KU LEUVEN

CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS15.13

JULY 2015

Wage policies, employment, and redistributive efficiency

Erwin OOGHE

Public Economics

Faculty of Economics
and Business

Wage policies, employment, and redistributive efficiency

Erwin Ooghe*

July, 2015

Abstract

I analyze whether wage policies—like minimum wages and wage subsidies—can add value to an optimal non-linear earnings tax scheme in a perfectly competitive labour market. Jobs in the labour market differ along two margins: intensity (labour effort) and duration (labour hours). Three key results follow. First, even though minimum wages destroy low performance jobs, they increase employment if the minimum wage is binding, but not too high. Second, minimum wages—and wage and labour controls more generally—can enhance redistributive efficiency. The underlying mechanism is their potential to deter mimicking and thus to relax the self-selection constraints in the optimal income tax problem. Third, wage and labour controls become superfluous if a wage-contingent earnings tax scheme—a tax scheme that depends non-linearly on earnings and wages—can be optimally set. Instead, wage and labour subsidies can be optimal in a wage-contingent tax scheme.

JEL codes: H2, I3, J2

Keywords: minimum wages, labour subsidies, wage subsidies, optimal taxation.

*Erwin Ooghe, Department of Economics, KU Leuven, Naamsestraat, 69, 3000 Leuven; e-mail to erwin.ooghe@kuleuven.be. This paper has benefited from discussions with Rolf Aaberge, Spencer Bastani, Vidar Christiansen, André Decoster, Laurence Jacquet, Arnaud Lefranc, Holger Lüthen, François Maniquet, Ronald Peeters, Andreas Peichl, Paolo Piacquadio, Erik Schokkaert, Kevin Spiritus, Tom Truyts, Frank Vandenbroucke, and Javier Vázquez-Grenno. I also thank audiences in Oslo, Brussels, Paris, Leuven, Mannheim, Louvain-la-Neuve, and Barcelona for helpful comments.

1 Introduction

Minimum wages divide economists. There is no agreement about the impact of minimum wages on employment and about the normative desirability of minimum wages, especially compared to alternative policy instruments. For each disagreement I provide some evidence, discuss the relevant literature, and summarize the contribution of this paper to the debate.

What is the impact of minimum wages on employment? The percentage of economists (AEA members) that generally agrees with the statement "A minimum wage increases unemployment among young and unskilled workers" is declining over time, but remains substantial: 68% in 1979, 62% in 1990, 46% in 2000, and 39% in 2011 (Kearl et al., 1979 and Fuller and Geide-Stevenson, 2014).

One possible explanation for the sharp drop in the nineties could be the 'new economics of the minimum wage' initiated by Card and Krueger (1995). Since then, two decades of intensive empirical minimum wage research seem to confirm the findings of Card and Krueger that minimum wages have little or no negative impact on employment (Brown, 1999, Doucouliagos and Stanley, 2009, and Belman and Wolfson, 2014, but also Neumark and Wascher, 2008 for a dissenting view).

Yet, the percentage of economists that is in agreement with the statement remains substantial in the light of the empirical state-of-the-art. The standard neoclassical textbook view—minimum wages decrease employment—is highly influential and alternative neoclassical views remain scarce.[1]

I follow Saez (2000, 2002) and introduce jobs with two margins: an intensive labour effort margin and an extensive labour hours margin. Similar to labour hours, labour effort decreases utility directly, but it also increases utility indirectly via earnings, because jobs that require more effort are assumed to be compensated by higher wages.[2]

As a first contribution I show that a small binding minimum wage always increases employment, even if it destroys desirable low performance jobs. The relationship be-

[1] Positive employment effects of minimum wages are usually obtained by introducing imperfect competition (Bhaskar, Manning, and To, 2002 for an overview), imperfect information (Manning, 1995, and Rebitzer and Taylor, 1995), or search frictions (Burdett and Mortensen, 1998, Flinn, 2006, and Ahn, Arcidiacono, and Wessels, 2011).

[2] Wage rates are endogenous via the choice of effort. There is evidence indeed that higher income tax rates induce people to move to jobs with less amenities and lower wage rates (Gentry and Hubbard, 2004, Blomquist and Selin, 2010, Powell, 2010, Aaberge and Flood, 2012, and Powell and Shan, 2012).

tween minimum wages and employment is not monotone however. The higher the minimum wage, the more and more some individuals will prefer not to work, and employment is therefore likely to decrease from some minimum wage level onwards.

Although this result is striking in a neoclassical setting with perfectly competitive labour markets—and therefore worth stressing in my view—it has hardly been observed before. Deltas (2007) also introduces effort as an additional choice margin alongside labour hours and reaches the same conclusion.[3] Yet, in contrast to Deltas (2007), minimum wages destroy low performance jobs in my model and therefore utility always decreases if the minimum wage is binding. This brings us to the next question.

Are wage controls desirable? The percentage of economists (again AEA members) that would like to eliminate or decrease the federal minimum wage in the United States was equal to 48% in 2005, while the percentage that prefers to keep it at the current level or to increase it was equal to 52% (Whaples, 2006).[4] Among those who favour an increase in minimum wages, redistribution is most frequently cited (Klein and Dompe, 2007).[5]

That minimum wages improve the well-being of the working poor may sound obvious at first sight, but it turns out to depend crucially on whether we measure well-being via earnings or utility. As mentioned before, a small binding minimum wage will indeed increase both wages and labour hours—and thus also the gross earnings—of the working poor in the current model. Yet, because a binding minimum wage destroys low performance jobs, the utility of bounded workers decreases for sure. So, how then can minimum wages enhance redistribution from a utility-based welfare point of view?

In theory, one distortion can counteract another distortion. If redistribution is based on distortionary non-linear earnings taxation, then the use of other distortionary policy instruments can make mimicking by higher ability types less attractive, relax the incentive constraints, and improve redistributive efficiency. Commodity taxes, public goods, in-kind transfers, and workfare can do the trick (Boadway, 2012, chapter 4), but

[3] To the best of my knowledge only one other paper by Fields (1997) shows that minimum wages may increase employment in a neoclassical setting if there is a covered and an uncovered sector.

[4] The original question (response) was whether the federal minimum wage in the United States should be eliminated (46.8%), decreased (1.3%), kept at the current level (14.3%), increased by 50 cents per hour (5.2%), increased by $1 per hour (15.6%), and increased by more than $1 per hour (16.9%).

[5] Redistribution was cited by 75% of the 95 respondents, closely followed by "equalizing an imbalance in bargaining skills" (60%).

minimum wages not (Allen, 1987).[6] The reason is that introducing minimum wages at the bottom do not prevent high (wage) types from mimicking because they use their own (high) wage to mimick low wage individuals.

I follow Stiglitz (1982) and introduce two ability types. Higher ability leads to a higher wage rate for the same effort. Because higher ability individuals are always better off, society wants to redistribute from the high to the low types. Earnings, wages, and labour hours are observable to the planner, but ability type and labour effort not.[7]

As a second contribution I show that wage controls can enhance redistributive efficiency. Minimum wages, for example, can be used to deter mimicking if high ability individuals would prefer to mimick the gross earnings of the low ability individuals with a lower wage rate—and thus more labour hours—compared to the low ability individuals. In general, wage and labour controls—like minimal labour requirements and maximum wages—can be optimal as well, depending on the combination of wages and labour hours that is preferred by the mimickers.

Lee and Saez (2012) and Gerritsen and Jacobs (2014) also show that minimum wages can enhance non-linear earnings tax redistribution in perfectly competitive labour markets. Let me stress some crucial differences. Both papers assume a rationing mechanism—efficient rationing in Lee and Saez (2012) and general rationing in Gerritsen and Jacobs (2014)—whereas in this paper efficient rationing follows automatically from the choices that individuals make.[8] In addition, the underlying reason why minimum wages can be useful is different. In Lee and Saez (2012) minimum wages allow for more redistribution to subsidized low income workers by blocking other potential

[6] Allen (1987) shows that minimum wages add no value to (*i*) an optimal non-linear earnings tax scheme (*ii*) absent other policy instruments in (*iii*) a competitive labour market where (*iv*) individuals choose along the intensive (labour effort) margin. Conversely, minimum wages can be useful if *ceteris paribus* (*i*) optimal income taxes are linear (Guesnerie and Roberts, 1987, Allen, 1987), (*ii*) other instruments, like monitoring of job search and job offer acceptance, are available (Boadway and Cuff, 1999, 2001), (*iii*) labour markets are not competitive, e.g., in case of search frictions (Hungerbühler and Lehmann, 2009) and monopsonistic labour markets (Cahuc and Laroque, 2014), and (*iv*) individuals choose also along another margin, e.g., an extensive participation margin (Lee and Saez, 2012) or a skill formation margin (Gerritsen and Jacobs, 2014).

[7] Full observability of wage rates avoids the so-called mixed observability assumption that according to Guesnerie and Roberts (1987, p. 498) limits the early literature on minimum wages and redistribution. Although extreme, it is not too far-fetched, especially for low earners. I provide examples of existing tax and transfer programs later on that use information on wage rates and labour hours alongside earnings.

[8] As in the neoclassical textbook model—and contrary to the current paper—minimum wages decrease employment in both papers.

workers to enter the low income job. In Gerritsen and Jacobs (2014) minimum wages are useful if the gains from additional skill formation dominate the losses caused by unemployment. Finally, Lee and Saez (2012) show that tax credits and minimum wages are complementary policy instruments. This result breaks down in the current paper.

Do there exist superior policies? Over 600 economists, including several Nobel prize winners, support the recent proposal to increase the minimum wage in the United States (Economic Policy Institute, 2014). Still, some prominent economists have argued that earned income tax credits—an earnings subsidy for poor families—outperform minimum wages and should therefore be promoted instead (Glaeser, 2013, Neumark, 2013, and Romer, 2013). Other economists claim that a wage subsidy—a subsidy per hour of work—offers a superior alternative (Phelps, 2013, Harris, 2014).

To analyze such alternative policy instruments, one has to add information on wage rates or labour hours to earnings. Although the need for extra information is obvious for wage subsidies, one could argue that non-linear earnings tax schemes are sufficiently flexible to allow for tax credits. In reality however, tax credits are far more complex as they may indeed require information on wage rates (Belgium and France) and labour hours (Ireland, New Zealand, and the United Kingdom) alongside earnings (OECD, 2011). Contrary to classical tax credits, these schemes subsidize workers with low earnings and either low wage rates or high labour hours.

I follow Beaudry, Blackorby, and Szalay (2009) and introduce a wage-contingent earnings tax scheme, i.e., a tax scheme that depends non-linearly on wages and earnings. The idea to devise earnings tax schemes contingent on wage or labour hours information goes back at least to Mirrlees (1971), but remained almost unexplored since.[9]

As a third contribution, I show that wage controls are superfluous—even potentially harmful—if a wage-contingent earnings tax scheme can be optimally set. Similar to wage controls, wage contingency allows to deter mimicking. For example, if mimickers choose the same earnings, but higher wage rates, then they can be taxed more heavily.

[9] Mirrlees (1971, p. 208) ends his seminal work on optimal income taxation as follows: "I conclude, for the present, that:–

(1) An approximately linear income-tax schedule [...] is desirable [...]; and in particular (optimal!) negative income-tax proposals are strongly supported.

(2) The income-tax is a much less effective tool for reducing inequalities than has often been thought; and therefore

(3) It would be good to devise taxes complementary to the income-tax [...] this could be achieved by introducing a tax schedule that depends upon time worked as well as upon labour-income."

In addition, we provide conditions that justify wage subsidies (a negative marginal tax rate on earnings for a given amount of labour hours) and labour subsidies (a negative marginal tax rate on earnings for a given wage rate). Although ultimately an empirical question, the case for labour subsidies seems a priori more likely.

Beaudry, Blackorby, and Szalay (2009) is the only paper, to the best of my knowledge, that also analyzes the optimal design of non-linear wage-contingent tax schemes.[10] They find that labour subsidies below a cutoff wage are optimal in an economy with two sectors—a formal market sector and an informal non-market (or household) sector—where the planner does not observe the valuation of nonmarket time.

Overview The paper is organized as follows. Section two introduces the job choice model and describes the impact of minimum wages on job choice. Section three looks at the value-added of wage and labour controls for optimal non-linear earnings taxation. Section four discusses wage-contingent earnings tax schemes. A final section five concludes.

2 Job choice and minimum wages

I introduce a job choice model in this section and analyze the impact of minimum wages on job choice. Section 2.1 introduces the main building blocks, section 2.2 analyzes the resulting job choice, and section 2.3 studies the impact of a binding minimum wage.

2.1 A job choice model

A job is described by its intensity (labour effort) k and duration (labour hours) ℓ, also known as the intensive and extensive margin; see Saez (2002, p. 1042).[11] Gross earnings y do not only depend on effort k and hours ℓ, but also on individual ability θ. We define earnings as $y = w\ell$, with $w = \theta k$ the gross wage rate.[12] So, individuals with a higher

[10]There exists a small related literature. Kesselman (1976) simulates the redistributive power of a linear labour-contingent earnings tax and compares it to a classical linear earnings tax. Allingham (1975), Wagstaff (1975), Dasgupta and Hammond (1980), and Blomquist (1981, 1984) show that wage taxation comes close to first-best lump-sum taxation and is therefore superior to earnings taxation. Nishimura (2004) and Tillmann (2005) discuss the implementation of fair (envy-free) allocations via non-linear labour-contingent tax schemes.

[11]My model has continuous, rather than discrete margins and is therefore closer to Saez (2000).

[12]Choosing a more general specification $w = f(\theta, k)$ leads to the same qualitative results as long as f is strictly increasing.

ability must do less effort for the same wage rate. For later use, note that earnings y, the wage rate w, and the amount of labour hours ℓ will be observable to the planner; ability θ and labour effort k not.

Individuals choose a job that maximizes utility. Each job is in perfectly elastic demand. The utility function $u : (c, k, \ell) \mapsto u(c, k, \ell)$ is the same for all individuals, twice differentiable, strictly increasing in net earnings c, strictly decreasing in labour effort k and labour hours ℓ, and strictly quasi-concave.

Two interpretations of labour effort broaden the scope of the paper. First, effort could refer to any job disamenity—such as stress, the risk of injury, or location—that is compensated for by a higher wage. This interpretation is common in the empirical literature that estimates wage rate responses to tax rates (Albouy, 2009, Powell, 2010, Aaberge and Flood, 2012, and Powell and Shan, 2012). Second, in a long-run perspective, effort could also be interpreted as educational effort that increases future wages. This interpretation is common in the public finance literature that devises jointly optimal educational subsidies and earnings taxes (Blomquist, 1982, 1984, Tuomala, 1986, Brett and Weymark, 2003, Bovenberg and Jacobs, 2005).

2.2 The 'best' job

It is convenient to describe a job by its wage rate w and amount of labour hours ℓ from now on. I define therefore a utility function $v : (c, w, \ell, \theta) \mapsto v(c, w, \ell, \theta) \equiv u(c, \frac{w}{\theta}, \ell)$, with $\frac{w}{\theta}$ the labour effort required by an individual with ability θ to do a job with wage rate w. The function v inherits all properties of the function u. In particular, because higher wages require more effort, the utility function v strictly decreases with wages.

Without government intervention, the best job for an individual with ability type θ follows from

$$\max_{c,w,\ell \geq 0} v(c, w, \ell, \theta) \text{ subject to } c \leq y = w\ell,$$

or equivalently,

$$\max_{w,\ell \geq 0} v(w\ell, w, \ell, \theta).$$

This leads—by assumption—to a unique and interior best job for each ability type θ, denoted $(w(\theta), \ell(\theta))$. To obtain uniqueness, I impose (in this section only) that the function $\bar{v} : (w, \ell, \theta) \mapsto \bar{v}(w, \ell, \theta) \equiv v(w\ell, w, \ell, \theta)$ is strictly quasi-concave in (w, ℓ) for each ability type θ.

To visualize the best job, it is convenient to introduce two related conditional problems. The first problem is the utility maximizing choice of a wage rate w conditional on an amount of labour ℓ; the second problem is the utility maximizing choice of an amount of labour hours ℓ conditional on a wage rate w. The conditional problems are

$$\max_{w \geq 0} v(w\ell, w, \ell, \theta) \text{ and } \max_{\ell \geq 0} v(w\ell, w, \ell, \theta),$$

and lead—again by assumption—to unique and interior best wages and labour hours, denoted $W(\ell, \theta)$ and $L(w, \theta)$.[13]

Lemma 1 links the job choice to the conditional wage and labour choices (all proofs can be found in the appendix).

Lemma 1. *We must have*

$$w(\theta) = W(\ell(\theta), \theta) \text{ and } \ell(\theta) = L(w(\theta), \theta), \tag{1}$$

as well as

$$W'_\ell(\ell(\theta), \theta) > 0, \; L'_w(w(\theta), \theta) > 0, \text{ and } W'_\ell(\ell(\theta), \theta) \times L'_w(w(\theta), \theta) < 1, \tag{2}$$

for each ability type θ.

Equation (1) follows from the first-order conditions of the global and conditional problems. It tells us that the 'best' job must lie at an intersection of the conditional wage and the conditional labour curve. Equation (2) follows essentially from the second-order conditions requiring that, at the optimum, both conditional curves are upward sloping and the conditional labour curve intersects the conditional wage curve from below.

Figure 1 illustrates the conditional wage curve, the conditional labour curve, and the best job at the (interior) intersection of both curves.

2.3 Minimum wages

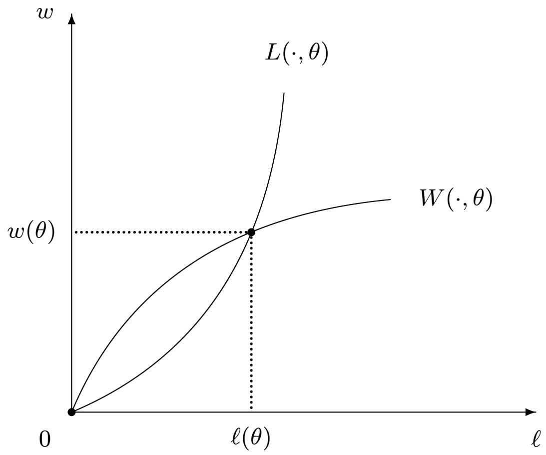
If we introduce a minimum wage in the job choice model, then the utility-maximizing job is given by

$$\max_{w \geq m, \ell \geq 0} v(w\ell, w, \ell, \theta), \tag{3}$$

with m the minimum wage level. Given a minimum wage, low performance jobs—jobs whose output per hour is lower than the minimum wage—are not offered anymore by profit-maximizing employers. This is in line with survey evidence among employers who

[13] Because $W(0, \theta) = L(0, \theta) = 0$, interiority of W and L holds only for strictly positive ℓ and w.

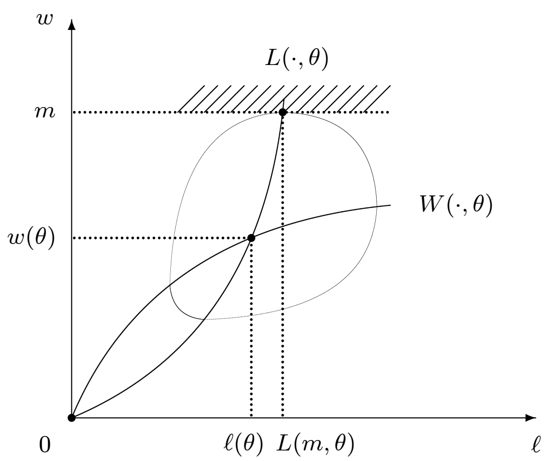
Figure 1: Job choice



state that performance standards and worker effort are among the main channels to cope with minimum wages (Hirsch, Kaufman and Zelenska, 2015).

Figure 2 repeats Figure 1, but adds a binding minimum wage $m > w(\theta)$ leading to the shaded set of feasible jobs. Figure 2 also shows the (strictly) convex better-than set that is just tangent to the shaded set.[14] The unique tangency between both sets must lie exactly where $L(\cdot, \theta)$ cuts the shaded set, because the conditional labour curve $L(\cdot, \theta)$ provides the best amount of labour conditional on a given wage rate. A binding minimum wage results therefore in a new best job $(m, L(m, \theta))$.

Figure 2: The impact of a binding minimum wage



Two remarks. First, binding minimum wages increase the amount of effort (from

[14] The better-than sets are defined as the sets $\{(w, \ell) | \bar{v}(w, \ell, \theta) = v(w\ell, w, \ell, \theta) \geq \alpha\}$ for different utility levels α. Because $\bar{v}$ is strictly quasi-concave (by assumption), the better-than sets are strictly convex.

$w(\theta)/\theta$ to m/θ in Figure 2).[15] In addition, $L'_w(w(\theta), \theta) > 0$ (Lemma 1) implies also that the amount of labour hours increases (from $\ell(\theta)$ to $L(m, \theta)$ in Figure 2). The positive impact of minimum wages on effort and labour hours could break down however for higher minimum wages because the corner solution (not working) may become optimal.[16] So, with a continuum of types, the relation between minimum wages on the one hand and total effort and total labour hours on the other hand is likely to be hump-shaped.[17] The higher the minimum wage, the more and more some individuals will prefer not to work such that total effort and total labour hours are likely to decrease from some minimum wage level onwards.

Second, irrespective of its impact on effort and labour hours, minimum wages always destroy low performance jobs and decrease therefore utility, because

$$\max_{w \geq m, \ell \geq 0} v(w\ell, w, \ell, \theta) \leq \max_{w \geq 0, \ell \geq 0} v(w\ell, w, \ell, \theta),$$

with a strict inequality if the minimum wage is binding. Proposition 1 summarizes.

Proposition 1. *A small binding minimum wage increases effort and labour hours. Yet, it also destroys desirable low performance jobs leading to a lower utility.*

3 Minimum wages and redistributive efficiency

Binding minimum wages reduce utility. The obvious next question is therefore: why should we introduce a minimum wage if it reduces utility? In this section I show that minimum wages—and wage and labour controls more generally—can improve redistributive efficiency by deterring mimicking. Section 3.1 provides an intuitive explanation. Section 3.2 looks again at the job choice problem of the individual, but now conditional on the instruments of the social planner. Section 3.3 formalizes the problem of the planner and presents the results.

3.1 An intuition

To understand the normative role of wage and labour controls in the simplest possible way, I follow Stiglitz (1982) and introduce two ability types, a low (L) and a high (H)

[15] Using the alternative interpretations of effort, minimum wages increase disamenities (or decrease amenities) and—in a long-run perspective—increase education. Admittedly, the empirical evidence for both effects is thin and inconclusive (Belman and Wolfson, 2014, chapter 6 for an overview).

[16] In addition—and for labour hours only—the conditional labour curve can be backward bending.

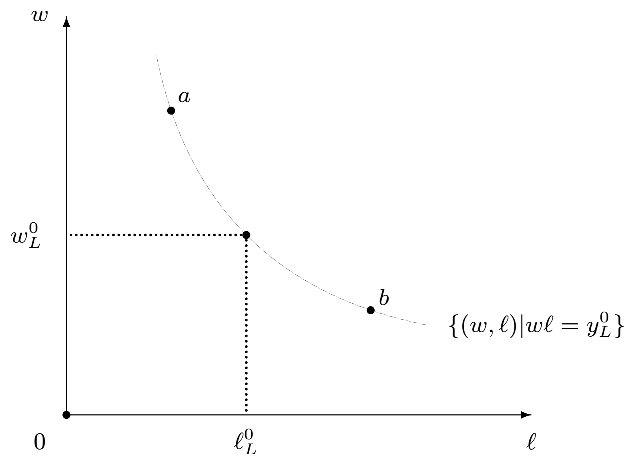
[17] Given a continuum of types, minimum wages will also lead to bunching, i.e., a mass of individuals hired at the minimum.

type. I add two standard assumptions. First, society wants to redistribute from high to low ability types because high ability types are always better off. Second, higher ability types always choose a job with higher gross earnings.[18]

Assume for the moment that only a non-linear earnings tax scheme is available for redistribution. With two types, the choice of an optimal earnings tax scheme is equivalent to the choice of optimal earnings bundles (c_L, y_L) and (c_H, y_H) subject to self-selection constraints—each type prefers the bundle that is intended for his type—and a budget constraint. Under the two standard assumptions described before, only the self-selection constraint of the high type will be binding at the second-best optimum (Stiglitz, 1982). In other words, the high type will be indifferent between both bundles.

Call (c_L^0, y_L^0) the optimal bundle for the low type. Because the low type can raise these gross earnings y_L^0 in different ways, let (w_L^0, ℓ_L^0) denote the preferred job of the low type to raise gross earnings y_L^0. Figure 3 plots the preferred job, together with the iso-earnings curve, i.e., all jobs that would lead to the same earnings as the low type.

Figure 3: Job choice of the mimicker and the low type



To mimick the low type, the high type must raise the same amount of earnings as the low type. Coincidentally, the high type may prefer to mimick the earnings of the low type on the basis of a higher wage rate and thus, lower labour hours (like job a in Figure 3) or vice-versa, using a lower wage rate and higher labour hours (like job b).

If case b applies, then a minimum wage set above the mimicker's preferred wage, will be binding for the mimicker. Because a binding minimum wage decreases utility (Proposition 1), it makes mimicking less attractive. The binding self-selection constraint

[18] To guarantee the last assumption, I will introduce a single crossing condition in the next section.

of the high type will be relaxed and redistribution can be improved. Similarly, if case a applies, a minimal labour requirement that is only binding for the mimicker will deter mimicking and enhance redistributive efficiency. In the next two sections I formalize the intuition with a focus on minimum wages.

3.2 Job choice conditional on earnings and minimum wages

The optimal tax-cum-minimum wage problem of the social planner is the following: what is the best earnings bundle for each type and the best minimum wage level subject again to self-selection constraints and a budget constraint. Before we can describe the planner's problem, it is crucial to understand the job that an individual will choose conditional on the policy instruments of the planner, being an earnings bundle and a minimum wage level.

For a given ability type θ, the best job (w, ℓ) conditional on net and gross earnings (c, y) and on a minimum wage level m is the solution to

$$\max_{w,\ell \geq 0} v(c, w, \ell, \theta) \text{ subject to } w\ell = y \text{ and } w \geq m. \tag{4}$$

Because the earnings constraint must be met with equality, one could equivalently define a utility function $\hat{v} : (c, w, y, \theta) \mapsto \hat{v}(c, w, y, \theta) \equiv v(c, w, \frac{y}{w}, \theta)$, with y/w the amount of labour hours necessary to raise gross earnings y for a given wage rate w. The solution for the best conditional wage, denoted $\hat{W}(c, y, m, \theta)$, then follows from

$$\max_{w \geq 0} \hat{v}(c, w, y, \theta) \text{ subject to } w \geq m, \tag{5}$$

and the solution for the best conditional labour amount, denoted $\hat{L}(c, y, m, \theta)$, is simply equal to $y/\hat{W}(c, y, m, \theta)$. The resulting indirect utility function is defined as $\hat{v}^* : (c, y, m, \theta) \mapsto \hat{v}^*(c, y, m, \theta) \equiv \hat{v}(c, \hat{W}(c, y, m, \theta), y, \theta)$.

The function $\hat{v}^*$ tells us how individuals with different ability types trade off net and gross earnings under the assumption that they optimally choose their best job. So, whether high type individuals will choose higher earnings depends on the marginal rate of substitution of consumption for earnings according to the indirect utility function $\hat{v}^*$, being

$$MRS\hat{Y}^*(c, y, m, \theta) \equiv -\frac{\hat{v}_y^{*\prime}(c, y, m, \theta)}{\hat{v}_c^{*\prime}(c, y, m, \theta)}.$$

If the higher ability type always requires less compensation in terms of net earnings for a small increase in gross earnings ($MRS\hat{Y}_\theta^{*\prime} < 0$), then the higher ability type will always choose a job with higher earnings.

I impose this single crossing condition for two reasons. It naturally links higher types with higher earnings; redistribution is then from the rich to the poor as usual. Second, this condition implies that it will never be optimal for the planner to 'pool', i.e., to provide the same bundle of earnings to the low and the high type. If pooling is not optimal and if second-best redistribution is assumed to be from the high to the low type, then only the self-selection constraint of the high type is relevant in the sequel.

3.3 The planner's problem

Pareto efficient redistribution from high to low ability types via non-linear earnings taxation and a minimum wage corresponds with the program

$$\max_{\{(c_i,y_i)\},m} \hat{v}^*(c_L, y_L, m, \theta_L) \;\; \text{subject to} \tag{6}$$

$$\hat{v}^*(c_H, y_H, m, \theta_H) \geq \underline{v}, \tag{7}$$

$$\hat{v}^*(c_H, y_H, m, \theta_H) \geq \hat{v}^*(c_L, y_L, m, \theta_H), \tag{8}$$

$$n_L(y_L - c_L) + n_H(y_H - c_H) \geq R_0, \tag{9}$$

with n_i the number of individuals with ability type $i = L, H$. Constraint (7) guarantees a minimal utility level $\underline{v}$ to the high type. Constraint (8) is the self-selection constraint of the high type. Constraint (9) is the budget constraint requiring that total tax revenues raise an exogenous amount of revenues R_0.

Let (c_L^0, y_L^0) and (c_H^0, y_H^0) denote the optimal redistribution scheme without minimum wages, i.e., the solution to the program in equations (6)-(9) with $m = 0$. Let w_L^0, w_M^0, and w_H^0 be the corresponding 'best' wage rates of the low type, the mimicker, and the high type.[19]

Proposition 2a. *Suppose $w_M^0 < \min\{w_L^0, w_H^0\}$ holds at the second-best optimum without minimum wages. Introducing a minimum wage relaxes the self-selection constraint and enhances redistributive efficiency.*

Proposition 2a tells us that it can be optimal to introduce a minimum wage, but it does not say anything about the resulting optimal minimum wage and the optimal earnings tax scheme. Let (c_L^*, y_L^*), (c_H^*, y_H^*), and m^* denote the optimal redistribution scheme and the optimal minimum wage, i.e., the solution to the program in equations (6)-(9). Let w_L^*, w_M^*, and w_H^* be the corresponding 'best' wage rates of the lows type,

[19] More precisely, $w_L^0 = \hat{W}(c_L^0, y_L^0, 0, \theta_L)$, $w_M^0 = \hat{W}(c_L^0, y_L^0, 0, \theta_H)$, and $w_H^0 = \hat{W}(c_H^0, y_H^0, 0, \theta_H)$.

the mimicker, and the high type.[20] Proposition 2b tells us that the optimal minimum wage will not only bind the mimicker, but also one of the other types. Proposition 2c shows that the marginal earnings tax rate remains to be zero for the rich and strictly positive for the poor, even in the presence of a minimum wage.

Proposition 2b. *Suppose $w_M^0 < \min\{w_L^0, w_H^0\}$ holds at the second-best optimum without minimum wages. At the second-best optimum with minimum wages the optimal minimum wage will be binding for the mimicker ($w_M^* = m^*$) and for the low type or the high type ($w_L^* = m^*$ or $w_H^* = m^*$).*

Proposition 2c. *Suppose $w_M^0 < \min\{w_L^0, w_H^0\}$ holds at the second-best optimum without minimum wages. At the second-best optimum with minimum wages the marginal tax rate for the high earner is zero and the marginal tax rate for the low earner is strictly positive.*

Also Lee and Saez (2012) and Gerritsen and Jacobs (2014) show that binding minimum wages can add value to a non-linear earnings tax in a perfectly competitive labour market. It is therefore worth stressing the main differences. First, I do not assume efficient rationing from the outset. Rather, efficient rationing follows automatically from the choices that individuals and firms make. Second, the reason why minimum wages can be useful is their potential to deter mimicking. Although a classical mechanism, it has not been described before.[21] Third, the marginal tax rate for low earners remains strictly positive and therefore subsidies to low earners (tax credits) and minimum wages are not complementary.

4 Wage-contingent tax schemes

We have seen that minimum wages can usefully supplement a non-linear earnings tax scheme. But if wages are observable, then tax schemes that depend non-linearly on wages and earnings can be used as well. Section 4.1 shows that wage-contingent taxation is superior to combining wage controls with non-linear earnings tax schemes. Section 4.2 introduces some new single crossing conditions. Section 4.3 characterizes the optimal wage-contingent tax scheme.

[20] So, here we have $w_L^* = \hat{W}(c_L^*, y_L^*, m^*, \theta_L)$, $w_M^* = \hat{W}(c_L^*, y_L^*, m^*, \theta_H)$, and $w_H^* = \hat{W}(c_H^*, y_H^*, m^*, \theta_H)$.

[21] In Lee and Saez (2012) minimum wages allow for more redistribution to subsidized low income workers by blocking other potential workers to enter the low income job. In Gerritsen and Jacobs (2014) minimum wages are useful if the gains from additional skill formation dominate the losses caused by unemployment.

4.1 Wage contingency is superior to wage controls

I follow Beaudry, Blackorby, and Szalay (2009) and introduce a wage-contingent earnings tax scheme T with $T(w, y)$ the tax (or subsidy, if negative) as a function of wages and earnings.

With a finite number of types, choosing an optimal wage-contingent tax scheme T is equivalent with directly choosing optimal extended bundles (c_L, w_L, y_L) and (c_H, w_H, y_H) for each type subject to self-selection constraints and a budget constraint. As usual, the self-selection constraints guarantee that the optimal extended bundles lie on the lower contour of the individual worse-than sets such that the optimal bundles can be decentralized via a wage-contingent earnings tax scheme T.

Pareto efficient redistribution from high to low ability types via a wage-contingent tax scheme corresponds therefore with the program

$$\max_{\{(c_i,w_i,y_i)\}} \hat{v}(c_L, w_L, y_L, \theta_L) \text{ subject to} \tag{10}$$

$$\hat{v}(c_H, w_H, y_H, \theta_H) \geq \underline{v}, \tag{11}$$

$$\hat{v}(c_H, w_H, y_H, \theta_H) \geq \hat{v}(c_L, w_L, y_L, \theta_H), \tag{12}$$

$$n_L(y_L - c_L) + n_H(y_H - c_H) \geq R_0. \tag{13}$$

The interpretation of the constraints is the same as before.[22]

Adding wage controls as constraints to the problem cannot improve redistribution, at the contrary. Because wages are now policy instruments, wage controls would restrict them and would therefore be harmless at best, but clearly harmful if binding. Proposition 3a summarizes.[23]

Proposition 3a. *If a wage-contingent earnings tax scheme can be optimally set, then binding wage controls can only harm redistribution and should not be introduced.*

4.2 Single crossing conditions

Before we characterize the optimal wage-contingent tax scheme, it is important to study single crossing conditions in more detail. Such conditions tell us how the different choices made by individuals with different ability types relate to each other. I look at two

[22] The single crossing conditions (discussed in the next section) guarantee that pooling—providing the same bundle (c, w, y) for both types—cannot be Pareto efficient. Again, only the self-selection constraint of the high type will be binding at the optimum.

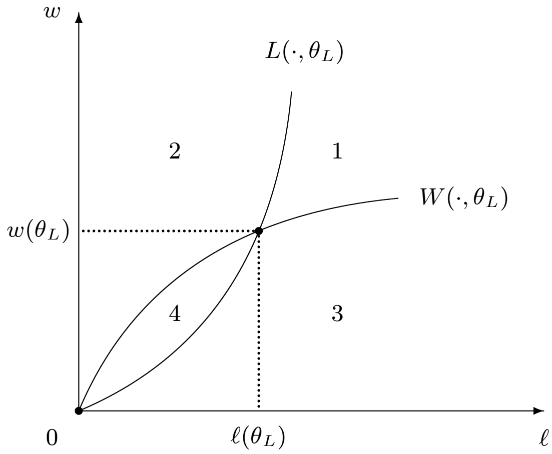
[23] Note that binding labour controls—binding restrictions on y/w—are harmful as well.

different ways to address this question and provide a summary at the end. The impatient reader can go directly to the summary in Section 4.2.3.

4.2.1 Conditioning on wages and labour hours

Figure 4 repeats Figure 1. The job choice at the intersection of the conditional curves is assumed to be the choice of the low ability type. The question is: where will the choice of the high ability type be?

Figure 4: Job choice conditional on wages and labour hours



Job choice will change with type depending on how the conditional wage and labour curves W and L change with type (so, depending on the sign of W'_θ and L'_θ).

Define the marginal rate of substitution of consumption for wage ($MRSW$) and the marginal rate of substitution of consumption for labour ($MRSL$) as

$$MRSW(c,w,\ell,\theta) = -\frac{v'_w(c,w,\ell,\theta)}{v'_c(c,w,\ell,\theta)} \quad \text{and} \quad MRSL(c,w,\ell,\theta) = -\frac{v'_\ell(c,w,\ell,\theta)}{v'_c(c,w,\ell,\theta)}. \tag{14}$$

Single crossing conditions restrict the sign of $MRSW'_\theta$ and $MRSL'_\theta$. For example, the condition $MRSW'_\theta < 0$ means that higher ability types always need to be compensated less in consumption for a small increase in wages, and thus effort. As a consequence, they will always choose higher (conditional) wages, ceteris paribus ($W'_\theta > 0$). Similarly, the condition $MRSL'_\theta > 0$ means that higher ability types always need to be compensated more in consumption for a small increase in labour hours and they will therefore always choose lower (conditional) labour hours, ceteris paribus ($L'_\theta < 0$).

We obtain four possible combinations of single crossing conditions. Each combination corresponds with one of the four cases indicated in Figure 4. Case 1 ($MRSW'_\theta < 0$ and

$MRSL'_\theta < 0$) is the classical case where higher ability types always end up with higher earnings as they choose to work more hours at higher wages. Case 4 ($MRSW'_\theta > 0$ and $MRSL'_\theta > 0$) is the opposite case where higher ability types have lower earnings through lower wages and lower labour hours. Case 2 ($MRSW'_\theta < 0$ and $MRSL'_\theta > 0$) and case 3 ($MRSW'_\theta > 0$ and $MRSL'_\theta < 0$) are ambiguous in terms of earnings.

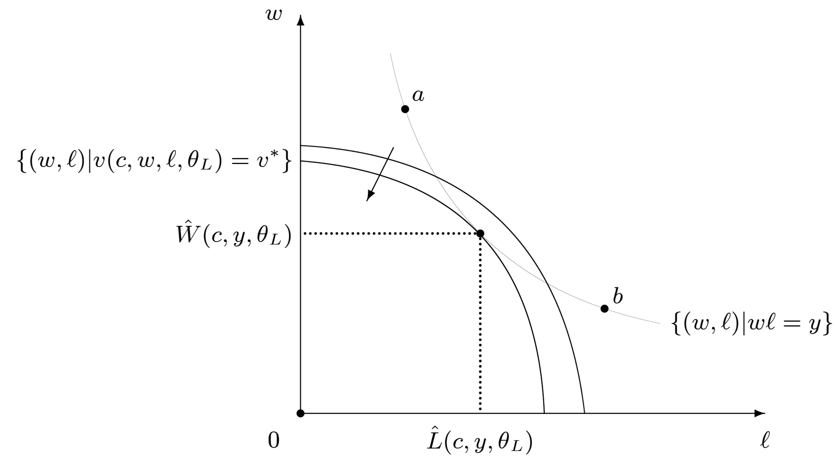
4.2.2 Conditioning on earnings

Proposition 3a tells us that wage controls are superfluous. I look here again at the problem described by equation (4) or (5), but now without the minimum wage constraint (i.e., $m = 0$). More precisely, given an amount of net and gross earnings c and y, what is the 'best' job (w, ℓ) for an individual with type θ to raise an amount of gross earnings equal to y? The solution, denoted $\hat{W}(c, y, \theta)$ and $\hat{L}(c, y, \theta)$ with slight abuse of notation,[24] follows from

$$\max_{w,\ell \geq 0} v(c, w, \ell, \theta) \text{ subject to } w\ell = y.$$

Figure 5 shows the iso-earnings curve, i.e., all possible combinations of wages and labour hours that raise the amount of gross earnings y. It also shows two iso-utility curves for the low type, being, all possible combinations of wages and labour hours that lead to the same utility level according to the low type (and given an amount of net earnings c). The best conditional job choice can be found where the iso-utility curve is just tangent to the iso-earnings curve.

Figure 5: Job choice conditional on earnings



[24] In the old notation, the solutions are $\hat{W}(c, y, 0, \theta)$ and $\hat{L}(c, y, 0, \theta)$; I simply drop the zero for the minimum wage in the sequel.

Using the definition of the function $\hat{v}$, the solution for the best conditional wage $\hat{W}(c, y, \theta)$ also follows from

$$\max_{w \geq 0} \hat{v}(c, w, y, \theta), \tag{15}$$

and $\hat{L}(c, y, \theta)$ is then simply equal to $y/\hat{W}(c, y, \theta)$.

The question is: where will the conditional choices of the high type locate compared to the low type? Because $\hat{W}'_\theta$ and $\hat{L}'_\theta$ have opposite signs by definition, I focus on $\hat{W}'_\theta$ only. The marginal rate of substitution of consumption for wages according to $\hat{v}$ ($MRS\hat{W}$) is defined as

$$MRS\hat{W}(c, w, y, \theta) = -\frac{\hat{v}'_w(c, w, y, \theta)}{\hat{v}'_c(c, w, y, \theta)} \tag{16}$$

Single crossing conditions restrict the sign of $MRS\hat{W}'_\theta$. We end up with two possible cases, denoted by a and b in Figure 5. In case a ($MRS\hat{W}'_\theta < 0$) individuals with a higher ability type always have a lower marginal rate of substitution of consumption for wages (according to $\hat{v}$), so they will always choose a higher wage rate conditional on earnings ($\hat{W}'_\theta > 0$) and thus a lower amount of labour hours ($\hat{L}'_\theta < 0$). The opposite ($\hat{W}'_\theta < 0$ and $\hat{L}'_\theta > 0$) holds in case b ($MRS\hat{W}'_\theta > 0$).

One can now address the final question: when will higher ability types choose higher earnings? Redefine the indirect utility function as $\hat{v}^* : (c, y, \theta) \mapsto \hat{v}^*(c, y, \theta) \equiv \hat{v}(c, \hat{W}(c, y, \theta), y, \theta)$ and the marginal rate of substitution of consumption for earnings according to $\hat{v}^*$ is

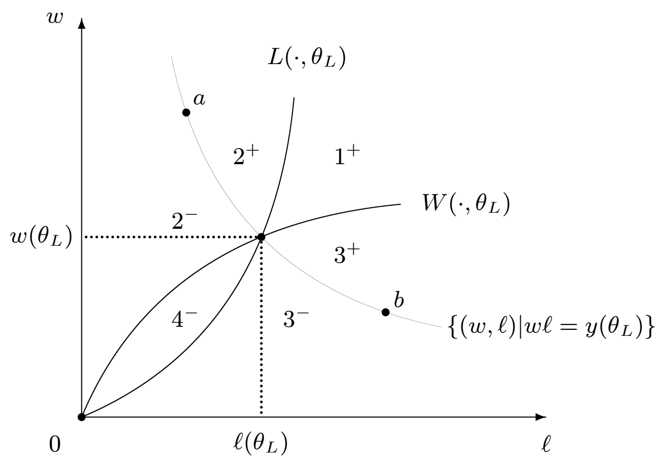
$$MRS\hat{Y}^*(c, y, \theta) = -\frac{\hat{v}^{*\prime}_y(c, y, \theta)}{\hat{v}^{*\prime}_c(c, y, \theta)},$$

We end up with two possible cases. If the higher ability type always requires less compensation in terms of net earnings for a small increase in gross earnings ($MRS\hat{Y}^{*\prime}_\theta < 0$), then the higher ability type will choose a job with higher earnings. If not ($MRS\hat{Y}^{*\prime}_\theta > 0$), then the higher ability type will choose lower earnings.

4.2.3 Summary

Figure 6 shows the job choice of the low ability type at the intersection of his conditional wage and labour curves. Figure 6 also draws the iso-earnings curve for the low type, i.e., all combinations of wages and labour that would lead to the same earnings as the low type (denoted by $y(\theta_L) = w(\theta_L)\ell(\theta_L)$). The iso-earnings curve divides the space in lower earnings (case –) and higher earnings (case +) compared to the low type.
Figure 6 has six zones (1^+, 2^-, 2^+, 3^-, 3^+ and 4^-) and two illustrative cases on the iso-earnings curve (a and b).

Figure 6: Job choice of the low and high ability type



As before, it is natural to assume that a high ability type always chooses a job with higher earnings (case + characterized by $MRS\hat{Y}^{*\prime}_\theta < 0$). This choice is not compatible with zone 4^- (characterized by $MRSW'_\theta > 0$ and $MRSL'_\theta > 0$) in Figure 6.

In addition, the marginal rates of substitution in equation (14) are linked to the marginal rate of substitution in equation (16). Using the definition of $\hat{v}$, it is easy to verify that

$$MRS\hat{W}(c, w, w\ell, \theta) = MRSW(c, w, \ell, \theta) - MRSL(c, w, \ell, \theta)\frac{\ell}{w}. \tag{17}$$

Differentiating both sides with respect to type, we obtain

$$MRS\hat{W}'_\theta(c, w, w\ell, \theta) = MRSW'_\theta(c, w, \ell, \theta) - MRSL'_\theta(c, w, \ell, \theta)\frac{\ell}{w}.$$

As a consequence, case 2 ($MRSW'_\theta < 0$ and $MRSL'_\theta > 0$) is not compatible with case b ($MRS\hat{W}'_\theta > 0$) and case 3 ($MRSW'_\theta > 0$ and $MRSL'_\theta < 0$) is not compatible with case a ($MRS\hat{W}'_\theta < 0$).

Of the sixteen potential combinations of single crossing conditions only four are left. Table 1 summarizes these combinations and indicates the corresponding cases.
In the next and final section, I will look at the properties of an optimal wage-contingent tax scheme in each of the four cases.

4.3 Results

I focus on a wage-contingent earnings tax scheme $T : (w, y) \mapsto T(w, y)$. For the interpretation however, it is convenient to also introduce an alternative, but informationally

Table 1: Possible combinations of single crossing conditions

$MRS\hat{Y}^{*\prime}_\theta < 0$	$MRSW'_\theta < 0$		$MRSW'_\theta > 0$	
case +	$MRSL'_\theta < 0$ case 1	$MRSL'_\theta > 0$ case 2	$MRSL'_\theta < 0$ case 3	$MRSL'_\theta > 0$ case 4
$MRS\hat{W}'_\theta < 0$ case a	$(1^+, a)$	$(2^+, a)$	(iii)	(i)
$MRS\hat{W}'_\theta > 0$ case b	$(1^+, b)$	(ii)	$(3^+, b)$	(i)

Combining cases (i) 4 with +, (ii) 2 with b, and (iii) 3 with a is not possible

equivalent labour-contingent earnings tax scheme $t : (y, \ell) \mapsto t(y, \ell)$, as was originally suggested by Mirrlees (1971). Let us discuss four policies.

To start with, $T'_w(w, y)$ and $t'_\ell(y, \ell)$ are the marginal tax rates for wages and labour hours conditional on earnings. For example, if $T'_w(w, y)$ is positive, then individuals with the same earnings, but slightly higher wages must pay more taxes. Because both tax rates are conditional on earnings, they will be especially useful to deter mimicking if mimickers prefer to raise the same earnings with a different wage-labour mix. For later use, note that the sign of $t'_\ell(y, \ell)$ is the opposite of $T'_w(w, y)$.[25] So, an earnings-conditional marginal wage tax ($T'_w(w, y) > 0$) corresponds with an earnings-conditional marginal labour subsidy ($t'_\ell(y, \ell) < 0$) at the optimum and vice-versa.

Two other policies of interest are a labour tax and a wage tax, whose signs depend on $T'_y(w, y)$ and $t'_y(y, \ell)$. For example, if $T'_y(w, y)$ is positive, then individuals with the same wage, but slightly higher earnings must pay more taxes. For a given wage rate, higher earnings are possible only by choosing higher labour hours, so, we call $T'_y(w, y) > 0$ a labour tax and $T'_y(w, y) < 0$ a labour subsidy. For the same reason, $t'_y(y, \ell) > 0$ is a wage tax and $t'_y(y, \ell) < 0$ a wage subsidy. For later use, note that the marginal wage tax $t'_y(y, \ell)$ can be expressed in terms of the wage-contingent tax scheme as[26]

$$t'_y(y, \ell) = T'_w(y/\ell, y)/\ell + T'_y(y/\ell, y). \tag{18}$$

Let us now go back to the problem of the planner displayed in equations (10)-(13) and show the main results for the earnings-conditional marginal wage and labour tax

[25] To see this, differentiate both sides of $t(y, \ell) = T(y/\ell, y)$ with respect to ℓ.

[26] To see this, differentiate both sides of $t(y, \ell) = T(y/\ell, y)$ with respect to y.

T'_w and t'_ℓ, the marginal labour tax T'_y, and the marginal wage tax t'_y.

Pareto efficient redistribution from the high to the low type usually implies that the high type remains undistorted at the margin. This result remains to be true for wage-contingent tax schemes.

Proposition 3b. *The earnings-conditional marginal wage tax (T'_w) and the marginal labour tax (T'_y) must be equal to zero for the high type at the optimum ($T'_w(w^*_H, y^*_H) = 0$ and $T'_y(w^*_H, y^*_H) = 0$).*

If high ability types always choose higher wages for the same earnings, then fiscally discouraging a choice of higher wages will deter mimicking and enhances redistributive efficiency. The deterring role played by wage controls in the previous section is taken over here by earnings-conditional wage or labour taxation. Proposition 3c summarizes.

Proposition 3c. *Taxing wages (resp. labour hours) conditional on earnings for the low type, i.e., $T'_w(w^*_L, y^*_L) > 0$ (resp. $t'_\ell(y^*_L, \ell^*_L) > 0$) will enhance redistributive efficiency if high ability types always prefer to work at higher wage rates (resp. higher labour hours) for the same earnings, i.e., if $MRS\hat{W}'_\theta < 0$ (resp. $MRS\hat{W}'_\theta > 0$) holds.*

If high ability types always choose higher labour hours for the same wage rate ($MRSL'_\theta < 0$), then it will be optimal to tax labour at the margin (and vice-versa). Similarly, if high ability types always choose higher wage rates for the same amount of labour hours ($MRSW'_\theta < 0$), then it will be optimal to tax wages at the margin (and vice-versa).

Proposition 3d. *Taxing (resp. subsidizing) labour for the low type, i.e., $T'_y(w^*_L, y^*_L) > 0$ (resp. $T'_y(w^*_L, y^*_L) < 0$) is optimal if high ability types always choose higher labour hours for the same wage rate, i.e., if $MRSL'_\theta < 0$ (resp. $MRSL'_\theta > 0$) holds.*

Proposition 3e. *Taxing (resp. subsidizing) wages for the low type, i.e., $t'_y(y^*_L, \ell^*_L) > 0$ (resp. $t'_y(y^*_L, \ell^*_L) < 0$) is optimal if high ability types always choose higher wage rates for the same amount of labour hours, i.e., $MRSW'_\theta < 0$ (resp. $MRSW'_\theta > 0$) holds.*

To finish this section, Table 2 plugs in the different results of Proposition 3 for the low earners in Table 1.

Although ultimately an empirical question, case b seems unlikely. Therefore, if case a holds, we are left with two possibilities, called $(1^+, a)$ and $(2^+, a)$ before. In both cases, an earnings-conditional wage tax—or equivalently, an earnings-conditional labour

Table 2: An overview of the results for the low earners

$MRS\hat{Y}^{*\prime}_{\theta} < 0$	$MRSW'_{\theta} < 0$		$MRSW'_{\theta} > 0$	
case +	$MRSL'_{\theta} < 0$ case 1	$MRSL'_{\theta} > 0$ case 2	$MRSL'_{\theta} < 0$ case 3	$MRSL'_{\theta} > 0$ case 4
$MRS\hat{W}'_{\theta} < 0$ case a	wage tax $\mid y^*_L$ wage tax labour tax	wage tax $\mid y^*_L$ wage tax labour subsidy	(iii)	(i)
$MRS\hat{W}'_{\theta} > 0$ case b	labour tax $\mid y^*_L$ wage tax labour tax	(ii)	labour tax $\mid y^*_L$ wage subsidy labour tax	(i)

Combining cases (i) 4 with +, (ii) 2 with b, and (iii) 3 with a is not possible

subsidy—is optimal. So, individuals who earn a low income with a higher wage rate (and thus a lower amount of labour hours) will pay more taxes. In addition, if case $(2^+, a)$ applies, then an unconditional labour subsidy is optimal, i.e., marginally subsidizing the earnings of individuals with both low earnings and low wages.[27]

Both results point in the direction of (earnings-conditional) labour subsidies rather than wage subsidies.[28] They can be used to justify the tax credit schemes implemented in Belgium and France (eligibility based on low earnings and a low wage rate) and Ireland, New Zealand, and the United Kingdom (eligibility based on low earnings and sufficient labour hours which implies a low wage rate for the same earnings). In contrast, classical tax credits—subsidizing the earnings of the working poor irrespective of the margin they use to earn more—can never be optimal.

[27] Figure 6 reveals indeed that in case $(2^+, a)$ the low type has both lower earnings and a lower wage rate compared to the high type.

[28] Because case a is needed to justify minimum labour requirements and case b to justify minimum wages, one could similarly state that both results point in the direction of minimal labour rather than minimum wage controls.

5 Conclusion

I analyze wage policies in a perfectly competitive labour market. Jobs in the labour market are described by their intensity (labour effort) and duration (labour hours). Some interesting implications result.

First, the theoretical relation between minimum wages and employment turns out to be more complicated than in the standard neoclassical textbook model. Employment first increases for sufficiently small binding minimum wages, but decreases afterwards. Although not the main aim of the current paper, it is a striking result that is worth investigating further in my view. For example, the demand side in the current version of the model is rather primitive as it assumes a perfectly elastic demand for each job. It could be interesting to develop a compensating wage differential model in the spirit of Rosen (1986) to allow for richer general equilibrium effects.

Second, wage policies can add value to an optimal non-linear earnings tax scheme. Not all wage policies are on the same footing, however. In theory, wage-contingent earnings taxation is superior to the combination of wage controls and classical earnings taxation. Yet, it is not clear at this stage whether such schemes can be implemented without excessive tax evasion. On the one hand, Banks and Diamond (2010) correctly point out that firms and workers could agree to report the same gross earnings to the tax authorities, but a different combination of wages and labour hours and share the reduced tax burden. On the other hand, Kleven, Kreiner, and Saez (2009) show that whistleblowing can be a successful mechanism to reduce tax evasion even with low penalties and low audit rates.

Finally, the current paper provides qualitative results only. Quantitative results could supplement the current analysis, but require a structural labour supply model that goes beyond the usual labour hours margin (see, e.g., Aaberge and Columbino, 2014 for an overview). Such an estimated structural model could allow for example to verify the impact of minimum wages on employment, to design optimal wage policies, and to quantify the realized welfare gains overl classical taxation. This is left for future research.

References

[1] Aaberge, R., Flood, L., 2012, U.S. versus Sweden: the effect of alternative in-work tax credit policies on labour supply of single mothers, mimeo (June, 17, 2012).

[2] Aaberge, R., Columbino, U., 2014, Labour supply models, in, O'Donoghue, C., ed., *Handbook of Microsimulation Modelling, Contributions to Economic Analysis* Volume 293, Bingley, UK: Emerald.

[3] Ahn, T., Arcidiacono, P., Wessels, W., 2011, The distributional impacts of minimum wage increases when both labor supply and labor demand are endogenous, *Journal of Business & Economic Statistics* 29(1), 12–23.

[4] Albouy, D., 2009, The unequal geographical burden of federal taxation, *Journal of Political Economy* 117(4), 635-667.

[5] Allen, S., P., 1987, Taxes, redistribution, and the minimum wage: A theoretical analysis, *Quarterly Journal of Economics* 102(3), 477–490.

[6] Allingham, M.G., 1975, Towards an ability tax, *Journal of Public Economics* 4, 361-376.

[7] Banks, J, Diamond, P., 2010, The case for direct taxation, chapter 6 in Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., Poterba, J., eds., *The Mirrlees Review*, Volume 1 (Dimensions of Tax Design), Oxford: Oxford University Press.

[8] Bhaskar, V., Manning, A., To, T., 2002, Oligopsony and monopsonistic competition in labor markets, *Journal of Economic Perspectives* 16(2), 155–174.

[9] Beaudry, P., Blackorby, C., Szalay, D., 2009, Taxes and employment subsidies in optimal redistribution programs, *American Economic Review* 99(1), 216-242.

[10] Belman, D., Wolfson, P.J., 2014, *What Does the Minimum Wage Do?* Upjohn Institute for Employment Research.

[11] Blomquist, S., 1981, Comparison of tax bases for a personal tax, *Scandinavian Journal of Economics* 83(3), 390–407.

[12] Blomquist, S., 1982, Should educational expenses be deductible? A comparison of taxes in a model where education is a choice variable, *Journal of Public Economics* 18, 65–83.

[13] Blomquist, S., 1984, The wage rate tax: An alternative to the income tax? *Scandinavian Journal of Economics* 86(2), 269–285.

[14] Blomquist, S., Selin, H., 2010, Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates, *Journal of Public Economics* 94, 878–889.

[15] Boadway, R., 2012, *From Optimal Tax Theory to Policy: Retrospective and Prospective views*, Munich lectures in Economics, MIT Press.

[16] Boadway, R., Cuff, K., 1999, Monitoring job search as an instrument for targeting transfers, *International Tax and Public Finance* 6, 317–337.

[17] Boadway, R., Cuff, K., 2001, A minimum wage can be welfare-improving and employment-enhancing, *European Economic Review* 45, 553–576.

[18] Bovenberg, L., Jacobs, B., 2005, Redistribution and education subsidies are Siamese twins, *Journal of Public Economics* 89, 2005–2035.

[19] Brett, C., Weymark, J., 2003, Financing education using optimal redistributive taxation, *Journal of Public Economics* 87, 2549–2569.

[20] Brown, C., 1999, Minimum wages, employment and the distribution of income, chapter 32 in Ashenfelter, O., Card, D., eds., *Handbook of Labor Economics*, Volume 3, North-Holland: Elsevier.

[21] Burdett, K., Mortensen, D., 1998, Wage differentials, employer size, and unemployment, *International Economic Review* 39(2), 257–273.

[22] Cahuc, P., Laroque, G., 2014, Optimal taxation and monopsonistic labor market: does monopsony justify the minimum wage? *Journal of Public Economic Theory* 16(2), 259–273.

[23] Card, D., Krueger, A., 1995, *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton: Princeton University Press.

[24] Dasgupta, P., Hammond, P., 1980, Fully progressive taxation, *Journal of Public Economics* 13, 141–154.

[25] Doucouliagos, H., Stanley, T., 2009, Publication selection bias in minimum-wage research? A meta-regression analysis, *British Journal of Industrial Relations* 47(2), 406-428.

[26] Deltas, G., 2007, Can a minimum wage increase employment and reduce prices in a neoclassical perfect information economy? *Journal of Mathematical Economics* 43, 657–674.

[27] Economic Policy Institute, 2014, Economist statement on the federal minimum wage, http://www.epi.org/minimum-wage-statement.

[28] Flinn, C., 2006, Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates, *Econometrica* 74(4), 1013–1062.

[29] Fuller, D., Geide-Stevenson, D., 2014, Consensus among economists—an update, *The Journal of Economic Education* 45(2), 131-146.

[30] Gentry, W., Hubbard, R., 2004, The effects of progressive income taxation on job turnover, *Journal of Public Economics* 88(11), 2301–2322.

[31] Gerritsen, A., Jacobs, B., 2014, Is a minimum wage an appropriate instrument for redistribution? CESifo working paper 4588.

[32] Glaeser, E., 2013, A better way to help the working poor, *Bloomberg View*, February, 19, 2013. http://www.bloombergview.com/articles/2013-02-20/use-tax-credit-not-minimum-wage-to-help-poor.

[33] Guesnerie, R., Roberts, K., 1987, Minimum wage legislation as a second best policy, *European Economic Review* 31, 490–498.

[34] Harris, L., 2014, Raise the minimum wage? No, subsidize wages instead, *Los Angeles Times*, February, 21, 2014. http://www.latimes.com/opinion/op-ed/la-oe-harris-minimum-wage-20140221-story.html

[35] Hirsch, B., Kaufman, B., and Zelenska, T., 2015, Minimum wage channels of adjustment, *Industrial Relations* 54(2), 199–239.

[36] Hungerbühler, M., Lehmann, E., 2009, On the optimality of a minimum wage: new insights from optimal tax theory, *Journal of Public Economics* 93, 464–481.

[37] Kearl, J., Pope, C., Whiting, G., Wimmer, L., 1979, A confusion of economists? *American Economic Review* 69(2), 28–37.

[38] Kesselman, J.R., 1976, Egalitarianism of earnings and income taxes, *Journal of Public Economics* 5, 285–301.

[39] Klein, D., Dompe, S., 2007, Reasons for supporting the minimum wage: asking signatories of the "Raise the Minimum Wage" statement, *Econ Journal Watch* 4(1), 125–167.

[40] Kleven, H, Kreiner, C., Saez, E., 2009, Why can modern governments tax so much? An agency model of firms as fiscal intermediaries, NBER working paper 15218.

[41] Lee, D., Saez, E., 2012, Optimal minimum wage policy in competitive labor markets, *Journal of Public Economics* 96, 739–749.

[42] Manning, A., 1995, How do we know that real wages are too high? *Quarterly Journal of Economics* 110(4), 1111–1125.

[43] Mirrlees, J., 1971, An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38(114), 175–208.

[44] Neumark, D., 2013, The minimum wage ain't what it used to be, the *New York Times*, December, 9, 2013. http://economix.blogs.nytimes.com/2013/12/09/the-minimum-wage-aint-what-it-used-to-be/?_r=2.

[45] Neumark, D., Wascher, W., 2008, *Minimum Wages*, Cambridge: MIT Press.

[46] Nishimura, Y., 2004, Tax implementability of fair allocations, *Review of Economic Design* 9, 31–41.

[47] Phelps, E, 2013, Nobel-winner Edmund Phelps on his plan to help low wage workers – without raising the minimum wage, *Washington Post*, December, 11, 2013. http://www.washingtonpost.com/blogs/wonkblog/wp/2013/12/11/nobel-winner-edmund-phelps-on-his-plan-to-help-low-wage-workers-without-raising-the-minimum-wage/

[48] Powell, D., 2010, Heterogeneity in income tax incidence: are the wages of dangerous jobs more responsive to tax changes than the wages of save jobs? RAND working paper WR-706-1.

[49] Powell, D., Shan, H., 2012, Income taxes, compensating differentials, and occupational choice: How taxes distort the wage-amenity decision, *American Economic Journal: Economic Policy* 4(1), 224–247.

[50] Rebitzer, J.B., Taylor, L.J., 1995, The consequences of minimum wage laws: some new theoretical ideas, *Journal of Public Economics* 56, 245–255.

[51] Romer, C., 2013, The business of the minimum wage, the *New York Times*, March, 2, 2013. http://www.nytimes.com/2013/03/03/business/the-minimum-wage-employment-and-income-distribution.html.

[52] Rosen, S., 1986, The theory of equalizing differences, chapter 12 in Ashenfelter, O., Layard, R., eds., *Handbook of Labor Economics*, Volume 1, North-Holland: Elsevier.

[53] Saez, E., 2000, Optimal income transfer programs: intensive versus extensive labor supply responses, NBER working paper 7708.

[54] Saez, E., 2002, Optimal income transfer programs: intensive versus extensive labor supply responses, *Quarterly Journal of Economics* 117(3), 1039–1073.

[55] Stiglitz, J.E., 1982, Self-selection and Pareto efficient taxation, *Journal of Public Economics* 17, 213–240.

[56] Tillmann, G., 2005, The equity-efficiency trade-off reconsidered, *Social Choice and Welfare* 24, 63–81.

[57] Tuomala, M., 1986, On the optimal income taxation and educational decisions, *Journal of Public Economics* 30, 183–198.

[58] Wagstaff, P., 1975, A Benthamite wages policy, *Review of Economic Studies* 42(4), 571–580.

6 Proof of Lemma 1

Recall the function $\bar{v}$ defined as $\bar{v}(w,\ell,\theta) = v(w\ell, w, \ell, \theta)$. The unique interior best job for a given ability type θ must satisfy the first and second-order conditions of the problem $\max_{w,\ell\geq 0} \bar{v}(w,\ell,\theta)$, being

$$\bar{v}'_w(w,\ell,\theta) = 0, \tag{19}$$

$$\bar{v}'_\ell(w,\ell,\theta) = 0, \tag{20}$$

and

$$\bar{v}''_w(w,\ell,\theta) < 0, \tag{21}$$

$$\bar{v}''_w(w,\ell,\theta)\bar{v}''_\ell(w,\ell,\theta) - [\bar{v}''_{w\ell}(w,\ell,\theta)]^2 > 0. \tag{22}$$

For later use, note that the second order conditions (21) and (22) imply

$$\bar{v}''_\ell(w,\ell,\theta) < 0, \tag{23}$$

at the optimum. For a given ability type θ, the unique interior best choice of a wage rate conditional on an amount of labour hours must satisfy the first and second-order conditions of the problem $\max_{w\geq 0} \bar{v}(w,\ell,\theta)$, being equations (19) and (21). Similarly, the unique interior best choice of an amount of labour hours conditional on a wage rate for a given ability type θ must satisfy the first and second-order conditions of the problem $\max_{\ell\geq 0} \bar{v}(w,\ell,\theta)$, being equations (20) and (23). Because the global solution $(w(\theta), \ell(\theta))$ satisfies equations (19)-(23), it must also be a solution to each of the conditional problems, so

$$w(\theta) = W(\ell(\theta),\theta) \text{ and } \ell(\theta) = L(w(\theta),\theta), \tag{24}$$

which is equation (1).

The conditional wage, denoted $W(\ell,\theta)$, must satisfy equation (19), so

$$\bar{v}'_w(W(\ell,\theta),\ell,\theta) = 0.$$

Differentiating both sides w.r.t. ℓ leads to

$$W'_\ell(\ell,\theta) = -\frac{\bar{v}''_{w\ell}(W(\ell,\theta),\ell,\theta)}{\bar{v}''_w(W(\ell,\theta),\ell,\theta)}. \tag{25}$$

Similarly, the conditional labour hours $L(w,\theta)$ must satisfy equation (20), so

$$\bar{v}'_\ell(w,L(w,\theta),\theta) = 0.$$

Differentiating both sides w.r.t. w leads to

$$L'_w(w,\theta) = -\frac{\bar{v}''_{w\ell}(w, L(w,\theta),\theta)}{\bar{v}''_{\ell}(w, L(w,\theta),\theta)}. \tag{26}$$

Evaluating expressions (25) and (26) at the global solution $(w(\theta), \ell(\theta))$ and using equation (24) leads to

$$W'_\ell(\ell(\theta),\theta) = -\frac{\bar{v}''_{w\ell}(w(\theta),\ell(\theta),\theta)}{\bar{v}''_{w}(w(\theta),\ell(\theta),\theta)} \quad \text{and} \quad L'_w(w(\theta),\theta) = -\frac{\bar{v}''_{w\ell}(w(\theta),\ell(\theta),\theta)}{\bar{v}''_{\ell}(w(\theta),\ell(\theta),\theta)}.$$

Multiplying both terms yields

$$W'_\ell(\ell(\theta),\theta) \;\times\; L'_w(w(\theta),\theta) = \frac{[\bar{v}''_{w\ell}(w(\theta),\ell(\theta),\theta)]^2}{\bar{v}''_{w}(w(\theta),\ell(\theta),\theta) \;\times\; \bar{v}''_{\ell}(w(\theta),\ell(\theta),\theta)}.$$

The second-order conditions (21), (22), and (23) establish therefore that

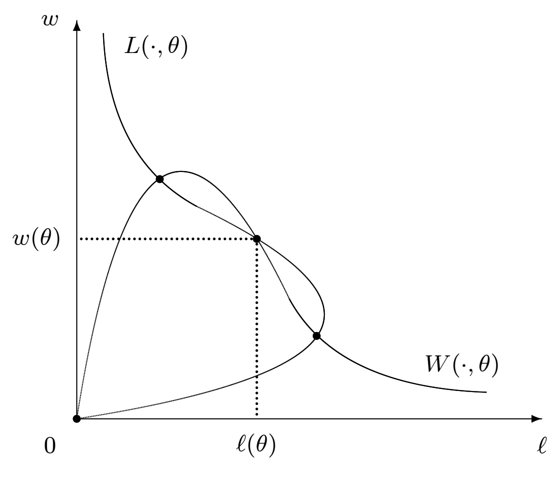
$$0 < W'_\ell(\ell(\theta),\theta) \;\times\; L'_w(w(\theta),\theta) < 1. \tag{27}$$

To obtain equation (2), it suffices to show that the combination $W'_\ell(\ell(\theta),\theta) < 0$ and $L'_w(w(\theta),\theta) < 0$ is not possible. We proceed by contradiction, i.e., suppose $W'_\ell(\ell(\theta),\theta) < 0$ and $L'_w(w(\theta),\theta) < 0$ holds at the optimum. Let L^{-1} be the inverse of L with respect to w; then, the right-hand side inequality in equation (2) can be rewritten as

$$W'_\ell(\ell(\theta),\theta) < L_w^{-1\prime}(\ell(\theta),\theta).$$

Because W'_ℓ and L'_w (and thus also $L_w^{-1\prime}$) are negative at the optimum, W must cut L from below. Figure 7 illustrates the conditional wage curve, the conditional labour curve, and the best job at the (middle) intersection of both curves.

Figure 7: Job choice - a contradiction



By assumption, we have $W(\ell,\theta)=0$ for $\ell=0$, $W(\ell,\theta)>0$ for $\ell>0$, and W continuous in ℓ. The same is true for L, i.e., $L(w,\theta)=0$ for $w=0$, $L(w,\theta)>0$ for $w>0$, and L continuous in w. Therefore, there exists at least two other crossings of the conditional curves, illustrated in Figure 7 by the intersections to the left and right of the middle one. However, the assumptions that the function $\bar{v}$ is strictly quasi-concave and $(w(\theta),\ell(\theta))$ is a unique and interior maximum implies that the function $\bar{v}$ must be single peaked in wages and in labour hours. Any other intersection of the conditional curves corresponds with a minimum or maximum in at least one direction, contradicting singlepeakedness.

Proof of Proposition 2

Before analyzing the program of the planner, consider the job choice problem in equation (5). The solution $\hat{W}(c,y,m,\theta)$ must satisfy the Kuhn-Tucker conditions

$$\hat{v}'_w(c,\hat{W}(c,y,m,\theta),y,\theta)+\delta = 0, \tag{28}$$

$$\delta(\hat{W}(c,y,m,\theta)-m) = 0, \tag{29}$$

with δ the Lagrange multiplier. We must have

$$\hat{v}^{*\prime}_m(c,y,m,\theta)=\hat{v}'_w(c,\hat{W}(c,y,m,\theta),y,\theta)\hat{W}'_m(c,y,m,\theta)=-\delta\hat{W}'_m(c,y,m,\theta)=-\delta, \tag{30}$$

where the first step follows from differentiating both sides of $\hat{v}^*(c,y,m,\theta)=\hat{v}(c,\hat{W}(c,y,m,\theta),y,\theta)$ with respect to m, the second step from equation (28), and the last step from differentiating equation (29) with respect to m. In words, $\hat{v}^{*\prime}_m(c,y,m,\theta)=-\delta$ tells us that the Lagrange multiplier is the marginal utility cost of increasing the minimum wage: it is zero if the minimum wage is not binding and strictly positive otherwise.

The program of the planner is

$$\max_{\{(c_i,y_i)\},m} \hat{v}^*(c_L,y_L,m,\theta_L) \text{ subject to}$$

$$\hat{v}^*(c_H,y_H,m,\theta_H)-\underline{v}\geq 0, \tag{α}$$

$$\hat{v}^*(c_H,y_H,m,\theta_H)-\hat{v}^*(c_L,y_L,m,\theta_H)\geq 0, \tag{β}$$

$$n_L(y_L-c_L)+n_H(y_H-c_H)-R_0\geq 0. \tag{γ}$$

I use α, β, and γ as the (strictly positive) Lagrange multipliers corresponding to the different constraints.[29] The first-order conditions (w.r.t. c_L, y_L, c_H, y_H, m) are

$$\hat{v}_c^{*\prime}(c_L, y_L, m, \theta_L) - \beta \hat{v}_c^{*\prime}(c_L, y_L, m, \theta_H) - \gamma n_L = 0, \quad (31)$$

$$\hat{v}_y^{*\prime}(c_L, y_L, m, \theta_L) - \beta \hat{v}_y^{*\prime}(c_L, y_L, m, \theta_H) + \gamma n_L = 0, \quad (32)$$

$$(\alpha + \beta)\hat{v}_c^{*\prime}(c_H, y_H, m, \theta_H) - \gamma n_H = 0, \quad (33)$$

$$(\alpha + \beta)\hat{v}_y^{*\prime}(c_H, y_H, m, \theta_H) + \gamma n_H = 0, \quad (34)$$

$$-\delta_L - (\alpha + \beta)\delta_H + \beta \delta_M = 0, \quad (35)$$

with

$$\begin{aligned} \delta_L &= -\hat{v}_m^{*\prime}(c_L, y_L, m, \theta_L), \\ \delta_M &= -\hat{v}_m^{*\prime}(c_L, y_L, m, \theta_H), \\ \delta_H &= -\hat{v}_m^{*\prime}(c_H, y_H, m, \theta_H), \end{aligned}$$

the Lagrange multipliers of the individual problem derived via equation (30).

Let (c_L^0, y_L^0) and (c_H^0, y_H^0) denote the second-best redistribution scheme without minimum wages ($m = 0$) and let $w_L^0 = \hat{W}(c_L^0, y_L^0, 0, \theta_L)$, $w_M^0 = \hat{W}(c_L^0, y_L^0, 0, \theta_H)$, and $w_H^0 = \hat{W}(c_H^0, y_H^0, 0, \theta_H)$ be the corresponding preferred wage rates of the low type, the mimicking high type, and the high type. The left-hand side of equation (35) is the change in the maximand—the utility of the low type in our problem—by changing the minimum wage. So, if $w_M^0 < \min\{w_L^0, w_H^0\}$ holds, then at $m = w_M^0$, we have

$$\left.\frac{\partial \mathcal{L}(\cdot)}{\partial m}\right|_{m=w_M^0} = \beta \delta_M = -\beta \hat{v}_m^{*\prime}(c_L^0, y_L^0, w_M^0, \theta_H) > 0,$$

with $\mathcal{L}(\cdot)$ the Lagrange function. In words, introducing a binding minimum wage enhances redistribution by relaxing the self-selection constraint β (Proposition 2a).

As a consequence of Proposition 2a, no binding minimum wage ($\delta_L = \delta_M = \delta_H = 0$) cannot be a solution. It is then easy to verify that the first-order condition (35) requires that we must have $\delta_M > 0$ and ($\delta_L > 0$ or $\delta_H > 0$) at the optimum. In words, the mimicker and at least one other type will be bound by the minimum wage at the optimum (Proposition 2b).

Finally, assume there exists a differentiable tax scheme τ as a function of earnings. The first-order condition for maximizing $\hat{v}^*(y - \tau(y), y, m, \theta)$ w.r.t. y links the marginal tax rate to the marginal rate of substitution (Stiglitz, 1982), being

$$\tau'(y) = 1 + \frac{\hat{v}_y^{*\prime}(c, y, m, \theta)}{\hat{v}_c^{*\prime}(c, y, m, \theta)}, \quad (36)$$

[29] All inequality constraints can be assumed to be binding (in case of α) or have to be binding (in case of β and γ) at the second-best optimum.

with $c = y - \tau(y)$. Using equation (36) together with equations (33)-(34) we obtain directly

$$\tau'(y_H) = 1 + \frac{\hat{v}_y^{*\prime}(c_H, y_H, m, \theta_H)}{\hat{v}_c^{*\prime}(c_H, y_H, m, \theta_H)} = 0.$$

Similarly, use equation (36) together with equations (31)-(32) to obtain

$$\tau'(y_L) = -\frac{\beta\gamma n_L}{\hat{v}_c^{*\prime}(c_L, y_L, m, \theta_H)}[MRS\hat{Y}^*(c_L, y_L, m, \theta_H) - MRS\hat{Y}^*(c_L, y_L, m, \theta_L)]. \quad (37)$$

The single crossing condition $MRS\hat{Y}_\theta^{*\prime} < 0$ implies that the right hand side of equation (37) is strictly positive. In brief, the high type is not distorted at the margin, while the low type faces a strictly positive marginal tax rate at the optimum (Proposition 2c).

Proof of Proposition 3

Consider the program

$$\max_{\{(c_i, w_i, y_i)\}} \hat{v}(c_L, w_L, y_L, \theta_L) \text{ subject to}$$

$$\hat{v}(c_H, w_H, y_H, \theta_H) - \underline{v} \geq 0, \quad (\alpha)$$

$$\hat{v}(c_H, w_H, y_H, \theta_H) - \hat{v}(c_L, w_L, y_L, \theta_H) \geq 0, \quad (\beta)$$

$$n_L(y_L - c_L) + n_H(y_H - c_H) - R_0 \geq 0. \quad (\gamma)$$

Using α, β, and γ as the Lagrange multipliers of the corresponding constraints, the first-order conditions (w.r.t. $c_L, w_L, y_L, c_H, w_H, y_H$) become

$$\hat{v}_c'(c_L, w_L, y_L, \theta_L) - \beta\hat{v}_c'(c_L, w_L, y_L, \theta_H) - \gamma n_L = 0 \quad (38)$$

$$\hat{v}_w'(c_L, w_L, y_L, \theta_L) - \beta\hat{v}_w'(c_L, w_L, y_L, \theta_H) = 0 \quad (39)$$

$$\hat{v}_y'(c_L, w_L, y_L, \theta_L) - \beta\hat{v}_y'(c_L, w_L, y_L, \theta_H) + \gamma n_L = 0 \quad (40)$$

$$(\alpha + \beta)\hat{v}_c'(c_H, w_H, y_H, \theta_H) - \gamma n_H = 0 \quad (41)$$

$$(\alpha + \beta)\hat{v}_w'(c_H, w_H, y_H, \theta_H) = 0 \quad (42)$$

$$(\alpha + \beta)\hat{v}_y'(c_H, w_H, y_H, \theta_H) + \gamma n_H = 0 \quad (43)$$

In addition, with a differentiable wage-contingent tax scheme, the first-order conditions for maximizing $\hat{v}(y - T(w, y), w, y, \theta)$ again link the marginal tax rates to the marginal rates of substitution, more precisely

$$T_w'(w, y) = \frac{\hat{v}_w'(c, w, y, \theta)}{\hat{v}_c'(c, w, y, \theta)} \text{ and } T_y'(w, y) = 1 + \frac{\hat{v}_y'(c, w, y, \theta)}{\hat{v}_c'(c, w, y, \theta)}, \quad (44)$$

using $c = y - T(w, y)$.

Using equation (44) together with equations (41)-(43) we obtain directly

$$T'_w(w_H, y_H) = \frac{\hat{v}'_w(c_H, w_H, y_H, \theta_H)}{\hat{v}'_c(c_H, w_H, y_H, \theta_H)} = 0 \text{ and } T'_y(w_H, y_H) = 1 + \frac{\hat{v}'_y(c_H, w_H, y_H, \theta_H)}{\hat{v}'_c(c_H, w_H, y_H, \theta_H)} = 0,$$

for the high type (Proposition 3b).

Using equation (44) together with equations (38) and (39), we get

$$T'_w(w_L, y_L) = -\frac{\beta \hat{v}'_c(c_L, w_L, y_L, \theta_H)}{\gamma n_L}[MRS\hat{W}(c_L, w_L, y_L, \theta_H) - MRS\hat{W}(c_L, w_L, y_L, \theta_L)]. \tag{45}$$

The sign of $T'_w(w_L, y_L)$ is inversely related to the sign of $MRS\hat{W}'_\theta$ (Proposition 3c).

Using equation (44) together with equations (38) and (40), we get

$$T'_y(w_L, y_L) = -\frac{\beta \hat{v}'_c(c_L, w_L, y_L, \theta_H)}{\gamma n_L w_L}[MRSL(c_L, w_L, \frac{y_L}{w_L}, \theta_H) - MRSL(c_L, w_L, \frac{y_L}{w_L}, \theta_L)], \tag{46}$$

where I use the definition of $\hat{v}$, being $\hat{v}(c, w, y, \theta) = v(c, w, y/w, \theta)$. The sign of $T'_y(w_L, y_L)$ is therefore inversely related to the sign of $MRSL'_\theta$ (Proposition 3d).

Finally, recall from equation (18) that $t'_y(y, \ell)$ is equal to $T'_w(y/\ell, y)/\ell + T'_y(y/\ell, y)$. Plugging in equations (45) and (46) and using equation (17), one obtains

$$t'_y(y_L, \ell_L) = -\frac{\beta \hat{v}'_c(c_L, \frac{y_L}{\ell_L}, y_L, \theta_H)}{\gamma n_L \ell_L}[MRSW(c_L, \frac{y_L}{\ell_L}, \ell_L, \theta_H) - MRSW(c_L, \frac{y_L}{\ell_L}, \ell_L, \theta_L)].$$

So, the sign of $t'_y(y_L, \ell_L)$ is inversely related to the sign of $MRSW'_\theta$ (Proposition 3e).