Optimal Income Taxation with Multidimensional Types

Kenneth L. Judd — Hoover Institution, NBER

Che-Lin Su — Northwestern University (Stanford Ph. D.)

PRELIMINARY AND INCOMPLETE

November, 2006

Introduction

- Optimal income taxation: Mirrlees
  - Heterogeneous productivity
  - Utilitarian (or redistributive) objective
  - Standard cases: clear pattern of binding IC constraints; tax rates in [0,1].
- Our criticism of Mirrlees - not enough heterogeneity
- Multidimensional heterogeneity
  - Little theory; special cases only
  - No clear pattern of binding IC constraints
  - Revelation principle still holds, producing a nonlinear optimization problem with IC constraints.
  - Clearly more realistic than 1-D models.

- This paper examines multidimensional heterogeneity
  - We take a numerical approach
    * This is not as difficult as commonly perceived.
    * Novel numerical difficulties arise for large problems since pooling outcomes imply failure of LICQ
  - Results
    * Optimal marginal tax rate at top can be negative
    * Binding incentive constraints are **not** local.
    * Increases in heterogeneity reduces optimal income redistribution
    * Intuition: Income is a less informative signal in complex models, so use it less.

Public Finance Conventional Wisdom

- Redistributive progressive taxation is usually related to income
  - One might learn about potential income from I.Q., number of degrees, and age, but the natural and supposedly most reliable indicator is income.
  - Mirrlees (1971) examines what the optimal nonlinear income tax schedule would look like
- Mirrlees (1971) makes simplifying assumptions:
  - Intertemporal problems are ignored even though an optimal tax would be tied to life-cycle income and initial wealth.
  - Differences in tastes and family are ignored.
  - The State has perfect information about the individuals in the economy, their utilities and, consequently, their actions. In practice, this is certainly not true for some kinds of self-employment income from self-employment, in particular work done for the worker himself and his family.

- Diamond (2006): The distinction between “ad hoc” restrictions on tax tools, and deriving those tools from an underlying technology is overdrawn.
  - If asymmetric information extends to private actors, then how can government cheaply track total individual transactions?
  - If we recast asymmetric information as infinite administrative costs, how can we cheaply get enough information to implement nonlinear taxation on total income?
  - Having a basic model deriving the tax structure is not a virtue if basic model has critical incompleteness.
- Mirrlees (1986): Computational issues loom large in optimal taxation
  - It is not always easy to devise models simple enough to be manageable and rich enough to be relevant.
  - Optimal tax theory has reached a stage where good theorems may be hard to come by.

Mirrlees Model

- N types of taxpayers.
- Two goods: consumption $(c)$and labour services (l).
- Taxpayer i's productivity is w_i; $0 < w_1, < ... < w_N$, i's pretax income is

$$y_i := w_i l_i, \quad i = 1, ..., N \tag{1}$$

- The *utilitarian social welfare function* $W : R^N \times R^N_+, \to R$ is

$$W(a) := \sum_i \lambda_i u^i(c_i, y_i/w_i), \tag{2}$$

  where λ_i is population frequency of type i.
- Resource constraint: $\sum_i \lambda_i c_i \leq \sum_i \lambda_i y_i$

- Each taxpayer can choose any (y_i, c_i) bundle offered by the government.
- Revelation principle: government constructs schedule s.t. type i will choose the (y_i, c_i) bundle
- Government problem

$$\begin{aligned} \max_{y_i, c_i} \quad & \sum_i \lambda_i u_i(c_i, y_i/w_i) \\ u_i(c_i, y_i/w_i) \geq & \; u_i(c_j, y_j/w_i), \forall i, j \\ \sum_i c_i \leq & \; \sum_i y_i \end{aligned} \tag{3}$$

- The zero tax commodity bundles, (c^*, l^*, y^*), are the solutions to

$$\max_l u_i(w_i l, l)$$

- Examples:

$$
\begin{aligned}
u(c,l) &= \log c - l^{1/\eta+1}/(1/\eta+1) \\
N &= 5 \\
w_i &\in \{1,2,3,4,5\} \\
\lambda_i &= 1/N
\end{aligned}
$$

- The zero tax solution is $l_i = 1$, $c_i = w_i$
- We compute the solutions for various w and η, and report the following:

$$
\begin{aligned}
y_i, &\quad i = 1,..,N, \\
\frac{y_i - c_i}{y_i}, &\quad i = 1,..,N, \quad \text{(average tax rate)} \\
1 - \frac{u_l}{wu_c}, &\quad i = 1,..,N, \quad \text{(marginal tax rate)} \\
l_i/l_i^*, &\quad i = 1,..,N, \\
c_i/c_i^*, &\quad i = 1,..,N,
\end{aligned}
$$

Five Mirrlees Economies

Table 1. $\eta = 1$

i	y_i	$\frac{y_i - c_i}{y_i}$	MTR_i^+	l_i/l_i^*	c_i/c_i^*
1	0.40	-2.87	0.63	0.40	1.56
2	1.31	-0.45	0.53	0.65	0.95
3	2.56	0.03	0.40	0.85	0.83
4	4.01	0.16	0.25	1.00	0.84
5	5.54	0.19	–	1.10	0.90

Table 2. $\eta = 1/2$

i	y_i	$\frac{y_i - c_i}{y_i}$	MTR_i^+	l_i/l_i^*	c_i/c_i^*
1	0.60	-2.09	0.68	0.60	1.87
2	1.54	-0.39	0.59	0.77	1.08
3	2.69	0.02	0.47	0.89	0.87
4	3.99	0.17	0.32	0.99	0.82
5	5.41	0.21	–	1.08	0.85

Table 3. $\eta = 1/3$:

i	y_i	$\frac{y_i - c_i}{y_i}$	MTR_i^+	l_i/l_i^*	c_i/c_i^*
1	0.70	-1.91	0.73	0.70	2.06
2	1.66	-0.38	0.64	0.83	1.15
3	2.77	0.02	0.53	0.92	0.90
4	3.99	0.17	0.38	0.99	0.82
5	5.33	0.23	–	1.06	0.82

Table 4. $\eta = 1/5$

i	y_i	$\frac{y_i - c_i}{y_i}$	MTR_i^+	l_i/l_i^*	c_i/c_i^*
1	0.80	-1.84	0.79	0.80	2.29
2	1.78	-0.39	0.71	0.89	1.24
3	2.85	0.02	0.61	0.95	0.93
4	4.01	0.19	0.48	1.00	0.81
5	5.25	0.26	–	1.05	0.77

Table 5. $\eta = 1/8$

i	y_i	$\frac{y_i - c_i}{y_i}$	MTR_i^+	l_i/l_i^*	c_i/c_i^*
1	0.87	-1.84	0.84	0.87	2.48
2	1.86	-0.41	0.77	0.93	1.31
3	2.91	0.02	0.69	0.97	0.95
4	4.02	0.20	0.58	1.00	0.80
5	5.19	0.28	–	1.03	0.73

Two-D Types - Productivity and Elasticity of Labor Supply

- $u^j(c,l) = \log c - l^{1/\eta_j+1}/(1/\eta_j + 1)$, $j \in \{1,2,...,N\}$
- w_i is productivity type $i \in \{1,2,...,N\}$
- No correlation between characteristics
- (c_{ij}, y_{ij}) is allocation for (i,j)-type taxpayer.
- Zero tax solution for type (i,j) is $(l^*_{ij}, c^*_{ij}, y^*_{ij}) = (1, w_i, w_i)$.
- Problem:

$$\max_{(y,c)} \quad \sum_{i=1}^N \sum_{j=1}^N \lambda_{ij}\, u^j(c_{ij}, y_{ij}/w_i)$$

$$u^j(c_{ij}, y_{ij}/w_i) - u^j(c_{i'j'}, y_{i'j'}/w_i) \geq 0 \quad \forall (i,j),(i',j')$$

$$\sum_{i=1}^N \sum_{j=1}^N c_{ij} \leq \sum_{i=1}^N \sum_{j=1}^N y_{ij}$$

- We choose the following parameters:
  - $N = 5$, $w_i = i$
  - $\lambda_{ij} = 1$
  - $\eta = (1, 1/2, 1/3, 1/5, 1/8)$.
  - Use zero tax solution (c^*, y^*) as starting point for NLP solver.

Table 6. $\eta = (1, 1/2, 1/3, 1/5, 1/8)$, $w = (1, 2, 3, 4, 5)$

(i,j)	c_{ij}	y_{ij}	$MTR_{i,j}$	$ATR_{i,j}$	l_{ij}/l^*_{ij}	c_{ij}/c^*_{ij}	Utility	
							Judd-Su	Mirrlees
$(1,1)$	1.68	0.42	0.28	-2.92	0.42	1.68	0.4294	.3641
$(1,2)$	1.77	0.62	0.32	-1.86	0.62	1.77	0.4952	.3138
$(1,3)$	1.79	0.65	0.51	-1.75	0.65	1.79	0.5378	.6601
$(1,4)$	1.83	0.77	0.50	-1.37	0.77	1.83	0.5700	.7830
$(1,5)$	1.86	0.86	0.43	-1.16	0.86	1.86	0.5940	.8760
$(2,1)$	1.86	0.86	0.60	-1.16	0.43	0.93	0.5308	.3751
$(2,2)$	2.03	1.39	0.50	-0.45	0.69	1.01	0.5973	.6180
$(2,3)$	2.07	1.50	0.56	-0.38	0.75	1.03	0.6512	.7189
$(2,4)$	2.16	1.74	0.46	-0.24	0.87	1.08	0.7006	.8181
$(2,5)$	2.20	1.83	0.46	-0.20	0.91	1.10	0.7413	.9085
$(3,1)$	2.20	1.83	0.55	-0.20	0.61	0.73	0.6053	.5496
$(3,2)$	2.47	2.49	0.43	0.00	0.83	0.82	0.7157	.7269
$(3,3)$	2.47	2.49	0.53	0.00	0.83	0.82	0.7878	.8158
$(3,4)$	2.55	2.68	0.52	0.04	0.89	0.85	0.8520	.9057
$(3,5)$	2.62	2.85	0.42	0.07	0.95	0.87	0.8965	.9672
$(4,1)$	3.36	4.00	0.16	0.15	1.00	0.84	0.7127	.7090
$(4,2)$	3.36	4.00	0.16	0.15	1.00	0.84	0.8794	.8664
$(4,3)$	3.36	4.00	0.15	0.15	1.00	0.84	0.9627	.9402
$(4,4)$	3.36	4.00	0.15	0.15	1.00	0.84	1.0461	1.0080
$(4,5)$	3.36	4.00	0.15	0.15	1.00	0.84	1.1017	1.0476
$(5,5)$	4.00	5.14	0	0.22	1.02	0.80	1.2439	1.1487
$(5,4)$	4.11	5.24	-0.05	0.21	1.04	0.82	1.1928	1.1331
$(5,3)$	4.34	5.43	-0.12	0.20	1.08	0.86	1.1188	1.0877
$(5,2)$	4.49	5.56	-0.11	0.19	1.11	0.89	1.0428	1.0286
$(5,1)$	4.87	5.87	-0.15	0.17	1.17	0.97	0.8933	.8901

Table 7. Binding IC$[(i,j),\ (i',j')]$

(i,j)	$(i'j')$	(i,j)	$(i'j')$
		$(4,1)$	$(3,2)$, $(3,3)$, $(3,5)$, $(4,2)$, $(4,3)$, $(4,4)$, $(4,5)$
$(1,2)$	$(1,1)$	$(4,2)$	$(4,1)$, $(4,3)$, $(4,4)$, $(4,5)$
$(1,3)$	$(1,2)$	$(4,3)$	$(4,1)$, $(4,2)$, $(4,4)$, $(4,5)$
$(1,4)$	$(1,3)$	$(4,4)$	$(4,1)$, $(4,2)$, $(4,3)$, $(4,5)$
$(1,5)$	$(1,4)$, $(2,1)$	$(4,5)$	$(4,1)$, $(4,2)$, $(4,3)$, $(4,4)$
$(2,1)$	$(1,4)$, $(1,5)$	$(5,1)$	$(4,1)$, $(4,2)$, $(4,3)$, $(4,4)$, $(4,5)$
$(2,2)$	$(1,5)$, $(2,1)$	$(5,2)$	$(4,1)$, $(4,2)$, $(4,3)$, $(4,4)$, $(4,5)$, $(5,1)$
$(2,3)$	$(2,2)$	$(5,3)$	$(5,2)$
$(2,4)$	$(2,3)$	$(5,4)$	$(5,3)$
$(2,5)$	$(2,4)$, $(3,1)$	$(5,5)$	$(5,4)$
$(3,1)$	$(2,3)$, $(2,5)$		
$(3,2)$	$(2,5)$, $(3,1)$, $(3,3)$		
$(3,3)$	$(3,2)$		
$(3,4)$	$(3,2)$, $(3,3)$		
$(3,5)$	$(3,4)$		

Comparisons

- Negative marginal rates at top in heterogeneous η case!
- Binding IC constraints
  - Some are not local in income space; appears to violate Assumption B in Guesnerie-Seade
  - More binding constraints than variables - LICQ problem?
- Less redistribution in heterogeneous η case
  - Total redistribution is less
  - Average tax rates are lower for top two productivity types
  - Marginal tax rates are lower for top two productivity types
  - All high ability types prefer heterogeneous world
- More output - both consumption and labor supply tends to be higher in heterogeneous economy

Numerical Issues

- LICQ (linear independence constraint qualification)
  - "The gradients of the binding constraints are linearly independent at the solution."
  - LICQ implies unique Karush-Kuhn-Tucker multipliers.
  - LICQ is a sufficient condition in convergence theorems for most algorithms.
  - Essentially a necessary condition for good convergence rate.
  - Will fail when there are more binding constraints than variables.
  - MFCQ fails in some cases, and shadow prices will be unbounded!
- Software and Hardware
  - AMPL - modelling language commonly used in OR
  - Desktop computers, primarily through NEOS

- Algorithms
  - FilterSQP was most reliable - robust to moderate LICQ failure
  - SNOPT was pretty reliable - robust to moderate LICQ failure
  - IPOPT stopped early - interior point method is too loose
  - MINOS often failed - relies strongly on LICQ
  - Others at NEOS failed
  - fmincon - no point in trying it
  - Lesson: try many different algorithms!
- Global optimization issues
  - Successful algorithms agreed
  - Small deviation examples and multiple restarts found same results

MPCC

- "Mathematical programming with complementarity constraints"

$$\begin{aligned} \max_{x} \quad & f(x) \\ g(x) &= 0 \\ h(x) &\geq 0, \; s(x) \geq 0 \\ s(x)\, h(x) &= 0, \; \text{componentwise} \end{aligned}$$

- If complementarity slackness conditions bind, then LICQ will generically fail in many problems

- "Stackelberg games" are MPCCs: choose all players' moves so as to maximize leader's objective subject to the followers' responses being consistent with equilibrium.

- Economics is full of MPCCs
  - All nonlinear pricing, optimal taxation, and mechanism design problems
  - Many empirical methods. Judd and Su (2006) shows
    - MPCC outperforms NFXP on Harold Zurcher problem
    - MPCC can estimate games; NFXP can't
- Algorithms
  - Several under development: Leyffer, Munson, Anitescu, Peng, Ralph
  - Su and Judd (2005) proposes hybrid approach combining lottery approach and MPCC methods to deal with global optimization problems

Three-Dimensional Types - Productivity and Labor Disutility

- Consider the utility function

$$u(c,l) = u(c, y/w) := \frac{(c-\alpha)^{1-1/\gamma}}{1-1/\gamma} - \psi \frac{(y/w)^{1/\eta+1}}{1/\eta+1}$$

- Possible heterogeneities: w, η, α, γ, and ψ
  - w - wage
  - η - elasticity of labor supply
  - α - the net of initial wealth and basic needs
  - γ - elasticity of demand for consumption
  - ψ - level of distaste for work

- Example: $N = 3$, $w_i \in \{2, 3, 4\}$, $\eta_j \in \{1/2, 1, 2\}$, $a_k \in \{0, 1, 2\}$, $\gamma = \psi = 1$

Two-Dimensional Types - Productivity and Age

- Dynamic OLG optimal tax
  - Individuals know life-cyle productivity
  - Mirrlees approach would have agent reveal type
  - Tax policy would be age-dependent
- Suppose age is not used
  - Better description of actual tax policies
  - Still a mechanism design problem - just (a lot) more incentive constraints

- Example:
  - No discounting, $u(c, l) = \log c - l^2/2$

Wage History

	Period		
Type	1	2	3
1	1	3	2
2	2	4	4
3	2	5	4
4	3	5	6

  - Wage patterns: four types, each lives three periods.
  - Consider four policies: Mirrlees I (see age and consumption), Mirrlees II (see only age), age-free Mirrlees (unobservable savings), linear $(-a + by)$
  - Total income patterns under three policies (observability in savings did not matter in this example)

Table 8: Aggregate Outputs for Each Type

	Total Income			Total Tax Paid			Total Utility		
Type	Mirr.	Nlin.	Lin.	Mirr.	Nlin.	Lin.	Mirr.	Nlin.	Lin.
1	4.72	5.43	5.65	–2.40	–1.36	–0.96	1.79	1.40	1.23
2	9.60	10.02	9.70	–0.03	0.07	–0.07	2.22	2.20	2.23
3	11.88	11.19	10.83	0.51	0.36	0.18	2.43	2.46	2.49
4	15.48	14.35	13.90	1.91	0.93	0.85	2.82	3.01	3.03

Table 9: Life-cycle patterns of income, taxes, and MTR

		OLG Model - Mirrlees			Nonlinear tax			Linear tax		
Type	age	y	Tax	MTR	y	Tax	MTR	y	Tax	MTR
1	1	0.31	–0.79	0.25	0.32	–1.01	0.25	0.42	–0.64	0.22
1	2	3.15	–0.79	0.16	3.55	0.24	0.10	3.46	0.02	0.22
1	3	1.25	–0.79	0.25	1.54	–0.59	0.12	1.75	–0.34	0.22
2	1	1.05	–0.01	0.15	1.05	–0.73	0.12	1.12	–0.48	0.22
2	2	4.32	–0.01	0.13	4.48	0.39	0.07	4.28	0.20	0.22
2	3	4.22	–0.01	0.15	4.48	0.39	0.07	4.28	0.20	0.22
3	1	1.05	0.17	0.00	1.02	–0.73	0.07	1.12	–0.48	0.22
3	2	6.59	0.17	0.00	6.29	0.79	0.09	6.10	0.60	0.22
3	3	4.22	0.17	0.00	3.85	0.29	0.12	3.59	0.05	0.22
4	1	1.99	0.63	0.00	1.54	–0.59	0.23	1.75	–0.34	0.22
4	2	5.52	0.63	0.00	4.90	0.47	0.12	4.83	0.32	0.22
4	3	7.96	0.63	0.00	7.90	1.05	0.01	7.30	0.87	0.22

Future Work and Conclusions

- Robustness
  - Other objectives - e.g., Rawls
  - Government expenditures
  - Labor participation decisions and fixed costs of working
  - Examine more of the parameter space
  - Empirically reasonable wage distributions
- Related policy issues
  - Optimal treatment of educational expenses
  - Deductibility of children, medical expenses, mortgage interest
  - Taxation of capital income and assets
  - Use wage rate when observable?
  - Allow some memory at option of taxpayer?
  - Marriage tax?

- Address computational issues
  - Resurrect LICQ by finding minimal but sufficient set of binding constraints
  - Hope that mathematicians solve the mathematical challenges we have described to them
  - Develop asymptotic approximation methods
  - Examine alternative formulations
    * Relaxations of ICs
    * Piecewise linear tax schedules
- Exploit third millenium computer technologies - Blue Gene, Red Storm, Thunderbird, Jaguar, TeraGrid, Condor, BOINC - that are far more powerful than second millenium technologies - abacus, sliderule, and desktops

Future Work and Conclusions

- Multidimensionality in taxpayer types significantly affects results
- Multidimensional problems require use of state-of-the-art computational methods but are feasible