Robust Bounds on Optimal Tax Progressivity

Anmol Bhandari[1] Jaroslav Borovička[2] Yuki Yao[3]

[1]University of Minnesota and NBER

[2]New York University and NBER

[3]University of Kent

How should we tax incomes?

Large literature on optimal tax design

- theory: Ramsey (1927), Mirrlees (1971)
- applications: Diamond and Saez (2011), Golosov et al. (2016), Heathcote et al. (2017)

Key predictions depend on hard to measure objects

- distribution of earning potentials (labor productivity)
- distribution of preferences (labor supply elasticity)

Optimal tax design acknowledging uncertainty about distribution of individual characteristics

- build on decision theory under ambiguity to model welfare consequences of statistical uncertainty about type distributions with Mirrlees (1971)
- quantify uncertainty using information from administrative and survey data

Key source of uncertainty

- tails of the productivity distribution with scarce information relative to their welfare implications

Main finding

- concerns for uncertainty call for substantially **lower** tax progressivity for high incomes

Framework

A continuum of households indexed with productivity type $z \sim F(z)$.

Households choose labor supply $n(z)$ subject to an income tax function $T(y)$ where $y(z) = zn(z)$.

A utilitarian government with Pareto weight function $\psi(z)$ chooses $T(y)$ to maximize social welfare.

- trades off redistributive motives and efficiency
- faces uncertainty about the type distribution $F(z)$ and considers a set of distributions statistically close to F

Given a labor income tax function $T(y)$, household of type z solves

$$\max_{c,n} U(c, n; z)$$

subject to the budget constraint

$$c = zn - T(zn)$$

Indirect utility function $\mathcal{U}(z; T)$ and decision rules $\mathcal{C}(z; T)$, $\mathcal{N}(z; T)$, $\mathcal{Y}(z; T)$.

Without uncertainty concerns, the utilitarian government solves

$$\max_{T} \int \psi(z)\mathcal{U}(z;T)dF(z) + V(G)$$

subject to the government budget constraint

$$\int T(\mathcal{Y}(z;T))dF(z) = G.$$

Uncertainty about household characteristics

The government is concerned that distribution $F(z)$ may be misspecified.

- it considers alternative distributions $\widetilde{F}(z)$ that are statistically close to $F(z)$

A measure of statistical closeness is the relative entropy (Kullback–Leibler divergence)

$$\mathcal{E}(F, \widetilde{F}) = \int m(z) \log m(z) \, dF(z)$$

- $m(z) = \frac{d\widetilde{F}(z)}{dF(z)}$ be the Radon–Nikodým derivative of $\widetilde{F}$ with respect to F

For a given benchmark F and entropy bound κ, the set of statistically close distributions is

$$\mathcal{F}(F, \kappa) = \left\{ \widetilde{F} : \mathcal{E}(F, \widetilde{F}) \leq \kappa \right\}$$

- the set $\mathcal{F}(F, \kappa)$ is large and the government does not put a prior on that set

A robust utilitarian government solves the max-min problem

$$\max_{T} \min_{\widetilde{F} \in \mathcal{F}} \int \psi(z) \mathcal{U}(z;T) d\widetilde{F}(z) + V(G)$$

subject to

$$\int T(\mathcal{Y}(z;T)) d\widetilde{F}(z) = G.$$

A robust utilitarian government solves the max-min problem

$$\max_{T} \min_{m:\widetilde{F}\in\mathcal{F}} \int \psi(z)\,\mathcal{U}(z;T)\,m(z)\,dF(z) + V(G)$$

subject to

$$\int T(\mathcal{Y}(z;T))\,m(z)\,dF(z) = G.$$

- utilitarian concern: low weight $m(z)$ on households with high contribution to welfare
- budgetary concern: low weight $m(z)$ on households with high contribution to the budget

Theoretical analysis

The optimal tax problem can be cast as a mechanism design problem (Mirrlees (1971))

- revelation principle allows to focus on direct mechanisms
- workers provide a report z' of their type z
- government offers a menu of allocations $(c(z'), y(z'))$ that incentivizes truthtelling, $z' = z$
- implied tax function $T(y(z)) = y(z) - c(z)$

The robust government solves

$$\max_{c,y} \min_{m:\widetilde{F}\in\mathcal{F}} \int \psi(z)\, U\left(c(z), \frac{y(z)}{z}\right) m(z)\, dF(z) + V(G)$$

subject to incentive compatibility constraints

$$U\left(c(z), \frac{y(z)}{z}\right) \geq U\left(c(z'), \frac{y(z')}{z}\right) \qquad \forall z, z'$$

and the government budget constraint

$$\int (y(z) - c(z))\, m(z)\, dF(z) = G.$$

The robust government solves

$$\min_{m:\widetilde{F}\in\mathcal{F}} \max_{c,y} \int \psi(z)\, U\left(c(z), \frac{y(z)}{z}\right) m(z)\, dF(z) + V(G)$$

subject to incentive compatibility constraints

$$U\left(c(z), \frac{y(z)}{z}\right) \geq U\left(c(z'), \frac{y(z')}{z}\right) \qquad \forall z, z'$$

and the government budget constraint

$$\int (y(z) - c(z))\, m(z)\, dF(z) = G.$$

Fixing $m(z)$ (fixing a distribution $\widetilde{F}(z)$), the problem is as in Mirrlees (1971), now under $\widetilde{F}(z)$.

- ex-post Bayesian interpretation of $\widetilde{F}(z)$ (**min** and **max** can be interchanged)

Incentive-compatibility constraints are type-by-type, do not depend on the distribution.

- misspecification concerns do not alter incentive compatibility

Optimal allocation and the minimizing 'worst-case' distribution determined jointly.

The worst-case distribution is given by $\widetilde{f}(z) = m(z) f(z)$ with

$$m(z) = \bar{m} \exp\left(-\frac{1}{\theta(\kappa)}\left[\psi(z)\mathcal{U}(z) + \mu T(y(z))\right]\right)$$

- utilitarian concern: lower weight on households with high welfare contribution $\psi(z)\mathcal{U}(z)$
- budgetary concern: lower weight on households who generate high tax revenue $T(y(z))$

We first focus on the theoretical characterization of top marginal tax rates.

- here, we for simplicity assume quasilinear utility

$$U(c,n) = c - \frac{n^{1+\gamma}}{1+\gamma}$$

- insights carry over to general separable preferences

We then provide a quantitative evaluation.

- concave utility, type distribution calibrated to data
- discipline the amount of uncertainty the planner faces

Optimal marginal tax schedule is given by the Diamond (1998)–Saez (2001) 'ABC' formula

$$\frac{T'(y(z))}{1-T'(y(z))} = \underbrace{(1+\gamma)}_{(A)} \underbrace{\frac{\widetilde{\Psi}(z)-\widetilde{F}(z)}{1-\widetilde{F}(z)}}_{(B)} \underbrace{\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)}}_{(C)}.$$

- (*A*): adverse effect of taxes on labor supply via labor supply elasticity
- (*B*): desire to redistribute

$$\widetilde{\Psi}(z) = \int^{z} \frac{\psi(\zeta)\widetilde{f}(\zeta)}{\int \psi(\xi)\widetilde{f}(\xi)\,d\xi} d\zeta$$

- (*C*): tradeoff between labor supply distortion at z and revenue from taxing types above z

Prototypical example

Assume planner puts zero welfare weight on top households

- $\psi(z) = 0$ for $z \geq \hat{z}$ where $\hat{z}$ is some threshold

Benchmark distribution in the tail above $\hat{z}$ is Pareto with shape parameter α

Without misspecification concerns, the tax formula for $z \geq \hat{z}$ simplifies to

$$\frac{T'(y(z))}{1 - T'(y(z))} = \frac{1+\gamma}{\alpha}.$$

- with a fat-tailed type distribution, taxes at the top are nonzero and quantitatively possibly large (Diamond–Saez)
- intuition: the tax revenue from types above z outweighs the labor supply distortion at z

With misspecification concerns, the tax schedule and distribution $\widetilde{F}(z)$ are determined jointly.

- Distribution $\widetilde{F}(z)$ pins down tax schedule by

$$\frac{T'(y(z))}{1-T'(y(z))} = (1+\gamma)\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)}$$

- Tax schedule determines distribution $\widetilde{F}(z)$ by

$$\begin{aligned} T(y(z)) &= T(y(\underline{z})) + \int_{y(\underline{z})}^{y(z)} T'(\eta)d\eta \\ m(z) &= \bar{m}\exp\left(-\frac{\mu}{\theta}T(y(z))\right) \end{aligned}$$

The optimal tax schedule is a fixed point of this argument.

Theorem 1.1

Assume preferences are quasilinear and $\kappa > 0$. Then the marginal tax rate vanishes to zero at the top:

$$\lim_{y \to \infty} T'(y) = 0. \tag{1.1}$$

Moreover, if the right tail of z is Pareto distributed with shape parameter α, then

$$\lim_{y \to \infty} \frac{d \log T'(y)}{d \log y} = -\frac{1}{2}. \tag{1.2}$$

Top rate level and the speed of convergence are independent of

- the magnitude of misspecification concerns κ
- shape of the Pareto tail α of the benchmark distribution
- labor supply elasticity γ

Results carry over to

- general (isoelastic) separable utility

$$U(c,n) = \frac{c^{1-\rho}}{1-\rho} - \chi \frac{n^{1+\gamma}}{1+\gamma}$$

- general welfare weights
- other statistical discrepancy functions

Quantitative application

Preferences and technology

- isoelastic preferences: $U(c,n) = \frac{c^{1-\rho}}{1-\rho} - \upsilon \frac{n^{1+\gamma}}{1+\gamma}$ with $\rho = 1$, $\upsilon = 1$, $\gamma = 2$
- government spending $V(G) = \bar{v}G$

Benchmark distribution F

- $\log z$ has exponentially modified Gaussian (EGM) distribution (Heathcote and Tsujiyama (2021))
- left tail of z distribution is lognormal (parameters μ, σ)
- right tail approximately Pareto (parameter α)

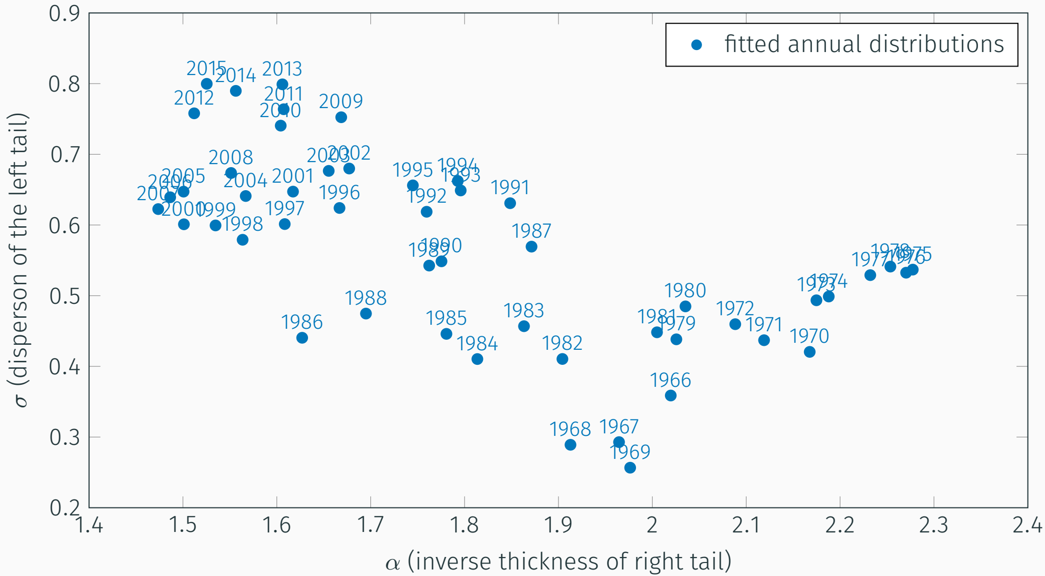
Entropy bound κ

- use time-series variation in observed income distributions (World Income Database)

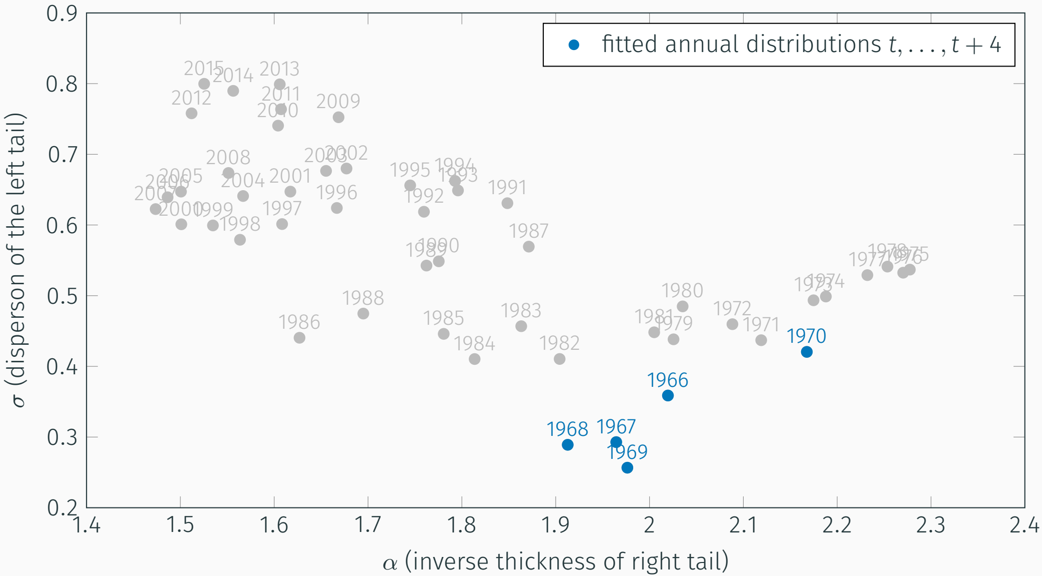
1. For each year t, we fit the EGM distribution to obtain parameters $(\mu_t, \sigma_t, \alpha_t)$.
2. For each 5-year window $\{t, \ldots, t+4\}$, we construct $\mathcal{F}(F_t, \kappa_t)$ as the set that
   - includes all fitted EGM distributions from years $\{t, \ldots, t+4\}$
   - has the smallest entropy radius κ_t
3. Baseline calibration uses the median of $\{\kappa_t\}$.

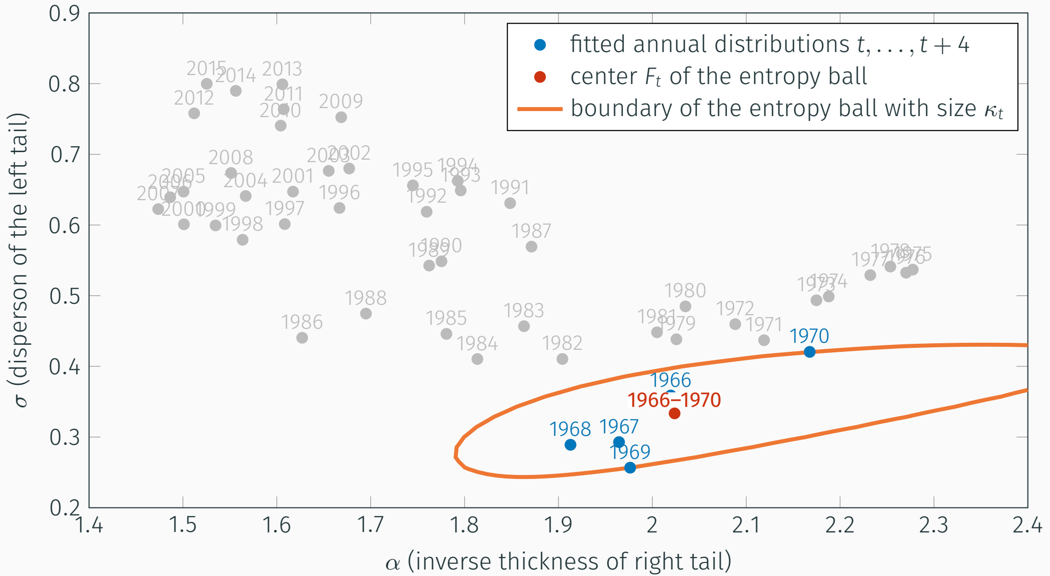
The set $\mathcal{F}(F_t, \kappa_t)$ is rich:

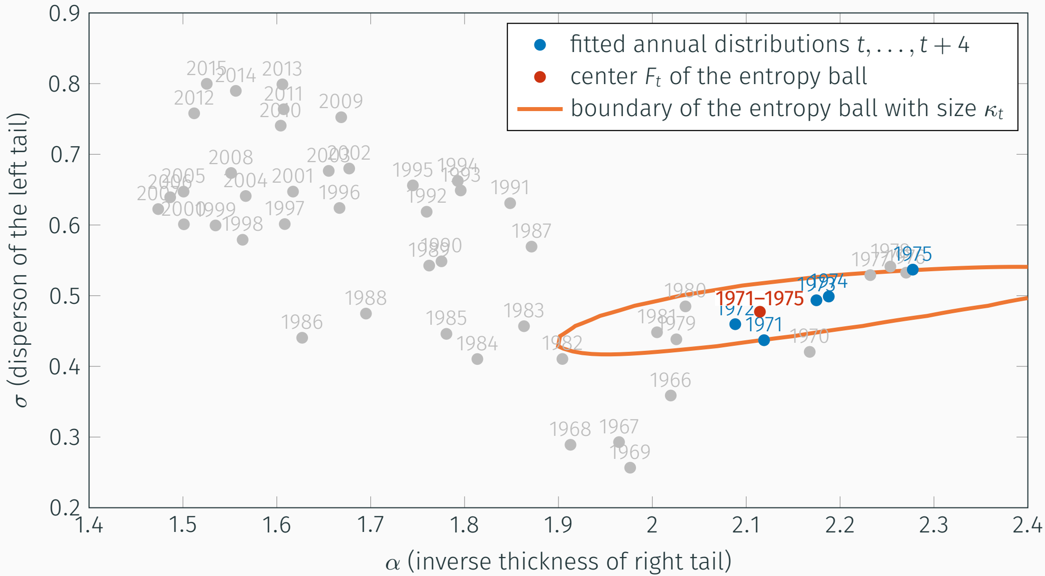
- it contains all distributions that are close to F_t
- not only the parameterized EGM family

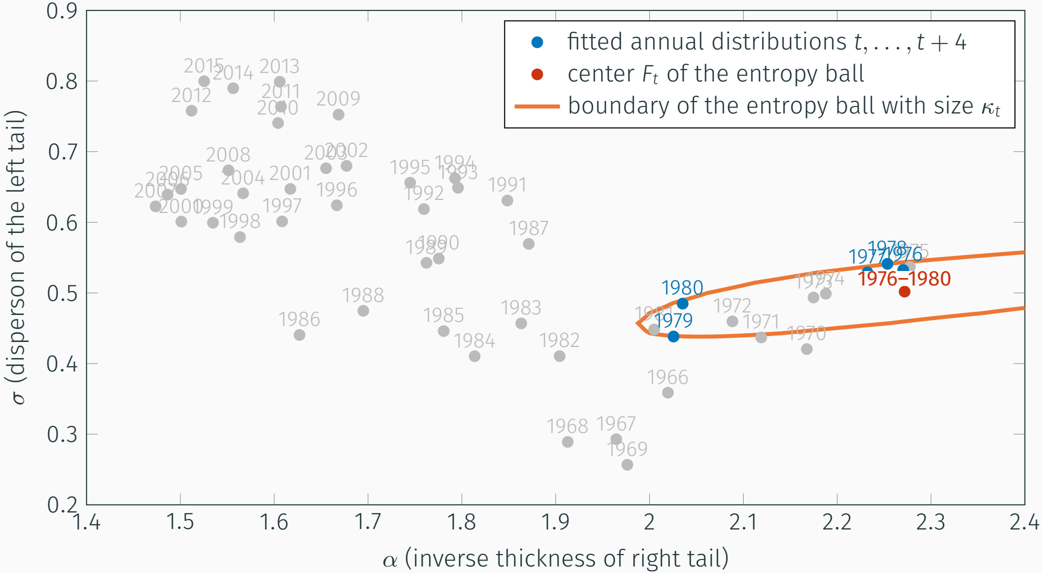

fitted annual distributions
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
2015
2014
2013
2012
2011
2010
2009
2008
2005
2004
2001
2002
1996
1997
1998
1999
1995
1994
1993
1992
1991
1987
1988
1986
1985
1984
1983
1982
1980
1979
1972
1971
1970
1974
1973
1966
1968
1967
1969

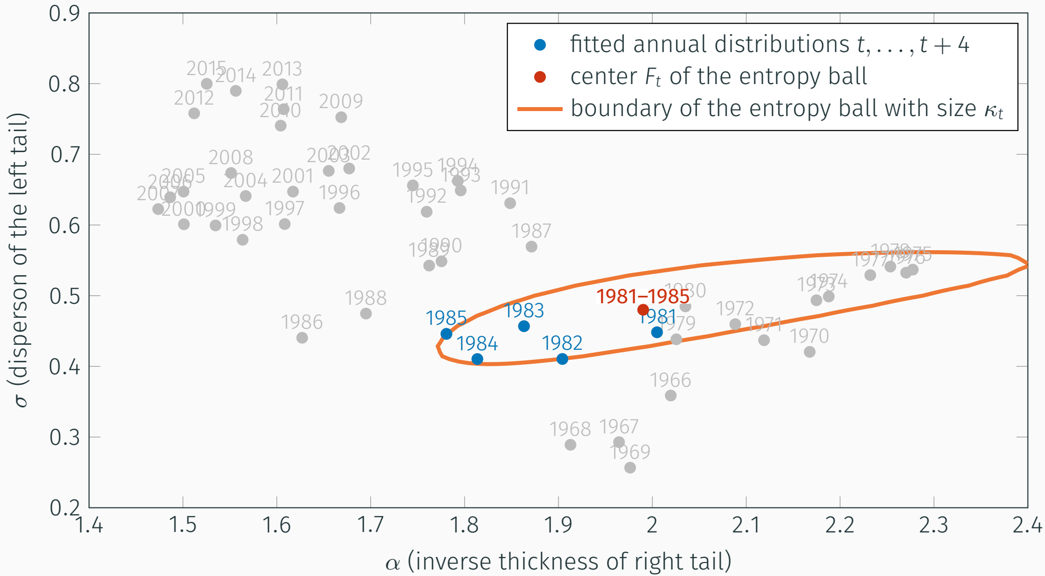
Quantifying uncertainty in income distributions



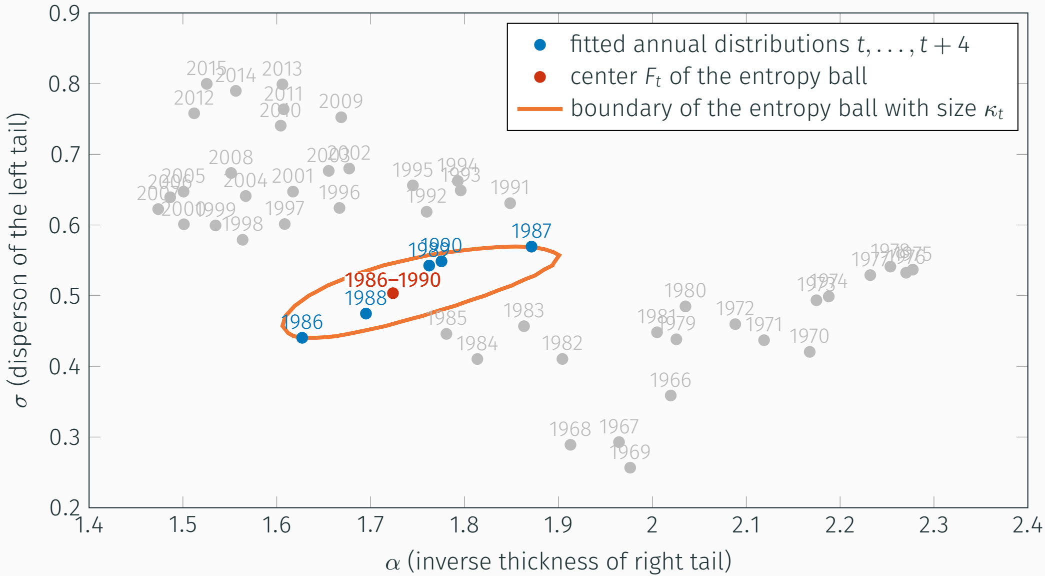

fitted annual distributions $t, \dots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
2015
2014
2013
2012
2011
2010
2009
2008
2005
2004
2001
2002
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1987
1988
1986
1985
1984
1983
1982
1981
1980
1979
1972
1971
1974
1973
1977
1975
1976
1970
1966
1966–1970
1968
1967
1969

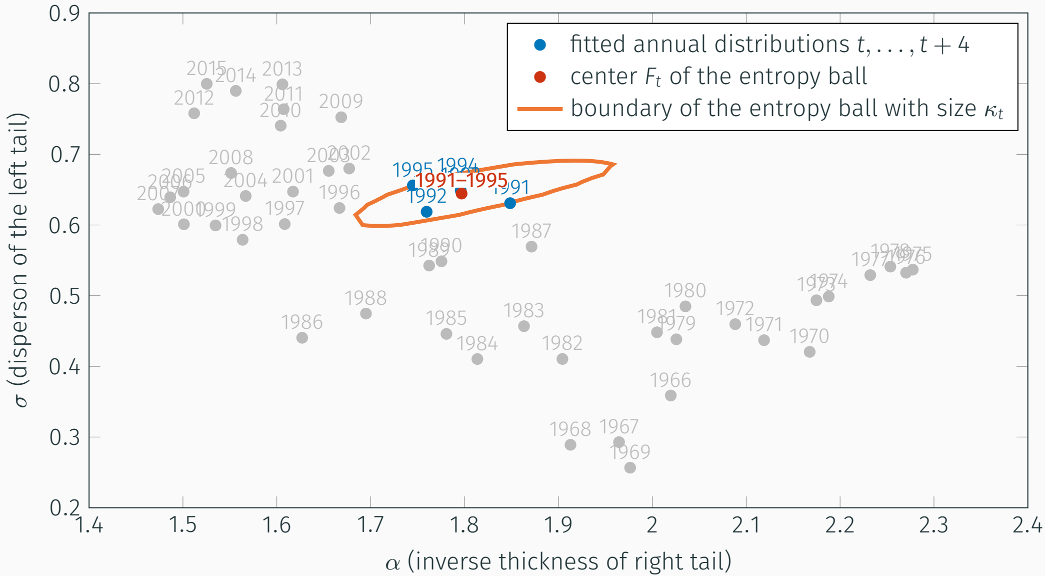

0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
σ (disperson of the left tail)
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
α (inverse thickness of right tail)
fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
1971–1975
1971
1972
1973
1974
1975
1966
1967
1968
1969
1970
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

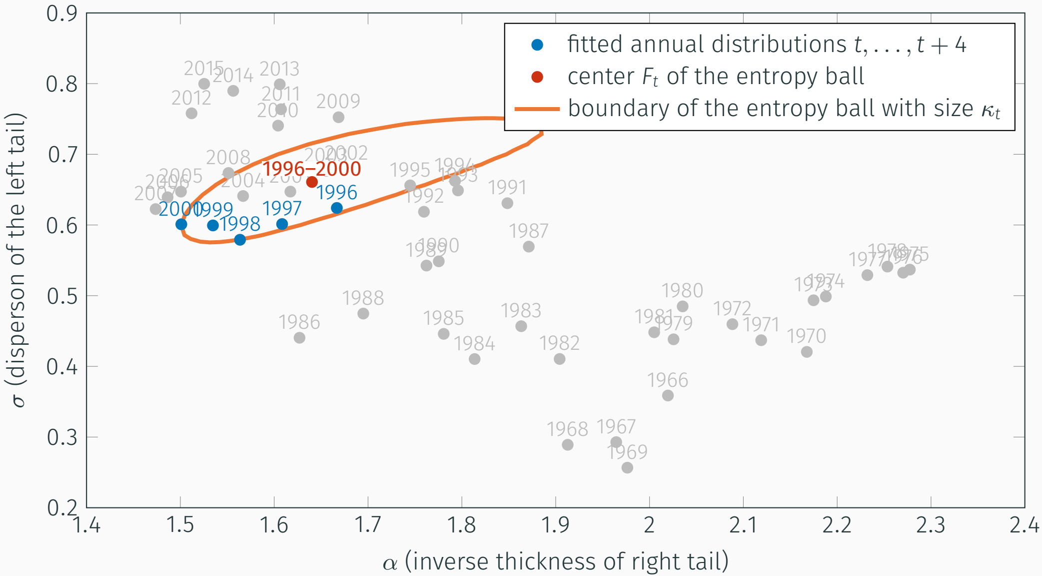

fitted annual distributions $t, \dots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
1976–1980
1980
1979
2015
2014
2013
2012
2011
2009
2008
2004
2001
1996
1997
1998
1995
1992
1991
1987
1988
1986
1985
1984
1983
1982
1966
1968
1967
1969
1972
1971
1970

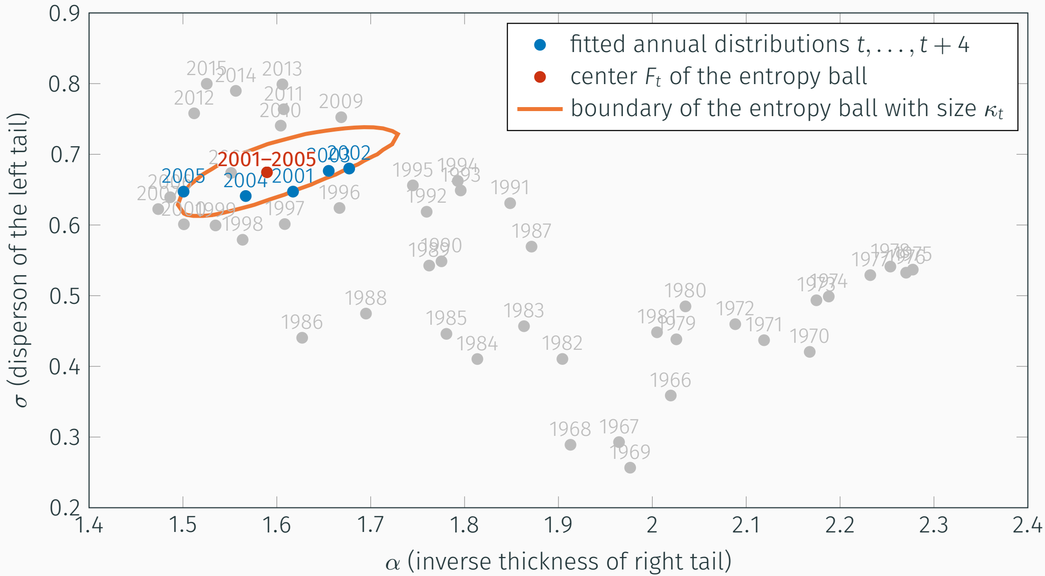


fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
1981–1985
1985
1984
1983
1982
1981
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4

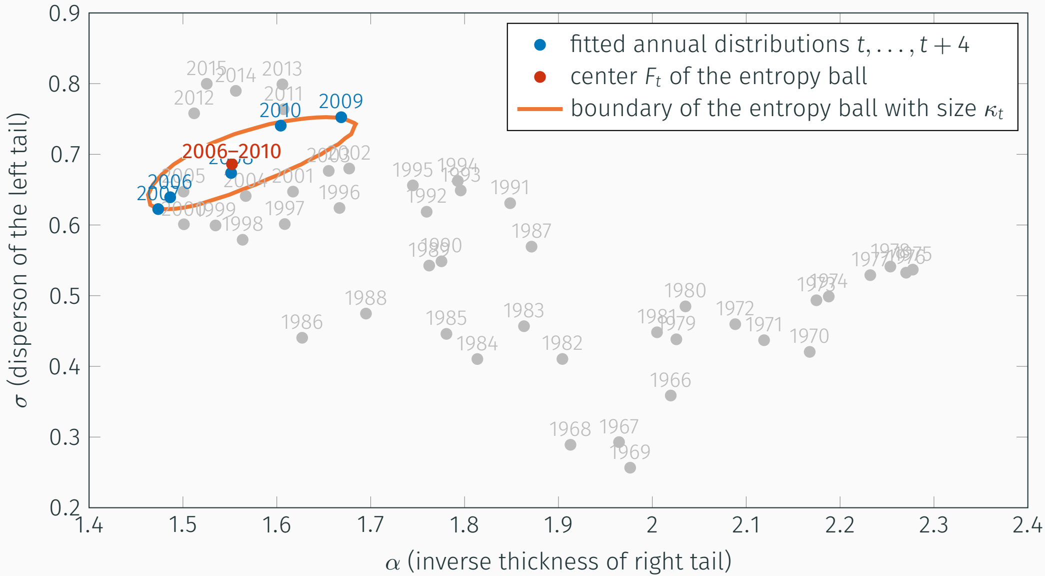

fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
1986–1990
1986
1987
1988
2015
2014
2013
2012
2011
2009
2008
2005
2004
2001
1996
1997
1998
1995
1991
1992
1985
1984
1983
1982
1980
1972
1971
1970
1966
1968
1967
1969

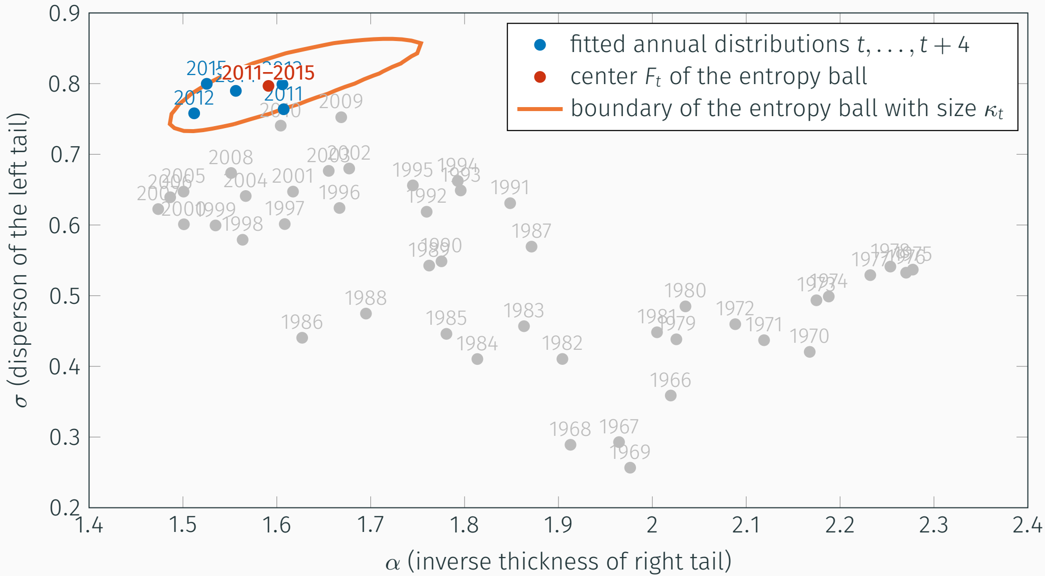

fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
1991–1995
1992
1991
1995
1994
2015
2014
2013
2012
2011
2009
2008
2001
2004
1996
1997
1998
1987
1988
1986
1985
1984
1983
1982
1980
1972
1971
1970
1966
1968
1967
1969


fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
1996–2000
σ (disperson of the left tail)
α (inverse thickness of right tail)

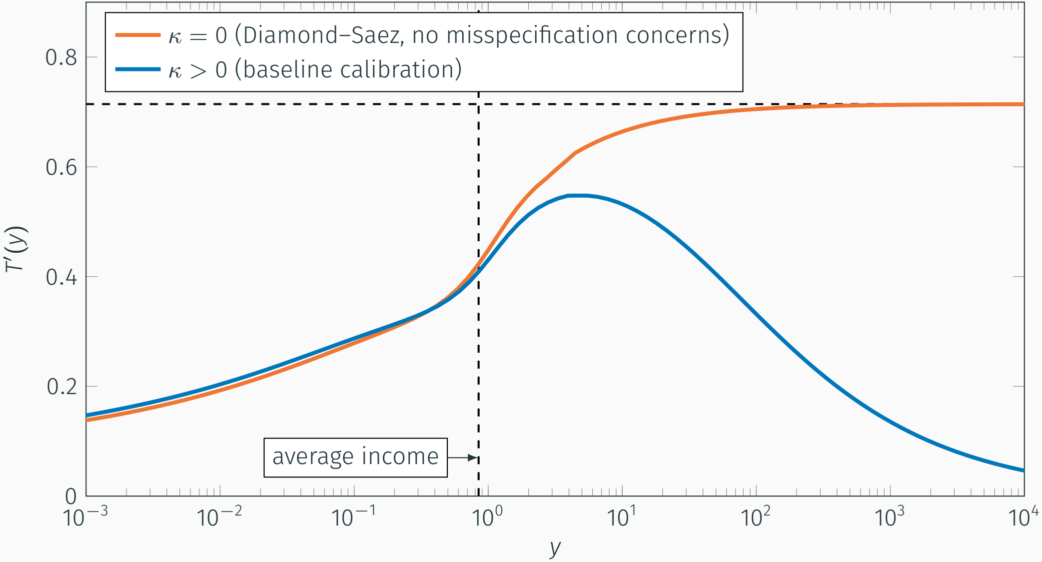

fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
σ (disperson of the left tail)
α (inverse thickness of right tail)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
2001–2005
2005
2004
2001
2002
2015
2014
2013
2012
2011
2009
1998
1997
1996
1995
1992
1991
1987
1988
1986
1985
1984
1983
1982
1980
1972
1971
1970
1966
1968
1967
1969

QUANTIFYING UNCERTAINTY IN INCOME DISTRIBUTIONS

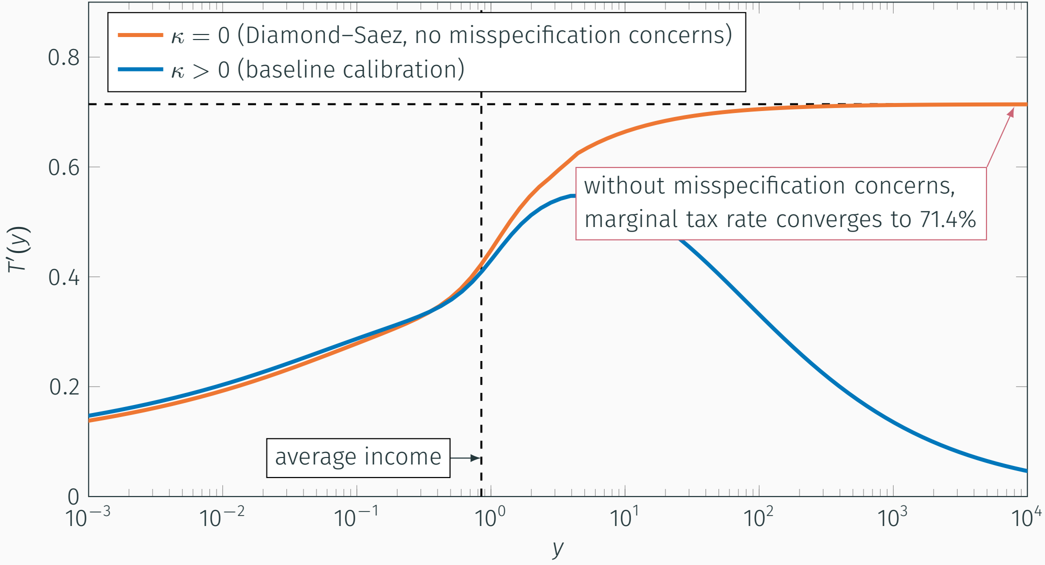


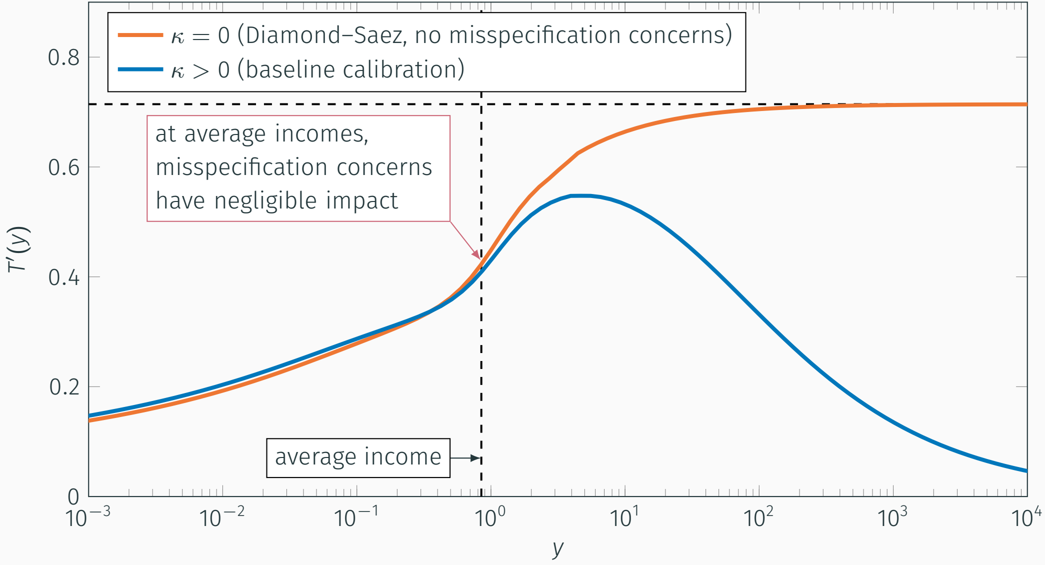


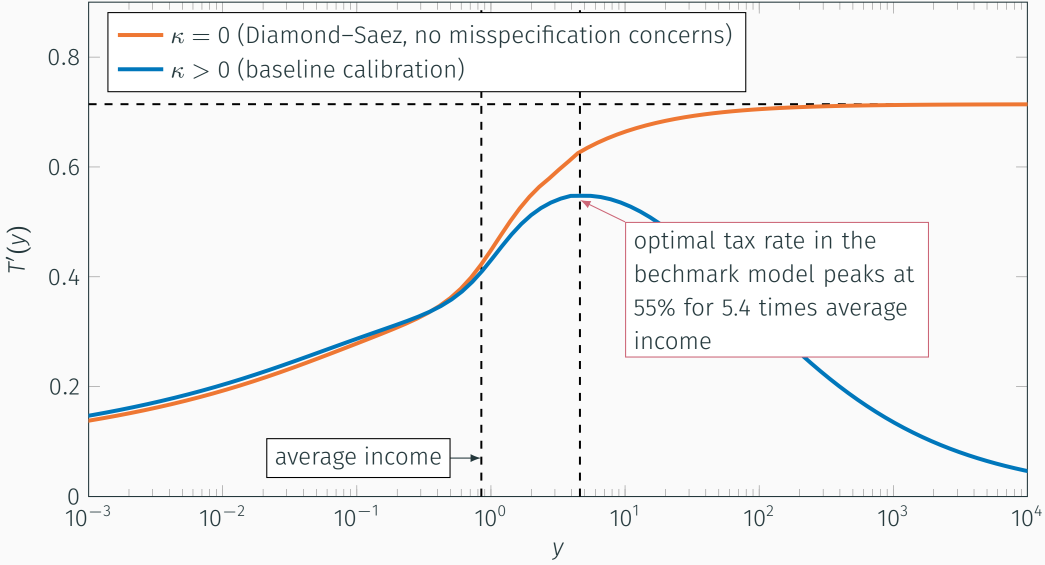
fitted annual distributions $t, \ldots, t+4$
center F_t of the entropy ball
boundary of the entropy ball with size κ_t
2011–2015
σ (disperson of the left tail)
α (inverse thickness of right tail)


$\kappa = 0$ (Diamond–Saez, no misspecification concerns)
$\kappa > 0$ (baseline calibration)
average income
$T'(y)$
0.8
0.6
0.4
0.2
0
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
y

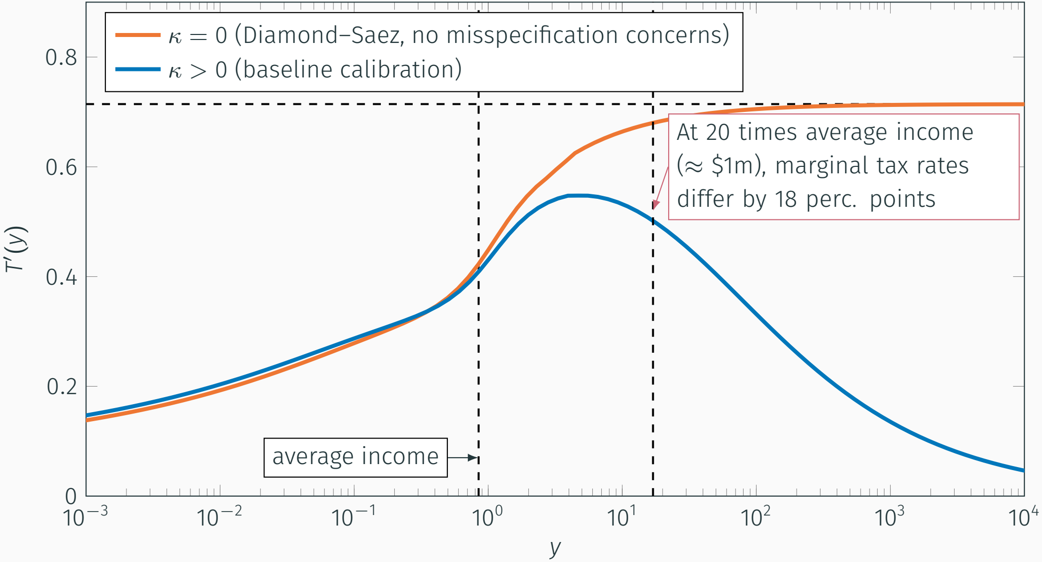
Optimal marginal tax schedules

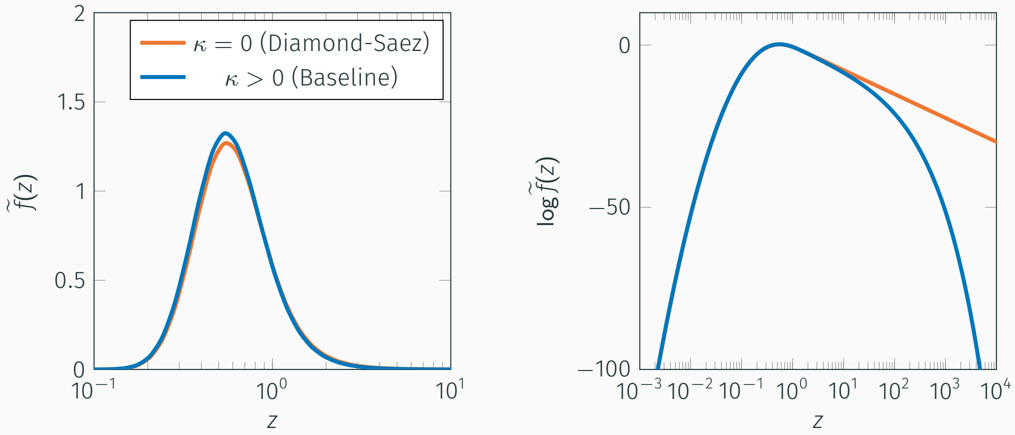



$\kappa = 0$ (Diamond–Saez, no misspecification concerns)
$\kappa > 0$ (baseline calibration)
at average incomes, misspecification concerns have negligible impact
average income
$T'(y)$
0.8
0.6
0.4
0.2
0
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
y


$\kappa = 0$ (Diamond–Saez, no misspecification concerns)
$\kappa > 0$ (baseline calibration)
0.8
0.6
0.4
0.2
0
$T'(y)$
optimal tax rate in the bechmark model peaks at 55% for 5.4 times average income
average income
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
y

Optimal marginal tax schedules





- Worst-case distribution thins out more quickly than the benchmark distribution
- The optimal marginal tax rates are lower at the right tail due to smaller tax gain above z.

The worst-case density is characterized by the distortion

$$m(z) = \bar{m} \exp\left(-\frac{1}{\theta}\left[\mathcal{U}(z) + \mu T(y(z))\right]\right)$$

Right tail of the type distribution

- dominated by budgetary concerns

Left tail of the type distribution

- without redistribution, we would have $\lim_{z\to 0} \mathcal{U}(z) = -\infty$, and $\lim_{z\to 0} m(z) = \infty$
- but redistributive transfers bound $\mathcal{U}(z)$ from below, and so $m(z)$ is bounded above
- this makes misspecification concerns at the bottom quantitatively small

Conclusion

Acknowledging distributional uncertainty points toward lower progressivity.

- especially at the top, where budgetary concerns (per household) are most severe
- the left tail is well insured, leading to only modest concerns
- insights robust to variation in underlying distributions and preferences

Magnitude of misspecification concerns can be disciplined using

- administrative data: time-series variability in income distributions
- survey evidence: measure distinguishability in finite samples

Bibliography

Literature I

Diamond, Peter A. (1998) "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88 (1), 83–95.

Diamond, Peter A. and Emmanuel Saez (2011) "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25 (4), 165–190.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski (2016) "Redistribution and Social Insurance," *American Economic Review*, 106 (2), 359–86, 10.1257/aer.20111550.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2017) "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 132 (4), 1693–1754.

Heathcote, Jonathan and Hitoshi Tsujiyama (2021) "Optimal Income Taxation: Mirrlees Meets Ramsey," *Journal of Political Economy*, 129 (11), 3141–3184.

Mirrlees, James A. (1971) "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38 (2), 175–208.

Ramsey, Frank P. (1927) "A Contribution to the Theory of Taxation," *Economic Journal*, 37 (145), 47–61.

Saez, Emmanuel (2001) "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68 (1), 205–229.