Optimal Taxation and Discrete Choice*

Laurence Ales, Christopher Sleet

Tepper School of Business, Carnegie Mellon University

Preliminary!

This Version: February 12, 2019.

Abstract

In this paper we derive optimal tax equations for discrete choice economies with smooth underlying shock distributions and potentially unstructured choice sets. This approach expands the set of applications for which optimal tax equations are available, unifies existing results on income and commodity taxation and links optimal tax analysis to the large literature on estimating discrete choice models. The optimal tax equations that emerge highlight the role of semi-elasticities of choice probabilities to utility perturbations in shaping optimal taxes and, in particular, the implied willingness of agents to substitute from high to low taxed choices as a force dampening tax rates. ***JEL Codes:*** *H24, H31, R13*

Key Words: *Local Shocks, Discrete Choice and Optimal Policy*

1 INTRODUCTION

Optimal income tax formulas capture the tradeoff between the distributional benefits and the economic distortions associated with income tax rate variation at an optimum. Following Mirrlees (1971); Saez (2001) and many others, micro-foundations for these formulas are supplied by models that assume a population of agents distributed across "smooth" income choice problems indexed by an agent's preference

*E-mail: ales@cmu.edu (Ales); csleet@andrew.cmu.edu (Sleet).

or productivity type. In these problems, an agent's income choice is a differentiable function of the marginal income tax rate she faces and her elasticity of income choice with respect to this rate is well defined. Aggregate elasticities defining distortions are then built from these individual level micro-elasticities. This framework permits structural calibrations that relate the aggregate elasticities found in tax formulas to a type distribution and agent utility parameters, with the latter pinning down individual micro-level choice elasticities.[1] In this paper, instead of assuming that individual income choice problems are smooth in tax rates, we assume that the distribution of agents across income-generating activity choices is smooth in tax levels.[2] Underlying this assumption is the requirement that agents are smoothly distributed across a (sufficiently rich) set of preferences. Choice sets and the dependence of individual utilities on choice may otherwise be general and unstructured and rich substitution patterns across choices in response to tax variation are permitted. We formulate optimal tax relations in such environments connecting them to flexible, empirical specifications used in the discrete choice literature. We show that under a simple logit distribution specification in which preference heterogeneity is generated by additive choice-dependence Gumbel shocks, optimal taxes depend only upon agent income and not on other dimensions of choice (and depend linearly under further restrictions). In more general settings, cross-choice variation in the willingness of agents to migrate to high or low tax choice alternatives in the face of tax perturbations is central to determining the dependence of taxes on incomes and other dimensions of choice. We describe how to implement optimal tax calculations in mixed logit settings and do so for a rich spatial model. Finally, we extend the analysis to settings in which for informational or administrative reasons the policymaker is restricted to taxing only income.

The approach to optimal tax analysis pursued in this paper has several advantages relative to the more typical approach that assumes agents are distributed across smooth individual choice problems. First, it permits derivation of optimal tax formulas for environments in which agent activity choice sets are unstructured and individual choice problems are non-smooth in tax parameters. In most of our analysis, agent choice sets are finite set of labels indexing production opportunities.

[1] Alternatively, aggregate elasticities may be treated as "sufficient statistics". However, since such statistics are not structural, they will generally vary with policy and optimal tax analysis requires that this variation is modeled and quantified.

[2] In doing so we build on and develop an agenda initiated by Saez (2002, 2004). We relate our contribution explicitly to these earlier contributions and the larger literature below.

Such choice sets are well suited to model locational, occupational or qualitative skill choices that impact incomes and production, but that lack natural orderings or topologies. Since more structured choice sets (describing, say, incomes) are also covered, the set of applications for which optimal tax formulas are available is correspondingly enlarged. Second, in relaxing assumptions on preferences that require locally unique, smooth-in-taxes individual choices, the framework permits a much richer set of substitution responses to (optimal) tax perturbations. In contrast, the classical Mirrlees model assumes that a tax perturbation at an income induces only agents in a small neighborhood of that income to modify their choices.[3] Empirical implementation of the approach requires the placing of some structure on the distribution of agents over preferences. We follow the empirical discrete choice literature and focus on mixed logit specifications, which combine additive choice-specific Gumbel preference shocks with low dimensional mixing shocks. We show that this specification is flexible enough to permit approximation of optimal tax policy in a wide range of discrete choice situations. Further this approach allows the use of both tax and other sources of wage variation in pinning down the parameters needed for (optimal tax) counterfactuals.[4]

We focus initially on settings with a finite (but potential large) number of mutually exclusive income-generating "activity" choices, and generalize later to settings with an uncountable choice set. Although choice sets may have some order or topological structure, we do not require this. Competitive firms hire distributions of choices and, in equilibrium, the pre-tax income associated with a choice is the associated marginal product. Agent utility is assumed to be a function of after-tax income, the choice itself and various idiosyncratic agent characteristics some of which the policymaker may be able to observe and tag. Taxes are a function of agent choice and publicly observable and taggable characteristics. We assume smoothness and monotonicity of agent utility with respect to its after-tax income argument, but not its choice argument, which in any case belongs to a finite set in our benchmark case. Given a mapping from choice and observable (and taggeable) characteristics

[3]In particular, we can dispense with single crossing type assumptions that are used to obtain tractable patterns of binding incentive constraints in the mechanism design approach to optimal taxation. Such assumptions are replaced with the requirement that choice distributions are smooth in taxes. However, this latter requirement is relatively mild in the sense that models without it can be analyzed as limiting cases of (slightly perturbed) models that do.

[4]An additional advantage of the approach pursued in this paper is it that it is closed to that pursued in the optimal commodity tax literature. Hence, it unifies commodity and income theory. This connection was recognized by Saez (2004) and is taken up by Scheuer and Werning (2016) in the (limiting) Mirrlees setting.

to after-tax incomes, agent choice in combination with the distribution of agents over characteristics induces a distribution of agents over choices. In the finite choice setting, we describe conditions on the underlying characteristic distribution that ensure smoothness of the choice distribution in after-tax incomes. With these conditions in place we derive optimal tax formulas. In these formulas, the deadweight loss component is cast in terms of choice distributions and their (semi) elasticities. Central in shaping these formulas is the implied willingness of agents to substitute from high to low tax choices in response to after-tax income perturbations.

The empirical discrete choice literature has identified (semi-)parametric conditions on agent preference distributions that deliver convenient smooth forms for choice distributions. In particular, this literature has highlighted the mixed logit model, which combines tractability (by having a simple parametric form for choice distributions conditional on a low dimensional mixing characteristic) with flexibility (by mixing over this characteristic and, hence, aggregating many conditional choice distributions). Further, the literature has developed strategies for identifying, estimating and testing the fit of such models. We extend existing approximation results and show that by adding a small amount of (extreme value) "noise" to agent preferences, a mixed logit model with choice distributions smooth in taxes can be built whose tax implications arbitrarily well approximate those of an underlying model. We show how our general tax formulas extend to the mixed logit case. A natural starting point and benchmark is the simple logit specification without mixing and with preferences that are additively separable in consumption and choice. This specification has been widely adopted in the spatial literature (our leading application).[5] In this case, optimal taxes depend *only* upon income and not on other dimensions of choice and depends linearly on income if agent utilities are log in consumption. More generally, non-linearity in income is governed by the coefficient of relative risk aversion. Dependence of taxes on incomes alone is a consequence of the substitution patterns implied by the simple logit specification. In this case, the proportional dispersion of agents across alternative choices in response to a choice-specific tax rise is independent of the taxed choice. Consequently, the deadweight loss term present in optimal tax formulas depend only upon the marginal utility of consumption (and, hence, on income), but not on other aspects of choice. More general mixed logit specifications relax the strong substitution restrictions present

[5]It is either adopted explicitly or implicitly via the inclusion of congestion externalities that induce the same pattern of choice responses across locations, see, e.g. Fajgelbaum and Gaubert (2018).

in the simple logit framework. However, the focus remains on the expected tax paid by agents that disperse from a choice contingent on a tax rise and the shaping of this term by the choice-conditional agent preference distribution.

Our initial analysis assumes that the policymaker can observe and tag all relevant dimensions of an agent's activity choice. This assumption is strong. We relax it in two steps. First, we extend our previous analysis to a setting in which the choice set is rich and uncountable. Recent work by Malmberg and Hössjer (2016) provides the mathematical foundations for this extension and formalizes the sense in which it approximates a model with a very large finite set of choices. We interpret the dimensions of the choice set as capturing multiple hedonic dimensions of choice. Permitting full conditioning of taxes on choice, enables the derivation of marginal tax formulas. In particular, if the choice set is single dimensional and choices are interpreted as incomes, new formulas for optimal marginal income taxes are derived. In a second step, we replace the assumption that taxes can be fully conditioned on choice with the assumption that they can be conditioned on income alone. Generically, in this (rich choice set) environment, choice is not measurable with respect to income and the policymaker can no longer fully control after-tax incomes. This requires a recasting of deadweight loss expressions and optimal tax formulas in terms of choice distribution/tax perturbation semi-elasticities, with the latter complex equilibrium objects.[6] We characterize these variables and show how they are used in optimal tax equations.[7]

REFERENCES

Fajgelbaum, P. and C. Gaubert (2018). Optimal spatial policies, geography and sorting. NBER WP 24632.

Malmberg, H. and O. Hössjer (2016). Continuous approximations to discrete choice models using point process theory. Unpublished paper.

[6]The complexity stems from the fact that the choice distribution depends upon after-tax income and taxes impact this directly and, unless the production function is linear in inputs, indirectly through their effect on equilibrium choices and, hence, pre-tax incomes. With full conditioning of taxes on choices, matters are simplified because the policymaker can be modeled as directly selecting after-tax incomes.

[7]Intermediate assumptions in which the policymaker can condition on incomes and some, but not all, other aspects of choice are possible. We leave this extension to further work.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(2), 175–208.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies 68*(1), 225–229.

Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. *The Quarterly Journal of Economics 117*(3), 1039–1073.

Saez, E. (2004). Direct or indirect tax instruments for redistribution: short-run versus long-run. *Journal of Public Economics 88*(3), 503–518.

Scheuer, F. and I. Werning (2016). Mirrlees meets diamond-mirrlees. Technical report, National Bureau of Economic Research.