Nonparametric Identification of Behavioral Responses to Counterfactual Policy Interventions in Dynamic Discrete Decision Processes

Victor Aguirregabiria[1]

Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215

August, 2004

Abstract: This paper deals with identification in Markov dynamic discrete decision processes. It shows the nonparametric identification of the behavioral responses to counterfactual policy interventions that modify the one-period utility function.

Keywords: Dynamic discrete decision processes; Identification; Counterfactual policy experiments.

JEL classification: C13, C35.

1 Introduction

In the context of dynamic discrete decision models, the difference between the utilities of two choice alternatives is not identified even when the discount factor and the distribution of unobservables are known (see Rust, 1994, pp. 3125-3130). This result contrasts with the case of static discrete choice models, where utility differences are identified and they can be used to evaluate behavioral responses to counterfactual changes in the utility function. Although differences between conditional choice value functions are identified in dynamic models (see Hotz and Miller, 1993, and Magnac and Thesmar, 2002), these value functions cannot be used to evaluate the behavioral effects of changes in one-period utilities. This paper shows the nonparametric identification of the difference between the value of choosing always the same alternative and the value of deviating one period from this policy. We prove that, given these values, one can identify the behavioral responses to policy interventions that modify one-period utilities.

2 Model

Time is discrete and indexed by t. At every period t an agent observes the vector of state variables s_t and chooses an action $a_t \in A = \{1, 2, ..., J\}$ to maximize the expected and discounted sum of current and future utilities $E\left[\sum_{j=0}^{\infty} \beta^j \; U(a_{t+j}, s_{t+j}) \mid a_t, s_t\right]$, where $\beta \in (0,1)$ is the discount factor, and $U(a_t, s_t)$ represents the utility at period t. The agent has uncertainty on future values of state variables. His beliefs about future states can be represented by a transition probability $p(s_{t+1}|a_t, s_t)$. These beliefs are *rational* in the sense that they are the true transition probabilities of the state variables. Let $V(s_t)$ be the value function of this problem. By Bellman principle of optimality this value function is the unique fixed-point of the contraction mapping:

$$V(s_t) = \max_{a \in A} \left\{ u(a, s_t) + \beta \int V(s_{t+1}) \;\; p(ds_{t+1}|a, s_t) \right\} \tag{1}$$

The optimal decision rule $\alpha(s_t)$ is the *arg max*$_{a\in A}$ of the term in brackets. From the point of view of the observing researcher there are two types of state variables, $s_t = (x_t, \varepsilon_t)$, where the vector x_t is observable to

[1] Tel.: +1-617-353-9583; fax: +1-617-353-4449; E-mail address: vaguirre@bu.edu.

the econometrician and the vector ε_t is unobservable. The one-period utility is additive separable between observable and unobservable variables: $U(a_t, x_t, \varepsilon_t) = u(a_t, x_t) + \varepsilon(a_t)$, where $\varepsilon_t(a)$ is the a-th component of the vector of unobservable state variables $\varepsilon_t = \{\varepsilon_t(a) : a \in A\}$. We follow Rust (1994) and consider the following assumptions on the joint distribution of the state variables.

ASSUMPTIONS: (1) The transition probability of the state variables factors as $p(s_{t+1}|a_t, s_t) = g(\varepsilon_{t+1})$ $f(x_{t+1}|a_t, x_t)$*; (2)* g *is the density of* ε_t *and it is absolutely continuous with respect to the Lebesgue measure in* R^J*; and (3)* x_t *has support* $X = \{x_{(1)}, x_{(2)}, ..., x_{(M)}\}$*, where* M *is a finite integer.*

Define the *integrated* value function $S(x_t) \equiv \int V(x_t, \varepsilon_t) g(d\varepsilon_t)$. Taking into account the Bellman equation in (1), we have that:

$$S(x_t) = \int \max_{a \in A} \left\{ \; u(a, x_t) + \varepsilon_t(a) + \beta \textstyle\sum_{x' \in X} f(x'|a, x_t) \; S(x') \; \right\} \; g(d\varepsilon_t) \tag{2}$$

The right-hand side of this equation is a contraction mapping in the integrated value function, and therefore $S(.)$ is the unique fixed point of this mapping (see Rust et al., 2002). Define also the integrated optimal decision rules or *optimal choice probabilities* $P(a|x_t) \equiv \int I\{\alpha(x_t, \varepsilon_t) = a\} g(d\varepsilon_t)$. Finally, define the *conditional choice value functions* $v(a, x_t) \equiv u(a, x_t) + \beta \sum_{x' \in X} f(x'|x_t, a) \; S(x')$.

3 Nonparametric identification of utilities

Suppose that there is a population of individuals who behave according to the previous model. We have a random sample of n individuals from this population. In the sample we observe individuals' decisions at some period t, and observable state variables at periods t and $t+1$. We are interested in the nonparametric estimation of the one-period utilities $\{u(a, x) : a \in A, x \in X\}$. Given the time-homogeneous Markov structure of the model and Assumptions (1) to (3), we can identify nonparametrically from these data the choice probabilities $\{P(a|x) : (a, x) \in A \times X\}$ and the transition probabilities $\{f(x'|x, a) : (a, x, x') \in A \times X \times X\}$.

The structure of the model implies two sets of restrictions on one-period utilities (see Magnac and Thesmar, 2002). The first set of restrictions comes from Hotz-Miller invertibility Proposition (Hotz and Miller, 1993). This Proposition establishes that there is a one-to-one relationship between the vector of value differences $\tilde{v}(x_t) \equiv \{v(a, x_t) - v(J, x_t) : a \in A_{-J}\}$ and the vector of choice probabilities $P(x_t) \equiv \{P(a|x_t) : a \in A_{-J}\}$, where $A_{-J} = \{1, 2, ..., J-1\}$. Let $Q(.)$ be this one-to-one mapping such that $\tilde{v}(x_t) = Q(P(x_t))$, and let $Q(a, P(x_t))$ be the $a-th$ element of this mapping, such that $v(a, x_t) - v(J, x_t) = Q(a, P(x_t))$. Taking into account the definition of v at the end of section 1, we can write these restrictions in matrix form as:

$$u(a) - u(J) + \beta \; (F(a) - F(J)) \; S \;\; = \;\; Q(a, P), \tag{3}$$

where $u(a)$ is a vector with the M utilities associated with alternative a; $F(a)$ is the $M \times M$ matrix of transition probabilities of x conditional to the choice of alternative a; S is the $M \times 1$ vector with the values $S(x)$; and $Q(a, P)$ is the $M \times 1$ vector with values $Q(a, P(x))$. An important property of the mapping Q is that it depends on the distribution of the unobservables but not on any other primitive of the model (i.e., discount factor, utilities and beliefs).

The second set of restrictions comes from the integrated Bellman equation in (2). Taking into account that $E\left(\max_{a\in A}\ v(a,x_t)+\varepsilon_t(a)\right)=\sum_{a\in A}\Pr\left(a|x_t\right)\ E\left(v(a,x_t)+\varepsilon_t(a)\mid x_t,\alpha(s_t)=a\right)$, we can re-write this Bellman equation in matrix form as:

$$S\ \ =\ \ \left(I-\beta\ \bar{F}\right)^{-1}\left(\ \bar{u}(P)+\bar{e}(P)\ \right) \tag{4}$$

$\bar{F}=\sum_{a\in A}P(a)*F(a)$ is the $M\times M$ matrix of unconditional transition probabilities, where $P(a)$ is the vector of choice probabilities $\{P(a|x):x\in X\}$, and $*$ is the element-by-element or Hadamard product. $\bar{u}(P)=\sum_{a\in A}P(a)*u(a)$ is the $M\times 1$ vector of expected utilities. $\bar{e}(P)=\sum_{a\in A}P(a)*e(a,P)$ is the $M\times 1$ vector of expected epsilons, where $e(a,P)$ is the vector $\{e(a,P(x)):x\in X\}$ and $e(a,P(x))\equiv E(\varepsilon_t(a)|x_t=x,\alpha(s_t)=a)$. A corollary of Hotz-Miller invertibility Proposition is that the conditional expectations $e(a,P(x))$ depend on the set of choice probabilities $P(x)=\{P(a|x):a\in A\}$ and on the distribution of the unobservables, but they not depend on the discount factor, utilities or beliefs.

If we solve expression (4) into equation (3), we get that for any $a\in A$, $u(a)-u(J)+\beta\ (F(a)-F(J))\left(I-\beta\bar{F}\right)^{-1}(\ \bar{u}(P)+\bar{e}(P)\ )=Q(a,P)$. This system of $M(J-1)$ equations represents all the restrictions that the model imposes on one-period utilities. It is straightforward to show that, without further restrictions, the utility differences $\{u(a)-u(J):a\in A_{-J}\}$ are not identified. Instead, we consider here the identification of the following set of value differences:

$$\tilde{u}(a)\equiv\left\{u(a)-\beta\ F(a)\ \ (I-\beta\ F(J))^{-1}\,u(J)\right\}-\left\{u(J)-\beta\ F(J)\ \ (I-\beta\ F(J))^{-1}\,u(J)\right\} \tag{5}$$

In the right-hand-side of this expression, the second term in brackets is a vector with the expected present values of choosing alternative J now and in the future. The first term in brackets is a vector with the present values of choosing alternative a today, and then choosing alternative J forever in the future. Therefore, $\tilde{u}(a)$ is the vector of values of deviating one period from the policy of choosing always alternative J.

PROPOSITION 1: Suppose that the discount factor and the distribution of unobservables are known. Then, the values $\{\tilde{u}(a):a\in A\}$ are nonparametrically identified. For any $a\in A$:

$$\tilde{u}(a)=Q(a,P)-\beta\ \ (F(a)-F(J)\ )\,(I-\beta\ F(J))^{-1}\left(\ \bar{Q}(P)+\bar{e}(P)\ \right), \tag{6}$$

where $\bar{Q}(P)=\sum_{a\in A}P(a)*Q(a)$.

Proof: If we multiply (element-by-element) the system of equations (3) by $P(a)$, we sum the result over a, and we solve for $\bar{u}(P)$, we have that: $\bar{u}(P)=u(J)+\bar{Q}(P)+\beta\ (F(J)-F)\ S$. Solving this expression into (4), rearranging terms, and taking into account that $(I-\beta F(J))$ is a non-singular matrix, we get: $S=(I-\beta\ F(J))^{-1}\left(\ u(J)+\bar{Q}(P)+\bar{e}(P)\ \right)$. Solving this expression in (3) and taking into account that $I+\beta F(J)\,(I-\beta F(J))^{-1}$ is equal to $(I-\beta F(J))^{-1}$, we get:

$$u(a)-u(J)+\beta\ \ (F(a)-F(J))\ \ (I-\beta\ F(J))^{-1}\left(\ u(J)+\bar{Q}(P)+\bar{e}(P)\ \right)=Q(a,P) \tag{7}$$

Rearranging terms and using the definition of $\tilde{u}(a)$, we obtain equation (6). The elements in the right hand side of this equation depend only on the discount factor, the distribution of the unobservables, choice probabilities, and transition probabilities. Therefore, under the conditions in the Proposition, $\tilde{u}(a)$ is identified.■

4 Counterfactual policy experiments

Proposition 2 shows that knowledge of the values $\{\tilde{u}(a) : a \in A\}$ can be used to identify the behavioral responses to policy interventions that modify one-period utilities.

PROPOSITION 2: Consider a policy intervention that modifies one-period utilities such that utilities after the intervention are $u^*(a) = u(a) + d(a)$. *Utility levels* $u(a)$ *and* $u^*(a)$ *are unknown, but the intervention* $d(a)$ *is known to the econometrician. Suppose that the discount factor, the distribution of unobservables, and the values* $\{\tilde{u}(a) : a \in A\}$ *are also known. Then, the (counterfactual) optimal choice probabilities after the intervention,* $P^* \equiv \{P^*(a) : a \in A\}$, *are identified. More specifically,* P^* *is the unique fixed point of a mapping* $\Phi(P) = \{\Phi(a, P) : a \in A_{-J}\}$ *such that:*

$$\Phi(a, P) = \int 1\left\{a = \arg\max_{k\in A} \left(\tilde{u}^*(k) + \beta\ F(k)\ \left(\bar{Q}(P) + \bar{e}(P)\right) + \varepsilon(k)\ \right)\right\} g(d\varepsilon)\ , \tag{8}$$

where $1\{.\}$ *is the indicator function.*

Proof: First, by definition:

$$\tilde{u}^*(a) \quad = \quad \tilde{u}(a) + \left\{d(a) - \beta\ F(a)\ \left(I - \beta\ F(J)\right)^{-1} d(J)\right\} - \left\{d(J) - \beta\ F(J)\ \left(I - \beta\ F(J)\right)^{-1} d(J)\right\} \tag{9}$$

And it is clear from this expression that the values $\tilde{u}^*(a)$ are known to the econometrician. Second, by Proposition 1(a) in Aguirregabiria and Mira (2002), the vector of choice probabilities P^* is the unique fixed point of a mapping $\Psi(P) = \{\Psi(a, P) : a \in A_{-J}\}$ such that:

$$\Psi(a, P) = \int 1\left\{a = \arg\max_{k\in A} \left(u^*(k) + \beta\ F(k)\ \left(I - \beta\bar{F}^*\right)^{-1} \left(\bar{u}^*(P) + \bar{e}(P)\right) + \varepsilon(k)\ \right)\right\} g(d\varepsilon) \tag{10}$$

Taking into account that (see the proof of Proposition 1 above) $\left(I - \beta\bar{F}^*\right)^{-1} \left(\bar{u}^*(P^*) + \bar{e}(P^*)\right) = \left(I - \beta F(J)\right)^{-1}$ $\left(u(J) + \bar{Q}(P^*) + \bar{e}(P^*)\right)$, it is straightforward to show that $\Psi(P^*) = \Phi(P^*)$. Therefore, the vector of optimal choice probabilities P^* is a fixed point of the mapping Φ. It remains to show that P^* is the unique fixed point of Φ. Suppose that Φ has two fixed points, say P_1^* and P_2^*. That would imply that there are two vectors of values that solve the the integrated Bellman equation: i.e., $S_1^* = \left(I - \beta F(J)\right)^{-1} \left(u^*(J) + \bar{Q}(P_1^*) + \bar{e}(P_1^*)\right)$ and $S_2^* = \left(I - \beta F(J)\right)^{-1} \left(u^*(J) + \bar{Q}(P_2^*) + \bar{e}(P_2^*)\right)$. However, this is not possible because the integrated Bellman equation is a contraction mapping. Therefore, P^* is the unique fixed point of Φ.

REFERENCES:

Aguirregabiria, V., and P. Mira (2002): "Swapping the nested fixed point algorithm: A class of estimators for Markov decision models," *Econometrica,* 70(4), 1519-43.

Hotz, J., and R. A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models," *Review of Economic Studies,* 60, 497-529.

Magnac, T. and D. Thesmar (2002): "Identifying dynamic discrete decision processes". *Econometrica*, 70(2), 801-816.

Rust, J. (1994): "Structural estimation of Markov decision processes," in R. E. Engle and McFadden (eds.) *Handbook of Econometrics Volume 4*, North-Holland.

Rust, J., J. F. Traub, and H. Wozniakowski (2002): "Is There a Curse of Dimensionality for Contraction Fixed Points in the Worst Case?," *Econometrica*, 70(1), 285-329.