Taxing the Job Creators: Efficient Progressive Taxation with Wage Bargaining

Nicholas Lawson*
Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS
nicholas.lawson@univ-amu.fr
May 6, 2014

Abstract

The standard economic view of the personal income tax is that it is a distortionary fiscal instrument that exists because it represents an equity-enhancing way of raising revenue to pay for public goods. In this paper, I analyze the efficiency role of taxation in a general equilibrium model of team production with wage bargaining. In this setting, the laissez-faire labour market equilibrium is inefficient, and income taxes that are progressive over most of the income distribution improve efficiency by offsetting the bargaining power held by managers at the top of the distribution; the efficient top tax rate is between 50% and 60%. This result is due to a “job-creation” effect: a positive marginal tax on high-income individuals offsets their incentive to work too hard in order to accumulate workers beneath them that they can exploit for rents; it is because of their job-creation activity that the “job creators” should be heavily taxed.

Keywords: optimal income taxation, progressive taxation, wage bargaining, team production

*I wish to thank participants of seminars at Princeton University, the Aix-Marseille School of Economics, and the Delhi School of Economics, for their input. I also thank Robin Boadway, Cecilia García-Peñalosa, Jacob Goldin, Marc Sangnier and Dean Spears for their helpful comments and suggestions. Any errors or omissions are the responsibility of the author.
1 Introduction

The standard economic view of the personal income tax is that it is a distortionary fiscal instrument that exists because it represents an equity-enhancing way of raising revenue to pay for public goods.\(^1\) This view has shaped the optimal income taxation literature that started with Mirrlees (1971) and is surveyed by both Mankiw, Weinzierl, and Yagan (2009) and Diamond and Saez (2011), and has persisted despite growing evidence that wages are not generally equal to marginal product (see, for example, Manning (2003) and Manning (2011)). In this paper, I present a model of production in hierarchical teams, in which lower-skill workers match with higher-skill managers. I show that, in this setting, wage bargaining introduces an efficiency role for taxation: income taxes that are progressive over most of the income distribution improve efficiency by offsetting the bargaining power held by managers at the top of the distribution.

Since Mirrlees (1971), the majority of the optimal income taxation literature has focussed on a competitive wage-setting environment; Piketty, Saez, and Stantcheva (2014) note that “There is relatively little work in optimal taxation that uses models where pay differs from marginal product.”\(^2\) Varian (1980) is one of the very few early examples that deviates from this setting, considering a case in which variation in income is generated by random luck rather than effort.

A literature looking at taxation in the context of search and matching models began to develop a few decades later, starting with several papers which focus on ex-ante identical populations: Boone and Bovenberg (2002) show how a linear wage tax can restore efficiency in a search and matching model, while Robin and Roux (2002) find that progressive taxation of workers can improve welfare by reducing the monopsony power of large firms.

An important contribution is made by Hungerbühler, Lehmann, Parmentier, and van der

---

\(^1\)For example, Blomquist, Christiansen, and Micheletto (2010) note that “The common view seems to be that marginal income taxes are purely distortive,” and Sandmo (1998) argues that “distortionary effects of taxation...can only be justified from a welfare economics point of view by their positive effects on the distribution of income.”

\(^2\)Outside of the economics profession, however, support for progressive taxes based on a belief that wages do not equal marginal product is more common. For example, Piketty, Saez, and Stantcheva (2014) argue that the 80-90\% marginal tax rates found in the U.S. and U.K. between the 1940s and 1970s were largely due to the fact that “policy makers and public opinions at that time probably considered...that at the very top of the income ladder, pay increases reflect mostly greed and socially wasteful activities rather than productive work effort,” and they note a revival of these ideas in the Occupy Wall Street movement which argues that “pay going to the top 1\% is unfair because it has come at the expense of the remaining 99\%.”
Linden (2006), who consider a distribution of individuals across skill levels, and examine the effect of taxes on vacancy creation with wage bargaining. They show that progressive taxes can reduce unemployment, with beneficial redistributional consequences that lead to a positive optimal tax rate at the top of the distribution of workers. Lehmann, Parmentier, and van der Linden (2011), meanwhile, extend the model to consider endogenous participation. However, both of these papers assume away inefficiency in the laissez-faire equilibrium. Furthermore, by focussing on a setting of directed segmented search, in which workers match with non-anthropomorphic firms in a continuum of separate labour markets, these papers ignore managers and executives, and therefore cannot say anything about income taxation at the upper end of the income distribution.

Finally, Piketty, Saez, and Stantcheva (2014) argue that most of the elasticity of taxable income that has been observed at high incomes comes from changes in bargaining over compensation rather than labour supply responses, and find a socially optimal top tax rate of 83%. They use a simple model of bargained wages to solve for an equation for the top tax rate based on elasticities of labour effort, tax avoidance, and bargaining with respect to taxes; however, there is limited evidence on the values of these elasticities, and Piketty, Saez, and Stantcheva (2014) do not present results relating to the rest of the income distribution.

No study that I am aware of considers an efficiency role for taxation in a model of the entire skill or income distribution, as in the standard Mirrleesian analysis. In this paper, I present the first such analysis using a model of team production in general equilibrium. Specifically, I use a model adapted from Antrás, Garicano, and Rossi-Hansberg (2006) which features endogenous hierarchical one-to-many matching, in which lower-skill workers match with higher-skill managers to produce according to the team’s ability to overcome problems encountered in production. This model represents two essential features of real-world labour markets: most individuals are employed in firms with two or more levels, so that workers at the bottom of the hierarchy answer to people higher up, and wages for the lower level are set by people at the top of the firm. I examine equilibrium outcomes under competitive taxation.

3 Related papers also include Jacquet, Lehmann, and van der Linden (2013), who consider both extensive and intensive labour supply responses, and Jacquet, Lehmann, and van der Linden (2014), who consider endogenous participation with Kalai bargaining. Another study, Hungerbühler and Lehmann (2009), focusses on the role for a minimum wage in a search and matching framework.

4 Boadway and Sato (2014) also studies taxation with different types of jobs, but in a different setting in which jobs are not hierarchical, but rather correspond to an extensive-margin decision between different occupations.
wage-setting and a simple form of wage bargaining, and show that given an underlying skill
distribution, wage bargaining generates a far more right-skewed income distribution, in which
rents extracted from middle-income workers are captured by the highest-skill managers.

I then consider the efficiency case for taxation. Since wage bargaining causes wages
to deviate from the competitive laissez-faire values, labour supplies are also distorted and
the equilibrium allocation is inefficient. However, a tax schedule that sets each individual’s
after-tax wage to its competitive laissez-faire value would restore efficiency to the labour
market, and I demonstrate that the efficient tax system is progressive over most of the
income distribution with a top marginal rate of 50-60%. Significant positive marginal taxes
at the top of the income distribution can serve an important efficiency role in offsetting
the bargaining power of the highest-skill managers. I also show that this result does not
follow simply from wage bargaining or complementarity of worker and manager labour effort,
but also requires a “job-creation effect”: positive taxes on high-skill managers are efficient
because increased manager effort allows them to accumulate more workers beneath them
that they can exploit for rents, making the manager’s “wage” per unit of labour supply
too high regardless of the level of worker effort. In other words, contrary to the common
argument that taxes at high incomes should be lowered to encourage job creation,\footnote{See
\cite{Krugman_2011} for a discussion of this point; Krugman points out that this argument
is dependent on high-income individuals not fully capturing the benefits that they produce for society. Work
by Alan Manning, as well as the equilibrium of the model in this paper, suggests the opposite conclusion.
\footnote{I do not model unemployment, as this would make the model excessively complicated. However, Boone
and Bovenberg (2002) demonstrate that, with low worker bargaining power, subsidies to workers and taxes
on firms are still optimal with unemployment due to search frictions.}
\footnote{The result of a declining optimal rate near the top of the income distribution is due to the assumption
of a finite top to the distribution; in the usual Mirrleesian analysis with competitive labour markets, the
optimal top tax rate is zero.}

we want
to tax the “job creators” \emph{because} they want to “create” too many jobs at their firm.\footnote{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetext{\protect\footnotetex}
production and describes the equilibrium under both competitive wage-setting and wage bargaining. Section 3 evaluates the efficiency case for taxation, presenting efficient taxes with wage bargaining and describing the theoretical properties underlying the results. Section 4 presents optimal taxes with diminishing marginal utility, and section 5 concludes the paper.

2 Model of Team Production

I consider a model of production in hierarchical teams adapted from Antràs, Garicano, and Rossi-Hansberg (2006), in which individuals from a skill distribution match in two-layer teams consisting of lower-skill workers and a higher-skill manager. This model is intended to represent two essential features of the real-world labour market: most individuals are employed in firms with two or more levels, so that workers at the bottom of the hierarchy answer to people higher up and, ultimately, to the executives at the top, and wages for the lower level are set, either through bargaining or subject to a competitive labour market, by people at the top of the firm.

I will begin by presenting and explaining the model, and then I will solve for the equilibrium under both competitive wage-setting and a form of wage bargaining.

2.1 Model Setup

The model features a continuum of agents with skill levels $z \in [0, 1]$, distributed according to a continuous function $F(z)$ and associated probability density function $f(z)$, who match in teams of a measure-one manager and measure $n$ of workers. The workers specialize in production, while the managers supervise the production process, and I will use subscripts $p$ and $m$ to denote quantities attached to workers ($p$ for production) and managers respectively. The matching process is endogenous, but the equilibrium must feature what I call perfect positive assortative matching: the matching process will take the form of a function, i.e. a one-to-one mapping between worker skill level $z_p$ and manager skill level $z_m$, and it will be hierarchical in the sense that less-skilled agents become workers and more-skilled agents become managers, with the lowest-skill worker matching with the lowest-skill manager and the highest-skill worker with the highest-skill manager. The proof of these results is discussed in appendix A.

8It is important to note that each agent represents an infinitesimally small space on the skill distribution, so the matching function will be continuous.
After agents form teams, wages are set, either competitively or as the result of a bargaining process, and workers and managers choose their labour supply, $L_p(z_p)$ and $L_m(z_m)$; from now on, I omit the $z$ arguments from labour supply to simplify the notation. Then, during production, each worker faces a problem of difficulty $d$ drawn from a uniform distribution over $[0, 1]$, and can solve any problem with difficulty less than or equal to their own skill level $z_p$. If the worker can’t solve the problem, they communicate it to the manager, subject to a communication cost $h \in (0, 1)$: the manager must spend $hL_p$ units of time on each problem that is forwarded to them, and can solve problems with $d \leq z_m$. If the problem is solved, $L_p$ units of output are produced by that worker, whereas workers with unsolved problems produce nothing. The manager therefore spends $hL_p(1 - z_p)$ units of time in expectation on each worker, and given that the manager faces a continuum of workers of measure $n$, the manager faces no uncertainty and has a managerial time constraint of $nhL_p(1 - z_p) = L_m$, and the team’s total output is $nL_pz_m$.

A utility function with no income effect is common in the optimal income tax literature; see, for example, Diamond (1998) and Persson and Sandmo (2005) and the description of such a specification as “standard” by Lehmann, Parmentier, and van der Linden (2011). To allow for diminishing marginal utility of income or a social taste for redistribution, I will therefore specify utility over consumption $C$ and labour supply as $U(C, L) = \frac{(C - \frac{1}{2}L)}{1 - \theta};$ this function exhibits zero income effect, and $\theta$ controls how fast marginal utility declines with income. Workers choose their labour supply $L_p$ to maximize utility, so a worker receiving a wage $w(z_p)$ will set $L_p = w(z_p)^{\frac{1}{1-\theta}}$. The manager chooses his labour supply $L_m$ and the skill level of worker $z_p$ he wishes to hire, which must be consistent with the equilibrium matching function; the manager receives total consumption of $C(L_m) = nL_p(z_m - w(z_p)) = L_m \frac{z_m - w(z_p)}{h(1-z_p)}$, and thus sets $L_m = \left( \frac{z_m - w(z_p)}{h(1-z_p)} \right)^{\frac{1}{1-\theta}}$, so that $r(z_m; z_p) \equiv C'(L_m) = \frac{z_m - w(z_p)}{h(1-z_p)}$ can be thought of as the manager’s “wage.”

I can then solve for the matching function given a particular wage function $w(z)$. If I denote $z^*$ for the cutoff skill level at which individuals are indifferent between being a worker or manager, and $m(z)$ as the skill level of the manager who supervises workers of skill $z$,
equilibrium in the labour market requires:

\[ \int_{0}^{z_p} f(z) dz = \int_{m(0)}^{m(z_p)} n(m^{-1}(z)) f(z) dz \quad \forall z_p \leq z^*. \]

Since \( n = \frac{L_m}{hL_p(1-z_p)} \), this can be rewritten as:

\[ \int_{0}^{z_p} f(z) dz = \int_{m(0)}^{m(z_p)} \left[ \frac{1}{h(1 - m^{-1}(z))} \right]^{\frac{1}{\gamma - 1}} \left[ \frac{z - w(m^{-1}(z))}{w(m^{-1}(z))} \right]^{\frac{1}{\gamma - 1}} f(z) dz \quad \forall z_p \leq z^* \]

and differentiating with respect to \( z_p \):

\[ f(z) = m'(z) \left[ \frac{1}{h(1 - z)} \right]^{\gamma - 1} \left[ \frac{m(z) - w(z)}{w(z)} \right]^{\frac{1}{\gamma - 1}} f(m(z)). \]

Therefore, the matching function is defined by:

\[ m'(z) = \left[ \frac{(h(1 - z))^{\gamma - 1}w(z)}{m(z) - w(z)} \right]^{\frac{1}{\gamma - 1}} f(m(z)). \]

(1)

### 2.2 Competitive Wage Setting

Next, the wage-setting mechanism must be described. The equilibrium will consist of two differential equations, one for the matching function and one for the wage function \( w(z) \).

As described above, in equilibrium, the manager’s choice of \( z_p \) must be consistent with the matching function, so in the competitive case I assume that the manager faces a wage function \( w(z) \) and must choose their preferred \( z_p \); thus, I differentiate the manager’s rents \( C \) with respect to \( z_p \) and set the derivative equal to zero,\(^{11}\) solving for:

\[ w'(z_p) = \frac{z_m - w(z_p)}{1 - z_p}. \]

Therefore, the equilibrium is defined by (1) and the equation describing the wage function:

\[ w'(z) = \frac{m(z) - w(z)}{1 - z} \]

along with the boundary conditions \( m(0) = z^* \), \( m(z^*) = 1 \), and \( C(L_m(z^*)) - \frac{1}{\gamma} L_m(z^*) = w(z^*)L_p(z^*) - \frac{1}{\gamma} L_p(z^*)^\gamma \), which ensures that individuals at \( z^* \) are indifferent between being a worker or a manager, and which simplifies to \( w(z^*) = \frac{z^* - w(0)}{h} \).

\(^{11}\)Essentially, the manager’s first-order condition tells us what the slope of the wage function must be for \( w(z) \) to be an equilibrium.
I solve these equations numerically, using a uniform skill distribution for now; here and for the remainder of the numerical analysis in the body of the paper, I assume a compensated elasticity of taxable income equal to 0.25,\(^\text{12}\) implying that $\gamma = 5$, and I use a population with 10001 mass points at $\{0, 0.0001, \ldots, 1\}$ as an approximation to a continuous distribution. For illustrative purposes, I assume a value of $h = 0.5$, in the middle of the permissible values;\(^\text{13}\) when I calibrate $h$ (along with the skill distribution) in later analysis, I end up with $h = 0.3850$ in the competitive case and $h = 0.9826$ in the bargaining case, suggesting that $h = 0.5$ is a reasonable compromise for this initial analysis.

The resulting wage and matching functions are displayed in Figure 1. $z^*$ takes a value of about 0.8, and to the right of $z^*$ the figure for $m(z)$ in panel (b) actually displays the inverse matching function $m^{-1}(z)$ while the wage function figure in (a) displays the manager’s “wage” $r(z_m)$. One important characteristic of the matching function is that it flattens out as $z$ approaches $z^*$, indicating that higher-skill managers are able to supervise more workers because both $L_m$ and $z_p$ are higher, the latter meaning that each worker can solve more problems and bothers the manager less frequently. Meanwhile, the wage function exhibits a kink at $z^*$: the wage rises more rapidly to the right of $z^*$. This confirms that it is an equilibrium for individuals below $z^*$ to become workers and those above $z^*$ to become managers: for a given skill level, the “wage” earned as a manager is higher than that earned as a worker for $z > z^*$ and vice-versa for $z < z^*$.

### 2.3 Wage Bargaining

Next, instead of perfectly competitive labour markets, I will consider a simple form of wage bargaining. Specifically, I assume fixed sharing of expected output, so $w = \beta z_m$, where $\beta$ is set in equilibrium to clear the labour market; this is also the outcome of a Nash bargain over the surplus, but the Nash bargaining solution becomes far more convoluted when non-zero taxes are introduced, so for simplicity I stick to the simple output-sharing specification.\(^\text{14}\)

---

\(^{12}\)An elasticity of 0.25 is suggested by Saez (2001), and Saez, Slemrod, and Giertz (2012) select it as the approximate midpoint of a range of plausible estimates from 0.12 to 0.4. I perform a sensitivity analysis in appendix E using a value of 1 for the elasticity of taxable income, and show that my results for optimal taxes with wage bargaining are only slightly lower, while optimal taxes with competitive wages are substantially lower.

\(^{13}\)The maximum feasible $h$ is about 0.916 with a uniform distribution and competitive wage-setting; above that values, communication is too costly to sustain a hierarchical equilibrium. With wage bargaining, the same constraint does not bind, but I impose a maximum of 1.

\(^{14}\)In essence, I abstract from any direct effects of taxes on wages; with Nash bargaining, wages shift with the marginal and average tax rates faced by the worker and the manager. As a result, efficient taxes with
With this output-sharing rule, the matching function simplifies to:

$$m'(z) = \left[ \frac{\beta(h(1-z))}{1 - \beta} \right]^{\frac{1}{1-\gamma}} \frac{f(z)}{f(m(z))}.$$  \hfill (2)

This equation defines the equilibrium, along with the wage equation $w(z) = \beta m(z)$, the boundary conditions $m(0) = z^*$ and $m(z^*) = 1$, and the condition of indifference at $z^*$, which simplifies to $\beta = \frac{z^*}{h+z^*}$.

Once again, I can solve this equation numerically given a flat skill distribution and $h = 0.5$, and the resulting wage and matching functions are displayed in Figure 2. The matching function looks similar to the competitive case, but the wage function exhibits a much more dramatic rightward skew; because managers are able to extract rents from their workers, and higher-skill managers have larger teams working below them, the highest-skill managers can receive very large returns. Thus, a model with wage bargaining can more easily explain a long right tail to the income distribution than can a model with competitive wage-setting.

3 Efficiency and Taxation

Because the laissez-faire equilibrium in the competitive case is efficient, and wages in the bargaining case deviate from the competitive values, it is clear that the allocation with wage bargaining will deviate less from zero, because taxes shift after-tax wages faster, but optimal taxes with diminishing marginal utility from income will tend to be more progressive, as progressive taxes induce indirect redistribution by increasing wages at lower incomes.
wage bargaining is not efficient; that is, it fails to achieve the first-best if worker utility is linear in consumption. This happens because, with wages distorted from their laissez-faire competitive values, labour supplies are also distorted from the efficient values; some workers (those with wages that are too low) supply too little effort, and some supply too much, working beyond the point when their contribution to society equals the marginal utility cost from effort.

To illustrate this, panel (a) of Figure 3 overlays the equilibrium wages from the competitive and bargaining cases, while panel (b) does the same with labour supplies. Both pictures tell a similar story: middle-skill workers are paid too little and thus work too little, while low-skill workers and especially the highest-skill managers receive excessive returns and work too hard. This pattern of inefficiency is not driven by the parameters chosen (recall that the output-sharing parameter $\beta$ is endogenously determined in equilibrium), but rather by the following logic: the highest-skill managers are able to supervise large teams and thus accumulate very large rents from the individuals at the top of the set of workers. Therefore, those top workers are underpaid, and because $\beta$ is set to make an individual at $z^*$ receive equal returns from being a worker or a manager, the lowest-skilled of the managers must also be underpaid, which implies that the lowest-skill workers are overpaid. Because of the convexity of the utility cost from labour supply, such a misallocation of labour can have significant efficiency consequences; in this case, if utility is linear in consumption (i.e. $\theta = 0$), the bargaining equilibrium features average utility that is 0.65% of mean consumption lower.
than the first-best.

Figure 3: Wage and Matching Functions with Uniform Distribution

(a) Equilibrium Wages

(b) Equilibrium Labour Supplies

This inefficiency implies that non-lump-sum taxes may not necessarily be distortionary, if the labour market equilibrium is already distorted. In fact, marginal taxes could serve an efficiency purpose: if a tax schedule can be chosen that sets each individual’s after-tax wage to the efficient value, then individuals will all choose the efficient labour supply and we will achieve the first-best. I consider this possibility in the following subsection.

3.1 Efficient Taxation with Wage Bargaining and Uniform Skill Distribution

Because the utility function exhibits zero income effects, each individual’s labour supply depends only on their after-tax wage: 
\[ L_p(z) = [(1 - t(y(z)))w(z)]^{\frac{1}{\gamma-1}} \] and 
\[ L_m(z) = [(1 - t(y(z)))r(z)]^{\frac{1}{\gamma-1}}, \]
where \( t(y(z)) \) is the tax rate assigned to an individual with labour market income \( y(z) \). As a result, the matching function with wage bargaining becomes:

\[
m'(z) = \left[ \beta(h(1-z))^{\gamma} \left( \frac{1 - t(y(z))}{1 - t(y(m(z)))} \right) \right]^{\frac{1}{\gamma-1}} \frac{f(z)}{f(m(z))}.
\] (3)

I assume, as is usual in the optimal taxation literature, that the government cannot observe skill levels of individuals, and thus can only tax them based on income. However, this is equivalent to assigning taxes to particular skill levels conditional on two constraints: an identification constraint requires that income increases with skill so that the government can identify skill levels from observations of income and impose the tax, and an incentive
compatibility constraint requires that when presented with the tax schedule as a function of income, each individual must prefer the income level that they would have chosen if faced only with the marginal tax rate assigned to them under the skill tax.\textsuperscript{15}

I therefore proceed to solve for the skill taxes that achieve the first-best, and then check to see if they satisfy the necessary constraints. I use a simple procedure in which I iterate between choosing the marginal taxes that match labour supply to the competitive value at each point along the skill distribution, and re-solving for equilibrium at the new taxes. I continue to use a flat skill distribution, and calculate the optimal marginal tax rates as displayed in Figure 4, which do satisfy both the identification and incentive compatibility constraints.

Figure 4: Efficient Taxes with Wage Bargaining and Uniform Distribution

(a) As Function of Skill

(b) As Function of Income

Panel (a) of Figure 4 presents optimal marginal taxes as a function of skill $z$, while panel (b) displays the optimal tax schedule as a function of income. The optimal taxes deviate from zero by a large amount, with small positive taxes at the bottom of the distribution, negative marginal rates in the middle, and rising tax rates at the top that reach as high as 60\%. Since positive marginal taxes are used to offset excessive bargaining power, taxes are especially high at the top end of the income distribution, where high-skill managers extract considerable rents from their moderate-skill workers.

\textsuperscript{15}The incentive-compatibility constraint is analogous to that in the standard Mirrleesian analysis, in which it must be the case that no individual wishes to “imitate” another worker and deviate from their prescribed income. In such a setting, income increasing with skill is a necessary condition for optimal taxation; see Mirrlees (1971).
This demonstrates that non-zero marginal income taxes, and in particular taxes that are progressive over much of the range of the income distribution, can actually increase efficiency when wages are not set equal to marginal product; highly progressive taxes can be justified without any motive for redistribution. However, the calculations were done using an arbitrary parametrization of the model; next, I will demonstrate that similar results are found when $h$ and the skill distribution are calibrated to the U.S. economy.

### 3.2 Efficient Taxes in Calibrated Model

In this subsection, I present efficient taxes with a non-uniform skill distribution; specifically, I calibrate the wage bargaining case of my model to match the income distribution measured by the 2012 March CPS, with both the CPS and model income distributions smoothed into a kernel density.\(^{16}\) I assume that the baseline tax system is represented by an approximation to the U.S. income tax: specifically, I use the assumption first made in Jacquet, Lehmann, and van der Linden (2013) of a linear tax at rate 27.9\% and a lump-sum transfer of $4024.90, which they argue is a good approximation to the real tax schedule of singles without dependent children according to the OECD tax database.

I start at $h = 0.5$ and a flat distribution, and then test small changes in $h$ and at 11 points along the skill distribution, $\{0, 0.1, ..., 1\}$, where the distribution is defined as a cubic spline across these points, and adjust $h$ and the spline nodes in the direction that reduces an objective function.\(^{17}\) In choosing the objective function, there is a conflict between matching the weight given to the center of the distribution and generating the long right tail to the income distribution that is observed empirically. The simple model presented above is unable to generate a right tail as long as that in the data, even in the case of the wage bargaining framework; to more accurately match the real-world income distribution, a model with more than two team layers would likely be required. However, as a compromise, I consider the differences between the resulting kernel densities multiplied by income squared; I choose $h$ and the values of the distribution nodes that minimize the sum of squares of this function. The resulting value of $h$ is 0.9826, and the skill distribution is displayed in Figure 5.

I then solve for efficient taxes using the methods described in the previous subsection: I first solve for the competitive equilibrium with this skill distribution, and then find the tax

---

\(^{16}\)I use a Gaussian kernel with a bandwidth of 0.3 times mean income.

\(^{17}\)To prevent very low densities, which could cause my matching algorithm to function poorly, I impose a minimum value of $f(z) = 0.2$ for the distribution.
rates that match the labour supply with wage bargaining to that obtained in the competitive laissez-faire equilibrium. However, the first-best is not incentive compatible in this setting, so the taxes are adjusted to ensure incentive compatibility according to a simple procedure described in appendix B. The result is displayed in Figure 6, as a function of real-world income in thousands of dollars. With the exception of a steeper drop in taxes at low incomes, from 0.6 to -0.6 rather than 0.2 to -0.4, the optimal tax schedule is quite similar to that presented in the flat-distribution case, and the welfare gain from moving from a 27.9% flat tax to the optimum is a substantial 2.59% of mean consumption. Therefore, these results confirm the robustness of the results in the flat-distribution case.

3.3 Theoretical Analysis of Efficient Progressive Taxation

The results above prove that there can be a strong efficiency case for marginal taxes that deviate considerably from zero, and even for taxes that are strongly progressive over most of the income distribution. However, with the interdependence of labour supply decisions through the matching function, the generality of this result is unclear: is it specific to the model I have presented? More generally, which features of my model lead to this conclusion?

In this subsection, I will focus on a simpler set of models, with the goal of explaining
the findings so far; in particular, I will explain what essential features of the model generate positive efficient marginal taxes at the top of the skill distribution. First of all, the idea that taxes can be used to offset pre-existing distortions in the labour market is simply an application of the Theory of the Second Best: introducing a new distortion may well improve welfare when the market is already distorted. To illustrate this, let me propose a very simple model of one representative team of one manager and one worker: assume that output is defined by \( Y = L_m L_p \), and that the manager pays the worker a wage \( w \) and retains the surplus, where utility is \( U(C, L) = C - \frac{1}{3} L^3 \). By assuming quasi-linear utility, I focus only on the efficiency role for taxation. The government may impose a marginal tax \( t_m \) on the manager and \( t_p \) on the worker. The worker’s labour supply is \( L_p = ((1 - t_p)w)^{\frac{1}{2}} \), and the manager’s problem is to maximize \( U = (1 - t_m)(L_m - w)L_p + b - \frac{1}{3} L^3_m \), where \( b \) is a lump-sum transfer or tax, which means \( L_m = ((1 - t_m)L_p)^{\frac{1}{2}} = (1 - t_m)^{\frac{1}{2}}((1 - t_p)w)^{\frac{1}{2}} \).

With competitive wages, \( w \) is set equal to marginal product, which is \( L_m \); therefore, \( w = (1 - t_m)^{\frac{1}{2}}(1 - t_p)^{\frac{1}{2}} \). In this case, solving for the tax rates that maximize social welfare \( W = Y - \frac{1}{3} L^3_p - \frac{1}{3} L^3_m \) is straightforward, and unsurprisingly I find that at the optimum \( t_p = t_m = 0 \).

If instead I assume a wage that is bargained to a value lower than marginal product,
there may be a role for taxation. To examine this possibility, let me assume that the wage is arbitrarily fixed at 0.8. Then social welfare can be written as follows:

\[ W = (0.8(1 - t_p))^{\frac{3}{4}}(1 - t_m)^{\frac{1}{2}} - \frac{1}{3}(0.8(1 - t_p))^{\frac{3}{4}} - \frac{1}{3}(1 - t_m)^{\frac{3}{2}}(0.8(1 - t_p))^{\frac{3}{4}}. \]

If I differentiate with respect to \( t_m \) and simplify, I find that:

\[ t_m^* = 0 \]

whereas solving for \( t_p \) gives me:

\[ t_p^* = -0.25. \]

Therefore, for efficiency purposes, the optimal corrective tax policy is a wage subsidy to the worker, while the marginal tax faced by the manager is zero.

In fact, this result is far more general than this specific example; to show this, let me consider a generalized version of the same model. Assume that there is some mass of agents from an unspecified skill distribution who match into teams, but importantly assume that these teams are of fixed size. The output of each team depends on the labour supplies and skill levels of the manager and workers: \( Y = Y(L_m, nL_p; z_m, z_p) \); the utility function is again assumed to be quasi-linear, with \( U = C - v(L) \). Focussing on one particular team, the social-welfare-maximizing levels of labour supply can be found by differentiating the social welfare function:

\[ W = Y(L_m, nL_p; z_m, z_p) - v(L_m) - n v(L_p) \]

which gives the following simple conditions:

\[ Y_m = v'(L_m) \]

\[ Y_p = v'(L_p) \]

where \( Y_m \) and \( Y_p \) denote the derivative of \( Y \) with respect to \( L_m \) and \( nL_p \) respectively. Meanwhile, given marginal taxes \( t_m \) and \( t_p \) and a wage \( w(t_m, t_p; z_m, z_p) \) that is bargained at the economy-wide level but may depend on taxes, individual choices satisfy the following conditions:

\[ (1 - t_m)Y_m(L_m, nL_p; z_m, z_p) = v'(L_m) \]

\[ (1 - t_p)w(t_m, t_p; z_m, z_p) = v'(L_p). \]
An inefficiency arises if \( w \) is not equal to \( Y_p \); however, if it is possible to set \((1-t_p)w = Y_p\) (that is, unless \( w \) decreases quickly as \( 1-t_p \) increases), then the worker’s labour supply will be restored to the efficient level. Then, note that \( w \) does not appear in \( Y_m \); the latter depends on the wage only through the worker’s labour supply \( L_p \), and therefore if \( t_p \) is set to restore \( L_p \) to the efficient level, \( L_m \) will also be efficient, and there is no need for a non-zero \( t_m \). Because the manager is assumed to be the residual claimant, when deciding on their labour supply, they take the worker’s decision as given and act as if they are maximizing total surplus. This result is summarized in the following proposition.

**Proposition 1.** With utility quasi-linear in consumption, and one-manager/n-worker teams of fixed size, where the manager is residual claimant:

(i) conditional on the worker’s effort choice, the manager’s effort choice is efficient;

(ii) if the wage-setting mechanism is such that a tax or subsidy on the worker can achieve the optimal worker labour supply, the efficient marginal tax faced by the manager is zero.

There is only one distortion, and so only one tax instrument is required to fix it. However, in the numerical illustrations above, I showed that the efficient tax schedule with wage bargaining includes not only wage subsidies for a significant fraction of the population of workers but also high tax rates for the highest-skill managers. What Proposition 1 tells us is that it is not wage bargaining or complementarity of labour inputs on their own that provides an efficiency argument for high tax rates at the top of the income distribution; something else must be responsible for this result. What is required is something that makes the manager’s decision inefficient even if worker labour supply is corrected, and in my model, the result of high taxes at the top of the income distribution is due to the fact that I do not hold team size fixed; team size depends not only on worker skill but also on the labour supplies chosen by the workers and the manager: \( n = \frac{L_m}{hL_p(1-z_p)} \). In the model above, if I allow for \( n'(L_m) > 0 \), then the manager’s labour supply satisfies:

\[
(1-t_m)(Y_m + (Y_p - w)L_p n'(L_m)) = v'(L_m)
\]

and if the rents \( Y_p - w \) obtained by the manager are positive, then \( t_m^* > 0 \).

This result can be explained as follows. If wages are set below marginal product, then a filled job is a scarce and valuable commodity for a manager or executive. If that manager can influence the number of workers under their control, i.e. if they can work harder and thereby
obtain more workers that can be exploited for rents, then the manager’s “wage” or return to each unit of effort is too high even if worker labour supply is efficient, and a positive marginal tax will reduce the manager’s labour supply towards the efficient level. Thus, in contrast to the usual public discussion about job creation, we want to tax high-income individuals or “job creators” not in spite of their job-creation, but because of their job-creation activities.

This result is related to the “Hosios condition” in the context of search and matching models: Hosios (1990) identified the condition on worker bargaining power that would need to be satisfied for the equilibrium to be efficient, and subsequent papers have examined optimal tax policy when the Hosios condition is not satisfied. A prominent example is Boone and Bovenberg (2002), who find that a tax on firms and a subsidy to workers is efficient if the workers’ bargaining power is too low, as in that case there would be both insufficient search by workers and excessive vacancy creation by firms. However, this idea has never been put in the current framework, with labour supply decisions and individual managers rather than non-anthropomorphic firms, and I am therefore able to provide new insight on the efficiency role of taxation across the wage distribution.

4 Optimal Taxation with Diminishing Marginal Utility

I now return to my hierarchical matching model and focus on optimal taxation in a setting with diminishing marginal utility of income; specifically, I assume \( \theta = 1 \), which implies \( U(C, L) = \ln \left( C - \frac{1}{\gamma} L^\gamma \right) \). I will present results using skill distributions which have been calibrated to the U.S. economy; results with a uniform skill distribution can be found in appendix D.

I allow the government to choose a non-linear continuously differentiable tax schedule \( T(y) \), for the purpose of financing lump-sum transfers and potentially a quantity of required government spending denoted by \( G \). Therefore, the government’s problem will be to choose the tax function to maximize average utility subject to their budget constraint:

\[
\max_{T(y)} W = \int_0^1 U(C(z), L(z)) f(z) dz \quad s.t. \int_0^1 T(y(z)) f(z) dz = G
\]

where I allow \( T(0) \) to represent any lump-sum transfer or tax.

---

\(^{18}\)A similar result is found by Cahuc and Laroque (2014), who consider a monopsonistic labour market with only an extensive margin.
As discussed before, labour supply depends only on the after-tax wage, and the matching function is altered to account for this, with (3) giving the matching function in the wage bargaining case, and the following equation for the competitive case:

\[ m'(z) = \left[ \frac{(h(1-z))^\gamma w(z)}{m(z) - w(z)} \left( \frac{1 - t(y(z))}{1 - t(y(m(z)))} \right) \right]^{\frac{1}{\gamma - 1}} \frac{f(z)}{f(m(z))}. \]

In the bargaining case, I use the same distribution and value of \( h \) as in section 3.2, and I use the same procedure for the competitive case to find \( h = 0.3850 \) and the skill distribution presented in Figure 7.

Figure 7: Calibrated Skill Density \( f(z) \) with Competitive Wages

To solve for optimal taxes in this setting, I need to evaluate the welfare impact of changing taxes at each point in the distribution. However, there is a complication: not only do marginal tax rates depend on income, but in the current matching environment the income of any individual depends on marginal tax rates across the entire income distribution, since they all impact the matching equilibrium. Rather than iterating between solutions for the income distribution and the marginal tax rates,\(^\text{19}\) I use the same basic method as in section 3 to solve for optimal taxes \( t(z) \) directly as a function of skill. A skill tax that satisfies the

\(^{19}\)Such an approach is computationally time-consuming, but more importantly, when I attempted this method, I encountered severe difficulties in ensuring convergence.
identification and incentive compatibility constraints can be implemented as an income tax, and since the set of income taxes is the implementable subset of skill taxes, the optimal implementable skill tax will also be the optimal income tax.

To apply the skill-tax method in this case is somewhat more complicated than in the efficient taxation analysis, however, as the laissez-faire competitive equilibrium does not provide an optimal baseline with log utility. I will therefore use a perturbation method, in which I consider a small alteration to the tax schedule at one point, which at the optimum should have zero impact on welfare. I consider a population of \( N \) individual mass points, denoted by \( n = \{1, \ldots, N\} \), with mass \( f(z_n) \) at skill levels \( z_n = \{z_1, z_2, \ldots, z_N\} \); in my case, \( N = 10001 \) with mass points at skill levels \( \{0, 0.0001, \ldots, 1\} \), but a perfectly continuous case is the limit as \( N \to \infty \). I assume that the government chooses \( N \) marginal tax rates, one for each individual, where the first applies to income up to and including the lowest-skill individual’s income \( y_1 \), and each subsequent tax rate \( t_n \) applies to the income in \( (y_n-1, y_n) \). \(^{20}\) To simplify notation, managers’ “wages” \( r(z_m) \) will be denoted \( w \) just like production workers. Then let me consider the effect on individuals across the distribution when the government changes one of the tax rates \( t_i \); I will separately consider the impact on individuals \( \{1, \ldots, i-1\} \) and on \( \{i, \ldots, N\} \).

**Impact of \( t_i \) on Individuals \( n = \{1, \ldots, i-1\} \):**

When the government raises the tax rate on individual \( i \), there are only two effects on individuals at lower skill (and income) levels: there will be a change in the lump-sum transfer \( T(0) \), and their wages may change as the matching function adjusts. Utility of individual \( n \) is:

\[
U_n = \log \left( w_n L_n - T(w_n L_n) - \frac{1}{\gamma} L_n^\gamma \right)
\]

so if I denote \( b = T(0) \), and write marginal utility \( \left[ w_n L_n - T(w_n L_n) - \frac{1}{\gamma} L_n^\gamma \right]^{-1} \) as \( MU_n \), the effect of a change in \( t_i \) is:

\[
\frac{dU_n}{dt_i} = MU_n \left[ \frac{db}{dt_i} + \left( L_n \frac{dw_n}{dt_i} + w_n \frac{dL_n}{dt_i} \right) (1 - t_n) - L_n^\gamma \right].
\]

Using the individual’s first order condition \( L_n^\gamma = w_n (1 - t_n) \), this simplifies to:

\[
\frac{dU_n}{dt_i} = MU_n \left[ \frac{db}{dt_i} + (1 - t_n) L_n \frac{dw_n}{dt_i} \right].
\]

\(^{20}\)In fact, I apply each tax rate \( t_n \) to \( (y_n-1 + \epsilon, y_n + \epsilon) \), where \( \epsilon \) is very small, so that I can evaluate the derivative \( \frac{dw_n}{dt_i} \) without having to be concerned about behavioural changes shifting an individual into a different tax bracket.
Impact of $t_i$ on Individuals $n = \{i, \ldots, N\}$:

Individuals at or above the skill level of individual $i$ also receive a change in the lump-sum transfer and face a change in wages, but they also pay higher taxes; since the tax $t_i$ applies to income from $(y_{i-1}, y_i]$, the change in the after-tax income of each $n$ is $y_i - y_{i-1}$, which I will denote as $\Delta y_i$. Therefore, the effect of a change in $t_i$ is:

$$\frac{dU_n}{dt_i} = MU_n \left[ \frac{db}{dt_i} - \Delta y_i + (1 - t_n)L_n \frac{dw_n}{dt_i} \right].$$

Total Effect of $t_i$ on Welfare:

If I denote welfare as $W = \sum_{n=1}^{N} f(z_n)U_n$, then the total impact of $t_i$ on welfare is:

$$\frac{dW}{dt_i} = \sum_{n=1}^{N} MU_n f(z_n) \left[ \frac{db}{dt_i} + (1 - t_n) L_n \frac{dw_n}{dt_i} \right] - \sum_{n=i}^{N} MU_n f(z_n) \Delta y_i. \tag{4}$$

Finally, I need to solve for $\frac{db}{dt_i}$. If $X$ is used to denote the total tax revenues collected by the government, $\frac{db}{dt_i} = \frac{dX}{dt_i}$. And $\frac{dX}{dt_i}$ can be written as:

$$\frac{dX}{dt_i} = \sum_{n=1}^{N} f(z_n) t_n \frac{dy_n}{dt_i} + \Delta y_i \sum_{n=i}^{N} f(z_n) \frac{dy_n}{dt_i} \tag{5}$$

where $\frac{dy_n}{dt_i}$ can be expressed as:

$$\frac{dy_n}{dt_i} = L_n \frac{dw_n}{dt_i} + w_n \frac{dL_n}{dt_i} = \frac{dw_n}{dt_i} + \frac{w_n}{\gamma - 1} \left( w_n^{\frac{\gamma - 1}{\gamma}} - w_n^{\frac{\gamma - 1}{\gamma}} \right) = \frac{\gamma}{\gamma - 1} L_n \frac{dw_n}{dt_i} + \frac{1}{\gamma - 1} \frac{y_n}{t_n} \frac{dt_n}{dt_i}$$

where $\frac{dt_n}{dt_i} = 1$ for $n = i$ and is zero otherwise. Therefore, $\frac{dX}{dt_i}$ is:

$$\frac{dX}{dt_i} = \sum_{n=1}^{N} f(z_n) \frac{\gamma}{\gamma - 1} t_n L_n \frac{dw_n}{dt_i} + \Delta y_i \sum_{n=i}^{N} f(z_n) - f(z_i) \frac{y_i}{\gamma - 1} \frac{t_i}{t_n} E(MU) \tag{5}$$

If I define $\sum_{n=1}^{N} f(z_n) = N$ and $\sum_{n=i}^{N} f(z_n) = N_i$, (4) and (5) can be combined to give:

$$\frac{dW}{dt_i} = \Delta y_i N_i \left[ E(MU) - E(MU|z_n \geq z_i) \right] - f(z_i) \frac{y_i}{\gamma - 1} \frac{t_i}{t_n} E(MU)$$

$$+ \sum_{n=1}^{N} f(z_n) L_n \frac{dw_n}{dt_i} \left[ \frac{\gamma}{\gamma - 1} t_n E(MU) + (1 - t_n) MU_n \right]. \tag{6}$$
This equation can easily be understood as the sum of three effects. The first term in (6) is the redistribution effect; the total tax revenues collected are multiplied by the marginal welfare gain from taxing high incomes and redistributing to everyone through a lump-sum transfer. The second term is the distortionary effect of the tax, the lost tax revenues from the reduced labour supply of individual $i$, and these first two terms represent the standard tradeoff in optimal taxation between redistribution and distortionary effects. However, the final term is a new component, a wage-shifting effect: the effect of the tax $t_i$ on wages is valued both for its redistribution effect, where it is weighted by each $MU_n$, and for its efficiency effect, where multiplied by $E(MU)$. This term provides an alternative way of expressing the distortion-offsetting effects of taxation: if a particular individual’s wage is too high, then taxing them will tend to increase average wages by shifting the matching function in an efficiency-enhancing direction.\footnote{A simple thought experiment shows why taxes must shift wages if they are not equal to marginal product, even in the fixed-team-size models of section 3.3: a manager’s “wage” is the sum of their actual contribution to society plus the rents they collect from workers divided by their labour supply. If a tax is imposed on the workers, they will work less and thus provide fewer rents to the manager, changing the hourly return the latter receives.}

If I write $R = \Delta y_i N_i [E(MU) - E(MU|z_n \geq z_i)]$ for the redistribution term and $S = \sum_{n=1}^{N} f(z_n) L_n \frac{d\ln w_n}{dt_i} \left[ \frac{\gamma - 1}{\gamma - 1} t_n E(MU) + (1 - t_n) MU_n \right]$ for the wage-shifting effect, I find that at the optimal tax rate $t_i$ it must be true that:

$$R + S = f(z_i) \frac{y_i}{\gamma - 1} \frac{t_i}{1 - t_i} E(MU)$$

and rearranging, this gives:

$$t_i = \frac{(\gamma - 1)(R + S)}{f(z_i)y_i E(MU) + (\gamma - 1)(R + S)}.$$  \hspace{1cm} (7)

Equations (6) and (7) look like a new set of “sufficient statistics” conditions for welfare analysis of taxation, but their practical applicability is limited by the fact that they require us to be able to measure marginal utilities, as well as individual wages $w_n$ and changes in those wages with taxation. Observation of wages is generally ruled out in analyses of optimal taxation; in the usual competitive labour market setting, wages are equivalent to skill levels, and thus observation of wages would make redistributional lump-sum taxes feasible. However, (6) and (7) can be used with any specific model, regardless of the wage-setting mechanism; by simulating the model, we can calculate the sufficient statistics and plug them into (6) to obtain the effect of changing $t_i$ on social welfare.
I use these equations to calculate the optimal tax schedule with my calibrated skill distributions: I find the optimal tax rate \( t_i \) for each individual subject to the identification and incentive-compatibility constraints. In practice, I use an iterative procedure and polynomial smoothing of the tax schedule, as described in further detail in appendix C.

The optimal taxes are presented for \( G = 0 \) and a positive \( G \), in this case the amount that balances the government budget given a baseline marginal tax of 27.9% and a minimum income of $4024.90. The competitive results are presented in Figure 8; the optimal marginal tax rates roughly follow an inverted-U shape. Moving from the baseline flat tax to the optimum produces welfare gains equivalent to 1.34% and 2.72% of mean consumption, due to gains from redistribution.

Figure 8: Optimal Tax Schedule with Log Utility and Competitive Wages

The reason for the inverted-U shape of the optimal tax schedule is simple: it is primarily driven by a redistribution effect that is generally in the shape of an inverted-U itself, as can be seen in Figure 9, which displays the values of \( R \) and \( S \) at baseline taxes in the \( G = 12.21 \) case (results are similar when \( G = 0 \)). The gains from redistribution are zero at the top and bottom of the income distribution, because a marginal tax at the top raises no revenue and a tax at the bottom cannot perform any redistribution; however, gains from redistribution are positive in between. There is a sharp spike upwards in the gains from redistribution at
the cutoff skill level $z^*$, because the income distribution becomes thinner at that point, but this is largely offset by the wage-shifting effect, which is positive at low incomes but drops abruptly to a large negative value above $z^*$. The latter occurs because a positive tax at any point in the distribution reduces labour supply at that point, shifting the matching function accordingly; thus, a tax on workers below $z^*$ has beneficial effects on welfare because it shifts workers to higher-skill managers and increases their wages, which a tax on managers has the opposite effect.

Figure 9: Values of $R$ and $S$ for Competitive Wages, Baseline Taxes and $G = 12.21$

(a) Value of $R$  
(b) Value of $S$

The results with wage bargaining can be found in Figure 10. The results are now a cross between the efficient tax with wage bargaining found in Figure 6 and the competitive results with log utility above: for the thick part of the income distribution, at low-to-moderate incomes, optimal taxes are still V-shaped, but after rising to about 60% at a fairly high income, the optimal marginal rate declines to below 40% at the top. This decline at the top, however, may not be a robust finding, as it is dependent on the assumption of a finite top to the income distribution, which is a necessary component of my model; in typical competitive models, a zero top tax rate will be optimal even with diminishing marginal utility of income, as a positive marginal tax at the very top of the distribution raises no revenue.\footnote{Saez (2001) points out that a lognormal income distribution will also tend to lead to a zero optimal asymptotic top tax rate; a Pareto tail to the income distribution generates a positive optimal asymptotic top rate.} The resulting welfare gains from moving to the optimal tax are 0.93% and 1.74% of mean consumption.
Figure 10: Optimal Tax Schedule with Log Utility and Wage Bargaining

The reason for this roughly S-shaped result can be found in the forms of $R$ and $S$ displayed in Figure 11. The gains from direct redistribution are positive but small for workers below $z^*$ (who occupy a very small space on the income distribution), but large and (aside from a spike just above $z^*$) inverted-U-shaped for managers, justifying high taxes at relatively high incomes but declining rates at the very top. Meanwhile, the wage-shifting effect takes a U-shape over most of the distribution, and this explains why marginal taxes do not go to zero at the top: high taxes at the very top of the distribution, by offsetting the bargaining power held by those highest-skill managers, improve efficiency.

The results in this subsection demonstrate that alternative models of wage-setting, or put differently, alternative estimates of the redistribution and wage-shifting gains from taxation, can lead to dramatically different results for optimal taxation. A competitive wage-setting environment leads to optimal taxes that are of a roughly inverted-U shape, whereas wage bargaining implies that taxes should be rising over much of the distribution and positive at the top, even though I assume a finite income distribution. In appendix D, I present results with a uniform skill distribution, and find very similar results, with even higher tax rates at the top in the wage bargaining case. I also redo the analysis with a considerably higher elasticity of taxable income of 1 in appendix E, and once again I find that my conclusions are
largely unaltered; optimal taxes are lower, but more so in the case with competitive wages, where optimal taxes are declining over much of the distribution.

5 Conclusion

In this paper, I have presented a new model of wage bargaining within teams in general equilibrium; this model demonstrates that a highly right-skewed income distribution can be generated without a skewed skill distribution when rents from workers are captured by high-skill team managers. I demonstrate that in this setting, marginal taxes that deviate significantly from zero can play an important role in improving efficiency, and I show that the efficient taxes are highly progressive over most of the income distribution. I also demonstrate that this result hinges on a “job-creation” effect, in which high-skilled managers exert too much effort in trying to accumulate workers and the rents that come with them. Finally, I apply an optimal income tax analysis to the model, and show that wage bargaining dramatically alters the optimal tax schedule to feature high tax rates near the top of the distribution.

Along with the numerical results, I present a new perturbation method for calculating optimal taxes that can be used with any model of the labour market. In so doing, I show that the welfare impact of taxes depends on a direct redistribution effect, the distortionary effects of taxes on the income of marginal individuals, and a new wage-shifting term.

Given the small range of papers which attempt to address issues of the use of income
taxes to offset labour market distortions, I believe this subject holds the promise of numerous important future contributions to our understanding of the welfare consequences of income tax policy.

A Proof of Perfect Positive Assortative Matching

In this section, I will describe the proof that the equilibrium of my model features perfect positive assortative matching, which I define as a one-to-one mapping between worker and manager skill with a single cut-off skill level $z^*$ between lower-skill workers and higher-skill managers. I will start with competitive wage-setting, and then present the proof with wage bargaining.

In the competitive labour market, much of the proof is identical to that in Antràs, Garicano, and Rossi-Hansberg (2005), so I will only discuss my deviations from that proof, and readers are directed to that paper for the details. I first need to prove that the mapping between worker skill $z_p$ and manager skill $z_m$ will be one-to-one, so that I can use a matching function. There is only one small modification to the proof in appendix B of Antràs, Garicano, and Rossi-Hansberg (2005); for a manager who hires $I$ different types of workers, the maximization problem is described by:

$$U = \max_{n_i} \sum_{i=1}^{I} n_i L_i (z_m - w(z_i)) - \frac{1}{\gamma} L^\gamma + \lambda [L - h \sum_{i=1}^{I} n_i L_i (1 - z_i)]$$

and therefore the first-order conditions for each $i$ are:

$$L_i (z_m - w(z_i)) - \lambda h L_i (1 - z_i) = 0$$
$$-n_i L_i w'(z_i) + \lambda h n_i L_i = 0.$$

However, the $L_i$ values cancel out, and then the first-order conditions are exactly identical to those in Antràs, Garicano, and Rossi-Hansberg (2005), and the proof proceeds as in that paper. Therefore, I can use a matching function $z_m = m(z_p)$ to describe the matching equilibrium.

Next, I must prove that $m(z)$ exhibits perfect positive assortative matching; that is, that $m'(z) > 0$ and the equilibrium is hierarchical with a single cutoff $z^*$. Again, with one minor modification, the proof of Theorem 1 in Antràs, Garicano, and Rossi-Hansberg (2005) applies: first, substitute the manager’s “wage” $r(z_m; z_p)$ for $\Pi(z_m, z_p)$, and the proof of $m'(z) > 0$ is exactly the same. The rest of the proof of Theorem 1 in Antràs, Garicano, and Rossi-Hansberg (2005) also goes through unchanged (though the solution derived there for $m(z)$ is altered in my analysis). Therefore, I can conclude that the matching function features perfect positive assortative matching in competitive equilibrium.

The proof with wage bargaining is somewhat different. With a fixed output-sharing rule of $w = \beta z_m$, all workers want to work for the highest-skill manager; and the manager gets the same rent of $(1 - \beta) z_m$ per worker no matter what their skill level, but must spend more time $h(1 - z_p)$ on lower-skill workers, so they also strictly prefer the highest-skill worker possible. Therefore, the equilibrium must feature a one-to-one mapping between $z_p$ and $z_m$, with positive assortative matching between individuals in the set of workers and those in the set of managers; in the interior of the set of workers, the matching function follows equation (2). Additionally, there must be workers at the bottom of the distribution (otherwise, some worker is working for a zero-skill manager and receiving zero income), and managers at the top (otherwise, the top individual could hire someone with $z = 1 - \epsilon$ and receive arbitrarily large rents as $\epsilon$ goes to zero), so the equilibrium must either
feature perfect positive assortative matching with a single cutoff $z^*$, or consecutive disjoint sets of workers and managers.

To prove that the latter is impossible, consider a situation in which individuals with skill up to $z_1$ are workers, those between $z_1$ and $z_2$ are managers, those between $z_2$ and $z_3$ are workers, and so on with any number of alternative blocks of workers and managers. Denote wages as $w_{01}(z)$ on the first segment and $w_{23}(z)$ on the third segment, with the manager’s wage denoted as $r_{12}(z)$ on the segment in between. If this is an equilibrium, it must be true that:

$$w_{01}(z_1) = r_{12}(z_1)$$
$$\lim_{z \uparrow z_1} w'_{01}(z) < \lim_{z \downarrow z_1} r'_{12}(z)$$
$$\lim_{z \downarrow z_2} r'_{12}(z) < \lim_{z \uparrow z_2} w'_{23}(z).$$

The first equation requires that individuals at $z_1$ are indifferent between being workers and managers, and can be written simply as $\beta z_2 = (1-\beta) z_1$. The two inequalities ensure that individuals marginally above and below the relevant cutoffs become workers or managers as required.

First, we can observe that the first inequality is always satisfied:

$$\lim_{z \uparrow z_1} w'_{01}(z) = \left[\frac{\beta h(1-z_1)\gamma}{1-\beta}\right]^{\frac{1}{\gamma-1}}$$
$$= (1-\beta) \left[\frac{z_1(1-z_1)}{z_2}\right]^{\frac{1}{\gamma-1}}$$
$$< (1-\beta) \frac{1-\beta}{h} = \lim_{z \downarrow z_1} r'_{12}(z).$$

Therefore, the second inequality needs to be disproved. I proceed as follows:

$$\lim_{z \downarrow z_2} w'_{23}(z) = \left[\frac{\beta h(1-z_2)\gamma}{1-\beta}\right]^{\frac{1}{\gamma-1}}$$
$$= (1-\beta) \left[\frac{z_1(1-z_2)}{z_2}\right]^{\frac{1}{\gamma-1}}$$
$$< (1-\beta) \frac{1-\beta}{h(1-z_1)} = \lim_{z \uparrow z_2} r'_{12}(z).$$

This disproves the second inequality, and proves that the only possible equilibrium features perfect positive assortative matching with a single cutoff $z^*$. And we know that such an equilibrium exists, because the proof of the first inequality applies to any $z_1$, including the $z^*$ in the perfect positive assortative matching equilibrium.

**B Solution Method for Efficient Taxes with Calibrated Distribution**

In the case with a calibrated skill distribution and wage bargaining in the efficient taxation analysis, the first-best taxes are not incentive compatible, as the marginal tax rates decline too quickly below

\footnote{A similar indifference condition must be satisfied at $z_2$, but this condition is not relevant to the current proof.}
Therefore, at each iteration, I apply an adjustment in which taxes are made to fit inside the boundaries imposed by the constraints, starting at $0.625z^*$ and moving both right and left from there; the choice of $0.625z^*$ makes the integral of the change in the tax schedule close to zero.

C Details of Solution Procedure for Optimal Taxes

To solve for the optimal tax schedule, I use equation (7) using an iterative procedure. For one round of iteration, I go through each individual one at a time and find the optimal tax rate within the bounds imposed by the identification and incentive-compatibility constraints; then I re-solve for the labour market equilibrium. In practice, the tax schedule adjusts very gradually towards the optimum, because the bounds imposed by the constraints are narrow but shift with the tax schedule.

I perform some number $q$ of rounds at one time, and only re-solve the $S$ term after all $q$ rounds in order to save time; I solve $S$ at 101 points, $z = \{0, 0.01, ..., 1\}$, and use 7th-order smoothed polynomials on each side of the threshold skill level $z^*$ to approximate the function (in the wage bargaining case, a cubic spline is used for the approximation above $z^*$ as a polynomial does not adequately capture the dramatic increase in $S$ at the top). At the end of each block of $q$ rounds, I also smooth the tax schedule using a polynomial best fit, choosing 13th-order polynomials on each side of $z^*$, as otherwise $S$ becomes unstable and poorly behaved; then I go through the tax rates and adjust them as needed to ensure that they fit inside the bounds imposed by the constraints. I allow $q$ to decline over time as the tax schedule converges, and the process concludes when $q$ reaches 6 and the squared sum of shifts in the tax schedule drops below 0.005 (which corresponds to a shift of about 0.0007 per individual).

D Optimal Taxes with Uniform Distribution

In this appendix, I present the optimal taxes with a uniform skill distribution and $h = 0.5$. I use the same procedure as in section 4 to solve for the optimal taxes in both the competitive and wage bargaining frameworks, for each of $G = 0$ and $G = 0.12$; the latter is a bit less than 20% of average income prior to taxes in both the competitive and bargaining frameworks, which is meant as a rough approximation to total tax revenues as a percent of GDP in the U.S. The results for the competitive case can be found in Figure 12, where we can observe that, as before, the optimal marginal tax rates roughly follow an inverted-U shape. As in the calibrated analysis, this is largely due to a redistribution effect in the shape of an inverted-U, as shown in Figure 13, which displays the values of $R$ and $S$ at zero taxes in the $G = 0$ case (results are very similar when $G = 0.12$). Moving from zero marginal taxes to the optimal taxes generates welfare gains that are equivalent to 4.06% and 8.63% of mean consumption respectively in the two cases, due to gains from redistribution.

The optimal marginal taxes for the wage bargaining case are presented in Figure 14. The results are generally similar to those in Figure 10, V-shaped at the bottom, then rising above 60% before declining to about 40% at the top. The values of $R$ and $S$ are displayed in Figure 15, and tell a broadly similar story to those in Figure 11 earlier. The welfare gains from shifting from zero marginal taxes to the optimum are 6.21% and 9.92% of mean consumption in the two cases.

The results in this appendix demonstrate that the findings in section 4 with the calibrated skill distribution are robust to the assumption of a uniform distribution.

---

24In the competitive analysis, I start the adjustment at $z^*$ and then move left and then right from there. In the main wage bargaining analysis, I start at $0.625z^*$, whereas I start at $0.55z^*$ with ETI = 1.
Figure 12: Optimal Tax Schedule with Log Utility and Competitive Wages

Figure 13: Values of $R$ and $S$ for Competitive Wages, Zero Taxes and $G = 0$

Value of $R$  
Value of $S$
Figure 14: Optimal Tax Schedule with Log Utility and Wage Bargaining

Figure 15: Values of $R$ and $S$ for Wage Bargaining, Zero Taxes and $G = 0$

Value of $R$  

Value of $S$
E  Sensitivity Analysis on Elasticity of Taxable Income

In this appendix I redo all of the numerical analysis of sections 3 and 4 using a higher value of the elasticity of taxable income (ETI). The correct value of this elasticity has been the subject of a significant controversy, with many of the earlier estimates being considerably larger than the 0.12-0.4 range stated as plausible by Saez, Slemrod, and Giertz (2012). A dramatic example is Feldstein (1995), who finds that the ETI is at least one; therefore, in this appendix, I assume an elasticity of one, or a value of $\gamma = 2$.

I begin with the flat-skill-distribution analysis. Figure 16 presents the efficient taxes with quasi-linear utility and wage bargaining, and a comparison of this figure with Figure 4 demonstrates that the conclusion about efficient taxes with wage bargaining is unaffected by a higher value of ETI; in fact, the top tax rate now increases to above 70%. The welfare gains in this case increase to 3.54% of mean consumption, as compared to 0.68% with ETI = 0.25.

Figure 16: Optimal Tax Schedule with ETI = 1, Quasi-Linear Utility and Wage Bargaining

Next, I calibrate $h$ and the skill distribution to the U.S. income distribution using the same procedure as before. With a lower value of $\gamma$, it is easier for both models to generate a long right tail to the income distribution. The efficient taxes with wage bargaining can be found in Figure 17. As before, the overall shape of the tax schedule is similar to the ETI = 0.25 case, and the optimal top tax again appears to be higher; the welfare gain from implementing the efficient tax schedule is a very large 7.93% of mean consumption.

Finally, the optimal taxes with log utility are in Figures 18 and 19. The level of the optimal tax with wage bargaining is somewhat lower than with ETI = 0.25, but the drop is much smaller than with competitive wages, and the optimal tax at the top remains above 30%, with a peak of over 40%. The reason for this is quite simply that the efficiency role of marginal taxes with wage bargaining does not depend directly on the responsiveness of individual behaviour to taxes; what matters is how far the wage deviates from the efficient level, and this deviation remains large with a higher ETI. The welfare gains are 2.59% and 3.32% with competitive wages, and 3.47% and 4.28% with wage bargaining, in each case significantly larger than when ETI = 0.25.

Furthermore, some research has indicated that the ETI may not be a constant, with Gruber and Saez (2002) finding larger values at higher income levels.

25
This appendix has demonstrated that the case for efficient progressive taxes at medium-to-high incomes is not sensitive to the elasticity of taxable income; the optimal tax schedule with wage bargaining is only moderately altered, even while optimal taxes with competitive wages drop significantly. Furthermore, the welfare gains from optimal taxes increase substantially with a higher value of ETI. The central reason for efficient progressive taxes is the distortion of wages from their efficient level, and this is not directly impacted by the responsiveness of individual behaviours to taxes.
Figure 19: Optimal Non-Uniform Taxes with ETI = 1, Log Utility and Wage Bargaining

References


