Robust bounds on optimal tax progressivity

Anmol Bhandari* Jaroslav Borovička† Yuki Yao‡

* University of Minnesota and NBER

† New York University and NBER

‡ University of Kent

February 15, 2024

Preliminary draft.

Abstract

We study the problem of a robust planner who designs an optimal taxation scheme for a heterogeneous population in presence of uncertainty about the shape of the distribution of underlying types. Low-income workers are well insured under the optimal scheme, and so concerns about the left tail of the type distribution are negligible. On the other hand, the planner fears misspecification of the right tail of the type distribution emerging from budgetary concerns. Even when the tail of the distribution is Pareto, arbitrarily small misspecification concerns lead to zero marginal taxes at the top. A quantitatively calibrated model shows that a plausible degree of uncertainty leads to an optimal tax scheme with substantially reduced marginal tax rates for high-income earners and a peak marginal tax rate much lower than in the model without uncertainty.

Bhandari acknowledges support from the Heller–Hurwicz Economic Institute.

1 Introduction

It is well-known from the theory of optimal taxation (Mirrlees (1971)) that the shape of the marginal tax curve crucially depends on the shape of the underlying labor productivities and labor supply preferences.[1] Despite increasingly available surveys and administrative micro-data on income, determining the joint distribution of skills and preferences that result in a specific distribution of reported income remains a complex task. The problem is more severe in the tails, where sample sizes are extremely small and the data suffer from a variety of measurement issues.[2]

In this paper, we study optimal taxation of labor income when the government acknowledges uncertainty about the underlying distribution of skills and preferences. We find that these concerns generally lead to lower progressivity. More strikingly, we find that the celebrated Diamond (1998) and Saez (2001)'s high top marginal tax rate result is overturned. By introducing a minimal degree of uncertainty to the Diamond–Saez setting, we prove that the top marginal tax approaches zero at a rate that is bounded away from zero and independent of structural parameters.

Methodologically, we build on Hansen and Sargent (2008)'s notion of "robustness" which aims to construct policies that work well not just for a single distribution but across a set of distributions. The concerns about the shape of the type distribution manifest as a min-max game between a government that chooses a nonlinear tax schedule to maximize a given welfare criterion and its alter ego that adversely perturbs the joint distribution of skills and preferences away from a given benchmark distribution subject to a penalty. Motivated by measurement concerns, we use a penalty that scales in a statistical measure of distance between the two distributions, i.e., relative entropy. The min-max problem allows the government to design a tax schedule that is *robust* with respect to a set of distributions that are hard to distinguish from each other using an available finite data sample.

While our work has elements in common with public finance and Hansen–Sargent robustness literature, there are several distinct differences. Similar to the classic Mirrlees problem, we examine a static environment where individuals vary in skills and preferences, and they supply labor given a non-linear tax schedule. The government devises this schedule to maximize a social welfare criterion while adhering to a budget constraint. Our approach differs in the explicit modeling of the uncertainty the government faces regarding the type-distribution. Instead of assuming certainty as in the traditional Mir-

[1] See, for example, Golosov, Troshkin, and Tsyvinski (2016).

[2] For instance, even if earnings could be measured well, say in tax records, in order to map earnings to skills one needs to take a stand on income that is not reported, as well as what part of the reported income is unrelated to labor supply. The problem gets even more difficult if we want to jointly estimate productivity and labor supply elasticities for such individuals.

rlees framework, our government uses the min-max formulation discussed previously.[3] If we set the penalty in the min-max problem to be large enough, our problem converges to the standard Mirrlees problem in which the planner perfectly knows the shape of the type distribution—a feature that gives us a convenient point of departure. We label the problem involving an infinite penalty as the "baseline problem", and we refer to the type distribution in this baseline as the "benchmark" distribution. Similarly, we use terms like "robust problem" and "worst-case distribution" to denote the planner's problem with a finite penalty and the adverse distribution chosen by the minimizing agent, respectively.

Compared to the robustness or more generally the literature on ambiguity[4], we focus on a different type of uncertainty. While the existing literature primarily deals with the distribution of alternative potential outcomes or states of nature a decision-maker might face under uncertainty, our government is concerned about the uncertainty about the shape of the cross-sectional distribution. Households who have private information about their types of skills and preference face no uncertainty. Moreover, our formulation expands upon the typically used one-parameter penalty specification, allowing us to express varying degrees of uncertainty in different segments of the underlying distribution. For example, we can formally articulate when the government is more uncertain about the extreme ends ("tails") of the distribution than the central part.

Our formalization allows us to examine the optimal nonlinear tax schedule using standard mechanism design techniques and use the revelation principle. This enables us to equivalently recast the (inner) problem of choosing a nonlinear tax schedule equivalently as selecting an optimal allocation, subject to truth-telling constraints. To compare our results with existing literature, we begin with a scenario where the type-distribution is one-dimensional (skills), and the penalty is scalar. We then broaden our scope to more complex situations. For this simplified setting, Diamond–Saez show that in the baseline problem, the marginal tax rate for a particular skill level is determined by the so-called ABC formula. This formula consists of three components: A) a term dependent on labor supply elasticities, B) a term tied to the hazard rate at a given skill level, and C) a term that relates to the shape of the distribution above that skill level. These three factors strike a balance between the efficiency costs of labor supply distortion caused by the marginal tax imposed at a given skill level, and the benefits derived from redistributing additional income collected from workers above that skill level in a non-distortionary fashion.

[3] There are a few papers that bring statistical concerns to an optimal taxation problem. See Lockwood, Sial, and Weinzierl (2021) and Chang and Wu (2023). Our work differs from these studies in modeling statistical concerns over an infinite dimensional object—the entire type distribution—using a non-Bayesian robust control approach, while they model uncertainty over a finite set of parameters that are handled with Bayesian techniques. Those papers found that parametric uncertainty in their environment leads to a more progressive tax system while we find the optimal tax schedule to be less progressive relative to the baseline.

[4] See, for example, Adam and Marcet (2011), Hansen and Sargent (2012), Karantounias (2013), Bhandari, Borovička, and Ho (2022), and Hansen and Miao (2022).

We illustrate that in our context, the marginal tax rate adheres to a modified Diamond–Saez ABC formula. The difference lies in the replacement of the hazard rates and distributions with the worst-case distribution, which is determined as part of the min-max problem. This ex-post Bayesian representation helps us analyze the differences in the optimal tax schedule with and without concerns for misspecification of the type-distributions by breaking it down into two logical components: 1) what influences the shape of the worst-case distribution, and 2) how the tax schedule is pinned down by the Diamond–Saez ABC formula given the shape of the worst-case distribution. Since the worst-case distribution is an endogenous object that relies on the shape of the tax function, the solution to the robust optimal tax problem is a fixed-point that simultaneously characterizes the optimal tax schedule and the worst-case distribution.

Our first theoretical result is the analysis of the top tax rate in the case where the benchmark distribution has a Pareto tail. In the baseline, when the penalty is infinite and the planner is therefore certain that the benchmark distribution is the correct type distribution, the Diamond–Saez ABC formula shows that the top tax rate approaches a positive finite value that depends on the labor supply elasticity and the Pareto tail parameter. This tax rate is quantitatively quite large—around 70–75% for reasonable elasticity and tail parameters. Intuitively, the thick tail of the productivity distribution means that the government can always collect sufficient revenues from the right of any given threshold productivity to offset the cost of distortions due to an increase in the marginal tax rate at that threshold. For similar reasons, for distributions with bounded support or, more generally, with thin tails, the top tax rate approaches zero.

Now consider the problem faced by a government which is concerned that the underlying bechmark Pareto distribution is misspecified. We first show that the ratio of densities of the worst-case distribution and the benchmark Pareto distribution has an exponential tilting expression familiar from the robustness literature. The worst-case density shifts mass away from worker types who are valuable to the government but the magnitude of this reweighting is disciplined by the entropy penalization of the statistical discrepancy between the benchmark and the worst-case distributions. In our context, an individual's value to the government is determined by two components: first, their utilitarian contribution to the welfare objective function, which depends on the individiual's indirect utility under the optimal allocation, and second, their contribution to easing the government budget constraint, which hinges on the net tax revenue the government collects from that individual weighted by the marginal social value of a unit of consumption.

Moving mass away from the right tail of the productivity distribution is costly to the planner and hence desired by the adverse player in the min-max problem because each high-productivity individual generates substantial tax revenue that a redistributive government can use to provide transfers to low-skilled individuals. While tax revenues raised from a particular productivity type rely on the product of productivity level and the mass of agents at that productivity level and decrease as we move further into the right tail, the entropy cost of shifting mass from the right tail diminishes much more rapidly because it scales only with the density. Consequently, in the minimization part of the min-max problem, gains from skimming the density in the right tail grow as we move further into the tail, regardless of the penalty value. We demonstrate that the optimal top tax rate gradually reduces to zero, with the worst-case distribution approaching a distribution with an exponentially decaying density, or a thin tail, in spite of the benchmark distribution being a Pareto that has a polynomial decay in the tail, or, more generally, any other fat-tailed distribution. Thus, the mechanism formalizes the practical concern of policymakers to deal with welfare-relevant aspects of tail behavior that are very hard to detect.

The asymptotic top tax rate is necessarily a statement about limits, and may not be relevant if marginal tax rates approach the zero limit very slowly. To investigate that, we study the elasticity of the marginal tax rate with respect to income. We show that in the limit as productivity becomes large, the worst-case distortion is dominated by planner's concerns about the amount of tax revenue raised from those types. The slower the decay in the marginal tax rate, the higher the revenue collected from the very productive individuals. But higher revenue means these individuals are more valuable to the government, and the adverse agent in the min-max problem has stronger incentives to shift the worst-case density away from them. As their mass becomes smaller, the redistributive gains from taxing high-productivity individuals decline, and the marginal tax rate dictated by the Diamond–Saez formula falls faster. We show that this fixed point is resolved in a unique value such that the elasticity of the marginal tax rate with respect to income equals minus one half. Remarkably, this asymptotic decay rate of the marginal tax rate is independent of any primitive parameters, such as preferences of the household or the government or those characterizing the distribution of skills.

To examine the complete tax schedule rather than just the top tax rate, we resort to a numerical solution. The benchmark distribution is calibrated as an exponentially modified Gaussian (EMG) distribution for the logarithm of productivity. We rely on the parameters reported by Heathcote and Tsujiyama (2021), who estimated them using data from the Survey of Consumer Finances. Although the top tax rate is not influenced by the specific value of the penalty parameter that controls the degree of misspecification concerns about the shape of the productivity distribution, the overall shape of the tax function is. As suggested in the literature (see Anderson et al. (2003)), we calibrate the penalty parameter to target a specific level of detection error probability. This probability measures the failure rate of a likelihood ratio test in distinguishing two alternative distributions (in our case, the worst-case and benchmark distributions), given a specific sample size of draws from the distribution. The detection error probability ranges from zero, when the distributions are easily distinguishable, to one half, when it is impossible to differentiate between the distributions. We use the sample size from the Survey of Consumer Finances and target a detection error probability of 10% to calibrate our penalty parameter, a number frequently used in the literature.

We evaluate our findings against the baseline scenario, in which the penalty parameter is infinite. We observe significant effects of the presence of misspecification concerns on the marginal tax rate. For instance, households earning more than $500,000 face a baseline tax rate exceeding 70% in the absence of misspecification concerns, but under our preferred calibration, the optimal tax rate peaks at 55% for earnings around $250,000. This rate falls to 40% for households with incomes of $5 million, ultimately diminishing to zero, as suggested by our theory. These lower tax rates result in an approximately 2% increase in output but also lead to about a 7% reduction in transfers to the lowest-income households. When targeting lower detection error probabilities, the maximum marginal tax rates can drop as low as 30%, leading to a 10% surge in aggregate output and a substantial 50% cut in transfers.

Finally, we consider a case in which the government is uncertain not just about the distribution of labor productivity but the joint distribution of labor productivity and labor supply elasticities. The theoretical tools to analyze taxation with multidimensional types is in a nascent stage, and we therefore only study this case numerically in a Ramsey fashion via a splines-based approximation of the nonlinear tax function.[5] Similar dynamics are at play in this scenario, and the worst-case distribution not only slims down the right tail in the skill dimension but also endogenously strengthens the correlation between labor supply elasticity and labor productivity. As observed earlier, mass is shifted away from types who contribute significantly to tax revenues. Consequently, the optimal tax rates that emerge are lower than those in the baseline model, which disregards uncertainty about types.

The rest of the paper is structured as follows. Section 2 describes the simple model with one-dimensional types and scalar penalty. Section 3 contains our main theoretical results about the top marginal tax rate and the rate of convergence. Section 4 uses a calibrated economy to study the full tax schedule. Section 5 extends the environment to type-dependent penalty, and Section 6 to multi-dimensional types. Section 7 concludes.

[5]See Golosov and Krasikov (2023) for a theoretical approach to a two-dimensional Mirrlees setting applied to the taxation of couples.

2 Model

The economy is populated by a continuum of workers indexed by their productivity z distributed according to density $f(z)$ with a continuous support $[\underline{z}, \bar{z}] \subseteq \mathbb{R}_+$ where $\bar{z}$ may be infinite. Productivity types are private information of the worker. A worker with productivity z supplying labor n produces income $y = zn$. The worker solves the utility-maximization problem

$$\max_{c,n} U(c,n) \quad \text{s.t. } c = zn - T(zn)$$

where $U(c,n)$ is a strictly concave and differentiable utility function representing worker's preferences over consumption and hours worked, and $T(y)$ is the tax levied on income y. Taking the tax function as given, worker's optimal choice of labor supply yields the condition

$$U_c(c,n)\left(1 - T'(zn)\right)z + U_n(c,n) = 0. \tag{1}$$

Denote the optimal choice of consumption and labor $\mathcal{C}(z;T)$ and $\mathcal{N}(z;T)$, respectively, the resulting output $\mathcal{Y}(z;T) = z\mathcal{N}(z;T)$, and the associated indirect utility function $\mathcal{U}(z;T)$. For the rest of the paper, we drop T as an explicit argument of functions $\mathcal{C}(\cdot)$, $\mathcal{N}(\cdot)$, $\mathcal{Y}(\cdot)$, and $\mathcal{U}(\cdot)$.

The government is in charge of choosing the tax schedule T as a function of observed income y. Taxes are levied for redistribution purposes and to pay for an exogenous government expenditure B. The government welfare objective is given by

$$\mathbb{E}[\psi\mathcal{U}] = \int_{\underline{z}}^{\bar{z}} \psi(z)\,\mathcal{U}(z)\,f(z)\,dz$$

where $\psi(z)$ is a Negishi (1960) welfare weighting function that satisfies $\mathbb{E}[\psi] = 1$. For example, $\psi(z) = 1$ implies a utilitarian planner, while $\psi(z) = \delta_{\underline{z}}(z)/f(z)$ where $\delta_{\underline{z}}(z)$ is the Dirac delta function yields the Rawlsian welfare criterion. In the absence of model misspecification concerns, the government chooses the tax schedule so as to maximize the welfare objective subject to the budget constraint

$$\mathbb{E}[T(\mathcal{Y})] \geq B.$$

We study the optimal taxation problem in a situation when the government is concerned that the underlying distribution of productivity types $f(z)$ is misspecified. In the spirit of Hansen and Sargent (2001a,b), the government contemplates a set of alternative type distributions $\widetilde{f}(z)$ that are statistically close to the 'benchmark' distribution $f(z)$. We denote $m(z) = \widetilde{f}(z)/f(z)$ the likelihood ratio between the benchmark and the alternative distribution, and $\widetilde{\mathbb{E}}[\cdot]$ the expectation operator under the distribution $\widetilde{f}(z)$. By construction,

$\mathbb{E}[m] = 1$. For any integrable function $X(z)$, the Radon–Nikodým theorem implies

$$\widetilde{\mathbb{E}}[X] = \int_{\underline{z}}^{\bar{z}} X(z)\widetilde{f}(z)\,dz = \int_{\underline{z}}^{\bar{z}} X(z)\,m(z)\,f(z)\,dz = \mathbb{E}[mX].$$

The degree of statistical distinguishability of the two distributions $f(z)$ and $\widetilde{f}(z)$ is represented by their relative entropy

$$\mathbb{E}[m\log m] = \int_{\underline{z}}^{\bar{z}} m(z)\log m(z)\,f(z)\,dz.$$

The relative entropy is nonnegative, and is equal to zero if an only if $m \equiv 1$ with probability one. Alternative distributions $\widetilde{f}(z)$ that are statistically easier to distinguish from $f(z)$ yields a larger relative entropy.

The government desires to choose a tax function that would perform well across the set of alternative type distributions $\widetilde{f}(z)$ that are statistically not too distinct from the benchmark distribution $f(z)$. This leads to the maxmin problem

$$\max_{T} \min_{\substack{m \\ \mathbb{E}[m]=1}} \mathbb{E}[m\psi\mathcal{U}] + \theta\mathbb{E}[m\log m] \quad \text{s.t. } \mathbb{E}[mT(\mathcal{Y})] \geq B. \tag{2}$$

The first term in the objective function is equal to the government welfare $\widetilde{\mathbb{E}}[\psi\mathcal{U}]$ evaluated under the alternative distribution $\widetilde{f}(z)$. The second term is an entropy penalty that penalizes distributions with a large statistical distance from the benchmark distribution. The degree of penalization is controlled by the parameter θ. The budget constraint dictates that the tax function must balance the budget under the alternative distribution, $\widetilde{\mathbb{E}}[T(\mathcal{Y})] \geq B$.

The government problem (2) leads to an optimal tax function that is robust to misspecifications of the type distribution that adversely affect the government objective. The problem can be interpreted as a two-player game in which the government faces an malevolent nature that chooses alternative distributions with the most adverse welfare consequences for the contemplated tax function.

Since $\widetilde{f}(z) = m(z)f(z)$, the likelihood ratio $m(z)$ plays the role of a weighting function that over- or underweighs the alternative distribution relative to the benchmark. The government problem implies that the malevolent nature exploits both the direct welfare impact as well as the budgetary consequences of adversely chosen distributions. On the one hand, it desires to impose a high $m(z)$ for types with low welfare impact $\psi(z)\mathcal{U}(z)$ to lower $\mathbb{E}[m\psi\mathcal{U}]$. On the other hand, it strives for adverse budgetary consequences by underweighing types for whom $T(\mathcal{Y}(z))$ (net tax payers) is positive while, vice versa, overweighing those for whom $T(\mathcal{Y}(z))$ is negative (net tax recipients), so that the budget constraint $\mathbb{E}[mT(\mathcal{Y})] \geq B$ binds more strongly. At the same time, alternative adverse distributions $\widetilde{f}(z)$ chosen by nature cannot be too distinct from the benchmark so as not to incur a large penalty $\theta\mathbb{E}[m\log m]$.

In line with the literature, we call the distribution $\widetilde{f}(z)$ that solves the minimax problem in (2) the worst-case distribution. A larger value of θ implies that the worst-case distribution is statistically closer to the benchmark. As $\theta \to \infty$, the entropy penalty becomes prohibitive, model misspecification concerns vanish, and we obtain $\widetilde{f} \equiv f$.

2.1 Mirrleesian formulation

In order to solve for the optimal tax function, we follow Mirrlees (1971), and characterize the constrained optimal allocation under the restriction that the tax function can only depend on worker's income. As usual, we cast the problem as a mechanism design problem, focusing on incentive-compatible mechanisms in which workers truthfully reveal their types.

The government offers to workers a menu of allocations $(c(z), y(z))$ indexed by z. Worker of type z chooses a reporting strategy $\hat{z}$ that entitles to consumption $c(\hat{z})$ in exchange for providing output $y(\hat{z})$ that requires labor input $y(\hat{z})/z$. The reporting strategy therefore solves the announcement problem

$$\max_{\hat{z}} U\left(c(\hat{z}), \frac{y(\hat{z})}{z}\right).$$

Incentive-compatibility requires that the optimal report satisfies $\hat{z} = z$. The first-order necessary condition evaluated at $\hat{z} = z$ yields

$$U_c\left(c(z), \frac{y(z)}{z}\right) c'(z) + U_n\left(c(z), \frac{y(z)}{z}\right) \frac{y'(z)}{z} = 0. \tag{3}$$

Totally differentiating the utility function with respect to z at the allocations $(c(z), y(z))$ and plugging in the optimal reporting strategy condition derived in (3), we obtain

$$\frac{dU}{dz} = -U_n\left(c(z), \frac{y(z)}{z}\right) \frac{y(z)}{z^2}. \tag{4}$$

This is a condition on the utility gradient the menu $(c(z), y(z))$ has to satisfy to be locally incentive-compatible (IC). When this condition holds, the worker has no incentives to misrepresent the true type by an infinitesimal deviation. We assume that the function

$$v(c,n) = -n\frac{U_n(c,n)}{U_c(c,n)} \tag{5}$$

is increasing in n for each fixed c. This implies a single-crossing property under which the local incentive-compatibility constraint also implies global incentive compatibility, and allocations $(c(z), y(z))$ that satisfy incentive compatibility are also strictly increasing in z.

Since the IC constraint is type-by-type and does not depend on the underlying distribution, the model misspecification concern on the side of the planner does not alter its form. We can therefore rely on the convenient Hamiltonian formulation of the allocation problem.

The planner is solving

$$\max_{c,y} \min_{m} \int_{\underline{z}}^{\bar{z}} \psi(z)\, U\left(c(z), \frac{y(z)}{z}\right) m(z) f(z)\, dz + \theta \int_{\underline{z}}^{\bar{z}} m(z) \log m(z) f(z)\, dz \tag{6}$$

subject to the IC constraint (4) and the budget constraint

$$\int_{\underline{z}}^{\bar{z}} (y(z) - c(z))\, m(z) f(z)\, dz \geq B. \tag{7}$$

We provide a solution to this problem using a Hamiltonian formulation that yields a modification of the Diamond (1998) and Saez (2001) elasticity formula for the marginal tax rate. Treating $\mathcal{U}$ as the state variable, λ as its co-state, and y and m as control variables, we form the constrained Hamiltonian

$$\begin{aligned} H(\mathcal{U}, y, m, \lambda) \quad = \quad & \psi(z)\,\mathcal{U}(z)\, m(z) f(z) + \theta m(z) \log m(z) f(z) - \chi m(z) f(z) \\ & -\lambda(z)\, U_n\left(c(z), \frac{y(z)}{z}\right) \frac{y(z)}{z^2} + \mu\left[y(z) - c(z)\right] m(z) f(z). \end{aligned} \tag{8}$$

Here, χ and μ are multipliers on the constraints $\mathbb{E}[m] = 1$ and (7), respectively, and $c(z)$ is defined implicitly from the definition of the utility function as $c(z) = C(\mathcal{U}(z), y(z))$.

We derive a general characterization of the problem in detail in Appendix A.1. Here we provide the analysis of a special case, in which we assume that preferences are quasilinear with isoelastic labor disutility

$$U(c, n) = u(c) - v(n) = c - \frac{n^{1+\gamma}}{1+\gamma}. \tag{9}$$

The first-order condition with respect to $m(z)$ in (8) together with the restriction $\mathbb{E}[m] = 1$ yields a characterization of the worst-case distortion in the form of an exponential tilting formula

$$m(z) = \frac{\exp\left(-\frac{1}{\theta}\left[\psi(z)\,\mathcal{U}(z) + \mu T(y(z))\right]\right)}{\int_{\underline{z}}^{\bar{z}} \exp\left(-\frac{1}{\theta}\left[\psi(z)\,\mathcal{U}(z) + \mu T(y(z))\right]\right) f(\zeta)\, d\zeta}, \tag{10}$$

where $T(y(z)) = y(z) - c(z)$ represents the effective tax the allocation imposes on worker of type z. The remaining optimality conditions then imply the formula for the marginal tax

$$\frac{T'(y(z))}{1-T'(y(z))} = (1+\gamma)\,\frac{\widetilde{\Psi}(z)-\widetilde{F}(z)}{1-\widetilde{F}(z)}\,\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} \tag{11}$$

where $\widetilde{F}(z)$ is the cumulative distribution function of the worst-case density, and $\widetilde{\Psi}(z)$ is planner's cumulative welfare weight

$$\begin{aligned}
\widetilde{F}(z) &= \int_{\underline{z}}^{z} \widetilde{f}(\zeta)\,d\zeta = \int_{\underline{z}}^{z} m(\zeta)\,f(\zeta)\,d\zeta \\
\widetilde{\Psi}(z) &= \int_{\underline{z}}^{z} \frac{\psi(\zeta)\,\widetilde{f}(\zeta)}{\int_{\underline{z}}^{\bar{z}} \psi(\xi)\,\widetilde{f}(\xi)\,d\xi}d\zeta.
\end{aligned}$$

The lump sum portion of the tax $T(y(\underline{z}))$ imposed on the least productive worker is then determined so as the whole tax scheme raises sufficient revenue to satisfy the government budget constraint

$$\int_{\underline{z}}^{\bar{z}} T(y(z))\,dz = B.$$

The optimal marginal tax formula is analogous to that of Diamond (1998) and Saez (2001), except that now, it depends on the endogenously determined distribution $\widetilde{f}(z)$ represented by the distortion $m(z)$ in (10). The first term on the right-hand side of (11) captures distortionary effects of taxation on the labor supply, indicating that marginal taxes should be lower when the inverse of the labor supply elasticity γ is low. The second term represents the desire for redistribution, and is bounded above by one. Marginal taxes will be strictly positive when $\widetilde{\Psi}(z) > \widetilde{F}(z)$, indicating a planner that puts higher welfare weights on lower worker types in the first-order stochastic dominance sense. Finally, the third term is determined by the shape of the tail of the type distribution, and it represents the tradeoff that an increase in the marginal tax $T'(y(z))$ causes at a particular z. This marginal tax has an adverse distortionary effect on the labor supply of all workers with type exactly at z, leading to a total output loss $z\widetilde{f}(z)$, while generating the benefit of raising extra revenue in lump sum fashion from all workers with type above z, whose mass is $1-\widetilde{F}(z)$.

The form of the distortion (10) reveals that the model misspecification concerns of the robust planner have a redistributive and a budgetary component. The numerator of the expression for $m(z) = \widetilde{f}(z)/f(z)$ in (10) indicates that the robust planner underweighs worker types who, under the optimal tax policy, receive allocations with high weighted utility $\psi(z)\mathcal{U}(z)$ or those with high net contributions to the planner's budget, $\mu T(y(z))$. The Lagrange multiplier μ converts the tax revenue to utility units under the government welfare function.

The planner uses the tax policy to maximize the social welfare function. The worst-case distribution putting more weight on types with a low $\psi(z)\mathcal{U}(z)$ and less weight on types with a high $\psi(z)\mathcal{U}(z)$ reflects the concern that the chosen tax function achieves lower government welfare $\widetilde{\mathbb{E}}[\psi\mathcal{U}]$ than that measured under the benchmark model, $\mathbb{E}[\psi\mathcal{U}]$. Since insurance is not perfect

At the same time, the government needs tax revenue to achieve the desired redistribution and tax collection goal. Underweighing worker types who deliver high tax revenue $\mu T(y(z))$ and overweighing those who deliver low tax revenue $\mu T(y(z))$ reflects concerns that worker types who contribute substantially to the budget are less abundant than under the benchmark model, making it more challenging to achieve the desired redistribution.

The parameter θ controls the entropy penalty in the planner's problem (2) and hence the degree of model misspecification concerns. A small value of θ reflects more substantial concerns, which leads to stronger exponential tilting in (10). As $\theta \to \infty$, model misspecification concerns vanish, and the worst-case density $\widetilde{f}(z)$ approaches the benchmark model density $f(z)$ in the statistical sense expressed by the entropy penalty $\mathbb{E}[m \log m]$.

Importantly, the tax function $T(y(z))$ and the worst-case distortion $m(z)$ are determined jointly as an outcome of the minimax problem (2). Given the tax function, the worst-case density delivers the lowest possible penalized objective in (2), and vice versa, taking the worst-case density as given, the tax function maximizes the planner's welfare. The solution is a saddle point in the objective function that constitutes an equilibrium in a two-person game between the benevolent government and the malevolent nature.

3 Optimal marginal tax rates at the top

In this section, we provide an analytical characterization of the asymptotic behavior of the tax rate in (11) in the presence of model misspefication concerns.

When the planner is utilitarian with $\psi(z) \equiv 1$, then, due to the quasilinear form of preferences in (9), the motive for redistribution is absent. Equivalent to the case without misspecification concerns, we obtain that marginal taxes are zero, $T'(y(z)) = 0$. Redistributive concerns are therefore induced by a decreasing welfare weighting function $\psi(z)$.

Here we focus on the case in which there exists a $\hat{z}$ such that $\psi(z) = 0$ for all $z \geq \hat{z}$. The planner hence puts a zero welfare weight on the right tail of the worker distribution. In this case, $\widetilde{\Psi}(z) = 1$ for $z \geq \hat{z}$, and the tax formula becomes

$$\frac{T'(y(z))}{1 - T'(y(z))} = (1+\gamma)\,\frac{1 - \widetilde{F}(z)}{z\widetilde{f}(z)} \tag{12}$$

Since the second term in the tax formula (11) is bounded above by one, the current case yields the highest possible tax rate in the tail of the type distribution across all alternative welfare functions. The worst-case distortion for $z \geq \hat{z}$ then becomes

$$m(z) = \bar{m} \exp\left(-\frac{\mu}{\theta} T(y(z))\right), \tag{13}$$

where $\bar{m}$ is a normalization constant that assures $\mathbb{E}[m] = 1$, corresponding to the reciprocal of the denominator in (10). The misspecification concerns in the right tail of the distribution therefore do not involve the welfare of high-type agents, and only reflect the concerns about the budgetary consequences of not having sufficiently many high-type workers who contribute to the budget.

In Appendix C, we provide the analysis of the more general case that relaxes the assumption of quasilinear preferences in (9) and also treats more general welfare weighting functions so that the welfare concern term $\psi(z)\mathcal{U}(z)$ in (10) is not zero in the right tail. It turns out that in most cases, the budgetary concern $\mu T(y(z))$ dominates, and when it does not, the additional welfare concern further reinforces our results.

3.1 Zero marginal taxes at the top

In the absence of model misspecification concerns, $m(z) \equiv 1$, which implies that the worst-case distribution in (12) corresponds to the exogenously specified benchmark model. The limiting tax rate then depends on the shape of the benchmark distribution. When the distribution $f(z)$ is sufficiently thick-tailed, then the marginal tax rate determined (12) has a strictly positive limit. For example, in the case of the Pareto distribution with shape parameter α, we have $(1 - F(z)) / (zf(z)) = \alpha^{-1}$. On the other hand, bounded or thin-tailed distributions, such as normal or lognormal, imply a zero limit. This has lead to widely differing policy prescriptions about the range of recommended marginal tax rates the planner should impose on top incomes.[6]

When model misspecification concerns are present, we characterize the shape of the worst-case density

$$\widetilde{f}(z) = m(z) f(z) = \bar{m} \exp\left(-\frac{\mu}{\theta} T(y(z))\right) f(z). \tag{14}$$

The distribution $\widetilde{f}(z)$ remains continuous, which implies that $\lim_{z \to \bar{z}} \widetilde{F}(z) = 0$. When $\bar{z}$ is finite, the conclusion about the top marginal tax rate is the same as without model

[6] For example, Diamond and Saez (2011) find a mid-range estimate for the top marginal tax of 73%, based on labor supply elasticity $\gamma^{-1} = 0.25$ and a Pareto distribution of types z with shape parameter $\alpha = 1.875$ (under the given elasticity of labor supply, this translates to a Pareto distribution of incomes $y(z)$ with shape parameter $\alpha_y = 1.5$).

misspecification concerns, and $\lim_{z\to\bar{z}} T'(y(z)) = 0$. To see this, notice that in order for the top marginal tax rate in (12) to be different from zero, we need $\lim_{z\to\bar{z}} z\widetilde{f}(z) = 0$. In this case, we can apply L'Hôpital's rule to obtain

$$\lim_{z\to\bar{z}} \frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} = \lim_{z\to\bar{z}} -\frac{1}{1+z\frac{d\log \widetilde{f}(z)}{dz}} = 0,$$

where the conclusion follows from the fact that $\lim_{z\to\bar{z}} \log \widetilde{f}(z) = -\infty$ for a finite $\bar{z}$, which means that the derivative must diverge to infinity as well.

We therefore focus on the more interesting case when $\bar{z} = \infty$. It turns out that the marginal tax rate at the top still asymptotically converges to zero.

Assumption 1. *There exists a $\hat{z}$ such that the type density $f(z)$ under the benchmark distribution is continuously differentiable on $[\hat{z}, \bar{z})$, and $zf(z)$ is strictly decreasing on $[\hat{z}, \bar{z})$, with*

$$\lim_{z\to\bar{z}} \frac{d\log f(z)}{d\log z} < -1,$$

with the limit possibly being $-\infty$.

Theorem 3.1. *Assume that worker's preferences are given by the quasilinear form (9), the type distribution satisfies Assumption 1 with $\bar{z} = \infty$, and $\theta < \infty$. Then the marginal tax rate vanishes to zero at the top:*

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

We formally prove the theorem in Appendix B.1. The proof requires a technical treatment of the existence of the limit but conditional on its existence, the result is intuitive. Denote $T'_{rat}(y_{rat}(z))$ the opitmal marginal tax rate in the model without model misspecification concerns, $\theta = \infty$. Then

$$\lim_{z\to\infty} \frac{T'_{rat}(y_{rat}(z))}{1-T'_{rat}(y_{rat}(z))} = (1+\gamma) \lim_{z\to\infty} \frac{1-F(z)}{zf(z)} = (1+\gamma) \lim_{z\to\infty} \frac{1}{-\frac{d\log f(z)}{d\log z} - 1} < \infty,$$

where the second equality follows from an application of L'Hôpital's theorem, and the final inequality is implied by Assumption 1. This yields $\lim_{z\to\infty} T'_{rat}(y_{rat}(z)) < 1$.

Further, the single-crossing property (5) implies that the optimal incentive compatible scheme implies that output $y(z)$ is strictly increasing in worker's type z, and since the marginal tax is strictly positive, $m(z)$ in (14) is strictly decreasing. This implies that, for any $z \geq \hat{z}$,

$$\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} < \frac{1-F(z)}{zf(z)}$$

and hence also $\lim_{z\to\infty} T'(y(z)) \leq \lim_{z\to\infty} T'_{rat}(y_{rat}(z))$.

The limiting tax rate under model misspecification cannot, however, be positive. If it were converging to $\tau_\infty > 0$, then the output function $y(z)$ implied by the optimal labor choice (3) for the case of quasilinear preferences (9) would asymptotically behave as

$$y(z) = \left(1 - T'(y(z))\right)^{\frac{1}{\gamma}} z^{\frac{1+\gamma}{\gamma}} \approx \tau_\infty^{\frac{1}{\gamma}} z^{\frac{1+\gamma}{\gamma}}, \tag{15}$$

and the worst-case distortion as

$$m(z) = \bar{m} \exp\left(-\frac{\mu}{\theta} T(y(z))\right) \approx \bar{m} \exp\left(-\frac{\mu}{\theta}(1-\tau_\infty)\tau_\infty^{\frac{1}{\gamma}} z^{\frac{1+\gamma}{\gamma}}\right). \tag{16}$$

In that case, an application of L'Hôpital's rule to the tax formula (12) yields

$$\lim_{z\to\infty} \frac{T'(y(z))}{1-T'(y(z))} = \lim_{z\to\infty} (1+\gamma) \frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} = \lim_{z\to\infty} \frac{1+\gamma}{\frac{\mu}{\theta} z \frac{d}{dz} T(y(z)) - \frac{d\log f(z)}{d\log z} - 1} = 0, \tag{17}$$

because the first term in the denominator diverges to ∞. This is a contradiction to the assumption that $\lim_{z\to\infty} T'(y(z)) = \tau_\infty > 0$, and hence the tax rate has to covnerge to zero.

The striking result is that the marginal tax rate at the top converges to zero irrespective of the degree of model misspecification concerns. Since we abstracted from welfare concerns at the top of the income distribution, the planner only cares about the budgetary consequences associated with taxing top incomes. From this perspective, the robust planner is concerned that there are fewer high-productivity workers that can be taxed.

The form of the distortion in (13) indicates that the concerns grow proportionally with the marginal social value of the tax revenue $\mu T(y(z))$ that the worker with a given productivity z contributes to the budget. The distortion $m(z)$ therefore becomes more severe as z increases, effectively generating a thinner tail in the worst-case distribution $\widetilde{f}(z)$ in (14). This consequently implies a lower and vanishing optimal marginal tax as $z \to \infty$, since the tradeoff of an increase in the marginal tax $T'(y(z))$ at z that compares the extra benefit of taxing workers above z with the cost of distorting labor supply of workers at z becomes less favorable under the worst-case distribution.

While we have shown the vanishing marginal tax rate result for the case of quasilinear utility and no planner's welfare concerns in the right tail of the type distribution, these results carry over to more general cases. In particular, we show in Appendix C that the result also when utility from consumption is concave, and in the presence of welfare concerns at the top of the type distribution when $\Psi(z)$ only asymptotically converges to one.

The intuition for these generalizations is straightforward. These generalizations yield the more general form of the distortion of the type distribution represented by expression

(10). The case of concave utility leads to a different marginal social value of public funds μ and to a different optimal tax function trajectory $T(y(z))$ but since μ remains strictly positive, we only need to show that the tax revenue $T(y(z))$ under the optimal tax continues to diverge to ∞ as $z \to \infty$.

Adding welfare concerns corresponds to a nonzero $\psi(z)$ function as $z \to \infty$. Since $\mathcal{U}(z)$ is increasing in z, this can only lead to a more strongly decreasing distortion $m(z)$ since the budgetary concerns and welfare concerns of not having sufficiently many workers with high types who contribute substantially both to the budget as well as to the welfare objective reinforce each other.

3.2 Rate of decay of the tax rate

Theorem 3.1 shows that the marginal tax rates at the top vanish to zero irrespective of the underlying type distribution. However, the theorem does not determine the rate of convergence. In quantitative applications, the rate of convergence matters because it sharpens information about the practical importance of the asymptotic behavior for finite values of productivity z.

In this subsection, we derive this rate of convergence. We first state it in the form of a theorem, and then provide a derivation of the result that leads to a specific differential equation that will be helpful in numerical implementation of the full characterization of the optimal tax scheme.

In order to simplify the formula, strengthen Assumption 1 and assume that the tail of the benchmark density $f(z)$ is given by a Pareto distribution with shape parameter α. This is the prototypical choice that yields strictly positive marginal taxes in absence of model misspecification.

We characterize the following result directly in income space, treating $y = y(z)$ as the endogenous income of each worker. Previous results imply that $\lim_{z\to\infty} y(z) = \infty$, so that limits $z \to \infty$ and $y \to \infty$ are equivalent.

Theorem 3.2. *Assume that worker's preferences are given by the quasilinear form (9), underlying productivity has a right tail that is Pareto distributed with shape parameter α, and $\theta < \infty$. Then the limiting tax rate satisfies*

$$\lim_{y\to\infty} \frac{\mu}{\theta} \left[T'(y)\right]^2 y = \gamma. \tag{18}$$

In particular, the limiting elasticity of the marginal tax with respect to income, if the limit exists, is equal to

$$\lim_{y\to\infty} \frac{d \log T'(y)}{d \log y} = -\frac{1}{2}. \tag{19}$$

The proof of the theorem is provided in Appendix B.3. Expression (18) depends on one endogenous object, the marginal social value of wealth, which needs to be determined separately. The expression implies that the marginal tax rate has to decay to zero at rate

$$\log T'(y) \approx -\frac{1}{2}\log y,$$

which implies that the elasticity of the marginal tax rate with respect to income has a limit, this limit has to be equal to $-\frac{1}{2}$.

In Section 3.4 and Appendix C, we again show how this result generalizes when we relax the assumptions of the theorem. Broadly speaking, the elasticity expression (19) remains robust but the rate of decay of the marginal tax rate to zero may be even faster when, for example, planner's welfare concerns for the top income earners are sufficiently strong, or when the type distribution under the benchmark density is already sufficiently thin-tailed to begin with.

Remarkably, the elasticity of the marginal tax rate with respect to income (19) does not depend on any of the parameters of the model. In order to understand the result, assume that for incomes $y \geq \hat{y}$, the marginal tax rate $T'(y)$ is sufficiently well approximated by a constant elasticity function with elasticity ν. Taking a $\bar{y} \gg \hat{y}$, we can write

$$T(\bar{y}) = T(\hat{y}) + \int_{\hat{y}}^{\bar{y}} T'(y)\, dy \approx T(\hat{y}) + \int_{\hat{y}}^{\bar{y}} (1+\nu)\, y^{\nu} dy = T(\hat{y}) - \hat{y}^{1+\gamma} + \bar{y}^{1+\nu}.$$

This means that asymptotically, for high-productivity types $\bar{z}$ with income $\bar{y} = y(\bar{z})$, the worst-case distortion behaves as

$$m(\bar{z}) = \bar{m}\exp\left(-\frac{\mu}{\theta}T(\bar{y})\right) \approx \widetilde{m}\exp\left(-\frac{\mu}{\theta}\bar{y}^{1+\nu}\right) \tag{20}$$

where $\widetilde{m}$ absorbs the contribution of terms $T(\hat{y}) - \hat{y}^{1+\gamma}$. At the same time, the output function $y(z)$ in (15) implies that as $T'(y) \to 0$, we can approximate $y(z) \approx z^{\frac{1+\gamma}{\gamma}}$. Applying L'Hôpital's rule as in (17), we obtain

$$\lim_{z\to\infty} \frac{T'(y(z))}{1 - T'(y(z))} = (1+\gamma)\lim_{z\to\infty} \frac{1}{\frac{\mu}{\theta} z T'(y(z))\, y'(z) - \frac{d\log f(z)}{d\log z} - 1} \tag{21}$$

The denominator on the right-hand side is then dominated by the the first term,

$$\frac{\mu}{\theta} z T'(y(z))\, y'(z) \approx \frac{\mu}{\theta}(1+\nu)\frac{1+\gamma}{\gamma} z^{\frac{1+\gamma}{\gamma}(\nu+1)}.$$

Comparing the elasticities of the left-hand and right-hand side of (21) with respect to the type z yields

$$\frac{1+\gamma}{\gamma}\nu = -\frac{1+\gamma}{\gamma}(\nu+1),$$

from which we obtain $\nu = -\frac{1}{2}$.

Intuitively, expression (21) indicates that the optimal rate of decay of the tax rate balances two forces, the effect on the tax revenue $T(y(z))$ collected from high worker types z, and the effect this tax revenue has on the worst-case distortion $m(z)$. If the decay rate was higher (a more negative ν), then the tax revenue $T(y(z))$ would grow more slowly as $z \to \infty$. This would consequently diminish the budgetary concerns of the model misspecification in (20), the worst-case density $\widetilde{f}(z)$ would be less distorted with a thicker tail, and the optimal tax formula would indicate a more gradual decay of the tax rate. The chain of arguments is reversed if the decay rate was lower (a less negative ν).

The elasticity choice $\nu = -\frac{1}{2}$ thus uniquely identifies the asymptote of the saddle point between the maximization problem that seeks the optimal tax rate, and the minimization problem that finds the model misspecification with the most adverse consequences for the planner.

3.3 Phase diagram

We formalize the proof of Theorem 3.2 by deriving a differential equation for the marginal tax rate $T'(y)$. After differentiating the tax formula (12) with respect to the productivity z and a sequence of algebraic manipulations, we obtain the differential equation

$$-\frac{T''(y)\,y}{1-T'(y)} = -\left[2-\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]^{-1}\left[\frac{\mu}{\theta}\left[T'(y)\right]^{2}y-\gamma+\gamma\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]. \quad (22)$$

This equation depends on z only implicitly through $y(z)$, and we therefore can treat the equation directly as a function of income y. This is a first-order differential equation for the unknown marginal tax $T'(y)$, and the unique strictly positive solution is pinned down by the terminal condition $\lim_{y\to\infty} T'(y) = 0$. More detail concerning the analysis of this differential equation is provided in Appendix B.2.

The left-hand side of the differential equation is the elasticity of the take-home rate with respect to income

$$-\frac{T''(y)\,y}{1-T'(y)} = \frac{d\log\left(1-T'(y)\right)}{d\log y}.$$

This elasticity must converge to zero as $y \to \infty$. In order for the right-hand side of (22) to converge to zero, the last bracket has to converge to zero. Since $\lim_{y\to\infty} T'(y) = 0$, we obtain equation (18) in Theorem 3.2.

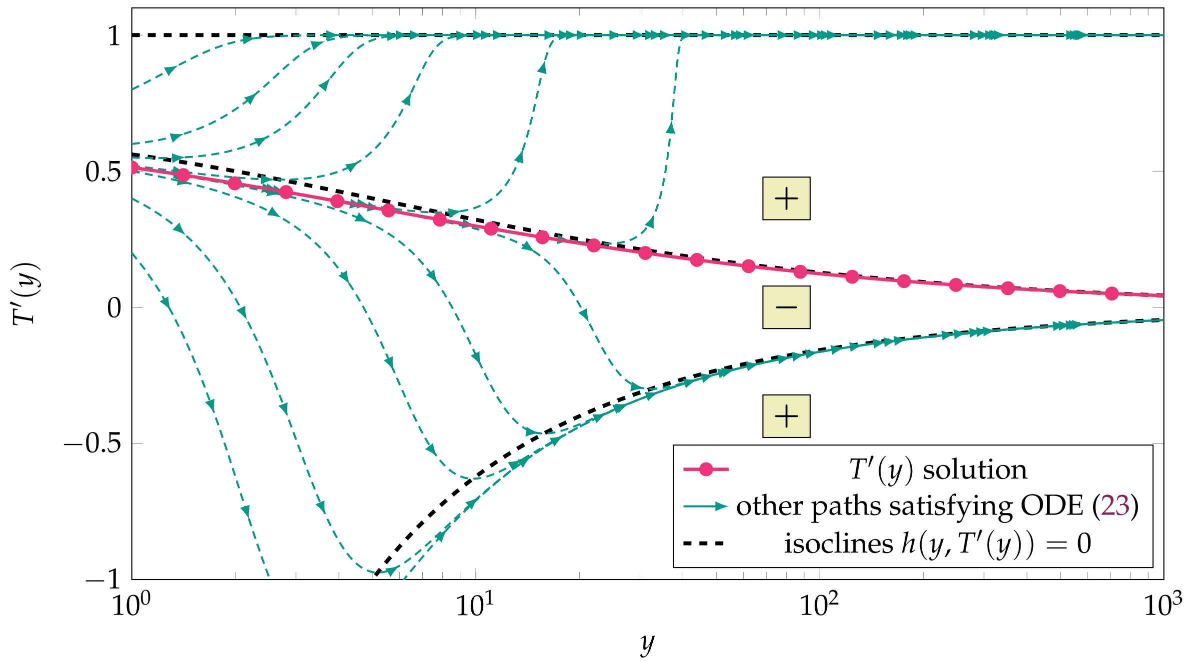

Figure 1: Phase diagram for differential equation (23). The dashed solid lines correspond to isoclines $h(y, T'(y)) = 0$, and the sign of $h(y, T'(y))$ in between these lines is depicted in the boxes. The magenta line with circles corresponds to the unique strictly positive solution that satisfies the terminal condition. Green dashed lines are other trajectories that satisfy (23). The parameters are $\mu = 1$, $\theta = 1$, $\alpha = 1.5$, $\gamma = 2$.

Let us denote

$$h\left(y,\tau\right) = \frac{1-\tau}{y}\left[2 - \frac{1+\gamma+\alpha}{1+\gamma}\tau\right]^{-1}\left[\frac{\mu}{\theta}\tau^2 y - \gamma + \gamma\frac{1+\gamma+\alpha}{1+\gamma}\tau\right].$$

Then the differential equation (22) can be rewritten as

$$T''\left(y\right) = h\left(y, T'\left(y\right)\right). \tag{23}$$

In Figure 1, we plot the phase diagram for this differential equation. The differential equation exogenously fixes the marginal social value of resources μ, which must be determined separately and jointly with the lump-sum tax on the lowest worker type $T\left(y\left(\underline{z}\right)\right)$ so that the planner's budget constraint holds. We reiterate that this characterization holds for the right tail of the type distribution that has a Pareto density and for which the planner has no welfare concerns. This solution for the right tail can then be combined with that for the rest of the type distribution that possibly has a different shape and for which welfare concerns are nonzero, using the general expression for the marginal tax rate given by (10)–(11). The marginal social value of resources μ connects the solutions.

The black dashed lines are the isoclines for which the slope of the marginal tax curve is equal to zero. Since the solution must satisfy the terminal condition $\lim_{y\to\infty} T'(y) = 0$ and must be strictly positive, these isoclines bound the solution into the positive part of the region denoted with the minus sign.

The solution is depicted with the magenta line with circles. Taking this trajectory from the perspective of an initial value problem, the solution constitutes an unstable saddle path. Starting from any other initial condition, the trajectories satisfying equation (23) converge to one of two stable saddle paths visible in the graph, so they either converge to one, or become negative. This also verifies that the terminal condition $\lim_{y\to\infty} T'(y) = 0$ pins down a unique strictly positive solution. In addition, this solution for the marginal tax rate must be strictly declining.

3.4 Generalizations

The preceding analysis studies the case when workers have quasilinear preferences and the planner has no welfare concerns for high-type workers in the right tail of the productivity distributions. In this subsection, we briefly discuss generalizations of these results, with detailed calculations provided in Appendix C. The central insight is that the marginal tax converging to zero at an exponential rate equal to (at least) $-\frac{1}{2}$ is a robust result that holds in a range of extensions.

3.4.1 Concave separable preferences

We first consider the case of general separable isoelastic preferences of the form

$$U(c,n) = \frac{c^{1-\rho}}{1-\rho} - \psi \frac{n^{1+\gamma}}{1+\gamma} \tag{24}$$

where ρ is is the inverse of the consumption elasticity. We still focus on the right tail of the type distribution for which we assume no welfare concerns on the planner's side, $\psi(z) = 0$. In this case, the optimality conditions for the Hamiltonian (8) yield the optimal tax formula in the form

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

where

$$\widetilde{f}(z) = \kappa^{-1}\,(c(z))^{\rho}\, m(z)\, f(z) \tag{25}$$

is now the inverse marginal utility weighted density under the worst-case model, with normalization constant

$$\kappa = \int_{\underline{z}}^{\bar{z}} (c(\zeta))^{\rho} \, m(\zeta) \, f(\zeta) \, d\zeta = \widetilde{\mathbb{E}}\left[c^{\rho}\right],$$

and $c(z) = y(z) - T(y(z))$. The function $\widetilde{F}(z)$ is the corresponding cumulative distribution function associated with density $\widetilde{f}(z)$ defined in (25).

Theorem 3.3. *Assume that worker's preferences are given by the separable form (24), underlying productivity has a right tail that is Pareto distributed, and $\theta < \infty$. Then the marginal tax rate vanishes at the top,*

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

More specifically, the limiting tax rate satisfies

$$\lim_{y \to \infty} \frac{\mu}{\theta} \left[T'(y)\right]^2 y = (\gamma + \rho) \, \widetilde{\mathbb{E}}\left[c^{\rho}\right], \tag{26}$$

and hence the limiting elasticity of the marginal tax with respect to income is equal to

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

This theorem generalizes Theorems 3.1 and 3.2 to the case of concave marginal utility of consumption. Setting $\rho = 0$ recovers the special quasilinear case. As in the quasilinear case, the zero limiting marginal tax is preserved in cases when the density that characterizes the right tail of the distribution is thinner than Pareto, even though the decay rate may then be faster than $-\frac{1}{2}$.

Intuitively, decreasing marginal utility from consumption effectively reduces the elasticity of labor supply with respect to the productivity. But since the zero limiting marginal tax result does not depend on the labor supply elasticity to start with, it is also robust with respect to the introduction of more general separable utility form (24). The only difference is expression (26) that adjusts for the consumption elasticity, with the last term $\widetilde{\mathbb{E}}\left[c^{\rho}\right]$ converting goods to their marginal social value units.

3.4.2 Welfare concerns at the top

When the welfare function does not assign zero weights $\psi(z)$ to top productivity types, the determination of the limiting tax must also take into account the distortions under the worst-case model that are due to welfare concerns. To illustrate the consequences, we consider here the case of a utilitarian planner with $\psi(z) \equiv 1$. The worst-case distortion then takes the form

$$m(z) = \bar{m} \exp\left(-\frac{1}{\theta}\left[\mathcal{U}(z) + \mu T(y(z))\right]\right), \tag{27}$$

and the marginal tax formula can be expressed as

$$\frac{T'(y(z))}{1 - T'(y(z))} = (1+\gamma)\,\frac{\widetilde{\Psi}(z) - \widetilde{F}(z)}{z\widetilde{f}(z)}, \tag{28}$$

where $\widetilde{f}(z)$ is the inverse marginal utility weighted worst-case density (25), $\widetilde{F}(z)$ the corresponding cumulative distribution function, and the cumulative welfare weight $\widetilde{\Psi}(z)$ specializes to

$$\widetilde{\Psi}(z) = \int_{\underline{z}}^{z} \widetilde{f}(\zeta)\, d\zeta.$$

Comparing $\widetilde{F}(z)$ and $\widetilde{\Psi}(z)$, we have $\widetilde{\Psi}(z) > \widetilde{F}(z)$ whenever $\rho > 0$, which reflects the redistributive motives of the utilitarian planner. When $\rho = 0$, the redistributive motive is absent, and $\widetilde{\Psi}(z) = \widetilde{F}(z)$.

The worst-case distortion $m(z)$ in (27) now combines the contributions of welfare and budgetary concerns. The distortion from the utility term $\mathcal{U}(z)$ reflects the concern that there are fewer high-type workers in the distribution, which directly adversely affects the planner's objective function. Since both $\mathcal{U}(z)$ and $T(y(z))$ are strictly increasing in z, both concerns imply a strictly decreasing $m(z)$. Which of the two terms dominates depends on the curvature of the utility function.

Theorem 3.4. *Assume that worker's preferences are given by the separable form (24), underlying productivity has a right tail that is Pareto distributed, the planner is utilitarian with $\psi(z) \equiv 1$, the curvature of the utility function is $\rho > 0$, and $\theta < \infty$. Then*

$$\lim_{y\to\infty} T'(y) = 0,$$

and

$$\lim_{y\to\infty} \frac{d \log T'(y)}{d \log y} = \min\left(-\frac{1}{2}, \rho - 1\right).$$

The theorem shows that the direct welfare concern dominates when the preferences are sufficiently elastic, $\rho < \frac{1}{2}$. In this case, the limited curvature of the utility function implies that the utility term $\mathcal{U}(z)$ grows faster than the tax revenue that determines the budgetary concern. However, the decay rate $-\frac{1}{2}$ derived for the benchmark model constitutes the slowest rate of decay we can anticipate.

3.4.3 Power divergence functions

The objective function (6) of the robust planner penalizes deviations from the benchmark using an entropy penalty, also known as the Kullback–Leibler divergence. While entropy is a natural penalty choice from a statistical perspective, the tendency toward lower progressivity holds for more general divergence functions. Here, we consider the Cressie and Read (1984) class of power divergence functions analyzed, for example, in Almeida and Garcia (2017) or Borovička et al. (2016). The class of divergences is given by $\mathcal{E}_\eta(m) = \mathbb{E}\left[\phi_\eta(m)\right]$ with

$$\phi_\eta(m) = \frac{m^{1+\eta} - 1}{\eta(1+\eta)},$$

where $m(z) = \widetilde{f}(z)/f(z)$ and $\eta \in \mathbb{R}$. Divergences $\mathcal{E}_\eta(m)$ for $\eta \in \{-1, 0\}$ are constructed by appropriate limiting arguments, yielding the entropy $\mathcal{E}_0(m) = \mathbb{E}[m \log m]$ in the limit as $\eta \to 0$. Relative to entropy, power divergences $\mathcal{E}_\eta(m)$ for $\eta > 0$ penalize relatively more the deviations in the left tail of the distribution, while for $\eta < 0$ they penalize more strongly the right tail.

Appendix C provides more detail on power divergences. It also shows that replacing $\theta\mathcal{E}_0(m)$ with $\theta\mathcal{E}_\eta(m)$ in the planner's objective function (6) leads, in the case when utilitarian concerns are absent in the right tail of the productivity distribution, to the same optimal tax formula (12) but the worst-case distortion (13) is now given by

$$m(z) = \left[\frac{\eta}{\theta}\left(\chi - \mu T(y(z))\right)\right]^{\frac{1}{\eta}} \tag{29}$$

where χ is the Lagrange multiplier on the constraint $\mathbb{E}[m] = 1$. As in the entropy case, the worst-case distortion is decreasing because the expression for the optimal marginal tax implies that the marginal tax is positive.

Specializing to the case of quasilinear utility and Pareto-distributed productivity under the benchmark distribution, the differential equation for the optimal marginal tax (22) now takes the form

$$-\frac{T''(y)\,y}{1-T'(y)} = -\left[2 - \frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]^{-1}\left[-\frac{\mu}{\lambda}\frac{\left[T'(y)\right]^2 y}{\mu T(y) - \chi} - \gamma + \gamma\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right].$$

This leads to the following result.

Theorem 3.5. *Assume that worker's preferences are given by the quasilinear form (9), the type distribution satisfies Assumption 1 with $\bar{z} = \infty$, the divergence penalty in the planner's problem is $\theta\mathcal{E}_\eta(m)$, and $\theta < \infty$. Then the optimal marginal tax rate $T'(y(z))$ for an agent with type z under the robust planner is lower than under the planner without model misspecification concerns.*

Moreover, assume in addition that the underlying productivity has a right tail that is Pareto distributed with shape parameter α. When $\eta \geq 0$, then the marginal tax rate at the top satisfies

$$\lim_{y \to \infty} T'(y) = 0. \tag{30}$$

When $\eta < 0$, then the marginal tax rate at the top is given by

$$\lim_{y \to \infty} T'(y) = \tau_\eta = \frac{1+\gamma}{1+\gamma+\widetilde{\alpha}} \qquad \text{with } \widetilde{\alpha} = \alpha - \frac{1+\gamma}{\gamma}\frac{1}{\eta} > \alpha. \tag{31}$$

Expressions (30) and (31) show that the asymptotic top marginal tax is continuous in the parameter η that indexes the power divergences. As $\eta \nearrow 0$, it also converges to the entropy case with zero asymptotic marginal tax.

Details of the derivation are provided in Appendix C. To provide intuition, consider first the case $\eta > 0$. In this case, the worst-case distortion formula (29) implies that $\chi - \mu T(y) > 0$. This means that $T(y)$ must be bounded from above, and since the marginal tax $T'(y)$ has to be positive, it must converge to zero.

When $\eta < 0$, expression (29) implies that $\chi - \mu T(y) < 0$, so $T(y)$ can be unbounded. Conjecturing that the limiting marginal tax $\tau_\eta \in (0,1)$, an asymptotic approximation analogous to that in (15) and (16) implies that the worst-case distortion (29) asymptotically behaves as a power function of z. This means that the worst-case density $\widetilde{f}(z) = m(z) f(z)$ behaves asymptotically as a Pareto density, with an adjusted shape parameter $\widetilde{\alpha}$. This shape parameter becomes arbitrarily large as $\eta \nearrow 0$, when the divergence function approaches entropy, and, in this case, the asymptotic marginal tax rate approaches zero.

4 Quantitative application

In Section 3, we provided a theoretical characterization of the tail behavior of the optimal tax function when the planner is concerned about misspecification of the type distribution. We now focus on a quantitative evaluation of the whole tax function. Specifically, we are interested in a plausible calibration of the magnitude of the misspecification concerns, and implications for the relative distortions across the type distribution. We envision that the planner has available survey evidence in the form of a random sample of I observations of workers' productivities $\{z_i\}_{i=1}^{I}$, and ask how informative this sample is about the underlying productivity distribution.

4.1 Quantifying model misspecification concerns

In order to quantify the magnitude of model misspecification concerns, we use detection error probabilities proposed by Anderson et al. (2003). Consider a random sample of independent draws $\left\{z_i^B\right\}_{i=1}^I$ from the benchmark distribution $f(z)$, and a sample $\left\{z_i^A\right\}_{i=1}^I$ drawn from the alternative distribution $\widetilde{f}(z;\theta)$ determined as the worst-case distribution under parameter θ. We evaluate the probability that the random sample $\left\{z_i^B\right\}_{i=1}^I$ is assigned a higher likelihood under the alternative distribution than under the correct benchmark distribution,

$$P\left(\sum_{i=1}^{I}\log\widetilde{f}\left(z_i^B;\theta\right)>\sum_{i=1}^{I}\log f\left(z_i^B\right)\right)=P\left(\sum_{i=1}^{I}\log m\left(z_i^B;\theta\right)>0\right),$$

and, conversely, the probability that the random sample $\left\{z_i^A\right\}_{i=1}^I$ is assigned a higher likelihood under the benchmark distribution than under the correct alternative distribution,

$$P\left(\sum_{i=1}^{I}\log f\left(z_i^A\right)>\sum_{i=1}^{I}\log\widetilde{f}\left(z_i^A;\theta\right)\right)=P\left(\sum_{i=1}^{I}\log m\left(z_i^A;\theta\right)<0\right).$$

The detection error probability is then defined as the average of the two probabilities above,

$$d\left(\theta,I\right)=\frac{1}{2}\left(P\left(\sum_{i=1}^{I}\log m\left(z_i^B;\theta\right)>0\right)+P\left(\sum_{i=1}^{I}\log m\left(z_i^A;\theta\right)<0\right)\right).$$

The detection error probability expresses the chance that the likelihood ratio leads to the errorneous conclusion about which of the two distributions generated the random sample.

The construction implies that $0\leq d\left(\theta,I\right)\leq\frac{1}{2}$, achieving the upper bound when $f(z)$ and $\widetilde{f}(z;\theta)$ are identical. The function $d\left(\theta,I\right)$ is increasing in θ, reflecting the fact that a higher entropy penalty θ implies more similar benchmark and worst-case distributions. In other words, more severe model misspecification concerns captured by a *lower* value of θ lead the planner to consider as plausible a worst-case distribution $\widetilde{f}(z;\theta)$ that is statistically more distinct from the benchmark distribution $f(z)$, leading to a smaller detection error probability $d\left(\theta,I\right)$. The detection error probability is also decreasing in the sample size I, as long as $f(z)$ and $\widetilde{f}(z;\theta)$ are statistically distinguishable.

In order to calibrate θ and the magnitude of the misspecification concerns, we fix a plausible level for the detection error probability $d\left(\theta,I\right)=\bar{d}$. We fix the sample size I to match the sample size of a given survey, and then infer $\theta\left(\bar{d},I\right)$ from the implicit equation

$$d\left(\theta\left(\bar{d},I\right),I\right)=\bar{d}.$$

A more informative survey with a larger sample size I will lead to a higher value $\theta\left(\bar{d}, I\right)$ for a given fixed $\bar{d}$, reflecting the fact that larger sample sizes achieve the same detection error probabilities for distributions that are statistically more alike.

4.2 Model calibration

We base our benchmark model calibration on the results in Heathcote and Tsujiyama (2021), who use labor income data from the Survey of Consumer Finances (SCF) to infer the productivity distribution. Heathcote and Tsujiyama (2021) argue that the SCF provides substantially more information about the right tail of the productivity distribution than other household surveys like the Current Population Survey (CPS).

They also show that the labor income distribution in SCF is very well approximated by using the exponentially modified Gaussian (EGM) distribution for the logarithm of the productivity $x = \log z$. The EGM distribution describes the sum of a normal and an exponential random variable, with density given by

$$\begin{aligned} f_x\left(x; \mu, \sigma, \alpha\right) &= \frac{\alpha}{2} e^{\frac{\alpha}{2}\left(2\mu + \alpha\sigma^2 - 2x\right)} \operatorname{erfc}\left(\frac{\mu + \alpha\sigma^2 - x}{\sqrt{2}\sigma}\right) \\ \operatorname{erfc}\left(x\right) &= \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \end{aligned}$$

This distribution implies that productivity $z = \exp\left(x\right)$ has support $\left(0, \infty\right)$, the left tail follows the log-normal distribution with parameters μ, σ, and the right tail is asymptotically Pareto distributed with shape parameter α. We can therefore invoke the theoretical results derived in Section 3 for the case of the Pareto-distributed right tail.[7] As in Heathcote and Tsujiyama (2021), we choose $\alpha = 2.2$ and $\sigma^2 = 0.142$. The first moments of the logarithm and the level of the productivity distribution are given by

$$E\left[\log z\right] = \mu + \frac{1}{\alpha} \qquad E\left[z\right] = \exp\left(\mu + \frac{1}{2}\sigma^2\right) \frac{\alpha}{\alpha - 1}.$$

We set μ to normalize the average productivity to $E\left[z\right] = 1$.

We further assume a utilitarian planner endowed with a concave separable utility function (24), with parameters $\rho = 1$ implying logarithmic utility from consumption, $\psi = 1$, and $\gamma = 2$. In the benchmark model, we set $B = 0$, implying that the government only taxes for redistributive purposes.

In order calibrate θ, we target the detection error probability of about 10%. For example, Hansen and Sargent (2010) argue that $\bar{d} = 0.2$ is a conservative choice, reflecting that $f\left(z\right)$

[7] Truncating the Gaussian component of the distribution $f_x\left(x\right)$ for a sufficiently high value of x so that the tail has an exact Pareto distribution is quantitatively inconsequential.

and $\widetilde{f}(z;\theta)$ are sufficiently hard to distinguish. In general, the literature uses value in the range between 5% and 20%. We choose $\theta = 20$, which yields $\bar{d} = 0.09$ for a sample size of $I = 3,500$, which corresponds to the number of households in the SCF survey. In what follows, we also provide comparative statics exercises with respect to alternative values of the parameter θ.

4.3 Asymptotic behavior of the marginal tax rate

We start with the following result that characterizes the marginal tax rate for the left and right tail of the productivity distribution.

Lemma 4.1. *When the logarithm of productivity z follows the exponentially modified Gaussian with parameters (μ, σ, α), the following results hold.*

1. *In the rational model ($\theta = \infty$), the asymptotic marginal tax rate is given by*

$$\lim_{y\to\infty} T'(y) = \frac{(\gamma+\rho)(1+\gamma)}{\alpha(\gamma+\rho)+\gamma(1+\gamma)}.$$

2. *Under model misspecification concerns ($\theta < \infty$),*

$$\lim_{z\to\infty} T'(y) = 0. \tag{32}$$

$$\lim_{y\to\infty} \frac{d\log T'(y)}{d\log y} = \min\left(-\frac{1}{2}, \rho - 1\right) \tag{33}$$

3. *For the left tail of the productivity distribution,*

$$\lim_{y\to 0} T'(y) = 0,$$

irrespective of the value of θ.

A sketch of the proof is provided in Appendix C. The first result evaluates the tax formula (28) that reflects planner's utilitarian concerns for the workers in the right tail of the productivity distribution, and the concave shape of the utility function. The limit aligns with the case when the tail is exactly Pareto distributed, since the contribution of the log-normal component in the right-tail vanishes. The second result restates results shown in Theorem 3.4.

Finally, the third result follows from the log-normal shape of the productivity distribution at zero. To understand the reason why the marginal tax rate is zero irrespective of the presence of misspecification concerns, we note that because the marginal utility of consumption is infinite at zero consumption, the planner optimally provides a finite, strictly

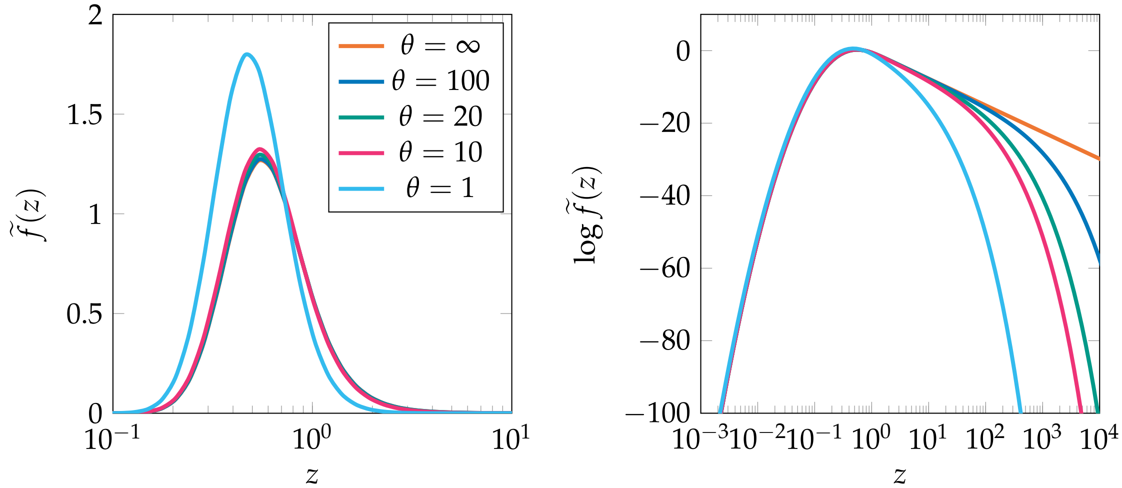

Figure 2: Worst-case distributions $\widetilde{f}(z)$ for alternative levels of misspecification concerns given by θ. The case $\theta = \infty$ corresponds to the rational benchmark for which $\widetilde{f}(z) = f(z)$.

positive transfer $T(\underline{z})$ to workers with zero productivity $\underline{z} = 0$. Such a transfer has infinite marginal social value against a finite social marginal cost of resources. Consequently, $\mathcal{U}(\underline{z})$ is finite, and the distortion $m(z)$ in (27) is bounded and bounded away from zero in the neighborhood of $\underline{z} = 0$. Since, as we show in the appendix, the limiting tax for the left tail is zero in the rational model, a perturbation of the tax rate formula (28) by a finite $m(z)$ will not alter the zero limiting marginal tax rate in the model with misspecification concerns.

4.4 Worst-case distributions and marginal tax rates

Figure 2 plots the worst-case productivity distributions $\widetilde{f}(z)$ across different choices of the parameter θ. The orange line with an infinite entropy penalty $\theta = \infty$ corresponds to the rational case for which $\widetilde{f}(z) = f(z)$, with subsequent lines representing increasing misspecification concerns as θ decreases.

The left panel shows the distribution in the proximity of the mean under the benchmark model, which is equal to $E[z] = 1$, while the right panel focuses on the broad range of productivities z and plots the density in logarithms to highlight distortions in the tail. In the log-log plot in the right panel, the straight orange line in the right tail reflects the Pareto shape of the right tail under the benchmark distribution $f(z)$.

As is apparent from the right graph, even modest levels of misspecification concerns lead to sharp distortions in the right tail of the distribution. This is in line with theoretical results from Section 3 that show that an arbitrarily small amount of misspecification concerns leads to a thin-tailed worst-case distribution. At the same time, the center of the distribution gets noticeably distorted only as misspecification concerns become substan-

quantiles \ θ	∞	100	20	10	1
$\widetilde{q}\,(0.01)$	0.291	0.290	0.289	0.287	0.252
$\widetilde{q}\,(0.05)$	0.394	0.393	0.391	0.388	0.252
$\widetilde{q}\,(0.25)$	0.629	0.628	0.622	0.614	0.500
$\widetilde{q}\,(0.50)$	0.919	0.916	0.902	0.885	0.675
$\widetilde{q}\,(0.75)$	1.490	1.478	1.434	1.384	0.933
$\widetilde{q}\,(0.95)$	4.348	4.241	3.826	3.438	1.597
$\widetilde{q}\,(0.99)$	12.568	11.949	9.601	7.765	2.460

Table 1: Quantiles of the distribution of productivity z under the worst-case distribution for alternative values of θ. The case $\theta = \infty$ corresponds to the rational case.

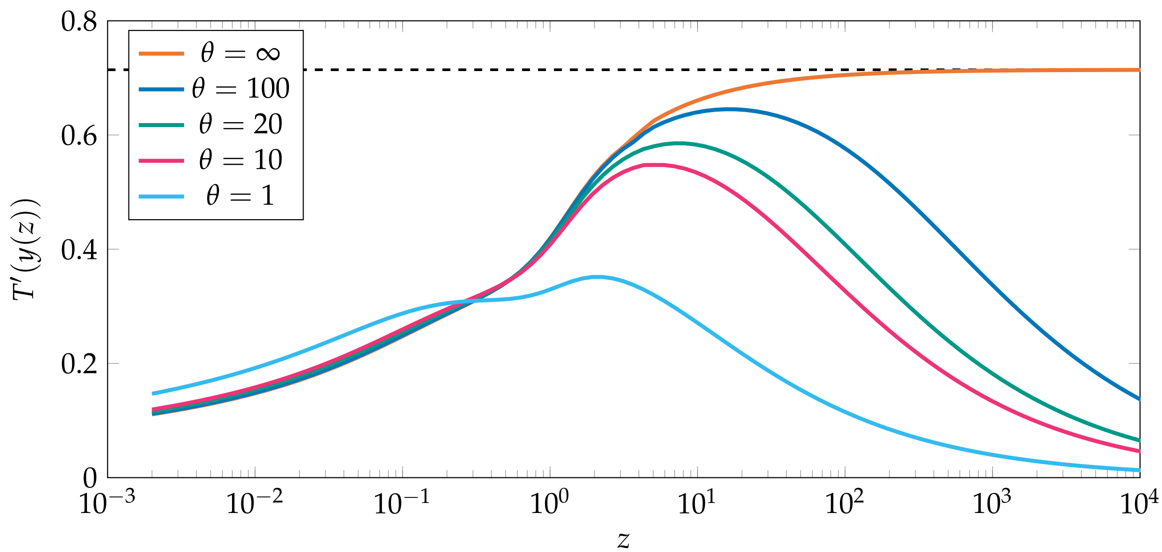

Figure 3: Optimal marginal tax schedules for alternative levels of misspecification concerns. The dashed line corresponds to the limiting marginal tax rate for the rational case.

tial ($\theta = 1$), and the left tail remains essentially undistorted. These conclusions are also confirmed in Table 1 that tabulates the quantiles of the productivity z under the alternative worst-case distributions $\widetilde{f}\,(z)$.

Figure 3 shows the optimal marginal tax rate schedule for alternative levels of the model misspecification concerns. The orange line represents the marginal tax rate for the rational case. In line with the literature, since the underlying productivity distribution exhibits a Pareto tail, the asymptotic tax rate $\lim_{z\to\infty} T'\,(y\,(z))$ is positive and quantitatively large, at 71.4%.[8] The tax rate asymptotes to zero as $z \to 0$, in line with Lemma 4.1.

When misspecification concerns are present, the shape of the optimal tax schedules looks notably different. While for wage levels around the mean ($E\,[z] = 1$), the optimal marginal tax looks similar to that under the rational case, it starts departing quickly for

[8]In Heathcote and Tsujiyama (2021), the computed marginal tax rates in the right tail asymptote to zero because they truncate the distribution and focus on numerical solutions for the truncated case.

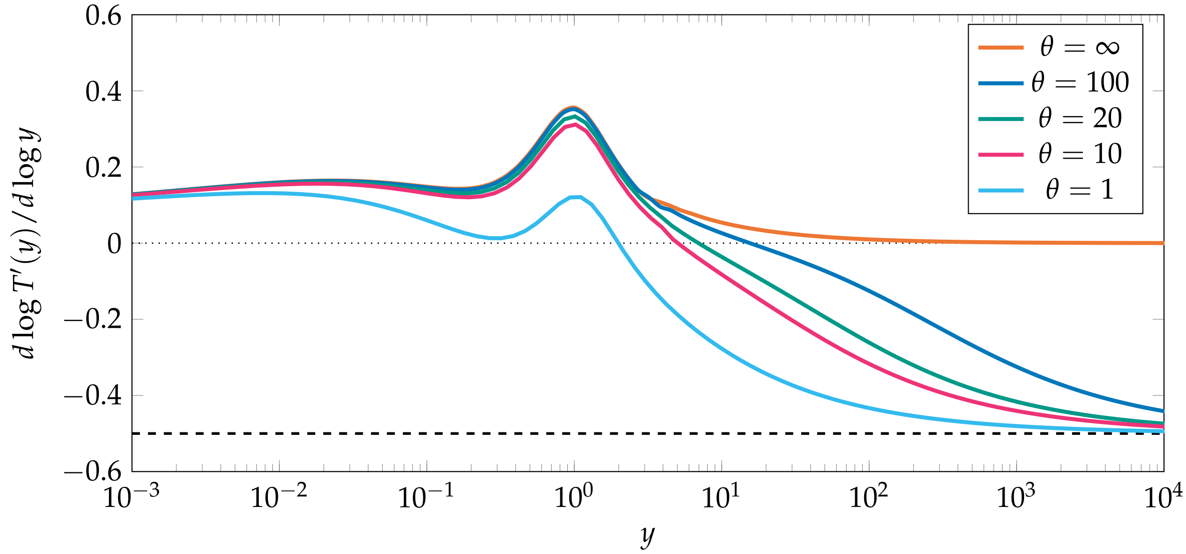

Figure 4: Growth rates of optimal marginal tax schedules, $d \log T'(y) / d \log y$, for alternative levels of misspecification concerns. The dashed black line corresponds to the theoretical limit, equal to $-1/2$.

higher wage levels. For the benchmark case of $\theta = 20$, the marginal tax peaks at 58.5% for wages equal to 7.5 the average wage, and starts declining thereafter.

Theoretical results from Section 3 show that not only should marginal tax rates decline to zero as $z \to \infty$, but the asymptotic rate of decline is also pinned down. From Lemma 4.1, given that $\rho = 1$, we have that the rate of decline should converge to $-\frac{1}{2}$. Figure 4 verifies this result numerically. The orange line, representing the rational case, is above zero and converges to zero, reflecting an increasing marginal tax rate schedule converging to a positive limiting tax rate. On the other hand, across all levels of misspecification concerns, the decay rate indeed asymptotically converges to the theoretically predicted value.

4.5 Insurance provision and budgetary concerns

The decision problem of the robust planner trades off utilitarian and budgetary concerns, reflected in the worst-case distortion

$$m(z) = \bar{m} \exp\left(-\frac{1}{\theta}\left[\mathcal{U}(z) + \mu T(y(z))\right]\right). \tag{34}$$

On the one hand, the planner is concerned that there are more agents in the left tail of the productivity distribution, who receive low utility $\mathcal{U}(z)$ and generate low (in fact negative) net tax revenue $T(y(z))$. The planner can diminish the utilitarian concerns by providing more insurance, thus raising $\mathcal{U}(z)$ and lowering $m(z)$. This insurance comes in the form

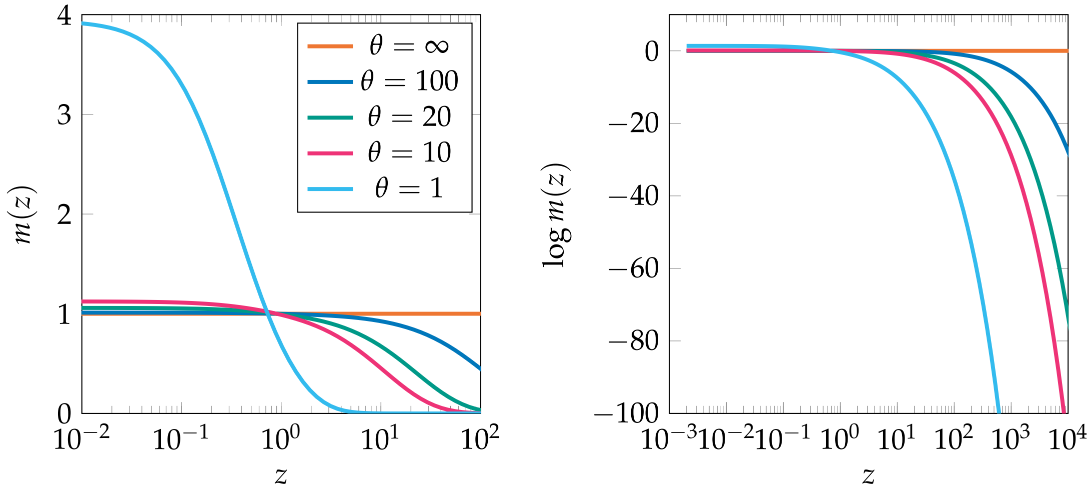

Figure 5: The likelihood ratio $m(z) = \widetilde{f}(z)/f(z)$ representing the distortion of the worst-case distribution relative to the benchmark distribution, plotted for alternative levels of misspecification concerns given by θ. The case $\theta = \infty$ corresponds to the rational benchmark for which $m(z) = 1$.

of transfers, and hence at the cost of a lower net tax revenue $T(y(z))$. On the margin, the optimal tax schedule designed by the planner trades off a unit of tax revenue at the marginal social value μ against a unit of consumption transferred to the low-productivity agent at marginal value $\mathcal{U}'(z)$.

Any transfers provided to low-productivity workers must come from those in upper parts of the productivity distribution. As shown in Section 3, when the utility function is sufficiently concave ($\rho > \frac{1}{2}$), budgetary concerns dominate the shape of the distortion (34) in the right tail of the productivity distribution. Following the tax formula (28), the planner chooses the marginal tax rate for a particular z as a result of a tradeoff between the tax distortion imposed on the $\widetilde{f}(z)$ agents at z against the social benefit of raising lump sum revenue from the (marginal utility weighted) mass of agents with productivities above z. Since the misspecification concern increases with z, the planner fears that the mass of agents above z is lower relative to agents residing exactly at z who are distorted by the marginal tax at z, and hence opts for lower marginal tax rates in the tail.

This desire to lower marginal taxes at the top combined with concerns about the higher prevalence of low-productivity and lower prevalence of high-productivity workers increases the marginal social value of a unit of tax revenue μ, pushing toward lower overall redistribution.

Figure 5 represents these implications quantitatively by plotting the shape of the distortion $m(z) = \widetilde{f}(z)/f(z)$. The plots reveal that in the left tail of the productivity distribution, the utilitarian and budgetary concerns reflected in the shape of $m(z)$ are minimal for the benchmark choice $\theta = 20$. Without any redistribution scheme, $\lim_{z\to 0} \mathcal{U}(z) = -\infty$,

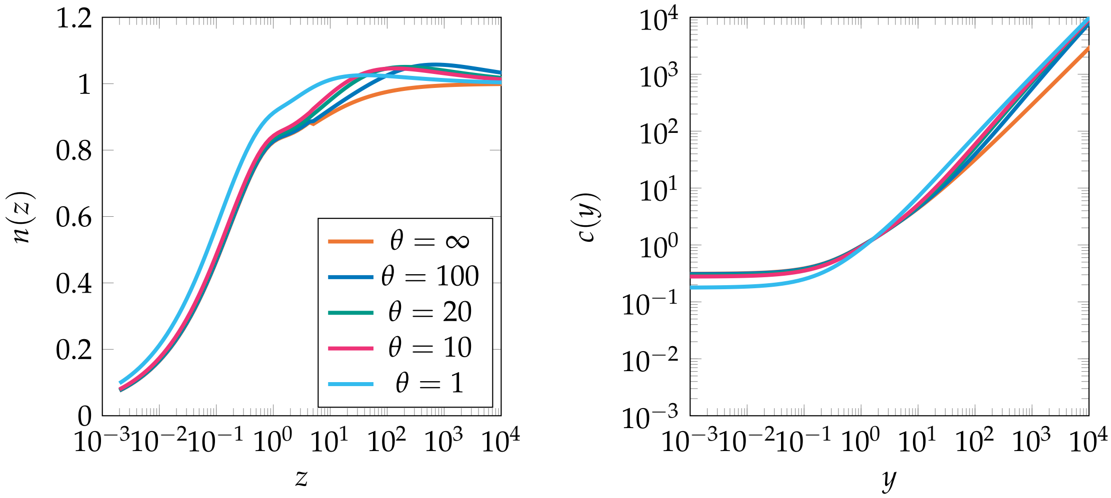

Figure 6: Labor supply under optimal tax schedules (left panel), and consumption as a function of income (right panel) under alternative levels of misspecification concerns.

and consequently $\lim_{z\to 0} m(z) = \infty$, as the utilitarian concerns about the low-productivity workers dominate. However, the optimal tax schedule insures the low-productivity workers sufficiently, leading to bounded and quantitatively modest distortions of the left tail. Only when the misspecification concerns are substantial ($\theta = 1$), the planner's concerns about insufficient tax revenue to insure low-productivity workers start increasing more notably.

On the other hand, in the right tail of the productivity distribution, the concerns about loss of tax revenue from high-productivity workers are severe. The planner, concerned about the distortion of the labor supply at the top, lowers marginal tax rates asymptotically to zero, yet the marginal tax rate declines sufficiently slowly to make the tax revenue $T(y(z))$ from each high-productivity worker grow without bound as $z \to \infty$. Consequently, the planner's concerns that these high-productivity workers are less prevalent than assumed under the benchmark distribution also grow without bound, leading to $\lim_{z\to\infty} m(z) = 0$.

The left panel in Figure 6 depicts the labor supply under the optimal tax schedule for alternative levels of the misspecification concerns. The lower marginal tax rates when misspecification concerns are present generally increase labor supply. At the same time, the right panel shows that low productivity households who produce little output are well insured by the optimal tax scheme.

Table 2 summarizes these insights in the form of moments under the benchmark and worst-case distributions. The benchmark mean $E[z]$ is identical and normalized to one across parameterizations since the distribution $f(z)$ is exogenously specified by the calibrated EGM distribution in Section 4.2. The worst case means $\widetilde{E}[z]$ decrease as misspec-

moments \ θ	∞	100	20	10	1
$E[z]$	1.000	1.000	1.000	1.000	1.000
$\widetilde{E}[z]$	1.000	0.986	0.951	0.918	0.657
$E[y]$	0.824	0.828	0.838	0.848	0.918
$\widetilde{E}[y]$	0.824	0.815	0.792	0.770	0.579
μ	1.220	1.229	1.261	1.297	1.726
T_0	-0.311	-0.307	-0.292	-0.278	-0.178
$\max_y T'(y)$ (%)	71.4	64.5	58.6	55.0	35.2
$\arg\max_y T'(y)$	∞	15.541	6.955	4.609	1.992
$E[T]$	0.000	0.008	0.023	0.037	0.110
$\widetilde{E}[T]$	0.000	0.000	0.000	0.000	0.000

Table 2: Moments under the benchmark distribution $f(z)$ and the worst-case distribution $\widetilde{f}(z)$ under optimal tax schedules for alternative values of θ.

ification concerns increase, reflecting shifts toward more adversely slanted productivity distributions subjectively perceived by the robust planner.

Objective means of the income distribution $E[y(z)] = E[zn(z)]$ increase with increasing misspecification concerns, as lower marginal taxes increase labor supply, plotted in the left panel of Figure 6. Nevertheless, the mean of the income distribution under the worst case distribution, $\widetilde{E}[y]$, declines.

As mentioned above, since increases in misspecification concerns are manifested in increases in the marginal social value of tax funds μ, reflecting planner's fear about more severe scarcity of tax revenue.

Finally, the bottom part of Table 2 reports statistics for the tax schedule. The lump sum transfer to the lowest productivity worker, $T_0 = T(y(\underline{z}))$, expressed as a share of average income $E[y]$, is roughly 33–37%, and rather stable across the a range of the values of the parameter θ. Only when misspecifications concerns and the fear of lack of tax revenue worsen severely, the transfer is reduced more substantially.

While the lump sum transfer to the lowest productivity worker is rather insensitive to the degree of misspecification concerns for a range of values of θ, the peak marginal tax rate changes substantially. For the case without misspecification concerns, the top marginal tax rate asymptotes at 71.4%, for the benchmark calibration $\theta = 20$, the marginal tax rate peaks at 58.6% at an income $\arg\max_y T'(y)$ corresponding to 8.3 of the average income, and starts declining thereafter.

Since the planner's optimization problem involves running a balanced budget under the endogenously determined subjective distribution, and we chose $B = 0$, this implies $\widetilde{E}[T] = 0$ for all choices of θ. However, the tax revenue under the benchmark distribution generally differs from zero. Since the worst-case distribution is pessimistically biased, then, compared to the benchmark distribution, the planner underestimates the amount

I	100	500	1000	3500	10000
$\theta = 100$	0.462	0.452	0.437	0.377	0.304
$\theta = 20$	0.432	0.298	0.247	0.090	0.010
$\theta = 10$	0.348	0.172	0.092	0.007	0.000
$\theta = 1$	0.001	0.000	0.000	0.000	0.000

Table 3: Detection error probabilities for models with alternative choices of misspecification concerns parameterized by θ, and for alternative sizes of the random samples I.

of tax revenue the given tax schedule raises under that benchmark distribution. For the preferred choice $\theta = 20$, the extra surplus generated by the tax policy is about 2.7% of total labor income in the economy.

If we interpret the benchmark distribution as the true productivity distribution under the data-generating measure, then the optimal tax policy of the robust planner indeed generates this surplus. This raises the question of what the planner does with the surplus resources. Our model is static so it does not speak to intertemporal tradeoffs but it is plausible to envision a dynamic extension of this model in which the planner also manages the accumulated debt or assets over time. Hansen and Sargent (2012, 2015), Kwon and Miao (2017), Ferriere and Karantounias (2019), or Karantounias (2023) are important contributions in this direction that continue optimal dynamic policies in representative agent frameworks in which the planner is ambiguous about the stochastic path of the aggregate economy. Introducing dynamic debt management into our framework faces novel challenges but is a natural way of moving forward.

4.6 Detection error probabilities

We calibrated the baseline magniftude of the misspecification concerns to achieve a detection error probability $\bar{d} = 0.09$ when the benchmark and worst-case distributions are compared using random draws of samples with size $I = 3500$, corresponding to the size of the SCF sample. In Table 3 we compare the detection error probabilities across alternative values of θ and I. In line with the theoretical construction outlined in Section 4.1, the detection error probabilities decline in θ and I.

For $\theta = 100$, the misspecification concerns are negligible and the two distributions $\widetilde{f}(z)$ and $f(z)$ statistically close to each other, so that the detection error probabilities are close to their upper bound of 0.5, at which the two distributions are indistinguishable. On the other hand, for $\theta = 1$, the worst-case distribution is extremely distinct from the benchmark, and even a small random sample can convincingly distinguish the two distributions.

5 Type-dependent model misspecification concerns

In the baseline model, we used a constant parameter θ to express the magnitude of the uncertainty about the underlying productivity distribution. However, surveys and other evidence can provide differential amount of information about the distribution in different parts of its support. For example, the government may be less confident about both the left and right tails of the distribution of the productivity distribution, compared to its center, due to lower response rates in the survey, undersampling, or issues with measurement.

To incorporate this idea, we extend the baseline model with a type dependent penalty parameter $\theta(z)$. We postulate a penalty function of the form

$$\mathcal{E}(m,\theta) = \mathbb{E}\left[\theta\left(m\log m - m + 1\right)\right] = \int_{\underline{z}}^{\bar{z}} \theta(z)\left(m(z)\log m(z) - m(z) + 1\right) f(z)\,dz, \quad (35)$$

where $\theta(z)$ is a strictly positive function. The minimax problem (2) augmented with the type-dependent penalty parameter $\theta(z)$ now reads

$$\min_{\substack{m \\ \mathbb{E}[m]=1}} \max_{T} \mathbb{E}\left[m\psi\mathcal{U}\right] + \mathbb{E}\left[\theta\left(m\log m - m + 1\right)\right] \qquad \text{s.t. } \mathbb{E}\left[mT(\mathcal{Y})\right] \geq B. \quad (36)$$

Key properties and implications of the state dependent penalty are summarized in the following lemma.

Lemma 5.1. *The entropy penalty $\mathcal{E}(m,\theta)$ in (35) satisfies the following properties.*

1. *When $\theta(z) = \overline{\theta}$ is constant, $\mathcal{E}(m,\theta)$ reduces to the entropy penalty used in the baseline model*
$$\mathcal{E}(m,\theta) = \overline{\theta}\mathbb{E}\left[m\log m\right].$$
2. *For any given strictly positive function $\theta(z)$, the entropy penalty satisfies $\mathcal{E}(m,\theta) \geq 0$, and is minimized for $m = 1$.*
3. *Denote μ the Lagrange multiplier on the budget constraint $\mathbb{E}\left[mT(\mathcal{Y})\right] \geq B$, and $v(z) = \psi(z)\mathcal{U}(z) + \mu T(y(z))$. Then, for a z and z' for which $\theta(z) = \theta(z')$, if $v(z) > v(z')$ then $m(z) < m(z')$ is decreasing in $v(z)$ is constant across all worker types, When $\psi(z)\mathcal{U}(z)$ is constant across worker types,*
4. *If $v(z)$ is constant across all worker types, then there the minimizing probability measure is not distorted, $m(z) = 1$.*

Proof of Lemma 5.1. When $\theta(z) = \overline{\theta}$, then definition (35) implies

$$\mathcal{E}\left(m,\overline{\theta}\right) = \overline{\theta}\mathbb{E}\left[m\log m\right] + \overline{\theta}\left(1 - \mathbb{E}\left[m\right]\right) = \overline{\theta}\mathbb{E}\left[m\log m\right]$$

since $\mathbb{E}[m] = 1$, showing property 1. To show 2., notice that

$$m(z) \log m(z) \geq m(z) - 1$$

with equality if and only if $m(z) = 1$. Premultiplying by $\theta(z)$, and integrating over z yields

$$\int_{\underline{z}}^{\bar{z}} \theta(z) \left(m(z) \ln m(z) - m(z) + 1\right) f(z)\, dz = \mathcal{E}(m, \theta) \geq 0,$$

and equality is attained if and only if $m(z) = 1$ almost everywhere on the support of the distribution of z.

Taking the first-order condition with respect to $m(z)$ in (36) yields

$$\psi(z) \mathcal{U}(z) + \mu T(y(z)) - \chi + \theta(z) \log m = 0,$$

where χ is the Lagrange multiplier on the constraint $\mathbb{E}[m] = 1$, and μ is the multiplier on the budget constraint. This can be reorganized as

$$m(z) = \exp\left(-\frac{1}{\theta(z)} \left(\psi(z) \mathcal{U}(z) + \mu T(y(z)) - \chi\right)\right). \tag{37}$$

The multiplier χ can be found from the restriction $\mathbb{E}[m] = 1$, by solving the implicit equation

$$\int_{\underline{z}}^{\bar{z}} \exp\left(-\frac{1}{\theta(z)} \left(\psi(z) \mathcal{U}(z) + \mu T(y(z)) - \chi\right)\right) f(z)\, dz = 1.$$

Equation (37) directly implies property 3. Finally, when $v(z) = \bar{v}$ is constant, then the unique solution to (37) is equal to $\chi = \bar{v}$, and $m(z) = 1$, confirming property 4. ■

The function $\theta(z)$ can now be calibrated to capture heterogenous precision of information about the shape of the productivity distribution in different parts of its support, assigning a smaller penalty to the parts of the support where misspecification concerns are stronger.

[TO BE COMPLETED]

6 Multidimensional uncertainty

[TO BE COMPLETED]

7 Conclusion

The design of optimal tax schedules crucially depends on assumptions about the underlying distribution of types of taxed individuals. These types are typically unobservable, and measurement of observable quantities like earnings hinges upon surveys with often small samples, especially in the tails that matter most.

In this paper, we tackle this uncertainty about the underlying type distribution by studying optimal taxation in the presence of concerns that the underlying distribution is misspecified. The robust approach we employ allows us to avoid making parametric assumptions about the structure of uncertainty, and instead focus on nonparametric misspecifications of the underlying type distribution that the planner fears to be most consequential for the welfare function.

In stark contrast to the existing literature, we find that marginal taxes should optimally decline to zero, even when the underlying benchmark distribution has fat tails, like in the Pareto case. Misspecification concerns do not only have a theoretical asymptotic impact, they decrease marginal tax rates of high income individuals in quantitatively important ways, against a modest decrease in redistribution toward low-productivity workers.

We calibrate the plausible degree of misspecification concerns using detection error probabilities that capture the ability to distinguish alternative distributions using sampled data, with sample sizes corresponding to those in available surveys. While doing so, we also provide a theoretical extension of entropy penalties widely used in the literature to allow for type-dependent penalization that captures differential amount of available information about the shape of the underlying distribution in different parts of the support of the type distribution.

Our main theoretical results and the benchmark quantitative application involve a static model of labor income taxation in the presence of a one-dimensional unobserved type. This sharply theoretically and quantitatively analyzed problem opens up avenues toward applications in which uncertainty about the underlying type distribution can be even more pressing. One such example that we provide involves multidimensional unobserved distribution of worker types in which the planner lacks information on the joint distribution of productivity and labor supply elasticity.

Other important directions are left for follow-up work. A natural candidate involves a dynamic extension where optimal taxation of heterogeneous workers is combined with an optimal debt management problem. Another prominent application in which uncertainty about the underlying distribution is substantial involves taxation of wealth. Insights developed in this work will help analyze such cases as well.

Appendix

A Derivations of optimal tax formulas

A.1 Derivation of the tax formula

$$\max_{c,y}\min_{m}\int_{\underline{z}}^{\bar{z}}\psi(z)\,U\left(c(z),\frac{y(z)}{z}\right)m(z)\,f(z)\,dz+\theta\int_{\underline{z}}^{\bar{z}}m(z)\log m(z)\,f(z)\,dz$$

subject to the IC constraint

$$\frac{dU}{dz}=-U_n\left(c(z),\frac{y(z)}{z}\right)\frac{y(z)}{z^2}.$$

and the budget constraint

$$\int_{\underline{z}}^{\bar{z}}\left(y(z)-c(z)\right)m(z)\,f(z)\,dz\geq B.$$

Treating $\mathcal{U}$ as the state variable, λ as its co-state, and y and m as control variables, we can form the constrained Hamiltonian

$$\begin{aligned}H(\mathcal{U},y,m,\lambda)\;&=\;\psi(z)\,\mathcal{U}(z)\,m(z)\,f(z)+\theta m(z)\log m(z)\,f(z)-\kappa m(z)\,f(z)\\&\quad-\lambda(z)\,U_n\left(c(z),\frac{y(z)}{z}\right)\frac{y(z)}{z^2}+\mu\left[y(z)-c(z)\right]m(z)\,f(z).\end{aligned}$$

Here, κ and μ are multipliers on the constraints $\mathbb{E}[m]=1$ and (7), respectively, and $c(z)$ is defined implicitly from the definition of the utility function

$$\mathcal{U}(z)=U\left(c(z),\frac{y(z)}{z}\right)$$

as $c(z)=C(\mathcal{U}(z),y(z))$.

The optimality condition with respect to output choice is

$$\begin{aligned}0\;=\;&H_y=\mu\left[1-C_y(U(z),y(z))\right]m(z)\,f(z)-\lambda(z)\,U_n\left(C(\mathcal{U}(z),y(z)),\frac{y(z)}{z}\right)\frac{1}{z^2}\\&-\lambda(z)\left[U_{nc}\left(C(\mathcal{U}(z),y(z)),\frac{y(z)}{z}\right)C_y(\mathcal{U}(z),y(z))+U_{nn}\left(C(\mathcal{U}(z),y(z)),\frac{y(z)}{z}\right)\frac{1}{z}\right]\frac{y(z)}{z^2}\end{aligned}\tag{38}$$

with respect to the distortion is

$$0=H_m=\psi(z)\,\mathcal{U}(z)\,f(z)+\theta\left[\log m(z)+1\right]f(z)-\kappa f(z)+\mu\left[y(z)-C(\mathcal{U}(z),y(z))\right]f(z)$$

and the costate dynamics restriction yields

$$\begin{aligned}\frac{d\lambda(z)}{dz}\;&=\;-H_U=\left[\mu C_U(\mathcal{U}(z),y(z))-\psi(z)\right]m(z)\,f(z)+\\&\quad+\lambda(z)\,U_{nc}\left(C(\mathcal{U}(z),y(z)),\frac{y(z)}{z}\right)\frac{y(z)}{z^2}C_U(\mathcal{U}(z),y(z)).\end{aligned}\tag{39}$$

The transversality condition is

$$\lambda\left(\bar{z}\right)\mathcal{U}\left(\bar{z}\right)=0$$

and since the problem is unrestricted at the left end of the type distribution (the choice $U\left(\underline{z}\right)$ is unrestricted), we also have $\lambda\left(\underline{z}\right)=0$.

The condition (39) can be integrated up over the whole range of types $\left(\underline{z},\bar{z}\right)$:

$$\begin{aligned}\int_{\underline{z}}^{\bar{z}}\frac{d\lambda\left(z\right)}{dz}dz &= \lambda\left(\bar{z}\right)-\lambda\left(\underline{z}\right)=0=\mu\int_{\underline{z}}^{\bar{z}}C_U\left(\mathcal{U}\left(z\right),y\left(z\right)\right)m\left(z\right)f\left(z\right)dz-\int_{\underline{z}}^{\bar{z}}\psi\left(z\right)m\left(z\right)f\left(z\right)dz\\ &\quad+\int_{\underline{z}}^{\bar{z}}\lambda\left(z\right)U_{nc}\left(C\left(\mathcal{U}\left(z\right),y\left(z\right)\right),\frac{y\left(z\right)}{z}\right)\frac{y\left(z\right)}{z^2}C_U\left(\mathcal{U}\left(z\right),y\left(z\right)\right)dz\end{aligned}$$

which can be solved for the Lagrange multiplier μ

$$\mu=\frac{\int_{\underline{z}}^{\bar{z}}\psi\left(z\right)m\left(z\right)f\left(z\right)dz-\int_{\underline{z}}^{\bar{z}}\lambda\left(z\right)U_{nc}\left(C\left(\mathcal{U}\left(z\right),y\left(z\right)\right),\frac{y(z)}{z}\right)\frac{y(z)}{z^2}C_U\left(\mathcal{U}\left(z\right),y\left(z\right)\right)dz}{\int_{\underline{z}}^{\bar{z}}C_U\left(\mathcal{U}\left(z\right),y\left(z\right)\right)m\left(z\right)f\left(z\right)dz} \tag{40}$$

This Lagrange multiplier can be interpreted as the marginal social value of public funds to the planner.

With an expression for μ at hand, we can solve for $\lambda\left(z\right)$ forward by integrating (39) on $\left(z,\bar{z}\right)$. For that purpose, denote the terms in (39) as follows:

$$\begin{aligned}H_{U,S}\left(z\right) &= \mu C_U\left(\mathcal{U}\left(z\right),y\left(z\right)\right)f\left(z\right)-\psi\left(z\right)m\left(z\right)f\left(z\right)\\ H_{U,N}\left(z\right) &= U_{nc}\left(C\left(\mathcal{U}\left(z\right),y\left(z\right)\right),\frac{y\left(z\right)}{z}\right)\frac{y\left(z\right)}{z^2}C_U\left(\mathcal{U}\left(z\right),y\left(z\right)\right).\end{aligned}$$

The term $H_{U,N}\left(z\right)$ is only present when preferences $U\left(c,n\right)$ are non-separable. We can then rewrite (39) as

$$\lambda'\left(z\right)=H_{U,S}\left(z\right)+\lambda\left(z\right)H_{U,N}\left(z\right).$$

This equation has the solution

$$\lambda\left(z\right)=-\int_{z}^{\bar{z}}H_{U,S}\left(\zeta\right)\exp\left(-\int_{z}^{\zeta}H_{U,N}\left(\xi\right)d\xi\right)d\zeta.$$

To simplify notation, let us simplify the arguments of functions above, and write, for example,

$$U_{nc}\left(z\right)=U_{nc}\left(C\left(\mathcal{U}\left(z\right),y\left(z\right)\right),\frac{y\left(z\right)}{z}\right).$$

Further, we utilize the following notational simplifications:

$$\begin{aligned}
\frac{1+\varepsilon^{u}}{\varepsilon^{c}} &= 1+n\frac{U_{nn}+wU_{cn}}{U_{n}} \\
\left(1-T'\left(y\left(z\right)\right)\right)z &= -\frac{U_{n}\left(z\right)}{U_{c}\left(z\right)} \\
C_{U}\left(\mathcal{U}\left(z\right),y\left(z\right)\right) &= \frac{1}{U_{c}\left(z\right)} \\
C_{y}\left(\mathcal{U}\left(z\right),y\left(z\right)\right) &= -\frac{1}{z}\frac{U_{n}\left(z\right)}{U_{c}\left(z\right)}
\end{aligned}$$

The first equation follows from the definition of uncompensated and compensated labor supply elasticities, ε^{u} and ε^{c}, respectively. The second equation is the worker's individual optimality condition for the choice of labor supply and consumption given a particular tax schedule. The last two follow from the definition of the implicit function $C\left(\mathcal{U}\left(z\right),y\left(z\right)\right)$.

With this notation and substitutions, we can reorganize the optimality condition (38) as

$$-\mu\left[T'\left(y\left(z\right)\right)\right]f\left(z\right)=\frac{\lambda\left(z\right)}{z}U_{c}\left(z\right)\left(1-T'\left(y\left(z\right)\right)\right)\left[1+n\left(z\right)\frac{U_{nn}\left(z\right)+w\left(z\right)U_{nc}\left(z\right)}{U_{n}\left(z\right)}\right]$$

which then yields

$$\frac{T'\left(y\left(z\right)\right)}{1-T'\left(y\left(z\right)\right)}=-\frac{1+\varepsilon^{u}\left(z\right)}{\varepsilon^{c}\left(z\right)}\frac{\lambda\left(z\right)}{zm\left(z\right)f\left(z\right)}\frac{U_{c}\left(z\right)}{\mu}. \tag{41}$$

To obtain the Lagrange multiplier μ, simplify equation (40) to obtain

$$\mu=\frac{\int_{\underline{z}}^{\bar{z}}\left[\psi\left(\zeta\right)-\frac{\lambda\left(\zeta\right)}{zm\left(\zeta\right)f\left(\zeta\right)}\frac{n\left(\zeta\right)U_{nc}\left(\zeta\right)}{U_{c}\left(\zeta\right)}\right]m\left(\zeta\right)f\left(\zeta\right)d\zeta}{\int_{\underline{z}}^{\bar{z}}\frac{1}{U_{c}\left(\zeta\right)}m\left(\zeta\right)f\left(\zeta\right)d\zeta}$$

Finally, the expression for $\lambda\left(z\right)$ can be simplified as

$$\begin{aligned}
H_{U,S}\left(z\right) &= \left[\frac{\mu}{U_{c}\left(z\right)}-\psi\left(z\right)\right]m\left(z\right)f\left(z\right) \\
H_{U,N}\left(z\right) &= \frac{n\left(z\right)U_{nc}\left(z\right)}{zU_{c}\left(z\right)} \\
\lambda\left(z\right) &= -\int_{z}^{\bar{z}}H_{U,S}\left(\zeta\right)\exp\left(-\int_{z}^{\zeta}H_{U,N}\left(\xi\right)d\xi\right)d\zeta.
\end{aligned}$$

For the special case when the utility function is quasilinear of the form (9), we obtain $U_{c}\left(z\right)=1$, $H_{U,N}\left(z\right)=0$, and

$$\mu=\int_{\underline{z}}^{\bar{z}}\psi\left(\zeta\right)m\left(z\right)f\left(\zeta\right)d\zeta,$$

yielding

$$\lambda\left(z\right)=-\int_{z}^{\bar{z}}\left[\mu-\psi\left(\zeta\right)\right]m\left(\zeta\right)f\left(\zeta\right)d\zeta.$$

The tax formula (41) then yields (11) when we notice that $\left(1+\varepsilon_{u}\right)/\varepsilon_{c}=1+\gamma$.

B Proofs for the baseline model

Throughout this section, we restrict our attention to the analysis of the baseline model, analyzed in Sections 3.1–3.3. In particular, we assume the quasilinear utility function

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

an unbounded type space, $\bar{z} = \infty$, and restrict our attention to the characterization of the optimal marginal tax on the interval $[\hat{z}, \bar{z})$ on which the planner has no utilitarian concerns, $\Psi(\hat{z}) = \widetilde{\Psi}(\hat{z}) = 1$, and for which Assumption 1 holds. In the proof of Theorem 3.2, we additionally assume that the shape of the type distribution for $z \geq \hat{z}$ under the benchmark model is proportional to the Pareto distribution.

B.1 Proof of Theorem 3.1

Before proving Theorem 3.1, we start with three preliminary lemmas.

Lemma B.1. *Optimal output chosen by individual workers under a given tax scheme $T(y)$ satisfies:*

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

Proof. Under the given tax scheme $T(y)$, the optimality condition (1) of a worker of type z for the special case of quasilinear preferences (9) implies labor supply $n(z)$ given by

$$\left(1 - T'(zn(z))\right) z = n(z)^{\gamma}$$

which can be rewritten as

$$y(z) = \left(1 - T'(y(z))\right)^{\frac{1}{\gamma}} z^{\frac{1+\gamma}{\gamma}}. \tag{42}$$

Differentiating with respect to z:

$$\begin{aligned} y'(z) &= -\frac{1}{\gamma}\left(1 - T'(y(z))\right)^{\frac{1-\gamma}{\gamma}} z^{\frac{1+\gamma}{\gamma}} T''(y(z))\, y'(z) + \frac{1+\gamma}{\gamma}\left(1 - T'(y(z))\right)^{\frac{1}{\gamma}} z^{\frac{1}{\gamma}} \\ &= -\frac{1}{\gamma}\frac{T''(y(z))}{1 - T'(y(z))} y(z)\, y'(z) + \frac{1+\gamma}{\gamma}\frac{y(z)}{z} \end{aligned}$$

yields

$$\left[1 + \frac{1}{\gamma}\frac{T''(y(z))}{1 - T'(y(z))} y(z)\right] y'(z) = \frac{1+\gamma}{\gamma}\frac{y(z)}{z}$$

and hence

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

■

Lemma B.2. *Let $f(z)$ and $\widetilde{f}_i(z)$, $i = 1, 2$ be density functions related by $\widetilde{f}_i(z) = m_i(z) f(z)$ where $m_i(z)$ are strictly positive functions representing changes of measure, and $m_1(z)/m_2(z)$ is strictly decreasing. Then*

$$\frac{1-\widetilde{F}_1(z)}{z\widetilde{f}_1(z)} < \frac{1-\widetilde{F}_2(z)}{z\widetilde{f}_2(z)}.$$

Proof. The expression can be rewritten as

$$\begin{aligned}
\frac{1-\widetilde{F}_1(z)}{z\widetilde{f}_1(z)} &= \frac{\int_z^{\overline{z}} m_1(\zeta) f(\zeta)\, d\zeta}{z m_1(z) f(z)} = \frac{\int_z^{\overline{z}} \frac{m_1(\zeta)}{m_1(z)} f(\zeta)\, d\zeta}{z f(z)} && (43)\\
&< \frac{\int_z^{\overline{z}} \frac{m_2(\zeta)}{m_2(z)} f(\zeta)\, d\zeta}{z f(z)} = \frac{\int_z^{\overline{z}} m_2(\zeta) f(\zeta)\, d\zeta}{z m_2(z) f(z)} = \frac{1-\widetilde{F}_2(z)}{z\widetilde{f}_2(z)}, && (44)
\end{aligned}$$

where the inequality comes from the fact that for any $\zeta > z$,

$$\frac{m_1(\zeta)}{m_2(\zeta)} < \frac{m_1(z)}{m_2(z)}.$$

■

Lemma B.3. *Define*

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

Then

$$\widetilde{\phi}'(z) = \frac{d}{dz}\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} = -\frac{1}{z} - \widetilde{\phi}(z)\left[\frac{1}{z} - \frac{\mu}{\theta}T'(y(z))\, y'(z) + \frac{f'(z)}{f(z)}\right].$$

Proof. By direct computation:

$$\begin{aligned}
\frac{d}{dz}\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} &= \frac{-\widetilde{f}(z)\, z\widetilde{f}(z) - \left(1-\widetilde{F}(z)\right)\left(\widetilde{f}(z) + z\widetilde{f}'(z)\right)}{\left(z\widetilde{f}(z)\right)^2}\\
&= -\frac{1}{z} - \frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)}\left[\frac{1}{z} + \frac{\widetilde{f}'(z)}{\widetilde{f}(z)}\right] = -\frac{1}{z} - \widetilde{\phi}(z)\left[\frac{1}{z} + \frac{d}{dz}\log\widetilde{f}(z)\right].
\end{aligned}$$

Using $\widetilde{f}(z)$ from expression (14), we obtain the last line of the lemma. ■

Proof of Theorem 3.1. We restrict attention to $z \geq \hat{z}$ for which the planner's welfare weight is zero, $\psi(z) = 0$, and for which Assumption 1 holds. In this case, the tax formula (12) is given by

$$\frac{T'(y(z))}{1-T'(y(z))} = (1+\gamma)\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} = (1+\gamma)\,\widetilde{\phi}(z). \tag{45}$$

The single-crossing property (5) implies that $y(z)$ is strictly increasing in z. Since the tax rate $T'(y(z))$ is strictly positive, the tax function $T(y(z))$ strictly increases in z. This means that $m(z)$ in (14) is strictly decreasing, and, by Assumption 1, $\widetilde{f}(z)$ is also strictly decreasing for sufficiently large z. Lemma B.2 then implies that

$$\frac{1-\widetilde{F}(z)}{z\widetilde{f}(z)} < \frac{1-F(z)}{zf(z)}, \tag{46}$$

and hence for all $z > \hat{z}$, the marginal tax rate $T'(y(z))$ must be strictly lower than the marginal tax rate $T'_{rat}(y_{rat}(z))$ in the model without model misspecification.

An application of L'Hôpital's rule to the tax rate under the model without model misspecification implies

$$\lim_{z\to\infty} \frac{T'_{rat}(y(z))}{1-T'_{rat}(y(z))} = (1+\gamma)\lim_{z\to\infty}\frac{1-F(z)}{zf(z)} = \lim_{z\to\infty}\frac{1}{-\frac{d\log f(z)}{d\log z}-1} < \infty$$

by Assumption 1. This yields $\lim_{z\to\infty} T'_{rat}(y(z)) < 1$, and combined with $T'(y(z)) < T'_{rat}(y_{rat}(z))$ implied by (46), we obtain that for sufficiently large z, $T'(y(z))$ must be bounded away from one, $T'(y(z)) < 1-\bar{\varepsilon}_T$, and $\widetilde{\phi}(z)$ is bounded,

$$\widetilde{\phi}(z) = \frac{1}{1+\gamma}\frac{T'(y(z))}{1-T'(y(z))} < \frac{1}{1+\gamma}\frac{1-\bar{\varepsilon}_T}{\bar{\varepsilon}_T} = K_\phi.$$

Consequently, the optimal allocation formula (42) implies that $\lim_{z\to\infty} y(z) = \infty$.

From the optimal tax formula (45), we obtain $1-T'(y(z)) = \left(1+(1+\gamma)\widetilde{\phi}(z)\right)^{-1}$, and then we can rewrite expression (42) as

$$y(z) = \left(1+(1+\gamma)\widetilde{\phi}(z)\right)^{-\frac{1}{\gamma}} z^{\frac{1+\gamma}{\gamma}}.$$

Differentiating this expression with respect to z yields

$$\begin{aligned} y'(z) &= -\frac{1}{\gamma}\frac{(1+\gamma)\widetilde{\phi}'(z)}{\left(1+(1+\gamma)\widetilde{\phi}(z)\right)^{\frac{1+\gamma}{\gamma}}} z^{\frac{1+\gamma}{\gamma}} + \frac{1+\gamma}{\gamma}\frac{1}{\left(1+(1+\gamma)\widetilde{\phi}(z)\right)^{\frac{1}{\gamma}}} z^{\frac{1}{\gamma}} \\ &= \frac{1+\gamma}{\gamma}\frac{1}{\left(1+(1+\gamma)\widetilde{\phi}(z)\right)^{\frac{1+\gamma}{\gamma}}} z^{\frac{1}{\gamma}}\left[2+\widetilde{\phi}(z)\left(2+\gamma-\frac{\mu}{\theta}zT'(y(z))y'(z)+z\frac{f'(z)}{f(z)}\right)\right], \end{aligned}$$

where the second line uses Lemma B.3 to substitute in for $\widetilde{\phi}'(z)$. Since $y'(z) > 0$, the last bracket must be strictly positive.

We now show that $\lim_{z\to\infty} T'(y(z)) = 0$. Assume that it is not. Then there exists $\underline{\varepsilon}_T > 0$ such that for an arbitrarily large $\bar{z}$ there exists a $z > \bar{z}$ such that $T'(y(z)) > \underline{\varepsilon}_T$. Take such a z for which $T'(y(z)) > \underline{\varepsilon}_T$. Since $f'(z) < 0$, we have

$$0 < 2+\widetilde{\phi}(z)\left(2+\gamma-\frac{\mu}{\theta}zT'(y(z))y'(z)+z\frac{f'(z)}{f(z)}\right) < 2+K_\phi\left(2+\gamma-\frac{\mu}{\theta}zy'(z)\underline{\varepsilon}_T\right),$$

which yields a bound on $zy'(z)$:

$$zy'(z) < \frac{\theta}{\underline{\varepsilon}_T \mu}\left(2 + \gamma + 2K_{\phi}^{-1}\right) = K_y. \tag{47}$$

Since, from the result in Lemma B.1, we have that

$$zy'(z) = \frac{(1+\gamma)\, y(z)}{\gamma + \frac{T''(y(z))}{1-T'(y(z))} y(z)} < K_y,$$

we can derive a restriction on $T''(y(z))$:

$$T''(y(z)) > \left(1 - T'(y(z))\right) \frac{(1+\gamma)\, y(z) - K_y \gamma}{K_y y(z)} > \bar{\varepsilon}_T \frac{(1+\gamma)\, y(z) - K_y \gamma}{K_y y(z)}. \tag{48}$$

Recall that we can find an arbitrarily large z for which this inequality holds. Since $\lim_{z\to\infty} y(z) = \infty$, we can find such a z that is sufficiently large to satisfy $(1+\gamma)\, y(z) > K\gamma$, denote it $\check{z}$. Then $T''(y(\check{z})) > 0$, and, consequently, $T'(y(z))$ is increasing at $\check{z}$. Hence the inequality $T'(y(z)) > \underline{\varepsilon}_T$ holds also for z in the right neighborhood of $\check{z}$, so that the lower bound on $T''(y(z))$ given in (48) also hold for z to the right of $\check{z}$, and this argument can then be extended for any $z \geq \check{z}$. This then implies

$$\begin{aligned}
\lim_{z\to\infty} T'(y(z)) &= T'(y(\check{z})) + \int_{y(\check{z})}^{\infty} T''(\xi)\, d\xi > T'(y(\check{z})) + \int_{y(\check{z})}^{\infty} \bar{\varepsilon}_T \frac{(1+\gamma)\,\xi - K_y\gamma}{K_y \xi} d\xi \\
&> T'(y(\check{z})) + \bar{\varepsilon}_T \frac{(1+\gamma)\, y(\check{z}) - K_y\gamma}{K_y y(\check{z})} \int_{y(\check{z})}^{\infty} d\xi = \infty,
\end{aligned}$$

which contradicts the bound $T'(y(z)) < 1 - \bar{\varepsilon}_T$. Therefore, the marginal tax must converge to zero,

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

Is is worth noting that misspecification concerns enter the proof by way of a finite bound on $zy'(z)$ in (47). In the absence of misspecification concerns, $\theta = \infty$, so that $K_y = \infty$ in (47), and there does not exist a $y(z)$ for which the right-hand side in (48) is positive, implying we cannot guarantee a strictly positive lower bound on $T''(y(z))$. ■

B.2 Analysis of the ODE for the optimal marginal tax rate

We now derive and analyze the differential equation (22) that characterizes the behavior of the optimal marginal tax rate. This will provide intuition for the subsequent proof of Theorem 3.2 in the next subseection. We restrict our attention to the case when the benchmark type distribution is Pareto, as the typical case that leads to nonzero top marginal taxes in absence of model misspecification concerns.

Lemma B.4. *When z is Pareto distributed with shape parameter α under the benchmark model, the worst-*

case density satisfies

$$\frac{d}{dz}\log\widetilde{f}(z)=\frac{d}{dz}\log m(z)+\frac{d}{dz}\log f(z)=-\frac{\mu}{\theta}T'(y(z))\,y'(z)-(\alpha+1)\,\frac{1}{z}.$$

Proof. Since

$$\begin{aligned}m(z) &= \bar{m}\exp\left(-\frac{\mu}{\theta}T(y(z))\right)\\ f(z) &= \frac{\alpha}{z^{\alpha+1}}\end{aligned}$$

we have, by direct computation,

$$\frac{d}{dz}\log\widetilde{f}(z)=-\frac{\mu}{\theta}T'(y(z))\,y'(z)-(\alpha+1)\,\frac{1}{z}.$$

■

Proposition B.5. *When the type distribution under the benchmark model is Pareto with shape parameter α, the optimal marginal tax $T'(y)$ obeys the differential equation*

$$-\frac{T''(y)\,y}{1-T'(y)}=-\left[2-\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]^{-1}\left[\frac{\mu}{\theta}\left[T'(y)\right]^{2}y-\gamma+\gamma\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]. \tag{49}$$

Proof. We start with the optimal tax formula

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

Differentiating this formula with respect to z, and using Lemma B.3,

$$\frac{T''(y(z))\,y'(z)}{\left(1-T'(y(z))\right)^{2}}=(1+\gamma)\left[-\frac{1}{z}-\widetilde{\phi}(z)\left[\frac{1}{z}-\frac{\mu}{\theta}T'(y(z))\,y'(z)+\frac{f'(z)}{f(z)}\right]\right].$$

Combining $y'(z)$ terms, we have

$$\left[\frac{T''(y(z))}{T'(y(z))\left(1-T'(y(z))\right)}-\frac{\mu}{\theta}T'(y(z))\right]y'(z)\,z=-\widetilde{\phi}(z)^{-1}-1-z\frac{f'(z)}{f(z)}.$$

We can now use Lemma B.1 to substitute out $y'(z)$ and obtain

$$\left[\frac{T''(y(z))}{T'(y(z))\left(1-T'(y(z))\right)}-\frac{\mu}{\theta}T'(y(z))\right]\frac{(1+\gamma)\,y(z)}{\gamma+\frac{T''(y(z))}{1-T'(y(z))}y(z)}=-\widetilde{\phi}(z)^{-1}-1-z\frac{f'(z)}{f(z)}.$$

We can now multiply by the denominator of the compound fraction on the left-hand side, use the expression for $\widetilde{\varphi}(z)$, and combine terms that contain $T''(y(z))$:

$$\left[2-\frac{\gamma-z\frac{f'(z)}{f(z)}}{1+\gamma}T'(y(z))\right]\frac{T''(y(z))}{1-T'(y(z))}y(z)=\frac{\mu}{\theta}\left[T'(y(z))\right]^2 y(z)-\gamma+\gamma\frac{\gamma-z\frac{f'(z)}{f(z)}}{1+\gamma}T'(y(z)).$$

Finally, for the case of the Pareto density,

$$-z\frac{f'(z)}{f(z)}=1+\alpha.$$

Substituting this expression in, we notice that the resulting differential equation does not depend explicitly on z. Since $y(z)$ is strictly monotonic, we can drop the z argument and rewrite the equation as a differential equation for $T(y)$, yielding the expression in the statement of the proposition.
■

We now study the phase diagram of the differential equation (49), which we can rewrite as

$$T''(y)=\frac{1-T'(y)}{y}\left[2-\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]^{-1}\left[\frac{\mu}{\theta}\left[T'(y)\right]^2 y-\gamma+\gamma\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right].$$

The resulting phase diagram is depicted in Figure 1.

Define the right-hand side of the above equation as a function $h:[0,\infty)\times(-\infty,1)\to\mathbb{R}$:

$$h(y,\tau)=\frac{1-\tau}{y}\left[2-\frac{1+\gamma+\alpha}{1+\gamma}\tau\right]^{-1}\left[\frac{\mu}{\theta}\tau^2 y-\gamma+\gamma\frac{1+\gamma+\alpha}{1+\gamma}\tau\right].$$

We study the function on $(y,\tau)\in(0,\infty)\times(-\infty,1)$. For simplicity, we assume that

$$\frac{1+\gamma+\alpha}{1+\gamma}<2,$$

so that the first bracket in the definition of $h(y,\tau)$ is never zero for $\tau\in(-\infty,1)$. This does not change any conclusions about asymptotic behavior of the optimal tax.

For a given $y\in(0,\infty)$, we first find the isoclines by solving for $\bar{\tau}(y)$ such that $f(y,\bar{\tau}(y))=0$. This $\bar{\tau}(y)$ solves the cubic equation

$$(1-\bar{\tau}(y))\left[\frac{\mu}{\theta}y\bar{\tau}(y)^2+\gamma\frac{1+\gamma+\alpha}{1+\gamma}\bar{\tau}(y)-\gamma\right]=0$$

with three solutions

$$\begin{aligned}\bar{\tau}_{1,2}(y)&=\frac{-\gamma\frac{1+\gamma+\alpha}{1+\gamma}\pm\sqrt{\left(\gamma\frac{1+\gamma+\alpha}{1+\gamma}\right)^2+4\frac{\mu}{\theta}\gamma y}}{2\frac{\mu}{\theta}y}\\ \bar{\tau}_3(y)&=1,\end{aligned}$$

where $\bar{\tau}_1(y)$ denotes the root with the minus sign. The isoclines are depicted with the black dashed lines in Figure 1. Asymptotically, the differential equation has two steady states

$$\begin{aligned}\lim_{y\to\infty}\bar{\tau}_{1,2}(y) &= 0\\ \lim_{y\to\infty}\bar{\tau}_{3}(y) &= 1.\end{aligned}$$

We can order the three isoclines as

$$\bar{\tau}_1(y) < 0 < \bar{\tau}_2(y) < \bar{\tau}_3(y),$$

and then, as depicted in the phase diagram,

$$\begin{aligned}f(y,\tau) &> 0 \qquad \tau < \bar{\tau}_1(y)\\ f(y,\tau) &< 0 \qquad \bar{\tau}_1(y) < \tau < \bar{\tau}_2(y)\\ f(y,\tau) &> 0 \qquad \bar{\tau}_2(y) < \tau < \bar{\tau}_3(y).\end{aligned}$$

The result from Theorem 3.1,

$$\lim_{z\to\infty} T'(y(z)) = \lim_{y\to\infty} T'(y) = 0,$$

is a transversality condition that pins down the unique optimal marginal tax function $T'(y)$. This optimal path is depicted in the phase diagram with the red solid line with bullet markers. It follows from the proofs of theorems 3.1 and 3.2 that any other path for $T'(y)$ that satisfies equation (49) either converges to one, or becomes negative for some sufficiently high y, both of which violate conditions that the optimal marginal tax function has to satisfy.

B.3 Proof of Theorem 3.2

Using the insights from the phase diagram, we now turn to the proof of Theorem 3.2.

Proof of Theorem 3.2. We investigate the limiting behavior of the differential equation (49). The left-hand side of this equation is the elasticity of take-home rate $1-T'(y)$ with respect to income, or

$$\frac{d\log(1-T'(y))}{d\log y} = -\left[2-\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]^{-1}\left[\frac{\mu}{\theta}\left[T'(y)\right]^2 y-\gamma+\gamma\frac{1+\gamma+\alpha}{1+\gamma}T'(y)\right]. \tag{50}$$

Since $\lim_{y\to\infty}(1-T'(y)) = 1$, if the limit as $y\to\infty$ of the left-hand side of the above equation exists, it has to be zero. Assume for now that this limit exists. Since $\lim_{y\to\infty} T'(y) = 0$, the first bracket on the right-hand side converges to a positive number as $y\to\infty$. For the same reason, the last term of the second bracket converges to zero as well. Hence the only way how the second bracket converges to zero is when

$$\lim_{y\to\infty}\frac{\mu}{\theta}\left[T'(y)\right]^2 y = \gamma. \tag{51}$$

In the remainder of the proof, we prove that the limit indeed exists. Assume it does not, so that there exists an ε such that for any $\check{y}$, there exists a $y \geq \check{y}$ such that

$$\left| \frac{\mu}{\theta} \left[T'(y)\right]^2 y - \gamma \right| > \varepsilon > 0. \tag{52}$$

Further, the analysis of the phase diagram implies that along the optimal path, $T'(y)$ monotonically decreases to zero. This means that $T''(y) < 0$, and the elasticity in (50) is strictly positive. In addition, for any arbitrarily small ε_τ, there exists a $\check{y}_\tau$ such that $0 < T'(y) < \varepsilon_\tau$ for all $y \geq \check{y}_\tau$.

Assume first that inequality (52) holds as $\frac{\mu}{\theta}\left[T'(y)\right]^2 y - \gamma > \varepsilon$. Then the elasticity in (50) can be bounded as

$$\frac{d \log\left(1 - T'(y)\right)}{d \log y} = \frac{-T''(y)\, y}{1 - T'(y)} < -\left[2 - \frac{1+\gamma+\alpha}{1+\gamma} T'(y)\right]^{-1} \left[\varepsilon + \gamma \frac{1+\gamma+\alpha}{1+\gamma} T'(y)\right] < 0,$$

which is a contradiction with $T''(y) < 0$, given that $T'(y) > 0$ and $\lim_{y \to \infty} T'(y) = 0$.

On the other hand, assume that inequality (52) holds as $\frac{\mu}{\theta}\left[T'(y)\right]^2 y - \gamma < -\varepsilon$. Then the elasticity in (50) can be bounded as

$$\begin{aligned} \frac{d \log\left(1 - T'(y)\right)}{d \log y} &= \frac{-T''(y)\, y}{1 - T'(y)} > -\frac{1}{2}\left[\frac{\mu}{\theta}\left[T'(y)\right]^2 y - \gamma + \gamma \frac{1+\gamma+\alpha}{1+\gamma} T'(y)\right] \\ &> -\frac{1}{2}\left[\frac{\mu}{\theta}\left[T'(y)\right]^2 y - \gamma\right] - \frac{\gamma}{2}\frac{1+\gamma+\alpha}{1+\gamma}\varepsilon_\tau > \frac{\varepsilon}{2} - \frac{\gamma}{2}\frac{1+\gamma+\alpha}{1+\gamma}\varepsilon_\tau > \frac{\varepsilon}{4}, \end{aligned}$$

where the last inequality follows from the fact that ε_τ can be taken to be arbitrarily small when we restrict our attention to sufficiently large $y \geq \check{y} \geq \check{y}_\tau$. We therefore obtain

$$T''(y)\, y < -\frac{\varepsilon}{4}\left(1 - T'(y)\right).$$

As a consequence

$$\begin{aligned} \frac{d}{dy}\left[\frac{\mu}{\theta}\left[T'(y)\right]^2 y - \gamma\right] &= \frac{\mu}{\theta}\left[2T''(y)\, T'(y)\, y + \left[T'(y)\right]^2\right] \\ &< \frac{\mu}{\theta}\left[-\frac{\varepsilon}{2}\left(1 - T'(y)\right) T'(y) + \left[T'(y)\right]^2\right] \\ &= \frac{\mu}{\theta} T'(y)\left[-\frac{\varepsilon}{2} + \left(1 + \frac{\varepsilon}{2}\right)\left[T'(y)\right]\right], \end{aligned}$$

which becomes negative for sufficiently large y because $\lim_{y \to \infty} T'(y) = 0$. Denote this y as y^*. This means that inequality $\frac{\mu}{\theta}\left[T'(y)\right]^2 y - \gamma < -\varepsilon$ continues to hold as y increases above y^*. Integrating up the inequality

$$d \log\left(1 - T'(y)\right) > \frac{\varepsilon}{4} d \log y$$

for $y \geq y^*$ yields

$$\left(1 - T'(y)\right) - \log\left(1 - T'(y^*)\right) > \frac{\varepsilon}{4}\left(\log y - \log y^*\right)$$

and hence

$$T'(y) < 1 - \left(\frac{y}{y^*}\right)^{\varepsilon/4}\left(1 - T'(y^*)\right).$$

Since $y \to \infty$, the right-hand side must ultimately become negative, which violates the restriction

$T'(y) > 0$.

We have thus shown that condition (52), which is equivalent to a violation of equation (51), cannot hold simulatenously with other restrictions on the optimal marginal tax rate. Either the marginal tax rate function would have to become increasing, or become negative. From the perspective of the phase diagram in Figure 1, analyzed in Appendix B.2, if condition (52) holds for a sufficiently large y, it must be that the given $T'(y)$ is either on a trajectory above the optimal path that crosses the $\bar{\tau}_2(y)$ isocline and converges to one, or becomes negative, crossing the $\bar{\tau}_1(y)$ isocline.

Finally, equation (51) implies that the marginal tax rate has to decay as

$$\log T'(y) = -\frac{1}{2} \log y + o(y)$$

where $o(y)$ converges to a constant as $y \to \infty$. Hence, differentiating this expression with respect to $\log y$, and taking the limit as $y \to \infty$, this limit, if it exists, must be given by expression (19). ■

C Generalizations of the baseline model

[TO BE COMPLETED]

References

Adam, Klaus and Albert Marcet. 2011. Internal Rationality, Imperfect Market Knowledge and Asset Prices. *Incompleteness and Uncertainty in Economics* 146 (3):1224–1252. 2

Almeida, Caio and René Garcia. 2017. Economic Implications of Nonlinear Pricing Kernels. *Management Science* 63 (10):3361–3380. 22

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association* 1 (1):68–123. 4, 24

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho. 2022. Survey Data and Subjective Beliefs in Business Cycle Models. 2

Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman. 2016. Misspecified Recovery. *Journal of Finance* 71 (6):2493–2544. 22

Chang, Minsu and Chunzan Wu. 2023. When in Doubt, Tax More Progressively? Uncertainty and Progressive Income Taxation. *mimeo* . 2

Cressie, Noel and Timothy R. C. Read. 1984. Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society, Series B (Methodological)* 46 (3):440–464. 22

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. 1, 9, 10

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. 12

Ferriere, Axelle and Anastasios G. Karantounias. 2019. Fiscal Austerity in Ambiguous Times. *American Economic Journal: Macroeconomics* 11 (1):89–131. 33

Golosov, Mikhail and Ilia Krasikov. 2023. The Optimal Taxation of Couples. *mimeo* . 5

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski. 2016. Redistribution and Social Insurance. *American Economic Review* 106 (2):359–86. 1

Hansen, Lars Peter and Jianjun Miao. 2022. Asset Pricing under Smooth Ambiguity in Continuous Time. *Economic Theory* 74 (2):335–371. 2

Hansen, Lars Peter and Thomas J. Sargent. 2001a. Acknowledging Misspecification in Macroeconomic Theory. *Review of Economic Dynamics* 4 (3):519–535. 6

———. 2001b. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66. 6

———. 2008. *Robustness*. Princeton, N.J.: Princeton, N.J. 1

———. 2010. Fragile Beliefs and the Price of Uncertainty. *Quantitative Economics* 1 (1):129–162. 25

———. 2012. Three Types of Ambiguity. *Journal of Monetary Economics* 59 (5):422–445. 2, 33

———. 2015. Four Types of Ignorance. *Journal of Monetary Economics* 69:97–113. 33

Heathcote, Jonathan and Hitoshi Tsujiyama. 2021. Optimal Income Taxation: Mirrlees Meets Ramsey. *Journal of Political Economy* 129 (11):3141–3184. 4, 25, 28

Karantounias, Anastasios G. 2013. Managing Pessimistic Expectations and Fiscal Policy. *Theoretical Economics* 8 (1):193–231. 2

———. 2023. Doubts about the Model and Optimal Policy. *Journal of Economic Theory* 210:105643. 33

Kwon, Hyosung and Jianjun Miao. 2017. Three Types of Robust Ramsey Problems in a Linear-Quadratic Framework. *Journal of Economic Dynamics and Control* 76:211–231. 33

Lockwood, Benjamin B., Afras Sial, and Matthew Weinzierl. 2021. Designing, Not Checking, for Policy Robustness: An Example with Optimal Taxation. *Tax Policy and the Economy* 35:1–54. 2

Mirrlees, James A. 1971. An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* 38 (2):175–208. 1, 8

Negishi, Takashi. 1960. Welfare Economics and Existence of an Equilibrium for a Competitive Economy. *Metroeconomica* 12 (2–3):92–97. 6

Saez, Emmanuel. 2001. Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies* 68 (1):205–229. 1, 9, 10