Optimal Policy with Endogenous Wages

Fatima Djalalova
Advisor: Professor Casey Rothschild

Submitted in Partial Fulfillment of the
Prerequisite for Honors in Economics

April 2024
Abstract

This paper incorporates the effects of government spending on wage distribution into the standard model of optimal linear taxation. Our model captures that government spending can increase wages of consumers, and that this rate of increase can vary across individuals. We characterize optimal tax and spending policy in settings with homogenous wage effects and settings with heterogeneous wage effects. We show results both for optimal levels of taxes and spending, and for the distribution of that spending across projects. We show that, under a natural condition, wage endogeneity implies that government spending and tax rates are higher than standard models would suggest. Notably, government spending (and taxes) are higher than in standard models even when spending only increases the wage of high-skill types and the government has a Rawlsian objective function that puts zero welfare weight on high-skill types. We also show that government spending across projects should be distorted away from efficiency towards projects that differentially benefit the wages of the lower-skill type.
Acknowledgements

I have no words to express my deepest gratitude to my thesis advisor, Professor Casey Rothschild. His incredible generosity with his time, knowledge, and wisdom has supported me through the most academically challenging moments and made the thesis process the most favorite part of my senior year. Professor Rothschild’s positive attitude, kindness, and passion for economics make each lecture and conversation with him incredibly inspiring.

I thank Professor Eric Hilt and all of the Economics Research Seminar participants for the great suggestions and support. I learned so much from all the presentations and discussions.

I thank Professor Susan Skeath for being my committee member and for being the first professor to introduce me to—and make me love—microeconomic theory. I thank my honors visitor, Professor Stanley Chang, for teaching me the mathematics and research skills that helped me immensely in my thesis work.

I thank Professor Courtney Coile, Professor Alex Diesl, and Professor Kyung Park for their fantastic support, their valuable advice, and for teaching some of my most favorite classes at Wellesley. I thank Professor Olga Shurchkov for the wonderful Case Fellow summer research experience. I am incredibly grateful to all of the professors with whom I had a chance to take a class.

I thank my dear friends for their encouragement and their positive energy.

I thank my family for their support, their love, and for always being with me despite the distance.
Contents

1 Introduction ........................................ 4
2 Efficiency motive in taxation .................... 7
3 Redistributive motive in taxation ............... 15
4 Efficiency-redistribution trade-off .............. 20
5 Optimal spending across projects ............... 26
6 Conclusion and discussion ....................... 33
References ........................................... 34
Appendix .............................................. 37
1 Introduction

How does the level and type of government spending affect the distribution of wages, and what does this effect imply for the optimal tax policy? The standard tax models lack the tools to study this question since they assume a fixed distribution of wages. In these models, the government that wishes to influence efficiency and redistribution has only two instruments at its disposal: taxes and transfers. Yet, the government spending policy—funded by the tax revenue—itself can affect efficiency and redistribution. For example, policies targeted at improving early childhood education and nutrition pay for themselves through increased human capital and earnings potential in adulthood, and simultaneously help to decrease socioeconomic inequality (Gray-Lobe et al., 2023; Heckman et al., 2013; Hendren & Sprung-Keyser, 2020). Conversely, policies that provide private school vouchers improve efficiency at the higher end of the income distribution (Romer, 2002), but can reduce educational outcomes for disadvantaged students (Abdulkadiroğlu et al., 2018).

This study of optimal tax policy explicitly recognizes the impact of government spending on individuals’ wages. In standard models, higher tax rates disincentivize effort, but can serve as a corrective tool for redistribution. With endogenous wages, higher tax rates have additional effects on efficiency and distribution because higher tax revenues are used to fund wage-influencing government spending. Suppose that government spending increases wages, and that wages of low-income individuals increase at a higher rate. On the one hand, this provides a reason to increase taxes—relative to a standard model without wage effects—because the higher wages facilitated by higher taxes are welfare improving. On the other hand, higher tax rates lead to a more compressed distribution of wages, reducing the benefits of redistributive taxation, and thus potentially call for less progressive tax rates. These countervailing forces make the analysis non-trivial, and motivate us to expand the standard model to incorporate variable wages. In developing our model, we are primarily interested in assessing the difference from the tax rate formulas that assume fixed wages, and in examining the welfare consequences of government’s choices of public investments.

This paper begins with a simple representative agent Ramsey model which isolates the efficiency motive for taxation. In this model, the government can invest in technology that improves wages according to an exogenous wage elasticity parameter. Suppose, for instance, that government spends on the school lunch program. Then government spending increases individuals’ consumption directly (acts as an effective lump-sum transfer) and indirectly (improves wages in adulthood). The endogeneity of wages calls for an alternative definition of the standard tax rate. We rely on the concept of a self-confirming policy equilibrium (SCPE), developed by Rothschild and Scheuer (2011). An SCPE describes the tax policy
that would emerge in the same economy if the planner incorrectly believed in exogeneity of wages. The naive planner perceives the wage distribution as fixed, and solves for the optimal tax rate that indeed induces the wage distribution that the planner started with. In this section, the SCPE tax rate coincides with the standard zero tax rate. A sophisticated planner recognizes the extra effect of taxation on productivity and selects a tax rate that is strictly above the standard tax rate. Intuitively, the improvement in wages results in an expanded government budget set that shifts social preferences towards non-zero tax rates. Proposition 1 characterizes the optimal tax policy that emerges in this simple model, and Proposition 2 establishes that the optimal tax rate always exceeds the standard tax rate.

We proceed to introduce heterogeneity in skill to study the redistributive motive. This extension assumes that individuals differ in their skill (low- and high- ability) and mass, that government spending increases wages of both types proportionately, and that the social welfare weights are exogenously given. Lemma 2 characterizes the optimal linear tax, which now depends on both the productivity effects of government spending, and the desire for redistribution. Lemma 3 describes the SCPE tax rate that emerges under the assumption of fixed wages, and Proposition 3 shows that the optimal tax rate strictly exceeds the standard tax unless the naive planner has more optimistic beliefs about the direct consumption effects of government spending. In this context, optimism refers to the divergence between the true and naive social planner’s view about the effect of government spending on consumption.

Next we consider a model with heterogeneous effects of government spending. In this model, government spending affects both the productivity and the distribution of wages. Suppose that government spends on technology with higher wage elasticity for the high-types (such as private school vouchers). Then government spending increases wages, but also widens the wage gap as the earning ability of high-types increases at a faster rate. Contrary to the standard analysis, taxation in this model increases efficiency but redistributes resources to the rich. While this section lacks analytical results, we provide numerical simulations that compare the SCPE and Pareto optimal tax rates under different assumptions on wage elasticity parameters. We also consider a scenario with Rawlsian weights and government spending that benefits only the high-types, and find that the sophisticated planner invests in this project if the direct effects on consumption are sufficiently large so as to increase the resources available for redistribution.

Our most general model analyzes the optimal allocation of government budget across projects with different wage elasticities. When government projects have homogeneous effects on wages, the optimal allocation of resources across the projects depends solely on the productivity effects of the projects, as shown in Proposition 4. When government spending projects have heterogeneous effects on wages, the optimal allocation of resources also depends
on the taste for redistribution. Proposition 5 shows that, in the presence of a redistributive
motive, the optimal allocation of spending distorts away from efficiency towards projects
that differentially benefit the wages of the low-types.

**Related literature.** This paper closely relates to the literature on optimal linear income
taxation with endogenous wages. Feldstein (1973) numerically solves for the optimal linear
tax in a two-type model when wages depend on the aggregate labor supply in the economy,
and finds that the optimal progressivity of the tax rate does not depend on the assumption
of variable wages. Allen (1982) argues that the result in Feldstein (1973) is sensitive to the
assumption of a Cobb-Douglas production function. Using a general production function,
Allen (1982) presents the conditions under which endogenous wages substantially change the
optimal tax rate formula. Zeckhauser and Schelling (1977) use a two-type model to generate
a series of examples where neglecting the general equilibrium effects of taxation leads to
unintended negative consequences on welfare.

The research on nonlinear taxation with endogenous wages further informs our analysis.
Stiglitz (1982) uses mechanism design approach to characterize optimal nonlinear tax in a
two-type model when wages depend on the overall supply of labor. He finds that, unless
the individuals are perfect substitutes, the optimal tax schedule involves a positive and
negative marginal tax rates on low- and high-skill individuals, respectively—challenging the
Mirrlees (1971) zero tax rate at the top result. He explains that the lower tax rate on high-
skill individuals encourages supply of high-skill labor, decreasing the returns to the high-
skill labor and increasing the returns to low-skill labor, resulting in indirect redistribution.
Rothschild and Scheuer (2013) derive similar results when they extend the Stiglitz (1982)
setup to a two-sector Roy model of occupational choice. Using a similar model with two-
dimensional skill heterogeneity, Rothschild and Scheuer (2016) introduce the wedge between
private and social returns to effort when individuals can engage in rent-seeking. They find
that the optimal tax is below the Pigouvian tax rate so as to keep returns to rent-seeking
low and discourage further rent-seeking activity. Other papers that introduce endogenous
wages into the Mirrleesian setting include Ales et al. (2015). Some studies, such as Sachs
et al. (2020), instead use Chetty (2009)-style sufficient statistics approach to characterize
optimal nonlinear taxes in endogenous wage setting.

Our analysis tracks the seminal contributions by Ramsey (1927), Mirrlees (1971), Dia-
mond (1998), and Saez (2001) (Piketty and Saez (2013) provide a survey of optimal income
taxation literature). To examine how the tax rate changes under the assumption of en-
dogenous wages, we rely on the concept of self-confirming policy equilibrium introduced by
Rothschild and Scheuer (2011). Rothschild and Scheuer (2011) and Rothschild and Scheuer
(2013) use the SCPE to describe the tax policy that solves the following fixed point problem:
given a fixed distribution of wages, we solve for the optimal tax rate that indeed results in
the distribution of wages that we started with. This concept connects to learning in games
literature, such as Fudenberg and Levine (1993), Sargent (2008), and Fudenberg and Levine
(2009), where the self-confirming equilibrium occurs when a player has wrong beliefs about
the off-path play (in our case, the planner incorrectly believes that wages are exogenous), but
since the player starts at an equilibrium where incorrect beliefs lead to the same outcome as
under the correct beliefs (the observed wage distribution is indeed the distribution generated
by the current tax code), the player does not get a chance to observe the off-path play (the
planner perceives the tax rates as optimal, so they do not ever learn that changing the tax
rate would in fact change the wage distribution).

Our article further connects to the literature on tax incidence (e.g. Harberger (1964); Saez and Zucman (2023); Kotlikoff and Summers (1987) and Fullerton and Metcalf (2002)
provide a summary of literature) and the distributional consequences of government policies
(e.g. Hendren and Sprung-Keyser (2020); Bénabou (1996); Bénabou (2002); Finkelstein et
al. (2018); Hoxby (2001); Hoxby (1996)).

Outline of the paper. Section 2 presents a simple model with no heterogeneity to
isolate efficiency motive in taxation. Section 3 introduces the redistributive motive, and
describes how assumptions on parameters affect the comparison of the standard and optimal
linear tax rates. Section 4 presents the model with heterogeneous effects on wages. Section
5 compares the optimal allocation of budget across projects with homogeneous and hetero-
genous effects on wages. Section 6 concludes. All derivations and proofs are relegated to
the Appendix.

2 Efficiency motive in taxation

A simple model without heterogeneity isolates the efficiency implications for optimal taxation
of spending-dependent wages. This section presents the results for the one-type model,
defines the self-confirming policy equilibrium (SCPE), and uses the social indifference curves
and the government budget constraint to graphically visualize the optimal tax policy.

A simple model without heterogeneity. Consider an economy consisting of a unit
mass of identical individuals. Agents have quasi-linear and iso-elastic preferences of form
\( u(c, l) = c - \frac{kl^\gamma}{\gamma} \), where \( k > 0 \) normalizes the disutility of labor relative to consumption,
\( \gamma > 1 \), and the wage elasticity of earnings is constant \( \varepsilon_{y,w} \equiv \frac{\gamma}{\gamma - 1} \). This form for preferences
simplifies the analysis of the effect of tax rate on labor supply by assuming no income effects
and constant elasticity of substitution.

While there is no scope for redistribution of income, a benevolent social planner sets a
linear tax rate $t \in [0, 1]$ and uses the collected tax revenue $g$ to invest in technology that increases individuals’ earnings capacities $w(g)$. The examples of such public investments include infrastructure, public health, education, rule of law, national defense, and measurement standards. Wages increase according to $w(g) = w_0 g^\beta$, where $\beta \in (0, \frac{1}{\varepsilon_{l,w}})$ denotes the wage elasticity of government spending, and $w_0$ is normalized as the wage when the level of government spending is $g = 1$. Government spending may also increase consumption directly by $T(g) = \nu g$, for some $\nu \in [0, 1]$, where we use the notation $T(g)$ to highlight its similarity to an (effective) lump-sum transfer.

The consumers take marginal tax rate $t$, government spending $g$, and hence wages $w \equiv w(g)$ as given, and solve

$$\max_{c,l} u(c,l) \quad \text{s.t.} \quad c \leq \nu g + (1 - t)w(g)l,$$

which yields the standard first-order condition for individual choice

$$h'(l) = (1 - t)w(g),$$

where $h(l) = \frac{kl}{\gamma}$ measures the disutility from labor, and the solutions

$$\{c^*(t,g), l^*(t,g)\} = \left\{\nu g + (1 - t)w(g)l^*(t,g), \left[\frac{(1 - t)w(g)}{k}\right]^{\frac{1}{1-\gamma}}\right\}.$$  

As discussed in Diamond (1998), the assumption of quasi-linearity ensures that the lump-sum transfers have no effect on efficiency. The level of effective transfers $\nu g$ does not distort the agent’s labor supply. Indeed, the labor supply solely depends on the after-tax-and-government-spending wages, and the wage elasticity of labor supply $\varepsilon_{l,w} \equiv \frac{1}{\gamma-1}$ represents both the compensated and uncompensated elasticity. This optimal choice of labor and consumption allows the planner to trace out the feasible tax policies as follows:

**Definition 1** A *feasible linear tax allocation* is an allocation $\{c^*(t,g), l^*(t,g)\}$ and a tax policy $(t,g)$ such that

(i) $\{c^*(t,g), l^*(t,g)\}$ solves program (1) given $(t,g)$ and $w \equiv w(g)$,
(ii) the government budget balances

\[ g = tw(g)l^*(t, g). \tag{4} \]

Condition (ii) describes a fixed point problem for government spending. Given a linear tax \( t \), the level of government spending \( g \) determines the wages \( w(g) \) and hence the level of effort \( l^*(t, w(g)) \). The tax revenue at this level of output, given by the right-hand side of equation (4), has to be equal to the level of \( g \) that induced that level of output. A stable solution to this fixed point problem exists if and only if \( \beta < \frac{1}{\varepsilon_{y,w}} \). As shown formally in Appendix A, for \( \beta \geq \frac{1}{\varepsilon_{y,w}} \), there exists no stable fixed point other than \( g = 0 \). This phenomenon connects to the discussion in Malcomson (1986) that the level of tax revenue may not be continuous or have an interior maximum when we introduce variable wages. Therefore, a feasible tax allocation exists if and only if the following assumption holds.

**Assumption 1** The wage elasticity parameter \( \beta \in (0, 1/\varepsilon_{y,w}) \).

Under Assumption 1, the budget balance requirement in equation (4) allows to express the government spending as a function of the tax rate, and gives rise to the following lemma.

**Lemma 1** The set of feasible linear tax allocations can be parametrized by the level of linear tax \( t \) as \((t, \hat{g}(t))\), where \( \hat{g}(t) \) is a continuous function of \( t \) given by

\[ \hat{g}(t) = \left[ \frac{(1-t)t^{\gamma-1}w_0^\gamma}{k} \right]^{\frac{1}{\gamma-\beta\gamma-1}} = [tw_0l^*(t, w(1))]^{\gamma-\beta\gamma-1} = [tw_0l^*(t, w(1))]^{\frac{1}{\gamma-\beta\gamma-1}}. \tag{5} \]

The feasible level of government spending \( \hat{g}(t) \) is a quasi-concave function that attains a unique maximum at \( t_{Laffer} = \frac{1}{\varepsilon_{y,w}} \).

Expression (5) characterizes the unique level of government spending consistent with a feasible linear tax allocation. Figure 1(a) plots the level of government spending \( \hat{g}(t) \) against the tax rate. Unlike the Laffer curve that assumes exogenous wages, the budget curve with endogenous wages \( \hat{g}(t) \) is not strictly concave. Figure 1(b) sheds light on the origin of convexity for low levels of tax rates by adding a labor supply curve (in red). Contrary to the standard negative effect of (a positive) tax rate on labor supply, the endogeneity of wages produces an interval on which the level of effort increases. This interval is given by \([0, t_{effort}]\), where the tax rate \( t_{effort} = \beta > 0 \) maximizes the level of effort, i.e. \( \frac{dl^*(t,w(\hat{g}(t)))}{dt}|_{t=t_{effort}} = 0 \). Intuitively, for low levels of taxes, a small increase in the tax rate results in net increase in after-tax wages, thereby incentivizing effort and leading to an increase in marginal revenue collected. This effect on marginal revenue switches to negative at \( t_{effort} = \beta > 0 \), a tax rate.
that uniquely maximizes the level of after-tax wages (and hence the effort). Remarkably, the revenue maximizing tax rate \( t^{Laffer} = \frac{1}{\epsilon_{y,w}} \) coincides with that for the standard model with exogenous wages. This feature reflects that, at the peak of the Laffer curve, a local change in the tax rate bears no effect on government spending, so wages remain locally fixed and produce no feedback effects on the budget.

![Figure 1: Government budget set](image)

Faced with a continuum of feasible linear tax allocations \( \{(t, \hat{g}(t)) : t \in [0, 1]\} \), a planner seeks to select the tax policy that maximizes social welfare. In the representative agent model, the social and individual preferences coincide, so the indirect utility function \( v(t, g) = u(c^*(t, g), l^*(t, g)) \) measures the welfare in the economy. A planner who is informed about the effect of government spending on wage thus solves

\[
\max_{t, g} v(t, g) \quad \text{s.t.} \quad g \leq tw(\hat{g})l^*(t, w(\hat{g})),
\]

which leads to the following definition:

**Definition 2** A *Pareto optimum with linear taxes* is a feasible tax allocation \((t^{PO}, g)\) that solves program (6).

By Lemma 1, the informed planner’s objective in (6) can be re-written as

\[
\max_t v(t, \hat{g}(t)) = \max_t \left[ \nu \hat{g}(t) + (1 - t)w(\hat{g})l^*(t, w(\hat{g})) - h(l^*(t, w(\hat{g}))) \right].
\]

The informed planner selects the tax rate that maximizes individual utility while taking into account the effect of tax \( t \) on government spending \( \hat{g}(t) \), the distribution of wages \( w(\hat{g}(t)) \),
and the optimal choice of labor supply by consumers \( l^*(t, w(\hat{g})) \). The next proposition shows that this maximum exists and characterizes the optimal tax rate.

**Proposition 1** The first-order conditions for problem (7) characterize the optimal tax rate as

\[
\tau_{PO} = \frac{(\beta - \nu + 1) - \sqrt{(\beta - \nu + 1)^2 - 4\beta(1 - \nu \varepsilon_{y,w})}}{2(1 - \nu \varepsilon_{y,w})} > 0,
\]

with the properties that

(i) \( \tau_{PO}(\nu = 0) = \tau_{effort} = \beta \),

(ii) \( \tau_{PO} < \tau_{Laffer} \),

(iii) \( \tau_{PO} \) increases in \( \nu \),

(iv) \( \tau_{PO} \) increases in \( \beta \).

Figure 2 shows that the optimum occurs at the tangency of the government spending and social indifference curves. The shape of the social indifference curves reflects that the individual utility increases in government spending, but decreases in tax rate.

![Figure 2: Optimal tax policy](image)

Appendix C contains the technical details of properties outlined in Proposition 1, but we provide the intuition for key results here. Property (ii) states that the optimal tax rate \( \tau_{PO} \) is bounded above by the revenue maximizing tax rate from Lemma 1. Indeed, we expect the optimal tax not to exceed the government spending maximizing tax rate \( \tau_{Laffer} \), because for each tax rate \( t_1 > \tau_{Laffer} \), there is a tax \( t_2 < \tau_{Laffer} \) such that \( g(t_1) = g(t_2) \), but \( t_1 \).
causes greater distortions to labor supply. This bound on the optimal tax rate allows us to constrain the analysis to the upward sloping part of the government budget curve and derive property (iii).

![Figure 3: Illustration of Proposition 1](image)

Figure 3(a) shows that a higher \( \nu \) raises the marginal benefit of government spending, flattening the social indifference curves (in dashed red). The budget constraint (in blue) remains the same because it is not affected by the level of transfers, so the optimum shifts in the direction of higher tax rates. Applying a similar argument to analyze an increase in \( \beta \) is complicated by the fact that \( \beta \) affects not only the slope of the indifference curve, but also the slope and the level of the budget constraint. An increase in \( \beta \) affects the budget constraint through two channels: the effect on the level \( g \) if wages remained fixed, and the effect on the slope of \( g \) due to re-optimized tax policy parameters. Since the level effect is always positive, the effect on the tax policy depends only on the change in the slopes of indifference curve and the budget constraint relative to each other. For this purpose, we define the compensated budget constraint to be the budget constraint absent the level effects. Figure 3(b) shows that the compensated budget constraint (in dashed blue) becomes steeper, while the indifference curve (in dashed red) becomes flatter. Therefore, the optimum moves in the direction of higher tax rates. Appendix C shows that an uncompensated increase in \( \beta \) also increases \( t^{PO} \). Figure 4 summarizes the comparative statics results from numerical simulations: the optimal tax increases in parameters \( \nu \) and \( \beta \), and does not exceed \( t^{Laffer} \) (this figure assumes \( \gamma = 3 \), so \( t^{Laffer} = \frac{1}{\varepsilon_{yw}} \approx 0.667 \)).
Figure 4: Numerical simulation of $t^PO$ vs. $\beta$, by levels of $\nu$

A social planner who is unaware of the endogeneity of wages, believes that government spending increases individual well-being only through direct effects on consumption $T(g) = \nu^{sc} g$, for some $\nu^{sc} \in [0, 1]$. Note that we use the notation $\nu^{sc}$ to indicate that the naive social planner’s view about the effect of government spending on consumption can differ from the true effect of that spending. We refer to any divergence between the two as an optimism gap. The naive planner thus solves

$$\max_{t,g} u(c^*(t;w), l^*(t;w)) \text{ s.t. } g \leq twl^*(t;w),$$ \hspace{1cm} (9)

taking the wage $w$ as given. Since this uninformed planner lives in the economy with endogenous wages, a natural question arises as to what wages does the planner take as given, and how can the planner remain unaware of the endogeneity of wages if their tax policy can directly affect the level of wages? These questions motivate the definition of self-confirming policy equilibrium (SCPE), developed by Rothschild and Scheuer (2011):

**Definition 3** A **self-confirming policy equilibrium (SCPE)** is a feasible linear tax allocation $(t^{SC}, g)$ that solves program (9) for $w = w(g)$.

In SCPE, the observed wage level is the wage level induced by the planner’s tax policy. Therefore, the SCPE planner remains uninformed about the endogeneity of wages since their tax policy precisely induces the wages that they started with. In this and the next sections, the SCPE coincides with the solution to the standard tax problem. In Section 3, the SCPE becomes non-trivial since the standard tax rate formula will end up depending on the level of government spending.

Program (9) has a standard solution $t^{SC} = 0$, which is also an SCPE since it results in a feasible, albeit degenerate, tax allocation with no government spending $g(t^{SC}) = 0$. Since
this planner always sets \( t^{SC} = 0 \), they do not ever learn that wages increase in response to government spending, and naively believe that they are at an equilibrium. Therefore, the SCPE taxes are below the optimal tax rates when the planner is motivated only by the efficiency concerns.\(^3\)

Comparison of the SCPE and the optimal tax rates, and their connection to effort maximizing tax rate, gives rise to the more general proposition.

**Proposition 2** In a one-type model, where preferences are given by a quasi-concave twice differentiable utility function \( u(c, l) \), with \( u_c > 0 \) and \( u_l < 0 \), the following hold:

(i) for any \( \nu^{SC} \in [0, 1] \), the SCPE tax rate \( t^{SC} \) corresponds to the tax rate \( t \in [0, 1] \) that maximizes labor supply \( l(t) \),

(ii) for \( \nu = 0 \) (no direct effect of government spending on consumption), the optimal tax rate \( t^{PO} \) corresponds to the tax rate \( t \in [0, 1] \) that maximizes labor supply \( l(t, \hat{g}(t)) \),

(iii) for \( \nu > 0 \) (direct effect of government spending on consumption), the optimal tax rate \( t^{PO} \in [0, 1] \) strictly exceeds the tax rate \( t \in [0, 1] \) that maximizes labor supply \( l(t, \hat{g}(t)) \).

By property (i) of Proposition 2, a naive planner treats the labor supply maximizing tax rate as optimal, regardless of direct consumption effects of spending. Property (ii) reveals that a sophisticated planner also treats the labor supply maximizing tax rate as optimal, provided that government spending has no direct effect on consumption. Yet, when government spending increases consumption directly, a sophisticated planner strictly prefers a tax rate that distorts the supply of labor. Note that the tax rates in parts (i) and (ii) do not coincide since planners have different beliefs about the effort maximizing tax rates, in particular \( t^{SC} = 0 \) for all \( \nu^{SC} \in [0, 1] \), while \( t^{PO}(\nu = 0) = \beta > 0 \) by Proposition 1. Together with property (iii), this result implies that \( t^{PO} > t^{SC} \) for all \( \nu, \nu^{SC} \in [0, 1] \), that is, optimal taxes are above the SCPE regardless of the optimism gap.

**Introducing heterogeneity.** This section presents our main result, Proposition 2, that the optimal taxes are above the standard taxes in a setting without heterogeneity. The next two sections study how heterogeneity in earning ability and in the effect of government spending affects our main result.

\(^3\)A careful reader might object that the result for the optimal tax rates is driven by the definition of wages as \( w(g) = w_0 g^\beta \), which requires \( t > 0 \) for \( w(g) > 0 \). Simulations in Appendix E use the alternative definition of wages as \( w(g) = w_0 (1 + g)^\beta \), and show that the optimal taxes are always above the standard taxes for sufficiently high levels of \( \beta \), provided that there is no optimism gap.
3 Redistributive motive in taxation

This section introduces heterogeneity in skill into the simple model described above. The difference in earning ability provides a motive for redistribution. The informed planner achieves redistribution via the lump-sum transfers $\nu g(t)$ and improves efficiency via government spending projects that improve earnings according to wage elasticity parameter $\beta$. Proposition 3 presents the conditions on the level of transfers and wage elasticity under which the Pareto optimal tax rates exceed SCPE tax rates.

**Heterogeneity in skill.** Suppose that agents differ in skill, low ($\theta_L$) and high ($\theta_H$), and that the mass of low-skill individuals is $\alpha \in (0, 1)$. The wages at $g = 1$ are normalized to be equal to skill, so $w_{0,i} = \theta_i$, with the wage ratio $r_0 \equiv \frac{w_{0,H}}{w_{0,L}}$. Let the individual Pareto weights $\psi_L, \psi_H \geq 0$, such that $\psi_L > \psi_H$ and $\alpha \psi_L + (1 - \alpha) \psi_H = 1$, be exogenously given. Define the total weight on low-types as $\chi \equiv \alpha \psi_L$. Then $\chi \in [0, 1]$, and $1 - \chi = 1 - \alpha \psi_L = (1 - \alpha) \psi_H$ is the total weight on high-types. As in the simple model, the planner can set a linear tax $t$ to raise government spending $g$ for investment in technology that improves earnings potential according to $w_i(g) = w_{0,i} g^\beta$, where we assume that the wage elasticity $\beta$ is the same for both types.

The solutions to the consumer problem remain the same

$$\{c^*_i(t, g; w_{0,i}), l^*_i(t, g; w_{0,i})\} = \left\{ \nu g + (1 - t)w_i(g)l^*_i(t, g; w_{0,i}), \left[ \frac{(1 - t)w_i(g)}{k} \right]^{1-\gamma} \right\}, \quad (10)$$

except that now we highlight the dependence on the initial level of wages that differ by type. Per Definition 1, a feasible linear tax allocation is an allocation $\{c^*_i(t, g; w_{0,i}), l^*_i(t, g; w_{0,i})\}$ in expression (10), and a tax policy $(t, g)$ such that the government budget balance, given by

$$g = t[\alpha w_L(g)l^*_L(t, g) + (1 - \alpha) w_H(g)l^*_H(t, g)], \quad (11)$$

and wages $w_i(g) = w_{0,i} g^\beta$ are internally consistent.

The results regarding the constraint on wage elasticity $\beta < \frac{1}{\varepsilon_{w,w}}$, and parametrization of feasible tax policy in Lemma 1 carry over to this section. To establish the connection to results in the previous section, normalize the disutility of labor relative to consumption as $k' = k \left[ \alpha + (1 - \alpha) r_0^{1-\gamma} \right]^{1-\gamma}$, and choose $w_0 = w_{0,L}$ in the budget constraint for the one-type problem in expression (4). For simplicity, write $w_L(g) = w(g; w_{0,L})$ and $l^*_L(t, g) = l^*(t, g; w_{0,L})$. Using that $\frac{w_Ll^*_H}{w_Ll^*_L} = r_0^{\gamma-1}$, the budget constraint for the simple problem in
expression (4) can be re-written as

\[
g = tw_L(g)l^*_L(t, g; k')
\]

\[
= tw_L(g)\left[\frac{(1-t)w_L(g)}{k'}\right]^{\frac{1}{k' - 1}}
\]

\[
= [\alpha + (1 - \alpha)r_0^{\frac{1}{k' - 1}}]tw_L(g)\left[\frac{(1-t)w_L(g)}{k}\right]^{\frac{1}{k' - 1}}
\]

\[
= [\alpha + (1 - \alpha)r_0^{\frac{1}{k' - 1}}]tw_L(g)l^*_L(t, g; k)
\]

\[
= t[\alpha w_L(g)l^*_L(t, g; k) + (1 - \alpha)r_0^{\frac{1}{k'}}w_L(g)l^*_L(t, g; k)]
\]

\[
= t[\alpha w_L(g)l^*_L(t, g; k) + (1 - \alpha)w_H(g)l^*_H(t, g; k)],
\]

(12)

which coincides with the budget constraint for our two-type problem in expression (11). Thus, given a budget balance requirement for the two-type problem in expression (11), we can construct a bijection to the budget balance requirement for the one-type problem in expression (4) by setting \( k' = k\left[\alpha + (1 - \alpha)r_0^{\frac{1}{k' - 1}}\right]^{1 - \gamma} \). In short, the two-type budget constraint is a multiple of a one-type budget constraint. This observation implies that that analysis from the previous section applies, so a feasible tax allocation exists if and only if \( \beta < \frac{1}{\varepsilon_{y, w}} \), and, per Lemma 1, the set of feasible tax policies can be parameterized as \( \{(t, \hat{g}(t)) : t \in [0, 1]\} \), where

\[
\hat{g}(t) = \left[\frac{(1-t)t^{\gamma-1}w_0^{\gamma}}{k'}\right]^{\frac{1}{1 - \beta \gamma - 1}}
\]

\[
= \left[\alpha + (1 - \alpha)r_0^{\frac{1}{k' - 1}}\right]^{1 - \gamma}\left(\frac{(1-t)t^{\gamma-1}w_0^{\gamma}}{k}\right)^{\frac{1}{1 - \beta \gamma - 1}}
\]

\[
= \left(\alpha + (1 - \alpha)r_0^{\frac{1}{k' - 1}}\right)tw_0^{\gamma}(t; w_0)\left(\frac{t^{\gamma-1}}{1 - \beta \gamma - 1}\right).
\]

(13)

Given this continuum of feasible tax allocations \( \{(t, \hat{g}(t)) : t \in [0, 1]\} \), a planner who is informed about the effect of government spending on wages solves

\[
\max_{t,g} \chi u_L(c^*_L, l^*_L) + (1 - \chi)u_H(c^*_H, l^*_H)
\]

s.t. \( g \leq t[\alpha w_L(g)l^*_L(t, g) + (1 - \alpha)w_H(g)l^*_H(t, g)] \).

(14)

As in Definition 2, the Pareto optimal linear tax is a feasible allocation \( (t, \hat{g}(t)) \) that solves program (14). Applying Lemma 1, and using the expression for the indirect utility \( v_i(t, g) \equiv u(c^*_i(t, g), l^*_i(t, g)) = \nu \hat{g}(t) + (1 - t)w_i(\hat{g}(t))l^*_i(t, \hat{g}(t)) - h(l^*_i(t, \hat{g}(t))) \), we can re-write problem

16
as
\[
\max_t \chi [\nu \hat{g}(t) + (1 - t)w_L l^*_L - h_L] + (1 - \chi) [\nu \hat{g}(t) + (1 - t)w_H l^*_H - h_H],
\]
where we suppress the arguments of \(w_i(\hat{g}(t)), l^*_i(t, \hat{g}(t)), \) and \(h(l^*_i(t, \hat{g}(t)))\) for simplicity. Since \(h_H = h_L = w_H l^*_H w_L l^*_L = \gamma r_0^{\gamma - 1}\), then some re-arrangement of terms in expression (15) results in
\[
\max_t \nu \hat{g}(t) + [\chi + (1 - \chi) r_0^{\gamma - 1}] [w_L l^*_L - h_L] + tw_L l^*_L [\alpha - \chi] [1 - r_0^{\gamma - 1}].
\]
(16)

This form for the objective illuminates the key components of social well-being that the planner takes into account: transfers \(\nu \hat{g}(t)\), individual utility \([\chi + (1 - \chi) r_0^{\gamma - 1}] [w_L l^*_L - h_L]\), and the value of redistribution \(tw_L l^*_L [\alpha - \chi] [1 - r_0^{\gamma - 1}]\). The last term \(tw_L l^*_L [\alpha - \chi] [1 - r_0^{\gamma - 1}]\) stems from the wedge between the mass and the social weight of low-type, and hence reflects the desire for redistribution. Under the utilitarian weights \(\chi = \alpha\), this last term vanishes, and we have the standard one-type problem as in program (7), multiplied by a constant \([\alpha + (1 - \alpha) r_0^{\gamma - 1}]\). In this case, the solution is the same as in Proposition 1 since we can construct a bijection with the one-type problem through the parameter \(k\), and the one-type optimal tax formula does not depend on \(k\). Whenever \(\chi > \alpha\) and government spending has direct effects on consumption \(\nu > 0\), the tax policy serves an extra redistributive purpose, as shown in the next lemma.

**Lemma 2** The first-order conditions for problem (16) characterize the Pareto optimal tax rate as
\[
t^{PO}(\chi) = \left(\frac{\nu - (\beta + 1)\mu}{2(\nu \varepsilon_{,w} - \mu)}\right) + \frac{\sqrt{(\nu - (1 - \beta)\mu)^2 + 4\beta\mu\nu\varepsilon_{i,w}}}{2(\nu \varepsilon_{,w} - \mu)} > 0,
\]
(17)

where \(\mu(\chi) = \frac{\chi + (1 - \chi) r_0^{\gamma - 1}}{\alpha + (1 - \alpha) r_0^{\gamma - 1}}\), with the properties that

(i) \(t^{PO} < t^{Laffer}\),

(ii) \(t^{PO}(\nu = 0) = t^{effort} = \beta\),

(iii) \(t^{PO}\) increases in \(\nu\),

(iv) \(t^{PO}\) increases in \(\beta\),

(v) \(t^{PO}\) increases in \(\chi\), provided that \(\nu > 0\).
Note that the only difference between Lemma 2 and Proposition 1 is that Lemma 2 includes a new term $\mu$—that captures the taste for redistribution—in the optimal tax rate formula, and lists an additional result, property (v). Property (v) establishes that the optimal tax rate increases in the redistributive motive. Whenever $\chi > \alpha$ and government spending directly increases consumption $\nu > 0$, we expect the optimal tax to exceed the one motivated by efficiency alone. Property (v) further reflects that redistribution occurs solely via effective lump-sum transfers. With no direct consumption effects of government spending $\nu = 0$, as in property (ii), the optimal tax rate reduces to effort maximizing tax rate $t^{PO} = \beta$, regardless of the desire for redistribution. Together, properties (ii) and (v) reveal that $\beta$ only affects the productivity effects of spending, while $\nu$ only affects the distribution consequences of spending.

Having characterized the Pareto optimal policy, we proceed to study how it compares with the SCPE policy. A social planner who is unaware of the effect of government spending on wages solves

$$
\max_{t,g} \chi u_L(c^*_L, l^*_L) + (1 - \chi)u_H(c^*_H, l^*_H)
\text{s.t. } g \leq t[\alpha w_L l^*_L(t; w_L) + (1 - \alpha) w_H l^*_H(t; w_H)],
$$

(18)

with effective lump-sum transfers $T(g) = \nu^{SC} g$, as in Section 2. Note that this problem appears identical to the informed planner’s problem in expression (14), except that the uninformed planner naively believes that the wages are fixed and that the only benefit of government spending lies in its direct effects on consumption.

**Lemma 3** The first-order conditions for problem (18) characterize the SCPE tax rate as

$$
t^{SC}(\chi) = \frac{1}{1 + \frac{\nu^{SC} \varepsilon_{l,w}}{\nu^{SC} - \mu}},
$$

(19)

where $\mu(\chi) = \frac{\chi(1-\chi)r_0^\gamma}{\alpha(1-\alpha)r_0^\gamma}$, with the properties that

(i) $t^{SC} > 0$ iff $\chi > \alpha$ and $\nu^{SC} > \mu(\chi)$,

(ii) $t^{SC}$ increases in $\nu^{SC}$, $\chi$, and $r_0$, provided that $\nu^{SC} > \mu(\chi)$.

(iii) $t^{SC}$ decreases in $\varepsilon_{l,w}$, provided that $\nu^{SC} > \mu(\chi)$.

To confirm that the tax rate described in Lemma 3 is indeed an SCPE, we check if the wage distribution $w(\hat{g}(t^{SC}))$ remains the same under the tax policy $(t^{SC}, \hat{g}(t^{SC}))$. Since
the tax rate in equation (19) depends only on relative wages, the tax policy \((t^{SC}, \hat{g}(t^{SC}))\) is consistent with the wage distribution \(w(\hat{g}(t^{SC}))\) that it induces. So \((t^{SC}, \hat{g}(t^{SC}))\) is the unique SCPE. The properties of the SCPE tax rate discussed in Lemma 3 are standard. A redistributive motive \(\chi > \alpha\), and government spending that substantially increases direct consumption \(\nu^{SC} > \mu\) are the necessary and sufficient conditions for the SCPE tax rate to be positive. This feature differs from the optimal tax rate, which is positive for all \(\nu > 0\), because the naive planner does not know about the endogeneity of wages and therefore finds the tax system wasteful unless a sufficient amount of spending enters consumption directly. Other than the condition on effective lump-sum transfers \(\nu^{SC} > \mu\), the SCPE and Pareto optimal tax rates share similar properties. Both tax rates increase in the total weight on low-types \(\chi\) and in the wage ratio \(r_0\) for redistributive purposes, and decrease in wage elasticity of labor supply \(\varepsilon_{l,w}\) for efficiency purposes.

Lemma 3 establishes that the SCPE tax rate can be positive in the presence of the redistributive motive. Contrast this result with the SCPE tax rate from Section 2, which remained at zero for all levels of \(\nu^{SC}\) due to homogeneity in consumers. Therefore, the comparison between the Pareto optimal and SCPE tax rates now depends on the effect of government spending on consumption \(\nu\) and government spending elasticity of wage \(\beta\). The following proposition formalizes the relationship between the Pareto optimal and SCPE tax rates.

**Proposition 3** The comparison of the Pareto optimal and the SCPE tax rates depends on the optimism gap—the divergence between the true and naive social planner’s view about the effect of government spending on consumption \(\nu\) — and the level of wage elasticity \(\beta\):

(i) (no optimism gap) \(t^{PO}(\nu) > t^{SC}(\nu^{SC} = \nu)\) for all \(\nu \in [0, 1]\),

(ii) (moderate optimism gap) For all \(\beta > 0\), there is \(\nu \geq 0\) such that \(t^{SC}(\nu^{SC} = 1) < t^{PO}(\nu)\) for all \(\nu \geq \nu\),

(iii) (extreme optimism gap) There exists \(\beta \in (0, \frac{1}{\varepsilon_{l,w}})\) such that for all \(\beta \geq \beta\), we have \(t^{SC}(\nu^{SC} = 1) < t^{PO}(\beta, \nu = 0)\).

To illustrate Proposition 3, Figure 5 plots Pareto (in blue) and SCPE (in red) tax rates against \(\nu\). Figure 5(a) illustrates the case when the naive planner has correct beliefs about the usefulness of government spending for lump-sum transfers, \(\nu^{SC} = \nu\). With no optimism gap, the optimal tax rate strictly exceeds the SCPE tax rate since the sophisticated planner recognizes the extra benefit of spending on wages and sets a positive tax rate even if \(\nu = 0\). In contrast, the naive planner views government spending as wasteful when \(\nu < \mu\),
and does not undertake a project unless a considerable proportion of spending enters individual consumption directly. Under the optimism gap $\nu^{SC} = 1$, there exist levels of $\nu$ with $t^{SC}(\nu^{SC} = 1) > t^{PO}(\nu)$, as shown in Figure 5(b). The naive planner believes that government spending translates one-to-one into consumption, and sets a positive tax rate that strictly exceeds the optimal tax rate on $[0, \nu)$. Provided that $\beta > 0$, the Pareto optimal tax rate catches up with the SCPE tax rate at some $\nu \in [0, 1]$ and remains above the SCPE afterwards.

![Figure 5: Pareto optima vs. SCPE (illustration of Proposition 3)](image)

Separability of efficiency and redistribution. The notable feature of the model above lies in homogeneity of wage elasticity $\beta$, so the productivity motive does not interact with the motive for redistribution. This analysis applies when government spending increases productivity of both types proportionately, without affecting the relative wages. Most of the government spending projects, however, have varying effects on individuals of different earning ability. For example, spending on universal preschool increases adult earnings for low-income groups at a higher rate. The next section expands our analysis to government spending projects with heterogeneous effects.

## 4 Efficiency-redistribution trade-off

This section introduces heterogeneity in the effect of government spending to the model in Section 3. In this model, government spending can affect efficiency and the desire for redistribution at the same time (for example, investments in projects that yield greater benefits for high-types increase efficiency and the desire to redistribute). While this section lacks analytical results, we provide numerical simulations and a policy-relevant example.
**Heterogeneity in skill and the effect of government spending.** Suppose that the government spending elasticity of wage varies by type, $\beta_i \in [0, 1]$, with the wedge between the wage elasticities denoted as $\Delta \equiv \beta_H - \beta_L$. This difference in wage elasticities implies that government spending affects relative wages, and hence the desire for redistribution. Figure 16 in Appendix I illustrates that the order of wages may switch at the optimum under exogenous social weights. To address the concern of endogenously determined relative wages, we consider two solutions: use endogenous social weights, or proceed with exogenous social welfare weights and make sure to check that wages for the type with lower welfare weights is indeed the lower wage type at the optimum for the parameters that we assume. We choose the second solution because it provides more tractability and enables us to compare the results across sections. The numerical simulation presented in Figure 17 in Appendix I confirms that wages remain exogenously ordered for the parameters assumed in this section. Appendix J reformulates the problem with endogenous social weights, and numerically checks that results with endogenous weights remain qualitatively similar.

The solution to the consumer problem $\{c^*_i(t, g; w_{0,i}), l^*_i(t, g; w_{0,i})\}$ remains the same as in expression (10) from Section 3. By Definition 1, a feasible linear tax allocation is an allocation $\{c^*_i(t, g; w_{0,i}), l^*_i(t, g; w_{0,i})\}$ from expression (10), and a tax policy $(t, g)$ such that the government budget balance given by expression (11), and wages $w_L(g) = w_{0,L} g^{\beta}, \ w_H(g) = w_{0,H} g^{\beta + \Delta}$ are internally consistent.

The planner who is aware of the effect of government spending on wage distribution solves

$$\max_{t,g} \chi u_L(c^*_L, l^*_L) + (1 - \chi) u_H(c^*_H, l^*_H)$$

s.t. \[ t(\alpha w_L(g)l^*_L(t, w_L(g)) + (1 - \alpha) w_H(g)l^*_H(t, w_H(g)) \leq g. \quad (20) \]

Per Definition 2, the socially optimal linear tax is a feasible allocation $(t, g)$ that solves program (20). While there is no explicit functional form for $\hat{g}(t)$ due to differences in exponents in expressions for wages $w_i(g)$, the budget constraint in program (20) implicitly describes the set of feasible allocations as

$$\{(t, g) : t \in [0, 1] \text{ and } g = t(\alpha w_L(g)l^*_L(t, w_L(g)) + (1 - \alpha) w_H(g)l^*_H(t, w_H(g))\}. \quad (21)$$

Figure 6 plots the set of feasible allocations (in blue) and the social indifference curves (in red). Notably, the social indifference curves are no longer convex, they become concave at the higher levels of the tax rate. Otherwise, the analysis from the previous sections holds, and the Pareto optimum occurs at the tangency of the government budget curve and the social indifference curve (this plot assumes that $\beta_L = 0.1$ and $\beta_H = 0.2$).
Figure 6: Pareto optimum

An SCPE is a feasible \((t^{SC}, g^{SC})\) pair satisfying the following equation

\[
(t^{SC}, g^{SC}) = \arg \max_{t,g} \chi u_L(c^*_L, l^*_L) + (1 - \chi) u_H(c^*_H, l^*_H) \\
s.t. \quad g \leq t[\alpha w_L(g^{SC})]l^*_L(t, w_L) + (1 - \alpha) w_H(g^{SC})l^*_H(t, w_H)],
\]

(22)

where \(T = \nu^{SC} g\) for some \(\nu^{SC} \in [0,1]\). To understand the importance of the restriction of wages to \(w_i(g^{SC})\), consider the problem without this requirement

\[
\max_{t,g} \chi u_L(c^*_L, l^*_L) + (1 - \chi) u_H(c^*_H, l^*_H) \\
s.t. \quad g \leq t[\alpha w_L(g^{SC})]l^*_L(t, w_L) + (1 - \alpha) w_H(g^{SC})l^*_H(t, w_H)].
\]

(23)

Problem (23) resembles problem (18) from the previous section, and has a solution

\[
t = \frac{1}{1 + \frac{\nu^{SC}}{(\nu^{SC} - \mu)}},
\]

(24)

where \(\mu = \frac{\chi + (1 - \chi)r^\gamma}{\alpha + (1 - \alpha)r^\gamma}\). Note that the tax formula in expression (24) depends on relative wages \(r = r_0 g^\delta\), which change in response to the level of spending \(g\), and thus does not necessarily describe an SCPE tax rate. In other words, given the current wage distribution, the planner sets the tax rate as above, but that tax rate can produce a different wage distribution, and hence the level of government spending. The SCPE needs to solve the fixed point problem for the wage distribution, where the observed wage distribution implies an SCPE tax policy that indeed induces the wage distribution that the planner started with. The formulation in expression (22) foresees this issue and ensures that the wage distribution
appears fixed under the SCPE policy by imposing a restriction that \( w_i(g^{SC}) \).

Appendix K outlines the numerical method that solves program (22), but we provide the intuition for the solution here. If a naive planner finds themselves at the Pareto optimum \((t^{PO}, g(t^{PO}))\), they observe the wage distribution \( w_L(g(t^{PO})), w_H(g(t^{PO})) \). Since the naive planner is not aware of the endogeneity of wages, they believe that the budget constraint is given by the dashed blue curve in Figure 7(a). The notional budget curve (in dashed blue) through the optimum is flatter than the true budget curve since the uninformed planner is not aware that a decrease in the tax rate implies a decrease in wages and government spending. The notional indifference curve (in dashed red) through the optimum is steeper than the true indifference curve since the naive planner does not know about the indirect benefits of government spending, and hence values it less than the sophisticated planner. Figure 7(b) zooms into the region around the optimum to highlight that the slopes of the notional budget curve and indifference curve are not tangent to each other, so the naive planner believes that the tax policy \((t^{PO}, g(t^{PO}))\) is not optimal. Since the slopes of the notional budget and indifference curves are respectively flatter and steeper than the true curves, the naive planner believes that the optimum is located to the left of the original optimum, i.e. that they can improve social welfare by decreasing taxes. Yet, when the naive planner lowers the tax rate, they discover that their budget no longer balances since the wage distribution has changed. So the naive planner continues to iterate in this manner until the equilibrium predicted by the notional budget and social indifference curves yields a feasible tax policy (located on the true budget curve).

![Figure 7](image_url)

(a) Notional budget and social indifference curves  
(b) Zoom-in around the Pareto optimum

Figure 7: Informed (solid curves) vs. uninformed (dashed curves) analyses

Figure 8 shows the SCPE tax policy that emerges as a result of naive planner’s learning
described in the previous paragraph. At \((t^{SC}, g(t^{SC}))\), the naive planner observes the wage distribution \(w_L(g(t^{SC})), w_H(g(t^{SC}))\). They believe that the budget and the social indifference curves are given by the dashed blue and red curves, respectively. Upon observing the wages \(w_L(g(t^{SC})), w_H(g(t^{SC}))\), they solve program (22), which yields a tax policy \((t^{SC}, g(t^{SC}))\), which indeed induces the wage distribution that they started with \(w_L(g(t^{SC})), w_H(g(t^{SC}))\). Therefore, the naive planner sets a tax policy \((t^{SC}, g(t^{SC}))\), and they believe that this tax policy is optimal, since, for them, the wages remain fixed, so they have no incentive to change their tax policy and hence do not ever learn that a change in their tax policy could result in a wage distribution different from the one they started with.

![Figure 8: SCPE vs. Pareto optimum](image)

In the analysis above, the SCPE tax rate is below the socially optimal tax rate. Under different assumptions on wage elasticity parameters and the level of transfers, it is possible to have cases where the SCPE tax rate is above the Pareto optimal tax rate (Figure 9(a)), or where the Pareto optimal tax is always above the SCPE tax rate (Figure 9(b)). In Figure 9(a), we set \(\beta_L = 0, \beta_H = 0.1\), so the government spending only increases the wages of high-types. In Figure 9(b), we set \(\beta_L = 0.35, \beta_H = 0.25\), so the wage elasticity of low-types is greater than the wage elasticity of high types.
Distribution and efficiency goals are not separable: an example. To study the linkage between efficiency and distribution, we compare the optimal and SCPE tax rates under the assumption of Rawlsian weights. Suppose that the planner only cares about the well-being of low-types, so $\chi = 1$, and that the government spending project only improves the wages of high-types, so $\beta_L = 0$.

Consider the naive planner with optimistic beliefs $\nu^{SC} = 1$. Motivated by redistribution through lump-sum transfers, this planner sets a positive tax rate. The naive planner is not aware that by setting a positive tax rate, they increase spending that benefits the high-types, and hence increases socioeconomic inequality. We would like to know how the awareness about the distributional effects of spending affects the tax policy. The sophisticated planner solves

$$\max_{t,g} u_L(c_L^*, l_L^*) \quad \text{s.t.} \quad g \leq t[\alpha w_{0,L} l_L^*(t; w_{0,L}) + (1 - \alpha) w_H(g) l_H^*(t, w_H(g))] \quad (25)$$

Figure 10 plots the optimal tax rates (in blue) against the SCPE tax rate (in red), and reveals that the optimal tax rates can be positive, and even exceed the SCPE tax rates, as long as a significant proportion of government spending enters consumption directly. While the informed planner does not care about the direct effect of government spending on well-being of high-types, they are aware that increasing the productivity of high-types results in greater output, and hence more resources available for effective lump-sum transfers to the low-types. Therefore, for sufficiently high levels of transfers, the optimal tax rate exceeds the SCPE tax rate. This example illustrates that the efficiency and redistribution motives in taxation can be interlinked, that is, “the way we slice the pie determines the size of the
5 Optimal spending across projects

This section extends the preceding analysis to a setting with multiple government spending projects. We consider the optimal allocation of budget across projects with homogeneous effects on wages, as in Section 3, and heterogeneous effects on wages, as in Section 4. For simplicity, we assume that the government faces a decision between two projects, but the following analysis can be extended to any (finite) number of government projects with constant elasticity of substitution.

**Homogeneous effects on wages.** Suppose that the planner uses the tax revenue to invest in projects \( g_1 \) and \( g_2 \), where wages are given by the constant elasticity of substitution (CES) government spending aggregator

\[
  w_i(g_1, g_2) = \left( \phi_{1,i} g_1^\rho + \phi_{2,i} g_2^\rho \right)^{\frac{\beta}{\rho}},
\]

with share parameters \( \phi_{1,i}, \phi_{2,i} \in (0, 1) \), elasticity of substitution \( \sigma \equiv \frac{1}{1-\rho} > 0 \), and homogeneity of degree \( \beta \in (0, 1/\varepsilon_{y,w}) \).

For now, we suppose that government spending has homogeneous effects on individuals, i.e. there is \( \zeta \in \mathbb{R}_{>0} \) such that \( \frac{\phi_{2,i}}{\phi_{1,i}} = \zeta \) for all \( i \in \{L, H\} \). The proportional increase in wages implies that the government spending project \( g_j \) elasticity of wage, given by

\[
  \frac{d \ln(w_i)}{d \ln(g_j)} = \frac{\beta g_j^\rho}{g_1^\rho + \zeta g_2^\rho},
\]

is constant across individuals, and that the relative wages are fixed at

\[
  r_0 \equiv \frac{w_H(g_1, g_2)}{w_L(g_1, g_2)} = \left( \frac{\phi_{1,H}}{\phi_{1,L}} \right)^{\frac{\beta}{\rho}}.
\]

Otherwise, we keep the same setup as before, with the

\[\text{Assumption 1 applies to this section.}\]
mass of low-types $\alpha \in (0, 1)$, total weight on low-types $\chi \in [0, 1]$, and effective lump-sum transfers $T(g; \nu) = \nu g$ for some $\nu \in [0, 1]$.

The planner sets a linear tax rate $t \in [0, 1]$ and divides the collected tax revenue $g$ between spending on projects $g_1$ and $g_2$. Hence, the planner solves

$$\max_{t, g_1, g_2} \chi u_L(c^*_L, l^*_L) + (1 - \chi) u_H(c^*_H, l^*_H)$$

s.t. $g_1 + g_2 \leq t[\alpha w_L(g_1, g_2)l^*_L(t, g_1, g_2) + (1 - \alpha) w_H(g_1, g_2)l^*_H(t, g_1, g_2)], \quad (27)$

where $c^*_L(t, g_1, g_2), l^*_L(t, g_1, g_2)$ solve the respective consumer problems.

As formulated in expression (27), the planner simultaneously chooses the optimal levels of taxes $t$ and investment in each of the projects, $g_1$ and $g_2$. To separate the optimal level of taxes (and overall spending $g$) from the optimal allocation of budget across projects, let $\eta \in [0, 1]$ denote the proportion of government spending $g_1$ on project $g_1$. Then $g_1 = \eta g$ and $g_2 = (1 - \eta)g$, and the specification for wages in expression (26) simplifies to

$$w_i(g_1, g_2) = w_i(g; \eta) = g^\beta [\phi_{1,i} \eta^\rho + \phi_{2,i} (1 - \eta)^\rho]^\frac{\beta}{\rho} = g^\beta \eta^\beta \phi_{1,i}^\beta \left[1 + \zeta \left(\frac{1 - \eta}{\eta}\right)^\rho\right]^\frac{\beta}{\rho}. \quad (28)$$

So we can recast problem (27) as

$$\max_{\eta} \left\{ \max_{t, g} \chi u_L(c^*_L, l^*_L) + (1 - \chi) u_H(c^*_H, l^*_H), \right. \right.$$

s.t. $g \leq t[\alpha w_L(g; \eta)l^*_L(t, g; \eta) + (1 - \alpha) w_H(g; \eta)l^*_H(t, g; \eta)], \quad (29)$

where the inner problem represents the intensive margin problem of choosing the overall level of government spending $g$ and tax rates $t$—as in problem (14) with one type of government spending—and the outer problem represents the extensive margin problem of choosing the spending ratio $\eta$ across projects.

To solve the inner problem, we follow the same steps as in Section 3. Per Definition 1, for any $\eta \in [0, 1]$, a feasible linear tax allocation is an allocation $\{c^*_L(t, g; \eta), l^*_L(t, g; \eta)\}$ from expression (10), and a tax policy $(t, g; \eta)$ that balances the government budget given by the constraint in problem (29). Under Assumption 1, Lemma 1 holds, and we can parametrize the set of feasible tax policies as $\{(t, \hat{g}(t); \eta) : t \in [0, 1]\}$, where

$$\hat{g}(t; \eta) = \left[ tw_L(1; \eta)l^*(t; w_L(1; \eta)) [\alpha + (1 - \alpha)r_0^{\frac{\gamma}{\beta - 1}}]\right]^{-\frac{\gamma - 1}{\beta - 1}}. \quad (30)$$
Therefore, we can re-write the planner’s inner problem as

\[
\max_t \nu \hat{g}(t) + \chi [(1 - t) w_L(\hat{g}(t)) l_L(t) - h(l_L)] + (1 - \chi) [(1 - t) w_H(\hat{g}(t)) l_H(t) - h(l_H)]
\]

where the proportional increase in wages enables us to reduce the two-type problem to one-type problem through the term \(\chi + (1 - \chi) r_{0}^{\gamma - 1}\), which keeps track of the social weight and the relative wages. The first-order conditions for problem (31) characterize the optimal tax rate as identical to that in Lemma 2:

\[
t^{PO} = \frac{(\nu - (\beta + 1)\mu) + \sqrt{(\nu - (1 - \beta)\mu)^2 + 4\beta\mu\nu\epsilon_{l,w}}}{2(\nu\epsilon_{y,w} - \mu)} \geq 0,
\]

where \(\mu = \frac{\chi + (1 - \chi) r_{0}^{\gamma - 1}}{\alpha + (1 - \alpha) r_{0}^{\gamma - 1}}\). Note that the optimal tax rate in expression (32) does not depend on the allocation of spending across projects \(\eta\), or on the relative productivity of projects \(\zeta\). So the optimal tax policy \((t^{PO}, \hat{g}(t^{PO}); \eta) = (t^{PO}, \hat{g}(t^{PO}))\) is the same for all \(\eta \in [0, 1]\), and the outer problem in expression (29) becomes simply

\[
\max_\eta \nu g^{PO} + [(1 - t^{PO}) w_L(\eta; g^{PO}) l_L(\eta; t^{PO}) - h(l_L)] [\chi + (1 - \chi) r_{0}^{\gamma - 1}],
\]

where \(g^{PO} = \hat{g}(t^{PO})\).\(^5\)

**Proposition 4** The first-order conditions for problem (33) characterize the optimal ratio of spending as

\[
\frac{\eta^{PO}}{1 - \eta^{PO}} = \zeta^{-\sigma},
\]

with the properties that

(i) \(\eta^{PO} = \arg\max_\eta w_i(\eta; \hat{g}(t^{PO}))\),

(ii) \(\eta^{PO} = 1_{\zeta < 1}\) for \(\sigma \to \infty\),

(iii) \(\frac{\eta^{PO}}{1 - \eta^{PO}} = 1\) for \(\sigma \to 0\).

\(^5\)While government spending \(g(\eta; t^{PO})\) also depends on \(\eta\) through \(w_L(1, \eta)\) by expression (30), it is locally independent of \(\eta\) at the optimum \(\eta^{PO}\) because \(\eta^{PO}\) maximizes wages—and hence \(w_L(1, \eta)\)—by problem (33). In other words, at the optimum \(\eta^{PO}\), a local change in \(\eta\) has zero effect on the budget constraint (if \(g\) adjusts automatically per expression (30)).
The properties described in Proposition 4 match our expectations. Property (i) reflects that \( \eta \) affects individual well-being only through its effect on wages \( w_i(\eta; \hat{g}(tPO)) \), as suggested by expression (33). The homogeneity in effects of spending \( \zeta \) implies that both types \( i \in \{L, H\} \) share the same wage maximizing \( \eta^{PO} \). Property (ii) tells that, when the two types of government spending are perfect substitutes, it is optimal to invest all of the tax revenue into a project with the higher share parameter. Property (iii) tells that, when the two types of government spending are perfect complements, it is optimal to invest equal proportions of tax revenue into both projects, regardless of the value of the share parameters. Proposition 4 further reveals that the optimal allocation of budget does not depend on the desire for redistribution. The optimal allocation solely depends on productivity effects of the government spending projects, captured by the relative productivity of projects \( \zeta \) and the elasticity of substitution \( \sigma \), because projects have no effect on distribution. This result aligns with our choice to model the projects as having no effect on relative wages, and connects to the discussion at the end of Section 3 regarding the separability of efficiency and redistribution motives when spending has homogeneous effects. The next part of this section compares how the results above change when government spending projects have heterogeneous effects on wages.

**Heterogeneous effects on wages.** We keep the same setup as above, except that now government spending projects have varying effects on individuals, i.e. \( \phi_{2,L} \neq \phi_{2,H} \). Let \( \zeta_i \equiv \frac{\phi_{2,i}}{\phi_{1,i}} \). Then we can write the expression for wages as

\[
 w_i(g, \eta) = g^\beta [\phi_{1,i} \eta^\rho + \phi_{2,i}(1 - \eta)]^{\frac{\beta}{\rho}} = g^\beta \eta^\beta \phi_{1,i}^\beta \left[ 1 + \zeta_i \left( \frac{1 - \eta}{\eta} \right)^{\rho} \right]^{\frac{\beta}{\rho}}.
\]

So the government spending project \( g_j \) elasticity of wage for type \( i \), given by \( \frac{d \ln(w_i)}{d \ln(g_j)} = \frac{\beta g_j^\rho}{g_j^\rho + \zeta_1 g_j^\rho} \), depends on type \( i \). The heterogeneity in wage elasticity further implies that relative wages vary with policy parameters

\[
 \frac{w_H(g, \eta)}{w_L(g, \eta)} = \frac{\phi_{1,H}^\beta \left[ 1 + \zeta_H \left( \frac{1 - \eta}{\eta} \right)^{\rho} \right]^{\frac{\beta}{\rho}}}{\phi_{1,L}^\beta \left[ 1 + \zeta_L \left( \frac{1 - \eta}{\eta} \right)^{\rho} \right]^{\frac{\beta}{\rho}}} = r_0[\Delta(\eta)]^{\frac{\beta}{\rho}},
\]

where \( \Delta(\eta) \equiv \frac{1 + \zeta_H \left( \frac{1 - \eta}{\eta} \right)^{\rho}}{1 + \zeta_L \left( \frac{1 - \eta}{\eta} \right)^{\rho}} \). The wage ratio depends on an endogenous parameter \( \eta \), so assuming exogenous social weights only makes sense if the low-type has a lower-wage at the optimum. To address this concern, we adopt the following assumption.
**Assumption 2** The share parameters $\phi_{j,i} \in (0, 1)$ satisfy $\phi_{j,H} > \phi_{j,L}$ for all $j \in \{1, 2\}$.

Assumption 2 ensures that the high-types always have greater absolute advantage, mirroring the assumption on the wage ratio $r_0 > 1$ from Section 3. Note that this assumption does not restrict the values of relative advantage from spending projects $\zeta_i$. Given Assumption 2, we can proceed with exogenous social weights. The planner solves

$$
\max_{\eta} \left\{ \max_{t,g} \chi u_L(c^*_L, l^*_t) + (1 - \chi)u_H(c^*_H, l^*_H),
\text{s.t. } g \leq \{aw_L(g; \eta)l^*_L(t, g; \eta) + (1 - \alpha)w_H(g; \eta)l^*_H(t, g; \eta)\},
\right. \tag{34}
$$

where $c^*_i(t, g; \eta), l^*_t(t, g; \eta)$ solve the respective consumer problems. Using that $\frac{w_t l^*_t}{w_L l^*_L} = r_0^{\gamma - \eta} \Delta(\eta)[\alpha^{\gamma}]^{\frac{\beta}{\gamma - \eta - 1}}$, we can re-write the budget constraint from problem (34) as

$$
g = tw_L(g; \eta)l^*_L(t, g; \eta)[\alpha + (1 - \alpha)r_0^{\gamma - \eta} \Delta(\eta)[\alpha^{\gamma}]^{\frac{\beta}{\gamma - \eta - 1}}]. \tag{35}
$$

Except for the difference in the wage ratio, the budget constraint above is the same as in the model with homogeneous effects of spending. Thus, given $\eta \in [0, 1]$, Lemma 1 implies that the set of feasible tax policies is described by $\{(t, \hat{g}(t); \eta) : t \in [0, 1]\}$, where

$$
\hat{g}(t; \eta) = \left[ tw(1; \eta)l^*(t; w(1; \eta))[\alpha + (1 - \alpha)r_0^{\gamma - \eta} \Delta(\eta)[\alpha^{\gamma}]^{\frac{\beta}{\gamma - \eta - 1}}]\right]^{\frac{\gamma - 1}{\gamma - \beta - 1}}. \tag{36}
$$

So the planner’s inner problem simplifies to resemble expression (31):

$$
\max_t \nu \hat{g}(t) + [(1 - t)w_L(\hat{g}(t))l_L(t; w_L) - h(l_L)][\chi + (1 - \chi)r_0^{\gamma - \eta} \Delta(\eta)[\alpha^{\gamma}]^{\frac{\beta}{\gamma - \eta - 1}}]. \tag{37}
$$

The first-order conditions for problem (37) characterize the **Pareto optimal tax rate** as

$$
t^{PO} = \frac{(\nu - (\beta + 1)\mu) + \sqrt{(\nu - (1 - \beta)\mu)^2 + 4\beta\mu\nu l_{L, w}}}{2(\nu\varepsilon_{y, w} - \mu)} \geq 0,
$$

where $\mu = \frac{\gamma + (1 - \gamma)(\alpha + (1 - \alpha)r_0^{\gamma - \eta} \Delta(\eta)[\alpha^{\gamma}]^{\frac{\beta}{\gamma - \eta - 1}})}{\alpha + (1 - \gamma)(\alpha + (1 - \alpha)r_0^{\gamma - \eta} \Delta(\eta)[\alpha^{\gamma}]^{\frac{\beta}{\gamma - \eta - 1}})}$. In contrast to the case with homogeneous effects of spending, the optimal tax rate now depends on the allocation of spending across the projects.

While solving for $\eta^{PO}$ is complicated, the following lemma characterizes the interval for the plausible values of optimal spending ratio.

**Lemma 4** Suppose that $\eta^*_i = \arg\max_{\eta} w_i(\eta; \hat{g}(t^{PO}), \zeta_i)$. Then $\eta^{PO} \in [\underline{\eta}, \bar{\eta}]$, where $\underline{\eta} = \min\{\eta^*_i : i \in \{L, H\}\}$ and $\bar{\eta} = \max\{\eta^*_i : i \in \{L, H\}\}$. In particular

(i) $\zeta_L > \zeta_H$ iff $\eta^*_L < \eta^*_H$.  

30
(ii) If $\zeta_L \neq \zeta_H$ and $\chi \in (0,1)$, then $\eta^{PO} \in (\bar{\eta}, \tilde{\eta})$.

The intuition for Lemma 4 is straightforward. Suppose that $\eta^*_L < \eta^*_H$. Then, for both types, the wages $w_i(\eta;\hat{g}(t^{PO}), \zeta_i)$ are decreasing for $\eta > \eta^*_H$, and increasing for $\eta < \eta^*_L$. So the optimal $\eta$ must lie in the interval $[\eta^*_L, \eta^*_H]$. Note that the low-types prefer $\eta^*_L < \eta^*_H$ if and only if government spending $g_2$ benefits them more than the high-types, i.e. $\zeta_L > \zeta_H$. Moreover, as long as the planner cares about both types’ well-being $\chi \in (0,1)$, the optimal spending ratio would lie in the open interval $(\eta^*_L, \eta^*_H)$, so that there is no type whose preferences determine the social preferences. Lemma 4 significantly simplifies the analysis of local comparative statics of $\eta^{PO}$ in response to changes in the share parameters $\phi_{j,i}$ and in social weight $\chi$, and gives rise to the following proposition.

**Proposition 5** Let $\chi \in (0,1)$ and $\zeta_i > 0$. Suppose that the optimal spending ratio is given by $\eta^{PO}(t; \zeta_H, \zeta_L, \chi)$. Then

1. $\frac{\partial \hat{\eta}^{PO}(t^{PO} \zeta_i \chi)}{\partial \zeta_i} < 0$ for all $i \in \{L, H\}$,
2. $\frac{\partial \hat{\eta}^{PO}(t^{PO} \zeta_L \chi)}{\partial \chi} > 0$ if and only if $\zeta_L < \zeta_H$.

Appendix N formally proves Proposition 5, but the intuition for these results is simple and matches our expectations. Consider a setting with homogeneous effects of spending $\zeta_L = \zeta_H = \bar{\zeta}$ and the corresponding optimal tax rate $t^{PO}(\bar{\zeta}, \bar{\zeta}, \chi)$. Suppose that wages of one of the types become more dependent on government project $g_2$, so $\zeta_i > \bar{\zeta}$ for some $i \in \{L, H\}$. Property (i) implies that the planner then re-allocates a greater share of budget to $g_2$, so $\eta^{PO}$, the proportion of spending on $g_1$, has to decrease. Since this result holds for all types and does not depend on the social weights, property (i) reveals that changes in the share parameters only influence the efficiency motive in spending.

Property (ii) states, informally, that putting higher welfare weight on the low-skill type will lead the planner to re-allocate a larger fraction of spending on the project that provides greater relative gains to the low-types. This result is stated and proved only for the case where the optimal tax rate $t^{PO}$ is held fixed, but simulations in Figure 11 suggest that these comparative statics would hold if we allowed the optimal tax rate to adjust with the social welfare weight as well (that is, it suggests that we could replace the partial derivative in property (ii) with a total derivative as in property (i)—but we have not yet been able to prove this result analytically).
The results of this section have two important policy implications. First, without the information about the distributional consequences of spending projects, the planner may mis-allocate the budget across projects and unintentionally exacerbate socioeconomic inequality. For example, without the awareness of the distributional consequences, the planner views the optimal spending ratio to be represented by the light blue curve (utilitarian weights) in Figure 11. However, as soon as they learn that spending on $g_1$ increases the wage gap on the interval $\zeta_L > 1 = \zeta_H$, they choose to re-allocate the budget to the project with greater benefit to low-types, and hence select a strictly lower spending on $g_1$, represented by the dark blue curves ($\chi > \alpha$). Second, Proposition 5 helps to re-interpret the example with Rawlsian weights from Section 4. In that example, the informed planner chose to invest in the project that only benefited the high-types for the purpose of redistributing to low-types through effective lump-sum transfers. That result was driven by the restriction of the model to one type of government spending. With access to multiple government projects, the spending ratio $\eta$ itself becomes a tool for redistribution, and enables the informed planner to adjust their allocation of budget according to their redistributive goals.
6 Conclusion and discussion

This paper takes a step toward incorporating the effects of government spending on wage distribution in an optimal linear taxation framework. We provide natural conditions under which the optimal tax rates with endogenous wages are strictly higher than the standard tax rates. This desire for higher optimal tax rates stems from the productivity effects of government spending. Improved productivity increases social well-being, and also expands the resources available in the economy. Notably, even if government spending benefits only the upper-end of the society, it remains strictly optimal for the planner to invest in the project so as to increase the tax revenue available for redistribution.

The optimal allocation of budget across projects depends on the government spending elasticity of wage and the elasticity of substitution between the projects. When both types share the same wage elasticities, the spending ratio does not affect the distribution of wages, so the optimal allocation aims to maximize efficiency, and thus depends only on the productivity effects of spending and the elasticity of substitution. When wage elasticities differ across types, the optimal policy distorts away from the efficient allocation towards the project with greater relative benefit to the low-skill type. Together, these findings suggest that the distributional consequences of the government projects affect the overall level of productivity—and hence the optimal levels of spending and taxation—in the economy.
References


Appendix

A Proof of Assumption 1

Proof. Fix some \( t \in (0, 1) \), and consider a map from government spending onto itself \( \phi(g) = \hat{g} \circ w \), where \( w(g) = w_0 g^\beta \) and \( \hat{g}(w) = tw^*(w; t) \). So \( \phi(g) = tw(g)l^*(w(g); t) = t\left[ \frac{1-t}{k} \right]^{\frac{1}{\gamma-1}} w_0^\gamma g^{\frac{\beta \gamma}{\gamma-1}} \). Unless \( \frac{\beta \gamma}{\gamma-1} = 1 \), this mapping has exactly two fixed points \( g = 0 \) and \( \hat{g} > 0 \). We are interested in the fixed point with a positive level of government spending \( \hat{g} > 0 \). The fixed point at \( \hat{g} \) is stable iff \( \phi'(g) = \frac{\beta \gamma}{\gamma-1} t \left[ \frac{1-t}{k} \right]^{\frac{1}{\gamma-1}} w_0^\gamma g^{\frac{\beta \gamma}{\gamma-1} - 1} < 1 \) iff \( \frac{\beta \gamma}{\gamma-1} < 1 \). In other words, a stable solution to the fixed point problem (4) exists if and only if \( \beta < \frac{1}{\varepsilon_{y,w}} \).

Figure 12 plots the fixed point problem for government spending when Assumption 1 holds (in blue) and does not hold (in red). The downward crossing point (the point where the blue curve crosses the 45-degree line) represents a stable equilibrium because, if \( g < \hat{g} \), then \( g \) will tend to rise, and if \( g > \hat{g} \), then \( g \) will tend to fall. So the stable equilibrium represents an “attracting” fixed point. In contrast, the upward crossing point (this point is not pictured, but it occurs when the red curve crosses the 45-degree line) represents an unstable equilibrium because, if \( g < \hat{g} \), then \( g \) will tend to fall, and, if \( g > \hat{g} \), the \( g \) will tend to rise.

Figure 12: Fixed point problem for government spending when \( \beta < \frac{1}{\varepsilon_{y,w}} \) and \( \beta \geq \frac{1}{\varepsilon_{y,w}} \)

Figure 13 sheds light on the origins of unstable equilibria by plotting wages against government spending (for some a fixed level of \( t \in (0, 1) \)). Figure 13(a) illustrates the case when Assumption 1 holds. The green curve plots wages as a function of government spending, while the blue curve plots the level of wages required for the government budget to balance. The point where the two lines cross represents a stable equilibrium since increasing \( g \) results in an unbalanced budget, and decreasing \( g \) results produces slack in the budget. In contrast,
the crossing point in Figure 13(b) represents an unstable equilibrium since increasing \( g \) produces slack in the budget.

![Figure 13: Government spending vs wage.](image)

\[ g = tw(g)l^*(t, w(g)) = tw_0g^{\gamma} \left[ \frac{(1-t)w_0g^\beta}{k} \right]^{\frac{1}{\gamma-1}} = tw_0^{\gamma} \left[ \frac{(1-t)}{k} \right]^{\frac{1}{\gamma-1}} g^{\frac{\beta+\gamma}{\gamma-1}}. \]

We can use the government budget balance requirement

\[ g = tw(g)l^*(t, w(g)) = tw_0g^{\gamma} \left[ \frac{(1-t)w_0g^\beta}{k} \right]^{\frac{1}{\gamma-1}} = tw_0^{\gamma} \left[ \frac{(1-t)}{k} \right]^{\frac{1}{\gamma-1}} g^{\frac{\beta+\gamma}{\gamma-1}}. \]

to solve for \( \hat{g}(t) \):

\[ \hat{g}(t) = \left[ \frac{(1-t)\gamma-1}{k} w_0^{\gamma} \right]^{1-\beta y,w} = [tw_0l^*(t; w_0)]^{\gamma-1} \left[ tw_0l^*(t; w_0) \right]^{\frac{1}{1-\beta y,w}}. \]  

We have verified numerically that the function \( \hat{g}(t) = [tw_0l^*(t; w_0)]^{\frac{1}{1-\beta y,w}} \) is quasi-concave, so the maximum exists and is unique. The first-order condition is given by

\[
0 = \frac{d\hat{g}(t)}{dt} = \frac{1}{1-\beta \varepsilon_{y,w}} \hat{g}(t) \left( \frac{1}{t} + \frac{1}{l^*(t; w_0)} \frac{l^*(t; w_0)}{dt} \right)
= \frac{1}{1-\beta \varepsilon_{y,w}} \hat{g}(t) \left( \frac{1}{t} + \frac{1}{l^*(t; w_0)} \frac{l^*(t; w_0)}{(-1)(1-t)(\gamma-1)} \right)
= \frac{1}{1-\beta \varepsilon_{y,w}} \hat{g}(t) \frac{(1-t)(\gamma-1)-t}{t(1-t)(\gamma-1)}.
\]
Note that $\hat{g}(t) = 0$ iff $t = 1$ or $t = 0$, which are corner solutions. Using $(1 - t)(\gamma - 1) - t = 0$, we find an interior solution $t = \frac{\gamma}{\gamma - 1} = \frac{1}{\varepsilon_{y,w}}$.

### C Derivation of Proposition 1

The informed planner solves

$$\max_t v(t, \hat{g}(t)) = \max_t v \hat{g}(t) + (1 - t)w(\hat{g})l^*(t, w(\hat{g})) - h(l^*(t, w(\hat{g}))).$$

The first-order condition is given by

$$\nu \frac{d\hat{g}(t)}{dt} - w(\hat{g}(t))l(t, w) + (1 - t)\frac{dw}{dt}l(t, w) + (1 - t)w(\hat{g}(t))\frac{dl}{dt} - \frac{dh}{dt} = 0.$$ 

From the derivation of Lemma 1 in Appendix B, we have

$$\frac{d\hat{g}(t)}{dt} = \hat{g}(t)\frac{(1 - t)(\gamma - 1) - t}{(1 - t)(\gamma - 1)} = \hat{g}(t)\frac{\gamma - \gamma t - 1}{(1 - t)(\gamma - \beta \gamma - 1)}.$$ 

Since $w(\hat{g}(t)) = w_0[\hat{g}(t)]^\beta$, then

$$\frac{dw}{d\hat{g}} \frac{d\hat{g}}{dt} = \beta w(\hat{g}(t)) \frac{\gamma - \gamma t - 1}{(1 - t)(\gamma - \beta \gamma - 1)}.$$ 

$$= \beta w(\hat{g}(t)) \frac{\gamma - \gamma t - 1}{(1 - t)(\gamma - \beta \gamma - 1)}.$$
Since \( l(t, w(\hat{g}(t))) = \left[ \frac{w(\hat{g}(t)(1-t)}{k} \right]^{\frac{1}{\tau-1}} \), then

\[
\frac{dl}{dt} = -\frac{1}{\gamma - 1} \frac{l(t, w(\hat{g}(t)))}{1-t} + \frac{1}{\gamma - 1} \frac{l(t, w(\hat{g}(t)))}{w(\hat{g}(t))} \frac{dw}{dg} \\
= \frac{-l(t, w(\hat{g}(t)))}{(\gamma - 1)(1-t)} + \frac{1}{\gamma - 1} \frac{l(t, w(\hat{g}(t)))}{w(\hat{g}(t))} \frac{\beta w(\hat{g}(t))}{l(1-t)(\gamma - \beta \gamma - 1)} \gamma - \gamma t - 1 \\
= \frac{-l(t, w(\hat{g}(t)))}{(\gamma - 1)(1-t)} + \beta \frac{l(t, w(\hat{g}(t)))}{\gamma - 1} \frac{\gamma - \gamma t - 1}{t(1-t)(\gamma - \beta \gamma - 1)} \\
= \frac{l(t, w(\hat{g}(t)))}{(\gamma - 1)(1-t)} \left[ \frac{\beta(\gamma - \gamma t - 1)}{t(\gamma - \beta \gamma - 1)} - 1 \right] \\
= \frac{l(t, w(\hat{g}(t)))}{(\gamma - 1)(1-t)} \frac{\beta \gamma - \beta - t \gamma + t}{t(\gamma - \beta \gamma - 1)} \\
= \frac{l(t, w(\hat{g}(t)))}{(\gamma - 1)(1-t)} \frac{\beta - t)(\gamma - 1)}{t(\gamma - \beta \gamma - 1)} \\
= \frac{l(t, w(\hat{g}(t)))}{(1-t)} \frac{(\beta - t)}{t(\gamma - \beta \gamma - 1)}. 
\]

Since \( h(l) = \frac{k}{\gamma} l^\gamma \), then

\[
\frac{dh}{dt} = k[l(t, w(\hat{g}(t)))]^\gamma \frac{dl}{dt} \\
= k[l(t, w(\hat{g}(t)))]^\gamma \frac{l(t, w(\hat{g}(t)))}{(1-t)} \frac{(\beta - t)}{t(\gamma - \beta \gamma - 1)} \\
= k[l(t, w(\hat{g}(t)))]^\gamma \frac{(\beta - t)}{t(1-t)(\gamma - \beta \gamma - 1)}. 
\]

Hence

\[
- w(\hat{g}(t)) \frac{dl}{dt} l(t, w) + (1-t) \frac{dw}{dt} l(t) + (1-t)w(\hat{g}(t)) \frac{dl}{dt} \\
= -w(\hat{g}(t)) l(t, w) + \beta w(\hat{g}(t)) l(t) \frac{\gamma - \gamma t - 1}{t(\gamma - \beta \gamma - 1)} + w(\hat{g}(t)) l(t, w) \frac{(\beta - t)}{t(\gamma - \beta \gamma - 1)} \\
= -w(\hat{g}(t)) l(t, w) \left[ \frac{\beta(\gamma - \gamma t - 1)}{t(\gamma - \beta \gamma - 1)} + (\beta - t) - 1 \right] \\
= w(\hat{g}(t)) l(t, w) \frac{\gamma(\beta - t)}{t(\gamma - \beta \gamma - 1)}. 
\]
Now
\[
\frac{w(\hat{g}(t))l(t, w)}{t(\gamma - \beta \gamma - 1)} \frac{\gamma(\beta - t)}{t(\gamma - \beta \gamma - 1)} \frac{\gamma(\beta - t)}{t(\gamma - \beta \gamma - 1)} - \frac{d\nu}{dt}
\]
\[
= w(\hat{g}(t))l(t, w) \frac{\gamma(\beta - t)}{t(\gamma - \beta \gamma - 1)} - k[l(t, w(\hat{g}(t)))]^\gamma \frac{(\beta - t)}{t(1 - t)(\gamma - \beta \gamma - 1)}
\]
\[
= \frac{\gamma(\beta - t)}{t(\gamma - \beta \gamma - 1)} [w(\hat{g}(t))l(t, w) - \frac{k}{\gamma(1 - t)} [l(t, w(\hat{g}(t)))]^\gamma]
\]
\[
= \frac{\gamma(\beta - t)}{t(1 - t)(\gamma - \beta \gamma - 1)} [(1 - t)w(\hat{g}(t))l(t, w) - h(t)].
\]

Substitute the results into the first-order condition:
\[
\nu \hat{g}(t) \frac{\gamma - \gamma t - 1}{t(1 - t)(\gamma - \beta \gamma - 1)} + \frac{\gamma(\beta - t)}{t(1 - t)(\gamma - \beta \gamma - 1)} [(1 - t)w(\hat{g}(t))l(t, w) - h(t)] = 0.
\]
Assuming \( t \neq 0, 1 \) and \( \beta \neq \frac{\gamma - 1}{\gamma} \), we have
\[
\nu (\gamma - \gamma t - 1) \hat{g}(t) + \gamma(\beta - t) [(1 - t)w(\hat{g}(t))l(t, w) - h(t)]
\]
\[
= \nu (\gamma - \gamma t - 1) \hat{g}(t) + \gamma(\beta - t) (1 - t)wl \frac{\gamma - 1}{\gamma}
\]
since \((1 - t)w(\hat{g}(t))l(t, w) - h(t) = (1 - t)wl \frac{\gamma - 1}{\gamma}\). Then the first-order condition is
\[
\nu (\gamma - \gamma t - 1) \hat{g}(t) + (\gamma - 1)(\beta - t)(1 - t)wl = 0.
\]
Since \( wl = w_0g(t)^\beta [w_0g^\beta (1 - t)/k]^{\frac{1}{\gamma - 1}} = w_0^{\frac{\gamma}{\gamma - 1}} g(t)^{\frac{\beta}{\gamma - 1}} [(1 - t)/k]^{\frac{1}{\gamma - 1}} \), then
\[
0 = \nu (\gamma - \gamma t - 1) \hat{g}(t) + (\gamma - 1)(\beta - t)(1 - t)wl
\]
\[
= \nu (\gamma - \gamma t - 1) \hat{g}(t) + (\gamma - 1)(\beta - t)(1 - t) \frac{\gamma}{\gamma - 1} w_0^{\frac{\gamma}{\gamma - 1}} \hat{g}(t)^{\frac{\beta}{\gamma - 1}} k^{\frac{1}{\gamma - 1}}
\]
\[
= \nu (\gamma - \gamma t - 1) + (\gamma - 1)(\beta - t)(1 - t) \frac{\gamma}{\gamma - 1} w_0^{\frac{\gamma}{\gamma - 1}} \hat{g}(t)^{\frac{\beta}{\gamma - 1}} k^{\frac{1}{\gamma - 1}}.
\]
assuming \( \hat{g}(t) \neq 0 \). We note that
\[
\hat{g}(t)^{\frac{\beta}{\gamma - 1}} = [t^{\gamma - 1} \frac{1 - t}{k}]^{\frac{\beta}{\gamma - 1}} w_0^{-\frac{\beta}{\gamma - 1}} [w_0^{\frac{\gamma}{\gamma - 1}} \hat{g}(t)^{\frac{\beta}{\gamma - 1}}]^{\frac{\beta}{\gamma - 1}}
\]
\[
= t^{-1} \left[ \frac{1 - t}{k} \right]^{\frac{1}{\gamma - 1}} w_0^{\frac{\gamma}{\gamma - 1}}.
\]
So

\[
0 = \nu(\gamma - \gamma t - 1) + (\gamma - 1)(\beta - t)(1-t) \frac{\gamma}{\gamma - 1} w_0^{\gamma-1} g(t)^{\beta - \gamma t - 1} k^{-\gamma - 1} w_L^{\gamma-1} k^{\gamma-1} \\
= \nu(\gamma - \gamma t - 1) + (\gamma - 1)(\beta - t)(1-t) \frac{\gamma}{\gamma - 1} w_0^{\gamma-1} t^{-1} \left[ \frac{1-t}{k} \right]^{\gamma-1} w_L^{\gamma-1} k^{-\gamma - 1} \\
= \nu(\gamma - \gamma t - 1) + (\gamma - 1)(\beta - t) \frac{1-t}{t} .
\]

Multiply the equation above by \( t \):

\[
0 = \nu t(\gamma - \gamma t - 1) + (\gamma - 1)(\beta - t)(1-t) \\
= t^2[(\gamma - 1) - \nu \gamma] + t[\nu \gamma - \nu - (\beta + 1)(\gamma - 1)] + (\gamma - 1) \beta .
\]

The solution is

\[
t = \frac{-[\nu \gamma - \nu - (\beta + 1)(\gamma - 1)]}{2[(\gamma - 1) - \nu \gamma]} \\
- \frac{\sqrt{[\nu \gamma - \nu - (\beta + 1)(\gamma - 1)]^2 - 4[(\gamma - 1) - \nu \gamma](\gamma - 1) \beta}}{2[(\gamma - 1) - \nu \gamma]}.
\]

We simplify the expression above as

\[
t^{PO} = \frac{-[(\nu - \beta - 1)(\gamma - 1)]}{2[(\gamma - 1) - \nu \gamma]} - \frac{\sqrt{[\nu - \beta - 1)(\gamma - 1)]^2 - 4(\gamma - 1)^2 \beta(1 - \frac{\nu}{\gamma - 1})}}{2[(\gamma - 1) - \nu \gamma]} \\
= \frac{-[(\nu - \beta - 1) + \sqrt{[\nu - \beta - 1)^2 - 4\beta(1 - \frac{\nu}{\gamma - 1})}]}{2(1 - \frac{\nu}{\gamma - 1})} \\
= \frac{-[(\nu - \beta - 1) + \sqrt{[\nu - \beta - 1)^2 - 4\beta(1 - \nu \varepsilon_{y,w})}]}{2(1 - \nu \varepsilon_{y,w})} \\
= \frac{(\beta - \nu + 1) - \sqrt{(\beta - \nu + 1)^2 - 4\beta(1 - \nu \varepsilon_{y,w})}}{2(1 - \nu \varepsilon_{y,w})}.
\]

Note that

\[
t^{PO} = \frac{(\beta - \nu + 1) - \sqrt{(\beta - \nu + 1)^2 - 4\beta(1 - \nu \varepsilon_{y,w})}}{2(1 - \nu \varepsilon_{y,w})} \geq 0 \quad (39)
\]
since, if \( (1 - \nu \varepsilon_{y,w}) > 0 \), then \( \sqrt{(\beta - \nu + 1)^2 - 4 \beta (1 - \nu \varepsilon_{y,w})} < (\beta - \nu + 1) \), so both the numerator and denominator are positive, and \( t^{PO} > 0 \). Similarly, if \( (1 - \nu \varepsilon_{y,w}) < 0 \), then \( \sqrt{(\beta - \nu + 1)^2 - 4 \beta (1 - \nu \varepsilon_{y,w})} > (\beta - \nu + 1) \), so both the numerator and denominator are negative, and \( t^{PO} > 0 \).

(i) Note that \( t^*(\nu = 0) = \frac{\beta + 1 - \sqrt{(\beta + 1)^2 - 4 \beta}}{2} = \frac{\beta + 1 - |\beta - 1|}{2} = \beta \).

(ii) Substituting \( \nu = 1 \) and \( \beta = \frac{1}{\varepsilon_{y,w}} \) into expression (39) yields \( \lim_{\beta \to \frac{1}{\varepsilon_{y,w}}} t^*(\nu = 1) = \frac{1}{\varepsilon_{y,w}} \sqrt{\left(\frac{1}{\varepsilon_{y,w}}\right)^2 - 4 \varepsilon_{y,w}^2 + 4} = \frac{1}{\varepsilon_{y,w}} \sqrt{\left(\frac{1}{\varepsilon_{y,w}} - 2\right)^2} = \frac{1}{\varepsilon_{y,w}} \) \( \varepsilon_{y,w}^2 \).

(iii) Define

\[
t^*(\nu) = \arg \max_{t \in [0,1]} \nu g(t) + (1 - t)w(g)l(w(g)) - h(l).
\]

We will prove that \( t^*(\nu) \) is increasing in \( \nu \).

**Proof.** By part (iii) of Lemma 2, we know that the optimal tax rate is bounded above by \( \frac{1}{\varepsilon_{y,w}} \). So we constrain our analysis to tax rate in this range. Let \( t', t'' \in [0, 1/\varepsilon_{y,w}] \) with \( t'' > t' \), and \( \nu', \nu'' \in [0, 1] \) with \( \nu'' > \nu' \). Then clearly \( \nu''[g(t'') - g(t')] > \nu'[g(t'') - g(t')] \) (since government spending is strictly increasing on \([0, 1/\varepsilon_{y,w}]\)). Since \( \nu g(t) \) has increasing differences, Topkis’s monotonicity theorem implies that \( t^*(\nu) \) is weakly increasing in \( \nu \). 

(iv) Consider the derivative given by

\[ t'(\beta) = \frac{1 - \frac{(\beta - \nu + 1) - 2(1 - \nu \varepsilon_{y,w})}{\sqrt{(\beta - \nu + 1)^2 - 4 \beta (1 - \nu \varepsilon_{y,w})}}}{2(1 - \nu \varepsilon_{y,w})}. \]

Suppose \( 1 - \nu \varepsilon_{y,w} > 0 \). Then

\[
t'(\beta) > 0 \iff \frac{(\beta - \nu + 1) - 2(1 - \nu \varepsilon_{y,w})}{\sqrt{(\beta - \nu + 1)^2 - 4 \beta (1 - \nu \varepsilon_{y,w})}} < 1
\]

\[
\iff (\beta - \nu + 1) - 2(1 - \nu \varepsilon_{y,w}) < (\beta - \nu + 1)^2 - 4 \beta (1 - \nu \varepsilon_{y,w})
\]

\[
\iff - 4(\beta - \nu + 1)(1 - \nu \varepsilon_{y,w}) + 4(1 - \nu \varepsilon_{y,w})^2 < -4 \beta (1 - \nu \varepsilon_{y,w})
\]

\[
\iff - (\beta - \nu + 1) + 4(1 - \nu \varepsilon_{y,w}) < -\beta
\]

\[
\iff 1 - \varepsilon_{y,w} < 0,
\]

which is true since \( \varepsilon_{y,w} = \frac{\gamma}{\gamma - 1} > 1 \). Similarly, we can show that if \( 1 - \nu \varepsilon_{y,w} < 0 \), then \( t'(\beta) > 0 \iff 1 - \varepsilon_{y,w} < 0 \). So \( t'(\beta) > 0 \).
D Proof of Proposition 2

(i) Suppose that wages are fixed, and we want to find the optimal tax rate. A naive planner solves

$$\max_{t,g} v((1 - t)wl^*(t) + \nu g(t), l^*(t)) \text{ s.t. } g \leq twl^*(t).$$

We plug $g(t)$ into the objective to reduce the problem to maximization over one variable. By the envelope theorem $u_l = 0$, so the first-order condition is given by

$$\frac{dv}{dt} = u_c[wl^*(t)\frac{d(1 - t)}{dt} + \nu \frac{dg}{dt}] = 0.$$ 

Since $g(t) = twl^*(t)$, then $\frac{dg}{dt} = wl^*(t) + twl^*(t)$, so we can re-write the expression above as

$$u_c[-wl^*(t) + \nu(wl^*(t) + twl^*(t))] = 0.$$ 

Suppose $\nu = 1$, then

$$u_c[twl''(t)] = 0.$$ 

Since $u_c, w, l'(t) \neq 0$, then it must be that $t^* = 0$. Note that the optimal tax rate coincides with effort maximizing tax rate, i.e. $l''(t^*) = 0$.

(ii) Suppose that wages are endogenous, but the planner has no access to lump-sum transfers, i.e. $\nu = 0$. The planner solves

$$\max_{t,g} v((1 - t)w(g)l(g, t), l(w(g), t)) \text{ s.t. } g \leq tw(g)l(g, t).$$

We solve for $g(t)$, and plug $g(t)$ into the objective to reduce the problem to maximization over one variable. By the envelope theorem $u_l = 0$, so the first-order condition is given by

$$\frac{dv}{dt} = u_c l \frac{d[\nu w(g(t))(1 - t)]}{dt} = 0.$$ 

Since $u_c, l > 0$, then the maximum must occur at $t$ such that $\frac{d}{dt}[w(g(t))(1 - t)] = 0$. Note that this condition is exactly the condition for $t$ that maximizes the labor supply $l(w(g(t)), t)$ since $\frac{dl}{dt} = 0$ at $t$ such that $\frac{d}{dt}[w(g(t))(1 - t)] = 0$. 

44
(iii) Suppose that wages are endogenous, but the planner has access to lump-sum transfers, i.e. \( \nu > 0 \). The planner solves

\[
\max_{t,g} v((1 - t)w(g)l(w(g), t) + \nu g(t), l(w(g), t)) \quad \text{s.t.} \quad g \leq tw(g)l(w(g), t).
\]

The first-order condition is given by

\[
\frac{dv}{dt} = u_c \frac{d}{dt}[w(g(t))(1 - t) + \nu g(t)] = 0.
\]

Note that the first term \( u_c l^*(t) \frac{d}{dt}[w(g(t))(1 - t)] \) is identical to the first-order condition in part (ii). So we know that \( t^{effort} \) is the unique \( t \in [0, 1] \) that satisfies \( \frac{d}{dt}[w(g(t))(1 - t)] = 0 \). So showing that \( u_c \nu \frac{dg}{dt} > 0 \) at \( t^{effort} \) is sufficient to establish that \( t^* > t^{effort} \). Observe that \( t^{effort} < t^{Laffer} \) since up until \( t^{effort} \) labor supply increases in \( t \) so the “standard” Laffer effects start to appear only for \( t > t^{effort} \). So there some small \( \varepsilon > 0 \) such that \( g(t) \) increases on \([t^{effort}, t^{effort} + \varepsilon]\). So \( \frac{dg}{dt}\bigg|_{t=t^{effort}} > 0 \). So

\[
\frac{dv}{dt}\bigg|_{t=t^{effort}} = u_c l^*(t^{effort}) \frac{d}{dt}[w(g(t^{effort}))(1 - t^{effort})] + u_c \nu \frac{dg}{dt}\bigg|_{t=t^{effort}} = u_c \nu \frac{dg}{dt}\bigg|_{t=t^{effort}} > 0
\]

implies that \( t^* > t^{effort} \).

E  Alternative definition of wages

Suppose that the informed planner solves program 6, where wages are given by \( w(g) = u_0(1 + g)^{\beta} \). Figure 14 provides the results from numerical simulations. Note that the optimal taxes are positive as long as \( \beta \) is sufficiently high and there is no optimism gap.

![Figure 14: Numerical simulation of \( t^{PO} \) vs. \( \beta \), by levels of \( \nu \)](image)
Figure 14 further indicates that the condition on the level of $\beta$ for the problem to be well-defined may be different under the new definition of wages. This plot assumes $\gamma = 3$, so Assumption 1 suggests that $\beta \in (0, 0.667)$ in the model with $w(g) = w_0g^\beta$. However, since the level of optimal tax rate achieves its peak for $\beta < 0.6$, then a valuable next step would be to investigate the conditions on $\beta$ in the model with $w(g) = w_0(1 + g)^\beta$.

F Derivation of Lemma 2

The informed planner solves

$$\max_t \nu \dot{g}(t) + [\chi + (1 - \chi)\gamma^\gamma\\r] w_L l_L - h_L] + tw_L l_L[\alpha - \chi][1 - r^\gamma].$$ \hspace{1cm} (40)$$

The first-order condition is given by

$$\nu \frac{d \dot{g}(t)}{dt} + [\chi + (1 - \chi)\gamma^\gamma\\r] [-w_L(\dot{g}(t))l_L(t, w_L) + (1 - t) \frac{dw}{dt} l_L(t) + (1 - t)w_L(\dot{g}(t)) \frac{dl}{dt} - \frac{dh_L}{dt}] = 0.$$ 

Following the same steps as in Appendix C, except that now we account for the two new constants $\bar{\chi} \equiv \chi + (1 - \chi)\gamma^\gamma$ and $\bar{\alpha} \equiv \alpha + (1 - \alpha)\gamma^\gamma$, we find that

$$t^{PO} = \frac{-[\nu \gamma - \nu - (\beta + 1)\bar{\alpha}(\gamma - 1)]}{2[\bar{\chi}(\gamma - 1) - \nu \gamma]} - \frac{\sqrt{[\nu \gamma - \nu - (\beta + 1)\bar{\alpha}(\gamma - 1)]^2 - 4[\bar{\chi}(\gamma - 1) - \nu \gamma][\bar{\alpha}(\gamma - 1)\beta]}{2[\bar{\alpha}(\gamma - 1) - \nu \gamma]}.$$ 

Let $m \equiv \bar{\alpha}$, then we can write the numerator as

$$(\gamma - 1)(\nu - [\beta + 1]m) - (\gamma - 1)\sqrt{(\nu - [\beta + 1]m)^2 - 4\beta m[m - \frac{\nu \gamma}{\gamma - 1}]}.$$ 

Consider the term under the square root $(\nu - [\beta + 1]m)^2 - 4\beta m[m - \frac{\nu \gamma}{\gamma - 1}] = \nu^2 - 2\nu(\beta + 1)m + (\beta + 1)^2m^2 - 4\beta m^2 + \frac{4\beta m \nu \gamma}{\gamma - 1} = \nu^2 - 2\nu\beta m - 2\nu m - \beta^2 m^2 + 2\beta m^2 + m^2 - 4\beta m^2 + \frac{4\beta m \nu \gamma}{\gamma - 1} = (\nu^2 - 2\nu m + m^2) - 2\beta m^2 + 2\beta m(\frac{2\nu m \gamma}{\gamma - 1} - \nu) + \beta^2 m^2 = (\nu - m)^2 + 2\beta m(\nu - \frac{2\nu m \gamma}{\gamma - 1} - m) + \beta^2 m^2 = (\nu - m)^2 + 2\beta m(\nu + \frac{2\nu}{\gamma - 1} - m) + \beta^2 m^2.$

Substituting the result into the tax rate formula, we have

$$t^{PO} = \frac{(\nu - (\beta + 1)m) + \sqrt{(\nu - (1 - \beta)m)^2 + 4\beta m \nu \varepsilon_{L,w}}}{2(\nu \varepsilon_{g,w} - m)},$$ \hspace{1cm} (41)
(i) Substituting \( \nu = 1 \) and \( \beta = \frac{1}{\varepsilon_{y,w}} \) into expression (41) yields

\[
\lim_{\beta \to \frac{1}{\varepsilon_{y,w}}} t^*(\nu = 1) = \frac{1 - (\frac{1}{\varepsilon_{y,w}} + 1)\mu + \sqrt{(1 - (1 - \frac{1}{\varepsilon_{y,w}})\mu)^2 + 4 \frac{1}{\varepsilon_{y,w}} \mu \varepsilon_{l,w}}}{2(\varepsilon_{y,w} - \mu)}
\]

\[
= \frac{1 - (\frac{1}{\varepsilon_{y,w}} + 1)\mu + \sqrt{1 + (1 - \frac{1}{\varepsilon_{y,w}})\mu)^2}}{2(\varepsilon_{y,w} - \mu)}
\]

\[
= \frac{1 - (\frac{1}{\varepsilon_{y,w}} + 1)\mu + 1 + (1 - \frac{1}{\varepsilon_{y,w}})\mu}{2(\varepsilon_{y,w} - \mu)}
\]

\[
= \frac{2 - 2\mu}{2(\varepsilon_{y,w} - \mu)}
\]

\[
= \frac{1}{\varepsilon_{y,w}}.
\]

(ii) We have \( t^{PO}(\nu = 0) = \frac{-(1+\mu+\sqrt{(1-\beta)\mu)^2}}{2(1-\mu)} = \frac{-(1+\mu)(1-\beta)\mu}{2(1-\mu)} = \beta. \)

(iii) Same proof as in Appendix C.

(iv) Same proof as in Appendix C.

(v) **Proof.** Consider the planner’s objective

\[
V(t, \chi) = \nu \tilde{g}(t) + \frac{\gamma - 1}{\gamma} (1 - t)w_L(\tilde{g}(t))l_L(t; w_L(\tilde{g}(t)))[\chi + (1 - \chi)r_0].
\]

We know that \( t^{PO} = \arg \max_t V(t; \chi) \) and that \( t^{PO} < t^{Laffer} \). We want to find how the optimal tax rate \( t^{PO} \) changes in response to an increase in social weight on low-types \( \chi \). Note that the partial derivative of the social objective with respect to \( \chi \) is given by

\[
\frac{\partial V}{\partial \chi} = \frac{\gamma - 1}{\gamma} (1 - t)w_L(\tilde{g}(t))l_L(t; w_L(\tilde{g}(t)))(1 - r_0).
\]

Since \( r_0 > 1 \), then the sign of the cross-partial \( \frac{\partial V}{\partial \chi \partial t} \) is opposite of the sign of the derivative \( \frac{d}{dt} \left[ \frac{\gamma - 1}{\gamma} (1 - t)w_L(\tilde{g}(t))l_L(t; w_L(\tilde{g}(t))) \right] \). Note that \( \frac{d}{dt} \left[ \frac{\gamma - 1}{\gamma} (1 - t)w_L(\tilde{g}(t))l_L(t; w_L(\tilde{g}(t))) \right] \) is decreasing for \( t > t^{PO} \), since \( t^{PO} \) is the unique maximizer. So \( \frac{\partial V}{\partial \chi \partial t} > 0 \). So the optimal tax rate \( t^{PO} \) is increasing in \( \chi \). ■

The numerical simulations further confirm that \( t^{PO} \) increases in \( \chi \). For example, Figure 15 plots the optimal tax rate against the total weight on L-types \( \chi \) by levels of \( \nu \). This plot shows that the optimal tax rate increases in \( \nu \), and that the rate of increase is higher for higher levels of \( \nu \).
G Proof of Lemma 3

Proof. Suppose that the government spending induced by the tax policy is $g^{SC}$ (with corresponding wages $w_i = w(g^{SC}; w_{0,i})$), but that the planner is unaware of the effect of government spending on wages, and so naively treats $g^{SC}$, and hence $w_i = w(g^{SC}; w_{0,i})$, as fixed. The planner believes that the budget constraint is

$$g(t) = \alpha w_L L(t) + (1 - \alpha) w_H H(t)$$

$$= t w_L L(t) [\alpha + (1 - \alpha) \gamma^{-\frac{\gamma}{\gamma - 1}}],$$

where $\alpha w_L L + (1 - \alpha) w_H H$ represents the total resources available in the economy. Since the naive planner treats wages as tax-independent, then the planner believes that the effect of a small change in tax rate on transfers $T(g(t); \nu^{SC}) = \nu^{SC} g(t)$ is

$$\frac{dT}{dt} = \nu^{SC} (\alpha + (1 - \alpha) \gamma^{-\frac{\gamma}{\gamma - 1}}) \left[ w_L L(t) + tw_L L'(t) \right]$$

$$= \nu^{SC} (\alpha + (1 - \alpha) \gamma^{-\frac{\gamma}{\gamma - 1}}) \left[ w_L L(t) + tw_L \frac{-l_L(t)}{(\gamma - 1)(1 - t)} \right]$$

$$= \nu^{SC} (\alpha + (1 - \alpha) \gamma^{-\frac{\gamma}{\gamma - 1}}) w_L L(t) \left[ 1 - \frac{t}{(\gamma - 1)(1 - t)} \right].$$

Figure 15: Numerical simulation of $t^{PO}$ vs. $\chi$, by levels of $\nu$
Since the total weight on the low-types is $\chi \in [0, 1]$, the naive planner solves

$$
\max_t T(t) + \left[ \chi(1-t) w_L l_L(t) - h(l_L(t)) \right] + [1 - \chi][1 - \chi][(1-t) w_H l_H(t) - h(l_H(t))]
$$

$$
= \max_t T(t) + (\chi + (1 - \chi) r_0^\frac{\gamma-1}{\gamma})(1-t)\frac{\mu}{\gamma-1}(1-t)\frac{\gamma}{\gamma-1} - \frac{\mu}{\gamma-1}(1-t)\frac{\gamma}{\gamma-1} w_L^\frac{\gamma}{\gamma-1}.
$$

The necessary condition for the planner’s problem is

$$
\nu^{SC}(\alpha + (1 - \alpha) r_0^\frac{\gamma}{\gamma-1}) w_L l_L(t) \left[ 1 - \frac{t}{(\gamma-1)(1-t)} \right] = (\chi + (1 - \chi) r_0^\frac{\gamma}{\gamma-1})(1-t)\frac{\mu}{\gamma-1} w_L^\frac{\gamma}{\gamma-1}.
$$

Since $w_L l_L(t) = w_L^\frac{\gamma}{\gamma-1}(1-t)\frac{\mu}{\gamma-1}$, then, assuming $t \neq 1$, the equation above simplifies to

$$
\nu^{SC}(\alpha + (1 - \alpha) r_0^\frac{\gamma}{\gamma-1}) \left[ 1 - \frac{t}{(\gamma-1)(1-t)} \right] = (\chi + (1 - \chi) r_0^\frac{\gamma}{\gamma-1}).
$$

Re-arranging yields

$$
\frac{t}{1-t} = \frac{(\gamma-1)(\nu^{SC}(\alpha + (1 - \alpha) r_0^\frac{\gamma}{\gamma-1}) - (\chi + (1 - \chi) r_0^\frac{\gamma}{\gamma-1}))}{\nu^{SC}(\alpha + (1 - \alpha) r_0^\frac{\gamma}{\gamma-1})} = \frac{\nu^{SC} - \mu}{\nu^{SC} \varepsilon_{t,w}},
$$

where $\mu \equiv \frac{\chi(1-\chi) r_0^\frac{\gamma}{\gamma-1}}{\alpha + (1-\alpha) r_0^\frac{\gamma}{\gamma-1}}$ and $\varepsilon_{t,w} \equiv \frac{1}{\gamma-1}$. So $t^{SC} = \frac{\nu^{SC} - \mu}{\nu^{SC} \varepsilon_{t,w}}$.

(i) Note that $t^{SC} = \frac{\nu^{SC} - \mu}{\nu^{SC} \varepsilon_{t,w}} > 0$ iff $\nu^{SC} > \mu$. Given utilitarian weights $\chi = \alpha$, we have $\mu(\chi, \alpha) = 1 \neq \nu^{SC}$ for all $\nu^{SC} \in [0, 1]$, so $t^{SC} = 0$, as in Section 2. So we need to choose $\chi > \alpha$ such that $\nu^{SC} > \mu(\alpha, \chi)$ for $t^{SC} > 0$.

(ii) Note that $\frac{d}{d \nu^{SC}} \left[ t \frac{1-t}{1-t} \right] = \frac{d}{d \nu^{SC}} \left[ \nu^{SC} - \mu \right] = \frac{\nu^{SC} \mu}{\nu^{SC} \varepsilon_{t,w}} > 0$. Also $\frac{d \mu}{d \chi} = \frac{1-r_0^\frac{\gamma}{\gamma-1}}{\alpha + (1-\alpha) r_0^\frac{\gamma}{\gamma-1}} > 0$ and $\frac{d \mu}{d r_0} = -\frac{\chi-\alpha}{(\alpha + (1-\alpha) r_0^\frac{\gamma}{\gamma-1})^2} < 0$ for $\chi > \alpha$, so $\frac{d \mu}{d \mu} \left[ \frac{t}{1-t} \right] = -\frac{1}{\nu^{SC} \varepsilon_{t,w}} < 0$ implies that $t^{SC}$ increases in $\chi$ and $r_0$.

(iii) Note that $\frac{d}{d \varepsilon_{t,w}} \left[ \frac{t}{1-t} \right] = -\frac{\nu^{SC}(\nu^{SC} - \mu)}{\nu^{SC} \varepsilon_{t,w} + \nu^{SC} - \mu} < 0$ for $\nu^{SC} > \mu$.  

49
H Proof of Proposition 3

(i) Let $\nu \in [0, 1]$ and suppose that $\nu^{SC} = \nu$. By the formula for the optimal tax rate in Lemma 2, we have $t^{PO} = t^{SC}$ at $\beta = 0$. Since $t^{PO}$ increases in $\beta$ by property (ii) of Lemma 2, we have $t^{PO} > t^{SC}$ for $\beta > 0$.

(ii) Suppose $\nu^{SC} = 1$ and $\beta > 0$. By the argument above, we have $t^{PO}(\nu = 1) > t^{SC}(\nu^{SC} = 1)$. Since the function for the optimal tax rate in Lemma 2 is continuous and increasing in $\nu$, there is $\nu \geq 0$ such that for all $\nu \in (\nu, 1]$ we have $t^{PO}(\nu) > t^{SC}(\nu^{SC} = 1)$.

(iii) Suppose that $\nu^{SC} = 1$ and $\nu = 0$. By property (ii) of Lemma 2, we have $t^{PO}(\nu = 0) = \beta \in (0, \frac{1}{\varepsilon_{y,w}})$. Note that $t^{SC} < t^{Laffer} = \frac{1}{\varepsilon_{y,w}}$ for all $\nu^{SC} \in [0, 1]$ (since the revenue maximizing point does not depend on the assumption of variable wages by Lemma 1). Choose $\beta = t^{SC}(\nu^{SC} = 1) < \frac{1}{\varepsilon_{y,w}}$. Then, for all $\beta \in (\beta, \frac{1}{\varepsilon_{y,w}})$, we have $t^{SC}(\nu^{SC} = 1) < t^{PO}(\beta, \nu = 0)$.

I Exogenous social weights in Section 4

Figure 16: L-type and H-type wages at the optimum vs. $\beta_L$

Figure 16 shows that, for sufficiently high levels of $\beta_L$ (government spending elasticity of wage for low-types), the order of wages can switch at the optimum, i.e. $w_L > w_H$, if we use exogenous social weights.

We check that the assumption of exogenous weights does not affect the numerical results presented in Section 4 by plotting the wages for low- and high-types at the optimal tax policies. Note that we only need to check the simulation in Figure 9(b) since Figures 9(a) and 10 both assume that $\beta_H > \beta_L = 0$. 

50
J  Endogenous social weights in Section 4

To preserve the order of wages at the optimum, we can use endogenous welfare weights $G(u)$, with $G'(u) > 0$ and $G''(u) < 0$, to capture that the marginal social weight decreases in utility. In particular, we assume $G(u) = \log(u)$, as in Saez (2001).\(^6\)

The planner who is aware of the effect of government spending on wage distribution solves

$$
\max_{t,g} \alpha G(u_L(c^*_L, l^*_L)) + (1 - \alpha) G(u_H(c^*_H, l^*_H))
$$

$$
s.t. \quad g \leq t[\alpha w_L(g)l^*_L(t, w_L(g)) + (1 - \alpha) w_H(g)l^*_H(t, w_H(g))]. \tag{42}
$$

The numerical simulations show that results remain qualitatively similar to the results in Section 4. For example, Figure 18 plots SCPE and social optima under the assumption of endogenous social weights (using the same parameters as used to produce Figure 9), and both Figures 18 and 9 are qualitatively similar.

\(^6\)This special case of CRRA social welfare function $G(u) = \frac{u^{1-\xi}}{1-\xi}$ with $\xi = 1$, where $\xi = 1$ is assumed for tractability.
\( \beta_L = 0, \beta_H = 0.1 \)

\( \beta_L = 0.3, \beta_H = 0.2 \)

Figure 18: SCPE vs. Social optima

Figures 10 and 19 are also qualitatively similar.

Figure 19: SCPE vs. Social optima (\( \beta_L = 0, \beta_H = 0.3 \))

So the assumption of endogenous social weights does not influence the comparison between the SCPE and the optimal tax rates under the model parameters assumed.

K Dynamic programming algorithm for finding the SCPE in section 4

Let \( \epsilon > 0 \) be the level of tolerance.

- Take \( g_0 \) from the informed planner’s problem (which implies wages \( w_H, w_L \)).
• Given \( g_0 \), and hence the wages \( w_H, w_L \), solve for \( t_1 \), where

\[
 t_1 = \arg \max_t G(u(c_L, l_L)) + (1 - \alpha)G(u(c_H, l_H)),
\]

where \( l_i(t; w_i) = \left[ \frac{w_i(1-t)}{k} \right]^{-\frac{1}{\gamma-1}} \), \( c_i(T; w_i) = T + (1 - t)w_i l_i(t; w_i) \) and \( T = t[\alpha w_L l_L + (1 - \alpha)w_H l_H] \).

• Check if \( t_1 \) balances the informed planner’s budget. If not, let \( g_1 \) be the new level of government spending implied by \( t_1 \).

• Iterate until \( |g_n - g_{n+1}| < \epsilon \).

L Proof of Proposition 4

The planner solves

\[
\max_\eta \; \nu g^* + [(1 - t^*)w_L(\eta; g^*)l_L(\eta; t^*) - h(l_L)] \bar{\chi},
\]

with the first-order condition given by

\[
(1 - t^*) \left[ \frac{dw}{d\eta} l + w \frac{dl}{d\eta} \frac{dw}{d\eta} \right] - \frac{dh}{dl} \frac{dl}{d\eta} \frac{dw}{d\eta} = 0.
\]

Since \( m(g) = g^\beta [\beta_1 \eta^\rho + \beta_2 (1 - \eta)^\rho]^{\frac{\beta}{\rho}} \), then

\[
\frac{dm}{d\eta} = g^\beta \beta \rho \left[ \beta_1 \eta^\rho + \beta_2 (1 - \eta)^\rho \right]^{\frac{\beta-\rho}{\rho}} \left[ \rho \beta_1 \eta^{\rho-1} - \rho \beta_2 (1 - \eta)^{\rho-1} \right] \\
= g^\beta \beta \left[ \beta_1 \eta^\rho + \beta_2 (1 - \eta)^\rho \right]^{\frac{\beta-\rho}{\rho}} \left[ \beta_1 \eta^{\rho-1} - \beta_2 (1 - \eta)^{\rho-1} \right]
\]

so

\[
\frac{dw}{d\eta} = \frac{dm}{d\eta} \frac{dw}{dm} = w_0 g^\beta \beta \left[ \beta_1 \eta^\rho + \beta_2 (1 - \eta)^\rho \right]^{\frac{\beta-\rho}{\rho}} \left[ \beta_1 \eta^{\rho-1} - \beta_2 (1 - \eta)^{\rho-1} \right],
\]

\[
\frac{dl}{dw} \frac{dw}{d\eta} = \frac{1}{\gamma - 1} w^\frac{2 - \gamma}{\gamma} \left[ \frac{1 - t}{k} \right]^{-\frac{1}{\gamma-1}} \frac{dw}{d\eta}
\]

\[
= \frac{1}{\gamma - 1} w_0^\frac{2 - \gamma}{\gamma} \left[ g^\beta \beta \left[ \beta_1 \eta^\rho + \beta_2 (1 - \eta)^\rho \right]^{\frac{\beta-\rho}{\rho}} \right]^{\frac{\beta-\rho}{\rho}} \left[ \frac{1 - t}{k} \right]^{-\frac{1}{\gamma-1}} \frac{dw}{d\eta}
\]

\[
= \frac{\beta}{\gamma - 1} w_0^\frac{2 - \gamma}{\gamma} g^\frac{\beta}{\rho} \left[ \beta_1 \eta^\rho + \beta_2 (1 - \eta)^\rho \right]^{\frac{\beta-\rho}{\rho-1}} \left[ \beta_1 \eta^{\rho-1} - \beta_2 (1 - \eta)^{\rho-1} \right] \left[ \frac{1 - t}{k} \right]^{-\frac{1}{\gamma-1}},
\]
and
\[
\frac{dh}{dl} = k l^{\gamma - 1} \frac{dl}{dw}
= w (1-t) \frac{dl}{dw}
= w_0 g^\delta [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma (1-t) \frac{dl}{dw}
= \frac{(1-t)\beta}{\gamma - 1} w_0 \frac{1}{w_0} g^{\frac{\delta}{\gamma - 1}} [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma [\beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1] \left[ \frac{1-t}{k} \right]^{\frac{1}{\gamma - 1}}.
\]

Now
\[
\frac{dw}{w \eta} l = w_0 g^\delta [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma \left[ \beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1 \right] l
= \beta \left[ \frac{1-t}{k} \right]^{\frac{1}{\gamma - 1}} w_0 \frac{1}{w_0} g^{\frac{\delta}{\gamma - 1}} [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma [\beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1],
\]
and
\[
w \frac{dl}{dw} \frac{dw}{d\eta} = w \frac{\beta}{\gamma - 1} w_0 \frac{1}{w_0} g^{\frac{\delta}{\gamma - 1}} [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma [\beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1],
\]
so
\[
\frac{dw}{w \eta} l + w \frac{dl}{d\eta} \frac{dw}{d\eta} = \beta \left[ \frac{1-t}{k} \right]^{\frac{1}{\gamma - 1}} g^{\frac{\delta}{\gamma - 1}} [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma [\beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1] \left[ \frac{1}{w \gamma - 1} - \frac{1}{\gamma - 1} \right]
\]
Since
\[
\frac{dh}{dl} \frac{dl}{dw} \frac{dw}{d\eta} = \frac{(1-t)\beta}{\gamma - 1} w_0 \frac{1}{w_0} g^{\frac{\delta}{\gamma - 1}} [\beta_1 \eta^\rho + \beta_2 (1-\eta)^\rho]_\sigma [\beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1] \left[ \frac{1-t}{k} \right]^{\frac{1}{\gamma - 1}},
\]
then the first-order condition
\[
(1-t^*) \left[ \frac{dw}{d\eta} l + w \frac{dl}{d\eta} \frac{dw}{d\eta} \right] - \frac{dh}{dl} \frac{dl}{d\eta} \frac{dw}{d\eta} = 0
\]
implies
\[
\beta_1 \eta^\rho - \beta_2 (1-\eta)^\rho - 1 = 0,
\]
54
so
\[
\frac{\eta}{1 - \eta} = \left( \frac{\beta_2}{\beta_1} \right)^{\frac{1}{\rho - 1}},
\]
i.e.
\[
\eta = \frac{\left( \frac{\beta_2}{\beta_1} \right)^{\frac{1}{\rho - 1}}}{1 + \left( \frac{\beta_2}{\beta_1} \right)^{\frac{1}{\rho - 1}}}.
\] (43)

(i) Note that
\[
\eta^{PO} = \arg\max_{\eta} \nu g^* + [(1 - t^*)w_L(\eta; g^*)l_L(\eta; t^*) - h(l_L)]\bar{\chi} \\
= \arg\max_{\eta} \nu g^* + \frac{\gamma - 1}{\gamma}(1 - t)w_L(\eta; g^*)l_L(\eta; t^*)\bar{\chi} \\
= \arg\max_{\eta} \nu g^* + \frac{\gamma - 1}{\gamma}(1 - t)[w_L(\eta; g^*)]^{\frac{1}{\gamma - 1}}\left[ \frac{1 - t^*}{k} \right]^{\frac{1}{\gamma - 1}}\bar{\chi} \\
= \arg\max_{\eta} w_L(\eta; g^*)
\]
since all the other terms are positive constants.

(ii) Follows directly from expression (43).

(iii) Follows directly from expression (43).

M Proof of Lemma 4

Suppose that \( \eta^*_i = \arg\max_{\eta} w_i(\eta; g(t^{PO}), \zeta_i) \). Let \( \bar{\eta} = \min\{\eta^*_i : i \in \{L, H\}\} \) and \( \bar{\eta} = \max\{\eta^*_i : i \in \{L, H\}\} \). Then \( \frac{dw}{d\eta} > 0 \) for all \( \eta < \bar{\eta} \) and \( \frac{dw}{d\eta} < 0 \) for all \( \eta > \bar{\eta} \). So \( \eta^{PO} \in [\bar{\eta}, \bar{\eta}] \).

(i) This result follows immediately by applying Proposition 4 separately for L-types and H-types, i.e. \( \zeta_L > \zeta_H \) iff \( \eta^*_L = \zeta_L^\sigma < \zeta_H^\sigma = \eta^*_H \) since \( \sigma > 0 \).

(ii) Proof. Let \( \chi \in (0, 1) \). Suppose \( \zeta_L > \zeta_H \). Then \( \eta^*_L < \eta^*_H \) by the argument above. Hold \( t^{PO} \) fixed. By the first part of Lemma 4, the planner selects \( \eta^{PO} \in [\eta^*_L, \eta^*_H] \) that solves
\[
\max_{\eta} \nu g^{PO} + [(1 - t^{PO})w_L(\eta; g^{PO})l_L(\eta; t^{PO}) - h(l_L)](\chi + (1 - \chi)r_0^{\frac{1}{\sigma - 1}}[\Delta(\eta)]^{\frac{\sigma}{\sigma - 1}}).
\]
By the product rule, the first-order condition is given by

\[
\bar{\chi} \left[ (1 - t^{PO}) \left[ \frac{dw}{d\eta} + w \frac{dl}{dw} \frac{d\eta}{dl} \right] - \frac{dh}{dl} \frac{dw}{d\eta} \right] + \bar{\chi}'(\eta) \left[ (1 - t^{PO}) w_L(\eta; g^{PO}) l_L(\eta; t^{PO}) - h(l_L) \right] = 0.
\] (44)

Since \( \eta^*_L \) is optimal in the model with one type (so the first term \( \bar{\chi} \left[ (1 - t^{PO}) \left[ \frac{dw}{d\eta} + w \frac{dl}{dw} \frac{d\eta}{dl} \right] \right] = 0 \) at \( \eta^*_L \)), then \( \eta^*_L \) is optimal in the model with two types iff \( \bar{\chi}'(\eta) \left[ (1 - t^{PO}) w_L(\eta; g^{PO}) l_L(\eta; t^{PO}) - h(l_L) \right] = 0 \) iff \( \bar{\chi}'(\eta) = 0 \). Note that

\[
\bar{\chi}'(\eta) = \frac{d}{d\eta} \left( \chi + (1 - \chi) \Delta(\eta) \frac{\beta \gamma}{\rho(\gamma - 1)} \right) = (1 - \chi) \frac{\beta \gamma}{\rho(\gamma - 1)} \Delta'(\eta),
\]

where

\[
\Delta'(\eta) = r_0 \left[ \frac{\rho \eta^{\rho - 1} + \zeta_H \rho (1 - \eta)^{\rho - 1}}{\eta^\rho + \zeta_L (1 - \eta)^\rho} - \frac{[\eta^\rho + \zeta_L (1 - \eta)^\rho][\rho \eta^{\rho - 1} + \zeta_L \rho (1 - \eta)^{\rho - 1}]}{[\eta^\rho + \zeta_L (1 - \eta)^\rho]^2} \right]
\]

\[
= r_0 \left[ \frac{\rho \eta^{\rho - 1} + \zeta_H \rho (1 - \eta)^{\rho - 1}}{\eta^\rho + \zeta_L (1 - \eta)^\rho} - \frac{[\eta^\rho + \zeta_H (1 - \eta)^\rho][\rho \eta^{\rho - 1} + \zeta_L \rho (1 - \eta)^{\rho - 1}]}{[\eta^\rho + \zeta_L (1 - \eta)^\rho]^2} \right].
\]

So \( \bar{\chi}'(\eta)|_{\eta = \eta^*_L} = 0 \) if and only if \( \chi = 1 \), a contradiction to our assumption that \( \chi \in (0, 1) \). Similarly, we can show that the first-order condition in expression (44) equals zero for \( \eta = \eta^*_H \) iff \( \chi = 0 \). Since we assume \( \chi \in (0, 1) \), then \( \eta^{PO} \in (\eta^*_L, \eta^*_H) \). ■

N Proof of Proposition 5

(i) Proof. Fix welfare weights \( \chi \in (0, 1) \). Start with a model with homogeneous effects on wages, i.e. \( \frac{\phi_{1,i}}{\phi_{2,i}} = \zeta \) independent of \( i \). In this model, the wage ratio \( r_0 \equiv \frac{w_H}{w_L} \) does not depend on endogenous variables, and we can explicitly solve for the optimal policy parameters

\[
(t_0, \eta_0) = \arg \max_{t, \eta} \nu \hat{g}(t, \eta) + [(1 - t) w_L(\hat{g}(t, \eta)) l_L(t, \eta) - h(l_L)] [\chi + (1 - \chi) r_0^{-1}].
\]

Note that \( \eta_0 = \arg \max_{\eta} w_L(\eta; \hat{g}(t^{PO}, \zeta)) = \arg \max_{\eta} w_H(\eta; \hat{g}(t^{PO}, \zeta)) \), i.e. the wage maximizing \( \eta \) is the same for both types, by property (i) of Proposition 4.

Next we consider a new model with \( \frac{\phi_{1,H}}{\phi_{2,H}} \neq \frac{\phi_{1,L}}{\phi_{2,L}} \), but for which at \( t_0 \) and \( \eta_0 \), the wage ratio \( w_H/w_L \) is the same, i.e. \( \frac{w_H(t_0, \eta_0)}{w_L(t_0, \eta_0)} = r_0 \). Given the heterogeneity in the share parameters, tax policy \( (\eta_0, t_0) \) may no longer be optimal. We know that, for this
specific \( \eta_0 \), tax rate \( t_0 \) is optimal (since, at \( \eta_0 \), the wage ratio appears fixed at \( r_0 \) and the expression (32) from the homogeneous effects of \( g \) case tells that \( t_0 \) does not vary with \( \eta \)). We want to know if \( \eta_0 \) is optimal given \( t_0 \), and, if not, if the optimal \( \eta \) is above or below the original \( \eta_0 \).

To answer the questions described in the previous paragraph, let \( y \in \mathbb{R} \) and define

\[
\phi_{2,L}(y) = \phi_{2,L} + y \quad \text{and} \quad \phi_{1,L}(y) = \phi_{1,L}(1 + \zeta x(\eta_0)) - (\phi_{2,L} + y)x(\eta_0),
\]

where \( x(\eta) \equiv \left(\frac{1-\eta}{\eta}\right)^{\rho} \). We verify that these share parameters produce the desired properties:

- For all \( y \in \mathbb{R} \), we have \( \frac{w_H(t_0, \eta_0)}{w_L(t_0, \eta_0); y} = r_0 \) since

\[
\frac{w_H(t_0, \eta_0)}{w_L(t_0, \eta_0); y} = r_0 \frac{\Delta(y; \eta_0)}{\rho} \left( \frac{1 + \zeta x(\eta_0)}{1 + \phi_{1,L}(y)x(\eta_0)} \right)^{\frac{\rho}{\beta}} \left( \frac{\phi_{1,L}(y)}{\phi_{1,L}(y) + \phi_{2,L}(y)x_0} \right)^{\frac{\rho}{\beta}} \left( \frac{\phi_{1,L}(1 + \zeta x_0)}{\phi_{1,L}(1 + \zeta x_0) - \phi_{2,L}(y)x_0 + (\phi_{2,L} + y)x_0} \right)^{\frac{\rho}{\beta}} = r_0,
\]

where we write \( x_0 \equiv x(\eta_0) \) for simplicity.

- At \( y = 0 \), we are in the homogeneous effects of \( g \) case since

\[
\frac{\phi_{2,L}(y; \eta_0)}{\phi_{1,L}(y; \eta_0)} = \frac{\phi_{2,L} + y}{\phi_{1,L}(1 + \zeta x_0) - (\phi_{2,L} + y)x_0} = \frac{1}{\phi_{1,L}(1 + \zeta x_0) - x_0}.
\]

In particular

\[
\frac{\phi_{2,L}(0; \eta_0)}{\phi_{1,L}(0; \eta_0)} = \frac{1}{\phi_{1,L}(1 + \zeta x_0) - x_0} = \frac{1}{\frac{1+\zeta x_0}{\zeta} - x_0} = \zeta.
\]
• Increasing $y$ makes L-types benefit relatively more from $g_2$ since
\[
\zeta_L(y) = \frac{d}{dy} \left[ \frac{1}{\phi_2,L + y} \right] = \frac{d}{dy} \left[ \frac{1}{\phi_2,L(1 + x_0) - x_0} \right] = \frac{\phi_1,L(1 + x_0)}{(y x_0 - \phi_1,L x_0 + \phi_2,L x_0 - \phi_1,L)^2} > 0.
\]

So $\zeta_L(y)$ is increasing.

We will apply Lemma 4 to find how the interval for the optimal $\eta^{PO}$ changes in response to changes in $y$. Note that the share parameters for high-types (and hence their wages) do not depend on $y$, so $\eta^\ast_H = \eta_0 = \arg \max_\eta w_H(\eta; \hat{g}(t^{PO}, \zeta))$. To find $\eta^\ast_L$, we solve
\[
\max_\eta w_L(\eta, y; \hat{g}(t^{PO})) = \max_\eta g^\ast [\phi_1,L(\eta) \eta^\rho + \phi_2,L(\eta)(1 - \eta)^\rho]^{\frac{\beta}{\beta - \rho}}.
\]  

The first-order condition for problem (46) is given by
\[
\frac{\partial w_L}{\partial \eta} = \max_\eta g^\beta [\phi_1,L(\eta) \eta^\rho + \phi_2,L(\eta)(1 - \eta)^\rho]^{\frac{\beta - \rho}{\beta}} [\phi_1,L(\eta) \eta^{\rho - 1} - \phi_2,L(\eta)(1 - \eta)^{\rho - 1}].
\]

We are interested in the sign of $\frac{\partial w_L}{\partial \eta}$, which depends only on the sign of $[\phi_1,L(\eta) \eta^{\rho - 1} - \phi_2,L(\eta)(1 - \eta)^{\rho - 1}]$. Note that $\frac{d\phi_1,L(y)}{dy} < 0$ and $\frac{d\phi_2,L(y)}{dy} > 0$ by their definition in expression (45). So $[\phi_1,L(\eta) \eta^{\rho - 1} - \phi_2,L(\eta)(1 - \eta)^{\rho - 1}] < 0$, and $\frac{\partial w_L}{\partial \eta} < 0$. So, for $y > 0$, the wages for low-types are decreasing at the original $\eta_0$, and $\eta^\ast_L < \eta_0 = \eta^\ast_H$. So the new interval for the optimal tax rates is given by $\eta^{PO} \in [\eta^\ast_L, \eta^\ast_H]$. Since we assume $\chi \in (0, 1)$, then property (ii) of Proposition 4 implies that $\eta^{PO} \in (\eta^\ast_L, \eta^\ast_H)$. So $\eta^{PO} < \eta^\ast_H = \eta_0$. So we have shown that $\eta^{PO}$ is locally decreasing in $\zeta_L$. In the proof above, we could similarly show that $\eta^{PO}$ is locally decreasing in $\zeta_H$ if we instead kept $\zeta_L$ fixed and defined $\phi_{j,H}(y)$ so that $\zeta_H(y)$ is increasing. ■

(ii) **Proof.** Let $\chi \in (0, 1)$. Suppose that the optimal tax policy is given by $(t^{PO}, \eta^{PO})$. Hold $t^{PO}$ fixed. The planner solves
\[
\max_\eta \nu g^{PO} + \chi \left[ (1 - t^{PO}) w_L(\eta; g^{PO}) l_L(\eta; t^{PO}) - h(l_L) \right]_{w_L(\nu; t^{PO})} \\
+ (1 - \chi) \left[ (1 - t^{PO}) w_H(\eta; g^{PO}) l_L(\eta; t^{PO}) - h(l_H) \right]_{w_H(\nu; t^{PO})},
\]

where $g^{PO} = \hat{g}(\eta^{PO}, t^{PO})$. The first-order condition for the problem above is given by
\[
\chi \frac{\partial v_L}{\partial \nu} \frac{\partial w_L}{\partial \eta} + (1 - \chi) \frac{\partial v_H}{\partial \nu} \frac{\partial w_H}{\partial \eta}.
\]
Consider a small increase in $\chi$. Then the cross-partial of the expression above 
\[
\frac{\partial v_L}{\partial w_L} \frac{\partial w_L}{\partial \eta} - \frac{\partial v_H}{\partial w_H} \frac{\partial w_H}{\partial \eta} > 0
\]
iff $\frac{\partial w_L}{\partial \eta} > 0$ and $\frac{\partial w_H}{\partial \eta} < 0$ iff $\eta^{PO} \in (\eta^*_H, \eta^*_L)$ iff $\zeta_L < \zeta_H$ (by Lemma 4). So $\eta^{PO}$ increases in $\chi$ iff $\zeta_L < \zeta_H$ (as expected since $\zeta_L < \zeta_H$ implies that spending $g_1$ provides greater benefit to low-types, so we want to spend more on $g_1$ when we value the consumption of low-types more, i.e. $\chi$ increases). \hfill \blacksquare