)\delta',\theta) < \lambda w(\delta,\theta) + (1-\lambda) w(\delta',\theta)$*

Theorem 25. *$\alpha(\delta,\theta) = \alpha(\delta',\theta)$ if and only if $\delta' = \delta + k$ for some $k \in \mathbb{R}$*

Proof. If $\delta' \notin \{\delta + k : \quad k \in \mathbb{R}\}$ and $\alpha(\delta,\theta) = \alpha(\delta',\theta) \equiv \alpha$ then

$$w(\lambda\delta + (1-\lambda)\delta',\theta) - (\lambda\delta + (1-\lambda)\delta') \cdot \alpha < \lambda[w(\delta,\theta) - \delta\cdot\alpha] + (1-\lambda)[w(\delta',\theta) - \delta'\cdot\alpha]$$

contradicting either $\alpha(\delta,\theta) = \alpha$ or $\alpha(\delta',\theta) = \alpha$. □

[10] The linear aggregation is justified by the steady state assumption. Indeed, this corresponds to a linear aggregation of the steady state outcomes across the population.

Having established invertibility, inversion can be performed by standard convex optimization techniques. For example, applying gradient descent to Problem 10 yelds the simple tatonnement algorithm

$$\delta^{n+1} = \delta^{n} - \gamma_n \left(\alpha \left(\delta^{n}, \theta \right) - \alpha \left(\delta, \theta \right) \right)$$

where the payoff of each coordinate is decreased proportionally to its "excess demand".