

# Organization of Knowledge and Taxation

Marek Kapička and Ctirad Slavík\*

February 14, 2019

## Abstract

This paper studies how income taxation interacts with the organization of knowledge and production, and ultimately the distribution of wages in the economy. A more progressive tax system reduces the time that managers allocate to work. This makes the organization of production less efficient and reduces wages at both tails of the distribution, which increases lower tail wage inequality and decreases upper tail wage inequality. The optimal tax system is significantly less progressive than the current one in the United States, because a less progressive tax system also generates a less unequal wage distribution.

J.E.L Codes: E6, H2, D8, L23

Keywords: inequality, wages, knowledge based hierarchies, income taxation

## 1 Introduction

United States and other developed economies have recently experienced substantial changes in wage inequality. In particular, after 1986, the upper tail wage inequality (90/50 percentile ratio) has significantly increased, while the lower tail wage inequality (50/10 percentile wage ratio) has decreased (many empirical studies confirm this fact,

---

\*Marek Kapička, CERGE-EI. Email: marek.kapicka@cerge-ei.cz. Ctirad Slavík, CERGE-EI. Email: ctirad.slavik@cerge-ei.cz. CERGE-EI is joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Czech Academy of Sciences. Politických vězňů 7, Praha 1, 111 21, Czech Republic.

see for example [Garicano and Rossi-Hansberg \(2011\)](#) with data until 2015). Standard models that study optimal taxation either assume that the wage distribution is exogenous, or that it can be partially modified by human capital investment. This is true for papers that use mechanism design techniques such as [Mirrlees \(1971\)](#) and many followers, as well as for papers that use parametric tax functions [Heathcote et al. \(2016\)](#) and others. Neither of the approaches can explain the observed changes in the upper and lower wage inequality without artificially engineering “correct” changes in the underlying exogenous distributions of wages or abilities. In addition, the interaction between changes in wage inequality and changes in taxes is nonexistent or limited.

In this paper we study the interaction between taxation and wage inequality by using a theory that studies how society organizes and uses knowledge in production by using knowledge based hierarchies ([Garicano \(2000\)](#), [Garicano and Rossi-Hansberg \(2006\)](#)). The theory of knowledge based hierarchies can explain simultaneous changes in the upper tail and the lower tail wage inequality as follows. It takes time to coordinate and communicate knowledge among managers and production workers. A decrease in the required communication time makes good managers relatively more useful in solving tasks, which increases wages of good managers relative to less able managers (an increase in upper tail inequality). At the same time, a decrease in the communication time allows managers to supervise more production workers. Even the production workers at the bottom of the distribution benefit from being matched with better managers, which reduces lower tail wage inequality.

We augment the theory by endogenizing hours worked by the agents. If managers change their hours worked, time available for supervision of production workers changes. This has an effect similar to changes in the required communication time. It thus produces changes in the upper and lower wage inequality. Since distortive taxation changes equilibrium hours worked, it changes the wage distribution too, in addition to the standard effects on earnings. Furthermore, the theory is able to generate a number of additional predictions about the interaction of taxation and organization of firms, span

of control, firm growth and other variables.

We assume that the tax function that the government uses has a constant rate of progressivity as in [Benabou \(2002\)](#). This allows for closed form solutions (see also [Heathcote et al. \(2016\)](#) and [Kapicka \(2018\)](#)). A decrease in tax progressivity increases hours worked of everyone (analogously to a decrease in the communication time), which increases the upper tail wage inequality, but decreases the lower tail wage inequality. We parameterize the model for the U.S. economy and show that, in the optimum, the optimal progressivity wedge reduces from 0.161 to 0.041. This produces a more equal wage distribution at the bottom: the log of the ratio of the 50th to the 10th wage percentile is reduced from 1.069 to 0.988. In contrast, if the wage distribution is exogenous, it is optimal to increase progressivity significantly, from 0.161 to 0.408. In addition, ignoring endogenous wage changes has potentially large welfare losses.

## 2 Related Literature

There are several strands of literature that are related. A large volume of research provides a connection between the wage distribution and taxes through general equilibrium effects. [Meh \(2005\)](#), [Boháček and Zubrický \(2012\)](#) and [Bruggemann \(2017\)](#) follow [Quadrini \(2000\)](#) and [Cagetti and De Nardi \(2006\)](#), and consider tax reforms in Bewley-Aiyagari economies with entrepreneurial activity. Taxes affect workers' wages first through changes in capital accumulation and, second, through endogenous occupational choice. They do not consider optimal taxation, however. Optimal taxation in models with entrepreneurship is considered by [Albanesi \(2011\)](#) and [Shourideh \(2012\)](#) who, however, do not model workers and so there is no occupational choice. [Ales and Sleet \(2016\)](#) study optimal taxation of top CEOs. They assume that higher effort of top earners (CEOs) positively affects the productivity and profits of the firm. However, workers are not explicitly modelled as well, and therefore there is no direct channel through which taxation of top earners would influence the wage schedule of regular workers.

Several recent papers study optimal taxation in models with heterogeneous occupations and endogenous wages. [Rothschild and Scheuer \(2013\)](#) and [Scheuer \(2014\)](#) study optimal taxation in an occupational choice Roy model, [Slavík and Yazici \(2014\)](#) study optimal taxation in a growth model with skilled and unskilled labor and capital-skill complementarity, and [Ales et al. \(2015\)](#) study optimal taxation in a task-to-talent assignment model of the labor market. As in the aforementioned papers, the interaction between different occupations (or workers and entrepreneurs) in these papers is only through general equilibrium effects.

Models in which managers, or entrepreneurs, interact with workers and thus affect their wages directly, are less frequent. [Saez et al. \(2014\)](#) consider a model in which workers' wages are the result of bargaining between workers and CEOs. If top marginal rates are lower, then the CEOs will bargain more aggressively for higher compensation which increases wage dispersion. As a result, endogenous wages (which are a result of compensation bargaining) lead to higher optimal wage progressivity, in contrast to the present paper.

[Ales et al. \(2017\)](#) study a model similar to our model. They build upon [Rosen \(1982\)](#)'s assignment model of talent allocation within a firm and focus on the optimal taxation of top labor incomes. In contrast to our model, the potential impact of taxes on workers' wages is limited. Workers are ex-ante identical, receive the same consumption and their assignment to different managers is indeterminate. We relax the assumptions leading to the assignment indeterminacy, and study the relationship between taxes and wage inequality at both tails of the wage distribution. We are thus able to model both lower-tail and upper-tail income inequality, a key aspect of our paper.

Finally, [Lopez and Torres-Coronado \(2018\)](#) considers a framework very similar to ours, but with inelastic labor supply. They do not study labor income taxation, but instead focus on the role of firm-size-dependent policies.

### 3 Setup

There is a measure one of agents. Agents like to consume, and dislike to work. Their preferences are represented by an additively separable utility function

$$U(c) - V(\ell),$$

where  $c \geq 0$  is consumption,  $\ell \geq 0$  is time spent at work,  $U$  is increasing, concave and differentiable, and  $V$  is increasing, convex and differentiable. We assume that the utility function takes the form

$$U(c) = \log c, \quad V(\ell) = \kappa \frac{1}{1 + \eta} \ell^{1 + \eta}$$

for  $\eta > 0$  being the inverse of Frisch elasticity of labor, and  $\kappa > 0$ .

The technology is similar to [Garicano \(2000\)](#), [Garicano and Rossi-Hansberg \(2006\)](#) or [Geerolf \(2016\)](#). Agents differ in their knowledge,  $z \in [\underline{z}, \bar{z}]$ , exogenously given. The distribution of knowledge is  $G(z)$ , with  $G(\underline{z}) = 0$ ,  $G(\bar{z}) = 1$  and has density function  $g(z)$ . There is a continuum of tasks per period distributed according to  $F(z)$  defined on  $[0, \bar{z}]$ , with a density function  $f(z)$ . An agent with knowledge  $z$  solves all tasks in  $[0, z]$ , and produces  $F(z)$  per unit of time. An agent working  $\ell$  units of time thus produces  $\ell F(z)$ . Note that when  $\underline{z} > 0$ , there is a mass of problems  $F(\underline{z}) > 0$  that every agent knows how to solve, namely  $[0, \underline{z}]$ .

Rather than producing on their own, agents can form teams, where specialized agents (managers) solve harder tasks. We assume two things. First, agents do not know who knows what. The agents thus first try to solve a particular problem by themselves and, if they cannot, ask the manager for help. A worker with knowledge  $z_p$  asks for help with a measure  $1 - F(z_p)$  of tasks. The manager helps them to understand how to solve the problem, if he can, but the problem itself is then solved by the agent. Second, managers spend time  $h$  communicating over measure one of the delegated problems, regardless

of whether they know the answer or not (all the communication costs are incurred by the manager). The way to think about this assumption is that a worker approaches a manager with a problem that the worker cannot solve. The worker explains the problem to the manager at which point the manager incurs the time costs. After the problem has been explained the manager helps the worker solve the problem if he can. A problem needs to be explained to the worker only once; once it has been explained,  $x$  units of time produce  $x$  time what one unit of time produces.

Organizations have a team consisting of production workers and a manager.<sup>1</sup> Let  $z_m$  be the knowledge of the manager, and  $z_p$  be the knowledge of the production workers. Managers are able to advise with tasks easier than  $z_m$ . After receiving advice, workers produce output. The production of the team is  $\ell_p F(z_m)n$ , where  $n$  is the number of production workers and  $\ell_p$  are hours worked by the production workers.<sup>2</sup> The manager faces a time constraint that limits how many production workers he can supervise. If the manager works  $\ell_m$  units of time, his time constraint is

$$hn [1 - F(z_p)] = \ell_m, \quad (1)$$

where  $1 - F(z_p)$  is the measure of tasks that each production worker delegates. Constraint (1) shows one of the key properties of the model. Having production workers with higher knowledge allows the manager to form larger teams and multiply their production. That is, there is a *skill complementarity* between the knowledge of a manager, and knowledge of a worker. Dividing both sides of the constraint (1) by  $\ell_m$  shows that it is only the ratio of the communication time  $h$  and manager's hours  $\ell_m$  that matters for the economy. Both an increase in communication cost and a decrease in available time mean that the manager is able to manage a smaller team of workers. Thus,

**Remark 1.** *A decrease in manager's hours worked  $\ell_m$  is equivalent to an increase in the commu-*

---

<sup>1</sup>We provide a generalization of the model in which organizations consist on multiple levels of management in Appendix C.

<sup>2</sup>For each worker,  $F(z_p)$  problems are solved by the worker himself,  $F(z_m) - F(z_p)$  problems are solved by the manager, and  $1 - F(z_m)$  problems are unsolved.

nication costs  $h$ .

Let  $w(z_p)$  be the hourly wage rate of a production worker with skill  $z_p$ . The manager with knowledge  $z_m$  that employs  $n$  workers with skills  $z_p$  and works  $\ell_m$  hours earns rents  $R$  given by

$$R = n [F(z_m) - w(z_p)] \ell_p = \ell_p \ell_m \frac{F(z_m) - w(z_p)}{h [1 - F(z_p)]}. \quad (2)$$

The manager's rental rate per one hour of work is  $r = R/\ell_m$ . The manager's rent increases linearly with his own hours worked, but also linearly with hours worked of the production worker, because the manager keeps a fraction of output  $F(z_m) - w(z_p)$  from each hour that the worker spends by working. The equation (2) shows a second key complementarity in the model: there is *working time complementarity* between the hours worked of a worker, and hours worked of a manager.

Government taxes individual earnings by a tax function  $T(y)$  regardless of whether the earnings are earned by production workers or managers. We assume that the tax function  $T(y)$  exhibits a constant rate of progressivity (Heathcote et al. (2014), Heathcote et al. (2016), Kapicka (2018)),

$$T(y) = y - \lambda y^{1-\tau},$$

where the wedge  $\tau$  determines the progressivity of the tax system and the level parameter  $\lambda$  of the tax function is chosen in such a way that the government budget constraint holds,

$$\mathbb{E}_y T(y) = G,$$

where  $G$  is government consumption, exogenously given. To simplify notation, we introduce the retention function  $\Gamma(y) = \lambda y^{1-\tau}$  to be the after tax income as a function of pre-tax income.

### 3.1 The Equilibrium

We consider an equilibrium in which agents choose to become managers, production workers or self-employed. If they are self-employed and work  $\ell$  hours, they produce  $F(z)\ell$ . There is also a version of the model in which agents do not have the option of becoming self-employed. Many features of both models are similar. The version without self-employment is studied in Appendix B.

**Assignment.** We start the description of the equilibrium conditions by characterizing the assignment of workers to managers. In equilibrium, agents will sort themselves out into three categories. Agents with skills below  $z_1$  will become production workers. Agents with skills between  $z_1$  and  $z_2$  will be self-employed. Agents with skills above  $z_2$  will become managers. There is an assortative matching, where the worst production worker is matched with the worst manager, and the best production worker is matched with the best manager. Let  $m(z_p)$  be the skill of the manager that employs production workers of skill  $z_p$ . Then the equilibrium assignment satisfies

$$m(\underline{z}) = z_2, \quad m(z_1) = \bar{z}, \quad (3)$$

We do not justify here that the equilibrium assignment takes this form. The reader is referred to [Garicano \(2000\)](#) for such a justification.<sup>3</sup>

Let  $n(z_p)$  be the number of workers employed in a team consisting of production workers with skill  $z_p$  and a manager with skill  $m(z_p)$ . We require that the supply of production workers has to be equal to the demand for production workers by their managers. Equivalently, the demand for managers by their production teams has to be

---

<sup>3</sup>If  $\bar{z}$  is infinite, the second condition should be replaced by  $G(m(z_1)) = 1$ .



equal to the supply of managers:<sup>4</sup>

$$\int_{\underline{z}}^{z_p} \frac{g(t)}{n(t)} dt = \int_{z_0}^{m(z_p)} g(t) dt, \quad \underline{z} \leq z_p \leq z_1, \quad (4)$$

where the left-hand side is the demand for managers by production teams of workers with knowledge less than  $z_p$ <sup>5</sup>, and the right-hand side is the supply of the corresponding managers, who have knowledge between  $z_0$  and  $m(z_p)$ .

**Production workers.** Production workers have only one choice. They face a wage rate  $w(z)$  and choose hours worked  $\ell(z_p)$  to solve

$$\ell(z) \in \arg \max_{\ell} U [\Gamma(\ell w(z))] - V(\ell), \quad \underline{z} \leq z \leq z_1. \quad (5)$$

We also require that the best production worker must be indifferent between being a production worker and being self-employed. Given that an agent simply chooses hours worked given the wage if worker/productivity if self-employed, he will be indifferent if

$$w(z_1) = F(z_1). \quad (6)$$

**Managers.** Managers choose hours worked  $\ell_m$  and production worker skills  $z_p$ . The production worker's skill is chosen to maximize the hourly rental rate  $r = R/\ell_m$ , where the total rents  $R$  are given by (2). Given that a worker  $z_p$  matches with manager  $m(z_p)$ , the optimal choice satisfies

$$m^{-1}(z_m) \in \arg \max_{z \leq z_1} \ell(z) \frac{F(z_m) - w(z)}{h[1 - F(z)]}. \quad (7)$$

---

<sup>4</sup>The intuition for the market clearing condition (4) and the explicit derivation of its marginal version, equation (12) introduced below, are contained in Appendix A.

<sup>5</sup>If the size of the production team is  $n$  then we can think of each worker demanding  $1/n$  of a manager.

Solving the managers' problem (7) yields a differential equation in production workers' wages

$$w'(z_p) = \left[ \frac{f(z_p)}{1 - F(z_p)} + \frac{\ell'(z_p)}{\ell(z_p)} \right] [F(m(z_p)) - w(z_p)]. \quad (8)$$

Together with the boundary conditions (6) and (10), one can solve for the equilibrium relationship between  $m(z)$  and  $w(z)$ .

Let  $r(m(z_p))$  denote the resulting hourly rental rate of a manager  $m(z_p)$ . Having chosen production workers skills, the managers choose hours worked to maximize utility:

$$\ell(z) \in \arg \max_{\ell} U[\Gamma(\ell r(m(z)))] - V(\ell), \quad z_2 \leq z \leq \bar{z}. \quad (9)$$

We require that the worst manager with knowledge  $z_2$  must be indifferent between being a manager and self-employed, i.e.,  $r(z_2) = F(z_2)$ . Rearranging yields

$$w(\underline{z}) = \left[ 1 - \frac{h}{\ell(\underline{z})} (1 - F(\underline{z})) \right] F(z_2). \quad (10)$$

**Definition 1.** *The equilibrium consists of threshold values  $z_1$  and  $z_2$ , matching function  $m : [\underline{z}, z_1] \rightarrow \mathbb{R}_+$ , workers' wage rate  $w : [\underline{z}, z_1] \rightarrow \mathbb{R}_+$ , managers' rental rate  $r : [z_2, \bar{z}] \rightarrow \mathbb{R}_+$ , and hours worked  $\ell : [\underline{z}, \bar{z}] \rightarrow \mathbb{R}_+$  such that  $\ell$  satisfies (5) and (9),  $w$  satisfies (8), (6) and (10),  $m$  satisfies (4), and the government budget constraint holds.*

The equilibrium of this form does not always exist, only if  $h$  is high enough. What is required is that  $z_2 \geq z_1$ . Otherwise there are no self-employed individuals.

### 3.2 Equilibrium wage inequality

The model allows us to easily characterize the wage inequality in the model in several dimensions.

We define the *top wage inequality*  $v^{\text{top}}$  as a wage inequality among managers, i.e. the ratio of the rental rate of the best manager to the rental rate of the worst manager. The rental rate of the best manager  $\bar{z}$  is given by  $r(\bar{z}) = \frac{\ell(z_1)}{h}$ , while the rental rate of the worst manager can be deduced from (10). The top wage inequality equals

$$v^{\text{top}} = \frac{r(\bar{z})}{r(z_2)} = \frac{1}{F(z_2)} \frac{\ell(z_1)}{h} \geq \frac{\ell(z_1)}{h}.$$

The top wage inequality is reduced if the ratio of hours worked  $\ell(z_1)$  and time cost  $h$ , i.e.,  $\ell(z_1)/h$ , decreases. Also, if the threshold value  $z_2$  increases, there will be fewer managers, and the top wage inequality will increase.

The *bottom wage inequality*  $v^{\text{bottom}}$  is defined as the ratio of the wage rate of the best production worker to the wage rate of the worst production worker. The wage rate of the best production worker is given by (6), while the wage rate of the worst production worker can be obtained from (10). The bottom wage inequality is

$$v^{\text{bottom}} = \frac{w(z_1)}{w(\underline{z})} = \frac{F(z_1)}{F(z_2)} \frac{1}{1 - \frac{h}{\ell(\underline{z})}(1 - F(\underline{z}))} \leq \frac{1}{1 - \frac{h}{\ell(\underline{z})}(1 - F(\underline{z}))}.$$

The bottom wage inequality increases if the ratio of hours worked  $\ell(\underline{z})$  and time cost  $h$ , i.e.,  $\ell(\underline{z})/h$ , decreases. If the threshold value  $z_1$  increases, the bottom wage inequality increases, just like the top wage inequality. Note also that  $F(z_1) \leq F(z_2)$ , and so one can now put a simple upper bound on within-group inequality among production workers.

Finally, we define the *overall wage inequality*  $v^{\text{all}}$  as the ratio of the rental rate of the best manager and the wage of the worst production worker. It equals

$$v^{\text{all}} = \frac{r(\bar{z})}{w(\underline{z})} = \frac{1}{F(z_2)} \frac{1}{\frac{h}{\ell(z_1)} \left[ 1 - \frac{h}{\ell(\underline{z})}(1 - F(\underline{z})) \right]} \geq \frac{1}{\frac{h}{\ell(z_1)} \left[ 1 - \frac{h}{\ell(\underline{z})}(1 - F(\underline{z})) \right]}.$$

A decrease in hours worked of the most productive worker  $\ell(z_1)$  and an increase in hours worked of the least productive worker  $\ell(\underline{z})$  decrease overall wage inequality. An increase in time cost  $h$ , or a simultaneous decrease in hours worked  $\ell(z_1)$  and  $\ell(\underline{z})$  (as in

the version of the model we specify below) thus has an ambiguous effect on the overall wage inequality in the economy, which is to be expected, given that they affect upper-tail and lower-tail inequality in the opposite direction. Hours worked now have a direct effect on the overall wage inequality, in contrast to a standard model where wages are exogenous and hours worked only affect earnings inequality. Note also that one can put a lower bound on the overall wage inequality purely in terms of exogenous variables and the ratio  $h/\ell(\underline{z})$ .

### 3.3 Aggregates

Aggregate output in the economy consists of the production of workers and production of self-employed (recall that managers do not directly produce output):

$$Y = \int_{\underline{z}}^{z_1} \ell(t)F(m(t))g(t) dt + \int_{z_1}^{z_2} \ell(t)F(t)g(t) dt.$$

Aggregate consumption in the economy is the sum of total consumption of production workers, of self-employed, and of managers:

$$C = \int_{\underline{z}}^{z_1} \Gamma(w(t)\ell(t))g(t) dt + \int_{z_1}^{z_2} \Gamma(F(t)\ell(t))g(t) dt + \int_{z_2}^{\bar{z}} \Gamma(r(t)\ell(t))g(t) dt.$$

By Walras Law, the requirement that the government budget constraint holds can be expressed as  $C + G = Y$ .

## 4 Characterizing the Equilibrium

We now characterize the equilibrium of the model. First, it is easy to show that, given that the utility is logarithmic in consumption, income and substitution effects cancel out, and the agents choose hours worked that are independent of their skill. Everyone's

hours worked are given by  $\ell(z) = \bar{\ell}(\tau)$  where

$$\bar{\ell}(\tau) = \left( \frac{1 - \tau}{\kappa} \right)^{\frac{1}{1+\eta}}. \quad (11)$$

The fact that hours worked are constant across all agents allows us to substantially simplify the problem. It is only the ratio of communication costs and hours worked that matters in the economy. That is, the wage rate and rent rate schedule satisfy the following property:  $w(z; \ell(\tau), h) = w(z; 1, h/\bar{\ell}(\tau))$  and  $r(z; \ell(\tau), h) = r(z; 1, h/\bar{\ell}(\tau))$ , and the earnings of each agent are  $\bar{\ell}(\tau)$  times wages or rents.<sup>6</sup> We can then normalize hours worked to one, and characterize the equilibrium wage and rent distribution. Any changes in hours worked due to a change in taxes will manifest themselves as a change in the communication costs  $h$ .

Setting  $\bar{\ell}(\tau) = 1$  and differentiating the market clearing condition (4) with respect to  $z_p$  and using the value of  $n$  implied by the time constraint (1) yields a differential equation in  $m$ :

$$m'(z_p) = h [1 - F(z_p)] \frac{g(z_p)}{g(m(z_p))}. \quad (12)$$

Integrating, we find that the matching function  $m$  must satisfy

$$G(m(z_p)) = G(z_p) + h\rho(z_p) \quad (13)$$

where the function  $\rho$ , given by

$$\rho(z_p) = \int_{\underline{z}}^{z_p} [1 - F(z)] g(z) dz,$$

is a function of only exogenous factors, namely the distribution functions  $F$  and  $G$ . It is thus very easy to compute, and so is the function  $m$ . Moreover, the expression (13)

---

<sup>6</sup>The aggregate earnings and consumption can be expressed as  $Y(\bar{\ell}, h) = \bar{\ell}Y(1, h/\bar{\ell})$  and  $C(\bar{\ell}, h) = \bar{\ell}^{1-\tau}C(1, h/\bar{\ell})$ .

shows that the matching function  $m(z)$  directly depends only on the threshold value  $z_2$  and not on  $z_1$ .

The second condition in (3) then yields an equilibrium relationship between both threshold values  $z_1$  and  $z_2$ :

$$1 - G(z_2) = h\rho(z_1), \quad (14)$$

The equation implicitly defines a relationship  $z_2 = \phi_m(z_1)$  that must hold in any equilibrium. The function satisfies  $\phi_m(\underline{z}) = \bar{z}$  and, since  $G' = g > 0$  and  $\rho' > 0$ , the relationship between both threshold values must be negative,  $\phi'_m < 0$ .

A second relationship between both thresholds is obtained from the managers' problem of choosing workers' type. Specializing the differential equation (8) for the case with constant hours worked yields

$$w'(z_p) = \kappa(z_p) [F(m(z_p)) - w(z_p)], \quad (15)$$

where  $\kappa(z_p) = \frac{f(z_p)}{1-F(z_p)}$ . Integrating and using (10), we write the wage function as

$$w(z_p) = [1 - h(1 - F(\underline{z}))] F(z_2) e^{-\int_{\underline{z}}^{z_p} \kappa(z) dz} + e^{-\int_{\underline{z}}^{z_p} \kappa(z) dz} \int_{\underline{z}}^{z_p} e^{\int_{\underline{z}}^z \kappa(\bar{z}) d\bar{z}} \kappa(z) F(\tilde{m}(z, z_2)) dz, \quad (16)$$

where  $\tilde{m}(z, z_2)$  is a solution to (13) for an arbitrary value of  $z_2$ . Using (6) defines a second relationship between both thresholds  $z_1$  and  $z_2$ :

$$F(z_1) = [1 - h(1 - F(\underline{z}))] F(z_2) e^{-\int_{\underline{z}}^{z_1} \kappa(z) dz} + e^{-\int_{\underline{z}}^{z_1} \kappa(z) dz} \int_{\underline{z}}^{z_1} e^{\int_{\underline{z}}^z \kappa(\bar{z}) d\bar{z}} \kappa(z) F(\tilde{m}(z, z_2)) dz.$$

Write  $z_2 = \phi_w(z_1)$ . It is easy to see that  $\phi_w(\underline{z}) < 1$ . Differentiating yields

$$\phi'_w = \frac{e^{\int_{\underline{z}}^{z_1} \kappa(z) dz}}{[1 - h(1 - F(\underline{z}))] f(z_2) + \int_{\underline{z}}^{z_1} e^{\int_{\underline{z}}^z \kappa(\bar{z}) d\bar{z}} \kappa(z) F(\tilde{m}(z, z_2)) dz} \kappa(z_1) [1 - F(\tilde{m}(z_1, z_2))].$$

Since the first two terms are positive, the sign is determined by the last term. The derivative  $\phi'_w$  is then positive if  $F(\tilde{m}(z_1, z_2)) < 1$ , i.e. if  $z_2 < \phi_m(z_1)$ , and is zero otherwise. The function  $\phi_w$  thus increases and reaches maximum along the  $z_2 = \phi_m(z_1)$  lines. The intersection of functions  $\phi_m$  and  $\phi_w$  determines the equilibrium values of  $z_1$  and  $z_2$ . Figure 1 illustrates.

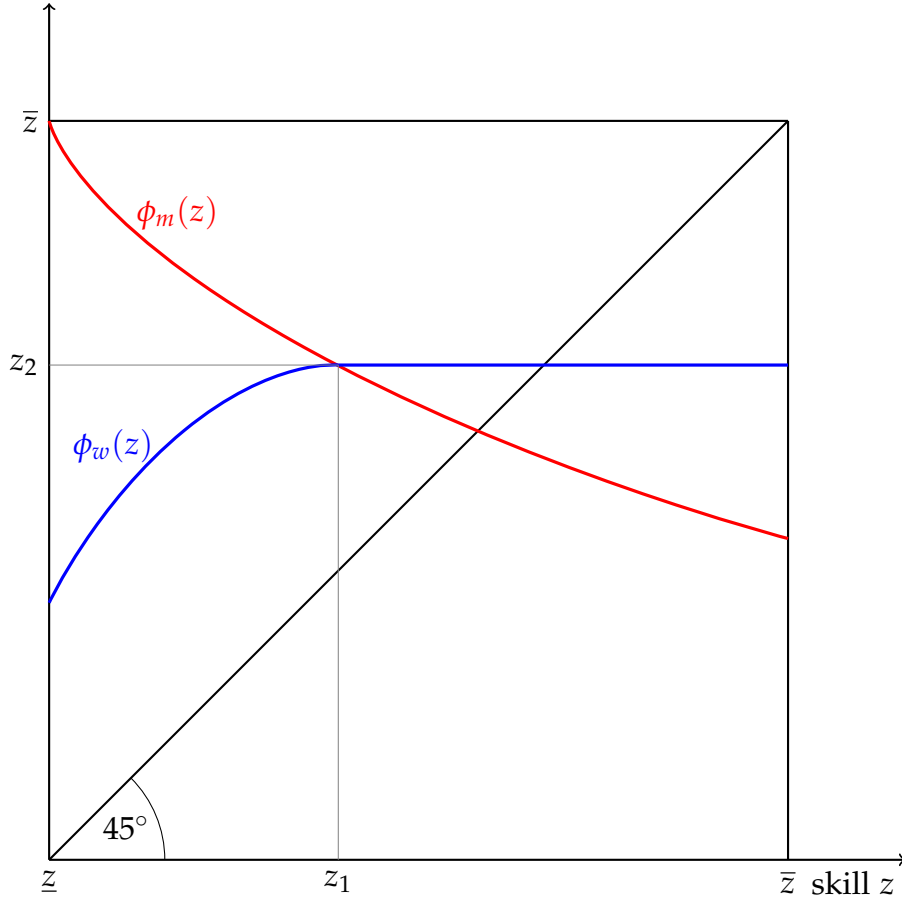


Figure 1: Functions  $\phi_m(z)$  and  $\phi_w(z)$ , and the determination of the equilibrium threshold values  $z_1$  and  $z_2$ .

#### 4.1 An Example

We assume that the underlying distributions  $F$  and  $G$  are both uniform, with  $\bar{z} = 1$ ,  $\underline{z} = 1 - \Delta$ . This configuration delivers a distribution of wages with a Pareto right tail,

see [Geerolf \(2016\)](#). The cumulative distribution functions are

$$F(z) = z, \quad G(z) = \frac{z + \Delta - 1}{\Delta}.$$

All agents thus knows problems with difficulty  $1 - \Delta$  or lower. The problem then has a closed form solution. The function  $\rho$  is given by

$$\rho(z) = \frac{1}{\Delta} \left( \frac{\Delta^2}{2} - \frac{(1-z)^2}{2} \right).$$

Equation (13) implies that the matching function is quadratic in  $z$ :

$$m(z) = z_2 + \frac{h}{2}\Delta^2 - \frac{h}{2}(1-z)^2.$$

Then condition (14) defines the first relationship between both thresholds:

$$\phi_m(z_1) = 1 - \frac{h}{2}\Delta^2 + \frac{h}{2}(1-z_1)^2,$$

with  $\phi'_m = -h(1-z_1)$ . Turning to the wage structure, equation (16) can be simplified substantially and produces a wage function that is again quadratic in the ability  $z$ :

$$w(z) = \frac{h}{2}(1-z)^2 - (1-z)h(\Delta + z_2) + z_2 + \frac{h}{2}\Delta^2.$$

Using (6), which can now be written as  $w(z_1) = z_1$ , implies a second relationship between both thresholds:

$$\phi_w(z_1) = \frac{(1-h\Delta)z_1 - \frac{h}{2}(1-z_1)^2 + h\Delta - \frac{h}{2}\Delta^2}{1 - (1-z_1)h}$$

Setting  $z_2 = \phi_m(z_1) = \phi_w(z_1)$  and simplifying yields a pair of equations in  $z_1$  and  $z_2$   $1 - z_2 = \frac{h}{2}\Delta^2 - \frac{h}{2}(1-z_1)^2$  and  $1 - h\Delta = hz_2 - h(1-z_1)$ . Solving, we get the a closed-



form solution for the equilibrium values of  $z_1$  and  $z_2$

$$z_1 = 1 - \frac{1}{h} + \frac{1}{h} \sqrt{(1 - h\Delta)^2 + 2(1 - h)}, \quad z_2 = \frac{2}{h} - \Delta - \frac{1}{h} \sqrt{(1 - h\Delta)^2 + 2(1 - h)}.$$

**Aggregates** Aggregate output and consumption are

$$Y = \frac{1}{\Delta} \left[ \int_{1-\Delta}^{z_1} m(t) dt + \int_{z_1}^{z_2} t dt \right]$$

$$C = \frac{\lambda}{\Delta} \left[ \int_{1-\Delta}^{z_1} w(t)^{1-\tau} dt + \int_{z_1}^{z_2} t^{1-\tau} dt + \int_{z_2}^1 r(t)^{1-\tau} dt \right].$$

Note that the last integral in the brackets can also be written as

$$\int_{z_2}^1 r(t)^{1-\tau} dt = h^\tau \int_{1-\Delta}^{z_1} [m(t) - w(t)]^{1-\tau} (1-t)^\tau dt$$

**Existence of equilibrium.** Equilibrium of this type exists if  $z_1 \leq z_2$ . Rearranging and simplifying, this is equivalent to

$$h \geq \frac{-1 - \Delta + 2\sqrt{(1 + \Delta)^2 - 3\Delta^2}}{(1 + \Delta)^2 - 4\Delta^2}.$$

Note that if  $\Delta = 1$ , the condition reduces to  $h \geq 3/4$ . In general, the minimum condition on  $h$  is nonmonotone, converges to 1 as  $\Delta$  converges to zero, and numerical simulations show that it does go much below 3/4.

If the condition is not satisfied then there are no self employed individuals. There is only one threshold value  $z_0$  that separates production workers and managers. The boundary conditions become

$$m(\underline{z}) = z_0, \quad m(z_0) = 1, \quad z_0 - w(\underline{z}) = h\Delta w(z_0).$$

The matching function  $m$  is again a quadratic function in  $z$ ,

$$1 - m(z) = m(z_0) - m(z) = \frac{h}{2} (1 - z)^2 - \frac{h}{2} (1 - z_0)^2,$$

where the first equality follows from the boundary condition  $m(z_0) = 1$ . Hence

$$m(z) = 1 + \frac{h}{2} (1 - z_0)^2 - \frac{h}{2} (1 - z)^2.$$

Solving for  $z_0$ , we get that  $z_0$  satisfies

$$z_0 = 1 + \frac{1}{h} - \frac{1}{h} \sqrt{1 + h^2 \Delta^2}.$$

Turning to the wage equation, conjecture that wage is likewise a quadratic function of knowledge:  $w(z) = A + B(1 - z) + \frac{C}{2}(1 - z)^2$ . Substituting into the differential equation for  $w$  together with the expression for  $m$  and equating coefficients yields  $A = 1 + \frac{h}{2}(1 - z_0)^2$ ,  $C = h$ , while  $B$  is not determined. To obtain  $B$ , use the boundary conditions to obtain  $B = h \frac{2(1 - z_0) - (1 + \Delta + h\Delta^2)}{1 + h(1 - z_0)}$ .

**Span of control.** We now characterize the distribution of the span of control, i.e. team size,  $n$ . The minimum value is  $\underline{n} = n(\underline{z}) = \frac{1}{h\Delta}$  and the maximum value is  $\bar{n} = n(1) = \frac{1}{h(1 - z_1)}$ . Thus, the span of control is distributed on  $[\underline{n}, \bar{n}]$ . A manager with skill  $z_m \in [z_1, 1]$  has a team size

$$n(z_m) = \frac{1}{h[1 - m^{-1}(z_m)]}.$$

Inverting, given  $n \in [\underline{n}, \bar{n}]$ , we express the manager's skill as  $z_m(n) = m \left(1 - \frac{1}{hn}\right)$ . The equilibrium distribution is

$$\begin{aligned} \text{Prob}(\tilde{n} \geq n) &= \text{Prob} \left[ z_m \geq m \left(1 - \frac{1}{hn}\right) \mid z_m \geq z_2 \right] \\ &= \frac{1}{1 - z_2} \left[ 1 - m \left(1 - \frac{1}{hn}\right) \right] \\ &= \frac{1}{1 - z_2} \left[ \frac{h}{2} \frac{1}{(hn)^2} - \frac{h}{2} (1 - z_1)^2 \right] \\ &= \frac{1}{2h(1 - z_2)} \left( n^{-2} - \bar{n}^{-2} \right). \end{aligned}$$

Let  $P_n(n) = 1 - \text{Prob}(\tilde{n} \geq n)$  be the distribution of team size. One can easily verify that  $P_n(\underline{n}) = 0$  and  $P_n(\bar{n}) = 1$ . The team size thus has a Pareto distribution with a Pareto parameter equal to 2. The density is

$$p_n(n) = \frac{1}{h(1 - z_2)} n^{-3}.$$

**Distribution of profits.** We can also characterize the manager's profits. The profit function is

$$R(m(z)) = \frac{1}{h} \frac{m(z) - w(z)}{1 - z} = \frac{1}{h} + z - z_1, \quad (17)$$

and so

$$R(z_m) = \left( \frac{1}{h} + m^{-1}(z_m) - z_1 \right) = \frac{1}{h} - z_1 + 1 - \sqrt{\frac{2}{h}(1 - z_m) + (1 - z_1)^2}, \quad (18)$$

where we have used  $m^{-1}(m) = 1 - \sqrt{\frac{2}{h}(1 - m) + (1 - z_1)^2}$ . We have  $\underline{R} = R(z_2) = \frac{1}{h} + \underline{z} - z_1$  and  $\bar{R} = R(\bar{z}) = \frac{1}{h}$ , and  $R \in [\underline{R}, \bar{R}]$ . Inverting the expression above,  $z_m(R) =$

$m(R + z_1 - 1/h)$  for  $R \in [\underline{R}, \bar{R}]$ , and we have

$$\begin{aligned} \text{Prob}(\tilde{R} \geq R) &= \text{Prob} \left[ z_m \geq m \left( R + z_1 - \frac{1}{h} \right) \mid z_m \geq z_2 \right] \\ &= \frac{1}{1 - z_2} \left[ 1 - m \left( R + z_1 - \frac{1}{h} \right) \right] \\ &= \frac{1}{1 - z_2} \left[ \frac{h}{2} \left( 1 - R - z_1 + \frac{1}{h} \right)^2 - \frac{h}{2} (1 - z_1)^2 \right] \end{aligned}$$

Let  $P_R(R) = 1 - \text{Prob}(\tilde{R} \geq R)$  be the distribution of profits. One can easily verify that  $P_R(\underline{R}) = 0$  and  $P_R(\bar{R}) = 1$ . The density is

$$p_R(R) = \frac{h}{1 - z_2} \left( 1 - R - z_1 + \frac{1}{h} \right).$$

**Aggregates**

$$Y = \frac{1}{\Delta} \int_{\underline{z}}^{z_1} m(t) dt + \frac{1}{\Delta} \int_{z_1}^{z_2} t dt.$$

**Calibration** Suppose that we have targets for fractions of the population  $G_1 = G(z_1)$  and  $G_2 = G(z_2)$ . Then we can calibrate the parameters  $h$  and  $\Delta$  as follows. Since  $z = \Delta G + 1 - \Delta$ , we can transform the equilibrium equations for  $z_1$  and  $z_2$  as

$$h\Delta = 2 \frac{1 - G_2}{1 - (1 - G_1)^2}, \quad h = 1 - h\Delta(G_1 + G_2 - 1).$$

The first can be solved for  $h\Delta$  and, given  $h\Delta$ , the second one can be solved for  $h$ .

## 5 Optimal progressivity

We now characterize the optimal value of the progressivity parameter  $\tau$ , and its determinants. Let  $\mathbb{E}[w|\tau] = Y[1, h/\ell(\tau)]$  be the average wage and rental rate, and  $\mathbb{E}[w^{1-\tau}|\tau] = C[1, h/\ell(\tau)]/\lambda$  be the average of wages and rents to the power  $1 - \tau$ . Putting back labor

supply  $\bar{\ell}(\tau)$  given by (11), we can write the resource constraint as

$$\lambda \bar{\ell}(\tau)^{1-\tau} \mathbb{E}[w^{1-\tau}|\tau] + G = \bar{\ell}(\tau) \mathbb{E}[w|\tau].$$

Solving for the equilibrium  $\lambda$  and substituting back to the expected utility yields the welfare for a given progressivity wedge  $\tau$

$$\mathcal{W} = \ln(\bar{\ell}(\tau) \mathbb{E}[w|\tau] - G) - \frac{1-\tau}{1+\eta} - \ln \mathbb{E}[w^{1-\tau}|\tau] + (1-\tau) \mathbb{E}[\ln w|\tau]. \quad (19)$$

The expression has a standard form, but the moments of the wage distribution are not exogenous, but depend on  $\tau$ .

**Model parameters.** In parameterizing the model, we set  $\kappa = 1$  and  $\eta = 2$ , producing a Frisch elasticity of labor of 0.5. We also set  $\Delta = 1$ . The progressivity of the current U.S. tax system can be approximated by  $\tau = 0.161$  as estimated by [Heathcote et al. \(2016\)](#) (see also [Guner et al. \(2014\)](#) for estimates of this and other tax functions). We choose the benchmark value of  $h$  so that the variance of log wages in the benchmark economy is 0.43. The resulting benchmark parameters are in [Table 1](#).

$\eta$	$\kappa$	$\Delta$	$h$	$\tau$
2.000	1.000	1.000	0.813	0.161

**Results.** We characterize the optimum for a case when there is no self-employment, and  $z_1 = z_2 \equiv z_0$ . The blue line in [Figure 2](#) plots the resulting distribution of wages in the benchmark economy. It differs significantly from the underlying distribution of abilities. Agents at both ends of the distribution gain significantly. Very low ability agents gain from being matched with a manager that solves some of their problems. Very high ability agents gain from being able to supervise a relatively large number of

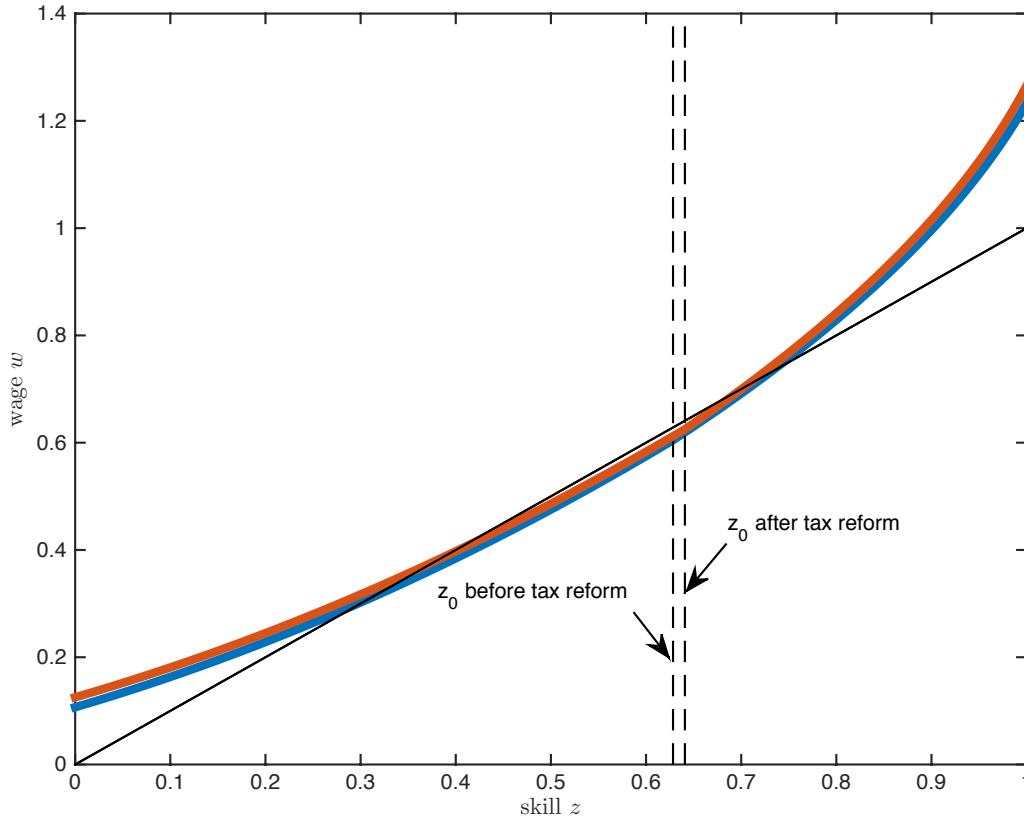


Figure 2: Distribution of wages, before tax changes (blue line) after tax changes (red line) and the distribution of abilities (black line)

production workers. Note that agents in the middle of the distribution lose relatively to autarchy where they only produce  $z$ . That's a consequence of the fact that this version of the model does not allow for self employment.<sup>7</sup> Table 5 shows the 90/50 and 50/10 ratio in the benchmark economy.

Consider now a change in the tax progressivity  $\tau$ . Start with a reform that counterfactually assumes that the distribution of wages is exogenous, and does not respond to changes in the tax system. The blue line in Figure 3 shows welfare as a function of the progressivity wedge  $\tau$ . The (incorrectly measured) welfare is maximized at a very high rate of progressivity  $\tau = 0.408$ . That is, if the wage distribution is exogenous, it is optimal to increase progressivity of the tax system significantly. The red line in Figure 3,

<sup>7</sup>We consider an extension with self-employment in the Appendix.

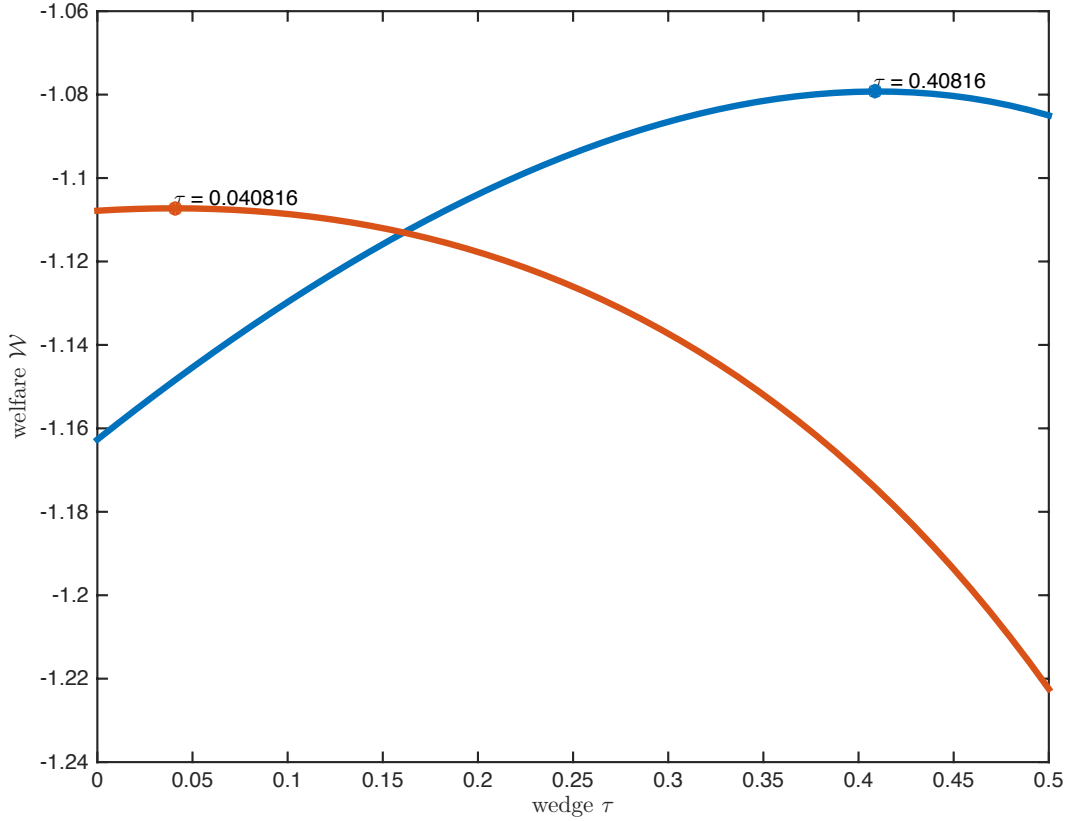


Figure 3: Welfare as a function of the progressivity wedge  $\tau$ , with endogenous wage distribution (red line) and exogenous wage distribution (blue line)

on the other hand, shows the welfare (19) when the distribution of wages is endogenous (both lines cross at the benchmark level of  $\tau$ ). The optimal level of progressivity is now 0.041, significantly lower than the benchmark value of  $\tau$ . The normative predictions are thus completely different.

The reason why the optimal progressivity is so low lies in the fact that wages adjust to changes in progressivity. In particular, a decrease in progressivity increases hours worked, which gives managers more time to supervise production workers (increase in hours worked is equivalent to a decrease in the communication time  $h$ ). This increases inequality at the very top, but decreases inequality at the bottom, as can be seen from the position of the red line in Figure 2 relative to the blue line. The decrease of inequality at the bottom is in particular reflected in higher wages for the least able agents. A decrease

in tax progressivity is thus endogenously counteracted by a less unequal distribution of wages at the bottom. Table 5 shows that the log of 50/10 ratio decreases significantly, from 1.069 to 0.988. It is the bottom of the distribution, not the top of it, that is critical for the welfare in the economy. The optimal progressivity then decreases relatively to a model with exogenous wages. Note also from Figure 3 that implementing a naive optimum with  $\tau = 0.408$  would lead to substantial welfare losses relative to the benchmark. The forces now work in the opposite directions, and the distribution of wages becomes more unequal at the bottom.

Table 2: Changes in the Wage Distribution

	Benchmark	Optimum
Variance log wages	0.430	0.388
log 90/50 ratio	0.740	0.737
log 50/10 ratio	1.069	0.988

## 6 Conclusions

In this paper, we study the effects of taxation in a model with knowledge based hierarchies. In the model, agents self-select into being workers, self-employed or managers based on their ability to solve tasks. Individual labor supply is endogeneous leading to important interactions between taxes and wage inequality. If taxes become more progressive, managers work less which decreases their wages but also the wages of their employees (workers). Using a parametric tax function, we provide a quantitative example in which these considerations lead a benevolent utilitarian government to use a less progressive tax schedule than the one currently in place in the United States. We leave the task of studying optimal taxation with more flexible tax functions (including the Mirrleesian approach of placing no ad-hoc restrictions on the labor income tax schedule) as a fruitful avenue for future research.



## References

- Albanesi, S. (2011). Optimal taxation of entrepreneurial capital with private information. Working paper, Columbia University. 3
- Ales, L., A. Bellofatto, and J. Wang (2017). Taxing atlas: Executive compensation, firm size, and their impact on optimal top income tax rates. *Review of Economic Dynamics* 26, 62–90. 4
- Ales, L., M. Kurnaz, and C. Sleet (2015). Technical change, wage inequality, and taxes. *American Economic Review* 105, 3061–3101. 4
- Ales, L. and C. Sleet (2016). Taxing top CEO. *American Economic Review* 106, 3331–3306. 3
- Benabou, R. (2002, March). Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency? *Econometrica* 70(2), 481–517. 3
- Boháček, R. and J. Zubrický (2012). A flat tax reform in an economy with occupational choice and financial frictions. *The Economic Journal* 122, 1313–1345. 3
- Bruggemann, B. (2017). Higher taxes at the top: The role of entrepreneurs. Technical report, Goethe University Frankfurt. 3
- Cagetti, M. and M.-C. De Nardi (2006). Entrepreneurship, frictions and wealth. *Journal of Political Economy* 114, 835–870. 3
- Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy* 108, 874–904. 2, 5, 8
- Garicano, L. and E. Rossi-Hansberg (2006). Organization and inequality in a knowledge economy. *Quarterly Journal of Economics* 121, 383–435. 2, 5

- Garicano, L. and E. Rossi-Hansberg (2011). Knowledge-based hierarchies: Using organizations to understand the economy. *Annual Review of Economics* 7, 1–30. 2
- Geerolf, F. (2016). A theory of pareto distributions. Technical report, UCLA. 5, 16
- Guner, N., R. Kaygusuz, and G. Ventura (2014, October). Income Taxation of U.S. Households: Facts and Parametric Estimates. *Review of Economic Dynamics* 17(4), 559–581. 21
- Heathcote, J., K. Storesletten, and G. Violante (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review* 104(7), 2075–2126. 7
- Heathcote, J., K. Storesletten, and G. Violante (2016). Optimal tax progressivity: An analytical framework. Working paper, Federal Reserve Bank of Minneapolis. 2, 3, 7, 21
- Kapicka, M. (2018). Quantifying the welfare gains from history dependent income taxation. Working paper. 3, 7
- Lopez, J. J. and J. Torres-Coronado (2018). Size-dependent policies, talent misallocation and the return to skill. Technical report, University of Houston. 4
- Meh, C. (2005). Entrepreneurship, wealth inequality, and taxation. *Review of Economic Dynamics* 8, 688–719. 3
- Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies* 38(2), 175–208. 2
- Quadrini, V. (2000). Entrepreneurship, saving and social mobility. *Review of Economic Dynamics* 3, 1–40. 3
- Rosen, S. (1982). Authority, control, and the distribution of earnings. *Bell Journal of Economics* 13(2), 311–323. 4

- Rothschild, C. and F. Scheuer (2013). Redistributive taxation in the Roy model. *Quarterly Journal of Economics* 128, 623–668. 4
- Saez, E., S. Stancheva, and T. Piketty (2014). Optimal taxation of top labor incomes: A tale of three elasticities. Technical report. 4
- Scheuer, F. (2014). Entrepreneurial taxation with endogenous entry. *AEJ: Economic Policy* 6, 126–163. 4
- Shourideh, A. (2012). Optimal taxation of wealthy individuals. Working paper, University of Pennsylvania. 3
- Slavík, C. and H. Yazici (2014). Machines, buildings, and optimal dynamic taxes. *Journal of Monetary Economics* 66, 47 – 61. 4

# Appendix

## A Market clearing conditions (4) and (12)

Consider first a situation in which there is a fixed number of workers per one manager, say one, i.e.  $n(t) = 1, \forall t \in [\underline{z}, z_p]$ . This means that the threshold  $z_1$  and the matching function  $m$  must be such that  $\forall z_p$ :

$$G(z_p) = G(m(z_p)) - G(z_1), \quad (20)$$

where the LHS represents the supply of workers with knowledge less or equal to  $z_p$  and the RHS the demand for these workers. This equation is almost equivalent to equation (4):

$$\int_{\underline{z}}^{z_p} g(t) dt = \int_{z_2}^{m(z_p)} g(t) dt, \quad \underline{z} \leq z_p \leq z_1. \quad (21)$$

If the number of workers of type  $z$  matched with the appropriate manager is  $n(z)$ , then we arrive at equation (4):

$$\int_{\underline{z}}^{z_p} \frac{g(t)}{n(t)} dt = \int_{z_2}^{m(z_p)} g(t) dt, \quad \underline{z} \leq z_p \leq z_1, \quad (22)$$

where once again, the LHS represents the supply of workers (per manager) and the RHS the demand for workers (per manager). Differentiating w.r.t  $z_p$  we arrive at (using the Leibnitz rule):

$$\frac{g(z_p)}{n(z_p)} = g(m(z_p))m'(z_p). \quad (23)$$

Once we plug in for  $n(z_p)$  we arrive at equation (12).

Perhaps a more intuitive way of writing the market clearing condition is:

$$\int_{\underline{z}}^{z_p} g(t) dt = \int_{z_2}^{m(z_p)} g(t) n(m^{-1}(t)) dt, \quad \underline{z} \leq z_p \leq z_1. \quad (24)$$

where, as above, the LHS represents the supply of workers with knowledge less or equal to  $z_p$  and the RHS the demand for these workers. Differentiating w.r.t  $z_p$  we get, as before:

$$g(z_p) = g(m(z_p)) n(z_p) m'(z_p). \quad (25)$$

Finally, notice that the masses of agents of each particular type are 0, which is why it does not make sense to require that the mass of type  $z_p$  equals the mass of type  $m(z_p)$  multiplied by  $n(z_p)$ . However, the market clearing condition must hold for any positive measure:

$$\int_{z_0}^{z_p} g(t) dt = \int_{m(z_0)}^{m(z_p)} g(t) n(m^{-1}(t)) dt, \quad \underline{z} \leq z_0 < z_p \leq z_1. \quad (26)$$

## B A version without self-employment

In equilibrium, the agent will sort themselves out into production workers and managers. There is a threshold level of skill  $z_0$  such that agents with skills below  $z_0$  will become production workers and agents with skills above  $z_0$  will become managers.

Let  $m(z_p)$  be the skill of the manager that employs production workers of skill  $z_p$  and  $n(z_p)$  be the number of workers employed. In equilibrium, the supply of production workers has to be equal to the demand for production workers:

$$\int_{\underline{z}}^{z_p} \frac{g(t)}{n(t)} dt = \int_{z_0}^{m(z_p)} g(t) dt, \quad \underline{z} \leq z_p \leq z_1,$$

where the left-hand side is the demand for managers by production workers with knowl-

edge less than  $z_p$  per manager, and the right-hand side is the supply of the corresponding managers, who have knowledge between  $z_0$  and  $m(z_p)$ . Differentiating with respect to  $z_p$  and using the expression for  $n$  again yields (12). We also require that the worst manager is matched with the worst worker, and that the best manager is matched with the best production worker:

$$m(\underline{z}) = z_0, \quad m(z_0) = \bar{z}. \quad (27)$$

The optimal choice of managers is still characterized by (8) and (2). However, the boundary conditions are different. We require that the worst manager with threshold knowledge  $z_0$  is indifferent between being a manager and self-employed. The worst manager earns a wage  $r(z_0)$ . If he became the production worker, he would have earned a wage  $w(z_0)$ . Since he supplies the same labor in both cases, he will be indifferent if and only if the wages in both sectors are equalized,  $r(z_0) = w(z_0)$ . Rearranging, the indifference condition becomes

$$F(z_0) - w(\underline{z}) = \frac{h}{\ell(\underline{z})} w(z_0) [1 - F(\underline{z})]. \quad (28)$$

**Equilibrium wage equality.** The hourly rents of the worst manager  $z_0$  are  $r(z_0) = w(z_0)$ , and the rents of the best manager  $\bar{z}$  are given by  $r(\bar{z}) = \ell(z_0) \frac{1-w(z_0)}{h[1-F(z_0)]}$ . The overall earnings inequality is

$$v^{\text{all}} = \frac{r(\bar{z})}{w(\underline{z})} = \frac{\ell(z_0)}{h} \frac{1 - w(z_0)}{1 - F(z_0)} \frac{1}{w(\underline{z})}.$$

The wage inequality will be reduced if the wage rate  $w(z_1)$  increases, since it will make the opportunity cost for the worst managers higher. This is due to a direct effect on the denominator, but there is also a magnifying effect that reduces wages at the top. The wage inequality among production workers is

$$v^{\text{bottom}} = \frac{w(z_0)}{w(\underline{z})} = \frac{\ell(\underline{z})}{h} \frac{F(z_0) - w(\underline{z})}{1 - F(\underline{z})} \frac{1}{w(\underline{z})}.$$

It is also reduced if the wage rate at the bottom increases due to the direct effect on the denominator, but the effect will be magnified by a reduction of wage  $w(z_1)$  (keeping other things equal).

Finally, the wage inequality among managers (the top earnings inequality) is

$$v^{\text{top}} = \frac{r(\bar{z})}{w(z_0)} = \frac{\ell(z_0)}{h} \frac{1 - w(z_0)}{1 - F(z_0)} \frac{1}{w(z_0)}.$$

### Closed form solutions with uniform distributions

Suppose again that the underlying distributions are uniform as before. Then the differential equations (8) and (12) become

$$m'(z) = h(1 - z), \quad w'(z) = \frac{m(z) - w(z)}{1 - z},$$

and the boundary conditions become

$$m(\underline{z}) = z_0, \quad m(z_0) = 1, \quad z_0 - w(\underline{z}) = h\Delta w(z_0).$$

The matching function  $m$  is again a quadratic function in  $z$ ,

$$1 - m(z) = m(z_0) - m(z) = \frac{h}{2}(1 - z)^2 - \frac{h}{2}(1 - z_0)^2,$$

where the first equality follows from the boundary condition  $m(z_0) = 1$ . Hence

$$m(z) = 1 + \frac{h}{2}(1 - z_0)^2 - \frac{h}{2}(1 - z)^2.$$

Solving for  $z_0$ , we get that  $z_0$  satisfies

$$z_0 = 1 + \frac{1}{h} - \frac{1}{h}\sqrt{1 + h^2\Delta^2}.$$

It is easy to verify that

**Lemma 1.**  $z_0$  is decreasing in  $h$ .

A decrease in the communication costs  $h$  thus decreases the fraction of individuals that are managers.

Turning to the wage equation, conjecture that wage is likewise a quadratic function of knowledge:  $w(z) = A + B(1 - z) + \frac{C}{2}(1 - z)^2$ . Substituting into the differential equation for  $w$  together with the expression for  $m$  and equating coefficients yields  $A = 1 + \frac{h}{2}(1 - z_0)^2$ ,  $C = h$ , while  $B$  is not determined. To obtain  $B$ , use the boundary conditions to obtain  $B = h \frac{2(1 - z_0) - (1 + \Delta + h\Delta^2)}{1 + h(1 - z_0)}$ .

**Existence of equilibrium.** The equilibrium exists if  $0 \leq z_0 \leq 1$ . It is easy to see from the expression for  $z_0$  that the equilibrium exists for all values of  $h$ .

## C Extension: An economy with $L$ layers of management

We now generalize the model to allow for organization with  $L \geq 1$  layers of management. We again set  $\bar{\ell} = 1$ . The set  $[\underline{z}, \bar{z}]$  is now partitioned into  $L$  connected subsets, separated by  $L$  thresholds  $z_0, z_1, \dots, z_{L-1}$ . Agents with knowledge in  $[\underline{z}, z_0]$  are production workers. Agents with knowledge in  $(z_0, z_1]$  are first level managers, agents with knowledge in  $(z_{l-1}, z_l]$  are managers of level  $l$ . Managers of level  $L$ , who are at the top of the hierarchy, have knowledge  $(z_{L-1}, \bar{z}]$ .

Consider an organization with  $n_0$  production workers with skill  $x_0$ . They solve a fraction  $F(x_0)$  of problems, and pass the remaining ones to level one managers. Level one managers with skill  $x_1$  are thus asked to solve  $n_0[1 - F(x_0)]$  unsolved problems, and they need  $hn_0[1 - F(x_0)]$  time for that. Since each manager has one unit of time, the organization needs  $n_1 = hn_0[1 - F(x_0)]$  managers of level one. The managers of level one solve problems easier than  $x_0$ , and thus pass  $n_0[1 - F(x_1)]$  problems to a manager of level two. Thus, the organization needs  $n_2 = hn_0[1 - F(x_1)]$  level-2 managers. The unsolved problems are passed up higher in the hierarchy, up to the highest level  $L$ . We assume that



there is only one top manager, and so his time constraint is  $1 = hn_0[1 - F(x_{L-1})]$ . The top level manager solves problems easier than  $x_L$ , and the production of the organization is  $n_0F(x_L)$ . The organization thus has a pyramidal structure with  $n_l \geq n_{l+1}$ .

The payoff of a top manager with knowledge  $x$  that employs a subordinate manager with skill  $y$  are

$$\begin{aligned} w(x) &= \max_{y_0, \dots, y_{L-1}} F(x)n_0 - w(y_0)n_0 - w(y_1)n_0n_1 \dots - w(y_{L-1})n_0n_1 \times n_{L-1} \\ &= \frac{F(x) - w(y_0) - h[1 - F(y_0)]w(y_1) \dots - h[1 - F(y_{L-2})]w(y_{L-1})}{h[1 - F(y_{L-1})]}, \end{aligned} \quad (29)$$

where  $w(y)$  is the equilibrium wage function.

## C.1 The Equilibrium

Define  $m(x)$  for  $x \leq z_{L-1}$  to be the skill of the manager that employs a subordinate with skill  $x$  (either a lower level manager, or a production worker). We extend the function on the whole space by defining  $m(x) = x$  for  $x \geq z_{L-1}$ . Similarly, let  $n(x)$  for  $x \geq z_0$  be the number of direct subordinates of a manager with skill  $x$ . We again extend the function on the whole domain by defining  $n(x) = 1$  for  $x \leq z_0$ . In equilibrium, the supply of production workers has to be equal to the demand for production workers:

$$\int_{\underline{z}}^x g(t) dt = \int_{z_2}^{m(x)} \frac{n(t)}{n(m(t))} g(t) dt, \quad x \in [\underline{z}, z_{L-1}].$$

The left hand side is the supply of subordinate workers (either production workers or managers) with knowledge less than  $x$ , while the right-hand side is the demand by the corresponding managers, who have knowledge between  $z_0$  and  $m(x)$ . They hire  $n(t)$  subordinates, and there is  $n(m(t))$  of them. Differentiating with respect to  $x$  yields

$$m'(x) = \frac{n(m(m(x)))}{n(m(x))} \frac{g(x)}{g(m(x))}$$

Since  $\frac{n_{l+1}}{n_l} = \frac{1-F(x_l)}{1-F(x_{l-1})}$  for  $l > 0$  and  $\frac{n_1}{n_0} = h[1 - F(x_0)]$ , we write

$$m'(x) = h[1 - F(x)] \frac{g(x)}{g(m(x))} \quad x \leq z_0 \quad (30)$$

$$= \frac{1 - F(x)}{1 - F(m^{-1}(x))} \frac{g(x)}{g(m(x))} \quad x \geq z_0. \quad (31)$$

## D An Economy with Exponential Distribution

Suppose that both distributions are exponential with the same shape parameter  $\lambda$ , but a different location parameter:

$$F(z) = 1 - e^{-\lambda z}, \quad z \geq 0$$

$$G(z) = 1 - e^{-\lambda(z-\underline{z})}, \quad z \geq \underline{z}$$

We will make the following transformation. Let  $G = 1 - e^{-\lambda(z-\underline{z})}$  be the cdf of each skill and  $\hat{m} = e^{-\lambda(m(z)-\underline{z})}$  be a transformation of  $m$ . Let also  $\delta = e^{-\lambda\underline{z}}$ . Then one can conveniently write the problem in terms of  $G \in [0, 1]$  and  $\delta$ . We have

$$\begin{aligned} \rho(G) &= \frac{\delta}{2} \left( 1 - (1 - G)^2 \right) \\ \hat{m}(G) &= \frac{h\delta}{2} \left[ (1 - G)^2 - (1 - G_1)^2 \right] \\ w(G) &= \frac{h\delta^2}{2} G^2 + h\delta [1 - \delta(1 - G_2)] G + [1 - \delta(1 - G_2)] (1 - h\delta) \end{aligned}$$

The thresholds are as follows.  $G_1$  solves a quadratic equation

$$h\delta G_1^2 + 2(1 - h\delta)G_1 - 1 \frac{1 - h}{h\delta} = 0,$$

and  $G_2$  solves

$$G_2 = 1 - \frac{h\delta}{2} \left[ 1 - (1 - G_1)^2 \right]$$

The wage indifference conditions can be expressed as  $w(0) = (1 - h\delta) [1 - \delta(1 - G_2)]$  and  $w(G_1) = 1 - \delta(1 - G_1)$ . The resource clearing conditions are  $\hat{m}(G_1) = 0$  and  $\hat{m}(0) = 1 - G_2$ .

The wage of self-employed can be written as

$$F(G_s) = \delta G_s + 1 - \delta$$

The wage of managers can be, after some algebra written (as a function of the corresponding rank of the production worker)

$$r(G_p) = 1 - \delta + \delta G_2 + \delta G_p$$

To express it as a function of the managers' rank  $G_m$ , note that  $G_m(G_p) = 1 - \hat{m}(G_p)$ . Inverting,  $G_p(G_m) = \hat{m}^{-1}(1 - G_m)$ . Since  $\hat{m}^{-1}(m) = 1 - \sqrt{m \frac{2}{h\delta} + (1 - G_1)^2}$ . Thus,

$$G_p(G_m) = 1 - \sqrt{(1 - G_m) \frac{2}{h\delta} + (1 - G_1)^2}$$

Hence

$$r(G_m) = 1 + \delta G_2 - \delta \sqrt{(1 - G_m) \frac{2}{h\delta} + (1 - G_1)^2}$$