ESSAYS IN OPTIMAL GOVERNMENT POLICY

ISBN: 978 90 361 0381 7

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 577 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

ESSAYS IN OPTIMAL GOVERNMENT POLICY

Essays over optimaal overheidsbeleid

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 21 februari 2014 om 9:30 uur

door

Adrianus Anthonius Franciscus Gerritsen
geboren te Utrecht

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor: Prof.dr. B. Jacobs

Overige leden: Prof.dr. A.J. Dur
Prof.dr. R.W. Boadway
Prof.dr. F. van der Ploeg

Aan mijn ouders
Adrianus (Adrie) Wilhelmus Gerritsen
Elisabeth (Bets) Francisca Gerritsen-Van Amerongen

Preface

Being a Ph.D. student has had a profound impact on my life, an impact that can hardly be overstated. It allowed me to learn more about economics, public finance, policy making, and political philosophy than I did in the preceding twenty three years combined. It provided me with a framework to analyze issues of an economic or moral nature, enabling me to think and form judgments critically and independently. In short, it allowed me to grow up intellectually. I am fully aware of how privileged I am for having been granted such opportunity, and am grateful beyond words to those who enabled it.

First and foremost, I owe a large debt of gratitude to Bas Jacobs, my supervisor. The way in which he welcomed me as part of his intellectual family has truly been heartwarming. Indeed, the term 'family' is hardly an exaggeration. His devotion to my academic development has been of an intensity and a selflessness commonly only observed among families. He invested his time and effort not just in improving my thesis, but in fostering a personal relationship that made collaboration with him a real pleasure. He had faith in me during times in which I had very little left of it myself, and whenever I was convinced a ceiling had been reached, he always managed to stimulate me to carry my work to a higher level. I could not have wished for a better or more involved supervisor.

I am happy to have had Ruud de Mooij as a second supervisor during the first year of my Ph.D., before he left for the IMF. I look back with fond memories to many coffees, lunches, and dinners I shared with him and Bas, and it is always a great pleasure to see Ruud at any of the international conferences or in Washington, D.C.

Ever since my first international conference, the 2010 meeting of the Canadian Public Economics Group in Quebec City, Robin Boadway has been, for me, the personification of the international community of public economists that greatly enhanced my experience as a Ph.D. student. I am therefore very excited and thankful that he agreed to be a member of my doctoral committee. I also very much appreciate the fact that Robert Dur, Rick van der Ploeg, Otto Swank, and Coen Teulings are willing to invest their scarce time in reading my dissertation and being members of my committee.

Four years would have been a long time to survive without income. I therefore thank The Netherlands Organization for Scientific Research, which provided the financial support for my Ph.D. project under NWO Vidi Grant 452-07-013. I'm grateful to the support staffs of the Tinbergen Institute and the Economics department at the Erasmus University Rotterdam, that made sure my Ph.D. project ran smoothly despite my own disorganized mind. Without diminishing the gratitude I owe to the rest, I especially want to mention Carolien Stolting (TI) and Milky Viola Gonzalez (EUR) for their help with getting my thesis published and defended, and Judith van Kronenberg (TI), whose unfailing cheerfulness is instrumental in making the Tinbergen Institute such a great place to work.

I consider myself extremely lucky for having had Floris Zoutman and Hendrik Vrijburg around me at the Erasmus University. While I knew Floris from my time as a Master student in Groningen, we had not been in touch for a year until happy coincidence brought us back together as office mates. It was thrilling to see how little effect time had had on the fervor, the excitement, the depth of our discussions. Floris is also the direct cause of my biggest disappointment during my period as a Ph.D. student: I vowed to use our four years as office mates to teach him an appreciation of Bob Dylan, at which I miserably failed. Neither did we succeed in building that fort of empty coffee boxes, despite a no less than heroic attempt. On the upside, together we did manage to solve any problem we set our mind to, as long as there was a white board around.

Being two years ahead, Hendrik has always been an important example to me. I greatly admire how he manages to soberly maintain his extraordinary involvement in everything he does, whether it concerns his research or teaching, his family including two young sons, or our common quest of discovering the most shady bar around any of the conference venues we visit. I have grown especially attached to our many coffees and lunches together, which never fail to provide either new insights or higher-order confusion regarding whatever topic that happens to be on the table. If the position of benevolent dictator ever becomes vacant, I would happily recommend both Floris and Hendrik.

During the final months of my Ph.D., Uwe Thuemmel and Alexandra Rusu joined the Erasmus University as the latest scions of tree Jacobs. I would like to wish them both all the best in their academic career. I am looking forward to seeing them more often in the years to come.

I am grateful to Jacques Siegers, of whom I was a research assistant during my time as a Bachelor student in Utrecht. I still fret at the idea of what I would have done without his encouragement when I faced a temporary spell of severe disillusionment with economics. Jacques frequently mentioned that the chief purpose of granting me a research assistantship was to stimulate my interest in economics and academic scholarship. I hope he agrees that this purpose has been served well.

My dissertation is dedicated to my parents who, from my earliest memories on, have stimulated my pursuit of knowledge. It was they, who taught me to read and write long before school would. It was they, who tirelessly tutored me whenever I had difficulties with a subject at school. It was they, who stimulated me to read, and read, and read. And, most importantly, it was they, who instilled in me an appreciation and respect for knowledge, the good, and the beautiful. I am also thankful to my brothers, Willem, Dirk and Tom, who helped me drown any of the superfluous brain cells that my parents provided me with.

My dear Bahar. More than anyone else, you know the highs and the lows I experienced for my Ph.D. What is more, you were always there to sit them through with me. When things went well, you shared in my excitement. When I felt discouraged, you would lift my spirits. I hope you know how grateful I am for your support – it is essential to everything I do.

Aart Gerritsen
München, Winter 2013/2014

‘It came into him life; it went out from him truth. It came to him short-lived actions; it went out from him immortal thoughts. It came to him business; it went from him poetry. It was dead fact; now, it is quick thought. It can stand, and it can go. It now endures, it now flies, it now inspires. Precisely in proportion to the depth of mind from which it issued, so high does it soar, so long does it sing.’

Ralph Waldo Emerson

Contents

Chapter 1

Introduction and summary

One might wonder why anyone is interested in what government ought to do, when 'is' is often so obviously far removed from 'ought.' In fact, I have been pondering that question myself throughout the course of the past four years. While this is not the place for a fully developed discussion of the matter, I believe John Maynard Keynes provided the key to an answer when closing *The General Theory Of Employment, Interest, And Money* with the words: 'soon or late, it is ideas, not vested interests, which are dangerous for good or evil.' Ideas are a necessary input to the legislation of government policy, and the public economist is primarily responsible for the supply and quality of these ideas. In this spirit, my Dissertation can be seen as a collection of ideas on optimal government policy, which I hope will, soon or late, find their way to policy circles and be dangerous for good. In this opening chapter, I aim to introduce these ideas to the reader in two separate parts. In the first part, I describe how my Dissertation fits within the body of ideas that makes up the orthodoxy of public economics. I do this by introducing a number of key concepts of public economics, paying special attention to how my Dissertation agrees with, or departs from, the orthodoxy. The second part of this introduction is devoted to a chapter-by-chapter summary of the most important ideas generated in my Dissertation.

Relation to the orthodoxy of public economics

Public economics

As its title suggests, my Dissertation is meant to contribute to the discussion of what government policy ought to be. This discussion is a central issue within the academic discipline of public economics. Public economics is generally considered to be a subdiscipline of economics, and is taught as such at universities. The central questions underpinning public economics, however, are radically different from the central question of economics. Whereas economics tries to explain human behavior, public economics is concerned with the determination of what government should do. The enormous ambition of this goal is apparent from its multidisciplinary nature. In order to establish what optimal government policy is, one first needs, as a bare minimum, determine what goals a government ought to strive for, what the behavioral responses to government policies are, and how these policies and their behavioral responses affect individuals' well-being. In other words, in determining optimal government policy, the public economist is inescapably conditioned by insights from political philosophy, economics, and the psychology of well-being. Below I briefly sketch which of these insights make up the orthodoxy of public economics, and how my Dissertation relates to them.

Political philosophy or what government ought to strive for

To determine what constitutes optimal government policy, it is necessary to know what government ought to strive for. This question plays a prominent role within political philosophy, and numerous theories have been offered to answer it. There are theories that argue government policy ought to be focused on so-called 'primary goods' (cf. Rawls, 1971), 'capabilities' (cf. Sen, 1992), minimal intervention (cf. Nozick, 1974), or on minimizing inequalities that originate from a limited number of factors for which individuals cannot be held responsible (Dworkin, 2000). By far the most influential theory, at least within public economics, is known as 'welfarism' (cf. Kaplow and Shavell, 2002). Welfarism argues that the ultimate goal of government policy ought to be the maximization of some sum of individuals' well-being.[1]

While welfarism is not wholly uncontroversial, as evident from the large number of competing theories of the 'right' goal of government policy, a number of important arguments are made in its favor. It could be argued that welfarism describes the goal for government policy that would be chosen by someone behind the so-called 'veil of ignorance' – a hypothetical prenatal state in which he is wholly ignorant about the particular life he would lead. Contrary to the other

[1]Welfarism does not necessarily imply that the goal of government should be an unweighted sum of individuals' well-being. Instead, it allows for worse-off individuals to have a larger weight than others and is therefore slightly more general than classical utilitarianism (Bentham, 1907; Mill, 1863; Sidgwick, 1874).

theories I mentioned, it can also be shown never to lead to a form of rule fetishism by favoring a state of the world in which every person is worse off just because it complies with the theory's prescribed goal. For these and other reasons I adopt welfarism as a normative guide to government policy in my Dissertation: I assume that optimal government policy maximizes a sum of individual well-being.[2]

Economics or the behavioral effects of government policy

An answer to the question what government ought to strive for is still far removed from an answer to the question what optimal government policy is. We first need to understand how government policy instruments can contribute to the goal of higher individual well-being. The behavioral effects of government policy are of crucial importance to this understanding. Suppose redistribution from the employed to the unemployed could be considered as raising aggregate well-being. In itself this does not imply anything for the desirability of higher unemployment benefits. While on the one hand such benefits would enhance aggregate well-being due to the redistribution, on the other hand it distorts the incentives of the unemployed to search for a new job. Such distortion leads to lower labor supply and hence lower tax revenue. Because of such behavioral effect, for every euro a government taxes away from the employed it can redistribute less than a euro towards the unemployed. Such a trade-off characterizes virtually every policy instrument. A minimum wage, for example, redistributes income from firms to employees. Although in itself this might be considered as improving aggregate well-being, it also leads to a behavioral response of firms, which will decide to hire less low-skilled labor in order to maintain profit margins.

Economics is particularly well suited to identify and quantify these trade-offs. After all, economics is concerned with the explanation and prediction of human behavior. Economists make extensive use of models for the identification of policies' behavioral consequences. Such models are highly stylized representations of reality, stripped of any non-essential elements. This approach is often criticized by non-economists, for both good and bad reasons. On the one hand, a high level of abstraction is essential to be able to say anything sensible about an extremely

[2]This is not to say that there are no problematic issues with welfarism. One of the most important concerns about welfarism is that it could be used as a justification of phenomena like racism as long as enough people derive well-being from it. This is not an issue in my Dissertation, as I assume that an individual's well-being is independent from other people. Indeed, I believe government policy ought to be concerned with so-called 'laundered' well-being – that part of an individual's well-being that is independent from others (cf. Harsanyi, 1982).

complex reality. On the other hand, an economist should always be able to defend his decisions to abstract from specific aspects of reality. One important abstraction, which is not always easy to defend but is nevertheless part of the orthodoxy of public economics, is the assumption that prices and wages are fully flexible. Under this assumption, prices and wages always adjust to equalize demand and supply – implying that involuntary unemployment cannot exist. In much of my Dissertation, I take issue with this assumption and determine how optimal government policy is affected by rejecting the notion of flexible wages.

Psychology of well-being

Finally, even when we accept welfarism and when economic science perfectly informs us about the exact behavioral effects of government policy, we still do not know what optimal government policy is. This requires insight into how policy, and its behavioral consequences, affect the well-being of people. The explanation of mental processes, such as well-being, is part of the academic discipline of psychology. The consistent application of welfarism thus demands solid knowledge of the psychology of well-being. An elegant trick to escape this responsibility, a trick that is routinely performed by the orthodoxy of public economics, is to assume that people always act to maximize their own well-being. Under this assumption no person ever behaves against his own interest, such that there is no need for government policy to 'correct' individual behavior. Moreover, this assumption allows economists to identify the well-being effects of government policy on the basis of observed behavior alone, without a need for any knowledge of mental processes.

The Dissertation

In my Dissertation, I explicitly challenge a number of assumptions that are deeply rooted in the orthodoxy of public economics. First of all, I reject the notion that human behavior is always and everywhere a matter of well-being maximization. This rejection is based on insights from psychology, behavioral economics, and neuroscience. Apart from these insights, introspection reveals that much of human behavior is motivated by factors that have little to do with the maximization of well-being. When behavior is not necessarily based on well-being maximization, it is possible that an individual is outside his well-being optimum. This creates a new role for government policy, namely to 'correct' individual behavior. In Chapter 2 of the Dissertation, I derive the consequences for optimal income taxation.

Chapters 3 and 4 focus on the minimum wage as a policy instrument. An important difference with earlier literature is that I explicitly model the educational decisions of individuals, a decision from which most other studies abstract. The effect of a minimum wage on this decision is ambiguous: a higher wage rate for low-skilled workers creates incentives to invest less in education, but higher unemployment among the low-skilled creates incentives to invest more in education. In Chapter 3, I analyze how the effect of a minimum wage on educational choice affects the desirability of a minimum wage, relative to income taxation.

Chapter 4 contributes to the literature on minimum wages by rejecting another often-made assumption. This assumption concerns the distribution among low-skilled workers of the involuntary unemployment caused by a minimum wage. Earlier literature commonly assumes that laid-off workers are the ones that derive least well-being from their job (so-called 'efficient rationing'), or, as in Chapter 3, that every low-skilled worker has an equal chance of becoming involuntarily unemployed (so-called 'uniform rationing'). However, there is hardly any theoretical or empirical justification for any of such assumptions. In Chapter 4, I do not impose any assumption on the distribution of unemployment, but analyze how the desirability of a minimum wage depends on this distribution. Moreover, I show how this desirability can be empirically tested without making assumptions on the distribution of unemployment.

Finally, in Chapter 5, I analyze the consequences of inflexible wages for optimal income taxation and labor-participation policies. Just as in the previous two chapters, I deviate from the orthodox assumption that wages perfectly adjust to equate demand and supply. The analysis of this chapter is particularly relevant, for example, for labor markets that are dominated by labor unions, whose wage demands are not aimed at equating demand and supply. Another situation in which the results of this chapter carries special relevance, is during times of economic recession. It is abundantly clear that involuntary unemployment tends to be extraordinarily high during recessions, which implies that during such times the typical assumption of flexible wages potentially disqualifies much of the insights generated by the orthodoxy of public economics.

Findings on optimal government policy

Optimal taxation and well-being (Chapter 2)

In Chapter 2 of my Dissertation I reject the commonly made assumption that individuals, when making decisions, always maximize their well-being. As is common in most of the economic literature, I do assume that individual behavior can be described as a maximization of utility. However, I reject the notion that utility is identical to well-being – that which ultimately makes life valuable. This implies that individuals can make 'mistakes' in their own behavior by acting against their own interest.

I subsequently determine how this affects the structure of optimal income taxation. Taxes can only be optimal if the net social gains of a small increase in the tax rate are nil. After all, if the net social gains would be strictly positive (negative), it would be socially beneficial to raise (lower) the tax rate. The social gains of taxation consist of what the revenues are used for: investments in public goods, for example, or redistribution towards low-income workers. The social costs consist of a lower income of the taxpayer, and a distortion of individual behavior.

This distortion of individual behavior is crucial for my research. A small increase of the income tax rate causes workers to work less. This reduction in labor supply leads to lower tax revenues, which constitutes a social cost. When individuals maximize their well-being, such a reduction of labor supply does not lead to changes in individual well-being. The reason for this is that individual well-being maximization implies that a small change in behavior does not affect well-being. This changes when I reject the assumption of individual well-being maximization. In that case, a reduction of labor supply also has a direct effect on the well-being of the individual. It leads to higher well-being if workers originally worked harder than good for them; it leads to lower well-being if workers were originally working less than good for them.

This creates an additional motive for the use of distortionary taxation. The income of workers that work 'too much' should be subject to a higher marginal tax rate to incentivize them to work less. Similarly, the income of workers that work 'too little' should be subject to a lower marginal tax rate to incentivize them to work harder. The same logic applies to taxes and subsidies on education or specific consumption goods. When people generally decide to invest less in education than what well-being maximization would prescribe, education subsidies should be higher compared to the benchmark case in which individuals maximize well-being.

Whether people indeed work 'too much' or 'too little' is an empirical matter. So in order to provide an answer to this question I require a measure of well-being. In my research I use the answer, on a scale from 1 to 7, to the question: "How dissatisfied or satisfied are you with your life overall?". Thanks to the British Household Panel Survey, I know the answer to this question for about 28,000 British individuals, followed over 12 years. On top of that, I know a great many other details of these individuals, including net income and the usual number of weekly working hours. On the basis of these data I determine how well-being depends on both net income and working hours, which allows me to determine the optimal number of working hours for every individual. Comparing this optimal number of hours to the actually worked hours, I conclude whether an individual works too much or too little.

The results of this empirical exercise lead me to the conclusion that British workers with a relatively low income work, on average, too little. They could enhance their own well-being by making more working hours each week. For workers with a relatively high income the opposite holds: they make too many hours each week, harming their own well-being. The policy implication of these results is that the British government should implement a lower marginal income tax for low-income workers, and a higher marginal income tax for high-income workers. Such a reform provides incentives to workers with a low income to make more hours, and to workers with a high income to make less hours, thereby correcting the well-being-suboptimal behavior of British workers.

Optimal minimum wages and education (Chapter 3)

Chapter 3 deals with the welfare effects of minimum-wage legislation. It deviates from most of the literature on minimum wages by explicitly modeling individuals' educational decisions. This innovation is motivated by the fact that the effect of a minimum wage on schooling is *a priori* ambiguous. On the one hand, a minimum wage raises the wages of low-skilled workers, thereby providing an incentive to invest less in education. On the other hand, a minimum wage leads to higher unemployment among the low-skilled as they becomes more expensive to hire. Such higher unemployment provides an incentive to invest more in education.

We[3] first show that the net effect of a minimum wage on education depends crucially on the ease with which firms can substitute high-skilled for low-skilled

[3]This chapter is joint work with Bas Jacobs.

labor. The easier this is, the more a firm will substitute high-skilled for low-skilled workers when the latter become more costly due to a minimum wage. The elasticity of substitution between high- and low-skilled workers – the percentage change in the number of high-skilled workers relative to low-skilled workers, as a response to a one-percent change in the high-skilled wage rate relative to the low-skilled wage rate – appears to be of crucial importance. If this elasticity is smaller than 1, a minimum wage leads to relatively little unemployment and therefore to less schooling. If this elasticity is larger than 1, a minimum wage leads to so much unemployment among the low-skilled that more people decide to become high-skilled worker, despite the increase in low-skilled wages. Since empirical estimates of this elasticity are generally larger than 1, my model indicates it is plausible that a minimum wage increase leads to more education.

Subsequently, we determine how the schooling decision affects the desirability of a minimum wage. We concentrate on the case of perfect competition on the labor market. Earlier studies showed that a minimum wage might improve economic efficiency in a labor market that is characterized by monopsony power. Our findings are conditional on the assumption that there is no monopsony power. In other words, we focus on the possible redistributive role of a minimum wage, rather than its role of correcting imperfect competition.

Abstracting from taxation, a binding minimum wage has two separate effects on social welfare. On the one hand it leads to higher wages for low-skilled workers, which firms pay for by offering lower wages to the high-skilled. This constitutes a positive welfare effect under the plausible assumption that an increase in low-skilled income is valued more dan a similar increase in high-skilled income. On the other hand, firms will be less inclined to hire low-skilled workers because of the higher wage costs. This leads to increased unemployment, which constitutes a negative effect on social welfare as the involuntary unemployed would rather have a job. The desirability of a minimum wage is therefore ambiguous and depends on the net welfare effect of increased redistribution and higher involuntary unemployment.

The social costs of a minimum wage increase once we introduce income taxation. While the social benefits still consist of the same redistributive gain, the costs of unemployment increase as higher unemployment now also leads to a loss of tax revenue. However, a minimum wage can still complement redistribution through income taxation as long as taxes cannot be perfectly conditioned on the wage rate or skill level. The reason for this is that a minimum wage directly increases the net wage of low-skilled workers and decrease the net wage of high-skilled workers.

Only if government is able to do condition taxes on skill level, it can mimic the redistributive gains of a minimum wage by lowering the tax rate for low-skilled workers, and raising the tax rate for high-skilled workers. In comparison to such a tax reform, a minimum wage still leads to increased low-skilled labor costs and thereby to higher unemployment, while this unemployment leads to higher education. The role of the minimum wage, in such case, does not consist of the redistribution of income, but of counteracting distortions on education, caused by a progressive tax structure. A minimum wage is then desirable if the social benefits of higher eduction are larger than the social costs of higher unemployment.

Optimal minimum wages and the distribution of unemployment (Chapter 4)

Chapter 4 builds on the results of the previous chapter, in which we concluded that a binding minimum wage is desirable if the social benefits of additional schooling outweigh the social costs of higher unemployment. In Chapter 4, we[4] show that this conclusion is generally valid under very weak assumptions. For example, we assume very general utility functions and allow individuals to decide not only on their working hours and educational decisions, but also on whether they are labor-market participants or voluntarily unemployed. Most importantly we relax assumption concerning the so-called rationing schedule. This schedule is a theoretical concept which describes how involuntary unemployment is distributed among the low-skilled. Earlier literature often assumes that rationing is efficient, in the sense that unemployment only affects those workers who derive least well-being from working as a low-skilled worker. Chapter 3, on the other hand, assumes that every low-skilled person faces the same probability of becoming unemployed. In Chapter 4, we abstain from such specific assumptions, and determine the desirability of a minimum wage under a general rationing schedule. In our view, this is an important step because there is hardly any theoretical or empirical foundation for any one specific schedule, while this rationing schedule does appear to be a crucial factor underlying the welfare effects of a minimum wage.

To determine whether a minimum wage is desirable in such a setup, we determine the consequences of a so-called net-income-neutral (NIN) minimum-wage increase. As part of this reform, government raises the minimum wage, while adjusting income tax rates such as to leave the net income of workers constant. The welfare consequences of such a NIN minimum-wage increase give a clear picture of

[4]This chapter is also joint work with Bas Jacobs.

the difference between redistribution through a minimum wage and redistribution through income taxation. As the net wages of workers remain constant, the NIN reform does not affect people's incentives to provide labor effort. Furthermore, as the rewards from working remains constant, the reform does not affect the labor-participation decision. It does, however, raise the costs for firms to hire low-skilled employees, thereby leading to lower demand for low-skilled labor. As low-skilled labor supply remains constant, while demand drops, involuntary unemployment rises. Higher unemployment, however, implies a lower expected gain from low-skilled work, providing incentives to invest more in education. The reform thus leads to both more involuntary unemployment and more high-skilled workers.

The reason why the rationing schedule is so crucial for the welfare effects of a minimum wage is now easy to understand. On the one hand, if much of the unemployment is concentrated among the low-skilled who are rather unemployed than high-skilled, a minimum wage mainly results in higher unemployment without affecting education. However, if unemployment is mainly concentrated among the low-skilled who are relatively indifferent between high- and low-skilled work, a minimum wage has a large positive impact on education. In the chapter we prove that a minimum wage can always be 'made' optimal by adopting the appropriate rationing schedule. In the same vein, a minimum wage can always be made undesirable. For this reason, the desirability of a minimum wage is, from a theoretical point of view, fundamentally ambiguous.

In order to escape this fundamental ambiguity, we rewrite the desirability condition of a minimum wage in terms of so-called *sufficient statistics*. These are statistics which are measurable empirically and allow us to forego the empirical identification of deeper model parameters such as the rationing schedule. We show that the desirability condition can be expressed in three sufficient statistics: the tax revenue associated with an additional high-skilled worker, the tax revenue losses associated with an additional unemployed person, and an elasticity which indicates how the number of high-skilled workers responds to an increase in unemployment.[5] We obtain data on the tax revenue gains and losses from schooling and unemployment from publications by the Organisation for Economic Co-operation and Development (OECD) for a large number of OECD countries. On the basis of these numbers we conclude that, for all OECD countries under consideration except the United States, a NIN increase in the minimum wage is only desirable

[5] Notice that we implicitly assume that direct utility losses from unemployment are negligible. Hereby we write the desirability condition in a way that is highly favorable for a minimum wage.

if a percentage point increase in the unemployment rate leads to a higher school enrollment rate of at least 0.6 percentage point. For the United States, enrollment rates should increase by at least 0.4 percentage point.[6]

Turning to the empirical literature on the effect of unemployment on school enrollment rates, we find that a percentage point increase of the unemployment rate is typically associated with an increase of enrollment rates of between 0.1 and 0.6 percentage points. Thus, from these findings we can conclude that a NIN minimum wage *decrease* is desirable for all countries under consideration – except possibly the United States. Such a reform leads to both more tax revenue and utility gains of individuals who manage to obtain a job due to the lower wage costs. A reduction of the minimum wage is thus part of a so-called Pareto-improving policy reform, leading some individuals to be better off and no one to be worse off. For countries who do not have a legally binding minimum wage – such as Germany, Austria, Italy, or the Nordic countries – implementing one is not desirable.

Optimal taxation and unemployment (Chapter 5)

In the last chapter of my Dissertation I consider the consequences of involuntary unemployment for optimal taxation. As in the previous chapter, I assume that it is *a priori* unknown how unemployment is distributed among the labor force. I show that the presence of involuntary unemployment leads to conclusions on taxation that are diametrically opposed to conventional wisdoms from public economics, in which unemployment is often assumed away. One of those wisdoms holds that increased taxation on the one hand enables a higher degree of redistribution, but on the other hand distorts labor supply. After all, higher taxation leads to lower net wages and thus a reduced incentive to work. Lower labor supply causes tax revenue to decline, which constitutes the welfare loss associated with the distortion. Optimal taxation therefore needs to balance this trade-off: the redistributive value of increased tax revenues versus the efficiency costs of lower labor supply.

Involuntary unemployment, however, implies that the supply of labor exceeds the demand for labor. If, on top of that, there is inefficient rationing, then some of the unemployed would derive more well-being from a job than some of the workers who managed to obtain a job. In that case, higher tax rates still lead to reduced incentives to supply work, causing some workers to work less or not at all. But

[6] See column 5 in Table 1 of Chapter 4 for specific values for all OECD countries for which enough data was available.

since the supply of labor already exceeded the demand of labor, this does not affect aggregate employment. A reduction in the labor supply of some workers leads to employment opportunities for those who were initially unemployed. Since aggregate employment is unaffected, there is no efficiency loss from a decline in tax revenue. On the contrary: efficiency improves as people who derive relatively little well-being from their job decide to work less, thereby creating employment for the unemployed that derive more well-being from working. Similar results apply to unemployment benefits. In the absence of involuntary unemployment, higher benefits lead to less labor supply and thereby to an efficiency loss. However, in the presence of involuntary unemployment, higher benefits lead workers to create jobs for the unemployed who derive more well-being from it. In other words, by raising taxes and benefits government substitutes voluntary unemployment for involuntary unemployment, generating an efficiency gain.

Contrary to the conventional equity-efficiency trade-off, taxes and benefits thus lead to both improved equity and improved efficiency. The policy implications are straightforward: raise income taxes for labor-market segments that are characterized by high involuntary unemployment, and use the additional revenue to increase unemployment benefits. From a dynamic perspective, government should raise income taxes and unemployment benefits in times of high-unemployment recessions, and lower them in more prosperous economic times. These conclusions also bear implications for government policy that is aimed at increasing labor-market participation. If the labor market is characterized by involuntary unemployment, it does not make sense to increase participation. Such policy would only lead people who are just as well off sitting at home, to enter the labor market and compete for jobs with people who do really want a job.

Furthermore, I show to what extent these policy implications are robust to allowing for endogenous educational decisions. The results are comparable to the previous chapter: if unemployment is concentrated among low-skilled workers who are indifferent between being low- or high-skilled, it could in principle be socially beneficial by promoting education. In that case, government might not want to entirely substitute voluntary for involuntary unemployment. I also show that the same policy implications obtain when the low-skilled wage rate is endogenously set by labor unions. If government commits to raising taxes when unemployment rises, labor unions will moderate their wage claims as they anticipate higher taxes in response to higher wages. As a consequence, such a commitment will lead to lower involuntary unemployment. Finally, I show that conventional wisdom on the incidence of taxation no longer holds in the presence of involuntary unemployment. According to this wisdom it is economically irrelevant whether labor income taxation is collected from employees or employers. In both cases the tax would lead to the same decrease in workers' net wages and increase in firms' labor costs. I show that this conventional wisdom is based on the assumption of flexible wages, an assumption that can hardly account for involuntary unemployment. In the presence of involuntary unemployment, higher employee taxes lead to both equity and efficiency gains (as described above), but higher employer taxes lead to even less labor demand and higher unemployment, and therefore to an efficiency loss.

Chapter 2

Optimal taxation when people do not maximize well-being

> ‘Those who know anything about the matter are aware that every writer, from Epicurus to Bentham, who maintained the theory of utility, meant by it ... pleasure itself, together with exemption from pain.’
> *John Stuart Mill in Mill (1863, p.8)*

> ‘The discrediting of *utility* as a psychological concept robbed it of its only possible virtue as an *explanation* of human behaviour in other than a circular sense, revealing its emptiness as even a construction.’
> *Paul A. Samuelson in Samuelson (1938, p.61)*

> ‘In the standard approach, the terms “utility maximization” and “choice” are synonymous. A utility function is always an ordinal index that describes how the individual ranks various outcomes and how he behaves (chooses) given his constraints (available options).’
> *Faruk Gul and Wolfgang Pesendorfer in Gul and Pesendorfer (2008, p.7)*

2.1 Introduction

Historically, the concept of ‘utility’ has been defined in at least two different ways.[1] The classical definition, implicit in the first quotation by John Stuart Mill, is intimately related to the well-being of its subject: utility is seen as the ultimate ‘good’

[1] This chapter is based on Gerritsen (2013*b*).

and, therefore, the natural aim of consequentialist public policy. The second definition, which is dominant in the field of economics since at least the seminal contribution on revealed preference theory by Paul A. Samuelson, and implicit in the second and third quotations, is directly related to individual behavior.[2] In economics, utility is simply defined as that of which the maximization leads to behavior, i.e., it is a rationalization of behavior. Normative economics and, more specifically, public economics and the literature on optimal taxation usually implicitly insist that the two definitions perfectly overlap. That is, individual behavior follows from utility maximization, and this same measure of utility is taken to be the ultimate 'good,' the aim of optimal policy. The implicit assumptions are generally (1) that individuals behave in a way that is consistent with their own well-being, and (2) that individuals' well-being ought to be the goal of public policy.

These assumptions are by no means uncontroversial, but criticism of either of the two assumptions follow from very different lines of reasoning. The latter assumption is a judgment based purely on moral and political philosophy, and is therefore a metaphysical judgment which I do not concern myself with in my Dissertation.[3] The former assumption, that behavior is consistent with maximizing well-being, is a judgment which can, in principle, be assessed empirically. Indeed, increasingly many studies reject the claim that voluntary choice is always and everywhere conducive to well-being. This rejection is based on a variety of insights from the fruitful interactions between economics and neuroscience (e.g., Camerer, Loewenstein and Prelec, 2005), psychology (e.g., Rabin, 1998), and happiness research (e.g., Clark, Frijters and Shields, 2008), as well as on straightforward introspection. Moreover, the divergence between utility and well-being is increasingly stressed by scholars of economic methodology (e.g., Hausman, 2011).

Notice that such rejection in no way invalidates positive economic analysis, which seeks to describe and predict individuals' economic behavior. It is true that, in much of the economic literature, individuals are represented as if they consciously maximize a certain utility function when making decisions. However,

[2]Later in life, especially since the publication of his *Foundations of Economic Analysis* in which he introduces the Bergson-Samuelson social welfare function (Samuelson, 1947), Samuelson came to view utility as more than the empty construct he alludes to in this quotation. For a historical account of his ideas on the Bergson-Samuelson function, see Backhouse (2013).

[3]Like most public economists, I am, however, convinced by and committed to a consequentialist moral philosophy that takes individuals' well-being as its ultimate end – see Chapter 1 of this Dissertation. Nevertheless, most of the theoretical results in this chapter could be reinterpreted in terms of a government that is motivated by other concerns than its subjects' well-being.

nothing in the positive economic analysis of behavior presupposes anything about the cognitive processes that underlie decision making. Economic analysis only presupposes that behavior could be captured in terms of maximizing a utility function – and sometimes not even that. In reality, this could come about because individuals consciously maximize their well-being, because they are genetically wired to behave in such a way, or because they are led to their decisions by their social environment. The nature of the cognitive processes underlying human behavior are irrelevant for its economic analysis. This makes the transition from utility as a positive concept to utility as a normative concept, to put it lightly, non-trivial. If individuals are genetically wired to behave as they do, there is little reason to equate utility with 'well-being' or, indeed, with 'the good.' Thus, even if individuals' decisions can be shown to be compatible with maximizing some function, it does not follow that this function carries the moral import that public economics tends to ascribe to it.[4]

The aim of this chapter is to reassess some standard results of optimal taxation by dropping the assumption that individuals necessarily maximize their own well-being. To avoid confusion, I use the concept of utility as defined in mainstream economics and refer to the aim of consequentialist government policy as well-being. Hence, individuals maximize utility, but not necessarily well-being. As a result, individuals tend to be away from their well-being-maximizing bliss point. Once I relax the assumption of well-being-maximizing agents, utility can no longer function as a moral guide and ordinary optimal tax calculations are flawed. Social welfare is no longer given by a sum of utility, but by a sum of well-being which, like utility, is assumed to depend on the individual's allocation of scarce goods.

The first contribution of this chapter is to derive optimal tax rates – direct and indirect, linear and nonlinear, on the intensive or extensive behavioral margin – and compare them with the standard case in which government maximizes utility. This yields optimal tax formulae that allow for straightforward interpretation. Taxes are set to equate marginal social costs and benefits, which are given by three separate terms: by (i) the social costs of an eroding tax base, (ii) the social benefits of increased redistribution, and (iii) the social costs or benefits of drawing people farther from or closer to their bliss point. The first term is identical to the distortive costs in the standard optimal tax formula. The second term is similar to the redistributional benefits in the standard formula, except that the benefits of

[4]For a further discussion on the irrelevance of cognitive processes for positive economic analysis, see Gul and Pesendorfer (2008).

redistribution are measured in terms of marginal well-being, rather than marginal utility. The third term is absent in the standard case when individuals are always on their bliss point. If, from a well-being perspective, individuals work too much (or consume too much of a certain good), this increases the marginal benefits of taxation as higher marginal tax rates will draw people closer to their bliss point. On the other hand, if individuals work too little (consume too little of a good), the marginal benefits of taxation are lower since higher marginal tax rates exacerbate the preexisting 'mistake' in individual behavior.

The main insight generated by the optimal tax formulae is thus that tax rates should be adjusted to correct for suboptimal individual behavior. The extent of the necessary correction can be expressed in terms of the difference between a utility-based marginal rate of substitution and a well-being-based marginal rate of substitution (MRS). If utility coincides with well-being, the utility-based MRS of, say, leisure for consumption, naturally equals the well-being-based MRS. However, if the well-being-based MRS exceeds the utility-based MRS, substituting leisure for consumption improves well-being even though, by the individual's first-order conditions, such substitution does not affect utility. Thus, a positive (negative) difference between the well-being based MRS and the utility based MRS of good A for good B provides a motive for positive (negative) marginal tax rates on good A.

The second contribution of this chapter is to determine empirically how important the corrective motive for income taxation is. For this, I first determine both the utility-based MRS and the well-being based MRS of leisure for consumption. It is relatively straightforward to measure people's utility-based MRS as it, by definition, equals observed net marginal wages. To measure the well-being based MRS, however, a direct measurement of well-being is necessary. For this, I use questionnaire data on subjective life satisfaction. Specifically, I use a panel dataset of British households to directly estimate well-being as a function of net income, proxying for consumption, and hours worked, among standard control variables and time- and person-fixed effects. Based on parametric and nonparametric estimation techniques I find that well-being is highly concave in consumption and single-peaked in working hours. On the basis of these estimations I construct a measure of the well-being based MRS of leisure for consumption. Comparison with the utility-based MRS indicates that low-income workers tend to work too little from a well-being perspective, whereas higher-income workers tend to work too much. Because well-being is highly concave in income, the implications for the optimal income tax schedule are especially pronounced for low-income workers. Compared to standard optimal tax simulations, and provided that workers are not demand-constrained – in which case lower marginal taxes would not help in raising their labor hours – my analysis endorses much lower marginal tax rates at the bottom of the income distribution.

To the best of my knowledge, this is the first attempt to calibrate optimal tax formulae using data on life satisfaction. There is a number of studies that determine optimal taxes when individuals' and government's preferences differ, but they stop short of empirically determining this difference. Similarly, there is a large number of studies on the determinants of well-being, but none of them integrate their findings into a proper public-finance setting. My contribution to the literature on optimal tax theory is closely related to the contributions by Kanbur, Pirttilä and Tuomala (2006) and Blomquist and Micheletto (2006), which both build on seminal work by Seade (1980). Kanbur, Pirttilä, and Tuomala, like Seade, derive the optimal nonlinear income tax schedule in a Mirrlees (1971) setting in which government maximizes something else than utility. Blomquist and Micheletto do the same in the setting of Stiglitz (1982). The conclusions from these studies are directly in line with mine, namely that the standard optimal tax schedule is supplemented with an additional term which depends on the difference between the individual's and government's MRS. I derive the same conclusion for a wider range of settings, including indirect taxation and taxation of discrete labor-supply decisions. In a broader context this study is related to the literature on optimal taxation in the presence of external effects, e.g., (e.g., Sandmo, 1975; Jacobs and De Mooij, 2011). Instead of an externality, individual behavior exhibits an internality, i.e., individuals do not take full account of their decisions' consequences for their own well-being. The optimal tax treatment of internalities, however, is very much comparable to optimal corrective taxation in the presence of externalities.

Since I attempt to merge optimal tax theory with the empirical study of life satisfaction, this chapter is also related to the empirical literature on the determinants of life satisfaction. This literature, by now too vast to comprehensively discuss here, has recently been reviewed by Clark, Frijters and Shields (2008). The study by Layard, Mayraz and Nickell (2008) is of special interest as they attempt to measure individuals' marginal well-being of income, which is a crucial ingredient of optimal tax formulae. However, to determine the nature of the trade-off between consumption and leisure – a crucial ingredient of any exercise in optimal taxation – one also needs to determine the marginal well-being of leisure. To my knowledge only few studies include work effort or hours in their analysis of the determinants of life satisfaction, and with mixed results. For example, Pouwels, Siegers and Vlasblom (2008) conclude that hours worked negatively affects life satisfaction but estimate an equation in which hours worked enter logarithmically, thereby imposing a diminishing marginal well-being of working hours. Knabe and Rätzel (2010) argue that such functional form is counterintuitive but fail to discern an effect of working hours when estimating an equation in which working hours enter quadratically. Booth and Van Ours (2008), using the same data as I, fail to find any relationship between hours worked and life satisfaction. I conjecture that this might be caused by not controlling for job changes and promotions. Indeed, if a promotion improves well-being and is generally followed by longer working days, as one would expect, the regression coefficient on hours worked might capture both a positive promotion effect and a negative hours effect. Controlling for this I indeed find a robust single-peaked relationship between life satisfaction and working hours.

The remainder of the chapter is organized as follows. The first three sections discuss optimal tax formulae when utility differs from well-being. The first section discusses linear income taxation, the second section nonlinear direct and indirect taxation, and the third section taxation of labor-market participation and education. The fourth section determines empirically the difference between the utility-based MRS and the well-being-based MRS and derives the implications for optimal nonlinear income taxation. The fifth section discusses the robustness of these results and the sixth and final section concludes.

2.2 Optimal linear income taxation

2.2.1 Individual utility maximization

To illustrate the basic intuition behind the optimal tax results, I first discuss optimal linear income taxation when government maximizes a sum of well-being. Optimal nonlinear taxation – which yields results very similar to the case of linear taxation – is left for the next section. Assume a mass-one population of individuals. Individuals are heterogeneous with respect to, and denoted by, their ability $n \in \mathcal{N} = [\underline{n}, \overline{n}]$; they are distributed over $\mathcal{N}$ according to cumulative distribution function $F_n \equiv F(n)$, with density $f_n \equiv F'(n)$. Utility of an individual n is denoted as u_n, which might be different from his well-being, denoted as g_n. Utility is assumed to be identical across individuals, and defined over consumption c_n and labor effort l_n:

$$u_n \equiv u\left(c_n, l_n\right), \quad u_c, -u_l > 0, \quad u_{cc} \leq 0, \quad u_{ll} < 0. \tag{2.1}$$

I assume that utility is increasing and concave in consumption, and decreasing and convex in labor effort. Subscripts n indicate that, in equilibrium, consumption and labor effort, and thus utility, can be written as functions of ability n. Individuals are constrained in their behavior by a budget constraint which stipulates that consumption should equal net income. I assume that ability corresponds to the gross wage rate per effective unit of labor, such that the budget constraint is given by:

$$c_n = (1-t)\, n l_n + T. \tag{2.2}$$

Here, t is the income tax rate and T is a non-individualized lump-sum transfer from government. Standard utility maximization implies that the individual's marginal rate of substitution (MRS) of leisure for consumption equals the net wage rate:

$$\frac{-u_{l,n}}{u_{c,n}} \equiv \frac{-u_l\left(c_n, l_n\right)}{u_c\left(c_n, l_n\right)} = (1-t)\, n. \tag{2.3}$$

Together with the budget constraint, this condition determines equilibrium values of consumption and labor effort for every individual $n \in \mathcal{N}$ as functions of the tax instruments. Denoting equilibrium labor effort as $l_n = l_n(t, T)$, I can define the Hicksian (compensated) labor supply elasticity as:

$$\varepsilon_n^c \equiv \left(\frac{\partial l_n}{\partial t} + n l_n \frac{\partial l_n}{\partial T}\right) \frac{1-t}{l_n} < 0. \tag{2.4}$$

Note that the elasticity ε_n^c is a compensated one in the sense that utility, not well-being, is assumed constant.

2.2.2 Social welfare maximization

I assume that individual well-being is, like utility, a function of consumption and labor effort and identical across individuals, such that it is given by:

$$g_n \equiv g\left(c_n, l_n\right). \tag{2.5}$$

I assume that social welfare, $\mathcal{W}$, is given by the integral, over all individuals, of well-being:

$$\mathcal{W} \equiv \int_{\mathcal{N}} g\left(c_n, l_n\right) \mathrm{d}F_n. \tag{2.6}$$

The function g could be, but is not necessarily, equal to a concave function of utility. If it is, such that I can write $g\left(c_n, l_n\right) = \hat{g}\left(u\left(c_n, l_n\right)\right)$, the model collapses to the standard exercise of optimal income taxation. Government income should equal expenditures and its budget constraint can thus be represented as:

$$\mathcal{B} \equiv \int_{\mathcal{N}} \left(tnl_n - T\right) \mathrm{d}F_n = 0. \tag{2.7}$$

Government can set the income tax rate t and the lump-sum transfer T. Optimal policy follows from maximizing social welfare, (2.6), subject to budget constraint (2.7). Denoting the shadow price of the budget constraint by λ, I can thus write for the optimum:

$$\frac{\mathrm{d}\mathcal{W}/\lambda}{\mathrm{d}x} + \frac{\mathrm{d}\mathcal{B}}{\mathrm{d}x} = 0, \quad x \in \{\tau, T\}. \tag{2.8}$$

I follow Diamond (1975) by defining α_n as the monetarized *social marginal well-being of income*,[5] i.e., the social gains associated with individual n receiving one unit of additional income:

$$\alpha_n \equiv \frac{g_{c,n}}{\lambda} + tn\frac{\partial l_n}{\partial T} - \frac{g_{c,n}}{\lambda}\left(\frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1\right)(1-t)\, n\frac{\partial l_n}{\partial T}. \tag{2.9}$$

The social marginal well-being of income consists of three elements. First, higher income leads to a direct welfare gain from higher consumption, $g_{c,n}/\lambda$. Second, for a positive tax rate, $t > 0$, and negative income effects on labor supply, $\partial l_n/\partial T < 0$, higher income leads to a smaller tax base and thus a loss in tax revenue given by $tn\left(\partial l_n/\partial T\right)$. Third, if the well-being-based MRS of leisure for consumption, $-g_{l,n}/g_{c,n}$, exceeds the utility-based MRS, $-u_{l,n}/u_{c,n}$, individual n supplies too much labor. In that case, the negative income effect on labor supply leads to an increase in well-being and thus to an increase in social welfare. Conversely if $-g_{l,n}/g_{c,n} < -u_{l,n}u_{c,n}$, individual n supplies too little labor and the negative income effect on labor supply leads to lower social welfare.

[5]Where Diamond refers to social marginal utility of income, I refer to social marginal well-being of income to stress the fact that social welfare is defined over well-being, not utility.

Substituting the derivatives of the welfare function and government's budget constraint into equation (2.8), and substituting for α_n, yields the following optimality condition for the lump-sum transfer:

$$\overline{\alpha} \equiv \int_{\mathcal{N}} \alpha_n \mathrm{d}F_n = 1. \tag{2.10}$$

Thus, the average social marginal well-being of income, $\overline{\alpha}$, should equal 1, which is the public resources required to marginally increase the lump-sum transfer. This result only differs from the standard result for the optimal lump-sum transfer in that the social marginal well-being of income includes an additional term associated with suboptimal individual behavior.

The optimality condition for the income tax rate t can be written as:

$$\int_{\mathcal{N}} nl_n \left(1 - \alpha_n + \frac{t}{1-t}\varepsilon_n^c - \frac{g_{c,n}}{\lambda}\left(\frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1\right)\varepsilon_n^c\right) \mathrm{d}F_n = 0. \tag{2.11}$$

The equation is entirely standard, except for the last term within brackets. This term gives the direct welfare gains of lower labor effort, resulting from a marginal increase of the tax rate. Naturally, if the welfare function is only a function of the utility function, then $-g_{l,n}/g_{c,n} = -u_{l,n}/u_{c,n}$, and the final term disappears.

2.2.3 Interpretation of the optimal income tax rate

I can now identify three sources of social costs and benefits from above optimality condition: a mechanical effect, a behavioral effect on the tax base, and a behavioral effect which corrects or worsens a possibly suboptimal labor supply decision by individuals. The first two effects are standard, the last one originates from the assumption that individuals do not maximize well-being.

Mechanical effect – The first mechanical effect of a marginal increase in the labor income tax consists of the revenue gain and is given by nl_n, which is the first term within brackets in equation (2.11). The second mechanical effect constitutes the direct welfare loss associated with the drop in individuals' net income due to a higher income tax. This effect is given by the second term in brackets, $-nl_n\alpha_n$.[6] The overall mechanical effect of a higher tax rate is thus given by $nl_n(1-\alpha_n)$.

[6] Here I somewhat stretch the meaning of 'mechanical' to include all effects that are not due to substitution effects. Naturally, α_n represents in part the welfare consequences of any income effects of a tax change.

Behavioral effect on the tax base – An increase in the income tax rate leads individuals to substitute leisure for consumption, eroding the tax base. This represents a welfare loss as long as the income tax rate is positive, $t > 0$. A one percentage-point increase in the tax rate leads to a compensated change in labor supply of $\varepsilon_n^c / (1-t)$ percent; a one percent decrease in labor supply leads to a tax revenue loss equal to tnl_n. Thus, the welfare effect associated with the behavioral effect on the tax base is measured by the product of the two, $nl_n \frac{t}{1-t} \varepsilon_n^c$, which, in equation (2.11), is the third term within brackets.

Behavioral effect on individuals' suboptimal labor supply – The substitution of leisure for consumption leads to an additional welfare gain (loss) by drawing individuals closer to (farther from) their well-being optimal labor supply. Again, a one percentage point increase in the tax rate leads to a decrease in labor supply of $\varepsilon_n^c / (1-t) < 0$ percent; a one percent decrease in labor supply leads to a (monetized) welfare loss due to lower consumption equal of $(1-t)\, nl_n g_{c,n}/\lambda$, and a (monetized) welfare gain of lower labor supply equal to $l_n g_{l,n}/\lambda$. Thus, the welfare effect associated with the behavioral effect on individuals' suboptimal labor supply is measured by the product of the behavioral effect and the net welfare gain. Some straightforward rearranging yields the final term in equation (2.11). This welfare effect of a higher tax rate is positive if the individual supplies too much labor, such that $-g_{l,n}/g_{c,n} > -u_{l,n}/u_{c,n}$. In that case a higher tax rate corrects the suboptimal behavior as it leads him to work less. The opposite holds if workers supply too little labor, such that $-g_{l,n}/g_{c,n} < -u_{l,n}/u_{c,n}$. In that case, higher taxation exacerbates individuals' suboptimal behavior as it incentivizes people to work even less. Naturally, this welfare effect disappears if well-being and utility coincide, such that the well-being based MRS equals the utility-based MRS.

In the tax optimum, these three effects, summed over the entire population $\mathcal{N}$, should equal zero such that no further increase or decrease in the income tax rate could lead to an increase in social welfare. This is exactly what is stated by condition (2.11). To rewrite the optimality condition in a more familiar form, I define the wedge on the labor supply of individual n as follows:

$$\omega_n \equiv \frac{t}{1-t} - \frac{g_{c,n}}{\lambda}\left(\frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1\right). \tag{2.12}$$

The wedge equals the difference between the monetized social welfare gains and private utility gains of one additional hour of labor, as a proportion of the net wage.

It consists of two terms. The first is standard and represents the tax revenue of an additional labor hour. The second term represents the move towards, or away from, the individual's bliss point. I define the income-weighted average of the labor wedge as:

$$\bar{\omega} \equiv \frac{\int_{\mathcal{N}} nl_n \omega_n \mathrm{d}F_n}{\int_{\mathcal{N}} nl_n \mathrm{d}F_n}. \tag{2.13}$$

Following Feldstein (1972), I define the distributional characteristic of the income tax base as the negative of the normalized covariance between labor income nl_n and the social marginal well-being of income:

$$\xi \equiv -\frac{\mathrm{cov}\left[nl_n, \alpha_n\right]}{\int_{\mathcal{N}} nl_n \mathrm{d}F_n \int_{\mathcal{N}} \alpha_n \mathrm{d}F_n} = 1 - \frac{\int_{\mathcal{N}} nl_n \alpha_n \mathrm{d}F_n}{\int_{\mathcal{N}} nl_n \mathrm{d}F_n} \in [0,1]\,, \tag{2.14}$$

where I made use of the first-order condition of the transfer to substitute for $\int_{\mathcal{N}} \alpha_n \mathrm{d}F_n = 1$. Since the distributional characteristic is a normalized covariance, it is necessarily the case that $-1 \leq \xi \leq 1$. However, as the social marginal well-being of income, α_n, is likely to be decreasing with income, nl_n, one could expect ξ to be positive. The larger the distributional characteristic of labor income, and hence the stronger the covariance between α_n and nl_n, the larger the redistributional gains from taxing labor income.

Substituting for ω_n and ξ into (2.11) and rearranging yields:

$$-\frac{\int_{\mathcal{N}} nl_n \omega_n \varepsilon_{lt,n} \mathrm{d}F_n}{\int_{\mathcal{N}} nl_n \mathrm{d}F_n} = \xi. \tag{2.15}$$

Furthermore, assuming that the Hicksian labor supply elasticity is equal for all individuals, such that $\varepsilon_n^c = \varepsilon^c$, I can write:

$$\bar{\omega} = \frac{\xi}{-\varepsilon^c}. \tag{2.16}$$

The formula for the optimal wedge on labor income is virtually identical to the standard formula for the optimal linear income tax derived under the assumption of well-being-maximizing individuals. The optimal labor wedge is determined by the redistributional gains of higher taxation, given by ξ, divided by the magnitude of the substitution effect, given by $-\varepsilon^c$. The only difference is that the wedge now includes a term that indicates the average difference between individuals' actual and well-being-maximizing labor supply. For a given optimal labor wedge, if individuals provide too much labor, taxes should be higher; if individuals provide too little labor, taxes should be lower.

Above derivation suggests a broader implication, which is confirmed in further analysis below. The standard calculation of optimal wedges largely carries over to the case in which well-being does not correspond to utility. However, the total wedge consists of the degree of suboptimal behavior, as well as the tax wedge.

2.3 Optimal non-linear taxation

2.3.1 Direct taxation – single good

In this section I derive optimal non-linear tax formulae in the case that utility does not necessarily correspond to well-being. Utility is still the same function of consumption and labor effort, given by equation (2.1). I assume taxes are conditioned on labor income nl_n, such that the taxes that an individual n pays are given by $T(nl_n)$. His budget constraint is therefore given by:

$$c_n = nl_n - T(nl_n)\,. \tag{2.17}$$

Utility maximization implies that the individual's utility-based marginal rate of substitution of leisure for consumption equals the marginal net wage rate:

$$\frac{-u_l(c_n, l_n)}{u_c(c_n, l_n)} = \left(1 - T'(nl_n)\right) n, \tag{2.18}$$

where $T'(\cdot)$ denotes the marginal tax rate. Together with the budget constraint, and for a given tax schedule $T(\cdot)$, this equation gives consumption and labor supply as a function of ability n.

The social welfare function is still given by equation (2.6). Government maximizes social welfare subject to a budget constraint and an incentive compatibility constraint. The budget constraint can be written as:

$$\mathcal{B} \equiv \int_{\mathcal{N}} T(nl_n)\,\mathrm{d}F_n = \int_{\mathcal{N}} (nl_n - c_n)\,\mathrm{d}F_n = 0, \tag{2.19}$$

where the second equation follows from substituting the individual's budget constraint. I solve for the government's maximization problem by deriving the optimal second-best allocation, which is decentralized by means of the optimal non-linear income tax schedule.[7] Incentive compatibility follows from the individual's first-order condition (2.18). Eliminating the marginal tax schedule from this condition by substituting in the derivatives of the budget constraint and the utility function, I can rewrite the incentive compatibility constraint as:[8]

$$\frac{\mathrm{d}u_n}{\mathrm{d}n} = \frac{-u_l\left(c_n, l_n\right) l_n}{n}. \tag{2.20}$$

The optimal allocation is obtained by maximizing social welfare, (2.6), with respect to c_n and l_n, subject to the budget constraint, (2.19), and the incentive compatibility constraint, (2.20). This can be seen as a problem of optimal control with control variables c_n and l_n, and state variable u_n. The Hamiltonian associated with this maximization problem is given by:

$$\mathcal{H} = \left(g\left(c_n, l_n\right) + \lambda\left(nl_n - c_n\right)\right) f_n - \theta_n \frac{-u_{l,n}\left(c_n, l_n\right) l_n}{n} + \mu_n\left(u_n - u\left(c_n, l_n\right)\right). \tag{2.21}$$

Here, λ, θ_n, and μ_n, are the shadow prices that belong to the resource constraint, the incentive compatibility constraint, and the constraint on utility. The first-order conditions are obtained by taking derivatives of the Hamiltonian. Together

[7]As proved by Mirrlees (1976), feasibility of the implementation requires the adoption of additional assumptions on the allocation and the utility function. Denoting gross labor earnings as $z_n \equiv nl_n$, and the marginal rate of substitution of gross labor earnings for consumption as $s(c_n, z_n, n) \equiv -u_l(c_n, z_n/n)/nu_c(c_n, z_n/n)$, I assume that the following necessary conditions hold:

$$\frac{\partial s(c_n, z_n, n)}{\partial n} < 0; \quad \frac{\mathrm{d}z_n}{\mathrm{d}n} > 0.$$

The first condition is the single-crossing condition which ensures that the marginal rate of substitution of gross income for consumption, evaluated at a fixed bundle of income and consumption, is decreasing in ability. The second, monotonicity, condition requires that, at the optimal allocation, gross income is monotonically increasing with ability.

[8]That is, total differentiating the utility function yields $\mathrm{d}u_n = u_c\left(c_n, l_n\right)\mathrm{d}c_n + u_l\left(c_n, l_n\right)\mathrm{d}l_n$. Total differentiating the budget constraint yields $\mathrm{d}c_n = \left(1 - T'\left(nl_n\right)\right)\left(n\mathrm{d}l_n + l_n\mathrm{d}n\right)$. Using this to substitute for $\mathrm{d}c_n$ in the derivative of the utility function, and substituting for $\left(1 - T'\left(nl_n\right)\right) n$ from the first order condition yields the incentive compatibility constraint.

with the boundary conditions on θ_n, and after rearranging, they are given by:

$$\frac{\partial H}{\partial c_n} = 0 : \quad \left(\frac{g_{c,n} - \lambda}{u_{c,n}}\right) f_n - \frac{\theta_n}{n} \frac{-u_{cl,n} l_n}{u_{c,n}} = \mu_n, \tag{2.22}$$

$$\frac{\partial H}{\partial l_n} = 0 : \quad \left(\frac{g_{l,n} + \lambda n}{u_{l,n}}\right) f_n + \frac{\theta_n}{n} \left(1 + \frac{u_{ll,n} l_n}{u_{l,n}}\right) = \mu_n, \tag{2.23}$$

$$\frac{\partial H}{\partial u_n} = \frac{\mathrm{d}\theta_n}{\mathrm{d}n} : \quad \frac{\mathrm{d}\theta_n}{\mathrm{d}n} = \mu_n, \tag{2.24}$$

$$\lim_{n \to \underline{n}} \theta_n = \lim_{n \to \overline{n}} \theta_n = 0. \tag{2.25}$$

Combining equations (2.22) and (2.23), and substituting in the individual first-order condition, (2.18), yields, after some rearranging, an expression for the optimal marginal income tax schedule:

$$\frac{T'(nl_n)}{1 - T'(nl_n)} = \frac{u_{c,n}\theta_n/\lambda}{n f_n} \left(1 + \frac{\mathrm{d}\ln\left(u_{l,n}/u_{c,n}\right)}{\mathrm{d}\ln l_n}\right) + \frac{g_{c,n}}{\lambda} \left(\frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1\right). \tag{2.26}$$

Hence, the optimal tax schedule is additive in two terms. The first term is identical to the optimal tax schedule in the case that individuals maximize well-being, and represents both the behavioral responses of higher taxation, and the redistributional gains. For a detailed interpretation of this first term, see Mirrlees (1971), Tuomala (1990), or Saez (2001), among others. The second term, which vanishes in the standard optimal tax exercise, is the term of interest for this chapter. This term is familiar from the case of linear taxation (e.g., from equation (2.12)), and indicates whether from a well-being perspective individual n works too much or too little. If $-g_{l,n}/g_{c,n} > -u_{l,n}/u_{c,n}$, an individual works too much, and the second term in equation (2.26) is positive, providing a motive for higher marginal income taxes to 'correct' individual behavior. On the other hand, if $-g_{l,n}/g_{c,n} < -u_{l,n}/u_{c,n}$, workers work too little, providing a motive for lower marginal tax rates.

Analogous to equation (2.12) for linear taxation, I can write the wedge on labor as:

$$\omega_n \equiv \frac{T'(nl_n)}{1 - T'(nl_n)} - \frac{g_{c,n}}{\lambda} \left(\frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1\right), \tag{2.27}$$

such that I can write for the optimal wedge on labor:

$$\omega_n = \frac{u_{c,n}\theta_n/\lambda}{n f_n} \left(1 + \frac{\mathrm{d}\ln\left(u_{l,n}/u_{c,n}\right)}{\mathrm{d}\ln l_n}\right). \tag{2.28}$$

Hence, the optimal labor wedge is virtually identical to the optimal marginal tax schedule derived in, e.g., Mirrlees (1971). The difference, if individuals do not maximize well-being, is that the wedge on labor includes both a tax wedge and a wedge due to suboptimal labor supply decisions.

An additional implication of condition (2.28) is that nonzero marginal income tax rates are optimal even in a first-best world in which government has full information on individuals' ability and markets are complete. If government can observe individuals' ability the incentive compatibility constraint is slack such that $\theta_n = 0$. Hence, by equation (2.28), optimal wedges in the first best equal zero. This implies that marginal taxes are set to perfectly correct individuals' suboptimal behavior. Thus, even if the second fundamental theorem of welfare economics holds, such that any feasible Pareto-efficient allocation can be implemented by means of individualized lump-sum taxes and transfers only, nonzero income-dependent taxation might still be optimal. After all, if individuals make decisions that do not correspond with their own well-being, Pareto efficiency ceases to be a compelling normative requirement.[9]

2.3.2 Indirect taxation – multiple goods

It is relatively straightforward to extend the model to include multiple taxable goods, yielding results which are similar in spirit to the ones derived under direct linear and non-linear taxation. For simplicity, I consider an additional good, x_n, upon which a commodity tax $t_x = t_x(x)$ can be conditioned. Utility is now given by:

$$u_n \equiv u(c_n, x_n, l_n). \tag{2.29}$$

The budget constraint takes account of the additional good purchases and taxation:

$$c_n + x_n + t_x(x_n) = nl_n - T(nl_n), \tag{2.30}$$

[9] Here I implicitly defined a Pareto-efficient allocation as one in which no person's *utility* can be increased without decreasing anyone else's utility. Naturally, one can alternatively define a Pareto-efficient allocation as one in which no person's *well-being* can be increased without decreasing anyone else's well-being. In that case, Pareto efficiency would still be a compelling normative requirement. However, that would imply a refutation of the fundamental theorems of welfare economics as complete markets, perfect competition, and local nonsatiation of preferences, in the absence of taxation, would no longer lead to a Pareto-efficient allocation.

which, together with the following two first-order conditions, describes equilibrium values for c_n, x_n, and l_n:

$$\frac{-u_l\left(c_n, x_n, l_n\right)}{u_c\left(c_n, x_n, l_n\right)} = \left(1 - T'\left(nl_n\right)\right) n, \tag{2.31}$$

$$\frac{u_x\left(c_n, x_n, l_n\right)}{u_c\left(c_n, x_n, l_n\right)} = 1 + t'_x\left(x_n\right). \tag{2.32}$$

Individuals' well-being is described by $g\left(c_n, x_n, l_n\right)$, such that welfare is given by:

$$W = \int_{\mathcal{N}} g\left(c_n, x_n, l_n\right) \mathrm{d}F_n, \tag{2.33}$$

which, to obtain optimal tax formulae, is maximized subject to a budget constraint and an incentive compatibility constraint:

$$\mathcal{B} \equiv \int_{\mathcal{N}} \left(T\left(nl_n\right) + t_x\left(x_n\right)\right) \mathrm{d}F_n = \int_{\mathcal{N}} \left(nl_n - c_n - x_n\right) \mathrm{d}F_n = 0, \tag{2.34}$$

$$\frac{\mathrm{d}u_n}{\mathrm{d}n} = \frac{-u_l\left(c_n, x_n, l_n\right) l_n}{n}. \tag{2.35}$$

Denote the wedge on consumption of good x_n as follows:

$$\omega_n^x \equiv \frac{t'\left(x_n\right)}{1 + t'\left(x_n\right)} + \frac{g_{c,n}}{\lambda}\left(\frac{g_{x,n}/g_{c.n}}{u_{x,n}/u_{c,n}} - 1\right). \tag{2.36}$$

Notice that, if $g_{x,n}/g_{c.n} > u_{x,n}/u_{c,n}$, individual n consumes too little of good x_n. Marginally substituting consumption of x_n for c_n leaves utility unchanged while improving well-being. Thus, even in the absence of indirect taxation, the wedge on x_n is positive if the individual consumes too little of the good, and negative if he consumes too much. Solving the maximization problem in the usual way yields the following expressions for the optimal wedges:

$$\omega_n = \frac{u_c\theta_n/\lambda}{nf_n}\left(1 + \frac{\mathrm{d}\log\left(u_{l,n}/u_{c,n}\right)}{\mathrm{d}\log l_n}\right), \tag{2.37}$$

$$\omega_n^x = \frac{u_c\theta_n/\lambda}{nf_n}\left(\frac{-\mathrm{d}\log\left(u_{x,n}/u_{c,n}\right)}{\mathrm{d}\log l_n}\right). \tag{2.38}$$

The optimal wedge on labor is given by the first equation and simply corresponds to equation (2.28). The optimal wedge on consumption of good x_n is given by the second equation. The right-hand side is the standard result for the optimal marginal tax schedule (e.g., Atkinson and Stiglitz, 1976). It indicates that the wedge on good x_n is proportional to the degree in which the marginal rate of substitution of x_n for c_n changes with labor effort – that is, to the relative complementarity of good x_n with labor. However, even though this is the case for the optimal wedge, the optimal tax schedule depends on whether individuals consume too much or too little of good x_n. If individuals consume too much, the optimal marginal tax schedule $t'_x(x_n) / (1 - t'_x(x_n))$ exceeds the optimal wedge on x_n; if individuals consume too little, the optimal marginal tax schedule is smaller than the optimal wedge.

2.4 The tax treatment of discrete decisions

So far, I only considered labor supply decisions on the intensive margin. The idea that well-being diverges from utility, and thus that individuals make well-being suboptimal decisions, might very well apply to discrete decisions as well. Especially decisions that are difficult to reverse and of a once-and-for-all nature – to follow higher education, to participate in the labor market – might not be made while taking all well-being consequences into account. Adolescents' decisions on higher education – on both participation in and type of higher education – is an obvious example. Both utility and well-being functions are likely to be affected by the decision, as well as an individual's intertemporal budget constraint. Such effects of education are difficult, if not impossible, to comprehend and it is therefore highly plausible that revealed preferences for higher education bear little relation to well-being.

The implications for optimal tax policy are, given the analyses in the previous two sections, theoretically straightforward. Imagine individuals neglect, on a net basis, the well-being benefits of participation in the labor market or of higher education. This provides a motive to stimulate these descrete decisions by lower participation taxes and higher education subsidies. These motives are seperate from and in addition to the standard motives for participation taxes and education subsidies. In this section I illustrate these points by developing a simple model that encompasses participation and schooling decisions. See Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation) for a more extensive model that also incorporates intensive labor supply decisions and involuntary unemployment but in which well-being and utility are assumed to be identical.

2.4.1 Individual utility maximization

Assume individuals can decide between three alternatives: (i) to be voluntarily unemployed and earn unemployment benefits b, (ii) to be a low-skilled worker, earn income w_L, pay taxes t_L, and incur disutility of participation $q_L(n)$, or (iii) to be a high-skilled worker, earn income w_H, pay taxes t_H, and incur disutility of schooling $q_H(n)$. Thus, disutility of labor is a function of individuals' ability.[10] I assume $q_H'(n) < q_L'(n) < 0$. Thus, disutility is decreasing in ability, and at a faster rate for high-skilled workers than for low-skilled workers, which is sufficient for participation and education decisions to be well-behaved. Utility for each seperate decision is given by:

$$u_U \equiv u\left(b\right), \tag{2.39}$$

$$u_L \equiv u\left(w_L - t_L\right) - q_L\left(n\right), \tag{2.40}$$

$$u_H \equiv u\left(w_H - t_H\right) - q_H\left(n\right). \tag{2.41}$$

There are two critical ability levels. An individual with ability level n_L is indifferent between being non-participant or low-skilled; an individual with ability level n_H is indifferent between being low-skilled or high-skilled. Hence, these critical levels are determined by the following two conditions:

$$n_L: \quad q_L\left(n_L\right) = u\left(w_L - t_L\right) - u\left(b\right), \tag{2.42}$$

$$n_H: \quad q_H\left(n_H\right) - q_L\left(n_H\right) = u\left(w_H - t_H\right) - u\left(w_L - t_L\right). \tag{2.43}$$

In equilibrium, individuals $n \in [\underline{n}, n_L)$ are unemployed, individuals $n \in [n_L, n_H)$ are low-skilled, and individuals with $n \in [n_H, \overline{n}]$ are high-skilled. Thus, the number of unemployed, low-skilled, and high-skilled are given by F_{n_L}, $F_{n_H} - F_{n_L}$, and $1 - F_{n_H}$. I furthermore define the following semi-elasticities:

$$\eta_{Li} \equiv \frac{dF_{n_L}/F_{n_L}}{di}, \quad \eta_{Hi} \equiv \frac{-dF_{n_H}/\left(1 - F_{n_H}\right)}{di}, \quad i \in \left\{b, t_L, t_H\right\}. \tag{2.44}$$

Thus, η_{Li} gives the relative change in the number of unemployed due to a marginal change in tax instrument i, and η_{Hi} gives the relative change in the number of high-skilled workers.

[10]It is convenient to model disutility of work, rather than labor earnings, as a function of ability. This way, all individuals of the same labor type earn the same income and I can ignore within-group income redistribution.

2.4.2 Social welfare maximization

For simplicity I assume that the utility of consumption $u(\cdot)$ corresponds to well-being. However, the disutility of participation $q_L(\cdot)$ and of education $q_H(\cdot)$ might not correspond to well-being. The well-being losses of participation and education are denoted $h_L(\cdot)$ and $h_H(\cdot)$. I denote the difference between utility and well-being losses as:

$$\Delta_i(n) \equiv q_i(n) - h_i(n), \quad i \in \{L, H\}. \tag{2.45}$$

If this difference is positive, individuals behave as if the well-being costs of participation (or education) are larger than they actually are. The social welfare function is thus given by:

(2.46)

$$\mathcal{W} = \int_{\underline{n}}^{n_L} u(b)\, dF_n + \int_{n_L}^{n_H} \left(u(w_L - t_L) - h_L(n)\right) dF_n + \int_{n_H}^{\overline{n}} u(w_H - t_H) - h_H(n)\, dF_n.$$

Government's budget constraint is given by:

$$\mathcal{B} = -\int_{\underline{n}}^{n_L} b dF_n + \int_{n_L}^{n_H} t_L dF_n + \int_{n_H}^{\overline{n}} t_H dF_n = 0. \tag{2.47}$$

I denote the shadow-price of the budget constraint – the marginal social value of government resources – as λ. The wedges on participation and education measure the social welfare gains associated with a marginal increase in participation and education:

$$\omega_L \equiv \left(t_L + b + \frac{\Delta_L(n_L)}{\lambda}\right), \tag{2.48}$$

$$\omega_H \equiv \left(t_H - t_L + \frac{\Delta_H(n_H) - \Delta_L(n_H)}{\lambda}\right). \tag{2.49}$$

Thus, the wedge on participation, ω_L, includes the tax benefits of an increase in participation, $t_L + b$, and the monetized difference between the utility costs and well-being costs of participation, $\Delta_L(n_L)/\lambda$. To understand why the latter term enters the wedge on participation, note from equation (2.42) that the utility costs of participation are set to equal the gains of increased consumption. If, for individual n_L, the well-being costs are smaller than the utility costs of participation, such that $\Delta_L(n_L) > 0$, the well-being gains outstrip the well-being losses of participation. Similarly, the wedge on education, ω_H, includes the tax benefits of increased education, $t_H - t_L$, and the monetized difference bewteen the utility costs and the well-being costs of education, $\left(\Delta_H(n_H) - \Delta_L(n_H)\right)/\lambda$.

Maximizing social welfare, (2.46), subject to the budget constraint, (2.47), with respect to b and t_H, and substituting in wedges, (2.48) and (2.49), and elasticities, (2.44), yields the following formulae for the optimal wedges:[11]

$$\omega_L = \left(\frac{u'(b)}{\lambda} - 1\right)/\eta_{Lb}, \tag{2.50}$$

$$\omega_H = \left(1 - \frac{u'(w_H - t_H)}{\lambda}\right)/\left(-\eta_{Ht_H}\right). \tag{2.51}$$

The optimal wedge in condition (2.50) consists of two terms. The first term in brackets measures the social welfare gain of distributing an additional unit of income to the unemployed, minus the resource costs of doing so. The second term gives the semi-elasticity of the number of unemployed with respect to the unemployment benefit. Thus, the optimal wedge is set according to a familiar logic (e.g., Gerritsen and Jacobs, 2013*a*, Chapter 4 of this Dissertation), and equals the redistributional gain of a participation tax, divided by the magnitude of the behavioral response. The same logic applies to the optimal wedge on eduation, (2.51). It equals the social benefits of redistributing away from the high-skilled (the bracketed term), divided by the negative semi-elasticity of the number of high-skilled with respect to an increase of the high-skilled tax.

As was the case in previous sections, however, participation and education wedges consist not only of the net participation tax and the net tax on high-skilled workers. They also take into account the difference between the well-being based and the utility-based assessment of participation and education. If net well-being of participation exceeds net utility of participation, too few people are participating in the labor market, providing a motive for lower participation taxes. If net well-being of education exceeds net utility of education, too few people become high-skilled, providing a motive for lower taxes on the high-skilled (or higher education subsidies).

[11] Naturally, to fully solve for all three seperate tax instruments (b, t_L, and t_H), I would need the first-order condition for t_L as well.

2.5 How far away from their bliss point are people?

Thus far I have established that, if individuals do not behave in a way that maximizes their well-being, optimal tax rules are adjusted in a straightforward manner. The formulae for the optimal wedges are largely in line with the standard formulations of optimal tax schedules. The crucial difference is that the wedge consists of the marginal increase in well-being, as well as the marginal increase in tax revenue, associated with an increase in labor effort (or consumption, participation, education). As shown, the total wedge consists of the sum of the tax wedge and the wedge between the well-being-based MRS and the utility-based MRS. While similar results have been obtained in a comparable context (e.g., Kanbur, Pirttilä and Tuomala, 2006; Blomquist and Micheletto, 2006), none of these studies came to a quantification of the latter wedge. On the one hand, it is straightforward to obtain empirical measures of the utility-based MRS as it *by definition* equals net relative marginal prices. On the other hand, to obtain a measure of the well-being-based MRS one first needs to determine what constitutes well-being, an issue likely to be contentious. In this section I use survey data on subjects' life satisfaction as approximation of their well-being. Using these data, I estimate a well-being function to determine the well-being-based MRS of leisure for consumption and compare it to the utility-based MRS. This provides an indication of how far away individuals are from their bliss point, i.e., whether individuals work too much or too little.

The focus of the empirical analysis is on the case of direct nonlinear taxation, discussed in Section 2.1. I am therefore interested in the wedge of equation (2.27).[12] More specifically, I focus on two related empirical results. First I determine the extent to which individuals tend to work too much, which I denote as Δ_n:

$$\Delta_n \equiv \frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1. \tag{2.52}$$

Notice that Δ_n is part of the second term of the wedge in equation (2.27). It has a straightforward interpretation: recall that $-u_{l,n}/u_{c,n}$ equals the net marginal wage rate of individual n, whereas $-g_{l,n}/g_{c,n}$ is the wage rate that individual n *should* have earned to justify the actual amount of hours that he or she is working. Hence, Δ_n gives the marginal wage rate an individual should have earned given his

[12]In principle, the same type of analysis could be carried out for the wedges in the case of direct linear taxation, (2.12), indirect taxation, (2.36), participation taxation, (2.48), or education subsidies, (2.49).

labor supply decision, relative to his actual marginal wage rate. If, say, $\Delta_n = \frac{1}{2}$, individual n works too much: he works as if he is earning fifty percent more than what he actually earns. On the other hand, if $\Delta_n = -\frac{1}{2}$, individual n works too little: he works as if his wage was only half his actual wage. Thus, Δ_n is a natural measure of the extent to which individuals work too much from a well-being point of view.

Second, I am interested in the total wedge on labor income, ω_n, the definition of which I repeat for convenience:

$$\omega_n \equiv \frac{T'(nl_n)}{1-T'(nl_n)} - \frac{g_{c,n}}{\lambda}\left(\frac{-g_{l,n}/g_{c,n}}{-u_{l,n}/u_{c,n}} - 1\right). \tag{2.27}$$

I am especially interested in the extent to which the total wedge deviates from the tax wedge on labor income, $T'/(1-T')$. This provides an indication of the extent to which wedges on labor income are habitually over- or underestimated by assuming that utility coincides with well-being. On the basis of preceding sections, I concluded that the full wedge, ω_n, rather than only the tax wedge, should be taken into account when judging the optimality of a tax system. The analysis below hence provides a qualification to studies that draw conclusions on the optimality of current tax systems on the basis of the tax wedge alone.

2.5.1 Data and strategy

Dataset and sample selection

The empirical analysis uses data from the British Household Panel Survey (BHPS), which includes information on various variables for a representative sample of individuals over consecutive years. Data are available for every year between 1991 and 2008, but a question on subjects' well-being has been available since 1996, with the exception of 2001. Thus, I am able to use data for the years 1996 to 2000 and 2002 to 2008, making for a raw sample of 27,699 unique individuals over a period of up to 12 years, with on average 6.2 years of data per individual.

In order to obtain a relatively homogeneous group of people without losing too many observations, and to limit the likelihood of omitted variable biases in my empirical analysis, I impose further restrictions on this sample. For homogeneity, I restrict the sample to heads of household, who are employed, without children, and of prime working-age between 25 and 59 years old.[13] In addition, I only

[13]For the definition of the head of household, the BHPS follows the General Household Survey,

include people if they have the same job function as in the preceding year, i.e., I exclude people whose function has changed, whether this was due to promotion or demotion, due to a change of company, or because of new entrance into the labor market. I do this because job changes are likely to have a direct impact on life satisfaction, while at the same time affecting the number of hours worked. As a result, without controlling for job changes, the effect of those changes on life satisfaction would be absorbed by the coefficient of the number of working hours.[14] Indeed, failure to control for changes of job function might well be an important reason why some previous studies on life satisfaction did not find a significant effect of hours worked (e.g., Booth and Van Ours, 2008). The remaining sample contains 4,194 unique individuals, with on average 3.2 observations per individual.

Measuring well-being

An individual's well-being is measured by the response when asked about satisfaction with his or her life. The specific question asked is:

> How dissatisfied or satisfied are you with your life overall?

Possible answers range from 1 to 7, with 1 labeled "Not satisfied at all," and 7 labeled "Completely satisfied." I assume that the answer to this question accurately reflects the well-being of the person answering the questionnaire, and is thus taken to be the empirical measure of g_n. Equivalently, the goal of optimal government policy is assumed to be the maximization of this specific measurement of well-being. This assumption can and should of course be subjected to criticism: the ideal measure of well-being is probably some combination of a choice-based measure (i.e., utility) and various non-choice-based measures such as life satisfaction (as argued by, e.g., Kőszegi and Rabin, 2008). However, since there are numerous studies on optimal taxation with utility as the sole measure of well-being, a sensible first step towards Kőszegi and Rabin's ideal is to compare the results of these studies to the case in which life satisfaction is taken as the sole measure of well-being.

i.e., the principal owner or renter of property, and (where there is more than one), the male taking precedence, and (where there is more than one potential head of household of the same sex), the eldest taking precedence.

[14] Instead of excluding job changers from the sample, I also directly controlled for changes in job function. This does not change results much. I prefer excluding these observations from the analysis entirely, because I cannot observe the reason for the job change, e.g., whether it was due to a promotion or a demotion, which is potentially important.

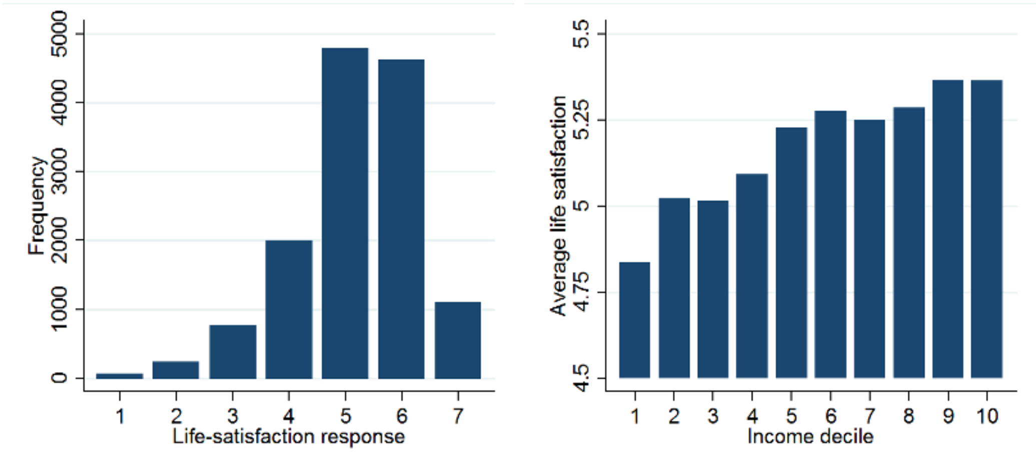

Figure 2.1: Histogram of life satisfaction scores and average life satisfaction by income decile

Figure 2.1 contains two panels that describe the data contained in the life satisfaction variable. The first panel illustrates the frequency at which a certain life-satisfaction score is given as answer. The second panel illustrates the average life-satisfaction score for each decile of net household income. Even without controlling for any other variables and without using any fixed effects, there appears to be a clear concave relationship between well-being and net household income.

Explanatory variables

Since ultimately I want to obtain a measure of the well-being based MRS of leisure for consumption, the most important explanatory variables are measures of consumption and work effort. As an approximation of consumption I choose to follow Layard, Mayraz and Nickell (2008) by using total real net household income. Naturally, one would like to use permanent income when explaining overall life satisfaction. However, in the absence of data on permanent income I need to rely on current income. Some, but most likely not all, of the bias that originates from my reliance on current income is eliminated by the sample restrictions on age. Income is measured at constant household costs, and includes income from labor, investments, benefits, pension, and transfers, net of taxes. I choose not to normalize the income variable by using equivalence scales to correct for the size of the household. The reason I do not do this is because the choice of the particular equivalence scale is always a controversial issue, and because for my main specification, in which income enters the well-being equation logarithmically, the equivalence scale is in any case absorbed by the marital-status dummies. The first panel in Figure 2.2

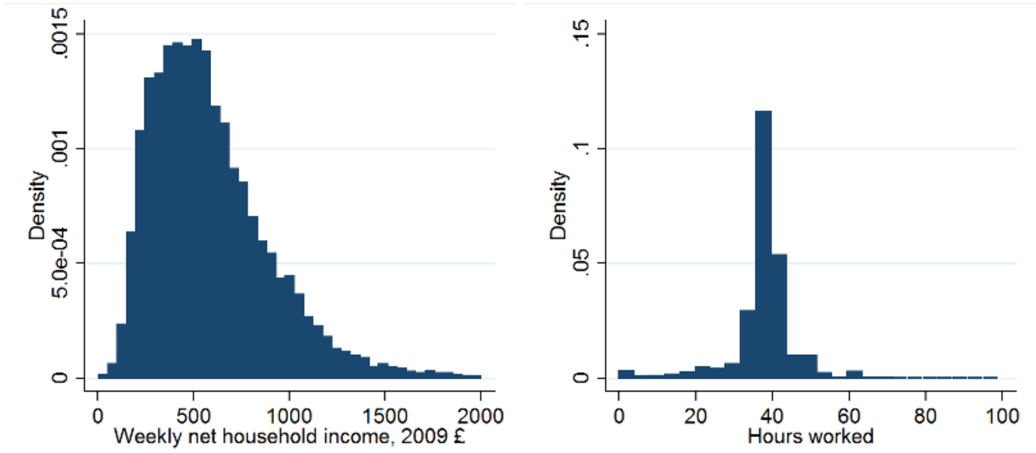

Figure 2.2: Densities of real net household income and hours worked

illustrates the density of net weekly household income.

The second crucial explanatory variable measures the number of hours a person works in a normal week. The second panel in Figure 2.2 illustrates the density of these weekly hours worked. Not surprisingly, the amounts of working hours are mostly concentrated around forty hours of work per week. Other explanatory variables I use include age dummies, subjective health evaluation dummies (answers ranging from 1, "very poor", to 5, "excellent"), year dummies, and marital status dummies. On top of that I include person-fixed effects to capture the influence of all person-specific time-invariant variables.

Empirical strategy

My main results follow from estimating the following linear equation:

(i) $$g_{it} = a_0 \ln c_{it} + a_1 l_{it} + a_2 l_{it}^2 + \sum_j b_j x_{jit} + c_t + c_i + v_{it},$$

where subscripts i and t denote individual i at time t, g_{it} denotes life satisfaction, c_{it} real net household income, and l_{it} hours worked. Furthermore, x_{jit} are control variables, c_t are time-fixed effects, c_i individual-fixed effects, and v_{it} is the error term. Note that I assume that the functional form of well-being is, apart from a constant, identical across persons and additive in nature. Moreover, it is assumed to be logarithmic in income and quadratic in hours worked. The latter assumptions on functional form are somewhat relaxed when I turn to semi-parametric analysis in the next section which is devoted to testing the robustness of my results. Crucially, I assume that observed changes in income and working hours are exogenous to life satisfaction. Naturally, this latter assumption is a source of concern in the absence of a proper quasi-experimental design. As discussed above, I tried to address these concerns by restricting my sample to exclude the most obvious cases of endogeneity.

The well-being based MRS of leisure for consumption is given by:

$$\frac{-g_{l,it}}{g_{c,it}} = \left(\frac{-a_1}{a_0} + 2\frac{-a_2}{a_0} l_{it}\right) c_{it}. \tag{2.53}$$

Hence, the estimation of equation (i) provides the first ingredient of the extent to which people work too much, Δ_{it}, and thus of the wedge on labor effort, ω_{it}. The second ingredient is given by the utility-based MRS. The first-order-condition of individual utility maximization, equation (2.18), indicates that the utility-based MRS is given by the person's net marginal wage rate:

$$\frac{-u_{l,it}}{u_{c,it}} = \left(1 - T'\left(n_{it} l_{it}\right)\right) n_{it}. \tag{2.54}$$

The wage rate n_{it} is calculated by dividing the *individual*'s gross labor income by the number of hours worked. The marginal tax rate is obtained by the following procedure. First, total taxes are determined by taking the difference between households' gross and net labor income, including income-dependent transfers and subsidies. Next, the resulting variable is smoothed over gross income and a numerical derivative is taken. This numerical derivative is taken to be the marginal tax rate. It is thus implicitly assumed that effective labor taxes are a function of household labor income. While this assumption is less accurate for moderate-to-high income workers as the income tax system in the United Kingdom is individual based, it is more accurate for low-income workers as eligibility for transfers and benefits generally depend on household income (see, e.g., Brewer, Saez and Shephard (2010)).[15] Figure 2.3 depicts the total tax schedule (first panel) and marginal tax schedule (second panel) found in this way. Due to the phasing out of transfers and benefits, marginal taxes are relatively high for low-income levels.

This provides all the ingredients needed to calculate Δ_{it}. In order to obtain a measure of the total wedge, ω_{it}, Δ_{it} needs to be multiplied by $g_{c,it}/\lambda$. For this, I assume there are no income effects, such that λ equals the simple average of the

[15] I also performed the exact same analysis while focussing on individual taxation. As expected, this resulted in much lower marginal tax rates at the bottom, since the phasing out of household-income dependent transfers and benefits are not taken into account. However, the general results of this section remain entirely intact.

Figure 2.3: Empirical total tax and marginal tax schedules

marginal well-being of income: $\lambda = \sum_{i,t} g_{c,it}/N$ where N is the total number of individuals.

2.5.2 Evidence on suboptimal behavior and the wedge on labor

The results of estimating equation (i) are given in Table 2.1. The first column shows results for the entire sample, while the second and third columns show the results of separate regressions for male and female respondents. For all regressions the coefficient on income is significant and around 0.18, which is to say that a percentage increase of net household income is associated with an increase in life satisfaction of (approximately) a hundredth of 0.18 point. While this effect seems rather small, it is in fact comparable to earlier results, for example from Layard, Mayraz and Nickell (2008). As can be seen from the first column, the coefficient on the linear working hours term is positive while the coefficient on the quadratic term is negative. This is suggestive of an inverted-'U' shaped relationship between life satisfaction and hours worked. It is easily verified that the top of this parabola is around 30 hours of work, after which every additional working hour corresponds with decreased life satisfaction. These findings are confirmed when the sample is restricted to male respondents, but loses its statistical significance when the sample is restricted to female respondents. The insignificant result for female respondents might well be due to the small sample size, especially considering the fact that, in the remaining sample, the average number of sampled years per person is less than three.

Table 2.1: Estimation results for equation (i)

	All	Male	Female
log income	0.182***	0.146***	0.217***
	(0.0332)	(0.0392)	(0.0641)
hours	0.00762**	0.00953**	0.00137
	(0.00346)	(0.00384)	(0.00790)
hours squared	-0.000126***	-0.000158***	-7.72e-06
	(4.62e-05)	(5.05e-05)	(0.000111)
Observations	13,529	9,908	3,621
R-squared (within)	0.033	0.031	0.065
Number of individuals	4,194	2,942	1,252

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: life satisfaction. All regressions include age dummies, subjective health dummies (on a scale from 1 to 5), year dummies, and marital status dummies, as well as person-fixed effects.

By using equation (2.53) and the estimation results of Table 2.1 I can determine, for every individual in the sample, the well-being based MRS of leisure for consumption. Substituting this and the empirical observation of the utility-based MRS into equation (2.52), I obtain for every person and year in my sample a value for Δ_{it}. As discussed above, this value indicates the extent to which he or she works too much. If positive, the person works too much from a well-being point of view; if negative, the person works too little. Since (corrective) taxation is conditioned on labor income, it is most informative to show how Δ_{it} varies over gross household labor income. The smoothed values of Δ_{it} are depicted in the first panel of Figure 2.4. This graph illustrates that up to a weekly gross labor income of around £850, individuals tend to supply too little labor. Conversely, individuals that earn more than that tend to supply too much labor from a well-being perspective.

Now I can readily determine the total wedge on labor effort, ω_{it}, as given in equation (2.27), by substituting for the empirical marginal tax schedule, the measure of overwork Δ_{it}, and the marginal welfare weights g_{it}/λ. The smoothed values of the total wedge are illustrated by the blue line in the second panel of Figure 2.4. The red line shows the tax wedge, which is normally taken as the total wedge on labor effort by studies that do not distinguish between utility and well-being. Naturally, the actual wedge is larger than the tax wedge for people that work

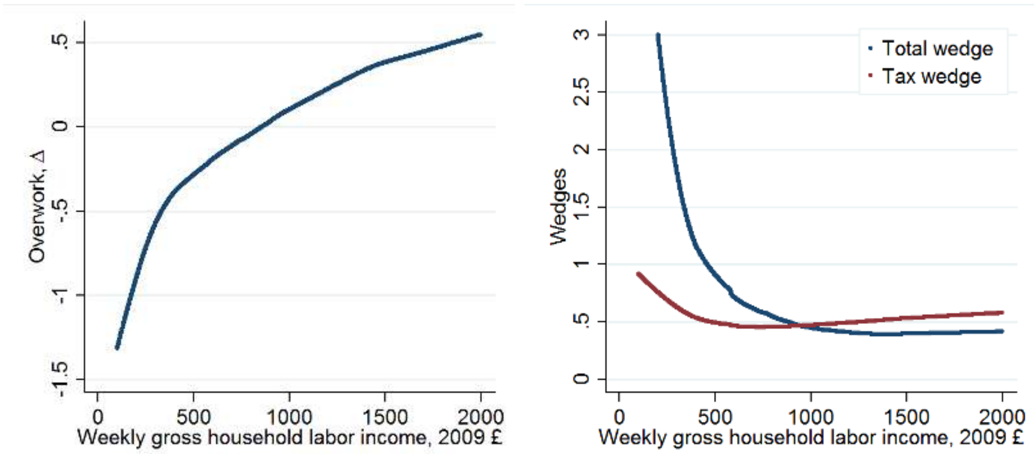

Figure 2.4: Overwork and wedges, full sample

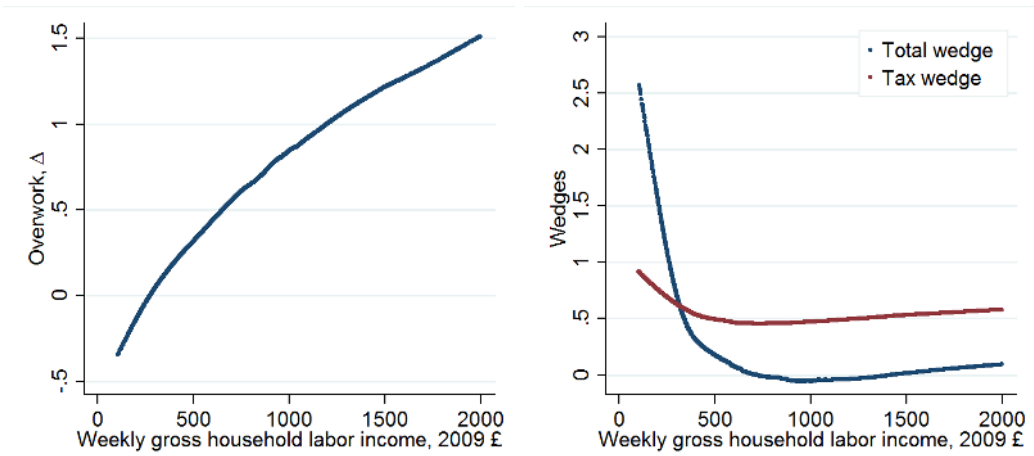

Figure 2.5: Overwork and wedges, male subsample

too little, and smaller than the tax wedge for people that work too much. As the marginal welfare weights, g_{it}/λ, are rapidly decreasing with income, the difference is more pronounced for low-income workers than for high-income workers. Low-income workers could increase their own well-being by simply working more, but for some reason do not act in the interest of their own well-being. As the total wedge measures the social welfare gain of increased labor effort, this wedge is larger than the tax wedge alone. For high-income workers the opposite holds: they work too much from a well-being perspective. While increased labor effort from high-income workers would raise tax revenue, it would also decrease their well-being. Therefore the total wedge for high-income workers is lower than the tax wedge alone would suggest.

Figure 2.5 illustrates the extent of overwork (first panel) and the total wedge on labor effort (second panel) if the analysis is restricted to male workers only. As is apparent, the previous results more or less carry over to a male-only sample. This time, however, only the very-low income workers work too little from a well-being perspective, whereas the rest works far too much. As a result, the total wedge on labor effort for median-to-high income workers is much lower than the tax wedge alone. In fact, the total wedge hovers around zero for a significant range of the income distribution. This implies that for these income groups the tax revenue gain associated with an increase in working hours would be completely offset by well-being losses.

2.5.3 Discussion of results

The results in Figure 2.4 convey a potentially important implication for applied studies of optimal taxation. If we do not distinguish between utility and well-being, which indeed we usually do not, we might misappreciate the actual wedge on labor effort by focussing solely on the marginal tax revenue gains of labor effort. The standard approach to applied optimal taxation is to determine the optimal wedge and compare this to the actually observed tax wedge. Policy recommendations are distilled from the difference between the optimal wedge and the actual tax wedge. However, as the analysis of Sections 1-3 shows, the optimal wedge should in fact be compared to the actual *total* wedge, not just the actual *tax* wedge. If the results in Figures 2.4 and 2.5 are correct, standard applied studies of optimal taxation underestimate the actual wedge for low-income workers and overestimate the actual wedge for high-income workers. As a result, their recommendations understate the required tax *decrease* for low-income workers, as well as the required tax *increase* for high-income workers.

As a concrete example, consider the tax reforms that have recently been suggested by Brewer, Saez and Shephard (2010) and Blundell and Shephard (2012) for the United Kingdom. Both studies call for a reduction of marginal tax rates for low-to-moderate earners. They conclude that marginal tax rates for low-income workers are currently so high that the distortions on intensive labor supply are too large to be justified by any redistributional gains. Taking into account the results from Figure 2.4, these recommendations hold *a fortiori*. After all, since low earners work too little from a well-being point of view, the total wedge on labor is even larger than what the tax wedge suggests. Consequently, marginal tax rates should be lowered even further than above studies suggest.

An important caveat pertains to the question why low-income workers make too little working hours. So far I assumed that it is a matter of supply, that low-income workers refuse to make more hours even though it would enhance their well-being. However, an alternative explanation is that low-income workers face demand restrictions due to above market-clearing wages, e.g., minimum wages, union wages, efficiency wages, or other forms of downward wage rigidity. In that case, decreasing marginal tax rates is not useful as the induced labor supply increase would not affect labor demand and therefore not translate in more actual hours worked. In fact, as I illustrate in Gerritsen (2013*a*, Chapter 5 of this Dissertation), demand restrictions provide a motive for higher, rather than lower, marginal tax rates. As is true for much of the literature on optimal taxation, the interpretation of the results of this chapter is only valid in the case of supply-determined labor markets.

2.6 Robustness

Potentially crucial to the above analysis is the specific functional form of the well-being function. In the previous section I simply assumed that well-being was additive in its arguments, logarithmic in income, and quadratic in working hours. In this section, I retain the assumption on additivity but attempt to determine the sensitivity of the results to the way in which income and working hours enter the well-being function. As theoretically very little can be said on the functional form of well-being, I apply semi-parametric regression techniques that allow for a large degree of flexibility with respect to the specific functional form.

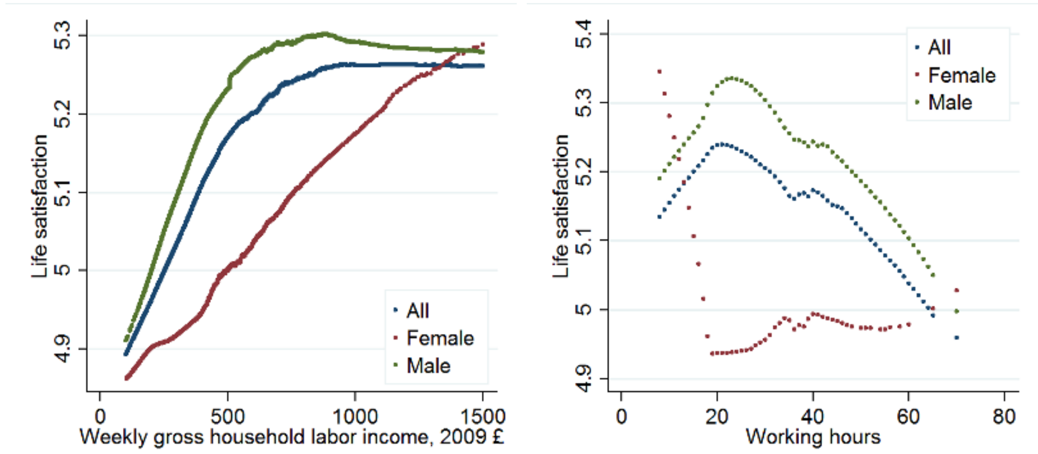

Figure 2.6: Non-parametric estimation results for income and working hours

2.6.1 Income

I first try to get a better understanding of the relationship between life satisfaction and income. For this, I estimate the following partially linear model:

$$g_{it} = \phi_c(c_{it}) + a_1 l_{it} + a_2 l_{it}^2 + \sum_j b_j x_{jit} + c_t + c_i + u_{it}, \tag{ii}$$

where all variables are the same as before, and $\phi_c(\cdot)$ is an unspecified function. The equation is estimated, using the algorithm developed by Lokshin (2006), who uses a locally weighted scatterplot smoothing (lowess) estimator to determine $\phi_c(\cdot)$. The resulting values of $\phi_c(c_{it})$, for different levels of c_{it}, are shown in the first panel of Figure 2.6. The blue line illustrates the estimated values for both male and female, and the green (red) line illustrates the estimated values if the sample is restricted to males (females) only.[16] These results are suggestive of a concave relationship between well-being and income, although seemingly linear for females.

However, even if the relationship between well-being and income is concave, it does not follow that a logarithmic specification is the correct one. For example, Layard, Mayraz and Nickell (2008) find in a similar setup that the relationship is slightly more concave than a logarithmic relation would imply. To determine whether the results of Section 4 are sensitive to the degree of concavity, I estimate

[16] Due to computational limitations because of the large number of individual dummies, I was forced to randomly discard a fifth of the sample when estimating equation (ii) for both male and female respondents. I therefore reiterated the same analysis many times, each time with highly similar results.

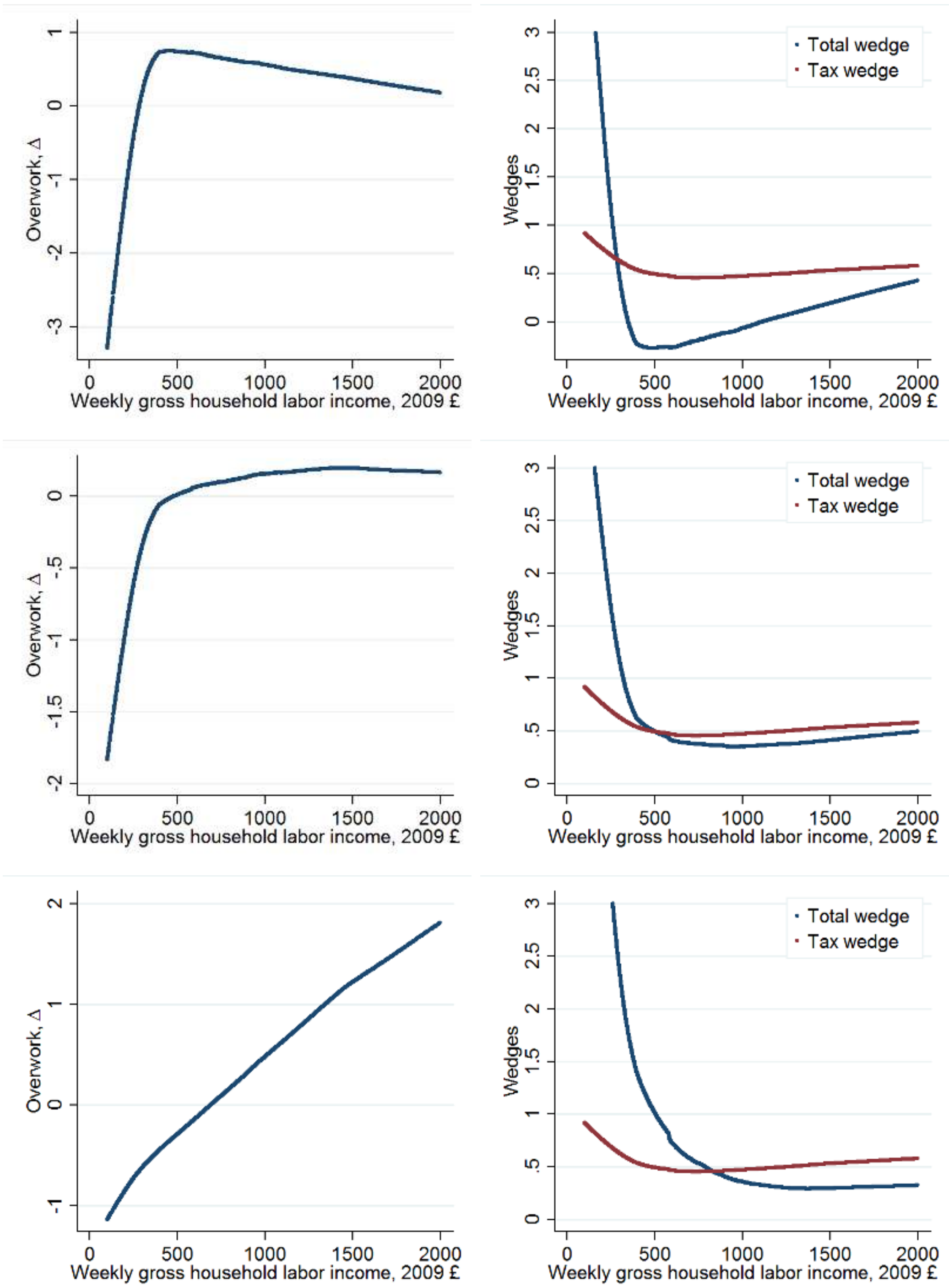

Figure 2.7: Overwork and wedges based on equation (iii), for $\rho = 0.1$ (upper panels), $\rho = 0.5$ (middle panels), and $\rho = 1.5$ (lower panels)

the following equation:

$$\text{(iii)} \qquad g_{it} = a_0 \left(\frac{c_{it}^{1-\rho} - 1}{1-\rho} \right) + a_1 l_{it} + a_2 l_{it}^2 + \sum_j b_j x_{jit} + c_t + c_i + u_{it},$$

for various values of ρ, allowing for varying degrees of concavity. On the basis of these estimations I derive the degree of overwork, Δ_{it}, and the total wedge, ω_{it}. Results for $\rho = \{0.1, 0.5, 1.5\}$ are depicted in Figure 2.7.[17] As can be seen, the conclusions on overwork and the total wedge remain in line with those of Section 4. Low-income workers work too little, while high-income workers work too much. Consequently, the total wedge on labor for low-income workers exceeds the tax wedge, whereas the total wedge for high-skilled workers is smaller than the tax wedge. Only in the case of a very low degree of concavity ($\rho = 0.1$), even the moderately poor seem to be working too much.

2.6.2 Working hours

Next, I further determine the results' sensitivity to the way working hours enter the well-being function. I estimate the following equation:

$$\text{(iv)} \qquad g_{it} = a_0 \ln c_{it} + \phi_h(l_{it}) + \sum_j b_j x_{jit} + c_t + c_i + u_{it},$$

with $\phi_h(\cdot)$ an unspecified function to be estimated non-parametrically. The resulting estimated values for $\phi_h(l_{it})$, separately for the full sample (blue), and the male (green) and female (red) subsamples, are shown in the second panel of Figure 2.6. For both the full sample and the male subsample, the relationship between life satisfaction and working hours appears to resemble an inverted 'U'. This corroborates the results of the parametric regressions in which the quadratic specification of working hours indicated a similar relationship. For female respondents no clear relationship is visible, which also corroborates earlier findings.

Even if the semi-parametric estimation indicates an inverted-'U' shaped pattern, this does not imply that the quadratic specification is correct. To test the robustness of my results, I therefore determine the marginal well-being of labor hours, $g_{l,it}$, by numerically taking the derivative of the estimated values of $\phi_h(l_{it})$. Together with the estimated value of $g_{c,it}$ from equation (iv), this allows me to

[17] To save on space, I only depict the results for these three values of ρ. The results remain broadly the same for any other positive value of ρ that I tried.

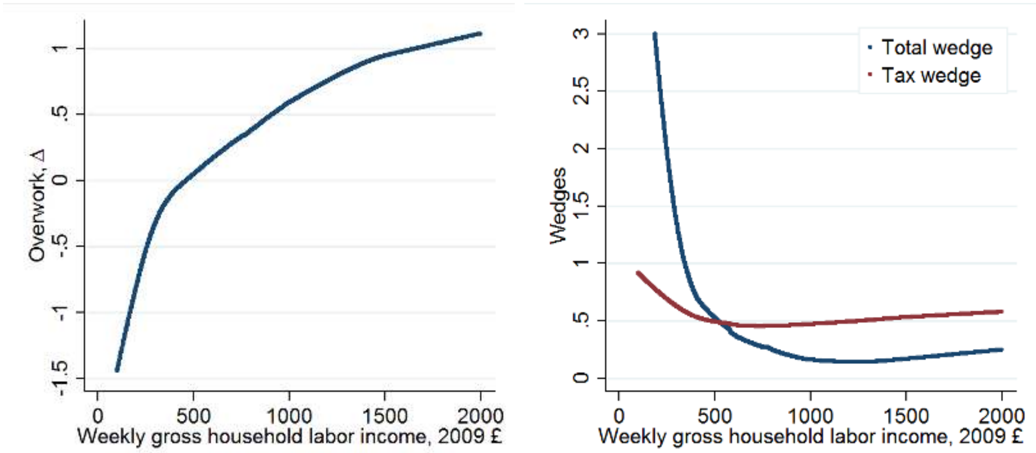

Figure 2.8: Overwork and wedges based on equation (iv)

determine values for Δ_{it} and ω_{it}. These values are illustrated in Figure 2.8. As before, low-income workers appear to be working too little, whereas high-skilled workers are working too much. Hence, the results of Section 4 appear insensitive to the way in which income and working hours enter the well-being function.

2.7 Concluding remarks

To the best of my knowledge, this is the first study that integrates the large empirical literature on the determinants of subjective well-being with the rigorous study of public finance. It is based on the notion that utility and well-being are not necessarily the same. Taking serious the potential divergence between utility and well-being, I find that the resulting optimal wedge on labor (or a specific good, labor participation, or education) is virtually identical to the one derived under conventional studies. However, the wedge itself now consists of the well-being consequences of drawing an individual farther from or closer to its well-being bliss point, as well as the standard tax wedge. Optimal marginal tax rates should be higher for workers that work too much from a well-being perspective, and lower for workers that work too little.

Using data on life satisfaction as a measure of a person's true well-being, I estimate the effect of income and working hours on well-being for a large panel of British individuals. On the basis of this estimation I conclude that low-income workers work less than optimal from a well-being point of view. Higher-income workers, on the other hand, work too much. Moreover, this finding is robust to varying assumptions on the functional form of well-being. Compared to standard derivations of optimal tax rates, and provided low-income labor does not face demand restrictions, these results thus endorse lower marginal tax rates at the lower end of the income distribution, and higher marginal tax rates at the higher end of the income distribution. Recommendations of recent studies, calling for a reduction of marginal tax rates for low earners in the United Kingdom, therefore hold *a fortiori*.

Perhaps more important, this study shows that it is possible to combine the rigor and emphasis on incentives that is so typical for the theory of optimal taxation, with an alternative measurement of well-being. My hope is that this might contribute to (i) more attention to economic incentives and optimal policy within applied studies of the determinants of subjective well-being, and (ii) a less dogmatic approach to well-being within public finance.

Chapter 3

Minimum wages and taxation in competitive labor markets with endogenous skill formation

3.1 Introduction

The desirability of a minimum wage has been fiercely debated by both policy makers and academics.[1] Proponents emphasize that a minimum wage leads to a higher income for low-income employees. Opponents mainly stress that it leads to higher unemployment rates as workers with productivity levels below the minimum wage find themselves unable to secure a job. As of yet, this debate has not been settled. Minimum wages were an important topic in the 2009 and 2013 federal elections in Germany, one of nine countries within the Organisation for Economic Co-operation and Development (OECD) without a statutory minimum wage (Immervoll, 2013), as well as in the American and French presidential elections of 2012. As noted by Cahuc and Laroque (2013), the OECD itself changed its appraisal of a minimum wage at least twice in the 1990s. The empirical literature on the effects of a minimum wage likewise seems to lack consensus. Some surveys report employment effects of a minimum wage to be absent or even positive (e.g., Card and Krueger, 1995), while in a more recent survey, Neumark and Wascher (2006) argue that the vast majority of the evidence points to a negative employment effect, albeit not always statistically significant.

The evidence on the effect of minimum wages on human capital investments,

[1] This chapter is based on joint work with Bas Jacobs, see Gerritsen and Jacobs (2013*b*).

whether it concerns adolescent education or on-the-job training, seems to be even more ambiguous. As we argue below, it is *a priori* unclear how minimum wages affect human capital investments. On the one hand, a higher minimum wage drives down the skill premium, thereby undermining incentives to develop skills. On the other hand, if mainly low-skilled wages are affected by the minimum wage, it will lead to higher unemployment among the low-skilled, thereby providing more incentives to develop skills so as to avoid involuntary unemployment.

The purpose of this chapter is to contribute to our understanding of the economic effects of minimum wages in two ways. From a positive perspective we determine how minimum wages affect the incentives to acquire skills, and identify conditions under which higher minimum wages lead to more skill formation. From a normative perspective we aim to contribute to the debate on the desirability of minimum wages by analyzing whether minimum wages are part of an optimal redistributive policy when skill formation is endogenous - and how this depends on the effect of minimum wages on skill formation. Importantly, we allow for income taxes as an alternative instrument to redistribute income.

We develop a general-equilibrium model with perfectly competitive labor markets. Firms demand both low-skilled and high-skilled labor. Individuals are heterogeneous with respect to their disutility of work. They optimally decide, first, to become either low-skilled or high-skilled, and, second, how many labor hours to supply. Individuals with little disutility of work have both an absolute and comparative advantage of working in high-skilled jobs, and thus end up becoming high-skilled, whereas high-disutility individuals become low-skilled. Minimum wages are binding for the low-skilled market segment, causing involuntary unemployment among the low-skilled only. As such, a minimum wage simultaneously discourages skill formation, by boosting low-skilled wages, and stimulates skill formation through higher unemployment.

We demonstrate that the net effect of a minimum wage on skill formation critically depends on the substitutability of high-skilled and low-skilled labor in the production function. Intuitively, if substitutability is high, a given increase in the minimum wage will cause firms to strongly substitute away from low-skilled labor, leading to a large increase in unemployment. If the substitutability is high enough, the increase in unemployment will outweigh the increase in the skill premium, and skill formation will rise. More specifically we show that in the absence of skill-dependent taxes and transfers, a minimum wage leads to more skill formation if the elasticity of substitution is larger than one, which seems to be the empirically plausible case.

The government maximizes a social welfare function featuring redistributive concerns. Due to informational constraints individualized lump-sum taxes are ruled out, such that the government needs to resort to distortionary income taxation and minimum wages to achieve its redistributive goals. The welfare effects of a minimum wage are studied in three different policy regimes, which are progressively more complex in the government's instrument set.

First, we determine the desirability of a minimum wage in the absence of income taxation. While this regime is obviously unrealistic, it helps in understanding the welfare consequences of a minimum wage in a relatively simple setup. In the absence of taxation, the social welfare gains of a minimum wage are a higher degree of income equality between low-skilled and high-skilled workers, and a higher degree of income equality among high-skilled workers as a higher minimum wage leads to lower high-skilled wages through general equilibrium effects. The social welfare losses of a minimum wage are given by higher inequality among low-skilled workers and the utility losses of laid off workers. Whether the gains outweigh the losses, and thus whether a minimum wage is optimally implemented, is ambiguous and crucially depends on initial inequalities, social redistributive preferences, and the minimum-wage elasticity of unemployment.

Second, we study the welfare effects of a minimum wage in a policy regime with skill-independent income taxation, i.e., taxation that cannot be conditioned on skill type. We believe this policy regime is of practical relevance as in reality redistributive taxes are hardly ever targeted on people's skill type. Even if some taxes and subsidies are targeted on skills, such as education subsidies, they are highly restricted in their capacity to target minimum-wage workers (i.e., in practice there are both educated minimum-wage earners and uneducated workers that earn more than the minimum wage). When the government sets taxes that are not conditioned on skill type, a minimum wage tends to lead to additional welfare losses as increased unemployment erodes the income tax base and therefore reduces tax revenue. However, taxes cannot be targeted well to deal with both inequality within skill groups and between high- and low-skilled workers. There might therefore be a role for the minimum wage in its capacity to redistribute income between skill groups if, for a given amount of redistribution, the welfare costs associated with a higher minimum wage (utility and tax revenue losses from higher unemployment) are sufficiently smaller than the welfare costs associated with higher income taxes (tax revenue losses from lower intensive labor supply). Minimum wages can in that case be seen to correct for the distributional imperfection of taxes that cannot be conditioned on skill type.

Third, we study the welfare effects of a binding minimum wage if the government can condition its tax instruments on skill type. In this policy regime, there is no distributional imperfection associated with taxes, and the government can use its tax instruments to achieve the exact same distributional effects of a minimum wage. That is, decreasing taxes on low-skilled income and increasing taxes on high-skilled income results in a higher net income for the low-skilled and lower net income for the high-skilled, just as a higher minimum wage would. This moreover leads to the same degree of distortion on the intensive labor supply margin. A minimum wage thus only differs from a distributionally equivalent tax-rate adjustment by causing higher unemployment and, as a direct result of this, more skill formation. Higher unemployment leads to utility losses and an erosion of the tax base. Higher skill formation, on the other hand, constitutes a welfare gain through higher tax revenues, provided that taxes are set progressively. A minimum wage is desirable if the benefits of higher skill formation outstrip the costs of higher unemployment. Thus, while in the policy regime with skill-independent tax instruments the role of a minimum wage is primarily to help redistribute income, in the regime with skill-dependent tax instruments its role is to reduce the inefficiency caused by taxation.

In most of this chapter, we assume that every low-skilled worker has an equal probability of becoming unemployed, i.e. that rationing occurs uniformly and on the extensive margin. The uncomfortable fact is that we do not really know in what way employment is decreased due to a minimum wage. We therefore also study a separate case in which unemployment is 'efficient,' implying that hours of work, rather than jobs, are rationed. In that case, there is no first-order utility loss associated with the unemployment caused by a marginally binding minimum wage. This ensures that a minimum wage is always optimal in the absence of taxation. However, in the presence of skill-independent taxes and transfers, the optimality of a minimum wage is still ambiguous. In the presence of skill-dependent taxes and transfers, a minimum wage is redundant as it can be perfectly mimicked by taxation.

Our work is most closely related to Lee and Saez (2012), Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation), and Gerritsen (2013*a*, Chapter 5 of this Dissertation). Lee and Saez (2012) also determine the desirability of a minimum wage in competitive labor markets, but focus on the extensive margin and assume efficient rationing, whereas we incorporate an intensive-margin labor-supply decision and focus mainly on uniform extensive-margin rationing. In Gerritsen and Jacobs (2013*a*), we determine how the desirability of a minimum wage depends on specific assumptions on the efficiency of rationing, and, contrary to this chapter, we focus purely on a policy regime with skill-dependent taxation and do not discuss optimal tax policy. Finally, Gerritsen (2013*a*) studies how optimal tax policy depends on the efficiency of rationing, but focusses on extensive-margin labor-supply decisions and does not discuss the minimum wage.

This chapter is structured as follows. Section 3.2 is devoted to a discussion of relevant literature. Section 3.3 introduces the theoretical model, the comparative statics of which are derived in Section 3.4. In Sections 3.5 and 3.6 we discuss the welfare effects of a minimum wage in the presence of skill-independent and skill-dependent tax instruments, respectively. In Section 3.7 we study the case of efficient rationing. We close the chapter with some concluding remarks.

3.2 Related literature

3.2.1 Theory

There are roughly two approaches to studying the implications of a minimum wage. One strand of the literature takes certain market imperfections in the labor market as given and determines how a minimum wage affects efficiency, employment, and/or social welfare. A popular assumption is that employers have a degree of monopsony power over wages, leading to inefficiently low wages. The classical argument is due to Robinson (1933). Indeed, it is straightforward to show that, in a partial-equilibrium setting with a monopsonist in the labor market, wages and employment are set inefficiently low. In that case, a binding minimum wage might be employed to ensure an efficient outcome. In a more realistic setting, however, this argument quickly becomes problematic. As Stigler (1946) argues, the optimum wage will vary with occupation, among firms and, often rapidly, through time. Therefore, "[a] uniform national minimum wage, infrequently changed, is wholly unsuited to these diversities of conditions".

More recent studies bring further nuance to the discussion. For example, Manning (2003), focussing on employment, considers a general-equilibrium model with heterogeneous firms and concludes that a minimum wage might have opposing employment effects for different firms, leaving the aggregate employment effect ambiguous. Bhaskar and To (1999) consider monopsonistic competition with exit and entry of firms, firm-specific job types and heterogeneous preferences for job types, and reach a similar conclusion. While a minimum wage increases employment per firm, it also forces some firms to exit the market, leaving aggregate employment and welfare outcomes ambiguous. Still, as Cahuc and Laroque (2013) show, with a sufficiently rich set of tax instruments the government can always reach the second-best competitive allocation without any need to resort to minimum wages.

The minimum wage is also studied in frameworks combining monopsony power with other market imperfections. For example, Rebitzer and Taylor (1995) study a model in which firms imperfectly monitor their employees and therefore set efficiency wages to motivate them not to shirk. If higher labor supply leads to costlier monitoring, they show that a minimum wage will increase employment over the short term, with ambiguous results over the long term. Cahuc, Saint-Martin and Zylberberg (2001) introduce a model where high- and low-skilled wages are bargained over between employers and unions that represent high-skilled workers. They show that a higher minimum wage might reduce the unions' bargaining power over the high-skilled wage, potentially leading to more employment for both low-skilled and high-skilled workers through general equilibrium effects. Flinn (2006) analyzes a matching model of the labor market and argue that if workers' bargaining power is too low for the Hosios (1990) efficiency condition to hold, a minimum wage might function as a crude measure to push labor market outcomes towards efficiency. In a similar vain, Hungerbühler and Lehmann (2009) find that a binding minimum wage might be part of an optimal redistributive policy as an indirect way to increase workers' bargaining power, even if the government uses non-linear income taxation to achieve its redistributive goals.

The second strand of the literature, which is closer in spirit to the present study, applies an optimal-taxation framework to competitive labor markets and heterogeneous workers with either continuous skill types as in Mirrlees (1971) or, more often, two skill types as in Stiglitz (1982). In the latter tradition, Allen (1987) and Guesnerie and Roberts (1987) show that a minimum wage might be optimal as part of a redistributive policy if the government is confined to linear taxation only. However, if non-linear taxation is available, a minimum wage is never optimal as it raises the attractiveness for high-skilled workers to imitate the low-skilled, thereby tightening the incentive-compatibility constraint. This approach has been criticized by Lee and Saez (2012) on informational grounds.

They argue that a government needs to be able to distinguish high-skilled from low-skilled workers in order to enforce a minimum wage, thereby making incentive-compatibility constraints irrelevant. The studies of Marceau and Boadway (1994) and Boadway and Cuff (2001) are liable to the same criticism. They combine a minimum wage with unemployment insurance and find that a minimum wage might still be optimal in combination with a non-linear tax schedule. Assuming that individuals can only apply for unemployment benefits if they are unable to find a job, Boadway and Cuff (2001) show that a minimum wage provides information about the bottom of the skill distribution, which can not be obtained by merely using taxes.

Almost any study takes skill levels of individuals as exogenously given. Two exceptions with endogenous skill formation on the extensive margin are Saint-Paul (1996) and Cahuc and Michel (1996). In Saint-Paul (1996), as in our model, an increase in low-skilled unemployment causes more individuals to become high-skilled. As he assumes perfect substitutability between high- and low-skilled labor, higher low-skilled unemployment might thereby lead to lower labor productivity and to even higher levels of unemployment in the case of real wage rigidity. The implementation of a binding minimum wage might thereby induce increasing returns to education and soaring low-skilled unemployment rates, reaching levels of up to a hundred percent. As we assume imperfect substitutability, such an extreme result is not attainable in our model. Cahuc and Michel (1996) develop an overlapping generations model with a high-skilled and a low-skilled production sector. Furthermore, high-skilled production exhibits positive externalities and, hence, serves as a catalyst of endogenous growth. They show that if a minimum wage increases human capital formation, this can lead to higher growth.[2] Our model exhibits similar extensive skill-formation as in Cahuc and Michel, although we analyze the effects of a minimum wage in an optimal-taxation setting without externalities.

Lee and Saez (2012) and Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation) are particularly closely related to the current study as they analyze the optimality of a minimum wage alongside optimal taxes and transfers in models with two skill types and competitive labor markets. Lee and Saez study the case in which rationing is efficient, such that new entrants are unable to find a job in a rationed low-skilled labor market. In that case, a binding minimum wage might be optimal to implement as it effectively mutes the distortionary effects of a transfer

[2]Naturally, as in the case of monopsonistic labor markets, this begs the question why the externalities are not internalized by appropriately setting taxes.

towards low-skilled workers. In Gerritsen and Jacobs (2013*a*), we derive a general optimality condition for a binding minimum wage that hold for any arbitrary rationing schedule, including but not restricted to efficient rationing. Calibration of this condition shows that a minimum wage decrease yields a Pareto-improvement in all countries under consideration, except possibly the United States. The current study distinguishes itself from Lee and Saez (2012) mainly by its focus on uniform rationing – i.e., a common probability of unemployment for every low-skilled worker – and from Gerritsen and Jacobs (2013*a*) by its focus on the skill formation and social welfare effects of a binding minimum wage under varying tax regimes.

3.2.2 Informational inconsistency

Following the seminal analysis of Mirrlees (1971), modern literature on public finance builds on the assumption that the fundamental source of heterogeneity across individuals is private information, unverifiable by the government. Typically, heterogeneity is assumed to originate from varying wage rates, or earning ability. The government can only observe total labor earnings, which is the product of the wage rate and the total number of hours worked. Since tax policy can only be conditioned on observables, and therefore not on earnings ability, a first-best outcome cannot be attained. However, to be able to implement a minimum wage, the government must observe individual wage rates. This leads to the problem of *informational inconsistency*: in the Mirrlees (1971) framework it contradicts the assumption that wage rates are private information and are thus not verifiable by the government. Indeed, information on individual wage rates theoretically enables the government to reach any desired redistribution without efficiency losses by implementing individualized lump-sum taxes and transfers. Consequently, studies that use the Mirrlees framework for the analysis of minimum wages usually make an – often implicit – *ad hoc* assumption that information required for the implementation of minimum wages cannot be used for taxes and transfers. Guesnerie and Roberts (1987, p.498), somewhat euphemistically, remark that "this is a somewhat mixed observability assumption."

The informational inconsistency appears when the source of heterogeneity can be defined in terms of observable variables. In the standard Mirrlees exercise, for example, exogenous earning ability can be defined as $n \equiv \frac{z}{l}$, z labor income, and l the number of hours worked. First best is not attainable because only z, and not l, is observable. If a minimum wage is introduced in this framework as, for example, in Boadway and Cuff (2001), first best is attainable since the wage rate must be observable and, obviously, the exogenous ability can be defined in terms of the wage rate $n \equiv w$.

The exact same inconsistency occurs when introducing minimum wage legislation in a model based on Stiglitz (1982), see for example Allen (1987), Marceau and Boadway (1994), Cahuc and Michel (1996), Aronsson and Koskela (2008), and Danziger and Danziger (2010). They all assume, contrary to Mirrlees, that workers with different wage rates are imperfect substitutes. Still, because in these models wage rates are generally exogenously given and the sole source of heterogeneity, we can again write $n \equiv w$, which implies first best is attainable once the government can observe wage rates. This result suggests that to be informationally consistent, we need to direct attention away from models in which the source of heterogeneity can be defined by the wage rate.

One way to do this is to introduce a labor-effort decision alongside an hours-of-work decision. Denoting labor effort as e, we can define the wage rate $w \equiv en$, or alternatively, earnings ability as $n \equiv \frac{w}{e}$. As long as the government cannot observe effort, exogenous ability cannot be defined by observables only, and first best is not attainable. This approach is taken by Deltas (2007).

An alternative approach is taken by Lee and Saez (2012). The model of Lee and Saez includes multiple job types and individuals supply one unit of labor if employed ($l = 1$). Thus, earnings are given by $z \equiv w$, which is verifiable by the government so that it can enforce a minimum wage and set income-tax policy. To avoid a first-best outcome, without violating informational consistency, individuals are assumed to be heterogeneous with respect to their costs of participation in a particular job, which are unrelated to earnings and thus unobserved. These costs of participation, θ, cannot simply be defined in terms of observables, making the first-best allocation infeasible. In the model below, we adopt, like Lee and Saez, disutility of work as the fundamental source of heterogeneity across individuals, safeguarding us from informational inconsistency.[3]

3.2.3 Empirics

There is a large empirical literature on the effects of minimum wages on employment, which has recently been surveyed by Neumark and Wascher (2006). Most

[3] Nevertheless, in Section 3.5 we do study the social welfare effects of minimum wages in the case that the government does not condition its tax instruments on skill type, and thus does not fully use the information at hand.

studies find that minimum wages reduce employment, although the estimates are not always significant. This is in line with our model of competitive labor markets, which predicts minimum wages to lead to a decline in labor demand. Still, there are some notable exceptions that find non-negative employment effects , which would be more difficult to reconcile with our model (e.g., Card and Krueger, 1995). Since the seminal contribution by DiNardo, Fortin and Lemieux (1996), a smaller literature emerged on the effects of minimum-wage legislation on the wage distribution (e.g, Lee, 1999; Teulings, 2000, 2003; Autor, Manning and Smith, 2010). This literature often emphasizes general-equilibrium effects on wages rates that exceed the minimum wage. Such general-equilibrium effects are taken into account in our model, which predicts higher minimum wages to lead to a compressed wage distribution by simultaneously raising low-skilled wages and reducing high-skilled wages.

When it comes to the effect of minimum wages on skill formation, empirical results are much scarcer and more ambiguous. A number of potential effects of higher minimum wages on skill formation are recognized. When minimum wages lead to a compression of wages, the net return of human capital investments will drop, leading to less skill formation. However, if employment opportunities decline for low-skilled jobs, skill formation might be boosted in order to avoid unemployment. These arguments hold for investments in education and, perhaps to a lesser degree, for employee-financed on-the-job training. However, for on-the-job training additional arguments play a role. On the one hand, employees might finance their training by accepting a lower hourly wage rate, the possibility of which is diminished by a higher minimum wage (e.g., Rosen, 1972). On the other hand, if training is firm-sponsored and labor markets are not perfectly competitive, a minimum wage might decrease the rents on low-skilled labor, leading to more investment in on-the-job training such that firms can obtain higher rents (e.g., Acemoglu and Pischke, 1999).

The studies that try to capture the effect of minimum wages on skill formation can be divided in those that explain school enrollment and related variables and those that explain on-the-job training. The evidence on either of the two human-capital variables is scarce and ambiguous. Moreover, most studies are unsuited to evaluate the distinct effects of minimum wages on skill formation – i.e., through a compressed wage structure and through higher unemployment – and analyze which effect dominates. Empirical studies of school enrollment are often flawed in this respect because they usually control for unemployment, such that estimates only show the direct minimum wage effect through the wage structure.[4] Studies of on-the-job training are often confounded, because minimum wages can impact the training decision in many different ways as it is usually a joint decision of employer and employee, each with their own incentives. For example, even if a minimum wage would lead to higher demand for on-the-job-training from the perspective of employees, an employer might be less interested in training its workers as the returns to training cannot be deducted from the wage rate of a minimum-wage worker. Data seem to be too weak to adequately take account of the different incentives.[5] Hence, amongst empirical ambiguity, we hope to contribute to our understanding of minimum wage legislation by theoretically identifying under what circumstances a minimum wage leads to more or less skill formation.

3.3 Model

We assume a unit mass of individuals and two job-types: high-skilled jobs and low-skilled jobs. We assume that wages on the high-skilled labor market are perfectly flexible to assure there is no unemployment among high-skilled workers. The government might, however, impose a binding minimum wage on the market for low-skilled labor. Unemployment will therefore be concentrated on the group of low-skilled workers. Thus, workers can either be unemployed low-skilled, employed low-skilled, or employed high-skilled workers. The fractions of each are denoted

[4]Studies that find that higher minimum wages lead to less schooling if controlling for the unemployment rate, include Cunningham (1981); Neumark and Wascher (1995*a*,*b*, 2003); Landon (1997); Chaplin, Turner and Pape (2003); Montmarquette, Viennot-Briot and Dagenais (2007); Pacheco and Cruickshank (2007). Interestingly, Cunningham (1981) only controls for the white unemployment rate and finds the schooling effect of minimum wages reversed for black youths. Similarly, Pacheco and Cruickshank (2007) find that the negative effect of higher minimum wages on enrollment rates is significant at a level of 1 percent when controlling for the unemployment rate, but only significant at a level of 10 percent if not. A number of studies do not find a significant effect of minimum wages on education, even when controlling for unemployment, see Ragan (1977); Ehrenberg and Marcus (1982); Card (1992); Crofton, Anderson and Rawe (2009). Only Mattila (1981) finds a positive effect of minimum wages on education, although this might be caused by the fact that she controls for the unemployment rate among people aged 35-44, which might be fairly irrelevant for students deciding whether to enroll for school. These findings suggest the importance of distinguishing the distinct effects of a minimum wage on skill formation. We express our hopes that future empirical research will give due attention to minimum wage effects stemming from a compressed wage structure and minimum wage effects stemming from higher unemployment.

[5]This is apparent in the contradictory findings. Negative effects of minimum wages on training are found by Hashimoto (1982); Schiller (1994); Neumark and Wascher (2001). Positive effects are found by Arulampalam, Booth and Bryan (2004); Dustmann and Schönberg (2009). Insignificant, or non-robust findings are presented by Mincer and Leighton (1981); Grossberg and Sicilian (1999); Acemoglu and Pischke (1999); Fairris and Pedace (2004).

by N^U, N^L, and N^H, respectively ($N^U + N^L + N^H = 1$). For short, we will denote the unemployed low-skilled as the unemployed. Similarly, the employed low-skilled workers are referred to as low-skilled workers. Type-specific variables are indexed with superscripts U, L, and H.

We assume that workers are heterogeneous with respect to their ability, θ. Rather than making the assumption, common in the optimal tax literature, that θ reflects the productivity per hour worked, we instead assume that θ reflects the utility cost per hour worked.[6] A higher θ implies that utility costs per hour worked are lower. This assumption is similar to Lee and Saez (2012), who also assume that more able workers have lower costs of participation, rather than higher labor productivity. Moreover, we assume that individuals with a higher ability enjoy a comparative advantage of performing high-skilled work, whereas individuals with a low level of ability enjoy a comparative advantage for low-skilled work. θ has support $[\underline{\theta}, \overline{\theta}]$ and follows a cumulative distribution function $G(\theta)$ with corresponding density function $g(\theta)$. We assume that $\underline{\theta} > 0$, which in our model implies that in the absence of unemployment insurance, individuals prefer being employed over being unemployed.

Individuals decide on the number or working hours and on whether to invest in human capital. The number of working hours is chosen to maximize utility, which is increasing in income and decreasing in the number of hours worked. Since more able high-skilled individuals have a lower cost of work, they supply more labor for a given wage rate. Hours worked are denoted by l^L for low-skilled workers and by l^H for high-skilled workers. Human-capital investment is made on the extensive margin, i.e., it is a discrete decision to become a skilled worker. The skilled wage is denoted by w^H, and the unskilled wage rate is denoted by w^L. If the individual invests in human capital, he earns $w^H l^H$, if not, he earns $w^L l^L$.

3.3.1 Individual optimization

Utility is denoted by V and is assumed to be separable and quasi-linear in consumption and working hours. Moreover, it exhibits a constant labor supply elasticity on the intensive margin, ε, which is assumed to be equal for both low and high-skilled

[6] Due to this assumption we manage to avoid the informational inconsistency to which we alluded in Section 3.2. It moreover ensures that hourly labor earnings are constant within groups and that a non-negligible share of the population earns the minimum wage, which conforms with reality. This cannot be the case if θ would reflect labor productivity. That is, had θ reflected exogenously given marginal productivity, any person with θ above the minimum wage would be hired and the mass of workers earning the minimum wage would be zero.

workers.

Initially, we assume that tax instruments are not differentiated according to skill type. Moreover, we restrict attention to linear instruments throughout this chapter. Hence, labor income is taxed at a common rate, t. In addition, tax revenue is rebated as a non-individualized lump-sum transfer, T. Later we explore the robustness of our results by allowing for skill-dependent instruments. Thus, with skill-independent tax policy, utility when unemployed, low skilled, and high skilled are given by:

$$V^U \equiv T, \tag{3.1}$$

$$V^L_\theta \equiv T + (1-t)w^L l^L - \frac{1}{\theta^\beta}\frac{(l^L)^{1+1/\varepsilon}}{1+1/\varepsilon}, \tag{3.2}$$

$$V^H_\theta \equiv T + (1-t)w^H l^H - \frac{1}{\theta}\frac{(l^H)^{1+1/\varepsilon}}{1+1/\varepsilon}. \tag{3.3}$$

Variables that depend on ability are denoted by a subscript θ. Note that there is no disutility of labor for unemployed workers, since they do not work. For employed workers, the marginal costs of labor supply are inversely related to ability, θ. As we assume that $\beta \in (0,1)$, individuals with a higher ability have a comparative advantage in doing high-skilled work. The higher is ability, the lower are the costs of labor effort in high-skilled jobs relative to the costs of labor effort in low-skilled jobs. This comparative advantage is stronger if β is lower. Since marginal utility of consumption is constant, households are risk-neutral with respect to the probability of becoming unemployed.[7] Each worker first decides to invest in human capital or not, and then, given the skill level, they optimally supply labor. We solve this optimization problem backwards.

Optimal labor supply for high- and low-skilled employed workers is given by:

$$l^H_\theta = (\theta(1-t)w^H)^\varepsilon, \tag{3.4}$$

$$l^L_\theta = (\theta^\beta(1-t)w^L)^\varepsilon. \tag{3.5}$$

Labor supply is an increasing function of the gross wage rate, decreasing in the tax rate and increasing with ability, θ. There are no income effects in labor supply,

[7]Allowing for risk-aversion would strengthen our result that a minimum wage leads to higher human capital accumulation (if the substitution elasticity is larger than one), see below. In that case, unemployment does not only raise skill formation by lowering expected utility of being low skilled, but also by increasing the variance in low skilled earnings. However, if unemployment is not concentrated on specific individuals, but instead spread uniformly across low-skilled workers, risk aversion would disappear, since the variance in low-skilled earnings would be nil.

which facilitates the analysis considerably. Substituting these expressions into the utility functions for low-skilled and high-skilled workers yields the following indirect utility functions:

$$V^U = T, \tag{3.6}$$

$$V_\theta^L = T + \frac{\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon}}{1+\varepsilon}, \tag{3.7}$$

$$V_\theta^H = T + \frac{\theta^{\varepsilon}((1-t)w^H)^{1+\varepsilon}}{1+\varepsilon}. \tag{3.8}$$

Individuals decide to invest in human capital if and only if their ability, θ, is such that their utility from being high-skilled is larger than or equal to the expected utility of being low-skilled.[8] We assume rationing is uniform so that the probability of becoming unemployed is equal for every low-skilled individual and does not depend on θ.[9] Hence, individuals decide to become high-skilled if $V_\theta^H \geq uV^U + (1-u)V_\theta^L$, where $u \equiv N^U/(N^L + N^U)$ is the unemployment rate amongst the low-skilled, defined as the share of the low-skilled population that is unemployed. Thus, we obtain a cut-off ability, Θ, for the individual who is indifferent between becoming skilled or staying unskilled:

$$\Theta = (1-u)^{\frac{1}{\varepsilon(1-\beta)}} \left(\frac{w^H}{w^L}\right)^{-\frac{1+\varepsilon}{\varepsilon(1-\beta)}}. \tag{3.9}$$

The cut-off ability decreases with both the skill premium, $\frac{w^H}{w^L}$, and the unemployment rate, u. A larger skill premium increases the benefits of being high-skilled, thereby leading to a decrease of the cut-off level of ability, Θ. Similarly, a higher unemployment rate increases the relative benefits of being skilled, since high-skilled workers are not affected by unemployment. Thus, a larger unemployment rate decreases the cut-off level of ability. The minimum wage therefore has an ambiguous

[8] Alternatively, we could speak of self-selection or sorting into skill levels. Our model is thus equivalent to an occupational-choice model with a high-skilled (high-wage) occupation and a low-skilled (low-wage) occupation.

[9] Our assumption of uniform rationing along the extensive margin, i.e., by laying off workers, is not innocuous. Rationing could as well be dependent on the ability level, θ, or could occur along the intensive margin by restricting hours. In Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation), we analyze the welfare consequences of a binding minimum wage in a more general setup in which individual unemployment rates may or may not depend on θ. Had rationing occurred along the intensive margin, it would be more efficient than rationing along the extensive margin. In an unrationed situation the marginal gain of an extra hour of work equals the marginal cost, such that a marginal level of intensive rationing does not have any welfare cost. We discuss the implications of such efficient rationing in a later section.

effect on skill formation. On the one hand, it lowers the skill premium. On the other hand, it raises unemployment among low-skilled workers. Note that the tax instruments do not affect skill formation. The reason for this is that skill-independent taxes symmetrically affect both the costs of skill formation (foregone low-skilled earnings) and the benefits of skill formation (high-skilled earnings). Individuals respond more elastically to wage differentials and unemployment rates if the elasticity of labor supply, ε, decreases, or if comparative advantage is weaker, i.e. β higher. Intuitively, a low ε and a high β make individuals more similar across skill types. As individuals are more similar, small changes in relative earnings translate into large changes in Θ.

For later reference, we note that

$$V_\Theta^H = V_\Theta^L - u\left(V_\Theta^L - V^U\right). \tag{3.10}$$

Hence, if unemployment is strictly positive, utility for the marginal high-skilled worker is below the utility of the marginal employed low-skilled worker: $V_\Theta^H < V_\Theta^L$. The reason is that the marginal high-skilled worker avoids low-skilled unemployment, and is willing to accept lower earnings in return.

By denoting total high-skilled labor supply and total low-skilled labor supply by L_S^H and L_S^L, the cut-off level, Θ, implies the following values for aggregate labor supply and group-sizes for high-skilled and low-skilled workers:

$$L_S^H \equiv \int_\Theta^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta), \tag{3.11}$$

$$L_S^L \equiv \int_{\underline{\theta}}^{\Theta} l_\theta^L \mathrm{d}G(\theta), \tag{3.12}$$

$$N^H \equiv 1 - G(\Theta), \tag{3.13}$$

$$N^L \equiv G(\Theta) - N^U. \tag{3.14}$$

Note that L_S^L is the *notional* aggregate low-skilled labor supply. In the presence of unemployment, not all low-skilled workers notionally supplying labor find employment.

3.3.2 Firms

There is a representative, competitive, profit-maximizing firm which produces output, Y, by employing aggregate high-skilled labor, L^H, and low-skilled labor, L^L, as factors of production. The price of output is normalized to unity. The firm operates a neoclassical constant-returns-to-scale production technology, which satisfies the Inada conditions:

$$Y = F(L^H, L^L), \quad F_H, F_L > 0, \quad F_{HH}, F_{LL} < 0, \quad F_{HL} > 0,$$
$$\lim_{H\to\infty} F_H = \lim_{L\to\infty} F_L = 0, \quad \lim_{H\to 0} F_H = \lim_{L\to 0} F_L = \infty. \tag{3.15}$$

The subscripts H and L of the production function denote partial derivatives with respect to L^H and L^L. The marginal products of labor are positive, but diminishing for each type of labor. Both inputs are essential and cooperant factors of production.

Firms demand labor, taking wage rates as given. The labor market is perfectly competitive and frictionless. The first-order conditions for profit maximization imply that the marginal labor productivities equal the wage rates of each type of worker:

$$F_H(L^H, L^L) = w^H, \tag{3.16}$$
$$F_L(L^H, L^L) = w^L. \tag{3.17}$$

These conditions, together with homogeneity of the production function, implicitly define the equilibrium factor ratio, L^H/L^L, as a function of the minimum wage, w^L. The Inada-conditions, joint with the cut-off ability level, Θ, in equation (3.9), imply that in equilibrium the numbers of high- and low-skilled individuals are both strictly positive, i.e., $\underline{\theta} < \Theta < \overline{\theta}$.

3.3.3 Labor market equilibrium

Labor-market equilibrium conditions for high-skilled and low-skilled workers are given by:

$$L^H = L^H_S = \int_{\Theta}^{\overline{\theta}} l^H_\theta \mathrm{d}G(\theta), \tag{3.18}$$
$$L^L = L^L_S - \int_{\underline{\theta}}^{\Theta} u l^L_\theta \mathrm{d}G(\theta) = (1-u)\int_{\underline{\theta}}^{\Theta} l^L_\theta \mathrm{d}G(\theta). \tag{3.19}$$

High-skilled labor demand should equal high-skilled labor supply, since the high-skilled wage adjusts to clear the labor market. Low-skilled labor demand equals low-skilled labor supply, minus the potential working hours of the unemployed.

The latter equality follows from the assumption of uniform rationing, i.e., independence of u from θ.

3.4 Comparative statics

We derive comparative statics to determine how unemployment and labor supply respond to a change in one of the policy variables. We do so by log-linearizing the model to find the (semi-)elasticities of the endogenous variables with respect to the policy variables: w^L, t, and T. These elasticities are an important ingredient of the government's optimization problem that we solve later. Equilibrium is described by equation (3.9) for Θ, the labor demand equations (3.16) and (3.17), and the two labor market clearing conditions (3.18) and (3.19).

We denote a relative change in variable x by $\tilde{x} \equiv \mathrm{d}\ln x = \mathrm{d}x/x$. Exceptions are variables that are already expressed in percentage terms: $\tilde{t} \equiv \mathrm{d}t/(1-t)$, $\tilde{u} \equiv \mathrm{d}u/(1-u)$, and $\tilde{N}^i \equiv \mathrm{d}N^i$, $i \in \{H, L, U\}$. As the latter variables are already expressed in percentage terms it is more convenient to write the elasticities of these variables as semi-elasticities. Loglinearization of the cut-off ability level, the first-order conditions for the firm, and the labor-market equilibrium conditions yields:

$$\tilde{\Theta} = \frac{1+\varepsilon}{(1-\beta)\varepsilon}(\tilde{w}^L - \tilde{w}^H) - \frac{1}{(1-\beta)\varepsilon}\tilde{u}, \tag{3.20}$$

$$\tilde{w}^H = \frac{1-\alpha}{\sigma}(\tilde{L}^L - \tilde{L}^H), \tag{3.21}$$

$$\tilde{w}^L = \frac{\alpha}{\sigma}(\tilde{L}^H - \tilde{L}^L), \tag{3.22}$$

$$\tilde{L}^H = -\frac{l^H_\Theta \Theta g(\Theta)}{L^H}\tilde{\Theta} + \varepsilon(\tilde{w}^H - \tilde{t}), \tag{3.23}$$

$$\tilde{L}^L = \frac{(1-u)l^L_\Theta \Theta g(\Theta)}{L^L}\tilde{\Theta} - \tilde{u} + \varepsilon(\tilde{w}^L - \tilde{t}), \tag{3.24}$$

where $\alpha \equiv F_H L^H/Y$ denotes the share of skilled labor earnings in total income and $\sigma \equiv -\mathrm{d}\ln(L^H/L^L)/\mathrm{d}\ln(F_H/F_L) = F_H F_L/(F_{HL}Y)$ is the elasticity of substitution between skilled and unskilled workers in production.

Combining these equations, and substituting for $\Theta g(\Theta)\tilde{\Theta} = -\tilde{N}^H$, yields a system of two equations: one for households and one for firms, relating changes in high-skilled employment and the low-skilled unemployment rate to changes in the minimum wage:

$$\text{(3.25)} \qquad \frac{(1-\beta)\varepsilon}{\Theta g(\Theta)}\tilde{N}^H = -\left(\frac{1+\varepsilon}{\alpha}\right)\tilde{w}^L + \tilde{u},$$

$$\text{(3.26)} \qquad \tilde{u} = \left(\frac{\sigma+\varepsilon}{\alpha}\right)\tilde{w}^L - \left(\frac{l_\Theta^H}{L^H} + \frac{(1-u)l_\Theta^L}{L^L}\right)\tilde{N}^H.$$

The first equation shows that, given the unemployment rate, a higher minimum wage reduces the number of high-skilled, because of a fall in the skill premium. Higher unemployment results in more skill formation, since individuals would like to avoid unemployment, which is concentrated among the unskilled. The second equation shows that, for a given number of high-skilled workers, a rise in the minimum wage increases the unemployment rate. As the minimum wage rises, firms start laying off low-skilled workers. A higher number of high-skilled workers increases the return to low-skilled labor and thus decreases the unemployment rate for a given minimum wage.

The equilibrium conditions can be solved for the changes in the number of high-skilled workers and the unemployment rate to find:

$$\text{(3.27)} \qquad \frac{\tilde{N}^H}{\tilde{w}^L} = \frac{\sigma-1}{\alpha\eta},$$

$$\text{(3.28)} \qquad \frac{\tilde{u}}{\tilde{w}^L} = \frac{\sigma+\varepsilon-(\sigma-1)\kappa}{\alpha} > 0,$$

where $\eta \equiv \frac{(1-\beta)\varepsilon}{\Theta g(\Theta)} + \frac{(1-u)l_\Theta^L}{L^L} + \frac{l_\Theta^H}{L^H} > 0$ and $\kappa = \left(\frac{(1-u)l_\Theta^L}{L^L} + \frac{l_\Theta^H}{L^H}\right)\eta^{-1} \in (0,1)$. The sign of (3.28) follows from $\sigma+\varepsilon-(\sigma-1)\kappa = (1-\kappa)\sigma+\varepsilon+\kappa > 0$.

As we can see from the first equation, there is a knife-edge condition that determines whether a rise in the minimum wage increases or decreases the amount of skill formation in the economy. If $\sigma < 1$, an increase in the minimum wage leads to less high-skilled workers. Intuitively, if high-skilled workers and low-skilled workers are poor substitutes, firms are less willing to substitute low-skilled workers for high-skilled workers. Therefore, employment of low-skilled workers does not fall enough compared to the drop in the skill premium to induce individuals to invest more in human capital. For $\sigma > 1$ the converse is true, and unemployment of low-skilled workers increases so much that individuals invest more in human capital, even though the skill premium has decreased. For $\sigma = 1$, an increase in the minimum wage has no effect on the share of high-skilled workers, since the effects of a lower skill premium exactly offsets the effect of a higher unemployment rate.

This result might be sensitive to a number of simplifying assumptions we made. First, the assumption of uniform rationing ensures that unemployment affects skill formation. Had rationing been more efficient and had low-skilled workers with ability Θ had a higher chance of obtaining a job than other low-skilled workers, unemployment would have had a smaller effect on skill formation. Second, the assumption of quasi-linear utility, or risk-neutrality, affects the effect of unemployment on skill formation. Had individuals been risk-averse, they would have been more averse to the possibility of unemployment, and a minimum wage would have had a more positive effect on skill formation. Finally, because in this simple setup low-skilled workers and high-skilled workers face the same tax rates, t, and transfer, T, these tax instruments do not affect the effect of a minimum wage on skill formation. This changes once we allow for skill-specific tax rates and transfers, as we show below.

What are plausible values for the substitution elasticity is an empirical question. Estimates of the substitution elasticity between high- and low-skilled workers are typically found to be larger than one. Katz and David (1999) find that a common estimate for σ is around 1.4, although much higher estimates are also reported. Hence, in the simple setup of our model, empirically plausible values for the substitution elasticity imply that the introduction of a minimum wage would typically lead to more skill formation. This finding is similar to the finding that, in response to a minimum-wage increase, total wage income of the affected group declines if labor demand elasticities for workers earning a minimum wage exceed unity (e.g., Dolado, Felgueroso and Jimeno, 2000; Freeman, 1996).

From the second equation follows that a minimum wage unambiguously increases the unemployment rate amongst the low-skilled. The first two terms in equation (3.28), σ and ε, represent labor demand and intensive labor supply responses to a higher minimum wage. An increase in the minimum wage leads to lower labor demand and higher intensive labor supply, both contributing to an increase in unemployment. The third term $-(\sigma-1)\kappa$ represents the human capital response. If $\sigma > 1$, the increase of the minimum wage leads to more skill formation, which renders this term negative, so that the unemployment effect diminishes. Intuitively, if more low-skilled workers transfer to the skilled sector, less of them are laid off.[10] Assuming $\sigma > 1$, studies that do not take into account human capital

[10] There is a large empirical literature dealing with the effect of a higher minimum wage on total employment. Although most of the evidence seems to point to negative employment effects (e.g., Neumark and Wascher, 2006), some present evidence of positive or non-negative employment

responses to a minimum wage might therefore underestimate the desirability of a minimum wage.

Proposition 3.1 *The minimum wage reduces (increases) the fraction of skilled workers (N^H) if the elasticity of substitution between low-skilled and high-skilled workers (σ) is smaller (larger) than 1. If $\sigma = 1$, a change in the minimum wage has no effect on the number of high-skilled workers. A higher minimum wage (w^L) increases the low-skilled unemployment rate (u). A higher minimum wage boosts the unemployment rate more if σ is higher and if the elasticity of low-skilled labor supply (ε) is higher.*

3.5 Optimal skill-independent policy

So far, we discussed the equilibrium effects of higher minimum wages on skill formation and unemployment. In the remainder of the chapter, we focus on the welfare effects of a minimum wage. To illustrate these welfare effects in the most basic setup, in this section we first discuss the social desirability of a minimum wage in the absence of tax policy. The derivation of optimal minimum wages, joint with optimal skill-independent taxes and transfers follows. A treatment of optimal minimum wages and skill-dependent taxation is postponed to the next section.

3.5.1 Government's objective

The government maximizes social welfare by optimally deciding on the minimum wage, the income tax rate, and the non-individualized lump-sum transfer. We rule out individualized lump-sum taxes and transfers. Consequently, the government has to resort to distortionary policy instruments to implement its redistributive goals. All individuals receive a lump-sum transfer T and, if employed, their labor earnings are taxed at a rate t. The informational requirement to implement this linear tax system is that the government observes aggregate labor earnings. We assume for now that the government is unable to distinguish high-skilled workers from low-skilled workers for tax purposes, which implies that we do not allow for group-specific lump-sum taxes and transfers, such as education subsidies.

effects (e.g., Card and Krueger, 1995). In our model, it can be shown to be theoretically possible that a minimum wage increases high-skilled employment by so much that the number of unemployed, N^U, decreases, even though the low-skilled unemployment rate, u, increases. However, calibration points out that such a positive employment effect would only happen under extreme parameter values with σ exceeding 10 or the unemployment rate, u, exceeding 80 percent.

Social welfare, $\mathcal{W}$, is a weighted sum of utilities:

$$\mathcal{W} \equiv N^U \Psi(V^U) + (1-u)\int_{\underline{\theta}}^{\Theta} \Psi(V_\theta^L)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi(V_\theta^H)\mathrm{d}G(\theta), \tag{3.29}$$

where $\Psi(\cdot)$ is a concave function of utility, with $\Psi'(\cdot) > 0$ and $\Psi''(\cdot) \leq 0$. Since utility is assumed to be quasi-linear in income, any social desire for redistribution enters through concavity of $\Psi(\cdot)$. Thus, if the government is utilitarian ($\Psi(V_\theta^i) = V_\theta^i$, $\Psi'(V_\theta^i) = 1$), the social objective exhibits no preference for redistribution. On the other extreme, if the government is Rawlsian, it only values the utility of the least well off (V^U in the presence of unemployment, $V_{\underline{\theta}}^L$ otherwise).

The government budget constraint states that government expenditures on the lump-sum transfer, T, and an exogenously given expenditure requirement, E, equal total tax revenue from labor earnings. Hence, the budget balance, denoted by $\mathcal{B}$ must equal zero in equilibrium:

$$\mathcal{B} \equiv t\left((1-u)\int_{\underline{\theta}}^{\Theta} w^L l_\theta^L \mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} w^H l_\theta^H \mathrm{d}G(\theta)\right) - T - E = 0. \tag{3.30}$$

By defining λ as the shadow price for the budget constraint, we can write down the Lagrangian associated with the government's optimization problem as:

$$\mathcal{L} \equiv \mathcal{W} + \frac{\mathcal{B}}{\lambda}. \tag{3.31}$$

The government's first-order conditions for the minimum wage, the transfer, and the tax rate are given by:

$$\begin{aligned} &(1-u)(1-t)\int_{\underline{\theta}}^{\Theta} l_\theta^L \Psi'(V_\theta^L)\mathrm{d}G(\theta) + (1-t)\int_{\Theta}^{\overline{\theta}} l_\theta^H \Psi'(V_\theta^H)\mathrm{d}G(\theta)\frac{\mathrm{d}w^H}{\mathrm{d}w^L} \\ &\quad -\left((1-u)\int_{\underline{\theta}}^{\Theta}(\Psi(V_\theta^L) - \Psi(V^U))\mathrm{d}G(\theta) + \lambda t w^L L^L\right)\frac{1}{1-u}\frac{\mathrm{d}u}{\mathrm{d}w^L} \\ &\quad +(\Psi(V_\Theta^H) - (1-u)\Psi(V_\Theta^L) - u\Psi(V^U))\frac{\mathrm{d}N^H}{\mathrm{d}w^L} = 0, \end{aligned} \tag{3.32}$$

$$N^U \Psi'(V^U) + (1-u)\int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H)\mathrm{d}G(\theta) - \lambda = 0, \tag{3.33}$$

$$(3.34)\qquad -(1-u)w^L\int_{\underline{\theta}}^{\Theta} l_\theta^L\Psi'(V_\theta^L)\mathrm{d}G(\theta)-w^H\int_{\Theta}^{\overline{\theta}} l_\theta^H\Psi'(V_\theta^H)\mathrm{d}G(\theta)$$
$$+\lambda(w^LL^L+w^HL^H)\left(1-\varepsilon\frac{t}{1-t}\right)=0,$$

where we substituted for the derivatives of the utility functions, labor hours worked, and for $g(\Theta)\mathrm{d}\Theta=-\mathrm{d}N^H$.

3.5.2 Optimal minimum wages without taxes and transfers

To highlight the main mechanisms at work, we first determine whether the introduction of a minimum wage above the market-clearing wage for low-skilled labor is desirable in the absence of taxation. We thus set $T=t=E=0$ in order to abstract from taxation, and $u=0$ to determine the desirability of a minimum wage in an initial equilibrium without unemployment. Note that the utility of the marginal low-skilled worker in this case exactly equals that of the marginal high-skilled worker.[11] Hence, we have $\Psi(V_\Theta^H)=\Psi(V_\Theta^L)$. To interpret equation (3.32), we follow Feldstein (1972) by introducing the distributional characteristics of low-skilled and high-skilled labor income:

$$(3.35)\qquad 0\le\xi^L\equiv 1-\frac{\frac{1}{G(\Theta)}\int_{\underline{\theta}}^{\Theta}\Psi'(V_\theta^L)w^Ll_\theta^L\mathrm{d}G(\theta)}{\frac{1}{G(\Theta)}\int_{\underline{\theta}}^{\Theta}\Psi'(V_\theta^L)\mathrm{d}G(\theta)\frac{1}{G(\Theta)}\int_{\underline{\theta}}^{\Theta}w^Ll_\theta^L\mathrm{d}G(\theta)}\le 1,$$

$$(3.36)\qquad 0\le\xi^H\equiv 1-\frac{\frac{1}{1-G(\Theta)}\int_{\Theta}^{\overline{\theta}}\Psi'(V_\theta^H)w^Hl_\theta^H\mathrm{d}G(\theta)}{\frac{1}{1-G(\Theta)}\int_{\Theta}^{\overline{\theta}}\Psi'(V_\theta^H)\mathrm{d}G(\theta)\frac{1}{1-G(\Theta)}\int_{\Theta}^{\overline{\theta}}w^Hl_\theta^H\mathrm{d}G(\theta)}\le 1.$$

ξ^i, $i=\{L,H\}$, is the negative normalized covariance between the social welfare weights and labor earnings. It measures the marginal social welfare gain expressed in monetary units as a fraction of labor income from redistributing one unit of income through a lower net wage rate in skill-group i. The distributional characteristic is positive for a government that values redistribution from rich to poor, as in that case social welfare weights are decreasing with income. For a government that does not value redistribution ($\Psi'(\cdot)=1$), ξ^i equals zero. Similarly, if there is no income inequality in either group, the distributional characteristic is also zero. The distributional characteristic increases with stronger social preferences for redistribution and with larger pre-tax income inequality.

[11] We henceforth refer to a low-skilled worker with ability Θ as 'the marginal low-skilled worker,' and to the high-skilled worker with ability Θ as 'the marginal high-skilled worker.'

By rearranging the first-order condition for the minimum wage, equation (3.32), and substituting in the distributional characteristics and the elasticity of unemployment, we find that it is desirable to implement a minimum wage if the following condition holds:

$$(1-\xi^L)\overline{\Psi'(V^L)} - (1-\xi^H)\overline{\Psi'(V^H)} > \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{w^L L^L} N^L \right) \frac{\tilde{u}}{\tilde{w}^L}, \tag{3.37}$$

where $\overline{\Psi'(V^L)} \equiv \int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L)\mathrm{d}G(\theta)/G(\Theta)$ and $\overline{\Psi'(V^H)} \equiv \int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^H)\mathrm{d}G(\theta)/(1 - G(\Theta))$ are the averages of the marginal social welfare of income of skilled and low-skilled workers. $\overline{\Psi(V^L)} = \int_{\underline{\theta}}^{\Theta} \Psi(V_\theta^L)\mathrm{d}G(\theta)/G(\Theta)$ is the average social welfare of low-skilled workers.

In the absence of unemployment taxation, a minimum wage has two first-order effects on social welfare. The left-hand side of inequality (3.37) shows the distributional benefits of a higher minimum wage, whereas the right-hand side shows the efficiency costs of a higher minimum wage. A higher minimum wage affects both inequality between the groups of low-skilled and high-skilled workers and inequality within the groups of low-skilled and high-skilled workers. The minimum wage reduces inequality between high-skilled and low-skilled workers through general-equilibrium effects on the wage structure. By raising the minimum wage, low-skilled employment declines, lowering high-skilled productivity and wages. This reduction in inequality is given by the first two terms of above condition. For a government with redistributive preferences, this general-equilibrium effect raises social welfare as the average social marginal utility of a low-skilled worker is larger than that of a high-skilled worker: $\overline{\Psi'(V_L)} > \overline{\Psi'(V_H)}$.

Besides between-group inequality, the minimum wage also affects inequality within the groups of high-skilled and low-skilled workers. Since high-skilled individuals with high ability make relatively many working hours, they suffer more from a decline in their wage rate than high-skilled individuals with a lower ability. This reduction in inequality within the group of high-skilled workers contributes to the desirability of the minimum wage as long as $\xi^H > 0$. However, by raising the low-skilled wage rate per hour worked, a higher minimum wage also increases inequality in low-skilled labor earnings if $\xi^L > 0$. Therefore, the minimum wage is less attractive for redistributive reasons if it generates more inequality among low-skilled workers, or if the government is strongly averse to inequality within the group of low-skilled workers.

A minimum wage has, overall, favorable distributional gains, since $(1-\xi^L)\overline{\Psi'(V^L)} =$

$\int_{\underline{\theta}}^{\Theta} \frac{\Psi'(V_\theta^L)}{\lambda} \frac{(1-u)l_\theta^L}{L^L} \mathrm{d}G(\theta) > \int_{\Theta}^{\overline{\theta}} \frac{\Psi'(V_\theta^H)}{\lambda} \frac{l_\theta^H}{L^H} \mathrm{d}G(\theta) = (1-\xi^H)\overline{\Psi'(V^H)}$. The second and third terms give weighted averages of the social marginal utility of income for low- and high-skilled workers. The inequality follows from the fact that, in the case of $u = 0$, the social marginal utility of every low-skilled worker is higher than the social marginal utility of any high-skilled worker. However, if unemployment is positive ($u > 0$) this inequality does not necessarily hold. For example, if $\xi^L >> \xi^H$ and unemployment is positive, the minimum wage might cause net distributional costs rather than benefits. Intuitively, in that case the increase in inequality *within* the group of low-skilled workers is not off-set by a reduction in inequality *between* low-skilled and high-skilled workers and *within* the group of high-skilled workers.

The second first-order welfare effect of a minimum wage increase is associated with the resulting increase in unemployment. This welfare effect is given by the right-hand side of equation (3.37). $(\overline{\Psi(V^L)} - \Psi(V^U))/w^L L^L$ measures the welfare loss due to larger unemployment in terms of total low-skilled income. For every laid off low-skilled worker society looses on average $\overline{\Psi(V^L)}$ of social welfare and gains $\Psi(V^U)$. Since the unemployed have lower utility levels than the employed, larger unemployment results in lower social welfare. Notice that any social welfare effect of the minimum wage on human capital investment is a second-order effect when there is no unemployment, since the utility of a marginal high-skilled worker in that case equals the utility of a marginal low-skilled worker: $V_\Theta^L = V_\Theta^H$.

The desirability of a minimum wage crucially depends on the elasticity of the unemployment rate with respect to the minimum wage, $\tilde{u}/\tilde{w}^L$, given by equation (3.28). In particular, a minimum wage raises unemployment more if the elasticity of substitution between skilled and unskilled workers, σ, is larger, the labor supply elasticity of low-skilled workers, ε, is larger, and, assuming $\sigma > 1$, if the human capital response, $(\sigma - 1)\kappa$, is smaller.

That unemployment results in a first-order welfare loss is an important deviation from Lee and Saez (2012) who assume efficient rationing. In Lee and Saez, the marginal laid-off worker has zero surplus from working, and is thus indifferent between working and being unemployed. Consequently, starting from a situation without unemployment, the social welfare loss of larger unemployment is only a second-order effect. In our model this does not hold, since every low-skilled worker prefers working over being unemployed. Contrary to Lee and Saez, individuals in our model incur disutility of work on the intensive margin, so that lay-offs are always inefficient. Later on, we briefly turn to the case of efficient rationing.

As we can see from equation (3.37), the desirability of a minimum wage depends on the redistributive preferences of the government. If it does not value redistribution – i.e., in the case of a utilitarian social welfare function – $\overline{\Psi'(V^L)} = \overline{\Psi'(V^H)} = 1$, and $\xi^L = \xi^H = 0$, such that the left-hand side of the inequality vanishes. Therefore, the government would not want to introduce a distortionary minimum wage as it produces no distributional benefits. If the government has Rawlsian preferences, the social welfare function without unemployment is given by $\mathcal{W} = V^L_{\underline{\theta}}$ and with unemployment is given by $\mathcal{W} = V^U$. In that case, a minimum wage always decreases social welfare, since the government only cares for the utility of the least well off. On both extremes of the spectrum of redistributive preferences – without any redistributive and with maximum redistributive preferences – a minimum wage is not desirable. However, for intermediate cases of redistributive preferences, this is not necessarily the case.

Proposition 3.2 *Starting from an undistorted initial equilibrium, the introduction of a minimum wage has ambiguous welfare effects for any redistributive, non-Rawlsian social welfare function. A minimum wage is more likely to be socially desirable if the elasticity of substitution between high-skilled and low-skilled workers (σ) is small, if the labor supply elasticity (ε) is small, if the welfare differential between the low-skilled employed and the unemployed, $\overline{\Psi(V^L)} - \Psi(V^U)$, is small, and if the general equilibrium effects on wages yield large distributional gains, such that $(1-\xi^L)\overline{\Psi'(V^L)} - (1-\xi^H)\overline{\Psi'(V^H)}$ is large. Distributional gains are higher with larger inequality between skilled and unskilled workers, $\overline{\Psi'(V^L)} >> \overline{\Psi'(V^H)}$, and with larger inequality within the group of high-skilled workers compared to low-skilled workers, $\xi^H >> \xi^L$. A minimum wage is never optimal if the social welfare function is Rawlsian or when it exhibits no preference for redistribution.*

To find the optimal minimum wage in the absence of taxes and transfers, we rewrite the first-order condition for the minimum wage (3.32) to obtain:

$$(1-\xi^L)\overline{\Psi'(V^L)} - (1-\xi^H)\overline{\Psi'(V^H)} = \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{w^L L^L} N^L\right)\frac{\tilde{u}}{\tilde{w}^L} - \left(\frac{\Psi(V^H_\Theta) - (1-u)\Psi(V^L_\Theta) - u\Psi(V^U)}{w^L L^L}\right)\frac{\tilde{N}^H}{\tilde{w}^L}. \tag{3.38}$$

Notice that this is only an optimality condition provided that the desirability condition (3.37) holds. Given that a binding minimum wage is indeed welfare increasing, the optimal minimum wage is set in such a way that the marginal redistributive gains due to lower income inequality between low-skilled and high-skilled workers (left-hand side) equals the marginal welfare losses of raising involuntary unemployment (right-hand side, first term) minus the marginal welfare gain (or loss) associated with the change in skill formation (right-hand side, second term). The first two terms are discussed above, the last one is new.

The second term on the right-hand side represents a positive externality from skill formation. If the government has redistributive preferences, and thus if $\Psi(\cdot)$ is strictly concave, we can establish that $\Psi(V_{\Theta}^{H})-(1-u)\Psi(V_{\Theta}^{L})-u\Psi(V^{U})>0$ if $u>0$, since we know from equation (3.10) that $V_{\Theta}^{H}-(1-u)V_{\Theta}^{L}-uV^{U}=0$. Intuitively, becoming high skilled can be seen as an insurance against unemployment. Due to concave social preferences, the government attaches a higher cost to the risk of becoming unemployed than risk-neutral individuals themselves. Therefore, in the presence of involuntary unemployment, the social value of skill formation exceeds its private value. A binding minimum wage, resulting in a positive unemployment rate, thus leads to an externality on skill formation. Clearly, there is no externality in the absence of unemployment as it vanishes for $u=0$ or when the government has no desire to redistribute income.[12] A higher minimum wage changes human capital formation if $\tilde{N}^{H}/\tilde{w}^{L}\neq 0$, and thus, as we have seen above, if $\sigma\neq 1$. As discussed, if $\sigma>1$, a higher minimum wage leads to more high-skilled workers. Due to the positive externality associated with skill formation, a higher minimum wage generates an additional welfare gain. If $\sigma<1$ a higher minimum wage leads to less skill formation, exacerbating the inefficiently low degree of skill formation.

3.5.3 Optimal minimum wages, taxes and transfers

Naturally, governments also employ tax instruments to redistribute income. Below, we derive how the welfare effect of a minimum wage are altered if the government optimally sets a skill-independent income tax rate, t, and a transfer, T.

Optimal transfer

The first order condition for the transfer, equation (3.33), can be rewritten to find that the social marginal benefits of a higher transfer, T, should equal its social

[12]By assuming quasi-linear utility functions we abstracted from risk aversion at the individual level. Risk aversion would reduce the positive externality from human capital investment, since individuals hedge against against labor market risk by investing more in human capital (see also Jacobs, Schindler and Yang, 2012). However, the positive externality will not disappear as long as the social welfare function is a (strictly) concave transformation of the individuals' private utility functions.

marginal costs:

$$(3.39)\quad \frac{\overline{\Psi'(\cdot)}}{\lambda} \equiv N^U \frac{\Psi'(V^U)}{\lambda} + (1-u)\int_{\underline{\theta}}^{\Theta} \frac{\Psi'(V_\theta^L)}{\lambda}\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \frac{\Psi'(V_\theta^H)}{\lambda}\mathrm{d}G(\theta) = 1,$$

where $\overline{\Psi'(\cdot)}$ denotes the average social marginal utility of income. The first three terms on the left-hand side give the increase in social welfare (expressed in monetary units) of the unemployed, low-skilled employed and high-skilled employed due to a marginally higher lump-sum transfer. This equals the transfer's resource costs on the right-hand side, equaling 1.

Optimal tax rate

In order to derive the optimal tax rate, we again follow Feldstein (1972) by introducing the distributional characteristic of total labor income:

$$(3.40)\quad \xi \equiv 1 - \frac{(1-u)\int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L) w^L l_\theta^L \mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H) w^H l_\theta^H \mathrm{d}G(\theta)}{(w^L L^L + w^H L^H)\left(N^U \Psi'(V^U) + N^L \overline{\Psi'(V_\theta^L)} + N^H \overline{\Psi'(V^H)}\right)} \geq 0$$

The interpretation of ξ is identical to the distributional characteristics of low-skilled and high-skilled earnings introduced in equations (3.35) and (3.36). It is the negative of the normalized covariance between the social welfare weights and labor earnings across the entire population.[13] We can establish a direct link between the distributional characteristics of skilled and unskilled labor income and the distributional characteristic for aggregate labor income:

$$(3.41)\quad (1-\xi) \equiv (1-\alpha)\frac{\overline{\Psi'(V^L)}}{\overline{\Psi'(\cdot)}}(1-\xi^L) + \alpha\frac{\overline{\Psi'(V^H)}}{\overline{\Psi'(\cdot)}}(1-\xi^H).$$

Consequently, one minus the distributional characteristic of aggregate labor is a weighted sum of one minus the distributional characteristics of skilled and unskilled labor – where the income shares α and $1-\alpha$ have been used as weights and a correction has been made for the differences in the average social marginal utility of income. For a government without redistributive preferences, $\xi = 0$, for a Rawlsian government, $\xi = 1$.

Labor income taxation distorts labor supply. The tax rate drives a wedge between the private and social benefits of work, leading to a substitution effect

[13] Note that labor earnings of the unemployed are zero, and do not feature in the numerator.

from consumption to leisure. The marginal deadweight loss associated with this distortion, given by $\frac{t}{1-t}\varepsilon$, is increasing in the elasticity of labor supply and the tax rate. The optimal tax rate is set such that the marginal redistributive gains of the tax rate equal its marginal efficiency costs. Rearranging the first order condition for t, equation (3.34), thus yields:

$$\frac{t}{1-t} = \frac{\xi}{\varepsilon}. \tag{3.42}$$

From this equation we derive the familiar result of optimal tax theory, that the optimal income tax rate is increasing in the distributional gain and decreasing in the labor-supply elasticity, see for example Atkinson and Stiglitz (1980).

Optimal minimum wage

Again, we first analyze the desirability of a minimum wage, when taxes are optimally set, by substituting for $u = 0$ in first-order condition (3.32). We furthermore substitute for the partial derivative $\mathrm{d}w^H/\mathrm{d}w^L$, implied by equations (3.21) and (3.22). A minimum wage is desirable if the following condition holds:

$$(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} - (1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} > \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{(1-t)w^L L^L \lambda} N^L + \frac{t}{1-t}\right)\frac{\tilde{u}}{\tilde{w}^L}. \tag{3.43}$$

As before, the first line gives the marginal redistributive gains of a minimum wage, the second line gives the welfare loss associated with higher unemployment, multiplied by the semi-elasticity of unemployment with respect to a higher minimum wage. The first term in the second line again gives the welfare loss associated with the direct utility drop of the workers who lose their job, this time normalized by low-skilled income net of taxes. With positive taxes, low-skilled workers do not reap the full benefits of a higher minimum wage as part of it is taxed away by the government. Hence, a given distributional gain requires a larger increase in the minimum wage and therefore increased unemployment is associated with a higher welfare loss.

The second term in the second line, $t/(1-t)$, is new and captures the welfare costs of a higher minimum wage associated with an erosion of the labor tax base. Unemployed workers do not pay income taxes, whereas employed workers do. $t/(1-t)(\tilde{u}/\tilde{w}_L)$ represents these losses in tax revenue from low-skilled workers as a result of a higher minimum wage. Tax revenue declines more if the increase in unemployment due to the minimum wage is larger. This is captured by the term

$\tilde{u}/\tilde{w}_L = (\sigma + \varepsilon - (\sigma - 1)\kappa)/\alpha$, the semi-elasticity of unemployment with respect to the minimum wage. This term has been extensively discussed above. $t/(1-t)$ is the tax wedge on low-skilled labor supply. The larger are tax distortions on labor supply – i.e., the larger is t – the costlier it is to raise the minimum wage. Thus the minimum wage exacerbates the distortions of the labor tax on low-skilled labor supply by further eroding the tax base.

Minimum wage versus income taxation

As equation (3.43) shows, the main benefit of a minimum wage is its capacity to redistribute income from high-skilled workers to low-skilled workers. Furthermore, it decreases inequality among the high-skilled but increases inequality among the low-skilled, and creates unemployment, causing a drop of utility and tax revenues from those who lose their jobs. A minimum wage is not the only means of redistributing income from high- to low-skilled workers. Redistribution can also be achieved through a higher income-tax rate, while rebating revenue in the form of higher transfers. A minimum wage will be optimal if, and only if, the marginal costs of redistribution through a minimum wage are smaller than the marginal costs of the same redistribution through higher income taxes, evaluated at the tax optimum.

To see whether this is indeed the case, we first rewrite the first-order condition for the optimal tax rate, equation (3.42), by substituting for ξ, using equation (3.41):

$$(3.44)\quad \frac{t}{1-t}\varepsilon = (\alpha - N^H)\frac{\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}}{\lambda} + (1-\alpha)\frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \alpha\frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H.$$

The left-hand side gives the marginal dead-weight loss of taxation. The right-hand side gives the distributional benefits of taxation by reducing inequality between skill groups (first-term), reducing inequality within the low-skilled group (second term), and reducing inequality within the high-skilled group (third term).

To determine if a binding minimum wage is optimal, we derive the net welfare effect of a marginal increase in the minimum wage, and compare this to the net welfare effect of a marginal increase in the income tax rate that achieves the exact same within-group redistribution as the minimum wage. Clearly, a binding minimum wage is optimal if, at the tax optimum without a minimum wage, this net welfare effect of a minimum wage increase outweighs the net welfare effects of the tax-rate increase. To see when this is the case, we substitute for $\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}$

from equation (3.44) into the desirability condition of a minimum wage, equation (3.43). The desirability condition of the minimum wage can then be written as follows:

$$\frac{t}{1-t}\left(\frac{\varepsilon}{\alpha-N^H}-\frac{\tilde{u}}{\tilde{w}^L}\right) > \left(\frac{\overline{\Psi(V^L)}-\Psi\left(V^U\right)}{(1-t)L^L w^L \lambda}N^L\right)\frac{\tilde{u}}{\tilde{w}^L} + \left(\frac{1-N^H}{\alpha-N^H}\right)\frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \left(\frac{N^H}{\alpha-N^H}\right)\frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H. \tag{3.45}$$

The first term on the left-hand side gives the efficiency costs of attaining a given between-group redistribution through an increase in the tax rate. The second term gives the efficiency costs of attaining the same between-group redistribution through higher minimum wages. The efficiency costs of a higher tax rate are given by the tax base erosion that takes place due to the distortion of intensive labor supply, $\frac{t}{1-t}\frac{\varepsilon}{(\alpha-N^H)}$. The denominator, $\alpha - N^H$, is of special interest: the smaller the high-skilled share of total income, relative to its population share, the more difficult it is to redistribute from high-skilled workers to low-skilled workers by using the income tax rate, t, which applies to both high-skilled and low-skilled workers.

The efficiency costs of a higher minimum wage are given by the tax-base erosion that takes place due to the distortion on the extensive employment margin, $\frac{t}{1-t}\frac{\tilde{u}}{\tilde{w}^L}$. Overall, a minimum wage might be more efficient in redistributing income from high-skilled workers to low-skilled workers because it directly raises low-skilled income and decreases high-skilled income. Uniform income taxes, on the contrary, cause net wages of both low-skilled and high-skilled workers to decline in order to redistribute the revenue back in the form of lump-sum transfers to both low-skilled and high-skilled workers.

However, for minimum wages to be optimal, this larger efficiency must outweigh its distributional losses relative to a tax increase. These losses are given by the right-hand side and consist of the direct welfare losses of laid-off individuals (first term) and the within-group distributional gains of a tax increase (last two terms). The first term we encountered and discussed before. The last two terms indicate the relative advantage of the income tax rate to achieve within-group inequality. In order to achieve a given between-group redistribution, the tax increase leads to more redistribution within the groups of high- and low-skilled workers than does the increase of the minimum wage.

In short, the minimum wage complements the income tax to reduce between-group income inequality. A minimum wage helps to directly redistribute income from high-skilled workers towards low-skilled workers without the tax-base erosion on the intensive margin associated with taxation, and thereby alleviates the distributional imperfection associated with the uniformity of the income tax.

Proposition 3.3 *Optimal labor-income taxes increase with the level of earnings inequality and decrease with the elasticity of intensive labor supply. Minimum wages are more distortionary if the government sets high taxes on labor earnings, since minimum wages erode the tax base by causing unemployment. Hence, a minimum wage is less desirable in the presence of taxes. The role of minimum wages in an optimal skill-independent tax-benefit system is to complement the tax-benefit system by reducing the distributional imperfections of the income tax. Minimum wages help to redistribute income between skill groups, so that income taxes can be better targeted at reducing inequality within skill groups.*

3.6 Optimal skill-dependent policies

So far, we assumed that the government cannot differentiate tax instruments according to skill type, whereas it did employ a minimum-wage policy, the enforcement of which requires knowledge on individuals' skill type. One may recognize this as an informational inconsistency. To implement and enforce a minimum wage, the government necessarily has information on the individuals' wage rates, but it does not use this information in determining optimal tax policy. This section, therefore, explores the implications of allowing the government to optimize a skill-dependent optimal tax and minimum-wage policy.[14] Before deriving expressions for optimal policy, we first repeat the comparative-statics analysis as the key elasticities of the model change due to the introduction of skill-specific instruments.

[14]We do not study participation-dependent policies. That is, the government is still assumed to be unable to condition taxes and transfers based on employment status. Allowing for separate unemployment benefits would necessitate the introduction of a participation margin and hence a second cut-off level for θ. To keep the model tractable we decided not to do so. In Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation), we do model the participation margin along with the skill decision. We do not, however, study optimal participation taxes or subsidies. See Gerritsen (2013*a*, Chapter 5 of this Dissertation) where this is done for a model with laobr supply on the extensive margin only.

3.6.1 Comparative statics again

We introduce income taxes, t^L and t^H, that depend on skill type. Moreover, a separate transfer, S, is given to high-skilled workers. Indirect utility is thus given by:

$$V_U = T, \tag{3.46}$$

$$V_\theta^L = T + \frac{\theta^{\beta\varepsilon}((1-t^L)w^L)^{1+\varepsilon}}{1+\varepsilon}, \tag{3.47}$$

$$V_\theta^H = T + \frac{\theta^{\varepsilon}((1-t^H)w^H)^{1+\varepsilon}}{1+\varepsilon} + S. \tag{3.48}$$

Note that the only changes as compared to equations (3.6)-(3.8) are the substitutions of t^L and t^H for t, and the inclusion of an extra transfer, S, for the high-skilled. As before, the critical value, Θ is determined by $V_\Theta^H = uV^U + (1-u)V_\Theta^L$, and, hence:

$$((1-t^H)w^H)^{1+\varepsilon}\Theta^\varepsilon - (1-u)((1-t^L)w^L)^{1+\varepsilon}\Theta^{\beta\varepsilon} = -(1+\varepsilon)S. \tag{3.49}$$

The other equilibrium conditions of the model are unaffected.

The comparative statics for Θ are different in the presence of skill-dependent tax instruments. We define $\rho \equiv \frac{S}{(1-t^H)w^H l_\Theta^H/(1+\varepsilon)+S}$, which gives the transfer to high-skilled workers, S, as a share of the total return to working as a high-skilled worker with ability Θ. The loglinearized equation for Θ is now given by:

$$(1-\beta-\rho)\varepsilon\tilde{\Theta} = (1+\varepsilon)(\tilde{w}^L - \tilde{t}^L) - (1-\rho)(1+\varepsilon)(\tilde{w}^H - \tilde{t}^H) - \rho\tilde{S} - \tilde{u}, \tag{3.50}$$

where $\tilde{S} \equiv \mathrm{d}S/S$. Thus, Θ increases in earnings for low-skilled workers and decreases in earnings for high-skilled workers and the unemployment rate.[15]

By redefining $\eta \equiv \frac{(1-\beta-\rho)\varepsilon}{\Theta g(\Theta)} + \frac{(1-u)l_\Theta^L}{L^L} + \frac{l_\Theta^H}{L^H}$, we can solve the log-linearized model to find the elasticities of the number of skilled workers with respect to the policy variables:

$$\eta\tilde{N}^H = \left(\frac{\sigma - 1 + (1+\varepsilon)(1-\alpha)\rho}{\alpha}\right)\tilde{w}^L + \tilde{t}^L - (1-(1+\varepsilon)\rho)\tilde{t}^H + \rho\tilde{S}. \tag{3.51}$$

[15]For high-skilled workers to be located at $\theta > \Theta$ and low-skilled workers at $\theta < \Theta$ we require that the difference between high-skilled utility and expected low-skilled utility is increasing in θ. Taking the derivative of the equilibrium condition with respect to θ, we thus obtain the second-order condition, which is necessary and sufficient for equation (3.49) to describe the equilibrium value of Θ: $(1-\beta-\rho S) > 0$.

Observe that, for $t^L = t^H$ and $S = 0$, the equation collapses to equation (3.27). A transfer to the high skilled alters the result that skill formation increases (decreases) in response to a higher minimum wage if $\sigma > 1$ ($\sigma < 1$). In particular, high-skilled workers now have an education subsidy or tax, S, which is unaffected by a higher minimum wage. Thus, while the minimum wage depresses high-skilled labor earnings, the effect of minimum wages on skill formation is cushioned due to the presence of non-wage income if $S > 0$. The exact opposite holds for $S < 0$, in which case wages make up for a larger share of net income for high-skilled workers.[16] Now that taxes are conditioned on skill type, they do affect skill formation. Quite naturally, skill formation increases with low-skilled taxes t^L and high-skilled transfers S, and decreases with high-skilled taxes t^H, provided that $(1+\varepsilon)\rho < 1$, which is what we assume.[17] Note that, if high-skilled workers receive higher transfers ($\rho > 0$ larger), the impact of the high-skilled tax rate on skill formation is lowered, as the transfers remain untaxed.

As before, we can solve the linearized model to find the elasticities of unemployment with respect to the policy variables:

$$\tilde{u} = \left(\frac{\sigma + \varepsilon - \kappa(\sigma - 1) - \kappa(1+\varepsilon)(1-\alpha)\rho}{\alpha}\right)\tilde{w}^L - (\kappa + \varepsilon)\tilde{t}^L + (\varepsilon + \kappa - \kappa(1+\varepsilon)\rho)\tilde{t}^H - \kappa\rho\tilde{S}. \tag{3.52}$$

Again, for $t^L = t^H$ and $S = 0$, this equation collapses to equation (3.28). The minimum wage has a smaller effect on unemployment if $S > 0$ ($\rho > 0$). The reason is that the impact of the minimum wage on skill formation is larger if $S > 0$. Intuitively, compared to the case in which $S = 0$, a minimum wage results in lower low-skilled labor supply, and (through input complementarity) higher low-skilled labor demand. Of course, the opposite holds if $S < 0$. Furthermore, low-skilled taxes decrease unemployment as it discourages low-skilled labor supply; high-skilled taxes increase unemployment as it discourages high-skilled labor supply (and the less so if S is larger); and transfers to high-skilled workers decreases unemployment as it encourages high-skilled labor supply.

[16]Naturally, to the extent that the subsidy S itself is negatively affected by the minimum wage – for example, through the high-skilled wages of teachers – this effect is smaller and might even disappear.

[17]If $(1+\varepsilon)\rho > 1$ a higher tax on high-skilled labor earnings leads to lower intensive high-skilled labor supply, lower low-skilled productivity, and higher unemployment, and thereby to more skill formation. This effect would then outweigh the direct negative effect on skill formation. However, notice that this presupposes a share of education subsidies in total high-skilled earnings, ρ, exceeding $1/(1+\varepsilon)$, which for plausible levels of ε would by highly unrealistic.

3.6.2 Optimal minimum wages, taxes and transfers

First-order conditions of optimal policy

This subsection presents the first-order conditions for the optimal minimum wage, low- and high-skilled income taxes, and lump-sum transfers for low-skilled and high-skilled workers. To interpret the optimal tax expressions, we introduce some simplifying notation. As usual, first-order conditions equate marginal distributional gains with marginal distortionary costs. Distortionary costs are represented by wedges multiplied by elasticities. Wedges in our model are defined as follows:

(3.53)
$$\Delta^H \equiv \frac{t^H}{1-t^H},$$
(3.54)
$$\Delta^L \equiv \frac{t^L}{1-t^L},$$
(3.55)
$$\Delta^U \equiv \frac{\overline{\Psi(V^L)} - \Psi(V^U)}{\lambda(1-t^L)L^L w^L} N^L,$$
(3.56)
$$\Delta^S \equiv \frac{\Psi(V^H_\Theta) - (1-u)\Psi(V^L_\Theta) - u\Psi(V^U)}{\lambda(1-t^L)L^L w^L} + \frac{t^H w^H l^H_\Theta - S - (1-u)t^L w^L l^L_\Theta}{(1-t^L)L^L w^L}.$$

These wedges measure the welfare gains of marginally higher intensive high-skilled labor supply, higher intensive low-skilled labor supply, lower unemployment, and higher skill formation. The interpretation is straightforward. The first two wedges measure revenue gains from higher labor supply as labor supply is distorted by income taxation. They are both expressed in terms of after-tax income. The third wedge gives the social welfare loss of higher unemployment due to a drop of utility – expressed in monetary terms as a fraction of net low-skilled earnings. The fourth wedge measures the welfare gains of higher skill formation. The first term gives the welfare gain of skill formation associated with the fact that the government is more averse to unemployment risk than individuals are. The second term gives the revenue gains (or losses) associated with larger skill formation. Whether there are revenue gains or losses depends on whether human capital formation is taxed or subsidized on a net basis, i.e., whether $t^H w^H l^H_\Theta - S - (1-u)t^L w^L l^L_\Theta \gtrless 0$. Both are expressed in terms of net low-skilled labor earnings.

Armed with the additional notation, we can express the first order-conditions as:

$$(3.57)\quad w^L : (1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} - \left(\frac{1-t^H}{1-t^L}\right)(1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} = \left(\frac{t^H - t^L}{1-t^L}\right)(1+\varepsilon) + (\Delta^U + \Delta^L)\frac{\tilde{u}}{\tilde{w}^L} - \Delta^S \frac{\tilde{N}^H}{\tilde{w}^L},$$

$$(3.58)\quad t^H : 1-(1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} = \Delta^H\varepsilon + (\varphi\Delta^U + \varphi\Delta^L)\frac{\tilde{u}}{\tilde{t}^H} - \varphi\Delta^S\frac{\tilde{N}^H}{\tilde{t}^H},$$

$$(3.59)\quad t^L : 1-(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} = \Delta^L\varepsilon + (\Delta^U + \Delta^L)\frac{\tilde{u}}{\tilde{t}^L} - \Delta^S\frac{\tilde{N}^H}{\tilde{t}^L},$$

$$(3.60)\quad S : \frac{\overline{\Psi'(V^H)}}{\lambda} - 1 = (\gamma\Delta^U + \gamma\Delta^L)\frac{\tilde{u}}{\tilde{S}} - \gamma\Delta^S\frac{\tilde{N}^H}{\tilde{S}},$$

$$(3.61)\quad T : 1 - \frac{N^U\Psi'(V^U) + N^L\overline{\Psi'(V^L)} + N^H\overline{\Psi'(V^H)}}{\lambda} = 0,$$

where we denoted $\varphi \equiv \frac{(1-t^L)L^L w^L}{(1-t^H)L^H w^H}$ as total low-skilled labor income relative to total high-skilled labor income, and $\gamma \equiv \frac{(1-t^L)w^L L^L}{N^H S}$ as total low-skilled labor income relative to high-skilled transfers.

Each expression implies that the net redistributive gains (left-hand side) optimally equal distortionary costs (right-hand side). The net redistributive gains always consist of the direct distributional impact (measured in monetary equivalents) of increasing the particular instrument under consideration, plus the impact of redistributing in lump-sum fashion any additional revenue. The efficiency costs are always determined by the behavioral responses on the intensive labor supply margins, the unemployment margin and the skill-formation margin. The first-order condition for T in equation (3.61) remains unaltered, and will not be discussed any further.

Optimal minimum wage

In equation (3.57), the first term on the right-hand side is new and the last term is modified as compared to 3.37. The first term on the right-hand side captures the marginal revenue gains (or losses) of the minimum wage, due to its effects on gross wage rates. It affects tax revenue from low-skilled and high-skilled workers differently if they face different tax rates. We say there is 'tax-rate progression' if taxes on skilled labor are higher than on unskilled labor, i.e., if $t^H > t^L$. The minimum wage increases low-skilled wages and lowers high-skilled wages, hence low-skilled labor supply increases and high-skilled labor supply falls. If there is tax rate progression, the minimum wage therefore causes a revenue loss, given that both low-skilled and high-skilled workers have the same labor-supply elasticity. These two effects exactly cancel out in the case of a flat tax rate, i.e., if $t^L = t^H$.

A second difference might originate from the last term, $-\Delta^S \tilde{N}^H/\tilde{w}^L$. This term could now turn negative if skill formation is so highly subsidized that revenue losses outweigh the social-insurance gains of larger skill formation. Generally, however, the government would want to redistribute from high-skilled to low-skilled workers, implying a positive net tax on skill formation, such that $\Delta^S > 0$. This also seems to be the empirically relevant case as most industrial countries tax skill formation on a net basis (OECD, 2011*a*). In that case, and assuming that a minimum wage boosts skill formation, $(\tilde{N}^H/\tilde{w}^L > 0)$, a higher minimum wage would yield higher tax revenues. If $\Delta^S < 0$, skill formation is subsidized on a net basis and, provided that minimum wages boost skill formation, higher minimum wages result in additional revenue losses.

The remainder of the optimal minimum-wage expression is unaffected, compared to the case with skill-independent taxes and transfers.

Proposition 3.4 *Minimum wages are less likely to be socially desirable under skill-specific instruments if there is tax rate progression ($t^H > t^L$). If a minimum wage increases skill formation, it is more likely to be desirable if skill formation is taxed on a net basis ($\Delta^S < 0$).*

Optimal tax-rate progression

The first-order conditions for t^H and t^L in equations (3.58) and (3.59) are similar: the left-hand side gives the net social welfare gains of redistributing a unit of resources by raising the income tax rate. The marginal redistributive gains of high-skilled taxes are larger than the redistributive gains of low-skilled taxes if (*i*) the average marginal social value of income of high-skilled workers is lower than that of low-skilled workers – i.e., if $\overline{\Psi'(V^H)} < \overline{\Psi'(V^L)}$) – and (*ii*) the government is more concerned about within-group income-inequality in the group of high-skilled workers than in the group of low-skilled workers – i.e., if $\xi^H > \xi^L$. While these conditions depend on the specific social welfare function, they seem intuitively plausible, so that taxes on high-skilled workers should be set higher on the basis of redistributive reasons – not considering the efficiency costs.

The right-hand sides in equations (3.58) and (3.59) give the efficiency costs of using either tax instrument in terms of lower intensive labor supply (first term), higher unemployment (second term), and higher skill formation (third term). The formal structure of the first-order conditions is very similar, the implications for the optimal values of the tax rates are not. Although either tax instrument reduces intensive labor supply, as indicated by the first terms $\Delta^H\varepsilon$ and $\Delta^L\varepsilon$, the other elasticities have opposite signs in both equations. In particular, the number of high-skilled workers increases with a higher low-skilled tax rate, but it decreases with a higher high-skilled tax rate. Similarly, the unemployment rate decreases with higher low-skilled taxation, whereas it increases with higher high-skilled taxation. This means that low-skilled taxes alleviate the distortions associated with the minimum wage by reducing unemployment as they stimulate high-skilled labor supply and discourage low-skilled labor supply. High-skilled taxes, on the other hand, exacerbate distortions of the minimum wage by raising unemployment. Whether there should be tax rate progression is therefore theoretically ambiguous.

Proposition 3.5 *Distributional concerns tend to call for tax-rate progression. Tax-rate progression is less desirable if minimum wages are set higher, as tax rate progression exacerbates the labor-market distortions of the minimum wage by increasing low-skilled labor supply. The case for tax rate progression is further weakened (strengthened) if skill formation is taxed (subsidized) on a net basis, i.e., if* $\Delta^S > 0$ *(*$\Delta^S < 0$*).*

Optimal subsidy on skill formation

The first-order condition for S in equation (3.60) equates the marginal redistributive costs of directly distributing resources towards high-skilled workers, $\overline{\Psi'(V^H)}/\lambda - 1$, with the marginal welfare gains of lower unemployment and larger skill formation. The distributional gains of providing higher transfers to the high-skilled is negative as it redistributes resources in the wrong direction. Indeed, using the first-order condition for T we can derive that $\overline{\Psi'(V^H)}/\lambda < 1$. Thus, for redistributive reasons, the government would like to tax the high-skilled. However, subsidies on skill formation reduce unemployment, since $\tilde{u}/\mathrm{d}S = -\kappa\rho/S < 0$. Hence, subsidies on skill formation alleviate the distortions associated with the minimum wage. Moreover, subsidies on skill formation naturally boost skill formation as $\tilde{N}^H/\mathrm{d}S = \rho/(S\eta) > 0$. If skill formation is distorted downwards (upwards), such that $\Delta^S > 0$ ($\Delta^S < 0$), subsidizing the high-skilled reduces (exacerbates) the distortion on skill formation. Consequently, it remains unclear whether skill formation should be subsidized on a net basis.

Proposition 3.6 *Subsidies (taxes) on skill formation result in distributional losses (gains), alleviate the distortions created by the minimum wage, and alleviate (exacerbate) distortions of skill formation if skill formation is taxed (subsidized) on a net basis.*

Minimum wage versus income taxation: a reinterpretation

As a final exercise, we ask the question which instruments are more desirable for income redistribution: minimum wages or income taxes? As can be inferred from the first-order conditions (3.58), (3.59), and (3.57), a properly designed combination of high-skilled and low-skilled income taxes can exactly replicate the distributional effects of a minimum wage. Hence, the question whether minimum wages are desirable in addition to optimal income taxes boils down to the question: do minimum wages entail larger or smaller distortions than income taxes to achieve the same marginal distributional benefits?

To answer this question, we can combine the optimal-tax expressions for the tax rates and the minimum wage to derive a new desirability condition for the minimum wage. This is essentially equivalent to determining the welfare effects of an increase in the minimum wage while offsetting its distributional effects by an appropriate adjustment of the tax rates (i.e., a low-skilled tax increase, combined with a high-skilled tax decrease). This yields the following desirability condition for the minimum wage:

$$\Delta^S \left(\frac{\tilde{N}^H}{\tilde{u}} \right) > \Delta^U + \Delta^L, \tag{3.62}$$

where $\tilde{N}^H/\tilde{u}$ denotes the partial effect of unemployment on high-skill labor supply. By substituting for the wedges and the elasticity, we find:

$$\left(\frac{t^H z^H_\Theta - S - (1-u)t^L z^L_\Theta}{(1-t^L)L^L w^L} \right) \frac{\Theta g(\Theta)}{(1-\beta-\rho)\varepsilon} > \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{\lambda(1-t^L)L^L w^L} N^L + \frac{t^L}{1-t^L} \right). \tag{3.63}$$

The left-hand side gives the welfare gain from higher skill formation. The right-hand side gives the welfare losses due to higher unemployment. These are the net welfare effects of a higher minimum wage, when the distributional effects are offset by an appropriate adjustment in income taxes. The ratio $\frac{\Theta g(\Theta)}{(1-\beta-\rho S)\varepsilon}$ represents the elasticity of the number of high-skilled workers with respect to a change in the unemployment rate.

The right-hand side gives the dead-weight loss of a minimum-wage increase, over and above the costs of a tax change that features the same distributional benefits. The first term represents the utility loss of those low-skilled individuals that lose their jobs because of a higher minimum wage. The second term expresses the marginal welfare loss associated with lower tax revenue, caused by higher unemployment. These welfare losses can be avoided by using the income tax rather than the minimum wage to redistribute income. The left-hand side gives the marginal welfare gains from the increase in skill formation caused by the higher unemployment rate.[18] There can only be welfare gains of a minimum wage if skill formation is taxed on a net basis, such that $\Delta^S > 0$. Indeed, if $\Delta^S > 0$, the minimum wage alleviates the net distortion on skill formation, caused by redistributive taxation, by raising human capital investment through higher unemployment. However, if skill formation is subsidized on a net basis, such that $\Delta^S < 0$, a minimum wage can *never* be socially desirable. Indeed, relative to a distributionally equivalent reform of the income tax rates, a minimum wage increase then only entails a dead-weight loss by causing both higher unemployment and higher skill formation.

The expression for the optimal minimum wage is obtained simply by substituting an equality sign for the inequality sign. The only difference with respect to the desirability condition is that the wedge on skill formation Δ^S now also contains the insurance benefit associated with skill formation, $\frac{\Psi(V_\Theta^H)-(1-u)\Psi(V_\Theta^L)-u\Psi(V^U)}{\lambda(1-t^L)w^L L^L}$, which was nil for $u = 0$. This term has been extensively discussed before. Notice that this optimality condition only holds if a binding minimum wage is desirable to start with, such that inequality (3.62) holds for $u = 0$.

Summing up, allowing for skill-specific taxes has some important ramifications regarding the desirability of a minimum wage. To see this, we compare equation (3.45) with equation (3.63). By allowing for skill-dependent tax rates and transfers, the government can directly redistribute income both within and between high-skilled and low-skilled workers. Unlike in the case of skill-independent tax rates, there is no benefit of having minimum wages to correct for a distributional imperfection of the income-tax system in reducing between-group inequality. The tax-benefit system can achieve exactly the same redistributive impact of a minimum wage, but without the subsequent increase in unemployment. This explains why the redistributive terms, that are still present in equation (3.45), are absent

[18]Note that, relative to a distributionally equivalent tax reform, a minimum wage increase has no direct effect on the incentives to invest in human capital, i.e., the minimum wage's effect on the skill premium is equivalent to that of the tax reform.

from equation (3.63).

However, there is now a new term in equation (3.63), which is associated with the distortion on skill formation. With skill-independent tax instruments, human capital formation is not distorted by uniform taxes or transfers. When taxes and transfers are skill-dependent, the tax-benefit system is no longer neutral with respect to skill formation. Hence, while tax instruments can be more accurately targeted to distribute from high-skilled workers to low-skilled workers, doing so generates a net tax on skill formation. This distortion was absent in equation (3.45), but shows up as Δ^S in equation (3.63). Increasing the minimum wage, and simultaneously offsetting the distributional impact through the income tax system, boosts skill formation by raising unemployment. Thus, minimum wages help to alleviate the distortions of the tax-benefit system on skill formation. Recall that this only holds if investment in human capital is indeed taxed on a net basis.

Proposition 3.7 *If skill formation is taxed on a net basis, a marginal increase in the minimum wage, compared to a distributionally equivalent tax reform, entails a social welfare loss from higher unemployment and a social welfare gain from higher skill formation. The minimum-wage increase is desirable if the social welfare gain outweighs the loss. Minimum wages are more desirable if unemployment has a larger effect on skill formation, if skill formation is more heavily taxed, and if the utility and tax revenue losses associated with higher unemployment are lower. Minimum wages are never desirable if human capital formation is subsidized on a net basis.*

3.7 Efficient rationing

3.7.1 Efficient versus inefficient rationing

A binding minimum wage leads to an oversupply of labor and hence to rationing on the labor market. Up to now we only discussed uniform rationing on the extensive margin, according to which every low-skilled individual has the same chance of getting fired. We now also discuss efficient rationing. If individuals are heterogeneous with respect to the disutility of extensive labor supply, rationing is efficient if persons with the highest disutility of work lose their job first. If workers are heterogeneous with respect to disutility of intensive labor supply, as is the case in our model, rationing is efficient if it occurs on the intensive margin, i.e. by restricting the number of hours people work. Every worker equalizes the marginal disutility of work with the marginal utility of higher income. Therefore, a marginal decrease in working hours, forced upon workers by a binding minimum wage, only has second-order effects on individuals' utilities.

Theory provides little guidance when it comes to the efficiency of rationing. In absence of a secondary or "black" market, in which the rationed good is traded, there is little reason to assume the rationed goods are acquired by the individuals who desire them most (Tobin, 1952). Empirically, as noted by Luttmer (2007), this has been confirmed by studies of the U.S. residential market for gas (Davis and Kilian, 2011), the gasoline market (Deacon and Sonstelie, 1989; Frech and Lee, 1987) and on the housing rental market (Glaeser and Luttmer, 2003). As there is no secondary market for jobs or hours of work, it is unlikely that rationing due to a minimum wage is efficient. The only more or less direct evidence for the efficiency of lay-offs due to a minimum wage is due to Luttmer (2007). He measures the change in the average (proxy of the) reservation wage of low-skilled workers after an increase in the minimum wage. For two out of four proxies, he finds a statistically significant drop in reservation wages. This could have been interpreted as evidence that workers with the lowest utility surplus of work are rationed first, were it not that in the sensitivity analysis, he finds significant increases in two proxies. Hence, he does not find convincing evidence that the efficiency of the job allocation changed due to a change in the minimum wage. He does, however, find some evidence that a higher minimum wage leads to lower employment. This evidence supports our assumption of uniform rationing on the extensive margin, according to which low-skilled workers are laid off without affecting the ability composition of workers. The assumption of rationing through lay-offs is further supported by a large body of evidence (Neumark and Wascher, 2006).

There is much less evidence on whether there is rationing on the intensive margin, let alone on its efficiency, and the evidence that exists seems to be conflicting. For example, Zavodny (2000) finds that a minimum wage reduces employment, but *increases* average hours worked, while Couch and Wittenburg (2001) find that a minimum wage reduces both employment and hours worked.

3.7.2 Model and comparative statics

If rationing occurs exclusively on the intensive margin, there is no unemployment. The intensive labor-supply decision of high-skilled workers remains unaltered. However, in the case of a binding minimum wage, low-skilled workers face a restriction on the number of hours they are allowed to work. We refer to this restriction as *under*employment. We denote the effective labor supply – the actual number of hours worked – as l^e, and the maximum number of hours an individual with ability θ is allowed to work as $\bar{l}_\theta$. The size of the minimum wage determines the aggregate number of hours that firms can feasibly employ. How this aggregate hour restriction translates into individual hour rations, and how these rations depend on θ is *a priori* unclear. We assume that the rations are efficient and derive the implications of this for the specific functional form of $\bar{l}_\theta$ below.

Low-skilled workers maximize utility, $V_\theta^L = (1-t)w^L l^e - \frac{1}{\theta^\beta}\frac{(l^e)^{1+1/\varepsilon}}{1+1/\varepsilon}$, with respect to effective labor supply, l^e, subject to the rationing constraint, $\bar{l}_\theta \geq l^e$.[19] We denote the Kuhn-Tucker multiplier for this constraint as $(1-t)w^L\mu$. In equilibrium, μ gives the shadow price of relaxing the rationing constraint in terms of the net wage, $(1-t)w^L$. The Lagrangian for the maximization problem of the individual can thus be written as:

$$\mathcal{L} = (1-t)w^L l^e - \frac{1}{\theta^\beta}\frac{(l^e)^{1+1/\varepsilon}}{1+1/\varepsilon} + (1-t)w^L\mu(\bar{l}_\theta - l^e). \tag{3.64}$$

We denote the optimal effective labor supply for an individual with ability θ as l_θ^e. It is determined by the first-order condition of the Lagrangian with respect to l^e and by the rationing constraint:

$$l_\theta^e = (\theta^\beta(1-\lambda)(1-t)w^L)^\varepsilon \quad \text{if } \mu = 0, \tag{3.65}$$

$$l_\theta^e = \bar{l}_\theta \quad \text{if } \mu > 0. \tag{3.66}$$

In the absence of rationing, the constraint is slack, such that $\mu = 0$, and the solution for effective labor supply reduces to the one obtained in previous sections: $l_\theta^e = l_\theta^L = (\theta^\beta(1-t)w^L)^\varepsilon$. We call l_θ^L notional labor supply, i.e., the number of hours the worker would optimally like to supply. If the constraint is binding, such that $\mu > 0$, effective labor supply is fully determined by the rationing constraint and $l_\theta^e = \bar{l}_\theta$. Notice from the first-order condition that a minimum wage acts as an implicit tax on labor supply through raising the shadow price of labor supply μ.

Thus, an individual would like to work l_θ^L hours, but if rationed is forced to work $\bar{l}_\theta < l_\theta^L$ instead. Without loss of generality we denote the hours restriction as a proportion of notional labor supply such that $\bar{l}_\theta \equiv (1-u_\theta)l_\theta^L$. We call u_θ the ra-

[19] We restrict attention to skill-independent tax instruments. We do not formally analyze the case of skill-dependent tax instruments as the results would be trivial. The effects of a higher minimum wage can be shown to be exactly mimicked by an increase in the low-skilled tax rate and a decrease in the high-skilled tax rate, leaving the minimum wage redundant (on this, see also Lee and Saez, 2012).

tioning schedule which may or may not depend on θ. It is important to distinguish u_θ from the unemployment rate as we have previously defined it. While in earlier sections u stands for the proportion of low-skilled individuals that are unemployed, in this section u_θ stands for the proportion of hours that are underemployed.

To determine individual labor supply we need to know the specific functional form of the rationing schedule. As discussed above, it is empirically unclear how rationing should depend on θ, but we assume in this section that the rationing schedule is efficient. This implies that the functional form of u_θ is such that the marginal utility of an extra hour of work is equal for every unskilled worker. Had this not been the case it would be efficiency improving to marginally decrease rationing of the high marginal utility worker and increase rationing of the low marginal utility worker. The marginal utility of being allowed to work an extra hour of work, in terms of the net wage, is given by the shadow price of labor supply μ. Substituting for $l^e_\theta = (1-u_\theta)(\theta^\beta(1-t)w^L)^\varepsilon$ in the first-order condition, we can write the shadow price as $\mu = 1-(1-u_\theta)^{\frac{1}{\varepsilon}}$. For rationing to be efficient, the shadow price should be independent of θ, and thus we require that:

$$\frac{\mathrm{d}\mu}{\mathrm{d}\theta} = \frac{1}{\varepsilon}(1-u_\theta)^{\frac{1}{\varepsilon}-1}\frac{\mathrm{d}u_\theta}{\mathrm{d}\theta} = 0. \tag{3.67}$$

This equation tells us that for rationing to be efficient, we necessarily have that $\mathrm{d}u_\theta/\mathrm{d}\theta = 0$. Hence, efficient rationing requires that the ration, as a proportion of the notionally supplied number of hours, is equal for every low-skilled worker. The crucial assumption underlying this result is that the compensated elasticity of labor supply, ε, is identical for every low-skilled worker. This assumption implies that substitution effects and thus dead-weight losses are identical for every worker facing the same ration u_θ. Throughout the remainder of this section we are exclusively interested in efficient rationing schedules and thus write $u_\theta = u$.

Substituting for effective labor supply, l^e_θ, we can write the indirect utility function for low-skilled workers as:

$$V^L_\theta = T + \left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)(1-u)\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon}. \tag{3.68}$$

Notice that, for $u = 0$, this collapses to the low-skilled utility in the case of extensive rationing. Furthermore, it can easily be shown that, in the absence of rationing, $\partial V^L_\theta/\partial u = 0$, such that a marginal increase in rationing does not affect low-skilled utility. However, for positive and increasing values of rationing,

$\partial V_\theta^L/\partial u$ is negative and decreasing.[20] As a direct consequence, in the absence of rationing, an increase in rationing only has a second-order effect on the cut-off ability level, Θ. Hence, a marginal increase in the minimum wage above the market-clearing wage only affects the human capital decision through a decrease in the skill premium, w^H/w^L. Only for higher levels of the minimum wage, further rationing causes an offsetting response in human capital. See the appendix for a full derivation of the comparative statics.

The comparative-statics equation for skill formation is now given by:

$$\tilde{N}^H = \left(\frac{\sigma-1}{\sigma} - (1-u)^{\frac{1}{\varepsilon}}\right)\frac{\sigma}{\eta'}\frac{\tilde{w}^L}{\alpha}, \tag{3.69}$$

where $\eta' > 0$ is a composite term describing the shape of the income distribution around ability Θ (see appendix). A higher minimum wage will lead to more skill formation if and only if $\frac{\sigma-1}{\sigma} > (1-u)^{\frac{1}{\varepsilon}}$. This is more likely to hold if σ and u are large and if ε is small.

Table 3.1 shows the critical levels of u, above which a higher minimum wage leads to higher skill formation and below which it leads to less skill formation. We show this for values of σ between 1.5 and 2.5. A value of 1.5 seems to be reasonable, although both lower and higher values are found in the literature (Katz and David, 1999). Notice that for $\sigma \leq 1$, a minimum wage always leads to lower skill formation. Furthermore, we choose values of ε between 0.2 and 0.4, which seems to be a reasonable range (see, e.g., Blundell and MaCurdy, 1999).

The critical values of the underemployment rate lie between 0.10 (for $\varepsilon = 0.2$ and $\sigma = 2.5$) and 0.36 (for $\varepsilon = 0.4$ and $\sigma = 1.5$). Empirical evidence on the degree of working-hour restrictions varies widely. There are studies observing employees working less hours than desired (e.g., Kahn and Lang, 1991; Dickens and Lundberg, 1993; Bloemen, 2008) and studies observing employees actually working more hours than they desire (e.g., Stewart and Swaffield, 1997; Böheim and Taylor, 2004). The largest

Table 3.1: Critical values for u

$u^* = 1 - \left(\frac{\sigma-1}{\sigma}\right)^\varepsilon$			
	$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 2.5$
$\varepsilon = 0.2$	0.20	0.13	0.10
$\varepsilon = 0.3$	0.28	0.19	0.14
$\varepsilon = 0.4$	0.36	0.24	0.18

[20] The first derivative is given by $\partial V_\theta^L/\partial u = \left((1-u)^{\frac{1}{\varepsilon}} - 1\right)\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon} < 0$, and the second derivative by $\partial^2 V_\theta^L/\partial u^2 = -\frac{1}{\varepsilon}(1-u)^{\frac{1-\varepsilon}{\varepsilon}}\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon} < 0$.

rationing proportion, which is based on a sample of low-income workers, is found by Dickens and Lundberg (1993) and is with 20 percent well within the range of above table. However, under more conservative underemployment rates of around 10 percent (as in Kahn and Lang, 1991), it is very unlikely that, with efficient rationing, a minimum wage leads to more skill formation.

3.7.3 Optimal policy

As there are no unemployed when rationing occurs on the intensive margin, the social welfare function simplifies to:

$$\mathcal{W} \equiv \int_{\underline{\theta}}^{\Theta} \Psi(V_\theta^L)\mathrm{d}G(\theta) + \int_{\Theta}^{\bar{\theta}} \Psi(V_\theta^H)\mathrm{d}G(\theta), \tag{3.70}$$

whereas the government' budget constraint is still given by equation (3.30). Forming the Lagrangian and taking derivatives we find that the first-order conditions for the tax rate and transfer do not change. We find the following first-order condition for the optimal minimum wage:

$$\begin{aligned} &(1-\xi^L)\overline{\Psi'(V^L)} - (1-\xi^H)\overline{\Psi'(V^H)} \\ &- \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)(1-\xi^L)\overline{\Psi'(V^L)}\left(\frac{w^L}{(1-u)}\frac{\mathrm{d}u}{\mathrm{d}w^L} - \varepsilon\right) \\ &- \lambda\frac{t}{1-t}\left(\frac{w^L}{1-u}\frac{\mathrm{d}u}{\mathrm{d}w^L} - \frac{(w^H l_\Theta^H - w^L(1-u)l_\Theta^L)}{L^L}\frac{\mathrm{d}N^H}{\mathrm{d}w^L}\right) = 0. \end{aligned} \tag{3.71}$$

The first line gives the redistributional gain of an increase in the minimum wage, which is the same as before. The second line gives the utility loss associated with more rationing due to a higher minimum wage. The third line gives the social welfare loss of an eroding tax base.

To focus on the desirability of a minimum wage, we analyze the first-order condition for $u = 0$. Note that in that case $\left(1-(1-u)^{\frac{1}{\varepsilon}}\right) = 0$ and $w^H l_\Theta^H - w^L(1-u)l_\Theta^L = 0$. Substituting this into the first-order condition, we obtain the following condition for a minimum wage to be desirable:

$$(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} - (1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} > \frac{t}{1-t}\frac{\tilde{u}}{\tilde{w}^L}. \tag{3.72}$$

This expression is almost identical to the analogue expressions in previous sections. The only term that is missing is the marginal utility loss from rationing which, as we discussed above, is only second-order under efficient rationing. Hence, in the absence of taxation, a marginal increase in the minimum wage above the market-clearing wage only has distributional gains, equal to the left-hand side of equation (3.72). With efficient rationing, and in the absence of taxation, a minimum wage is therefore unambiguously desirable. This confirms Lee and Saez (2012) who derive the same result in the case of efficient rationing on the extensive margin.

If there is a positive tax rate, a minimum wage erodes the tax base as more rationing leads to fewer workers paying taxes. This welfare loss is represented by the right-hand side of equation (3.72) and is increasing in the tax rate. In order to determine the desirability of a minimum wage at the optimal tax system we substitute for $\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}$ from the first-order condition for the tax rate, $t/(1-t) = \xi/\varepsilon$. This yields:

$$\frac{t}{1-t}\left(\frac{\varepsilon}{\alpha - N^H} - \frac{\tilde{u}}{\tilde{w}^L}\right) > \frac{N^L}{\alpha - N^H}\frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \frac{N^H}{\alpha - N^H}\frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H. \tag{3.73}$$

This condition is almost identical to equation (3.45), the only differences being the utility loss of unemployment, which drops out, and a slightly altered elasticity of underemployment $\tilde{u}/\tilde{w}_L = (\sigma + \varepsilon + \kappa')/\alpha$ (see appendix). The left-hand side gives the efficiency costs of redistributing between skill-groups by using income taxes instead of a minimum wage. In the case of income taxes, the efficiency costs consist of the tax-base erosion caused by downwardly distorted intensive labor supply of both high-skilled and low-skilled workers. In the case of a higher minimum wage, the efficiency costs consist of the tax-base erosion associated with underemployment. The right-hand side gives the within-group distributional advantage income taxes have over the minimum wage.

Proposition 3.8 *Similarly to the case with uniform unemployment on the extensive margin, if rationing is efficient, the role of a minimum wage in an optimal skill-independent tax-benefit system is to complement the tax-benefit system by reducing the distributional imperfections of the income tax. Minimum wages help to redistribute more income between skill groups, so that income taxes can be better targeted at reducing inequality within skill groups. The desirability of a minimum wage depends on whether this benefit outweighs the loss in tax revenue due to higher rationing. Contrary to the case with uniform unemployment on the extensive margin, there is no direct utility loss associated with a marginally binding minimum wage.*

3.8 Conclusion

This study indicates that the role and desirability of a minimum wage depends on the available tax instruments. If taxation cannot be conditioned on skill type or wage rate, a minimum wage might be useful as a means to directly redistribute income from high- to low-skilled workers. If taxation can be conditioned on skill type, such redistribution can also be achieved by appropriately setting taxes. The redistributive role of a minimum wage in that case vanishes. However, a binding minimum wage might still be useful to alleviate a distortion on skill formation caused by redistributive taxation. While a minimum wage exacerbates a tax distortion by raising unemployment, it simultaneously alleviates a distortion by raising skill formation. The net welfare effect determines the desirability of a minimum wage. These results bring to mind the discussion on optimal indirect taxation. In the absence of non-linear taxation, indirect taxation might be useful for redistribution (e.g., Atkinson and Stiglitz, 1980). If non-linear taxation is available (and with homogeneous preferences), indirect taxation loses its redistributive role. It might still help, however, in alleviating distortions caused by the non-linear tax schedule (e.g., Atkinson and Stiglitz, 1976).

An important factor determining the optimality of a minimum wage is the degree to which education increases as a result of higher unemployment among the low-skilled. This is in turn driven by how strongly the job chances of low-skilled workers on the skill margin are affected by unemployment. We assumed that every low-skilled worker's job chances are affected equally. Assuming, as Lee and Saez (2012) mostly do, that workers on the skill margin are hit by unemployment first would drastically improve the case for a minimum wage as it would lead to a larger increase in skill formation. The welfare consequences of different rationing schedules are discussed in more detail in Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation) and Gerritsen (2013*a*, Chapter 5 of this Dissertation). In Gerritsen and Jacobs (2013*a*), we study the desirability of a minimum wage in the presence of a more general rationing schedule and empirically calibrate the resulting desirability condition. Gerritsen (2013*a*) treats the wage floor as given, and derives implications for optimal tax policy.

3.A Efficient rationing

In the case of efficient rationing, indirect utility is represented by:

$$V_\theta^L = T + \left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)(1-u)\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon}, \tag{3.74}$$

$$V_\theta^H = T + \frac{\theta^\varepsilon((1-t)w^H)^{1+\varepsilon}}{1+\varepsilon}. \tag{3.75}$$

For the cut-off level of ability, Θ, we need $V_\Theta^L = V_\Theta^H$, which implies:

$$\Theta^{(1-\beta)\varepsilon} = \left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)(1-u)(1+\varepsilon)\left(\frac{w^H}{w^L}\right)^{-(1+\varepsilon)}. \tag{3.76}$$

The rest of the equilibrium conditions consist of the firms' first-order conditions and the market clearing conditions.

$$F_H(L^H, L^L) = w^H, \tag{3.77}$$

$$F_L(L^H, L^L) = w^L, \tag{3.78}$$

$$L^H = \int_\Theta^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta), \tag{3.79}$$

$$L^L = (1-u)\int_{\underline{\theta}}^{\Theta} l_\theta^L \mathrm{d}G(\theta). \tag{3.80}$$

Log-linearizing the equilibrium conditions around an initial equilibrium yields the following equations:

$$(1-\beta)\varepsilon\tilde{\Theta} = (1+\varepsilon)(\tilde{w}^L - \tilde{w}^H) - \frac{1-(1-u)^{\frac{1}{\varepsilon}}}{1-\frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}}\tilde{u}, \tag{3.81}$$

$$\tilde{w}^H = \frac{1-\alpha}{\sigma}(\tilde{L}^L - \tilde{L}^H), \tag{3.82}$$

$$\tilde{w}^L = \frac{\alpha}{\sigma}(\tilde{L}^H - \tilde{L}^L), \tag{3.83}$$

$$\tilde{L}^H = -\frac{l_\Theta^H \Theta g(\Theta)}{L^H}\tilde{\Theta} + \varepsilon(\tilde{w}^H - \tilde{t}), \tag{3.84}$$

$$\tilde{L}^L = \frac{(1-u)l_\Theta^L \Theta g(\Theta)}{L^L}\tilde{\Theta} - \tilde{u} + \varepsilon(\tilde{w}^L - \tilde{t}). \tag{3.85}$$

Moreover recall that $N^H = 1 - G(\Theta)$ and hence $\tilde{N}^H = -\Theta g(\Theta)\tilde{\Theta}$. The equations above can now be solved to express $\tilde{N}^H$ and $\tilde{u}$ in terms of the exogenous variables, $\tilde{w}^L$ and $\tilde{t}$:

$$\tilde{N}^H = \left(\frac{\sigma-1}{\sigma} - (1-u)^{\frac{1}{\varepsilon}}\right)\frac{\sigma}{\eta'}\frac{\tilde{w}^L}{\alpha}, \tag{3.86}$$

$$\tilde{u} = \left(\sigma + \varepsilon - \left(\frac{\sigma-1}{\sigma} - (1-u)^{\frac{1}{\varepsilon}}\right)\sigma\kappa'\right)\frac{\tilde{w}^L}{\alpha}, \tag{3.87}$$

where $\eta' = \left(\left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)\frac{(1-\beta)\varepsilon}{\Theta g(\Theta)} + \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)\left(\frac{l^H_\Theta}{L^H} + \frac{(1-u)l^L_\Theta}{L^L}\right)\right)$ and $\kappa' = \left(\frac{l^H_\Theta}{L^H} + \frac{(1-u)l^L_\Theta}{L^L}\right)\eta'^{-1}$. The interpretation of these two comparative-statics equations is similar to the case with extensive rationing. In particular, it shows that a higher minimum wage leads to more skill formation if and only if $(\sigma-1)/\sigma > (1-u)^{\frac{1}{\varepsilon}}$. Furthermore, it can be shown that rationing always increases due to a higher minimum wage.

The Lagrangian for the government's optimization problem is the following:

$$\mathcal{L}(w^L, t, T) = \int_{\underline{\theta}}^{\Theta} \Psi(V^L_\theta)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi(V^H_\theta)\mathrm{d}G(\theta) + \lambda(tw^L L^L + tw^H L^H - T - E). \tag{3.88}$$

This leads to the following first-order conditions:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial w^L} = {}& (1-t)\left(\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)(1-u)l^L_\theta \mathrm{d}G(\theta) - \frac{L^L}{L^H}\int_{\Theta}^{\overline{\theta}} \Psi'(V^H_\theta)l^H_\theta \mathrm{d}G(\theta)\right)\\ &- (1-t)\left(1-(1-u)^{\frac{1}{\varepsilon}}\right)\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)(1-u)l^L_\theta \mathrm{d}G(\theta)\left(\frac{\tilde{u}}{\tilde{w}^L} - \varepsilon\right)\\ &- \lambda t L^L\frac{\tilde{u}}{\tilde{w}^L} + \lambda(tw^H l^H_\Theta - tw^L(1-u)l^L_\Theta)\frac{\mathrm{d}N^H}{\mathrm{d}w^L} = 0,\end{aligned} \tag{3.89}$$

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial t} = {}& -\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)w^L(1-u)l^L_\theta \mathrm{d}G(\theta) - \int_{\Theta}^{\overline{\theta}} \Psi'(V^H_\theta)w^H l^H_\theta \mathrm{d}G(\theta)\\ &- \varepsilon\left(1-(1-u)^{\frac{1}{\varepsilon}}\right)\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)w^L(1-u)l^L_\theta \mathrm{d}G(\theta)\\ &+ \lambda\left(w^L L^L + w^H L^H - \varepsilon\frac{t}{1-t}(w^L L^L + w^H L^H)\right) = 0,\end{aligned} \tag{3.90}$$

$$\frac{\partial\mathcal{L}}{\partial T} = \int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V^H_\theta)\mathrm{d}G(\theta) - \lambda = 0. \tag{3.91}$$

The first-order condition for the minimum wage can be simplified as follows:

$$(3.92) \qquad (1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} - (1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} = \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda}\left(\frac{\tilde{u}}{\tilde{w}^L}-\varepsilon\right) + \frac{t}{1-t}\frac{\tilde{u}}{\tilde{w}^L} - \frac{t}{1-t}\frac{w^H l_\Theta^H - w^L(1-u)l_\Theta^L}{w^L L^L}\frac{\tilde{N}^H}{\tilde{w}^L}.$$

The left-hand side gives the distributional gain of a higher minimum wage. The right-hand side gives the costs associated with higher unemployment and the inability of low-skilled workers to react by altering their hours worked (first term) and the costs associated with exacerbating the tax distortion due to stronger rationing (second term) and lower high-skilled labor supply (third term). In the case of $u = 0$, the first and third terms only imply second-order welfare effects.

Rearranging, and substituting for ξ and ξ^L, yields the following expression for the optimal income tax:

$$(3.93) \qquad \frac{t}{1-t} = \frac{\xi}{\varepsilon} - \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)(1-\alpha)(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda}.$$

Hence, compared to the case with rationing along the extensive margin, there is an additional cost of taxation, which lowers the optimal tax level. In the case of extensive rationing, low-skilled workers coped with higher taxation by working less hours, thereby absorbing part of the direct utility costs. However, if these workers are intensively rationed, they will not reduce their working hours as they already work less than they would prefer. In the case of $u = 0$ this cost naturally vanishes.

Chapter 4

Is a minimum wage an appropriate instrument for redistribution?

4.1 Introduction

The public debate on increasing, decreasing, or implementing a minimum wage is not unlike the mythological hydra of Lerna. Not long after the debate is temporarily settled in one country it rears its head in another.[1] This regularity is suggestive of the lack of consensus on the desirability of a minimum wage. Proponents emphasize its positive effects on the earnings of low-skilled workers; opponents stress that a minimum wage tends to increase unemployment. As long as the debate is framed in terms of the trade-off between these two opposing welfare effects, this chapter's title question is bound to remain contentious even among economists that are not principally opposed to redistribution. This is why the merits and demerits of a minimum wage should not be judged in isolation, but in contrast to the merits and demerits of obtaining a similar degree of redistribution in a direct way by means of the income-tax system. A minimum wage can only be desirable if it is more effective than income taxation in supporting the earnings of low-skilled workers.

The central purpose of this chapter is to shed light on the question whether minimum wages are optimally employed alongside income taxes in order to redistribute income. To that end, we develop a model with perfectly competitive labor

[1] In the same analogy, the modest ambition of this chapter is to take up the role of Hercules.

markets, in which firms demand high- and low-skilled labor. Recent evidence suggests that minimum wages have important general-equilibrium effects on the entire wage structure (e.g., Teulings, 2000, 2003; Autor, Manning and Smith, 2010). Our analysis allows for such effects by taking into account multiple job types that are imperfect substitutes in production. For expositional reasons, we focus on the case of two job types (high- and low-skilled) but can easily generalize this to $n > 2$ job types with varying degrees of complementarity. Individuals decide on hours worked, participation, and education, that is, on being high- or low-skilled. Individuals are heterogeneous in a single dimension: their disutility of work, which we refer to as their ability. An individual's ability determines whether he optimally decides to be voluntarily unemployed, a low-skilled worker, or a high-skilled worker. A minimum wage, if it binds, only does so for the low-skilled segment of the labor market, thereby creating involuntary unemployment among low-skilled workers.

Our first contribution is to determine whether and when a binding minimum wage is part of an optimal redistributive policy. To that end, we consider a net-income-neutral (NIN) increase in the minimum wage. As the name suggests, the NIN minimum-wage reform raises the minimum wage, while keeping workers' net incomes constant by perfectly offsetting tax changes. Complementarity between job types implies that a minimum-wage increase affects the gross wage rates of all job types. We assume that taxation can be conditioned on job type or, equivalently, on the wage rate, the observability of which is required to enforce a minimum wage. This enables the government to neutralize the effects of a minimum-wage increase on the net wages of both low- and high-skilled workers.[2] As net incomes remain constant, the reform has no direct effects on labor supply. As it raises firms' wage costs for low-skilled workers, it does reduce low-skilled labor demand. This reduced demand compresses the low-skilled labor market in two distinct ways. On the one hand, low-skilled workers are pushed 'downwards' into unemployment. On the other hand, the increased probability of unemployment induces low-skilled workers to invest in education and move 'upwards' into the high-skilled segment of the labor market.

The effects of a NIN minimum-wage increase are thus twofold: it raises both

[2]This is similar in spirit to Kaplow (2008), who analyzes policy reforms (e.g., larger public good provision, changes in indirect taxes) by adjusting the tax schedule to keep individuals' utilities fixed. We analyze a minimum wage reform while adjusting the tax system to keep the net incomes constant of those who retain their jobs. Naturally, the optimality of a minimum wage does not depend on the particular tax reform chosen.

involuntary unemployment and education. The unemployment effect causes two separate welfare losses: laid-off workers who strictly prefer employment over unemployment suffer direct utility losses, and tax revenue declines if the government employs progressive taxation to redistribute from the employed to the unemployed. The education effect results in higher tax revenues and thus a welfare gain if high-skilled workers are taxed more heavily than low-skilled workers. A NIN minimum-wage increase is desirable if the welfare gains from more education outweigh the welfare losses from higher unemployment. This is true regardless of the policy mix in place. However, if this condition holds in the tax optimum without a minimum wage, a minimum wage must be part of the optimal policy mix and therefore be an appropriate instrument for redistribution. An alternative, but equivalent, way to state this is the following. If the optimal tax system is progressive and taxes education and participation on a net basis, the optimal tax system distorts both education and participation decisions downward. A NIN minimum-wage reform exacerbates the distortion on participation by increasing unemployment, but alleviates the distortion on education as low-skilled workers seek to avoid involuntary unemployment. The net welfare effect determines a minimum wage's desirability.

Our second contribution is to demonstrate that the net welfare effect of a minimum wage critically depends on the specific assumptions we make on exactly which low-skilled workers become unemployed – given heterogeneity among low-skilled workers. We propose a general *rationing schedule* that specifies the probability of unemployment at each ability level. We furthermore define the *unemployment incidence* at a specific ability level, as the relative degree to which higher unemployment is concentrated on this ability level. Therefore, a crucial difference with previous studies is that we do not need to make specific assumptions on how unemployment depends on the source of heterogeneity.

To illustrate what we mean by unemployment incidence, consider Figure 4.1. The horizontal line represents the range of ability types, $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$. In our model's equilibrium, there are two critical ability levels, Φ and Θ. For individuals with low ability, $\theta \in [\underline{\theta}, \Phi)$, the disutility of work is so large that it is optimal for them to be voluntarily unemployed. Individuals with intermediate ability, $\theta \in [\Phi, \Theta)$, supply labor as low-skilled workers. Individuals with high ability, $\theta \in \left[\Theta, \overline{\theta}\right]$, work as high-skilled workers. Now consider an increase in low-skilled unemployment due to a NIN minimum-wage increase. Panels (a)-(c) in Figure 4.1 depict, for three different types of unemployment incidence, the change in individual unemployment probabilities as a function of ability. In principle, there is an infinite number of possible assumptions on the exact unemployment incidence of a given increase in aggregate unemployment.

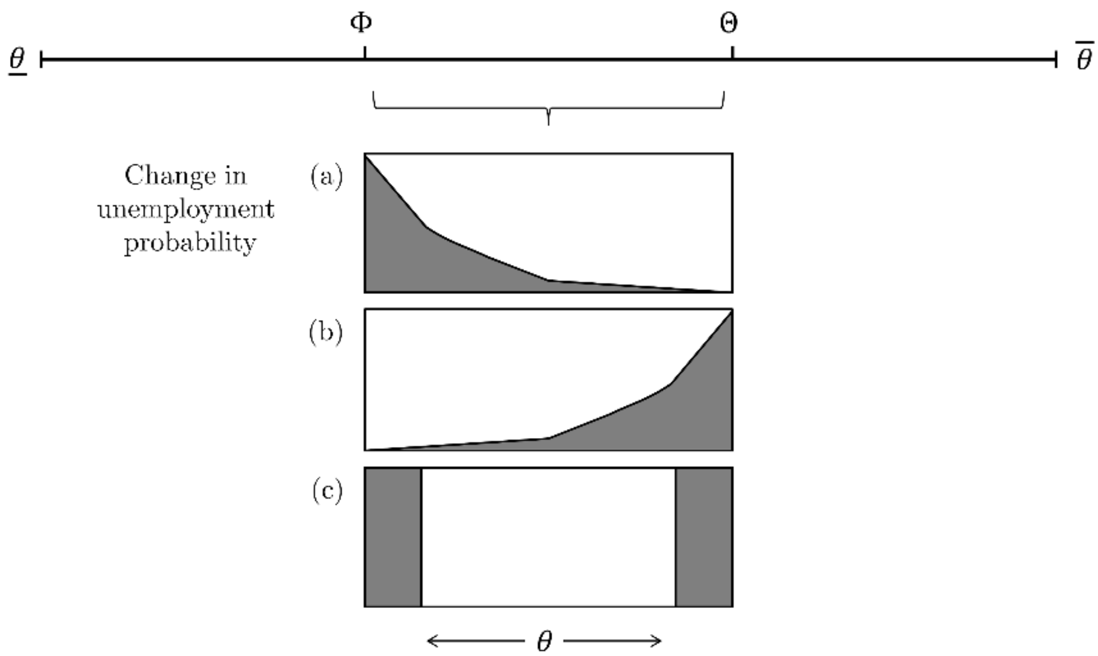

Figure 4.1: The unemployment incidence of a minimum wage increase: examples

Earlier literature recognizes that the unemployment incidence is important to assess the welfare effects of a minimum wage as it determines the utility losses of laid-off workers (Lott, 1990; Palda, 2000; Lee and Saez, 2012). For example, Lee and Saez (2012) assume that the unemployment incidence is as depicted in panel (c) of Figure 4.1. They dub this case 'efficient rationing' as unemployment is concentrated on those individuals that are indifferent between low-skilled work and their outside option – be it non-participation or high-skilled employment. Panels (a) and (b) depict cases of 'inefficient rationing' as part of the unemployment incidence falls on individuals that are on neither the participation margin, nor the skill margin, and therefore strictly prefer low-skilled employment over their outside option. We show that the unemployment incidence of the minimum wage is critical in two important ways: it determines not only the direct utility losses of the unemployed, but also the education response to a minimum-wage increase.

To see this, consider the first two panels. In panel (a), the lion share of the unemployment incidence falls on low-skilled workers who are close to the participation margin Φ and thus relatively indifferent between having a job or being non-participant. As a consequence, the education effect of higher unemployment will be relatively small. In panel (b), most of the unemployment incidence falls on low-skilled workers close to the skill margin Θ. Consequently, many low-skilled workers will decide to become high-skilled in order to avoid the increased probability of unemployment. For the same increase in aggregate unemployment, the case depicted in panel (b) leads to a larger increase in high-skilled employment than the case depicted in panel (a). A minimum wage is thus more likely to be desirable in case (b) than in case (a). More generally, the optimality of a minimum wage fundamentally depends on the incidence of unemployment. We prove that a minimum wage can always be 'made' optimal by making the appropriate assumptions on the unemployment incidence. Unfortunately, both theoretically and empirically, it is unclear how unemployment probabilities depend on workers' ability or disutility of work. Any specific assumption on the incidence of unemployment implies a certain degree of arbitrariness.

This brings us to the third contribution of this chapter, which is to reinterpret the desirability condition of a NIN minimum-wage increase in terms of three sufficient statistics. These sufficient statistics allow us to forego the need to determine deeper model parameters, in particular the unemployment incidence. We assume that the low-skilled workers in our model correspond to actual workers that did not complete upper secondary education. The three sufficient statistics are then given by the social costs of low-skilled unemployment, the social gains of more upper-secondary education, and the effect of low-skilled unemployment on upper-secondary enrollment or graduation rates. The only element of the sufficient statistics that cannot easily be measured are the direct utility losses of laid-off workers. In our calibration we ignore these direct utility losses, thereby biasing results in favor of a minimum-wage increase. Data on net tax-revenue losses of unemployment and gains of education can be found for a large number of member countries of the Organisation for Economic Co-operation and Development (OECD). Moreover, there is a sizable empirical literature on the effect of unemployment on school enrollment. We make conservative assumptions when calibrating the desirability condition, further biasing our findings in favor of a minimum wage. Our results indicate that a NIN minimum-wage *decrease* would result in a Pareto improvement for almost all the countries under consideration. That is, it would increase government revenue and enable some involuntarily unemployed to find a job, while leaving no one worse off due to the net-income neutrality of the reform. Only for the United States, using a highly conservative calibration, we cannot reject the desirability of a minimum-wage increase outright.

The remainder of this chapter is structured as follows. We discuss the relation to previous literature in Section 4.2. Section 4.3 introduces the theoretical model. The comparative statics of a NIN minimum-wage reform are developed in Section 4.4, while Section 4.5 provides the welfare analysis of the minimum-wage reform. Section 4.7 discusses the empirical application of the model and we conclude with some final thoughts.

4.2 Earlier literature

This chapter contributes to a small strand of the literature, which analyzes minimum-wage legislation in models of competitive labor markets.[3] The classic references are Guesnerie and Roberts (1987) and Allen (1987), who adopt versions of the Stiglitz (1982) model in which workers are predetermined to be high-skilled or low-skilled. Apart from their skill type, individuals are identical so there is no need to specify a specific rationing schedule. The government sets a minimum wage which might bind for the low-skilled segment of the labor market. It also uses non-linear income taxes to optimally redistribute from high- to low-skilled workers, but cannot directly observe a worker's skill type. The goverment is thus restricted by an incentive-compatibility constraint stipulating that high-skilled workers prefer not to mimic low-skilled workers. Since this is the only constraint preventing full income redistribution, a minimum wage can never be optimal as it tightens the incentive-compatibility constraint by making it more attractive for a high-skilled worker to mimic a low-skilled worker.

There are a number of drawbacks to this approach. First, workers cannot choose their skill type. For high-skilled workers it is impossible to obtain a low-skilled job; for low-skilled workers it is impossible to educate themselves to obtain a high-skilled job. A second drawback is that, in order to enforce a minimum wage, the government requires information on individual wage rates and, hence, on skill types. The very absence of such information, however, is assumed to be the reason for the government to resort to distortionary taxation in the first place.[4]

[3]Note that there is also a larger literature on minimum wages in non-competitive labor markets. In most of that literature, the effects of a minimum wage are studied in isolation from income tax policy. See, for example, Gerritsen and Jacobs (2013*b*, Chapter 3 of this Dissertation) and Lee and Saez (2012) for overviews and extensive references. A notable exception is Hungerbühler and Lehmann (2009), who find a role for a minimum wage alongside optimal non-linear labor income taxes if workers' bargaining power is inefficiently low and the government has no means to directly control bargaining power. Our results should therefore be interpreted with caution, keeping in mind that labor-market frictions might have a separate impact on the desirability of a minimum wage.

[4]Guesnerie and Roberts (1987, p.498) are perfectly aware of this, stating that the assumption

This approach thus suffers from an *informational inconsistency* in the optimal redistribution problem.

These two drawbacks are avoided both by this chapter and by Lee and Saez (2012). Lee and Saez, like us, assume individuals' decisions on participation and skill type are driven by their idiosyncratic disutility of work. Taxes are conditioned on skill type but the first-best allocation is unattainable because workers can decide on their skill type. As discussed above, Lee and Saez furthermore make the very specific assumption that unemployment is concentrated on those low-skilled workers with the lowest willingness to pay to remain a low-skilled worker. Under this assumption, workers that are indifferent between non-participation and low-skilled employment, or between low-skilled and high-skilled employment, are the first workers to be rationed, as illustrated by panel (c) of Figure 4.1. A minimum wage therefore enables the government to provide additional transfers to low-skilled workers without causing any distortion; anyone (be it a non-participant or a high-skilled worker) who tries to obtain a low-skilled job in response to these transfers faces certain unemployment, and thus no one tries. Consequently, a minimum wage is optimal if, in the optimum without a minimum wage, the government would like to redistribute more income towards low-skilled workers (away from high-skilled workers and non-participants), but is prevented from directly transferring more income towards the low-skilled because the distortions of doing so would become larger than the distributional gains.

The optimality condition of Lee and Saez crucially depends on their assumption on how the probability of unemployment is related to unobserved disutility of work. Since it is virtually impossible to empirically assess the plausibility of this assumption it remains unclear whether their results are of practical relevance. In this chapter we attempt to provide a solution to this concern by developing a model in which individuals decide on participation and skill type, without imposing restrictions on how the probability of unemployment depends on a low-skilled worker's disutility of work. We derive an optimality condition for a minimum wage that holds regardless of the incidence of unemployment.[5] In Section 4.5 we provide a proof that this condition harbors the result of Lee and Saez as a special case. We furthermore show that the effect of unemployment on education can

that the government cannot observe wages but can enforce minimum wages "is a somewhat mixed observability assumption."

[5]In a related study, Gerritsen (2013*a*, Chapter 5 of this Dissertation) determines how optimal tax rules depend on the unemployment incidence when the government does not directly set the wage floor.

function as a sufficient statistic, foregoing the need to make a specific assumption on the unemployment incidence. Thus, in contrast to earlier literature, our model enables us to derive a condition for the desirability of a minimum wage, based on sufficient statistics that can be determined empirically, without depending on a pre-determined skill distribution, informational inconsistencies, or very specific assumptions on the unemployment incidence of a minimum wage.

4.3 Model

This section describes in detail the behavior of individuals and firms. Individuals choose (i) whether to participate in the labor market as a high-skilled worker, participate as a low-skilled worker, or not to participate at all, and (ii) conditional on being high-skilled or low-skilled, how many labor hours to supply. Firms demand two types of labor, low-skilled and high-skilled, in competitive labor markets. In case of a binding minimum wage, low-skilled labor demand will be insufficient to match supply, causing involuntary unemployment among the low-skilled.

4.3.1 Preferences and budget constraints

Individuals differ in their ability θ, which has cumulative distribution $G(\theta)$, density $g(\theta)$, and support $\left[\underline{\theta}, \overline{\theta}\right]$, with $\overline{\theta} > \underline{\theta} \geq 0$. While in many studies of optimal redistribution ability determines a person's earning capacity, in our model it affects his disutility of work.[6] Specifically, take a worker with ability θ and skill type $i \in \{H, L\}$, where H stands for high-skilled and L for low-skilled work.[7] This worker's number of labor hours is denoted by l_θ^i. His disutility of work, denoted q_θ^i, is assumed to be a skill-specific function of labor hours and ability: $q_\theta^i \equiv q^i(l_\theta^i, \theta)$.

Disutility of work is increasing and convex in hours worked ($q_{\theta,l}^i, q_{\theta,ll}^i > 0$) and decreasing in ability ($q_{\theta,\theta}^i < 0$).[8] Disutility of work could simply represent

[6] We assumed heterogeneity in disutility of work, rather than earning capacity for two reasons. First, the ability to implement and enforce a minimum wage requires that the government has information on individual wages. If individuals are only heterogeneous with respect to their earnings ability, this information would enable the government to implement the first-best allocation. This is avoided by assuming heterogeneity with respect to disutility of work. Second, if workers are only heterogeneous with respect to their earning capacity, anyone with productivity below the minimum wage would simply become involuntarily unemployed. A minimum wage would, in that case, not raise any person's wage rate and as such not be an instrument for redistribution at all.

[7] Our results can readily be extended to a setting with more than two skill types. We briefly return to this point later.

[8] For variables that depend on ability, the first subscript θ always denotes the ability level.

the utility costs of sacrificing an hour of leisure, but could also encompass utility losses associated with educational effort. Its functional form therefore depends on whether the worker is high- or low-skilled. Specifically, disutility of a high-skilled worker is assumed to decrease with ability faster than disutility of a low-skilled worker, i.e., $q^H_{\theta,\theta} \leq q^L_{\theta,\theta}$. This ensures that individuals with high ability have a comparative advantage in high-skilled work, while individuals with low ability have a comparative advantage in low-skilled work.[9]

Utility is a twice differentiable function of consumption, c^i_θ, and disutility of work, $q^i(l^i_\theta, \theta)$. This function is assumed to be identical across individuals and given by:

$$(4.1) \quad V(c^i_\theta, q^i(l^i_\theta, \theta)), \quad V_c, -V_q > 0, \quad V_{cc}, V_{qq} \leq 0, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right].$$

Utility is increasing in consumption, but at a non-increasing rate. Naturally, utility is decreasing in the disutility of work, and is so at a non-decreasing rate. Consequently, since disutility of work is increasing in labor hours and decreasing in ability, utility is decreasing in labor hours and increasing in ability: $V_q q^i_{\theta,l} < 0$ and $V_q q^i_{\theta,\theta} > 0$.

The household budget constraint stipulates that consumption cannot exceed the sum of after-tax labor income:

$$(4.2) \qquad c^i_\theta \leq \omega^i l^i_\theta + T^i, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right].$$

Here, ω^i is the *net* wage rate per hour worked for skill type i, i.e., the wage that is left after income taxes are paid. Skill-dependent government transfers are given by T^i. If we denote t^i as the income tax rate and w^i as the gross wage rate for skill type i, we can alternatively write net wages as:

$$(4.3) \qquad \omega^i \equiv (1 - t^i) w^i, \quad i \in \{H, L\}.$$

When the minimum wage binds, the low-skilled gross wage, w^L, equals the minimum wage. Due to its analytical convenience we employ net wages and the minimum wage as government instruments in the remainder of this chapter. Taxes on low-skilled and high-skilled labor then follow residually from (4.3).

Other subscripts denote partial derivatives. Hence, $q^i_{\theta,x} \equiv \partial q^i_\theta / \partial x$.

[9] For example, q^H_θ can include the utility costs of educational effort required to become high-skilled. The assumption that $q^H_{\theta,\theta} \leq q^L_{\theta,\theta}$ could then be reinterpreted as an assumption that the disutility of educational effort to become high-skilled decreases with ability.

Implicit in our formulation are two assumptions. First, wages are identical for all workers with the same skill type, while they differ between workers with different skill types. Thus, workers of the same skill type are assumed to be perfect substitutes in production, whereas workers with different skill types are not.

Second, tax rates are identical for all workers with the same skill type, but differ between workers with different skill types. Allowing for fully non-linear income taxation – taxation that can be conditioned on total labor earnings, $w^i l^i_\theta$, as well as on skill type i – does not qualitatively alter our results. The reason is that a minimum wage is a skill-specific, linear policy instrument. Therefore, its effects on net labor earnings can be perfectly replicated by linear, skill-specific tax instruments. Consequently, judging the merits of a minimum wage relative to redistributive taxation does not require income taxes to be non-linear. We come back to this point in slightly more detail after deriving our main results.

So far, we discussed the utility function and budget constraints of the working population. When an individual is not employed – be it voluntarily or involuntarily so – he does not suffer any disutility of work, nor does he earn any labor income, but instead receives benefits equal to T^U. Hence, his utility is given by:

$$v^U \equiv V(T^U, 0). \tag{4.4}$$

4.3.2 Intensive labor supply

Individuals first decide to participate as a low-skilled or as a high-skilled worker or not to participate at all, then decide on how many hours of labor to supply. Using backward induction, we first solve for the optimal number of labor hours conditional on skill type. Every worker chooses his number of labor hours, l^i_θ, so as to maximize utility, (4.1), subject to the budget constraint, (4.2). Labor hours are implicitly given by equating the marginal rate of substitution of leisure for consumption with the net wage rate:

$$\frac{-V_q(c^i_\theta, q^i(l^i_\theta, \theta))q^i_l\,(l^i_\theta, \theta)}{V_c(c^i_\theta, q^i(l^i_\theta, \theta))} = \omega^i, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right]. \tag{4.5}$$

Jointly with the household budget constraint, this first-order condition determines an individual's intensive labor supply as a function of the net wage rate, ω^i, the government transfer, T^i, and his ability, θ:

$$l^i_\theta = l^i(\omega^i, T^i, \theta), \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right]. \tag{4.6}$$

4.3.3 Rationing schedule and expected indirect utility

Substituting optimal labor hours and the budget constraint into the utility function yields skill-dependent indirect utility, v^i_θ, as a function of the net wage rate, ω^i, the transfer, T^i, and ability, θ:

(4.7)

$$v^i_\theta \equiv v^i(\omega^i, T^i, \theta) \equiv V(\omega^i l^i(\omega^i, T^i, \theta) + T^i, q^i(l^i(\omega^i, T^i, \theta), \theta)), \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right].$$

Notice that the indirect utility function is *ex post*, in that it gives indirect utility for an individual with skill i, *conditional on obtaining a job.* However, in the case of a binding minimum wage, low-skilled workers face a positive probability of involuntary unemployment. When deciding to be low-skilled or not, an individual takes into account this probability of unemployment.[10] The *ex ante* expected indirect utility for a low-skilled individual depends on his probability of finding a job and his utility when unemployed, as well as his *ex post* indirect utility when obtaining a job. To formalize this, we first define the rationing schedule.

Definition 4.1 *The rationing schedule, $\{u_\theta\}$, assigns a probability of low-skilled unemployment to each level of ability $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$.*

We assume that this rationing schedule is differentiable with respect to θ. The relationship between the probability of unemployment and ability is, as we shall see, theoretically important, though empirically ambiguous. For that reason, we prefer to impose as few restrictions on the rationing schedule as possible. Instead, we determine the desirability of a minimum wage for any arbitrary rationing schedule.

Individuals decide to be either voluntarily unemployed, supply low-skilled labor, or supply high-skilled labor on the basis of the *expected* utility of the three options. Utility of being voluntarily unemployed is non-random and simply equals $V(T^U, 0)$. Similarly, as there is no unemployment among the high-skilled, utility of being high-skilled is also non-random and equals v^H_θ. Utility of an individual that decides to supply low-skilled labor, however, is random as he enjoys utility v^L_θ with probability $1 - u_\theta$, and utility $V(T^U, 0)$ with probability u_θ. Thus, the expected utility of becoming low-skilled equals:

(4.8)

$$v^{EL}_\theta \equiv v^{EL}(\omega^L, T^L, T^U, u_\theta, \theta) \equiv (1 - u_\theta) v^L(\omega^L, T^L, \theta) + u_\theta V(T^U, 0), \quad \theta \in [\underline{\theta}, \Phi).$$

[10]Thus, we assume individuals decide on their skill type before knowing with certainty whether they become employed or unemployed if low-skilled. Also, individuals cannot renege on their skill decision once unemployment is realized.

4.3.4 Extensive labor supply

We have assumed that individuals with low levels of θ have a comparative advantage in low-skilled work, whereas individuals with high levels of θ have a comparative advantage in high-skilled work. Therefore, in equilibrium, there are two critical ability levels, Φ and Θ, for which anyone with ability $\theta \in [\underline{\theta}, \Phi)$ becomes voluntarily unemployed, anyone with ability $\theta \in [\Phi, \Theta)$ wants to become a low-skilled worker, and anyone with ability $\theta \in [\Theta, \overline{\theta}]$ becomes a high-skilled worker. Figure 4.2 graphically illustrates this equilibrium. Notice that in the case of a binding minimum wage, individuals who decide to become low-skilled will either end up as low-skilled employed, or become involuntarily unemployed.

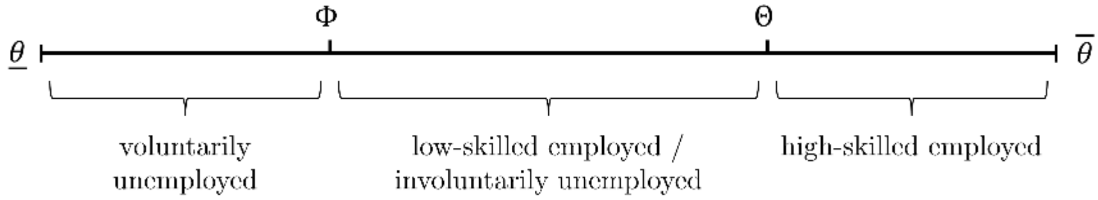

Figure 4.2: Graphical representation of an equilibrium

Participation decision

An individual with ability θ decides to be voluntarily unemployed if $v^U > v_\theta^{EL}$. Substituting for expected low-skilled utility from equation (4.8), this inequality can be written as $v^U > v_\theta^L$. The critical level of ability that separates non-participants from low-skilled workers, Φ, is therefore implicitly determined by:

$$V(T^U, 0) = v^L(\omega^L, T^L, \Phi). \tag{4.9}$$

Thus, the individual with ability Φ is indifferent between participating as a low-skilled worker and not participating at all.[11] Individuals with ability below Φ do not participate; those with ability above Φ do. As implied by equation (4.9), the critical level Φ can be written as a function of the policy instruments ω^L, T^L and

[11] For there to be any voluntarily unemployed individuals at all, a necessary condition is that individuals at the bottom of the ability distribution, i.e. those with ability $\underline{\theta}$, strictly prefer non-participation over participation: $V(T_U, 0) > v^L(\omega_L, T_L, \underline{\theta})$. Violation of this condition would imply everyone wants to participate. This case would have no bearing on our results and we disregard it in what follows. Furthermore, uniqueness of Φ is ensured by our assumption that low-skilled utility is strictly increasing in ability, i.e.: $v_{\theta,\theta}^L = V_q q_{\theta,\theta}^L > 0$ for all θ.

T^U:

$$\Phi = \Phi(\omega^L, T^L, T^U), \quad \Phi_{\omega^L}, \Phi_{T^L} < 0, \quad \Phi_{T^U} > 0. \tag{4.10}$$

Naturally, if the low-skilled net wage, ω^L, and transfer, T^L, decrease or unemployment benefits, T^U, increase, more individuals decide to become voluntarily unemployed and Φ increases as a consequence.

Skill decision

An individual with ability θ decides to become low-skilled rather than high-skilled if $v_\theta^{EL} > v_\theta^H$. The critical level of ability that separates the high-skilled from the low-skilled, Θ, is therefore implicitly determined by:

$$v^{EL}(\omega^L, T^L, T^U, u_\Theta, \Theta) = v^H(\omega^H, T^H, \Theta). \tag{4.11}$$

Thus, the individual with ability Θ is indifferent between being high- or low-skilled.[12] Individuals with ability below Θ (and above Φ) decide to be low-skilled; those with ability above Θ decide to be high-skilled.

Equation (4.11) implicitly determines the critical level Θ as a function of the government instruments ω^H, ω^L, T^H, T^L, and T^U, and the unemployment rate u_Θ:

$$\Theta = \Theta(\omega^H, \omega^L, T^H, T^L, T^U, u_\Theta), \quad \Theta_{\omega^H}, \Theta_{T^H}, \Theta_{u_\Theta} < 0, \quad \Theta_{\omega^L}, \Theta_{T^L}, \Theta_{T^U} > 0. \tag{4.12}$$

Notice that the higher is Θ, the lower is the number of high-skilled workers. Naturally, if expected low-skilled earnings rise or high-skilled earnings fall, less individuals decide to become high-skilled and Θ increases as a consequence. Therefore, the number of persons that are high-skilled is increasing in the high-skilled net wage rate, ω^H, and the high-skilled transfer, T^H. It is decreasing in the low-skilled net wage rate, ω^L, the low-skilled transfer, T^L, and unemployment benefits, T^U.

A crucial ingredient for our welfare analysis of minimum-wage legislation is the effect of unemployment on skill formation. The partial derivative of Θ with

[12] A necessary condition for this to hold is that, at the critical level Θ high-skilled utility is increasing in ability at a faster rate than expected low-skilled utility: $v_{\Theta,\theta}^H > v_{\Theta,\theta}^{EL} + v_{\Theta,u_\theta}^{EL} u_{\Theta,\theta}$. A sufficient condition for Θ to be unique is that this holds for every level of ability: $v_{\vartheta,\theta}^H > v_{\vartheta,\theta}^{EL} + v_{\vartheta,u_\theta}^{EL} u_{\vartheta,\theta}$ for all ϑ. We assume this is indeed the case. Allowing for multiple equilibrium levels of Θ would not compromise any of the economic insights, but only increase the complexity of the analysis. We briefly return to this point later.

respect to u_Θ, obtained by taking the derivative of equation (4.11), is given by:

$$\frac{\partial \Theta}{\partial u_\Theta} = -\frac{v^L_\Theta - v^U}{v^H_{\Theta,\theta} - v^{EL}_{\Theta,\theta} - v^{EL}_{\Theta,u_\Theta} u_{\Theta,\theta}} < 0. \tag{4.13}$$

Notice that the numerator gives the difference between low-skilled utility and unemployed utility for an individual with ability Θ, which is positive. It is assumed that high-skilled utility v^H_Θ is increasing in ability faster than expected low-skilled utility v^{EL}_Θ, such that the denominator is positive as well (see also footnote 12). Hence, the partial derivative of Θ with respect to u_Θ is negative. An increased probability of unemployment lowers the expected utility of being low-skilled, while it does not directly affect high-skilled utility. An increase in the unemployment rate u_Θ, *ceteris paribus*, therefore leads to more high-skilled workers and a lower level of Θ. Intuitively, when the low-skilled unemployment rate increases, some individuals escape the higher probability of low-skilled unemployment by becoming a high-skilled worker.

4.3.5 Aggregate labor supply

Given individuals' extensive labor supply decisions and their intensive labor supply decision, aggregate low-skilled labor supply is given by:

$$\int_\Phi^\Theta l^L_\theta \mathrm{d}G(\theta). \tag{4.14}$$

Notice that this aggregate labor supply is only notional, i.e., it includes labor hours of every low-skilled individual that would like to work, including the ones that are ultimately unable to find a job. Aggregate high-skilled labor supply is given by:

$$\int_\Theta^{\overline{\theta}} l^H_\theta \mathrm{d}G(\theta). \tag{4.15}$$

4.3.6 Firms

We assume workers of the same skill type are perfect substitutes in production, whereas high-skilled labor and low-skilled labor are imperfect substitutes in production. Denoting aggregate labor demand for high-skilled workers as L^H and for low-skilled workers as L^L, we can write the production function as:

$$Y \equiv F(L^H, L^L), \quad F_i > 0, \quad F_{ii} < 0, \quad F_{HL} > 0, \quad i \in \{H, L\}. \tag{4.16}$$

For notational convenience, we let F_i denote the partial derivative of F with respect to L^i. Marginal products of each labor type are positive but diminishing, and F is assumed to be homogeneous of degree one. A representative firm takes wages as given and decides on aggregate high- and low-skilled labor demand so as to maximize profits. This yields standard necessary (and sufficient) conditions for profit maximization:

$$F_i(L^H, L^L) = w^i, \quad i \in \{H, L\}\,. \tag{4.17}$$

Notice that labor demand is a function of gross wages. An increase in the gross wage of labor type i leads to lower demand for labor type i to realign marginal productivity and labor costs.

4.3.7 General equilibrium

The determination of equilibrium in a labor market with a binding minimum wage differs substantially from that in a labor market without a binding minimum wage. For high-skilled labor we assume the minimum wage is never binding. In that case, the high-skilled gross wage adjusts to equate labor supply and labor demand in the conventional way: higher demand leads to higher wages, higher supply to lower wages. In equilibrium, labor demand always equals labor supply:

$$L^H = \int_{\Theta}^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta). \tag{4.18}$$

If the low-skilled labor market is affected by a binding minimum wage, equilibrium can no longer be determined by wage-rate adjustments as the gross wage rate cannot be lowered. Instead, unemployment adjusts to ensure labor-market equilibrium: higher demand leads to lower unemployment, higher supply to higher unemployment. In equilibrium, labor demand equals *effective* labor supply, which is less than *notional* labor supply:

$$L^L = \int_{\Phi}^{\Theta} (1 - u_\theta) l_\theta^L \mathrm{d}G(\theta). \tag{4.19}$$

Notice that we have an infinity of unemployment probabilities $\{u_\theta\}$, but only one equilibrium condition. Thus, above equation simply provides the equilibrium condition for aggregate unemployment, but does not determine individual unemployment probabilities, $\{u_\theta\}$. Individual probabilities are determined by the rationing schedule. This goes directly to the core of our argument: theory points to higher unemployment due to higher minimum wages, but theoretically we know little about *which workers* become unemployed.

Substituting the labor-market equilibrium conditions into the first-order conditions of the firm, we get:

$$F_i\left(\int_{\Theta}^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta), \int_{\Phi}^{\Theta}(1-u_\theta) l_\theta^L \mathrm{d}G(\theta)\right) = w^i, \quad i \in \{H, L\}. \tag{4.20}$$

We can now define equilibrium in the private sector as follows.

Definition 4.2 *For given values of government instruments* $\{w^L, \omega^L, \omega^H, T^U, T^L, T^H\}$, *a competitive equilibrium in the private sector is defined as the allocation for which labor supply, participation and skill decisions* $\{l_\theta^L, l_\theta^H, \Phi, \Theta\}$, $\forall\theta \in [\underline{\theta}, \overline{\theta}]$, *the high-skilled wage rate* w^H, *and unemployment rates* $\{u_\theta\}$, $\forall\theta \in [\Phi, \Theta)$, *are such that the optimality conditions for labor supply, (4.6), participation, (4.10), skill formation, (4.12), and labor demand, (4.20), are satisfied.*

4.4 Equilibrium implications of a minimum-wage reform

4.4.1 A net-income-neutral minimum-wage reform

This section tracks the equilibrium implications of the decision to increase the minimum wage, while keeping net incomes of high- and low-skilled workers constant. That is, t^H and t^L endogenously adjust to changes in gross wages such that net wages ω^H and ω^L remain fixed. Transfers T^i remain constant as well. We refer to this policy as a *net-income-neutral (NIN) minimum-wage reform*, since it increases w^L while keeping $\{\omega^L, \omega^H, T^U, T^L, T^H\}$ constant.[13]

How do income taxes adjust under the NIN minimum-wage reform? Recall that $\omega^i \equiv (1-t^i)w^i$. Since $\mathrm{d}\omega^i = 0$, we must have that $\mathrm{d}t^i/(1-t^i) = \mathrm{d}w^i/w^i$. Furthermore, notice that linear homogeneity of the production function implies that profits are absent, such that $\mathrm{d}\left(F(L^H, L^L) - w^H L^H - w^L L^L\right) = 0$. Solving for the derivative and using the firm's first-order conditions, (4.17), yields $L^H \mathrm{d}w^H = -L^L \mathrm{d}w^L$.

[13] Comparative statics of a change in the minimum wage, for given tax rates, have straightforward interpretations but are mathematically tedious as we demonstrate in Gerritsen and Jacobs (2013*b*, Chapter 3 of this Dissertation). A NIN minimum-wage reform allows us to ignore the behavioral effects of changes in net wages.

Intuitively, as the minimum wage increase reduces low-skilled employment, labor productivity and gross wages of high-skilled workers decline since high- and low-skilled workers are cooperant factors of production, i.e., $F_{HL} > 0$. Thus, we can define a net-income-neutral minimum wage increase as follows.

Definition 4.3 *A net-income-neutral (NIN) minimum wage increase is a policy reform that raises the minimum wage,* $\mathrm{d}w^L > 0$*, while keeping net wage rates,* ω^L, ω^H*, and government transfers,* T^U, T^L, T^H*, constant. In order to keep net wages constant, income tax rates,* t^L, t^H *endogenously adjust, such that:*

$$\frac{\mathrm{d}t^L}{1-t^L} = \frac{\mathrm{d}w^L}{w^L}, \quad \frac{\mathrm{d}t^H}{1-t^H} = \frac{\mathrm{d}w^H}{w^H} = -\frac{w^L L^L}{w^H L^H}\frac{\mathrm{d}w^L}{w^L}. \tag{4.21}$$

Under the NIN minimum-wage reform the low-skilled tax rate rises to offset the higher low-skilled gross wage. Furthermore, the high-skilled tax rate declines to compensate for the negative general-equilibrium effect on the high-skilled gross wage.

4.4.2 Comparative statics of the NIN minimum-wage reform

To understand the welfare effects of the NIN minimum-wage reform, we first derive the comparative statics of the reform for the model's key variables. To that end, it is helpful to introduce the concept of the unemployment incidence.

Definition 4.4 *The unemployment incidence at ability level* θ*, denoted by* I_θ*, gives the marginal increase in unemployment at that ability level,* $g(\theta)\mathrm{d}u_\theta$*, as a fraction of the total increase in rationing,* $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$*:*

$$I_\theta \equiv \frac{g(\theta)\mathrm{d}u_\theta}{\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)} \in [0, \infty), \quad \theta \in [\Phi, \Theta). \tag{4.22}$$

What we refer to as the unemployment incidence, I_θ, is in fact a density function. I_θ measures the increase in unemployment at θ, $g(\theta)\mathrm{d}u_\theta$, relative to the increase in unemployment across all ability levels $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$. As with any density function, I_θ takes on a value between 0 and ∞ and the integral of the unemployment incidence equals one, i.e. $\int_\Phi^\Theta I_\theta \mathrm{d}\theta = 1$. If none of the incidence of increased unemployment is at ability level θ, then $I_\theta = 0$. If the incidence of increased unemployment is solely at ability level θ, then $I_\theta \to \infty$.

Lemma 4.1 *The comparative statics results of a NIN minimum-wage reform, given the equilibrium as described by Definition 4.2, and for given values of all other government instruments, are:*

$$\text{(4.23)} \qquad \mathrm{d}l^i_\theta = \mathrm{d}v^U = \mathrm{d}v^i_\theta = \mathrm{d}G(\Phi) = 0, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right],$$

$$\text{(4.24)} \qquad -\mathrm{d}G(\Theta) = -\frac{\partial \Theta}{\partial u_\Theta} I_\Theta \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta),$$

$$\text{(4.25)} \qquad \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta) = \left(\int_\Phi^\Theta \frac{l^L_\theta I_\theta}{L^L} \mathrm{d}\theta\right)^{-1} \left(\varepsilon^L_w \frac{\mathrm{d}w^L}{w^L} + \varepsilon^L_\Theta \mathrm{d}G(\Theta)\right).$$

Here, $\varepsilon^L_w \equiv \left(\frac{-F_{LL}L^L}{F_L}\right)^{-1} > 0$ *is the low-skilled labor-demand elasticity with respect to the low-skilled wage rate,* w^L*, and* $\varepsilon^L_\Theta \equiv \left(\frac{l^H_\Theta}{L^H} + \frac{(1-u_\Theta)l^L_\Theta}{L^L}\right) > 0$ *is the low-skilled labor-demand elasticity with respect to high-skilled labor supply,* $1 - G(\Theta)$*.*

Proof. From equations (4.4), (4.6), (4.7), and (4.10), follows that intensive labor supply, indirect utility, and the participation margin, do not depend on the minimum wage, which proves (4.23). Taking the derivatives of (4.11) and (4.20), and substituting for I_θ from (4.22), yields (4.24) and (4.25). ■

First, equations (4.23) demonstrate that a NIN minimum-wage reform is neutral with respect to intensive labor supply, indirect utility, and the number of voluntary unemployed. Intuitively, individuals' consumption and intensive labor supply decisions, and hence utility, depend on their net wage, not their gross wage. Provided that an individual remains (un)employed, his utility is unaffected by the NIN minimum-wage reform. Similarly, the critical level Φ equates low-skilled utility and unemployed utility and, therefore, also does not depend on the minimum wage.

Second, equation (4.24) gives the effect of the NIN minimum-wage reform on skill formation. The critical level Θ equates high-skilled utility with expected low-skilled utility. As such, it does not directly depend on the minimum wage, but does depend on the unemployment probability at the critical level, u_Θ. From equation (4.24) it follows that the change in Θ can be seen as the product of two terms. $\frac{\partial \Theta}{\partial u_\Theta}$ is the partial effect on Θ of a higher unemployment probability at ability level Θ, which is strictly negative and finite. Intuitively, when the chances of obtaining a low-skilled job diminish, individuals with ability Θ decide to become high-skilled instead. $I_\Theta \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$ gives the increase in the number of unemployed workers with ability Θ. Hence, if the minimum wage raises unemployment, and the unemployment incidence at ability level Θ is positive, the NIN minimum-wage reform leads to more high-skilled workers.

Third, equation (4.25) gives the effect of the NIN minimum-wage reform on unemployment. A minimum wage reduces low-skilled labor demand, thereby raising unemployment. This increase is seen to depend on three terms. First, it negatively depends on the labor-weighted unemployment incidence, $\int_{\Phi}^{\Theta}(l_{\theta}^{L}I_{\theta}/L^{L})\mathrm{d}\theta$. If the unemployment incidence mainly falls on workers with high intensive labor supply, fewer workers become unemployed for a given aggregate reduction in labor demand. Second, $\varepsilon_{w}^{L}\frac{\mathrm{d}w^{L}}{w^{L}}$ captures the direct unemployment effect of higher minimum wages. For a given relative increase in minimum wages, $\mathrm{d}w^{L}/w^{L}$, the larger the labor-demand elasticity, ε_{w}^{L}, the larger the increase in low-skilled unemployment. Finally, $\varepsilon_{\Theta}^{L}\mathrm{d}G(\Theta)$ gives the unemployment effect of a change in skill formation. If a minimum wage leads to more high-skilled workers, such that $\mathrm{d}G(\Theta) < 0$, the increase in unemployment is smaller. Since low-skilled labor productivity is increasing in the number of high-skilled workers, there is less need to reduce low-skilled labor demand if skill formation increases.

The unemployment incidence at the skill margin, I_{Θ}, is *the* critical determinant of the comparative statics of a NIN minimum-wage increase. This is illustrated by the following Corollary.

Corollary 4.1 *The comparative statics of a NIN minimum-wage increase,* $\mathrm{d}w^{L} > 0$*, depend on* I_{Θ} *as follows:*

$$\text{(4.26)} \qquad I_{\Theta} = 0: \quad -\mathrm{d}G(\Theta) = 0, \quad \int_{\Phi}^{\Theta} \mathrm{d}u_{\theta}\mathrm{d}G(\theta) > 0,$$

$$\text{(4.27)} \qquad I_{\Theta} \to \infty: \quad -\mathrm{d}G(\Theta) > 0, \quad \int_{\Phi}^{\Theta} \mathrm{d}u_{\theta}\mathrm{d}G(\theta) = 0,$$

$$\text{(4.28)} \qquad \forall\ I_{\Theta} \in (0,\infty): \quad \frac{-\mathrm{d}^{2}G(\Theta)}{\mathrm{d}I_{\Theta}} > 0, \quad \frac{\int_{\Phi}^{\Theta} \mathrm{d}^{2}u_{\theta}\mathrm{d}G(\theta)}{\mathrm{d}I_{\Theta}} < 0.$$

Proof. Solve equations (4.24) and (4.25) of Lemma 4.1 for $-\mathrm{d}G(\Theta)$ and $\int_{\Phi}^{\Theta}\mathrm{d}u_{\theta}\mathrm{d}G(\theta)$, and rearrange to obtain:

$$\text{(4.29)} \qquad -\mathrm{d}G(\Theta) = -\frac{\varepsilon_{w}^{L}}{\int_{\Phi}^{\Theta}\frac{l_{\theta}^{L}I_{\theta}}{L^{L}}\mathrm{d}\theta + \varepsilon_{\Theta}^{L}\frac{-\partial\Theta}{\partial u_{\Theta}}I_{\Theta}}\frac{\mathrm{d}w^{L}}{w^{L}}\frac{\partial\Theta}{\partial u_{\Theta}}I_{\Theta} \geq 0,$$

$$\text{(4.30)} \qquad \int_{\Phi}^{\Theta}\mathrm{d}u_{\theta}\mathrm{d}G(\theta) = \frac{\varepsilon_{w}^{L}}{\int_{\Phi}^{\Theta}\frac{l_{\theta}^{L}I_{\theta}}{L^{L}}\mathrm{d}\theta + \varepsilon_{\Theta}^{L}\frac{-\partial\Theta}{\partial u_{\Theta}}I_{\Theta}}\frac{\mathrm{d}w^{L}}{w^{L}} \geq 0.$$

Substitute for $I_\Theta = 0$, and the limit of $I_\Theta \to \infty$ into these equations to obtain the first two lines of the Corollary. Take derivatives to obtain the third line. ■

Corollary 4.1 shows how changes in skill formation ($-\mathrm{d}G(\Theta)$) and total unemployment ($\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$) depend on the unemployment incidence at the skill margin (I_Θ). Equation (4.26) indicates that if the incidence is 0, a NIN minimum-wage increase does not affect skill formation, but does lead to more unemployment. The intuition is straightforward: as no one at the skill margin is affected by the rise in unemployment, the minimum-wage increase does not provide incentives to become high-skilled.

Equation (4.27) indicates that the opposite occurs if the unemployment incidence is concentrated solely at ability level Θ. In that case, a NIN minimum-wage increase leads to more skill formation without affecting unemployment. Intuitively, if $I_\Theta \to \infty$, all individuals that would be affected by an increase in unemployment are located at the skill margin and decide to become high-skilled to escape low-skilled unemployment. Therefore, any reduction in low-skilled labor demand is completely offset by individuals moving from the low-skilled to the high-skilled sector, and no one becomes unemployed.

In general, equation (4.28) establishes that the increase in high-skilled employment is monotonically increasing in the unemployment incidence at the skill margin. For precisely that reason, the magnitude of the unemployment effect is monotonically decreasing in the unemployment incidence at the skill margin. Summing up, as I_Θ goes from zero to infinity, the change in skill formation monotonically goes from zero to some positive amount, and the change in unemployment monotonically goes from some positive amount to zero.

4.5 Welfare analysis

4.5.1 Social objectives and government budget constraint

We assume that social preferences are utilitarian such that social welfare, $\mathcal{W}$, is given by the unweighted sum of individual utilities:

(4.31)
$$\mathcal{W} \equiv \int_\Theta^{\overline{\theta}} v^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + \int_\Phi^\Theta v^{EL}(\omega^L, T^L, T^U, u_\theta, \theta)\mathrm{d}G(\theta) + G(\Phi)V(T^U, 0).$$

It is relatively straightforward to allow for a social welfare function with social welfare weights that decline with utility. This does not affect our results.

Ability is private information, which precludes a first-best outcome as the government cannot condition its taxes on $\theta \in [\underline{\theta}, \overline{\theta}]$. Instead, the government must rely on distortionary taxes on verifiable labor income, or introduce a minimum wage, to pursue its redistributional goals. As a result, the trade-off between equity and efficiency emerges. As we have done throughout this chapter, we assume that the government can condition tax rates and transfers on skill type $i \in \{H, L\}$. The government must observe wages, and thus skill levels, to implement and enforce a minimum wage.

The government obtains resources from income taxation, while it spends resources on transfers and some exogenous revenue requirement, R. Its budget $\mathcal{B}$ is thus given by:

(4.32)

$$\begin{aligned}\mathcal{B} \equiv &(w^H - \omega^H) \int_{\Theta}^{\overline{\theta}} l^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + (w^L - \omega^L) \int_{\Phi}^{\Theta} (1 - u_\theta) l^L(\omega^L, T^L, \theta)\mathrm{d}G(\theta) \\ &- (1 - G(\Theta))T^H - \int_{\Phi}^{\Theta} (1 - u_\theta)\mathrm{d}G(\theta)T^L - \left(\int_{\Phi}^{\Theta} u_\theta \mathrm{d}G(\theta) + G(\Phi)\right) T^U - R.\end{aligned}$$

Notice that tax revenue per hour worked for workers of skill type i is given by $t^i w^i = w^i - \omega^i$. Thus, the first line gives total revenue from income taxation, whereas the second line gives total expenditures on transfers and the exogenous revenue requirement. The government is required to balance its budget, so that $\mathcal{B} = 0$.

4.5.2 Desirability of a minimum-wage increase

For any private equilibrium, a NIN minimum-wage increase, $\mathrm{d}w^L > 0$, is desirable if the change in social welfare is positive. Denoting the Lagrange multiplier associated with the government budget constraint by λ, the desirability condition for an increase in the minimum wage is given by:

$$\frac{\mathrm{d}\mathcal{W}}{\lambda} + \mathrm{d}\mathcal{B} > 0. \tag{4.33}$$

In order to facilitate the interpretation of the desirability condition of the minimum wage, we define *wedges* on skill formation and unemployment as Δ_Θ and Δ_u. The welfare gain of an increase in high-skilled employment is given by Δ_Θ. It is the difference between the social and private value of increased skill formation, which equals the increase in tax revenue when a low-skilled worker with ability Θ

decides to become high-skilled:

$$\Delta_\Theta \equiv t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U. \tag{4.34}$$

When a low-skilled individual with ability Θ becomes high-skilled, the government receives tax revenue $t^H w^H l_\Theta^H - T^H$, but foregoes expected tax revenue $(1-u_\Theta)(t^L w^L l_\Theta^L - T^L)$ from employed low-skilled workers and $-u_\Theta T^U$ from unemployed low-skilled workers. For progressive tax systems, the tax wedge on skill formation tends to be positive, $\Delta_\Theta > 0$. In that case, more skill acquisition leads to higher tax revenue and, thus, higher social welfare.

Similarly, Δ_u is equal to the welfare loss of an additional involuntarily unemployed individual:

$$\Delta_u \equiv \int_\Phi^\Theta \left(\frac{v_\theta^L - v^U}{\lambda} + t^L w^L l_\theta^L - T^L + T^U \right) I_\theta \mathrm{d}\theta. \tag{4.35}$$

Unemployment affects both individuals' utility and tax revenue. Individuals that were previously employed and enjoyed utility v_θ^L become unemployed and enjoy utility $v^U \leq v_\theta^L$. Hence, welfare losses due to direct utility losses (in monetary equivalents) are given by the first term within brackets. Moreover, the government foregoes tax revenue, $t^L w^L l_\theta^L - T^L$, and has to pay additional unemployment benefits, T^U. These revenue effects are given by the remaining terms within brackets. The unemployment incidence, I_θ, determines the additional unemployment at ability level θ. Thus, the total wedge is given by the integral over utility and revenue costs, weighted by the unemployment incidence. If the government redistributes towards the unemployed, Δ_u tends to be positive as higher unemployment leads to both utility and revenue losses.

The following Proposition gives the main theoretical result of the chapter: it provides the desirability condition of a NIN minimum-wage increase.

Proposition 4.1 *A NIN minimum-wage increase enhances social welfare if and only if the resulting marginal benefits of increased high-skilled employment outweigh the marginal costs of increased unemployment:*

$$-\Delta_\Theta \mathrm{d}G(\Theta) > \Delta_u \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta). \tag{4.36}$$

Equivalently, a NIN minimum-wage increase enhances social welfare if and only if *the marginal benefits of higher unemployment are larger than its marginal costs:*

$$-\frac{\partial\Theta}{\partial u_\Theta} I_\Theta \Delta_\Theta > \Delta_u. \tag{4.37}$$

Proof. Take derivatives of (4.31) and (4.32) and substitute into (4.33). Rearrange and substitute (4.34) and (4.35) to obtain (4.36) (see the Appendix for the full derivation). Note from (4.12) that $\frac{\mathrm{d}\Theta}{\mathrm{d}u_\Theta} = \frac{\partial\Theta}{\partial u_\Theta}$ since $\left\{\omega^H, \omega^L, T^H, T^L, T^U\right\}$ remain constant. Substitute (4.24) and rearrange to obtain (4.37). ■

Lemma 4.1 established that the only comparative statics of a NIN minimum-wage increase consist of increases in involuntary unemployment and the number of high-skilled workers, while leaving intensive labor supply, participation and utility unaffected. In line with this finding, Proposition 4.1 establishes that a minimum-wage increase is desirable only if the benefits of having more high-skilled workers outweigh the costs of more unemployment. To see this, note that the left-hand side of inequality (4.36) gives the marginal benefits of increased skill formation while the right-hand side gives the marginal costs of higher unemployment.

Because a NIN minimum wage only affects welfare through unemployment, its desirability is equivalent to the desirability of higher unemployment. For this reason, we can rewrite the desirability condition as equation (4.37). Notice that the left-hand side gives the marginal benefits of unemployment. The benefits increase with the effect of aggregate unemployment on skill formation, given by $-\frac{\partial\Theta}{\partial u_\Theta} I_\Theta$. The incidence I_Θ determines the effect of an additional unemployed individual on u_Θ, whereas $\frac{\partial\Theta}{\partial u_\Theta}$ determines the effect of u_Θ on the number of high-skilled workers. The marginal benefits of unemployment are furthermore increasing with the marginal social gains of high-skilled employment, Δ_Θ. As the minimum wage leads to more high-skilled employment, such a policy is more likely to be desirable if the social benefits of high-skilled work, Δ_Θ, are high. The right-hand side of inequality (4.37) gives the marginal costs of unemployment, Δ_u. Since a minimum wage leads to higher unemployment, it is less likely to be desirable if the marginal social costs of higher unemployment, Δ_u, are large.[14]

Notice that if conditions (4.36) and (4.37) hold, a NIN minimum-wage in-

[14]Alternatively, the term $-\frac{\partial\Theta}{\partial u_\Theta} I_\Theta$ can be expressed in terms of an elasticity as $(H/L)\zeta_u$ which is the ratio of high-skilled individuals ($H \equiv 1 - G(\Theta)$) and low-skilled individuals ($L \equiv G(\Theta) - G(\Phi)$), multiplied with the semi-elasticity $\zeta_u \equiv \frac{-\mathrm{d}G(\Theta)/(1-G(\Theta))}{\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)/(G(\Theta)-G(\Phi))}$, which gives the relative change in the number of high-skilled workers with respect to a one-percent increase in all individual unemployment rates.

crease raises social welfare, but it never constitutes a Pareto improvement if the unemployment incidence is inefficient. After all, the individuals that become involuntarily unemployed are worse off because of the reform. Conversely, if conditions (4.36) and (4.37) do not hold, a NIN minimum-wage decrease yields a genuine Pareto improvement if it leads to both higher government revenue and positive utility benefits for the involuntarily unemployed that are able to obtain a job due to the reform.[15]

Proposition 4.1 demonstrates the critical importance of the unemployment incidence I_Θ for the desirability of minimum wages. To illustrate this graphically, turn back to Figure 4.1. In the first panel, none of the unemployment incidence is on workers with ability Θ. As a consequence, a NIN minimum-wage increase merely leads to more unemployment and is therefore unambiguously welfare decreasing. In that case, it would be desirable to reduce the minimum wage – in a net-income-neutral fashion by compensating tax changes. However, in the second panel of Figure 4.1, much of the incidence falls on workers with ability Θ. Consequently, a NIN minimum-wage increase has a large beneficial effect on high-skilled employment, relative to its adverse effect on unemployment. In that case, a minimum-wage increase might very well be desirable.

In the next section we calibrate the condition of Proposition 4.1 to get some idea of the desirability of a NIN minimum-wage increase (or decrease) in various OECD countries. Before doing so, however, we turn to the question of whether a binding minimum wage is part of the optimal policy mix.

4.5.3 A binding minimum wage as part of the policy optimum

Regardless of the initial equilibrium, Proposition 4.1 provides the condition under which a NIN minimum-wage increase leads to a welfare gain. However, we are especially interested in whether the minimum wage is also an appropriate instrument for redistribution, that is, whether it is part of the overall policy optimum. This must be the case if a NIN minimum-wage increase is desirable in the tax optimum without a minimum wage. In this subsection we therefore determine if condition (4.37) in Proposition 4.1 holds in the policy optimum without a binding

[15] Thus, the condition for a NIN minimum-wage decrease to be Pareto-improving is stronger than the simple negation of condition (4.37). Denoting the public revenue loss due to an additional involuntarily unemployed individual as $\Delta_\Phi \equiv t^L w^L l^L(\Phi) - T^L + T^U \leq \Delta_u$, this condition can be written as: $-\frac{\partial \Theta}{\partial u_\Theta} I_\Theta \Delta_\Theta \leq \Delta_\Phi$.

minimum wage. For this, we rely on the following Lemma, which establishes that, in the absence of a minimum wage, the optimal wedges on unemployment (Δ_u) and skill formation (Δ_Θ) are positive and finite.

Lemma 4.2 *If the social marginal value of income is decreasing in ability, then the wedges on unemployment and skill formation,* Δ_u *and* Δ_Θ*, are positive and finite in the policy optimum without a minimum wage.*

Proof. See Appendix. ■

The intuition for a positive unemployment wedge ($\Delta_u > 0$) is as follows. Recall that the utility loss of a marginal increase in unemployment is non-negative, such that it suffices to show that a marginal increase in unemployment generates a net tax-revenue loss. If the government attaches a larger social marginal value of income to the unemployed than to the employed, it will, in the tax optimum, redistribute resources from the employed to the unemployed. In that case, an increase in unemployment, *ceteris paribus*, causes revenue losses, implying that $\Delta_u > 0$. The intuition for $\Delta_\Theta > 0$ follows a similar logic. If the government attaches a lower social marginal value of income to high-skilled workers than to low-skilled workers, it will redistribute resources away from the high-skilled towards the low-skilled. In that case, an increase in high-skilled employment, *ceteris paribus*, generates a revenue gain, such that $\Delta_\Theta > 0$.

Lemma 4.2 allows us to formulate how the optimality of a minimum wage depends on the incidence of involuntary unemployment.

Corollary 4.2 *Given the policy optimum without a minimum wage, there exists a value of the unemployment incidence* $I_\Theta^* = \frac{\Delta_u}{\Delta_\Theta}\left(-\frac{\partial\Theta}{\partial u_\Theta}\right)^{-1}$ *for which the introduction of a binding minimum wage has no effect on social welfare. For any value of* $I_\Theta > I_\Theta^*$*, a binding minimum wage is part of the policy optimum. For any value of* $I_\Theta < I_\Theta^*$*, a (marginally) binding minimum wage is not part of the policy optimum.*

Proof. Lemma 4.2 establishes that Δ_u and Δ_Θ are positive and finite in the policy optimum without a minimum wage. Since there is no unemployment in this optimum, Δ_u and Δ_Θ are necessarily independent of I_Θ. Equation (4.13) establishes that $-\partial\Theta/\partial u_\Theta$ is positive and finite and independent of I_Θ. Together with the desirability condition (4.37) of Proposition 4.1, this implies that a marginally binding minimum wage has no welfare effect for $I_\Theta = I_\Theta^*$, is welfare increasing for all $I_\Theta > I_\Theta^*$, and is welfare decreasing for all $I_\Theta < I_\Theta^*$. ■

From Corollary 4.1 we know that if the unemployment incidence I_Θ goes to infinity, a minimum wage does not lead to unemployment at all, but only to more high-skilled workers. In that case, a minimum wage is strictly welfare enhancing and optimally applied alongside taxes and transfers. If I_Θ goes to zero, a minimum wage only leads to higher unemployment and does not affect the number of high-skilled workers, such that a marginally binding minimum wage is strictly welfare decreasing and not a part of the optimal policy mix.[16] Thus, a binding minimum wage is an appropriate instrument for redistribution if the unemployment incidence at Θ is large enough. That is, if the increased probability of unemployment falls, to a large degree, on low-skilled workers that are indifferent between being low-skilled or high-skilled. In that case, a binding minimum wage leads to a large increase of high-skilled employment, relative to the increase in unemployment. The social benefits of this increase in high-skilled employment then outweigh the social costs of higher unemployment.

4.5.4 Robustness

In deriving our results we made a number of assumptions that warrant a brief discussion. The results appear robust to relaxing these assumptions.

Number of skill types – We could extend the model by allowing for more than two skill types in production, in which case only the lowest skill type would be subject to a binding minimum wage. Our main results would be unaffected, provided that the government is able to tax each skill type separately. In particular, a NIN minimum-wage reform, which neutralizes the general-equilibrium effects of a minimum wage on the net returns of the additional factor inputs by appropriate tax adjustments, would result in a similar desirability condition for a minimum wage. With constant net factor returns, factor supply is unaffected, except to the extent that increased unemployment causes some low-skilled workers to supply labor as a different skill type. Again, the desirability of a minimum wage would be determined by the net balance of the costs of unemployment and the benefits of low-skilled workers deciding to become another skill type.

Uniqueness of Θ – Throughout the analysis, we assumed that the equilibrium skill margin, Θ, is unique. Uniqueness is implied by the assumption that the

[16]Notice that a *marginal* increase in the minimum wage is not welfare enhancing if $I_\Theta < I^*_\Theta$. This does not necessarily imply that larger increases of the minimum wage are not welfare enhancing, as that would depend on the total inframarginal unemployment incidence. This is the reason we employ the term 'marginally binding minimum wage.'

utility difference between high-skilled and low-skilled work is increasing in ability. This facilitates the analysis and graphical representation. However, even when Θ is not unique, the economic insights remain the same. Suppose that there are three critical ability levels at which individuals are indifferent between being low-skilled or high-skilled, Θ_a, Θ_b, and Θ_c, such that individuals with ability $\theta \in [\Phi, \Theta_a) \cup [\Theta_b, \Theta_c)$ become low-skilled, while individuals with ability $\theta \in (\Theta_a, \Theta_b] \cup \left[\Theta_c, \overline{\theta}\right]$ become high-skilled. The comparative statics of a NIN minimum-wage increase would again consist of higher unemployment and more skill formation. The extent to which it leads to more skill formation, however, now depends on the unemployment incidence at all three critical levels. Similarly, the desirability condition would feature the unemployment incidence and wedges on skill formation at all three critical levels. However, the intuition would remain the same: a NIN minimum-wage increase is only desirable if the resulting social benefits from skill formation outweigh the social costs of unemployment.

Social welfare function – Our results do not depend on the presumed utilitarian social welfare function, and would remain valid for any concave social welfare function. Naturally, a different social welfare function would affect the optimal second-best allocations, and therefore the optimal wedges on unemployment and skill formation (Δ_u and Δ_Θ). They would, however, still be positive and finite such that our results remain unaffected.

Non-linear taxation – Allowing for fully non-linear taxation, i.e., taxation conditional on both skill type and the level of labor earnings, would simply have brought more mathematical complexity without generating additional insights. Recall that the relative changes in tax rates to keep the net incomes of high-skilled and low-skilled workers constant should satisfy $\mathrm{d}t^i/(1-t^i) = \mathrm{d}w^i/w^i$ for $i \in \{H, L\}$. The NIN minimum-wage reform thus requires skill-specific, proportional tax changes. Hence, linear tax instruments are sufficient, since a minimum wage is also a linear policy instrument. Of course, any NIN minimum-wage reform under a linear tax system can be perfectly replicated using a non-linear tax system. The same welfare analysis carries over, leading to the same desirability condition for a NIN minimum-wage increase, except that the wedges (Δ_u and Δ_Θ) are determined by the non-linear tax schedule. With social marginal utility of income decreasing in ability, the government would still like to redistribute income from high-skilled to low-skilled workers and from low-skilled workers to non-participants. This implies, again, positive values for Δ_u and Δ_Θ in the optimum, confirming our results.

4.5.5 Comparison with Lee and Saez (2012)

Our findings can be seen to harbor the results of Lee and Saez (2012) as a special case. They analyze the desirability of a minimum wage under the specific assumption that the low-skilled workers who face unemployment are those that have the lowest willingness to pay to remain low-skilled employed. In terms of our model this implies that only individuals with ability Φ or Θ are rationed by a minimum wage, as depicted in the third panel of Figure 4.1. Under this assumption, dubbed 'efficient rationing' by Lee and Saez, a minimum wage is shown to be part of an optimal policy mix if, in the tax optimum without a minimum wage, the marginal social value of income of low-skilled workers exceeds the marginal value of public funds. In the Appendix we demonstrate that, in the special case of 'efficient rationing', this condition is equivalent to our desirability condition, (4.37), evaluated at the optimum without a minimum wage.

Intuitively, if unemployment only hits those workers with the lowest net benefits of remaining employed, the effect of a minimum-wage increase on the allocation of jobs is identical to the effect of a lower low-skilled transfer T^L. A lower transfer leads workers with the lowest willingness to pay to remain employed to decide on their outside option – be it unemployment or high-skilled employment. The only difference then, between a NIN minimum-wage increase and a decrease in the transfer T^L, is that the lower transfer leads to a transfer of resources from low-skilled workers to the government. If, in the optimum without a minimum wage, the net social value of such a transfer from low-skilled workers to the government is negative, then an unambiguous welfare gain can be made by a higher low-skilled transfer, combined with a NIN minimum-wage increase which leaves the original allocation of jobs unaltered. Such a reform would not affect the allocation of jobs, but would redistribute resources from the government to low-skilled workers.

Naturally, this result only holds if the incidence of unemployment is efficient – i.e., if the effect on the allocation of jobs of involuntary unemployment is identical to the effect of a lower low-skilled transfer. For any other rationing schedule the result of Lee and Saez breaks down. The plausibility of efficient rationing in the specific sense of Lee and Saez, however, could be criticized on both theoretical and empirical grounds. Theoretically, there is little to say in favor of any specific assumption on the incidence of unemployment, simply because it is not clear why and how the labor market would discriminate between workers which are identical in all respects but their disutility of work.[17] Moreover, there is no reliable empirical

[17]Furthermore, we have theoretical difficulties with the alleged 'efficiency' of the rationing

evidence on the relationship between unemployment and disutility of work.[18] We try to circumvent these problems in the next section.

4.6 Minimum-wage reform: an empirical application

4.6.1 A sufficient-statistics approach

We have seen, in Proposition 4.1 and Corollary 4.2, that the optimality of a minimum wage, or even the desirability of a minimum-wage increase, depends crucially on the unemployment incidence, I_Θ. A cynic could argue that a minimum wage can always be 'made' optimal by making the appropriate *ad hoc* assumptions on this unobservable incidence. Note, however, that the unemployment incidence enters the desirability condition, (4.37), solely because it partly determines the effect of higher unemployment on skill formation. If we have a measure of this unemployment-education effect, we can use this measure as a *sufficient statistic*, avoiding the need to calibrate I_Θ. In this section we illustrate how such a sufficient-statistics approach might help us to bring the desirability condition to the data.

To calibrate the effect of unemployment on high-skilled employment for a number of countries, we make the assumption that low-skilled workers in our model are individuals that did not complete upper-secondary education. This is a strong assumption: dropping out of secondary school does not condemn one to working for a minimum wage, and upper-secondary education is hardly a guarantee for a job at a higher wage rate. Nevertheless, schooling seems to be an important force

scheme of Lee and Saez. In their model, as in ours, unemployment realizations are made after individuals decide on their skill type. But, once individuals have decided on their skill type, rationing individuals with ability Θ is in fact *most* inefficient as their utility surplus over unemployment is the largest of all low-skilled workers. Once unemployment materializes and individuals can no longer renege on their skill decision, the most efficient rationing would be concentrated solely at ability Φ, rather than at both Φ and Θ.

[18] The only more or less direct evidence of the (in)efficiency of lay-offs due to a minimum wage, and thus indirectly of the relationship between lay-offs and ability, is given by Luttmer (2007). He measures the change in the average (proxy of the) reservation wage of low-skilled workers after an increase in the minimum wage. For two out of four proxies, he finds a statistically significant drop in reservation wages. This could be interpreted as evidence that workers with the highest reservation wages, and thus highest disutility of work, are rationed first. In a sensitivity analysis, however, he finds significant *increases* in reservation wages for the other two proxies, suggesting that rationing is inefficient. There is, however, plenty of evidence on misallocation due to price controls in markets for rental housing, gasoline, and natural gas. See, again, Luttmer (2007) for references.

driving both labor earnings and employment opportunities (e.g., Nickell, 1979; Card, 1999). Moreover, there is a sizable empirical literature on the effect of the low-skilled unemployment rate on enrollment rates for upper-secondary education. Denoting the enrollment rate as e, and the low-skilled unemployment rate as $\bar{u} \equiv \int_{\Phi}^{\Theta} u_\theta \mathrm{d}G(\theta)/(G(\Theta) - G(\Phi))$, such studies generally attempt to measure

$$\eta \equiv \frac{\mathrm{d}e}{\mathrm{d}\bar{u}}, \tag{4.38}$$

by regressing the enrollment rate on the unemployment rate (and various control variables) for a panel of regions.[19] We can use these estimates to avoid making specific assumptions on the unemployment incidence I_Θ. For this, we relate de to the change in the number of high-skilled individuals, $-\mathrm{d}G(\Theta)$, and d$\bar{u}$ to the (weighted) change in unemployment rates $\int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta)$.

First, denote the cohort size of people that are eligible for graduation from upper-secondary education by S. Assume that drop-out rates are negligible.[20] In that case, an increase in the enrollment rate of de increases the number of high-skilled individuals by Sde. Hence, we can write

$$\mathrm{d}e = \frac{-\mathrm{d}G(\Theta)}{S}. \tag{4.39}$$

Second, denote the number of low-skilled individuals as $L \equiv G(\Theta) - G(\Phi)$, and note that d$\bar{u}$ can be written as:[21]

$$\mathrm{d}\bar{u} = \frac{1}{L}\left(\int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta) + (u_\Theta - \bar{u})\mathrm{d}G(\Theta)\right). \tag{4.40}$$

The change in the average low-skilled unemployment rate does not directly correspond to the weighted sum of changes in unemployment probabilities, $\int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta)$, but also incorporates the effect of a change in high-skilled employment on the average unemployment rate, $(u_\Theta - \bar{u})\mathrm{d}G(\Theta)$. This implies that we need to make a

[19] Note that a drop in enrollment rates does not necessarily imply a drop in upper-secondary educational attainment levels, as people who drop out of high-school may decide to enter again later. Card and Lemieux (2001), however, analyze trends in both enrollment rates and educational attainment and conclude that dropping out of high school is, by and large, a once-for-all decision.

[20] By not allowing for drop out in upper-secondary education we overstate the effects of minimum wages on high-skilled employment, so that our results are biased in favor of minimum wages.

[21] We substituted for $\mathrm{d}\Phi = 0$, as a NIN minimum wage reform does not affect the participation margin.

specific assumption on $(u_\Theta - \bar{u})$ to write the desirability condition in terms of η. The smaller is $(u_\Theta - \bar{u})$, the more likely it is that a NIN minimum-wage increase is welfare enhancing. Therefore, we make the highly conservative assumption that $u_\Theta - \bar{u} = -0.5$, which is based on the theoretical minimum for the unemployment rate at the skill margin ($u_\Theta = 0$), and a fifty percent average unemployment rate among the low-skilled workers ($\bar{u} = 0.5$), which can be regarded as an empirical upper bound. The following Lemma establishes how η relates to the desirability condition for a NIN minimum-wage increase.

Lemma 4.3 *Assume that low-skilled workers are workers without upper-secondary education. Furthermore, assume that drop-out rates are negligible. Finally, assume that $u_\Theta - \bar{u} = -0.5$. Then we can rewrite the desirability condition for a NIN minimum-wage increase as:*

$$\eta > \eta^* \equiv \frac{L}{S}\left(\frac{\Delta_u/\Delta_\Theta}{1 + \frac{1}{2}\Delta_u/\Delta_\Theta}\right). \tag{4.41}$$

Proof. Substitute for $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$ in the desirability condition (4.36), by using (4.40). Substitute for $-\mathrm{d}G(\Theta)$ by using (4.39). Finally, substitute $u_\Theta - \bar{u} = -0.5$ and rearrange to obtain (4.41). ■

For given levels of L and S, which are readily available for almost any country, the desirability of a minimum-wage increase is thus seen to depend on three sufficient statistics: the effect of the unemployment rate on school enrollment rates, η, the welfare gain of schooling, Δ_Θ, and the welfare loss of unemployment, Δ_u.[22] Since upper-secondary education is typically completed around the age of 18, S is taken to be the size of the 18-year-old population cohort, which is reported for a number of OECD countries in the first column of Table 1.[23] The second column gives L, the size of the labor force that completed at most primary education. Below we discuss the calibration of the remaining sufficient statistics to determine if a NIN minimum-wage increase could be welfare enhancing.

[22]Our formulation captures the marginal impact of the increase in skilled employment through an increase in the enrollment rate in upper-secondary education. If the increase in the enrollment rate would be permanently higher, the number of high-skilled workers would steadily increase, whereas the number of unemployed workers would correspondingly fall over time. As the number of low-skilled workers diminishes, the marginal cost of more low-skilled unemployment would thus decrease, making minimum wages less harmful over time or maybe even desirable. However, our desirability condition does not permit the analysis of such non-marginal changes; it should be evaluated at the new allocation using updated values for e and L.

[23]The sample is restricted only by the availability of data on Δ_Θ and Δ_u.

4.6.2 The welfare gain of schooling

Following our assumption on skill formation, Δ_Θ measures the public revenue gain from one additional person with an upper secondary educational degree. For a number of countries these revenue gains are provided by OECD (2011*a*, pp. 172-73) and are reported in column 3 of Table 1. The OECD considered revenue gains from higher income taxes and employees' social-security contributions, lower transfers, and higher labor utilization, and the revenue losses from direct costs of financing education and the foregone taxes on earnings associated with education. Gains and losses are calculated over the entire life cycle and discounted at a three percent annual real interest rate to obtain the public net present value of an additional high-skilled worker: Δ_Θ.

4.6.3 The welfare loss of unemployment

As discussed in the previous section, Δ_u consists of both utility losses and public revenue losses associated with unemployment. Unfortunately, we have no empirical approximation for the direct utility losses and thus focus solely on the revenue losses. This implies that our empirical approximation of Δ_u is a potentially severe underestimation of the total welfare costs of unemployment. It also implies that if we find that $\eta < \eta^*$, a NIN minimum-wage decrease is not only welfare enhancing, but also constitutes a Pareto improvement. Such a reform would then lead to higher government revenue, in addition to higher utility for those unemployed individuals that manage to find a job thanks to the lower minimum wage (also see footnote 15).

Statistics on the revenue losses from low-skilled unemployment are extracted from OECD (2011*b*, p. 56). The OECD reports the participation tax rate of an individual moving from short-term unemployment to full-time work at 50 percent of the average wage. These values take into account the losses from lower income taxes and social-security contributions, and higher social, housing, family, and unemployment benefits, together with the gains from lower in-work tax benefits, if applicable. Multiplying these participation tax rates with the average minimum wage income[24], also obtained from the OECD, we calculate a value for Δ_u as shown

[24] Calculated as the minimum wage, relative to the average wage, multiplied by average wage income. For countries without a minimum wage we assume that the low-skilled workers that would become unemployed due to the minimum wage would earn 25 percent of the average wage. This percentage is our sample's lower bound of the minimum wage in terms of the average wage.

in the fourth column of Table 1.[25]

4.6.4 The effect of unemployment on schooling

A recent study by Clark (2011) estimates the impact of the youth unemployment rate among workers aged 18 and 19 on the enrollment rate for 16-year-olds for a sample of English regions between 1975 and 2005. He finds that a one percentage-point increase in the unemployment rate leads to a 0.32 percentage-point increase in the enrollment rate for boys, while for girls the increase amounts to 0.45 percentage point. He also provides a detailed survey of earlier estimates for the United Kingdom and concludes that his estimate is at least twice as large as those found in previous studies. In another recent study on the UK, Tumino and Taylor (2013) find an effect similar to that of Clark, namely an increase in enrollment of 0.48 percentage point.[26]

While most studies on the impact of unemployment on school enrollment focus on the UK, a few studies analyze the relationship for the United States, Spain, and Denmark. For the US, Card and Lemieux (2001) use variations over states and years to estimate the effect of unemployment on enrollment rates, and find that a one percentage-point increase in the unemployment rate raises school enrollment rates of 17-year-olds by 0.40 percentage point. They also determine the effect of the unemployment rate in the state of birth at age 17 on educational attainment, and find that a one percentage-point increase in the unemployment rate leads to a 0.17 percentage-point increase in the share of high-school graduates.[27] In a study on African-American students, Kane (1994) finds the effect to be as large as 0.6.

The disadvantage of the US studies is that data availability confines them to using the prime-age unemployment rate, which is arguably a worse proxy for low-skilled unemployment than the youth unemployment rate used in UK stud-

[25] One might wonder why the life-time value of the increase in tax revenues is taken for Δ_Θ, whereas Δ_u is a one-year cost of higher unemployment. If we would analyze a multiperiod life-cycle model with human capital formation, the marginal cost of a one-year increase in schooling would be the one-year expected forgone net earnings as a low-skilled worker, whereas the marginal benefits would consist of the discounted value of the increase in future net earnings. The corresponding wedge Δ_u would then be equal to the net cost of a one-year unemployment spell, whereas Δ_Θ would capture the discounted value of the net taxes on all future earnings increases.

[26] When splitting their sample in home-owner and non-home-owner families, they find an even larger effect on the enrollment rates of children from non home-owner families.

[27] This estimate is likely to suffer from attenuation bias because of interstate migration. After all, the unemployment rate in the state of birth is not likely to affect the schooling decision of a person who moved to another state. On the basis of interstate migration data, Card and Lemieux suspect this attenuation bias to be in the order of 10-25 percent.

ies. Similar to the UK studies, Petrongolo and San Segundo (2002) analyze the impact of youth unemployment on school enrollment in Spain and find that a one percentage-point increase in the unemployment rate leads to an increase in the enrollment rate for boys of 0.44.

The Appendix gives a further overview of the estimates we discussed. Summing up, the impact of a one percentage-point increase in the unemployment rate on the enrollment rate in upper-secondary education is found to be in a range of 0.1 to 0.6, while all estimates except for one are well below 0.6.

4.6.5 The desirability of a minimum-wage increase

Column 5 in Table 1 provides values of η^*, the right-hand side of the desirability condition (4.41). A NIN minimum-wage increase is only desirable if the effect of the low-skilled unemployment rate on enrollment rates exceeds this critical value, such that $\eta > \eta^*$. If, on the other hand, the effect of unemployment on enrollment is smaller than the value in column 5, a NIN minimum-wage decrease leads to a Pareto improvement. For all countries, values in column 5 range from 0.4 for the United States to 10.3 for Spain.[28]

It is useful to consider the two extreme cases in some more detail. In the United States, a NIN minimum-wage increase is only desirable if a one percentage-point increase in the unemployment rate leads to a higher enrollment rate of at least 0.4 percentage point. At the other extreme, a minimum-wage increase in Spain only enhances welfare if a percentage-point increase in the unemployment rate leads to a higher enrollment rate of at least 10.3 percentage point. The reasons for these differences between the United States and Spain are readily observable from Table 1. For the United States, we see that the public benefits of more workers with secondary education, (Δ_Θ), are relatively large. On top of that, the size of the labor force with only primary education (L) is relatively small, such that a percentage-point increase in the low-skilled unemployment rate is less costly. Spain, on the other hand, shows a relatively small public return to secondary education and a relatively large unskilled population, raising the costs of an increase in the unemployment rate.

On the basis of an empirical calibration of the desirability condition (4.41),

[28] As can be seen from the bottom row of Table 1, in France net tax revenues from a person completing upper-secondary education actually decline. Hence, regardless of the value of η, a NIN minimum-wage decrease leads to a Pareto improvement in France as lower unemployment and lower education both lead to higher public revenue.

Table 4.1: Calibrating the desirability condition

Country	S	L	Δ_Θ	Δ_u	η^*	Minimum wage
	(1)	(2)	(3)	(4)	(5)	(6)
United States	4245	14993	60	8	0.4	Y
Czech Republic	132	344	20	5	0.6	Y
Germany	969	6496	65	8	0.8	N
Hungary	126	585	33	6	0.8	Y
Austria	100	754	65	8	0.9	N
United Kingdom	813	6592	95	11	0.9	Y
Poland	548	1558	9	6	1.3	Y
Sweden	133	913	30	8	1.7	N
Norway	65	537	34	10	2.2	N
Canada	450	2537	25	13	2.3	Y
Denmark	68	766	45	11	2.5	N
Italy	606	9403	37	7	2.5	N
Slovenia	23	145	23	12	2.6	Y
Finland	67	452	16	8	2.7	N
Australia	299	3177	30	10	3.0	Y
Ireland	55	465	33	15	3.1	Y
Portugal	118	3761	43	8	5.3	Y
Spain	465	10213	15	9	10.3	Y
France	833	7327	-6	14	$\Delta_\Theta < 0$	Y

All values 2009 or latest. L and S are measured in thousands; Δ_Θ and Δ_u are measured at 2009 prices, in thousands of PPP equivalent USD. Δ_Θ is an average of male and female values using shares in age-18 cohorts as weights. Δ_u is the unweighted average of revenue losses from an additional unemployed minimum- wage earning single, single parent with 2 children, one-earner married with 2 children, and two-earner married with 2 children with a spouse earning 67 percent of the average wage. All data are available in a separate spreadsheet, available upon request from the authors.

Source: OECD (2011*a*, pp. 172-73), OECD (2011*b*, p. 56), Eurostat and national statistical offices.

we conclude that the effect of unemployment on enrollment should exceed 0.4 to be able to build a case for a higher minimum wage for the United States. That is, a percentage-point increase in the unemployment rate should, *ceteris paribus*, lead to at least a 0.4 percentage-point increase in enrollment rates. As this is well within the range of empirical estimates we found, we cannot reject that a NIN minimum-wage increase might be beneficial for the United States. However, given our very conservative assumptions when calibrating the desirability condition – most importantly the fact that we ignore direct utility losses of unemployment – we do not consider the case for a NIN minimum-wage increase in the United States very strong. For all other countries, a percentage-point increase in the unemployment rate should lead to at least a 0.6 percentage-point increase in enrollment rates. Since this is the upper bound of the empirical estimates, we conclude that the case for a NIN minimum-wage increase is weak for those countries. Instead, a *decrease* of the minimum wage, along with compensating tax changes to keep net wages constant – a NIN minimum-wage decrease – would lead to a Pareto improvement.[29]

4.7 Conclusion

Minimum-wage legislation distinguishes itself from redistributive income taxation by raising employers' labor costs, thereby reducing low-skilled employment. As some low-skilled workers will seek to avoid an increased probability of unemployment by acquiring more skills, a minimum wage leads to both more unemployment and more education. We show that the degree to which a minimum wage leads to additional education rather than unemployment is crucially governed by the unemployment incidence of the minimum wage. If the incidence falls mainly on those low-skilled workers who are relatively inclined towards high-skilled work, the education effect is large relative to the unemployment effect. If the incidence falls mainly on low-skilled workers that are inclined towards non-participation, a minimum wage mainly leads to higher unemployment.

The welfare consequences of a minimum wage are therefore theoretically ambiguous. On the one hand, it leads to lower social welfare as the newly unemployed suffer utility losses and pay less taxes. On the other hand, it leads to social welfare gains as high-skilled workers tend to pay more taxes than low-skilled workers. A

[29]Naturally, such a decrease of the minimum wage is only possible in countries that have a minimum wage. The final column of Table 1 indicates in which countries this is the case.

minimum wage is optimally employed alongside income taxation if and only if the gains from more education outstrip the losses from higher unemployment. This is the case only if the incidence of rationing falls to a large enough extent on low-skilled workers inclined towards high-skilled work, such that a given increase in the unemployment rate leads to enough additional education.

We used data on the net revenue gains and losses of education and unemployment to calibrate the desirability condition of a net-income-neutral minimum-wage increase. Naturally, the results from this exercise should be interpreted with caution. While at several junctions we have deliberately biased our empirical calibration in favor of a minimum wage, we did not take into account the effect that labor-market frictions might have on the desirability of a minimum wage. That being said, we find that for a one percentage-point increase of the low-skilled unemployment rate, we need an increase of school enrollment rates of around half a percentage-point for a minimum-wage increase to be desirable in the United States. This is well within empirical estimates of this effect and, accordingly, we cannot reject the hypothesis that a net-income-neutral minimum wage increase might be welfare enhancing in the United States. For any other country under consideration, the required unemployment-education effect should be well above empirical estimates of this effect. For these countries, the policy recommendation is to decrease the minimum wage, while adjusting income taxes to offset effects on net wages. Our model predicts that such a reform leads to a Pareto improvement: it leads to (i) additional government revenue due to lower unemployment, exceeding the loss in government revenue due to lower education, and (ii) higher utility for those formerly unemployed individuals that can find a job due to the lower minimum wage.

4.A Proof of Proposition 4.1

The change in welfare, $\mathrm{d}\mathcal{W}$, in response to a NIN minimum-wage increase, $\mathrm{d}w^L > 0$, is given by taking the derivative of welfare $\mathcal{W}$, and substituting for $\mathrm{d}\omega^i = \mathrm{d}T^i = \mathrm{d}\Phi = 0$:

$$\mathrm{d}\mathcal{W} = -\left(v^H(\omega^H, T^H, \Theta) - v^{EL}(\omega^L, T^L, T^U, u_\Theta, \Theta)\right)\mathrm{d}G(\Theta) + \int_\Phi^\Theta v_u^{EL}(\omega^L, T^L, T^U, u_\theta, \theta)\mathrm{d}u_\theta \mathrm{d}G(\theta). \tag{4.42}$$

The individual's first-order condition (4.11) for the skill choice implies that the first line is zero. Furthermore, use equation (4.8) to establish that: $v_u^{EL}(\omega^L, T^L, T^U, u_\theta, \theta) = -v^L(\omega^L, T^L, \theta) + V(T^U, 0)$. Substituting this result into equation (4.42), we get:

$$\mathrm{d}\mathcal{W} \equiv -\int_\Phi^\Theta (v^U - v_\theta^L)\mathrm{d}u_\theta \mathrm{d}G(\theta). \tag{4.43}$$

The change in the government budget, $\mathrm{d}\mathcal{B}$, is obtained by taking derivatives of $\mathcal{B}$, and substituting for $\mathrm{d}\omega^i = \mathrm{d}T^i = \mathrm{d}\Phi = 0$. Rearranging then yields:

$$\begin{aligned}\mathrm{d}\mathcal{B} = &\int_\Theta^{\bar\theta} l_\theta^H \mathrm{d}G(\theta)\mathrm{d}w^H + \int_\Phi^\Theta (1-u_\theta) l_\theta^L \mathrm{d}G(\theta)\mathrm{d}w^L \\ &- ((w^H - \omega^H) l_\Theta^H - T^H - (1-u_\Theta)\left((w^L - \omega^L) l_\Theta^L - T^L\right) + u_\Theta T^U)\mathrm{d}G(\Theta) \\ &- \int_\Phi^\Theta ((w^L - \omega^L) l_\theta^L - T^L + T^U)\mathrm{d}u_\theta \mathrm{d}G(\theta).\end{aligned} \tag{4.44}$$

From equation (4.21) we know that:

$$\mathrm{d}w^H = -\frac{L^L}{L^H}\mathrm{d}w^L = -\frac{\int_\Phi^\Theta (1-u_\theta) l_\theta^L \mathrm{d}G(\theta)}{\int_\Theta^{\bar\theta} l_\theta^H \mathrm{d}G(\theta)}\mathrm{d}w^L. \tag{4.45}$$

This eliminates the first two terms in equation (4.44). Furthermore, substitute for $(w^i - \omega^i) = t^i w^i$ to obtain:

$$\begin{aligned}\mathrm{d}\mathcal{B} = &- (t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U)\mathrm{d}G(\Theta) \\ &- \int_\Phi^\Theta (t^L w^L l_\theta^L - T^L + T^U)\mathrm{d}u_\theta \mathrm{d}G(\theta).\end{aligned} \tag{4.46}$$

Substituting equations (4.43) and (4.46) into condition (4.33), we get:

$$\begin{aligned}&-\int_\Phi^\Theta \left(\frac{v^U - v_\theta^L}{\lambda} + t^L w^L l_\theta^L - T^L + T^U\right)\mathrm{d}u_\theta \mathrm{d}G(\theta) \\ &-(t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U)\mathrm{d}G(\Theta) > 0.\end{aligned} \tag{4.47}$$

Using the definition of the unemployment incidence in equation (4.22) to substitute for $g(\theta)\mathrm{d}u_\theta = I_\theta \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$ yields:

$$(4.48)\qquad -\int_\Phi^\Theta \left(\frac{v^U - v_\theta^L}{\lambda} + t^L w^L l_\theta^L - T^L + T^U\right) I_\theta \mathrm{d}\theta \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$$
$$-(t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U)\mathrm{d}G(\Theta) > 0.$$

Finally, substituting the wedges Δ_Θ and Δ_u from equations (4.34) and (4.35) we obtain the final result of equation (4.36).

4.B Proof of Lemma 4.2

Without a minimum wage and, hence, without involuntary unemployment, social welfare is given by:

$$(4.49)\quad \mathcal{W} \equiv \int_\Theta^{\overline{\theta}} v^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + \int_\Phi^\Theta v^L(\omega^L, T^L, \theta)\mathrm{d}G(\theta) + G(\Phi)V(T^U, 0).$$

The government's budget constraint is given by $\mathcal{B}$:

(4.50)

$$\mathcal{B} \equiv \int_\Theta^{\overline{\theta}} (w^H - \omega^H) l^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + \int_\Phi^\Theta (w^L - \omega^L) l^L(\omega^L, T^L, \theta)\mathrm{d}G(\theta)$$
$$- (1 - G(\Theta))T^H - (G(\Theta) - G(\Phi))T^L - G(\Phi)T^U - R = 0.$$

Defining λ as the Lagrange multiplier of the budget constraint, we can set up the following maximization problem for the government:

$$(4.51)\qquad \max_{\{T^H, T^L, T^U, \omega^H, \omega^L\}} \mathcal{L} = \mathcal{W} + \lambda \mathcal{B},$$

which is subject to Θ, Φ, w^H, and w^L as determined by the individuals' and firms' first-order conditions in equations (4.6), (4.10), (4.12), and (4.20). Notice again that we chose net wages ω^i, rather than tax rates t^i, as the government's control variables.

The first-order conditions of this maximization problem are obtained by equating the partial derivatives of the Lagrangian to zero. These derivatives are given by:

(4.52)
$$\frac{\partial \mathcal{L}}{\partial \omega^H} = \int_{\Theta}^{\overline{\theta}} \left(V_{\theta,c}^H - \lambda \left(1 - \Delta_H \frac{\omega^H}{l_\theta^H} \frac{\partial l_\theta^H}{\partial \omega^H} \right) \right) l_\theta^H \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}\omega^H} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}\omega^H},$$

(4.53)
$$\frac{\partial \mathcal{L}}{\partial T^H} = \int_{\Theta}^{\overline{\theta}} \left(V_{\theta,c}^H - \lambda \left(1 - \Delta_H \omega^H \frac{\partial l_\theta^H}{\partial T^H} \right) \right) \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}T^H} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}T^H},$$

(4.54)
$$\frac{\partial \mathcal{L}}{\partial \omega^L} = \int_{\Phi}^{\Theta} \left(V_{\theta,c}^L - \lambda \left(1 - \Delta_L \frac{\omega^L}{l_\theta^L} \frac{\partial l_\theta^L}{\partial \omega^L} \right) \right) l_\theta^L \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}\omega^L} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}\omega^L},$$

(4.55)
$$\frac{\partial \mathcal{L}}{\partial T^L} = \int_{\Phi}^{\Theta} \left(V_{\theta,c}^L - \lambda \left(1 - \Delta_L \omega^L \frac{\partial l_\theta^L}{\partial T^L} \right) \right) \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}T^L} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}T^L},$$

(4.56)
$$\frac{\partial \mathcal{L}}{\partial T^U} = G(\Phi)(V_c^U - \lambda) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}T^U} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}T^U}.$$

We took a number of steps to arrive at these expressions. First, observe that changes in Θ and Φ have no direct effect on individuals' utility, but only affect welfare indirectly through their effect on the government budget. This is because individuals at Θ (Φ) are indifferent between low-skilled work and high-skilled work (voluntary unemployment). Second, we applied Roy's identity to rewrite derivatives of indirect utility in terms of direct utility, i.e., $v_{\theta,\omega^i}^i = l_\theta^i v_{\theta,T^i}^i$ and $v_{\theta,T^i}^i = V_{\theta,c}^i$, where $V_{\theta,c}^i$ gives the marginal utility of consumption of an individual with skill type i and ability θ. Third, from the firm's first-order conditions we substituted for $L^H \mathrm{d}w^H = -L^L \mathrm{d}w^L$, which implies that the effect on the government budget of any increase in one of the two gross wage rates is exactly offset by a decrease in the other gross wage rate. This is the reason $\mathrm{d}w^H$ and $\mathrm{d}w^L$ do not show up in the expressions. Fourth, and final, we defined wedges as follows: $\Delta_H \equiv \frac{t^H}{1-t^H}$ and $\Delta_L \equiv \frac{t^L}{1-t^L}$ are the wedges on high- and low-skilled net labor earnings; $\Delta_\Phi \equiv t^L w^L l^L(\Phi) - T^L + T^U$ is the wedge on the participation margin; $\Delta_\Theta \equiv t^H w^H l^H(\Theta) - T^H - t^L w^L l^L(\Phi) + T^L$ is the wedge on the skill margin. In other words, Δ_i, $i \in \left\{L^H, L^L, \Phi, \Theta\right\}$ denote the gains in government revenue from higher levels of $\omega^H l_\theta^H$ and $\omega^L l_\theta^L$, and lower levels of $G(\Phi)$ and $G(\Theta)$. Notice that Δ_Θ corresponds to (4.34) in the absence of a minimum wage and unemployment.

We can now rewrite the first-order conditions for each control variable. To facilitate the derivations, we adopt a number of additional notational conventions. We define uncompensated net-wage elasticities of intensive labor supply as $\varepsilon^{i,u}_{\theta,\omega^i} \equiv \frac{\partial l^i_\theta}{\partial \omega^i}\frac{\omega^i}{l^i_\theta}$, and income elasticities as $\varepsilon^i_{\theta,T^i} \equiv -\omega^i \frac{\partial l^i_\theta}{\partial T^i} > 0$. Applying the Slutsky equation we can write the compensated net-wage elasticity of labor supply as $\varepsilon^{i,c}_{\theta,\omega^i} = \varepsilon^{i,u}_{\theta,\omega^i} + \varepsilon^i_{\theta,T^i}$. We define the income-weighted average compensated elasticities of labor supply as:

$$\text{(4.57)} \qquad \overline{\varepsilon^L_{\omega^L}} \equiv \int_\Phi^\Theta \frac{w^L l^L_\theta}{w^L L^L}\varepsilon^{L,c}_{\theta,\omega^L}\mathrm{d}G(\theta) > 0, \quad \overline{\varepsilon^H_{\omega^H}} \equiv \int_\Theta^{\overline{\theta}} \frac{w^H l^H_\theta}{w^H L^H}\varepsilon^{H,c}_{\theta,\omega^H}\mathrm{d}G(\theta) > 0,$$

where the signs follow from the fact that compensated wage elasticities of labor supply are always positive.

The semi-elasticities of participation and skill formation are defined as follows:

(4.58)
$$\eta_{T^U} \equiv \frac{\mathrm{d}G(\Phi)}{G(\Phi)\mathrm{d}T^U}, \quad \eta_{T^L} \equiv \frac{-\mathrm{d}G(\Phi)}{(G(\Theta)-G(\Phi))\,\mathrm{d}T^L}, \quad \eta_{\omega^L} \equiv \frac{-\mathrm{d}G(\Phi)}{L^L\mathrm{d}\omega^L},$$

(4.59)
$$\zeta_{T^L} \equiv \frac{\mathrm{d}G(\Theta)}{(G(\Theta)-G(\Phi))\,\mathrm{d}T^L}, \ \zeta_{T^H} \equiv \frac{-\mathrm{d}G(\Theta)}{(1-G(\Theta))\,\mathrm{d}T^H}, \ \zeta_{\omega^L} \equiv \frac{\mathrm{d}G(\Theta)}{L^L\mathrm{d}\omega^L}, \ \zeta_{\omega^H} \equiv \frac{-\mathrm{d}G(\Theta)}{L^H\mathrm{d}\omega^H}.$$

The term η_j measures the change in $G(\Phi)$ due to a marginal change of $j \in \{T^U, T^L, \omega^L\}$, and ζ_k measures the change in $G(\Theta)$ due to a marginal change of $k \in \{T^L, \omega^L, T^H, \omega^H\}$. All semi-elasticities are defined to be positive.

Following Diamond (1975), we define γ^i_θ as the social marginal value of income for an individual with ability $\theta \in [\underline{\theta}, \overline{\theta}]$ and skill level $i \in \{H, L\}$. This term consists of the private marginal utility of income, minus the social value of the loss in tax revenue due to the income effect on labor hours. Normalizing in terms of resources, by dividing by λ, this yields:

$$\text{(4.60)} \qquad \gamma^U \equiv \frac{V^U_c}{\lambda}, \quad \gamma^i_\theta \equiv \frac{V^i_{\theta,c}}{\lambda} + \Delta_i\omega^i\frac{\partial l^i_\theta}{\partial T^i}, \quad \theta \in [\underline{\theta}, \overline{\theta}], \quad i \in \{H, L\}.$$

The average values for the social marginal value of income for high- and low-skilled workers are given by:

$$\text{(4.61)} \qquad \overline{\gamma^L} \equiv \frac{\int_\Phi^\Theta \gamma^L_\theta \mathrm{d}G(\theta)}{G(\Theta)-G(\Phi)}, \quad \overline{\gamma^H} \equiv \frac{\int_\Theta^{\overline{\theta}} \gamma^H_\theta \mathrm{d}G(\theta)}{1-G(\Theta)}.$$

Finally, we define the distributional characteristics of the income tax bases as ξ^i

(cf. Atkinson and Stiglitz, 1980, p.388):

$$\xi^L \equiv 1 - \frac{(G(\Theta) - G(\Phi)) \int_\Phi^\Theta w^L l_\theta^L \gamma_\theta^L \mathrm{d}G(\theta)}{\int_\Phi^\Theta w^L l_\theta^L \mathrm{d}G(\theta) \int_\Phi^\Theta \gamma_\theta^L \mathrm{d}G(\theta)} = -\frac{\operatorname{cov}\left[w^L l_\theta^L, \gamma_\theta^L\right]}{\overline{w^L l^L \gamma^L}} \tag{4.62}$$

$$\xi^H \equiv 1 - \frac{(1 - G(\Theta)) \int_\Theta^{\overline{\theta}} w^H l_\theta^H \gamma_\theta^H \mathrm{d}G(\theta)}{\int_\Theta^{\overline{\theta}} w^H l_\theta^H \mathrm{d}G(\theta) \int_\Theta^{\overline{\theta}} \gamma_\theta^H \mathrm{d}G(\theta)} = -\frac{\operatorname{cov}\left[w^H l_\theta^H, \gamma_\theta^H\right]}{\overline{w^H l^H \gamma^H}} \tag{4.63}$$

That is, ξ^i gives the negative of the normalized covariance between gross labor income and the social marginal value of income for skill type i. The larger is the term ξ^i, the more the social welfare weight γ_θ^i is decreasing with income $w^i l_\theta^i$, making the labor-income tax a more attractive instrument to redistribute income. Due to the normalization, the distributional term takes on a value between zero and one: $\xi^i \in [0, 1]$.

This, then, allows us to rewrite the first-order conditions for T^U, T^L, ω^L, T^H, and ω^H, respectively, as follows:

$$\gamma^U = 1 + \Delta_\Phi \eta_{T^U}, \tag{4.64}$$

$$\overline{\gamma^L} = 1 - \Delta_\Phi \eta_{T^L} + \Delta_\Theta \zeta_{T^L}, \tag{4.65}$$

$$(1 - \xi^L)\overline{\gamma^L} = 1 - \Delta_L \overline{\varepsilon_{\omega^L}^L} - \Delta_\Phi \eta_{\omega^L} + \Delta_\Theta \zeta_{\omega^L}, \tag{4.66}$$

$$\overline{\gamma^H} = 1 - \Delta_\Theta \zeta_{T^H}, \tag{4.67}$$

$$(1 - \xi^H)\overline{\gamma^H} = 1 - \Delta_H \overline{\varepsilon_{\omega^H}^H} - \Delta_\Theta \zeta_{\omega^H}. \tag{4.68}$$

Notice that if social welfare weights are decreasing in ability θ, we can write $\gamma^U > \overline{\gamma^L} > \overline{\gamma^H}$. This allows us to combine the first, second and fourth first-order conditions to find:

$$\Delta_\Phi \left(\eta_{T^U} + \eta_{T^L}\right) > \Delta_\Theta \zeta_{T^L}, \tag{4.69}$$

$$\Delta_\Phi \eta_{T^U} > -\Delta_\Theta \zeta_{T^H}, \tag{4.70}$$

$$\Delta_\Theta \left(\zeta_{T^L} + \zeta_{T^H}\right) > \Delta_\Phi \eta_{T^L}. \tag{4.71}$$

From the first line we establish that if $\Delta_\Phi < 0$, then $\Delta_\Theta < 0$. However, from the second line we see that if $\Delta_\Phi < 0$, then $\Delta_\Theta > 0$. This is a contradiction. Thus, in the optimum we must have $\Delta_\Phi > 0$. Notice, from the definition of Δ_u in equation (4.35), that if $\Delta_\Phi > 0$, we must necessarily have that $\Delta_u > 0$. From the third line, it immediately follows that $\Delta_\Theta > 0$. Combining the remaining first-order conditions, it is relatively straightforward to show that the optimal wedges on high- and low-skilled labor earnings are positive as well: $\Delta_H > 0$ and $\Delta_L > 0$. Furthermore, since individuals' marginal utility of income is strictly positive, we can readily deduce from the first-order conditions that all wedges are finite. This proves Lemma 4.2.

4.C Efficient rationing: Lee and Saez (2012)

We derive the desirability condition for the minimum wage when rationing is 'efficient', i.e., when the involuntary unemployed are those that have the smallest benefits of being low-skilled employed. In this case, the rationing schedule is no longer continuous. Instead, rationing is concentrated on the extremes of the low-skilled ability distribution, Φ and Θ. Individuals with ability Φ or Θ are indifferent between low-skilled employment and their outside option – voluntary unemployment for Φ, high-skilled employment for Θ. In response to a NIN minimum-wage reform, $\mathrm{d}w^L$, the change in welfare is therefore nil: $\mathrm{d}\mathcal{W} = 0$. Hence, the desirability condition of a NIN minimum-wage reform simplifies to $\mathrm{d}\mathcal{B} > 0$.

In the case of efficient rationing, the minimum-wage induced increase in high-skilled employment, relative to the increase in unemployment, is identical to the increase in high-skilled employment, relative to the increase in voluntary unemployment, due to a decrease in the low-skilled transfer T^L. In response to a NIN minimum-wage reform, $\mathrm{d}w^L$, this yields:

$$\text{(4.72)} \qquad \frac{-\mathrm{d}G(\Theta)/\mathrm{d}w^L}{\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)/\mathrm{d}w^L} = \frac{-\mathrm{d}G(\Theta)/\mathrm{d}T^L}{\mathrm{d}G(\Phi)/\mathrm{d}T^L}.$$

Since only individuals with ability Φ or Θ are rationed, there is no direct utility loss of unemployment and the loss of tax revenue is solely determined by the earnings of a worker with ability Φ. Hence, we can write Δ_u as:

$$\text{(4.73)} \qquad \Delta_u = \Delta_\Phi \equiv t^L w^L l_\Phi^L - T^L + T^U,$$

where Δ_Φ is the wedge on participation. Substituting above two equations, together with the elasticities as defined in (4.58) and (4.59), into the desirability condition for a NIN minimum-wage increase, (4.36), we get:

$$\text{(4.74)} \qquad \frac{\zeta_{T^L}}{\eta_{T^L}} > \frac{\Delta_\Phi}{\Delta_\Theta}.$$

The values for the wedges on participation and skill formation, Δ_Φ and Δ_Θ, in the tax optimum without a minimum wage, are given by the first-order condition (4.65). Substituting this condition into above desirability condition yields:

$$\overline{\gamma^L} > 1. \tag{4.75}$$

Here, $\overline{\gamma^L}$ is the low-skilled average marginal social value of income as defined in the proof of Lemma 4.2. The term $\overline{\gamma^L}$ exactly corresponds with g_1 in Lee and Saez (2012). Hence, we can see that our desirability condition is equivalent to their Proposition 2. A binding minimum wage is optimal if and only if the social marginal welfare weight of low-skilled workers ($\overline{\gamma^L}$ or g_1) exceeds unity. In the case of efficient rationing, the only difference between a NIN minimum-wage increase and a low-skilled transfer decrease, is that the minimum-wage increase does not redistribute away from the low-skilled. Hence, if redistribution away from the low-skilled is socially costly, a binding minimum wage is part of the policy optimum.

4.D The effect of unemployment on school enrollment

Table 4.A1 gives an overview of empirical studies on the impact of unemployment on school enrollment, as discussed in this chapter. Earlier UK evidence is surveyed in a similar overview by Clark (2011), in his Table 1. The first column indicates the study of interest; the second column indicates the country of analysis; the third column indicates the time span of the analysis; the fourth column indicates whether the schooling variable refers to enrollment rates (E), or high-school graduation rates (G), whether it refers to boys (b), girls (g), or both (bg), and the age group under consideration; the fifth column indicates to which age-group the unemployment variable refers; the final column gives the estimate of η.

Table 4.A1: Empirical estimates of η

Study	Country	Time	Schooling	Unemployment	η
Clark (2011)	UK	1975-2005	E, b, 16 y/o	18-19 y/o	0.32
			E, g, 16 y/o	18-19 y/o	0.45
Tumino and Taylor (2013)	UK	1991-2009	E, bg, 16 y/o	16-21 y/o	0.48
Kane (1994)	US	1973-1988	G, bg, 18-19 y/o	Total	0.60
Card and Lemieux (2001)	US	1968-1996	E, bg, 15-16 y/o	25-54 y/o	0.14
			E, bg, 17 y/o	25-54 y/o	0.40
		1954-1964	G, b	25-54 y/o	0.17
			G, g	25-54 y/o	0.18
Petrongolo and San Segundo (2002)	ES	1991	E, b, 16-17 y/o	16-24 y/o	0.44

The column on schooling indicates whether the dependent variable was the enrollment rate (E) or the high-school graduation rate (G), whether it concerned boys (b), girls (g), or both (bg), and the age group to which the schooling variable refers. Note that Kane (1994) uses the graduation-rate of blacks. Estimates of Clark (2011) are found in his Tables 2 and 3 on pages 533-534. Estimates for Tumino and Taylor (2013) are found in their Table 3 on page 22. Estimates for Kane (1994) are found in his text, page 890. Estimates for Card and Lemieux (2001) are found in their table 9.4 on page 467 and table 9.6 on page 471. Estimates for Petrongolo and San Segundo (2002) are found in their text, page 364.

Chapter 5

Equity and Efficiency in Rationed Labor Markets

5.1 Introduction

It is well-understood that labor market policies and institutions that fix wage rates above their market-clearing level lead to an inefficient allocation of production factors.[1] More people would like to work for prevailing wages than there are jobs, causing involuntary unemployment. It is less well-understood that such labor-market rationing entails two different types of inefficiency. First, there are potential labor market transactions which carry a positive surplus for both worker and firm, which are not executed due to inflexible wages. Second, without a secondary market for jobs, the labor market mechanism cannot ensure that the limited amount of jobs are allocated to the persons with the highest utility of work. That is, the market does not discriminate between two persons with identical productivity but different levels of participation costs or, similarly, different levels of reservation wages. In stark contrast with the first source of inefficiency, the second has received little attention. If the way in which labor markets are rationed is indeed inefficient, public policy that affects the degree of rationing – obviously minimum wage legislation, but also participation policy and taxation – requires reappraisal. This chapter is an attempt to fill this gap by providing a theoretical analysis of the implications for optimal government policies of inefficiently rationed labor markets.

The main idea of this chapter is best illuminated by considering a simple economy in which one type of labor is the only factor of production. Workers differ with

[1] This chapter is based on Gerritsen (2013*a*)

respect to their costs of participation, but are otherwise identical. If wages cannot adjust to equate demand and supply, and thus if firms and workers cannot freely transact, the number of jobs will be fixed below the number of potential workers. As firms have no incentive to take into account workers' costs of participation, and if workers cannot freely exchange jobs among themselves, some of the employed are likely to have higher costs of participation than some of the unemployed. The first inefficiency originates from the absence of free job exchange between firms and the unemployed, the second from the absence of free exchange between workers and the unemployed.

In such case, government would find it optimal to increase the relative rewards of being voluntary unemployed by increasing unemployment benefits, financed by higher labor income taxes. Workers that derive least utility from their job would decide to quit and reap the increased benefits of being unemployed. In clearing labor markets, this decision entails an efficiency loss as aggregate employment, and thus the tax base, would decline. In rationed labor markets there is no efficiency cost because aggregate employment will not fall as there were more potential workers than jobs in the first place. Instead there is an efficiency gain since, by giving up their job, workers who derive little utility from working create jobs for unemployed people who derive more utility from working. Hence, government can correct for the absence of a job market between workers and the unemployed by appropriately setting taxes. As a corollary, it is suboptimal to increase the degree of rationing by implementing a binding minimum wage, or to stimulate participation in rationed labor markets. Moreover, since the proposed tax reform entails transfers from the employed towards the unemployed, it improves equity as well as efficiency, such that its optimality is robust under redistributive social preferences.

I formally derive these results for the simple single-labor-type model with fixed wages, and consequently determine to what extent the policy implications survive in a richer context with additional types of labor, an endogenous wage floor, and different types of tax instruments.

First, results do not necessarily carry over to a richer general equilibrium framework with both skilled and unskilled labor as factors of production. If only the unskilled segment of the labor market is rationed, while the skilled segment clears, involuntary unemployment among the unskilled might encourage people to invest to become skilled workers. Increasing unskilled taxes and unemployment benefits to reduce the inefficiency of rationing now also affects people's skill decisions. The policy reform's effect on skill formation is theoretically ambiguous: the increased tax burden on unskilled workers increases the relative attractiveness of skilled labor, whereas the decline in unemployment makes unskilled labor supply more attractive. If the reform leads to more skilled labor supply, it can be seen to improve equity by transferring resources from the employed to the unemployed, efficiency because it eliminates inefficient rationing and because the increase in skilled labor supply implies an expansion of the tax base.

I derive a condition under which the policy reform that raises unskilled taxes and unemployment benefits indeed leads to more skilled labor supply. A crucial ingredient of this condition is how rationing depends on individuals' underlying costs of participation. The policy reform is more likely to lead to higher skilled labor supply if rationing is mainly concentrated among the unskilled with high costs of participation, i.e., those close to the margin between voluntary unemployment and unskilled employment. In such case, lower unemployment would not affect skill formation and it would be optimal to increase unskilled taxes and unemployment benefits to substitute voluntary unemployment for involuntary unemployment. If, on the other hand, rationing is concentrated among those unskilled workers who are relatively indifferent between being skilled or unskilled, decreased rationing leads to less skill formation as skilled individuals reap the benefits of improved probabilities of obtaining an unskilled job. Only in that case might rationing be part of a policy optimum. This conclusion complements the studies by Lee and Saez (2012) and Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation) who find that under such type of rationing it might be optimal to impose binding minimum wage legislation.

Second, the derived policy implications remain valid if the above-market-clearing wage rate is endogenously determined by a labor union. The intuition behind this result is that with endogenous wages, government can directly affect wage rates, and thus employment, through its income tax policies. I assume the union sets wages to maximize a goal function which features net wages and employment as complements. In that case it is true that an exogenous increase in the income tax rate leads unions to demand higher, compensatory, wages. However, regardless of the *level* of taxes, if government commits to raising them as a response to increased involuntary unemployment, it incentivizes the union to moderate its wage demands. After all, the union knows that higher wage demands now leads to increased taxes. Thus, the conclusion that labor income taxes should depend positively on the rate of unemployment carries over to the case of endogenous union wages.

Third, I show that the logic of the derived policy implications is only applicable to employees' income taxation – i.e., supply-side taxation. Higher income for employees leads to lower labor supply, but keeps employment constant since supply exceeds demand. However, higher labor income taxes for employers lead to higher labor costs, and thereby to lower labor demand. Since there was already too little demand, higher employer taxes lead to even less employment and higher unemployment. Thus, in rationed labor markets, the conventional incidence equivalence of supply- and demand-side labor income taxes no longer holds. Instead, increased employee taxes lead to efficiency gains, whereas increased employer taxes lead to efficiency losses.

An important ingredient of my analysis is the question to whom the limited amount of jobs is allocated. Unfortunately, the literature on involuntary unemployment provides little guidance when it comes to the efficiency of labor market rationing.[2] In general, in the absence of a secondary or "black" market in which a rationed good is traded, there is little reason to assume the rationed goods are acquired by the individuals who desire them most (Tobin, 1952). Empirically, as noted by Luttmer (2007), this has been confirmed by studies on the U.S. residential market for gas (Davis and Kilian, 2011), the gasoline market (Deacon and Sonstelie, 1989; Frech and Lee, 1987) and on the housing rental market (Glaeser and Luttmer, 2003). As there is no secondary market for jobs, it is unlikely that labor market rationing is efficient. This point has first been made by Lott (1990), and more recently repeated by Palda (2000), Luttmer (2007), Lee and Saez (2012), and Gerritsen and Jacobs (2013*a*,*b*, Chapters 3 and 4 of this Dissertation). Empirical evidence on the efficiency of labor market rationing is virtually non-existent. The only more or less direct evidence for the efficiency of lay-offs due to an increase in rationing is given by Luttmer (2007) who analyzes the effects of an increase in the minimum wage. He measures the change in the average (proxy of the) reservation wage of low-skilled workers after an increase in the minimum wage. He finds conflicting results for different reservation wage proxies and can therefore not find convincing evidence regarding the efficiency of labor market rationing. This chapter builds on these earlier studies in an attempt to provide more insight into optimal government policies, given that labor market rationing is not necessarily

[2] This is in stark contrast to the first source of efficiency, represented by the aggregate employment effects of above-market-clearing wage rates. See Neumark and Wascher (2006) for an extensive survey documenting empirical studies on the employment loss due to minimum wages and the literature survey in Gerritsen and Jacobs (2013*b*, Chapter 3 of this Dissertation) for theoretical studies on the welfare implications of minimum wages.

efficient.

In the next section I show that, in a simple model of labor market rationing, higher labor taxes improve both efficiency and equity. In Section 5.3 I derive conditions for which this still holds in the case of multiple labor types. In Section 5.4 I discuss endogenous wage setting and the differences in incidence of employee and employer taxes. I conclude with a discussion.

5.2 A simple model of rationing

5.2.1 Model and policy implications

Below I present a highly stylized partial equilibrium model of a rationed labor market to demonstrate the basic narrative of the chapter. To keep things as simple as possible, I assume there is a fixed amount of available jobs $\bar{m}$. For now, I leave the exact reason for this demand constraint out of consideration. I merely note that one possible explanation could be a combination of fixed wages and a technology featuring decreasing returns to labor. I denote the number of individuals by n, and assume that there are more individuals then jobs: $n > \bar{m}$. Since there are more potential workers than jobs, not every individual will succeed in securing a job. Ultimately, there will be $n_L = \bar{m}$ employed individuals and n_U unemployed individuals (be it voluntarily or involuntarily), such that $n_L + n_U = n$. For ease of expression I normalize the total number of individuals to one: $n = 1$.

Individuals are heterogeneous with respect to their disutility of work, c. This is distributed according to $G(c)$, with density $g(c) \equiv G'(c)$ and support $[0, \bar{c}]$, and can be seen as a combination of monetary costs associated with work (e.g. travel costs, costs of education, child care) and loss of leisure. An individual, if employed, inelastically supplies one unit of labor. He then earns after-tax income $w - t_L$, with w the wage rate and t_L a tax on income (if positive) or a government transfer (if negative), and suffers disutility of work. If unemployed, he does not suffer any work-related disutility and earns unemployment benefits $-t_U$. Government is assumed to be unable to distinguish the voluntary unemployed from the involuntary unemployed, such that there is no distinction in transfers for the two different types of unemployment. I assume utility is linear in consumption and disutility of work, yielding the following utility functions for employed and unemployed persons[3]:

$$V_c^L = w - t_L - c, \tag{5.1}$$

$$V_U = -t_U. \tag{5.2}$$

Variables that, in equilibrium, depend on disutility of work I denote with a subscript c.

An individual notionally supplies labor if $V_c^L > V_U$. Hence, there is a critical level of disutility, c^*, leaving an individual indifferent between working and not working. Every individual with disutility higher than c^* prefers to be unemployed, every individual with lower disutility prefers to be employed. This critical level is determined by $V_{c^*}^L = V_U$:

$$c^* = w - t_L + t_U. \tag{5.3}$$

I furthermore assume that absent taxation, i.e., in the case that $c^* = w$, labor supply exceeds labor demand such that $G(w) > \bar{m}$. This requires the adoption of a certain *rationing schedule* that prescribes which of the $G(c^*)$ individuals that would like to work obtain a job and which do not. I therefore define u_c as the probability of unemployment for an individual with disutility c. The density-weighted sum of individual unemployment probabilities should add up to the aggregate number of unemployed, that is: $\int_0^C u_c g(c)\mathrm{d}c = G(c^*) - \bar{m}$. Hence, in equilibrium I can write a person's unemployment probability as a function of the critical level c^*, the number of jobs $\bar{m}$, and his disutility of work c: $u_c = u(c^*, \bar{m}, c)$.

Lee and Saez (2012) and Gerritsen and Jacobs (2013*b*, Chapter 3 of this Dissertation) discuss two special cases of rationing in the light of minimum wage policies: efficient rationing and uniform rationing. Rationing is efficient if workers with the highest costs of participation become unemployed first, such that $u_c = 0$ for $c \in [0, c_{er}]$ and $u_c = 1$ otherwise, with $G(c_{er}) = \bar{m}$. Rationing is uniform if every worker faces the same probability of unemployment such that $u_c = \bar{u}$ for all $c \in [0, c^*]$. Since there is little reason to suspect either case is a reasonably accurate description of reality I intend not to impose too stringent assumptions on the rationing schedule.

Government's task is to collect taxes and pay out benefits, and finance some exogenous revenue requirement r, such that the following budget constraint is

[3] As explained below, the policy implications drawn in this section hold *a fortiori* for any concave transformation of these utility functions.

satisfied:

$$\mathcal{B} \equiv n_L t_L + n_U t_U - r = 0. \tag{5.4}$$

Any positive outlays to transfers or exogenous expenditures need to be offset by tax revenue. I assume, for now, that social preferences are utilitarian such that government, given linear utility functions, does not care about the distribution of income. Social welfare is therefore given by the simple sum of all individuals' utilities:[4]

$$\mathcal{W} \equiv \int_0^{c^*} (1 - u_c)(w - t_L - c)\mathrm{d}G(c) - n_U t_U. \tag{5.5}$$

By substituting in the budget constraint and $n_L = \bar{m}$, and after some rearranging, I obtain:

$$\mathcal{W} = \bar{m}w - r - \int_0^{c^*} (1 - u_c)c\mathrm{d}G(c), \quad G(c^*) \geq \bar{m} \tag{5.6}$$

Provided that $G(c^*) \geq \bar{m}$, the first two terms, $\bar{m}w$ and r, are constant. Hence, government's optimization problem simplifies to the minimization of total disutility of labor, $\int_0^{c^*}(1-u_c)c\mathrm{d}G(c)$, with respect to $c^* = w - t_L + t_U$, subject to the budget constraint. Trivially, total disutility is minimized by setting c^* as low as possible, i.e. such that $G(c^*) = \bar{m}$. This is obtained by raising the tax on labor income, t_L, redistributing the revenues in the form of higher unemployment benefits, $-t_U$. Hence, in the optimum, taxes and transfers are set such that supply equals demand and nobody is involuntarily unemployed.

It is not surprising that this policy is optimal if marginal utilities of consumption are constant and social preferences utilitarian. With such individual and social preferences, government is merely concerned with efficiency. Since the market mechanism is not likely to generate an efficient allocation of jobs in the presence of rationing, it is optimal for government to correct this inefficiency with its tax instruments. This conclusion strictly holds for any rationing schedule that is not perfectly efficient and weakly holds for an efficient rationing schedule. That is, had rationing been efficient, the government would be indifferent between the laissez-faire outcome and the outcome in which it actively rewarded/taxed the persons with the highest disutility out of the labor market – after all, the allocation of jobs would in both cases be identical.

[4] I can ignore any firm profits since they are fixed by the ration.

As it turns out, above outcome is independent of the assumption of utilitarian social preferences. Note that the policy reform suggested above requires a transfer of resources from the working population towards the unemployed, i.e., higher t_L, lower t_U, while keeping total employment, $\bar{m}$, and thus the tax base constant. Besides the efficiency gain, this reform therefore leads to a distributional gain for any government that values redistribution from individuals with high utility towards individuals with low utility. In fact, if marginal welfare is decreasing in utility, optimal taxes might very well push c^* below the point at which $G(c^*) = \bar{m}$, sacrificing production for a larger degree of equity.

Proposition 5.1 *Consider an economy with a single type of labor as sole factor of production in which labor supply exceeds demand. It is then optimal for the government to increase both taxes for workers and benefits for the unemployed, such that workers with a low utility of work voluntarily become unemployed, thereby creating jobs for unemployed individuals that derive more utility from working. This reform improves efficiency by reducing the inefficient ration, and improves equity by transferring resources from workers to the unemployed. Government optimally raises taxes and benefits to the point that involuntary unemployment is entirely replaced by voluntary unemployment.*

5.2.2 Graphical representation and discussion of results

Above analysis points to a little recognized inefficiency related to involuntary unemployment, originating from heterogeneity of individuals' disutility of labor participation. Some of the unemployed are bound to have a lower disutility of participation than some of the employed, since there is no market mechanism which ensures otherwise. Thanks to the model's uncomplicated nature, this basic narrative can easily be illustrated by familiar graphical representation of the labor market. Figure 5.1 illustrates the additional dead-weight loss created by inefficient rationing in the absence of taxes and benefits. Panel a. shows the extreme case of perfectly efficient rationing in which the dead-weight loss equals the conventional red triangle IV. Producer surplus is given by the blue area I. Individuals that succeed in obtaining a job are the ones that have the highest utility surplus of working, represented by the leftmost part on the labor supply curve. Worker surplus is therefore given by the blue areas II and III. The opposite extreme is illustrated in panel b. Only the people with the lowest positive utility surplus of working obtain a job, represented by the rightmost part of the labor supply curve

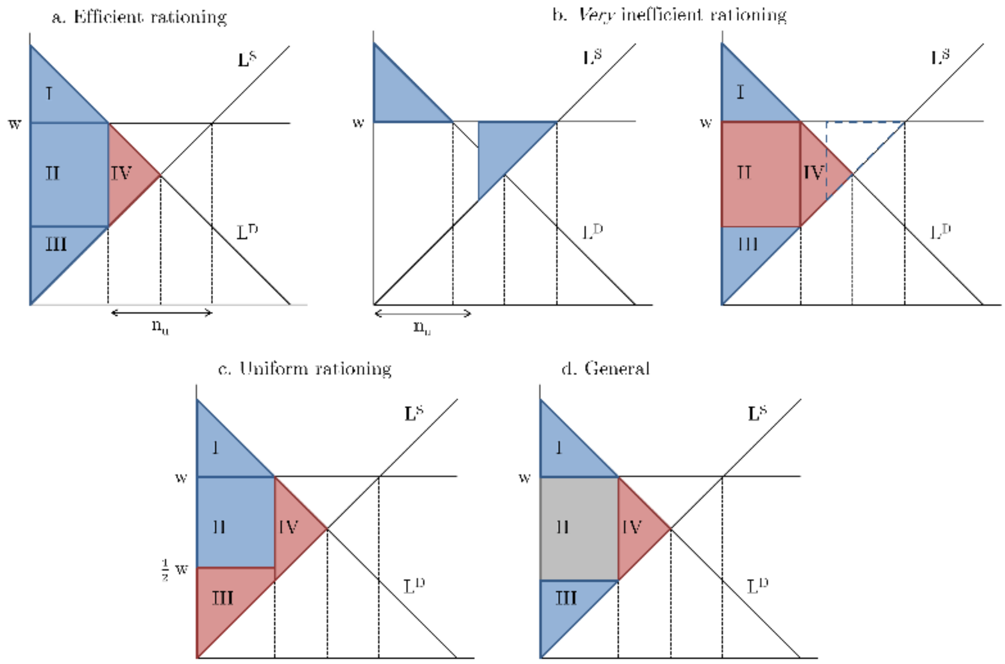

Figure 5.1: Deadweight loss for various rationing schedules

as shown in the first graph. As shown in the second graph, worker surplus now equals area III, generating an additional dead-weight loss, equal to area II, over and above the conventional dead-weight loss of area IV. As drawn in this figure, the additional dead-weight loss exceeds the conventional dead-weight loss.

The intermediate case of uniform rationing, in which every worker has an equal probability of unemployment such that $u_c = \bar{u}$ for all $c \in [0, c^*]$, is illustrated in panel c. Given a linear supply schedule, average disutility of work now equals $w/2$. Therefore I can illustrate worker surplus by the blue rectangle II which gives the number of workers multiplied by $w/2$. The additional dead-weight loss in this case equals the red area III. As drawn in the figure, the additional dead-weight loss again exceeds the conventional dead-weight loss. Palda (2000) provides some calculations of the size of the two different dead-weight losses based on a simple calibration of the uniform-rationing case. He shows that area III exceeds area IV especially for low levels of rationing. This owes to the fact that the conventional dead-weight loss of a marginal increase in unemployment is zero in the absence of rationing, making it a second-order welfare loss. The additional dead-weight loss, due to inefficient rationing, of a marginal increase in unemployment is strictly positive in the absence of rationing and hence represents a first-order welfare loss.

Panel d. of Figure 5.1 summarizes the earlier panels for the case of an *a* *priori* unknown rationing incidence. Area II gives the potential dead-weight loss of inefficient rationing. For any realistic rationing schedule in between the two extremes of efficient and very inefficient rationing, area II will be part worker surplus, part dead-weight loss. The proposed policy reform increases taxes for the working population and transfers for the unemployed population as long as involuntary unemployment is prevalent. Both aspects of the policy reform shift the labor supply schedule to the left, and does so until it intersects the labor demand schedule and the wage floor at the same point. At that point, the policy reform effectively brought the economy back to panel a., in which case area II represents a surplus divided between the employed and the unemployed.[5] Hence, the policy reform improves equity by transferring resources from workers to the unemployed, and improves efficiency by removing the inefficiency of the rationing schedule. The contrast with the standard analysis in public finance is striking. Without initial involuntary unemployment, the same reform would result in the same labor supply response. However, instead of generating an efficiency gain, this labor supply response would result in an efficiency loss, represented by the familiar dead-weight loss triangle of area IV. In the presence of involuntary unemployment, however, this dead-weight loss is already prevalent and lower labor supply improves, rather than worsens, efficiency.

Within the confinement of the model, Proposition 5.1 states that government optimally uses its tax instruments to remove any involuntary unemployment. Consequently, a binding minimum wage or other policies raising wages above market-clearing levels are never optimal. My findings also shed new light on participation policy which, by implication, is aimed at increasing the critical level, c^*. My analysis shows that, if labor markets are rationed, it is optimal to *de*crease participation. Higher participation leads, in the best case scenario of efficient rationing, to no welfare change at all, but in any other scenario it leads to welfare losses due to more inefficient rationing. Hence, the wisdom of participation policy crucially depends on whether the relevant segment of the labor market is rationed or not. If it is, increasing participation will merely lead some persons with a low utility surplus of work to take over jobs of persons with a higher utility surplus of work. Taking a cyclical interpretation of my results, the presence of inefficient

[5]One might be led to believe that area II equals the tax revenue of the government and thus constitutes the utility surplus of the unemployed, such that workers' utility is measured by area III. However, the labor supply schedule shifts leftwards not only because of higher taxes, but also because of higher unemployment benefits. Thus, tax revenue is necessarily smaller than area III, and part of this area in fact constitutes workers' utility.

rationing suggests the optimality of anticyclical tax wedges on labor supply. After all, the 'normal' distortionary effects of taxes and unemployment benefits are less relevant in an economic slump characterized by rigid wage floors and concurrent high involuntary unemployment.[6]

The model I discussed in this section is highly stylized, which naturally limits the value of any policy implications that might be drawn from it. This problem is partly confronted in the next section by expanding the model to include general equilibrium effects as individuals might decide to work in a skilled and unrationed segment of the labor market. Additional assumptions need to be made to derive similar results. In a later section, I also consider endogenous wage floors, additional tax instruments, and a dynamic setup of the model.

5.3 Rationing and skill formation

5.3.1 Model

Firms

In this section I assume there are two different types of labor - skilled and unskilled - that are imperfect substitutes in production. Wages for workers are fixed at a level above the market-clearing wage, which can be thought of as a consequence of minimum wage legislation, union wage setting, efficiency wages or some other form of institutionalized wage rigidity.[7] There is a representative, competitive, profit-maximizing firm which produces output, Y, by employing skilled and unskilled labor as inputs. I normalize labor supply such that every worker supplies one unit of labor, and I denote aggregate skilled labor as n_H and unskilled labor as n_L. I furthermore normalize the price of output to unity. Production is a function of skilled and unskilled labor and homogeneous of degree one:

$$(5.7) \qquad Y = F\left(n_H, n_L\right), \quad F_H, F_L > 0, \quad F_{HH}, F_{LL} < 0, \quad F_{HL} > 0.$$

[6] This cyclical interpretation echoes the results of Landais, Michaillat and Saez (2013), who argue in favor of anticyclical unemployment benefits for much the same reason.

[7] In this section, I do not explicitly model the specific source of rationing, but instead simply assume it as exogenously given. I briefly return to this point in the next section, in which I derive similar policy implications from a simple model with union wage setting.

Subscripts H and L of the production function denote partial derivatives with respect to n_H and n_L. The marginal products of labor are positive, but diminishing for each type of labor. Both inputs are essential and cooperant factors of production.

Firms demand labor, taking wages as given. The first-order conditions for profit maximization imply that the marginal productivities of labor equal the gross wage rates of each type of worker:

$$F_H(n_H, n_L) = w_H, \tag{5.8}$$

$$F_L(n_H, n_L) = w_L, \tag{5.9}$$

where w_H and w_L denote skilled and unskilled wage rates. As the unskilled wage is fixed above the market clearing wage, unskilled labor supply exceeds unskilled labor demand, n_L. This causes involuntary unemployment among the unskilled.

Notice that homogeneity of the production function allows us to rewrite the first-order conditions as:

$$F_H\left(\frac{n_H}{n_L}, 1\right) = w_H, \tag{5.10}$$

$$F_L\left(\frac{n_H}{n_L}, 1\right) = w_L. \tag{5.11}$$

The second condition implies that a fixed unskilled wage rate, w_L, determines the relative factor intensity, n_H/n_L, which, according to the first condition, determines the skilled wage rate, w_H. Thus, fixed unskilled wages imply fixed skilled wages, and unskilled employment can be written as a fixed proportion of skilled employment: $n_L = \gamma n_H$, with γ some constant.

Individuals

There is a continuum of individuals with mass normalized to one. As in the previous section, individuals are heterogeneous with respect to c, which determines their utility costs of participation and is distributed according to $G(c)$, with density $g(c)$ and support $[0, \bar{c}]$. Individuals decide whether to become skilled worker, unskilled worker, or voluntarily unemployed. Disutility of being unskilled employed equals c, disutility of being skilled employed equals $(1 + \beta)c$. I assume $\beta > 0$, such that for every individual it is more costly to become a skilled worker than it is to become an unskilled worker. As a result, only individuals with a low value of

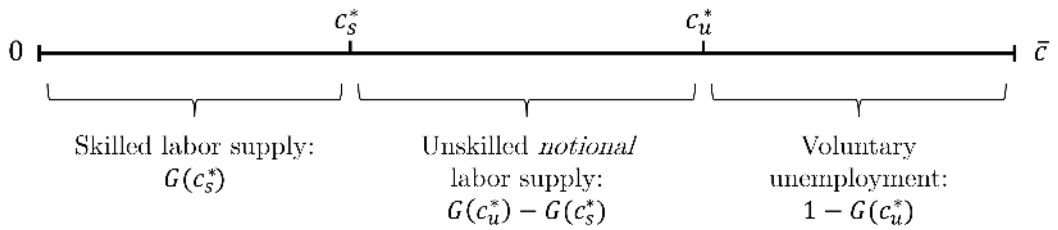

Figure 5.2: Graphical illustration of labor market equilibrium

c prefer to be skilled, while individuals with a high value of c prefer to remain unskilled.

I assume government observes individual labor earnings (or wage rates), such that it can set skill-specific tax instruments. Taxes for skilled workers, unskilled workers, and the unemployed are given by t_H, t_L, and t_U. Naturally, as government is required to run a balanced budget, one or more of these taxes might be negative and thus represent a transfer. Utility of the three types of individuals is given by:

$$V_c^H \equiv w_H - t_H - (1+\beta)\,c, \quad (5.12)$$

$$V_c^L \equiv w_L - t_L - c, \quad (5.13)$$

$$V_U \equiv -t_U, \quad (5.14)$$

Subscripts c are meant to emphasize the fact that utility of skilled and unskilled workers depend on their disutility of work.

There are now two critical levels of c. One critical level, denoted c_s^*, separates the skilled from the unskilled; the other critical level, denoted c_u^*, separates the voluntary unemployed from the unskilled employed. Individuals with $c \in [0, c_s^*)$ prefer to be skilled, individuals with $c \in [c_s^*, c_u^*)$ prefer to be unskilled employed, and individuals with $c \in [c_u^*, \bar{c}]$ prefer to be voluntarily unemployed. Thus, skilled labor supply equals $G(c_s^*)$, unskilled *notional* labor supply equals $G(c_u^*) - G(c_s^*)$, and the number of voluntarily unemployed individuals is given by $1 - G(c_u^*)$. This equilibrium is graphically illustrated by Figure 5.2.

The upper critical level of the disutility of participation, c_u^*, equates the utilities of being unemployed and unskilled employed: $V_{c_u^*}^L = V_U$. Hence:

$$c_u^* = w_L - t_L + t_U. \quad (5.15)$$

Any individual with $c > c_u^*$ decides to become voluntarily unemployed. Naturally,

c_u^* is increasing in the net income of the unskilled employed and in the tax rate of the unemployed.

Individuals decide to be skilled or unskilled before knowing whether they are actually able to secure a job as an unskilled worker. They do know their probability of obtaining a job. This probability is governed by the rationing schedule, which assigns a probability of involuntary unemployment, u_c, to every level of $c \in [c_s^*, c_u^*)$. The *expected* utility of being unskilled is given by the probability-weighted average of unskilled employed utility and unemployed utility, given by:

$$V_c^{EL} \equiv (1 - u_c)\, V_c^L + u_c V_U. \tag{5.16}$$

The lower critical value, c_s^*, is determined by equating skilled utility and expected unskilled utility: $V_{c_s^*}^H = V_{c_s^*}^{EL}$. Substituting for utility and rearranging yields:

$$(\beta + u_{c_s^*})c_s^* = w_H - t_H - w_L + t_L + u_{c_s^*}c_u^*. \tag{5.17}$$

This critical value is increasing in the income of skilled workers and decreasing in the income of unskilled workers. Moreover, as the last term in equation (5.17) indicates, the critical value is increasing in the unemployment probability multiplied by the upper critical value, $u_{c_s^*}c_u^*$. A higher unemployment probability reduces the expected income from being unskilled, the reduction given by the difference between unskilled employed income and unemployment benefits which equals c_u^*.[8]

Labor market equilibrium

As the skilled segment of the labor market is not rationed, skilled labor demand must, in equilibrium, equal skilled labor supply:

$$n_H = G(c_s^*). \tag{5.18}$$

The unskilled segment of the labor market is rationed and notional labor supply exceeds labor demand, the excess being involuntary unemployment. Unskilled labor demand thus equals the difference between unskilled labor supply and involuntary

[8] Additionally, the following second-order condition needs to be satisfied: $\chi(c_s^*) \equiv \mathrm{d}\left(V_{c_s^*}^H - V_{c_s^*}^{LE}\right)/\mathrm{d}c < 0$. That is, at the critical level c_s^*, skilled utility should be decreasing in c faster than expected unskilled utility. For the critical level to be unique it is sufficient to assume that $\chi(c) > 0$ for all c. This condition will be imposed to ensure uniqueness of c_s^*. For completeness, note that $\chi(c) = -(\beta + u_c) + \frac{\mathrm{d}u_c}{\mathrm{d}c}(c_u^* - c)$.

unemployment:

$$n_L = \int_{c_s^*}^{c_u^*} (1 - u_c)\mathrm{d}G(c). \tag{5.19}$$

Substituting the equilibrium values for n_H and n_L into the firm's first-order condition for unskilled labor, (5.9), yields:

$$F_L\left(G(c_s^*), \int_{c_s^*}^{c_u^*} (1 - u_c)\mathrm{d}G(c)\right) = w_L. \tag{5.20}$$

Thus, unemployment can be seen to result from the representative firm's first-order condition. Notice, however, that this condition only determines *aggregate* unemployment, $\int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c)$, whereas the continuum of individual unemployment rates, $\{u_c\}$, are determined by the rationing schedule.

I can now define equilibrium in the private sector as a set of values for c_s^*, c_u^*, and $\{u_c\}$, which satisfies individuals' and firms' optimization as described by equations (5.15), (5.17), and (5.20), along with a rationing schedule which determines how aggregate unemployment translates into individual unemployment probabilities.

Government

Government sets taxes in a way that balances its budget:

$$\mathcal{B} \equiv n_H t_H + n_L t_L + n_U t_U - r = 0, \tag{5.21}$$

where $n_U = 1 - n_H - n_L$ is the number of unemployed people and r is an exogenous revenue requirement. Again, I assume for now that government is utilitarian such that social welfare can be represented as a simple sum of individuals' utility:

$$\begin{aligned} \mathcal{W} \equiv\ & n_H(w_H - t_H) + n_L(w_L - t_L) - n_U t_U \\ & - (1+\beta)\int_0^{c_s^*} c\mathrm{d}G(c) - \int_{c_s^*}^{c_u^*} (1 - u_c)c\mathrm{d}G(c). \end{aligned} \tag{5.22}$$

5.3.2 The welfare analysis of a participation tax reform

Inelastic skill formation

In this subsection I analyze the welfare effects of a policy reform that raises unskilled taxes and unemployment benefits to substitute voluntary unemployment for involuntary unemployment. For expository purposes, I first perform the welfare analysis in the case of a fixed supply of skilled labor. Thus, the change in skilled labor supply due to the reform is assumed to be nil: $\Delta n_H = 0$.[9] Notice that this implies, by equation (5.20), that the number of unskilled employed, and thus the number of unemployed, remain constant: $\Delta n_L, \Delta n_U = 0$. The policy reform under consideration raises unskilled taxes by $\Delta t_L > 0$, and unemployment benefits by $-\Delta t_U > 0$, such that the government budget, (5.21), balances, and involuntary unemployment is substituted entirely by voluntary unemployment. As in the previous section, the transfer from the employed to the unemployed constitutes an equity improvement. Taking differences of equation (5.19), the reform thus implies the following:

$$(5.23) \qquad -\Delta G(c_u^*) = -\Delta \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c) = \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c).$$

Notice that the left-hand side of the first equation gives the increase in the number of voluntary unemployed, which must equal the change in the number of involuntary unemployed. This must necessarily hold as the total number of unemployed remains constant. The second equation imposes that the reform eliminates all involuntary unemployment by replacing it with voluntary unemployment.[10]

As $\Delta n_H, \Delta n_L, \Delta n_U = 0$, the tax base is fixed such that higher taxation does not lead to tax base erosion. To see this, notice that budget balance requires $\Delta \mathcal{B} = 0$. Thus, taking differences of equation (5.21) and substituting for $\Delta \mathcal{B} = 0$ yields:

$$(5.24) \qquad -n_U \Delta t_U = n_L \Delta t_L.$$

This equation implies that the increase in outlays on unemployment benefits, $-n_U \Delta t_U$, is equal to the mechanical revenue gains from increased taxes, $n_L \Delta t_L$. In standard full-employment models of public finance, the outlays on unemployment benefits would be lower than the mechanical revenue gains as lower labor supply leads to a smaller tax base. In the presence of involuntary unemployment, however, the tax base is determined by labor demand, not labor supply. Hence, the standard equity-efficiency trade-off is not present. In fact, efficiency improves

[9] This would for example be applicable if, for a large enough neighborhood around c_s^*, the density $g(c)$ is nil.

[10] Feasibility of this reform is ensured by the fact that c_u^*, and thus $-\Delta G(c_u^*)$, are increasing in t_L and decreasing in t_U.

as the inefficient rationing is removed. This can be seen by taking differences of equation (5.22), the utilitarian social welfare function:

$$\Delta \mathcal{W} = \int_{c_u^* + \Delta c_u^*}^{c_u^*} c \mathrm{d}G(c) - \int_{c_s^*}^{c_u^*} u_c c \mathrm{d}G(c) \geq 0. \tag{5.25}$$

As there is no tax base erosion, and productivity and wages remain constant, total production in the economy remains constant as well. As a result, the only welfare effects consist of changes in disutility of work. The first term gives the disutility of work of the unskilled workers that decide to become voluntarily unemployed due to the tax reform, i.e., every person with disutility of work $c \in (c_u^* + \Delta c_u^*, c_u^*)$. This constitutes a welfare gain as those people will no longer suffer this disutility of work. The second term is the disutility of work that is suffered by the previously unemployed people that take over the jobs of the now voluntary unemployed. Equation (5.23) implies that the additional number of voluntary unemployed exactly offsets the number of previously unemployed people. Thus, the total disutility of work (weakly) decreases due to the policy reform, such that equation (5.25) is greater than or equal to zero.[11]

Notice that the welfare analysis of the policy reform is identical to the analysis of Section 5.2. The only difference is that the constant producer surplus does not constitute firm profits, but instead represents the utility surplus of skilled workers. The reason that the above analysis is identical to the one of the previous section is of course that the number of skilled workers is constant, an assumption I relax below.

Elastic skill formation

The crucial question, then, is how the policy implications are affected by elastic skill formation which allows for $\Delta n_H \neq 0$. I again consider a policy reform that raises unskilled labor taxes and unemployment benefits, under a balanced budget, which eliminates involuntary unemployment. Remember that profit maximization and homogeneity of the production function implies that $n_L = \gamma n_H$. Hence, taking differences of this equation, the reform implies the following for the change in

[11]Notice that $\Delta \mathcal{W} = 0$ if and only if the rationing schedule is efficient. In that case, only the people who were previously involuntarily unemployed become voluntarily unemployed. In that case the allocation of jobs, and thus the total disutility of work, remains constant.

involuntary unemployment:

$$-\Delta G(c_u^*) + (1+\gamma)\Delta n_H = -\Delta \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c) = \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c). \tag{5.26}$$

Notice that the first term of the first equation again gives the increase in voluntary unemployment. Every additional voluntarily unemployed person yields a job for one involuntarily unemployed person. The second term is new and gives the effect of a change in skilled labor supply. Higher skilled labor supply leads to a direct decrease in low-skilled labor supply, and increases labor demand by γ as higher skilled labor supply increases productivity of unskilled workers. Both effects lead to lower involuntary unemployment. I again assume that government raises unskilled labor taxes and unemployment benefits until all involuntary unemployment is eliminated, as stated by the second equation.[12]

A sufficient condition for obtaining the same policy implication as under inelastic skill formation, is that the reform stipulated above leads to higher skilled labor supply. Since the ratio n_H/n_L is constant, higher skilled labor supply leads to higher unskilled labor demand. As a result, the policy reform does not only increase equity and reduce inefficient rationing, but yields an additional efficiency gain because of an expanding tax base. It is easiest to illustrate this result formally by considering the marginal effect of a budget-neutral increase in the unskilled tax rate on the size of the unemployment benefits. This is obtained by taking the derivative of the government budget, equation (5.21), and substituting for $\mathrm{d}\mathcal{B} = 0$:[13]

$$-n_U \mathrm{d}t_U = n_L \mathrm{d}t_L + ((t_H - t_U) + (t_L - t_U)\gamma)\mathrm{d}n_H. \tag{5.27}$$

Here, I used the fact that $\mathrm{d}n_L = \gamma \mathrm{d}n_H$ and $\mathrm{d}n_U = -\mathrm{d}n_L - \mathrm{d}n_H$. In contrast to the case of inelastic skilled labor supply, the additional transfer to the unemployed may be larger or smaller than the mechanical revenue gains from higher unskilled

[12] Feasibility of this policy reform requires that the left-hand side of equation (5.26) is increasing in the unskilled tax rate. This is necessarily the case if skilled labor supply increases in response to the policy reform, $\Delta n_H > 0$. If n_H decreases too strongly with the unskilled tax rate, eliminating involuntary unemployment with the unskilled tax rate and unemployment benefits alone is not feasible. Below, however, I concentrate on the case of increasing skilled labor supply, for which feasibility is not an issue.

[13] I present the marginal effects of a budget-neutral increase in unskilled labor taxes and unemployment benefits, even though the proposed policy reform is designed to eliminate all involuntary unemployment. This allows me to save on the notation of inframarginal effects, which are somewhat cumbersome, do not provide additional insights, and are not relevant for the results.

labor taxes. Assuming, as seems reasonable, that taxes on the employed are larger than taxes on the unemployed, $t_H, t_L > t_U$, the final term in equation (5.27) is positive for $\mathrm{d}n_H > 0$ and negative for $\mathrm{d}n_H < 0$. In other words, higher skilled labor supply leads to tax base expansion, whereas lower skilled labor supply leads to tax base erosion.[14]

The total efficiency gain from the policy reform can be derived by taking the derivative of equation (5.22), the utilitarian social welfare function:

$$\mathrm{d}\mathcal{W} = -\int_{c_s^*}^{c_u^*} (c_u^* - c)\mathrm{d}u_c \mathrm{d}G(c) + ((t_H - t_U) + (t_L - t_U)\gamma)\mathrm{d}n_H, \tag{5.28}$$

the derivation for which I used equation (5.27). The first term represents the efficiency gain associated with eliminating inefficient rationing. Notice that $c_u^* - c$ measures the utility difference between being unskilled employed and unemployed for a person with participation costs c. The efficiency gains of lower involuntary unemployment is given by the integral over this difference, multiplied by the change in unemployment. The second term represents the social welfare effect associated with the change in tax revenue due to expansion or erosion of the tax base. As derived above, both terms are unambiguously positive if the policy reform leads to higher skilled labor supply. If, on the other hand, the response of skilled labor supply is negative, it is ambiguous whether the policy reform leads to an efficiency gain. It then depends on whether the efficiency gain from eliminating involuntary unemployment outweighs the efficiency loss from an eroded tax base.

Summing up, if a budget-neutral increase in both unskilled taxes and unemployment benefits leads to higher skill formation, this policy reform improves equity by redistributing from the employed to the unemployed, and improves efficiency because it both eliminates the inefficient rationing schedule and expands the tax base. A positive skilled labor supply response is therefore a sufficient condition for the traditional equity-efficiency trade-off to break down. The important question that remains to be answered, then, is whether the policy reform indeed leads to higher skilled labor supply. This turns out to be ambiguous and depends on whether, for an individual with disutility c_s^*, the expected earnings of being unskilled increase or decrease as a result of the policy reform. This can be seen from

[14] Naturally, such tax base erosion could be avoided by simultaneously reducing taxes for skilled workers. However, while the inefficiency due to tax base erosion would be averted, this tax adjustment implies a transfer towards the skilled employed and hence an equity deterioration. This is exactly the classic equity-efficiency trade-off which thus might appear if the policy reform leads to lower skilled labor supply.

taking differences of equation (5.17):

$$\beta \Delta c_s^* = \Delta t_L - u_{c_s^*}(c_u^* - c_s^*). \tag{5.29}$$

Remember that c_s^* is the critical disutility level which separates the skilled from the unskilled. Hence, an increase in this critical level implies higher skilled labor supply. On the one hand, the policy reform tends to stimulate skilled labor supply as it taxes unskilled labor. This is illustrated by the first term of equation (5.29), which is positive. On the other hand, the policy reform discourages skilled labor supply since it removes involuntary unemployment and thus raises the probability of obtaining a job when supplying unskilled labor. This is illustrated by the second term of equation (5.29), which is the product of the before-reform unemployment rate at c_s^*, $u_{c_s^*}$, and the utility difference between obtaining an unskilled job and being unemployed, $w_L - t_L - c_s^* + t_U = c_u^* - c_s^*$. If the unskilled tax increase is larger than the earnings increase due to a lower probability of unemployment, skilled labor supply increases.

To obtain more insight into the conditions under which the policy reform leads to higher skilled labor supply, I can further substitute for Δt_L in equation (5.29). Before doing so, I define the average density between disutility c_1 and c_2 as $\bar{g}(c_1, c_2) \equiv (G(c_2) - G(c_1))/(c_2 - c_1)$, and the average involuntary unemployment rate as $\bar{u} \equiv \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c) / \int_{c_s^*}^{c_u^*} \mathrm{d}G(c)$. As I show in the Appendix by substituting for Δt_L in equation (5.29), the policy reform leads to more skilled labor supply if and only if the following condition is satisfied:

$$\frac{n_U}{n_U + n_L} > \left(\frac{u_{c_s^*}}{\bar{u}}\right) \frac{\bar{g}(c_u^*, c_u^* + \Delta c_u^*)}{\bar{g}(c_s^*, c_u^*)}. \tag{5.30}$$

The left-hand side gives the size of the unskilled labor tax increase associated with an increase in voluntary unemployment. If there is a relative large number of unemployed, a given increase in unskilled tax revenue leads to only a small increase in the unemployment transfer as it has to be shared among a large number of unemployed. In that case, the necessary unskilled tax increase is relatively large, encouraging skilled rather than unskilled labor supply. The right-hand side of the condition features two fractions. The first fraction gives the unemployment rate for a person with disutility level c_s^*, relative to the average unemployment rate. If $u_{c_s^*}$ is relatively large, the elimination of involuntary unemployment strongly raises the expected unskilled utility for a person with disutility c_s^*. As a result, skilled labor supply is more likely to decrease. On the other hand, if the average unemployment rate, $\bar{u}$, is relatively large, a large increase in unskilled taxes is required to eliminate involuntary unemployment, increasing the attractiveness of skilled labor supply. The same logic applies to the last term which gives the fraction of average population densities between disutility levels $[c_u^*, c_u^* + \Delta c_u^*]$ and $[c_s^*, c_u^*]$. If the density at the participation margin is relatively large, unemployment responds relatively strongly on a given increase in unskilled taxes. If the population density on $[c_s^*, c_u^*]$ is relatively large, involuntary unemployment is also large (for given unemployment rate), and a large increase in taxes is required to eliminate involuntary unemployment.

Proposition 5.2 *Consider a policy reform that substitutes voluntary unemployment for involuntary unemployment by increasing both unskilled taxes and unemployment benefits. In an economy with both skilled and unskilled labor as factors of production, in which unskilled labor is rationed, such a policy reform unambiguously improves both equity and efficiency if it leads to higher skilled labor supply. It improves equity by transferring resources from the employed to the unemployed. It improves efficiency because it both eliminates inefficient rationing and increases the tax base by increasing skilled and unskilled labor demand. The policy reform increases skilled labor supply only if the unemployment rate at the skill margin,* $u_{c_s^*}$*, is small enough for condition (5.30) to hold.*

5.3.3 Discussion

Proposition 5.2 underlines the crucial importance of the rationing schedule. On the one hand, the more inefficient is the rationing schedule, the higher are the efficiency gains from eliminating involuntary unemployment as is evident from the first term in equation (5.28). On the other hand, if rationing is so inefficient that much of the unemployment is concentrated on individuals that are relatively indifferent between skilled and unskilled work, eliminating unemployment might lead to efficiency losses due to declining skilled labor supply and an eroding tax base. This is illustrated by condition (5.30) and echoes the results of Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation), who find that a minimum wage can only be part of optimal policy if the resulting involuntary unemployment leads to sufficiently more skilled labor supply.

It is useful to compare my results to those of Lee and Saez (2012) who find that a finding minimum wage, and thus unemployment, might be optimal if rationing is efficient. Proposition 5.2, however, implies that if all rationing is concentrated on those with the highest disutility of work, rationing is optimally substituted by voluntary unemployment. The explanation of this seeming contradiction is that Lee and Saez define efficient rationing to be concentrated at both the participation margin and the skill margin – since those people at the skill margin have a low utility surplus of unskilled work, *relative to skilled work*.[15] Semantics aside, the relevant question is not so much which definition of efficient rationing makes more sense; rather it is what rationing schedule is the most plausible description of reality.

In providing a provisional answer to this question, it might be helpful to distinguish between the short run – in which education is fixed – and the long run – in which education is flexible. In the short run it is unlikely that involuntary unemployment affects skilled labor supply as skill formation tends to be a time-consuming investment. It is also unlikely that skilled workers will be rationed if they want to switch to unskilled employment. In the words of Lee and Saez, "it may be realistic that employers could preferentially hire the most qualified workers even for minimum wage jobs" (footnote 15, p744). This effectively implies that $u_{c_s^*} = 0$ and, according to Proposition 5.2, it is in that case both efficiency- and equity-enhancing for government, in response to a short-run rationing shock, to temporarily increase unskilled taxes and unemployment benefits.

In the longer run, students are likely to take into account the conditions on the unskilled labor market when making educational decisions. Indeed, it is found that a high rate of unemployment encourages school enrollment and discourages dropping out from high school (e.g., see the surveys of empirical estimates of the effect of unemployment on school enrollment and drop-out rates in Clark, 2011; Gerritsen and Jacobs, 2013*a*, Chapter 4 of this Dissertation). Thus, in the long run one could argue that involuntary unemployment contributes to skill formation which, if the effect is large enough, could then justify the existence of some degree of rationing. However, the same studies that find positive schooling effects of unemployment also generally find that a higher remuneration for unskilled work (the counterpart of higher rationing) discourages schooling. For the optimality of rationing it is crucial to know the simultaneous effect of higher rationing and higher unskilled wages on skill formation. To the best of my knowledge, however,

[15]The definition of efficient rationing, implicit in panel a. of Figure 5.1 could also be termed *ex-post* efficient rationing: given the skill decisions made by individuals, rationing is most efficient if it is concentrated on those unskilled workers with the highest cost of participation. The definition of Lee and Saez then refers to *ex-ante* efficient rationing: before any skill decision has been made, it is most efficient to ration those workers that are most willing to become unemployed or skilled, rather than unskilled employed.

this has never properly been tested empirically.

5.4 Further extensions

5.4.1 Endogenous union-set wages

So far, I assumed that involuntary unemployment was the result of an exogenous wage floor which exceeds the market-clearing wage. In reality, one could argue, wage floors are set endogenously by rational agents such as a labor-monopolizing union. If such agent cares for workers' net wages, which is presumably the reason why he sets a wage floor, a policy reform that increases unskilled labor taxes might induce him to increase the wage floor. As a result, the policy reform might reduce labor demand as well as labor supply, leading to an erosion of the tax base and causing the equity-efficiency trade-off to resurface. I show below that a higher level of labor taxes is indeed likely to raise wage demands of a union that cares about both net wages and employment. However, if government commits to raising taxes in response to an increase of involuntary unemployment, it can directly incentivize the union to moderate its wage demands.

I return to the model of Section 5.3, but now assume that unskilled wages are set by a labor union. As is commonly observed (e.g., Booth, 1995), I assume that the goal function of the union depends positively on (unskilled) net wages and employment, according to:

(5.31)

$$\Pi \equiv \Pi\left(w_L - t_L, n_L\right), \quad \Pi_{w_L-t_L}, \Pi_{n_L}, \Pi_{w_L-t_L,n_L} > 0, \quad \Pi_{w_L-t_L,w_L-t_L}\,\Pi_{n_L,n_L} < 0,$$

where subscripts denote partial derivatives with respect to net wages and employment. The union maximizes this function, subject to firm behavior which determines employment, n_L, and government behavior which determines the tax rate, t_L. I assume that firms' production function implies a constant elasticity of labor demand which equals $\varepsilon \equiv -\frac{\mathrm{d}n_L}{\mathrm{d}w}\frac{w_L}{n_L}$. First suppose that t_L is exogenously set by the government and thus independent of the level of unemployment. The labor union then sets wages according to:

$$\frac{\Pi_{w_L-t_L}}{\Pi_{n_L}}\frac{w_L}{n_L} = \varepsilon. \tag{5.32}$$

As long as the left-hand side is larger than ε, the union raises wage demands as the marginal gains from higher net wages outweighs the marginal employment costs.

For equation (5.32) to represent an equilibrium wage rate, I need to assume that the left-hand side is declining with the gross wage rate. This holds true if the complementarity between net wages and employment in the union's goal function is sufficiently strong.[16]

Equation (5.32) implies that gross wage demands increase in response to an income tax increase. The direct effect of a tax increase is to raise the union's marginal rate of substitution of net wages for employment, Π_{w-t_L}/Π_{n_L}, raising the left-hand side of equation (5.32) above ε. Consequently, the union raises the gross wage rate until equilibrium is restored. It would be wrong, however, to interpret this as discrediting the results of Sections 5.2 and 5.3. After all, I argued that income taxation should *not* be set exogenously but be determined by and depend positively on unemployment. That is, income taxation should be endogenous with respect to the union's decisions that affect the degree of unemployment. Abstracting from the assumption of an exogenously set wage rate, the equilibrium wage rate can now be seen to equal:

$$\frac{\Pi_{w_L-t_L}}{\Pi_{n_L}}\frac{w_L}{n_L}\left(1-\frac{\mathrm{d}t_L}{\mathrm{d}w_L}\right)=\varepsilon. \tag{5.33}$$

Notice that the only difference with equation (5.32) is the additional term $(1 - \mathrm{d}t_L/\mathrm{d}w_L)$ on the right-hand side. This term implies that government can directly influence the union's marginal gains of increasing its wage demands. Regardless of the level of taxation, if taxes depend positively on unemployment, $\mathrm{d}t_L/\mathrm{d}w_L > 0$, the union's gains from raising wages is diminished since only part of a gross wage increase is translated into a net wage increase. Thus, in the presence of an endogenously set wage floor, government can directly influence this wage floor by making taxes dependent on unemployment. According to the logic of previous sections, government should make taxes depend positively on unemployment in order to directly eliminate rationing.

5.4.2 Employer-paid vs employee-paid taxes

In conventional supply-determined models of public finance, it is irrelevant whether taxes are levied on the employee or on the employer. Regardless of the statutory

[16]Hence, second-order conditions require that $\frac{\Pi_{w_L-t_L}}{\Pi_{n_L}}\frac{w_L}{n_L}$ is decreasing in the wage rate w_L. As long as labor taxes are nonnegative, a sufficient condition for this is that $\frac{\Pi_{w_L-t_L}}{\Pi_{n_L}}\frac{w_L-t_L}{n_L}$ is decreasing in w_L, which implies that the union's elasticity of substitution of net wages for employment at the equilibrium is smaller than one.

incidence, the economic incidence of an income tax levied on employees is identical to that of one levied on employers. Naively applying this principle on the analysis of previous sections would imply that government could equivalently raise employers' taxes and unemployment benefits without facing a trade-off between equity and efficiency. It is, however, straightforward to see that this logic is false. In supply-determined models the statutory incidence is irrelevant because of flexible wages. If employers are taxed, they demand less labor, leading to lower wages and thus a shared burden of the tax. If employees are taxed, they reduce labor supply, leading to higher wages and an equivalently shared tax burden. Rationed labor markets, however, are characterized by inflexible wages, causing the statutory and economic incidence to coincide.

To see this, I expand the model of Section 5.3 by introducing an employer tax on unskilled labor, τ, and compare the comparative statics of changes in τ and t_L. The firm's first-order condition for unskilled labor demand is now given by:

$$F_L\left(n_H, n_L\right) = w_L + \tau, \tag{5.34}$$

I consider the comparative statics of changes in τ and t_L, keeping other policy parameters constant. The incidence of these two tax changes on unskilled workers is derived by taking the derivative of unskilled net labor earnings, $w_L - t_L$:

$$\mathrm{d}(w_L - t_L) = -\mathrm{d}t_L. \tag{5.35}$$

Hence, due to a fixed unskilled wage rate, unskilled workers carry all the incidence of an increase in the employee part of taxation. That is, net earnings decrease one-to-one with the unskilled employee tax rate. On the other hand, unskilled workers carry *none* of the incidence of an increase in the employer part of taxation. The reason for this is, again, the fact that the wage rate is constant such that no incidence can be shifted to the unskilled employed.

To determine the incidence of the tax changes on skilled workers, I need to derive the comparative statics for net earnings $w_H - t_H$. For this, note that homogeneity of the production function implies that profits are nil, such that $\mathrm{d}(F(n_H, n_L) - w_H n_H - (w_L + \tau)n_L) = 0$. From this derivation, I obtain the incidence of the tax changes on skilled workers:

$$n_H \mathrm{d}(w_H - t_H) = -n_L \mathrm{d}\tau. \tag{5.36}$$

Thus, as all of the incidence of higher unskilled employee taxes is carried by unskilled workers, skilled workers carry none of the incidence of unskilled employee taxes. However, skilled workers carry the full incidence of higher unskilled employer taxes. As can be seen from equation (5.36), skilled workers' net income declines one-to-one with the tax burden of the unskilled employer tax. [17]

Hence, the incidence of employee taxes and the incidence of employer taxes are very different in a rationed labor market.[18] But how do the different taxes affect efficiency? For expository purposes, I assume that skilled labor supply is inelastic. First, consider the government budget, now given by $\mathcal{B} \equiv n_H t_H + n_L (t_L + \tau) + n_U t_U - r$. Taking the total derivative yields:

$$\mathrm{d}\mathcal{B} = n_L (\mathrm{d}t_L + \mathrm{d}\tau) + (t_L + \tau - t_U)\, \mathrm{d}n_L. \tag{5.37}$$

The first term gives the mechanical revenue gains associated with a higher tax rate. Since employee and employer taxes have the same tax base, the mechanical revenue effects are identical for equal increases in either tax. The second term gives the revenue effect associated with a change in the tax base. Assuming that workers are taxed more heavily than the unemployed, such that $t_L + \tau > t_U$, an erosion of the tax base, $\mathrm{d}n_L < 0$, is associated with lower tax revenue. It is readily observable from equation (5.36) that the tax base is declining in the employer tax rate, $\mathrm{d}n_L/\mathrm{d}\tau = F_{LL}^{-1} < 0$, while it is independent from the employee tax rate, $\mathrm{d}n_L/\mathrm{d}\tau = 0$. Thus, echoing results of earlier sections, employee taxes do not affect labor demand. On the other hand, employer taxes lead to lower labor demand and thereby to an efficiency loss through tax base erosion.

The difference between the efficiency consequences of employer and employee taxes are even more dramatic when it comes to the inefficiency of the rationing schedule. Substituting for $n_L \equiv \int_{c_s^*}^{c_u^*} (1 - u_c) \mathrm{d}G(c)$ into equation (5.34) and taking the total derivative, I obtain the following equation for the change in involuntary unemployment:

$$\int_{c_s^*}^{c_u^*} \mathrm{d}u_c \mathrm{d}G(c) = -(1 - u_{c_u^*}) g(c_u^*) \mathrm{d}t_L + \frac{1}{-F_{LL}} \mathrm{d}\tau, \tag{5.38}$$

[17]Notice that if government can freely and separately set both employer and employee taxes, the fixed wage rate becomes irrelevant. Government can then effectively determine both employees' net wages and employers' labor costs, which is equivalent to setting a minimum wage along with employee income taxes. I shortly return to the optimality of minimum wages in my concluding remarks.

[18]A similar difference between the incidence of employer taxes and employee taxes has recently been documented empirically for France by Lehmann, Marical and Rioux (2013).

where I substituted for $\mathrm{d}c_u^* = -\mathrm{d}t_L$. The first term shows that involuntary unemployment is decreasing in the employee tax. The reason is that higher employee taxes lead to lower labor supply, and thus to a smaller difference between supply and demand. The second term, however, shows that higher employer taxes lead to even more involuntary unemployment. Higher employer taxes raise the costs of hiring unskilled labor, thereby reducing labor demand even further. Summing up, while higher unskilled employee taxes do not affect the tax base and improve efficiency by reducing rationing, higher employer taxes worsen efficiency as they lead to tax base erosion and more rationing.

5.5 Concluding remarks

My analysis of rationed labor markets stresses an inefficiency which has not received much attention. With too little jobs for too many potential workers, the market mechanism does not necessarily allocate jobs to the persons that derive most utility from the job. Because of this inefficiency, government might find it optimal to tax the individuals who derive least utility from working out of the labor market, while using the additional revenue to increase unemployment benefits. Proposition 5.1 establishes that, under fixed skill formation, such tax reform improves efficiency as well as equity. The efficiency improvement, originating from the reallocation of jobs, is in stark contrast to public-finance orthodoxy which predicts a trade-off between equity and efficiency as lower labor supply erodes the tax base. In rationed labor markets this orthodoxy is invalid as the tax base is determined by labor demand, which is smaller than labor supply.

I argue that the assumption of fixed skill formation might be appropriate for the short run, but less so for the long run. Relaxing this assumption, Proposition 5.2 establishes that the reform enhances both equity and efficiency if it leads to higher skilled labor supply. This holds true if rationing is relatively efficient in the sense that it is not concentrated on the unskilled workers that are relatively indifferent between skilled and unskilled work. I moreover show that the policy implication – government should raise taxes and unemployment benefits in response to higher unemployment – holds *a fortiori* if a binding wage floor is endogenously determined by a labor union. Such an unemployment-sensitive policy incentivizes a union to moderate its wage claims as it can rationally expect higher taxation in retaliation to higher wage demands. Finally, I show that the incidence of labor taxes for employer and employee are wildly different in the presence of labor rationing –

again in stark contrast to public-finance orthodoxy. Higher employee taxes improve efficiency by reducing involuntary unemployment, whereas higher employer taxes increase involuntary unemployment.

These results have implications for government policies beyond optimal taxation. Naturally, if rationing is optimally removed by appropriately setting taxes, government-imposed rationing arising from a binding minimum wage can never be optimal. Government can use the same information required to enforce a minimum wage to set taxes and transfers in such a way that the minimum wage is no longer binding, thereby removing the inefficiency of rationing while redistributing income from rich to poor. With efficient rationing, a minimum wage has the same efficiency properties as the tax-transfer schedule but inferior distributional properties. When rationing is not efficient, the tax-transfer schedule strictly dominates a minimum wage on both efficiency and distributional grounds.[19]

Policies that encourage participation are also ill-advised if the relevant labor market is rationed. Normally, participation policies are justified by pointing out that a positive income tax wedge distorts the participation decision. Encouraging participation might lead to efficiency gains as it partially offsets this tax distortion. However, if the relevant labor market segment is rationed, this argument has no bite and higher participation will lead to an efficiency loss as more individuals with relatively high costs of work begin competing with others for a fixed amount of jobs. Hence, in determining the wisdom of participation policies it is crucial to understand whether the additional labor supply can add to aggregate employment or whether it merely increases unemployment among workers who derive more utility from working.

The material point of this chapter is that the presence of involuntary unemployment drastically alters the efficiency properties of taxation. Discouraging labor supply, which would be inefficient in the case of full employment, turns out to be efficient in the presence of involuntary unemployment. This can also be interpreted in a cyclical sense. During times of high unemployment – for example due to a negative productivity shock combined with downward wage rigidity – the welfare costs of labor income taxation and unemployment benefits are relatively low. This suggests the optimality of anticyclical labor taxation and unemployment benefits.

[19] Gerritsen and Jacobs (2013*a*, Chapter 4 of this Dissertation) determine how the desirability of a minimum wage depends, among other things, on the rationing schedule and conclude, on the basis of an empirically grounded calibration, that minimum wages are optimally decreased in every OECD country under consideration except the United States.

5.A Appendix

This Appendix contains the full derivation of condition (5.30). Taking differences of equation (5.15) yields:

$$\Delta c_u^* = -(\Delta t_L - \Delta t_U). \tag{5.39}$$

Taking differences of equation (5.21) and imposing $\Delta \mathcal{B} = 0$, yields:

$$- n_U \Delta t_U = n_L \Delta t_L + \omega_H \Delta n_H, \tag{5.40}$$

where I defined $\omega_H \equiv (t_H - t_U - \Delta t_U) + (t_L + \Delta t_L - t_U - \Delta t_U)$ as the tax wedge on skilled labor supply. Combining these two equations by substituting for Δt_U yields:

$$\Delta c_u^* = -\left(1 + \frac{n_L}{n_U}\right) \Delta t_L - \frac{\omega_H}{n_U} \Delta n_H. \tag{5.41}$$

Next, notice that I can write $\Delta G(c_u^*) = \bar{g}(c_u^*, c_u^* + \Delta c_u^*)\Delta c_u^*$. Substituting this, along with Δc_u^*, into equation (5.26), I get:

$$- \bar{g}(c_u^*, c_u^* + \Delta c_u^*)\Delta c_u^* = \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c) - (1 + \gamma)\Delta n_H. \tag{5.42}$$

Substituting for Δc_u^* by using equation (5.41) yields:

$$\left(\frac{n_U + n_L}{n_U}\right) \Delta t_L = \frac{\int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c)}{\bar{g}(c_u^*, c_u^* + \Delta c_u^*)} - \left(\frac{1 + \gamma}{\bar{g}(c_u^*, c_u^* + \Delta c_u^*)} + \frac{\omega_H}{n_U}\right) \Delta n_H. \tag{5.43}$$

Finally, using this equation to substitute for Δt_L in equation (5.29), and substituting for $\bar{u} \equiv \int_{c_s^*}^{c_u^*} u_c \mathrm{d}G(c) / \int_{c_s^*}^{c_u^*} \mathrm{d}G(c)$ and $\int_{c_s^*}^{c_u^*} \mathrm{d}G(c) = \bar{g}(c_s^*, c_u^*)(c_u^* - c_s^*)$, yields:

$$\begin{aligned} \beta \Delta c_s^* + \left(\frac{n_U}{n_U + n_L}\right) \left(\frac{1 + \gamma}{\bar{g}(c_u^*, c_u^* + \Delta c_u^*)} + \frac{\omega_H}{n_U}\right) \Delta n_H = \\ (c_u^* - c_s^*) \left(\frac{\bar{g}(c_s^*, c_u^*)}{\bar{g}(c_u^*, c_u^* + \Delta c_u^*)} \left(\frac{n_U}{n_U + n_L}\right) \bar{u} - u_{c_s^*}\right). \end{aligned} \tag{5.44}$$

Imposing that $\Delta c_s^* > 0$, and, by implication, $\Delta n_H > 0$, yields inequality (5.30).

Postscript: of triangles, gaps, heaps

The research exhibited in this Dissertation allows us to answer one of the defining questions of the scientific discipline of economics, if not of science itself: How many Harberger triangles does it take to fill one Okun gap?[1] The first step towards answering this question came from James Tobin, who in 1977 formulated what is now generally known as the Tobin Conjecture: 'It takes a heap of Harberger Triangles to fill an Okun gap' (Tobin, 1977, p.468). However, far from providing a definitive answer to the Question, the Tobin Conjecture only generated more questions and confusion. Did Tobin provide sufficient argumentation to back up his Conjecture? Could it be proven? What constitutes a *heap*? Especially this last question sparked heated debates, which rage on to this day. Stephen Williamson, for example, evidently believes that $heap = 1$, as he states that

> [i]t takes a heap of Harberger triangles to fill a heap of Harberger triangles. ... The inefficiencies that arise in New Keynesian models are indeed identical to the ones which would be generated by a set of good-specific taxes. ... If I argue that Keynesian sticky wage/price distortions are large, and that tax distortions are small, that's a contradiction.[2]

According to Williamson, there is no difference between a triangle and a gap, which, if the Tobin Conjecture holds, necessarily implies that $heap = 1$. Paul Krugman, on the other hand, seems convinced that $heap = lot > 1$ when he states that:

> Macro Trumps Micro. Or, as the late James Tobin used to say, it takes a lot of Harberger triangles to fill an Okun gap. ... [I]t's a more

[1] I shall henceforth refer to this question as 'the Question.'

[2] Stephen Williamson. Gaps and Triangles. From his weblog, *New Monetarist Economics*, 12 September 2012.

> general observation that even bad microeconomic policies, which lead to substantial distortions in the use of resources, have a hard time doing remotely as much damage as a severe economic slump, which doesn't misallocate resources – it simply wastes them.[3]

In an ultimate attempt to bring the two rival camps closer together, N. Gregory Mankiw calls for more rigorous research on the matter to enable us to quantitatively determine the exact value of *heap* more precisely:

> That [i.e., the Tobin Conjecture] is a great slogan for the Keynesian team. But I agree ... that it would be better to go beyond quips and try to quantify the issue with real data and real models.[4]

Still, even this lucent voice of reason is seduced to speculation on the value of *heap* when he states, not without a hint of drama, that

> Harburger [sic] triangles loom larger now than they did in Tobin's day.[5]

Figure 5.1 gives a visual representation of the answer that my Dissertation provides to the Question. To determine the value of *heap*, we first need to determine *triangle* and *gap*. I define the Harberger triangle as the welfare loss associated with reducing employment through increased distortionary taxation. In the figure, this welfare loss is given by the red triangle of area IV. I define the Okun gap as the welfare loss associated with reducing employment through direct wage controls, causing involuntary unemployment. Or, equivalently, the welfare loss associated with a reduction of employment caused by a combination of a negative productivity shock and downward rigid wages. The size of the Okun gap, then, depends on the rationing schedule. If rationing is efficient, the Okun gap is simply equal to the Harberger triangle given by area IV. However, under inefficient rationing, the Okun gap equals the Harberger triangle plus part of area II. The reason is that under inefficient rationing the employment reduction due to wage controls is partly carried by workers who derive relatively much utility from working. In case of an employment reduction through taxation, on the other hand, only those who least value their job will opt out of the labor market.

This gives us a first rough answer to the Question: it takes at least one Harberger triangle to fill an Okun gap, and potentially more than one, the exact

[3] Paul Krugman. Macro Trumps Micro. From his weblog, *The Conscience of a Liberal*, 12 September 2012.

[4] N. Gregory Mankiw. Triangles vs Gaps. From his weblog, *Greg Mankiw's Blog*, 13 January 2012.

[5] *Ibid.*

amount crucially depending on the particular rationing schedule. The quarrel between Williamson and Krugman, rather than being a matter of fundamental discord, can therefore be reduced to a disagreement about which rationing schedule is most plausible. Williamson, by stating that $heap = 1$, implicitly believes that the rationing schedule is efficient. Krugman, by stating that $heap = lot > 1$, implicitly believes rationing to be inefficient. Indeed, as I mention multiple times in my Dissertation, there is little reason to suspect rationing is anywhere near efficient. In that sense, Krugman seems to have a stronger case than Williamson's.

It is possible to be more precise. Notice that preexisting distortions are not present in Figure 5.1.[6] This would imply that, for only a marginal decline of employment, an infinite number of Harberger triangles fit in one Okun gap. After all, the marginal dead-weight loss of taxation is nil in an undistorted market. To make matters more interesting, I consider the relative sizes of triangles and gaps associated with a reduction of employment, over and above any preexisting distortions in the economy. For this, consider Figure P.1, in which L^D depicts labor demand, L_0^S initial pre-tax labor supply (i.e., labor supply as function of the net wage), and L_1^S post-tax labor supply (i.e., labor supply as function of the gross wage). The red triangle represents the preexisting distortion in this economy, caused by a tax rate t on labor income of, in this case, fifty percent.

For simplicity, let us further focus on a marginal reduction of employment. First consider the Harberger triangle. An employment reduction due to increased taxation causes more voluntary unemployment. That is, people with the lowest utility surplus from work decide to exit the labor market. These people are represented by the position *eff* on the horizontal axis of Figure P.1. Their utility is not directly affected by the employment reduction as they are anyway indifferent between participation and non-participation. The welfare loss therefore only consists of tax-revenue losses, illustrated by the vertical red line at position *eff*, equaling tw. Now consider the Okun gap, which is the welfare loss associated with involuntary unemployment. Again, government loses tax revenue due to the reduction of employment, equaling tw. On top of that, the newly involuntarily unemployed might suffer a direct utility loss, the size of which depends on the rationing schedule. At one extreme, if the newly unemployed are represented by the position *eff* on the horizontal axis, rationing is efficient and there are no additional welfare losses. In that case, the Okun gap equals the Harberger triangle. At the other

[6] I furthermore ignored the possibility that unemployment might cause substitution towards different tax bases, e.g., due to skill formation as in Chapter 3-5 in my Dissertation. Such substitution might lead to both smaller triangles and gaps.

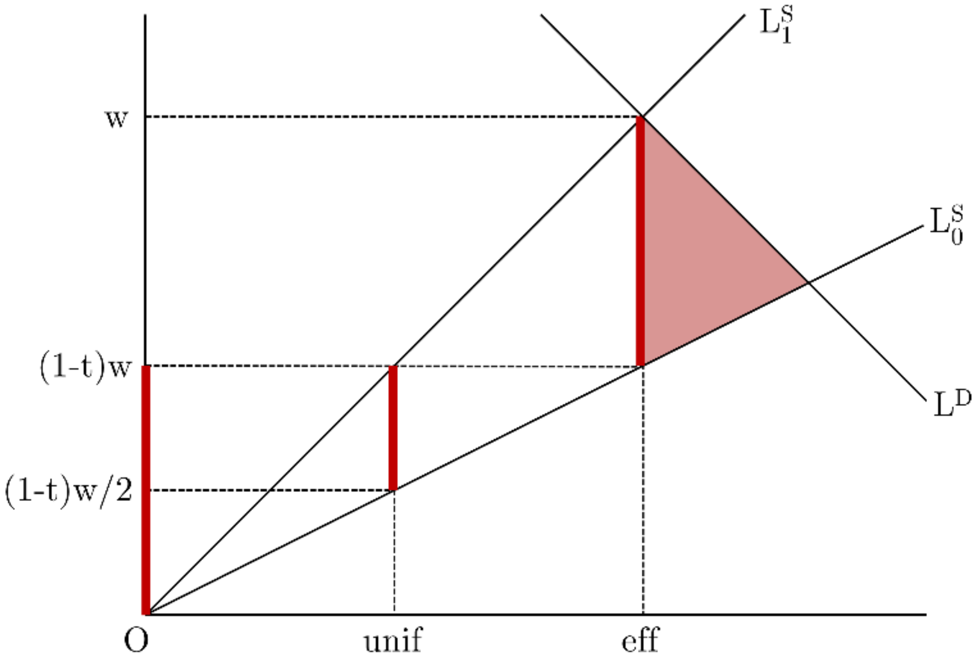

Figure P.1: Marginal Harberger triangle and Okun gap

extreme, if the newly unemployed people are represented by the origin of figure P.1, they suffer utility losses equal to $(1-t)w$. In that case, the Okun gap simply equals w and it takes $w/tw = 2$ Harberger triangles to fill the Okun gap.

Naturally, this conclusion depends on the implicit assumption that the labor supply curve goes through the origin, such that the maximum utility loss of an unemployed worker equals his net wage.[7] Given this assumption, we can determine an upper and a lower bound for the amount of Harberger triangles that fit into one Okun gap. As a lower bound, we know that it takes at least one Harberger triangle to fill an Okun gap. As an upper bound, we know that it takes at most $1/t$ Harberger triangles to fill an Okun gap. As the theoretical or empirical case for either the upper or the lower bound is not very strong, the true answer to the Question is likely to be found somewhere in between these two extremes. Another special case, which might garner more intuitive appeal than either extreme, is the case of uniform rationing – in which every worker faces the same probability of unemployment. Under uniform rationing, and with the additional assumption of a linear labor supply curve, the direct utility losses of the unemployed are

[7]It is easy to provide reasons why this might not be a good assumption, but it is less straightforward to evaluate whether the maximum utility losses should be larger or smaller than the net wage. Valuable leisure suggests the utility losses should be less than the net wage; the observation of charity workers and the idea that employment might be intrinsically valuable – or unemployment intrinsically hurtful – to people suggests the opposite.

Table P.1: How many Harberger triangles to fill an Okun gap? (I)

Tax rate	Upper bound	Uniform rationing	Lower bound
0.00	∞	∞	1.0
0.10	10.0	5.5	1.0
0.20	5.0	3.0	1.0
0.30	3.3	2.2	1.0
0.40	2.5	1.8	1.0
0.50	2.0	1.5	1.0
0.60	1.7	1.3	1.0
0.70	1.4	1.2	1.0
0.80	1.3	1.1	1.0
0.90	1.1	1.1	1.0

The upper-bound and uniform-rationing values are conditional on the assumption that the labor supply curve goes through the origin, and on the assumption that the reduction in employment does not enlarge other tax bases (e.g., due to skill formation). The uniform-rationing values are conditional on the assumption that the labor supply curve is linear. Upper-bound values are given by $1/t$, uniform-rationing values by (1+t)/2t, lower-bound values by 1.

represented by the vertical red line at position *unif* on the horizontal axis of Figure P.1.[8] The direct utility losses of uniform rationing equal half of the utility losses of upper-bound inefficient rationing: $(1-t)w/2$. This yields an Okun gap equal to $tw + (1-t)w/2$. The number of Harberger triangles that fit into this Okun gap, with a tax rate of fifty percent, thus equals $(1+t)/2t = 1.5$.

More generally, Table P.1 provides the number of Harberger triangles it takes to fill an Okun gap for varying levels of distortionary taxation. It is striking to see that both the upper-bound and uniform-rationing values relatively quickly converge to values below 2. The Organisation for Economic Co-operation and Development (OECD) has provided data, for most OECD countries, on the participation tax rate faced by low-income workers earning fifty percent of the average wage – representing a part of the working population that is most vulnerable to unemployment (OECD, 2011*b*). These participation tax rates are determined on the basis of an individual moving from either long-term or short-term unemployment to full-time work. Both long- and short-term measures of t are given in Table P.2 for a variety of countries. Along with the tax rates, the table provides the upper-bound (UB) and uniform-rationing (UR) values of the number of Harberger triangles required to fill an Okun gap.

[8]Both the Harberger triangle and the Okun gap would be larger with convex supply curves, e.g., in the case of a constant labor supply elasticity.

Table P.2: How many Harberger triangles to fill an Okun gap? (II)

	Long-term tax rates			Short-term tax rates		
Country	t	UB	UR	t	UB	UR
Greece	0.04	23.8	12.4	0.92	1.1	1.0
Italy	0.08	11.8	6.4	0.75	1.3	1.2
Slovak Republic	0.31	3.3	2.1	0.41	2.5	1.7
United States	0.31	3.2	2.1	0.58	1.7	1.4
Israel	0.41	2.5	1.7	0.74	1.4	1.2
Portugal	0.44	2.3	1.6	0.81	1.2	1.1
Spain	0.44	2.3	1.6	0.74	1.4	1.2
Estonia	0.46	2.2	1.6	0.62	1.6	1.3
Hungary	0.47	2.1	1.6	0.86	1.2	1.1
Australia	0.49	2.0	1.5	0.49	2.0	1.5
Korea	0.52	1.9	1.5	0.30	3.3	2.1
France	0.53	1.9	1.4	0.76	1.3	1.2
Poland	0.53	1.9	1.4	0.79	1.3	1.1
Canada	0.54	1.8	1.4	0.73	1.4	1.2
Ireland	0.57	1.7	1.4	0.64	1.6	1.3
New Zealand	0.62	1.6	1.3	0.62	1.6	1.3
Belgium	0.64	1.6	1.3	0.77	1.3	1.1
United Kingdom	0.64	1.6	1.3	0.65	1.5	1.3
Finland	0.65	1.5	1.3	0.82	1.2	1.1
Austria	0.66	1.5	1.3	0.78	1.3	1.1
Czech Republic	0.66	1.5	1.3	0.82	1.2	1.1
Sweden	0.68	1.5	1.2	0.88	1.1	1.1
Norway	0.70	1.4	1.2	0.86	1.2	1.1
Luxembourg	0.70	1.4	1.2	0.89	1.1	1.1
Germany	0.72	1.4	1.2	0.82	1.2	1.1
Iceland	0.73	1.4	1.2	0.84	1.2	1.1
Japan	0.75	1.3	1.2	0.67	1.5	1.2
Netherlands	0.76	1.3	1.2	0.82	1.2	1.1
Slovenia	0.79	1.3	1.1	0.91	1.1	1.0
Chile	0.80	1.2	1.1	0.93	1.1	1.0
Denmark	0.96	1.0	1.0	0.94	1.1	1.0

t gives the 2009 participation tax rate associated with moving from long- or short-term unemployment to full-time work at fifty percent of the average wage. Besides labor-income related taxation, it incorporates social, housing, and family benefits, and, for the short-term measure, unemployment benefits. Data shown are the averages of the participation tax rate for four demographic groups: single parent, two children; one-earner married couple, two children, spouse inactive; single; two-earner married couple, two children, spouse earning 67 percent of the average wage. UB and UR give the upper-bound and uniform-rationing values of the required number of Harberger triangles to fill one Okun gap.
Source: OECD (2011*b*) and author's calculations, available on request.

Two things stand out from Table P.2. First, and surprisingly, in the short run an Okun gap consists primarily of the Harberger triangle. The reason for this, as can be seen in the table, is that short-term participation tax rates are relatively large. This implies that most of the losses of moving from employment to unemployment are taken by the government, which sees its tax revenues dwindle. Analogous to this, the unemployed are to a large extent insured against the income losses from short-term unemployment, and consequently suffer little direct utility losses from a temporary employment reduction. Hence, even the upper-bound value of the number of Harberger triangles it takes to fill an Okun gap is below 2 for almost any country. Second, things are decidedly different when it comes to long-term unemployment. Participation tax rates are often much lower in the long run when the unemployed are no longer entitled to unemployment benefits. This implies that a potentially large part of the social losses associated with long-term unemployment is carried by the unemployed themselves, in the form of direct utility losses, rather than by the government. This observation becomes particularly worrisome when we observe, from Table P.2, that this holds especially for South-European countries – Greece and Italy, and to a lesser extent Spain and Portugal – that are currently in a deep economic crisis, with high unemployment rates that might persist for a long time. Thus, for the countries that currently suffer most from high unemployment, it may take many Harberger triangles to fill their Okun gap.

This brings us back to the Tobin Conjecture and the speculative remarks by Williamson, Krugman, and Mankiw. On the basis of my Dissertation, which introduced the concept of a rationing schedule, we can conclude that an answer to the Question heavily depends on the rationing schedule and on preexisting distortions. In the short run, Williamson's assertion, that the Okun gap *is* a Harberger triangle is approximately correct. However, for longer-lasting spells of unemployment, which cause the unemployed to lose their unemployment benefits, direct utility losses might take up a large part of the social losses associated with unemployment. In that case, Krugman might be very right by stating that during severe economic slumps it takes a lot of Harberger triangles to fit an Okun gap.

Nederlandse samenvatting (summary in Dutch)

In deze samenvatting van het proefschrift "Essays over optimaal overheidsbeleid" concentreer ik me op twee verschillende aspecten. Allereerst waag ik een poging om in een aantal pagina's aan te geven hoe het proefschrift aansluit op de orthodoxie van publieke economie. Hierbij ontkom ik er niet aan een korte beschouwing te geven van deze academische discipline, om op die manier het contrast tussen het proefschrift en de rest van de relevante literatuur helder weer te kunnen geven. In het tweede deel van de samenvatting richt ik me op de vraag wat het proefschrift ons leert over optimaal overheidsbeleid. Per hoofdstuk bespreek ik de concrete beleidsimplicaties die uit mijn onderzoek voortvloeien.

Relatie tot de orthodoxie van publieke economie

Publieke economie

Zoals de titel doet vermoeden, probeert het proefschrift een bijdrage te leveren aan de discussie over wat goed overheidsbeleid behelst. Deze discussie staat centraal binnen de academische discipline van de publieke economie. Publieke economie wordt doorgaans gezien als een van vele academische subdisciplines van economie en wordt ook op die manier gedoceerd aan universiteiten. De onderliggende vraagstelling van publieke economie verschilt echter radicaal met die van economie. Het doel van economie kan omschreven worden als de verklaring van menselijk gedrag, terwijl publieke economie zich de bepaling van optimaal overheidsbeleid ten doel heeft gesteld. De enorme ambitie van dit doel blijkt uit zijn multidisciplinaire karakter. Voordat iets zinnigs kan worden gezegd over wat optimaal overheidsbeleid inhoudt, moet er op zijn allerminst uitsluitsel zijn over wat een overheid zou moeten nastreven, over het effect van overheidsbeleid op menselijk gedrag, en over de nawerking van dit gedrag op het welzijn van mensen. In andere woorden, bij de bepaling van optimaal overheidsbeleid baseert de publieke econoom zich onontkoombaar op inzichten uit de politieke filosofie, de economie, en de psychologie van het welzijn. Hieronder schets ik in het kort welke van deze inzichten deel uitmaken van de orthodoxie van publieke economie, en in hoeverre dit proefschrift daarin meegaat of van afwijkt.

Politieke filosofie of wat behoort de overheid na te streven?

Om te bepalen wat optimaal overheidsbeleid inhoudt, moet eerst een antwoord worden geformuleerd op de vraag wat een overheid zou moeten nastreven. Deze vraag speelt een prominente rol binnen politieke filosofie en er bestaat een grote verscheidenheid aan theorieën die hier een antwoord op pogen te geven. Zo zijn er theorieën die stellen dat overheidsbeleid gericht hoort te zijn op zogenaamde 'primaire goederen' (cf. Rawls, 1971), 'capabilities' (cf. Sen, 1992), minimale interventie (cf. Nozick, 1974), of op de consequenties van een beperkt aantal aangeboren ongelijkheden voor welke individuen zelf niet verantwoordelijk gehouden mogen worden (cf. Dworkin, 2000). Verreweg de meest invloedrijke theorie, binnen publieke economie althans, staat bekend als welfarisme (cf. Kaplow and Shavell, 2002). Welfarisme stelt dat het uiteindelijke doel van overheidsbeleid de maximalisatie van het welzijn van individuen behoort te zijn.[1]

Gezien het aantal concurrerende theorieën over het 'juiste' doel van overheidsoptreden is welfarisme niet geheel onomstreden, maar er zijn een aantal belangrijke argumenten aan te dragen ter verdediging ervan. Zo kan beredeneerd worden dat dit het doel is dat iemand zou kiezen in de hypothetische situatie waarbij hij onwetend is over welk leven hij precies zal leiden. Ook kan aangetoond worden dat het, in tegenstelling tot eerder genoemde theorieën, nooit kan leiden tot regel-fetisjisme, waarbij een situatie waarin iedereen slechter af is wordt geprefereerd omdat het voldoet aan het door de theorie gestelde doel. Onder andere vanwege deze redenen, neem ik het door welfarisme gestelde doel van overheidsoptreden over in mijn proefschrift: bij de bepaling van optimaal overheidsoptreden neem ik zodoende aan dat de overheid een som van individueel welzijn behoort te maximaliseren.[2]

[1] Welfarisme impliceert niet dat het doel *per se* een ongewogen som van welzijn behoort te zijn, maar staat bijvoorbeeld toe dat het welzijn van mensen die relatief slecht af zijn zwaarder telt dan het welzijn van mensen die relatief goed af zijn. Hierdoor is het algemener dan utilitarisme (cf. Bentham, 1907; Mill, 1863; Sidgwick, 1874).

[2] Dit neemt niet weg dat er geen problemen zijn met welfarisme. Een van de belangrijkste bezwaren is de observatie dat een fenomeen als racisme te rechtvaardigen is aan de hand van welfarisme als hieraan maar genoeg mensen welzijn ontlenen. Aangezien ik in mijn proefschrift aanneem dat het welzijn van een individu onafhankelijk is van andere individuen zal een dergelijke

Economie of de gedragseffecten van overheidsbeleid

Een antwoord op de vraag wat een overheid na zou moeten streven is nog altijd ver verwijderd van een antwoord op wat optimaal overheidsbeleid is. Eerst moet bepaald worden in welke mate de beleidsinstrumenten van een overheid bijdragen aan het doel van hoger individueel welzijn. De gedragseffecten van overheidsoptreden zijn hierbij van cruciaal belang. Stel dat herverdeling van werkenden naar werklozen beschouwd kan worden als welzijnsverhogend. Dit zegt nog weinig over de wenselijkheid van hogere werkloosheidsuitkeringen. Hoewel deze aan de ene kant meer welzijn op zullen leveren vanwege de herverdeling, verstoort het de beslissing van werklozen om te zoeken naar een nieuwe baan. Een dergelijke verstoring zorgt voor lagere werkgelegenheid en daarmee voor lagere belastingopbrengsten. Vanwege dit gedragseffect kan de overheid, voor iedere euro die het door middel van belastingen ophaalt bij de werkenden, minder dan een euro herverdelen naar werklozen. Een soortgelijke *trade-off* karakteriseert vrijwel ieder beleidsinstrument. Een minimumloon, bijvoorbeeld, herverdeelt inkomen van bedrijven naar laaggeschoolden. Hoewel dit *an sich* wellicht als welzijnsverhoging kan worden gezien, leidt het tot een gedragseffect waarbij bedrijven, om hun winstmarges op peil te houden, minder laaggeschoolden aan zullen nemen.

De identificatie en kwantificatie van deze trade-offs is iets waar economie bij uitstek voor is geschikt. Economie is immers de wetenschap die als taak heeft om menselijk gedrag te verklaren en voorspellen. Voor de identificatie van gedragseffecten maken economen intensief gebruik van economische modellen. Deze modellen zijn bijzonder gestileerde weergaven van de werkelijkheid, waarbij alle niet-essentiële elementen buiten beschouwing worden gelaten. Deze aanpak wordt veelal bekritiseerd door niet-economen, om zowel slechte als goede redenen. Aan de ene kant is een hoge mate van abstractie van essentieel belang om iets zinnigs te kunnen zeggen over een extreem complexe werkelijkheid. Aan de andere kant moet de econoom zijn beslissingen om te abstraheren van specifieke elementen van de werkelijkheid altijd kunnen verdedigen. Een belangrijke abstractie, die behoort tot de orthodoxie van publieke economie en een prominente rol speelt in mijn proefschrift, is dat prijzen en loonvoeten flexibel zijn. Onder deze aanname zullen lonen en prijzen zich altijd aanpassen om te garanderen dat vraag en aanbod aan elkaar gelijk zijn. Dit impliceert ook dat onvrijwillige werkloosheid niet kan bestaan.

laakbare situatie niet voor kunnen komen. Inderdaad ben ik van mening dat overheidsbeleid gericht moet zijn op zogenaamde 'witgewassen' welzijn, dat wil zeggen, het gedeelte van welzijn dat onafhankelijk is van andere individuen (cf. Harsanyi, 1982).

Psychologie van het welzijn

Tot slot, zelfs wanneer we welfarisme accepteren en de economische wetenschap ons perfect weet te informeren over de precieze gedragseffecten van overheidsbeleid, weten we nog altijd niet wat optimaal overheidsbeleid behelst. Dit vereist namelijk inzicht in hoe beleid, en haar gedragseffecten, het welzijn van mensen beïnvloeden. De verklaring van mentale processen, waar welzijn onder valt, maakt deel uit van het vakgebied van de psychologie. De consequente toepassing van welfarisme vereist daarom een gedegen kennis van de psychologie van het welzijn. Een elegante truc om hier onderuit te komen, een truc die routineus door de orthodoxie van publieke economie wordt gehanteerd, is de aanname dat mensen met hun gedrag altijd hun eigen welzijn maximaliseren. Onder die aanname zal niemand 'verkeerd' handelen en hoeven mensen niet 'bijgestuurd' te worden door de overheid. Bovendien zorgt deze aanname ervoor dat economen op basis van geobserveerd gedrag de impact van overheidsbeleid op het welzijn van mensen kan bepalen, zonder kennis van mentale processen.

Het proefschrift

Met dit proefschrift probeer ik een eigen bijdrage te leveren aan publieke economie. Hierbij wijk ik op een aantal punten expliciet af van de orthodoxie. Allereerst verwerp ik de aanname dat menselijk gedrag altijd en overal een kwestie van welzijnsmaximalisatie is. Hierbij baseer ik me op inzichten uit de psychologie, gedragseconomie, en neurowetenschappen. Los van deze inzichten, wijst introspectie erop dat veel menselijk gedrag voortkomt uit factoren die weinig te doen hebben met de maximalisatie van welzijn. Als we niet langer aannemen dat gedrag gebaseerd is op welzijnsmaximalisatie kan het voorkomen dat individuen zich niet in hun private welzijnsoptimum bevinden. Dit creëert een nieuwe rol voor overheidsbeleid, en wel het 'corrigeren' van individueel gedrag. In het eerste inhoudelijke hoofdstuk van het proefschrift, hoofdstuk 2, bepaal ik de consequenties hiervan voor de optimale inkomstenbelasting.

Hoofdstukken 3 en 4 concentreren zich op het minimumloon als beleidsinstrument. Een belangrijk verschil met eerdere literatuur is dat ik expliciet de onderwijskeuzen van individuen modelleer, een beslissing waar andere studies over het minimumloon veelal van abstraheren. Het effect van een minimumloon op deze keuze is theoretisch evenwel ambigu: hogere lonen voor laaggeschoolden geven prikkels om minder te investeren in onderwijs, maar hogere werkloosheid onder laaggeschoolden geeft juist prikkels om meer te investeren. In hoofdstuk 3 analyseer ik hoe het effect van een minimumloon op onderwijskeuzes de wenselijkheid van een minimumloon, ten opzichte van inkomstenbelastingen, beïnvloedt.

Hoofdstuk 4 draagt bij aan de literatuur over het minimumloon door een andere veel gemaakte theoretische aanname te verwerpen. Deze aanname heeft betrekking op de verdeling van de door het minimumloon gecreëerde werkloosheid over de laagopgeleide bevolking. Eerdere literatuur neemt veelal aan dat ontslagen werknemers degenen zijn die het minste welzijn ontlenen van hun baan (zogenaamde 'efficiënte arbeidsrantsoenering'), of, zoals in hoofdstuk 3, dat iedere laagopgeleide werknemer een even grote kans op werkloosheid heeft (zogenaamde 'uniforme arbeidsrantsoenering'). De theoretische of empirische rechtvaardiging voor deze aannamen is echter flinterdun. In hoofdstuk 4 maak ik geen specifieke aanname over de verdeling van werkloosheid, maar analyseer ik hoe de wenselijkheid van het minimumloon afhangt van deze verdeling. Bovendien laat ik zien hoe deze wenselijkheid empirisch kan worden getest zonder aannamen over de verdeling van werkloosheid.

In hoofdstuk 5 analyseer ik tot slot de consequenties van een inflexibele loonvloer voor optimale belastingen en arbeidsparticipatiebeleid. Net als in de vorige twee hoofdstukken wijk ik hiermee af van de orthodoxe aanname dat lonen zich perfect aanpassen om vraag en aanbod gelijk te stellen. De analyse van dit hoofdstuk is bijvoorbeeld relevant in arbeidsmarkten die gedomineerd worden door vakbonden, wier looneisen niet gericht zijn op het gelijkstellen van vraag en aanbod. Een andere situatie waarin de analyse van bijzondere relevantie is, is in tijden van economische recessie. Het is overduidelijk dat onvrijwillige werkloosheid tijdens een recessie over het algemeen relatief hoog is, wat impliceert dat de orthodoxe aanname van flexibele lonen in een dergelijke situatie sterk aan relevantie inboet.

Bevindingen over optimaal overheidsbeleid

Optimale belastingen en welzijn (hoofdstuk 2)

In hoofdstuk 2 van mijn proefschrift stap ik af van de standaardaanname dat individuen, bij het maken van beslissingen, altijd hun eigen welzijn maximaliseren. Net als in de rest van de economische literatuur, ga ik er nog altijd vanuit dat menselijk gedrag beschreven kan worden als een maximalisatie van zogenaamd 'nut.' Maar ik verwerp de aanname dat het nut van een persoon identiek is aan zijn welzijn – dat wat zijn leven uiteindelijk waardevol maakt. Dit impliceert dat hij 'fouten' kan maken in zijn gedrag door iets anders dan zijn eigen welzijn na te streven.

Vervolgens bepaal ik hoe dit de optimale belastingstructuur beïnvloedt. Een belastingvoet is alleen optimaal als de netto sociale baten van een kleine verhoging van de belasting gelijk zijn aan nul. Immers, als de netto sociale baten strikt positief (negatief) zouden zijn, loont het voor de overheid om de belastingvoet te verhogen (verlagen). De sociale baten bestaan uit wat de overheid kan doen met de belastingopbrengst: investeren in publieke goederen bijvoorbeeld, of herverdelen naar mensen met lage inkomens. De sociale kosten bestaan uit een lager inkomen van de belastingbetaler, en een verstoring van individueel gedrag.

Deze verstoring van individueel gedrag is cruciaal voor mijn onderzoek. Een kleine verhoging van de belasting op arbeidsinkomen zorgt ervoor dat werknemers minder werken. Deze afname in arbeidsaanbod leidt op zijn beurt tot een lagere belastingopbrengst, wat een sociale kostenpost is. Als individuen hun eigen welzijn maximaliseren, leidt een dergelijke afname van arbeidsaanbod evenwel niet tot een verandering van individueel welzijn. De reden hiervoor is dat welzijnsmaximalisatie van het individu impliceert dat een kleine verandering in zijn gedrag geen invloed kan hebben op zijn welzijn. Dit verandert wanneer ik de aanname van individuele welzijnsmaximalisatie verwerp. In dat geval heeft een afname van arbeidsaanbod ook een direct effect op het welzijn van het individu. Het leidt tot een hoger welzijn als werknemers oorspronkelijk harder werkten dan goed voor hen is; het leidt tot een lager welzijn als werknemers oorspronkelijk minder hard werkten dan goed voor hen is.

Dit creëert een extra reden voor het gebruik van verstorende belastingen. Een situatie waarin werknemers 'te veel' werken pleit voor een hogere marginale belastingvoet om deze werknemers te prikkelen minder hard te werken. Een situatie waarin werknemers 'te weinig' werken pleit daarentegen voor een lagere optimale belastingvoet om deze werknemers te prikkelen harder te werken. Dezelfde logica gaat op voor belastingen of subsidies op onderwijs of specifieke goederen. Als mensen uit eigen beweging minder onderwijs genieten dan welzijnsmaximalisatie impliceert, dan is dit een reden voor hogere subsidies op onderwijs.

De vraag of mensen daadwerkelijk 'te veel' of 'te weinig' werken is een empirische kwestie. Om deze vraag te kunnen beantwoorden heb ik een maatstaf van welzijn nodig. In mijn onderzoek gebruik ik hiervoor het antwoord, op een schaal van 1 tot 7, op de vraag: "How dissatisfied or satisfied are you with your life overall?". Dankzij de British Household Panel Survey beschik ik over jaarlijkse antwoorden op deze vraag van ongeveer 28,000 Britse individuen, gevolgd over 12 jaar. Daarnaast beschik ik over een groot aantal andere variabelen, zoals netto inkomen en het aantal gewerkte uren per week. Op basis van deze data bepaal ik hoe het welzijn van individuen afhangt van zowel hun netto inkomen als het aantal gewerkte uren, wat me in staat stelt om per individu het optimale aantal arbeidsuren te berekenen. Na vergelijking met het werkelijke aantal gewerkte uren concludeer ik of een individu te veel of te weinig werkt.

De resultaten van deze empirische analyse leiden mij tot de conclusie dat Britse werknemers met relatief lage inkomens gemiddeld te weinig werken. Zij zouden hun welzijn kunnen verhogen door wekelijks meer uren te werken. Voor werknemers met hoge inkomens geldt het tegenovergestelde: deze werknemers stoppen wekelijks te veel uren in hun baan, waarmee ze hun eigen welzijn schaden. De beleidsimplicatie die hieruit volgt is dat de Britse overheid lagere marginale belastingtarieven voor lage inkomens, en hogere marginale belastingtarieven voor hoge inkomens zou moeten implementeren. Een dergelijke belastinghervorming geeft prikkels aan werknemers met lage inkomens om meer te werken, en aan werknemers met hoge inkomens om minder te werken. Op die manier corrigeert deze hervorming deels het welzijnssuboptimale gedrag van Britse werknemers.

Optimale minimumlonen en scholing (hoofdstuk 3)

Hoofdstuk 3 concentreert zich op de welvaartseffecten van een minimumloon. Hierbij wijkt het af van het merendeel van de literatuur over minimumlonen door de scholingsbeslissingen van individuen expliciet te modelleren. Deze innovatie is ingegeven door het feit dat het effect van een minimumloon op scholing *a priori* ambigu is. Aan de ene kant leidt een minimumloon tot een hoger loon voor laaggeschoolden, wat een prikkel geeft om minder te investeren in scholing. Aan de andere kant leidt een minimumloon tot hogere werkloosheid onder laaggeschoolden omdat het minder aantrekkelijk wordt hen aan te nemen. Deze hogere werkloosheid onder laaggeschoolden geeft een prikkel om juist meer te investeren in scholing.

We[3] laten allereerst zien dat het netto effect op scholing in belangrijke mate afhangt van hoe eenvoudig bedrijven hooggeschoolden kunnen substitueren voor laaggeschoolden. Hoe eenvoudiger dit is, hoe meer een bedrijf haar productieproces zal omschakelen naar hooggeschoolden wanneer een hoger minimumloon laaggeschoolden duurder maakt. De substitutie-elasticiteit tussen hoog- en laaggeschoolden – de procentuele verandering in de verhouding tussen hooggeschoolde

[3]Dit hoofdstuk is gezamenlijk werk met Bas Jacobs.

en laaggeschoolde werkgelegenheid, als reactie op een procentuele verandering in de verhouding tussen de hooggeschoolde en laaggeschoolde loonvoet – blijkt hierbij van groot belang. Als deze elasticiteit kleiner is dan 1, dan leidt een hoger minimumloon tot relatief weinig werkloosheid en daardoor tot minder scholing. Is deze elasticiteit groter dan 1, dan leidt een minimumloon tot dusdanig veel werkloosheid onder laaggeschoolden dat meer mensen besluiten zich om te scholen tot hooggeschoolde werknemer. Aangezien empirische schattingen van deze elasticiteit doorgaans hoger zijn dan 1, maakt mijn model het aannemelijk dat een minimumloon een positief scholingseffect heeft.

Vervolgens bepalen we hoe de scholingsbeslissing de wenselijkheid van een minimumloon beïnvloedt. Hierbij concentreren we ons op het geval van perfecte concurrentie op de arbeidsmarkt. Eerder studies hebben aangetoond dat een minimumloon economische doelmatigheid kan bevorderen in monopsonistische arbeidsmarkten. Onze bevindingen zijn conditioneel op de aanname dat er geen sprake is van dergelijke monopsonie. In andere woorden, we concentreren ons op de herverdelende rol van het minimumloon, niet op zijn rol ter correctie van imperfecte concurrentie.

Als we belastingen buiten beschouwing laten, blijkt dat een bindend minimumloon twee verschillende effecten heeft op de sociale welvaart. Aan de ene kant leidt het tot hogere lonen voor laaggeschoolden, welke bedrijven betalen door lagere lonen te bieden aan hooggeschoolden. Dit behelst een positief welvaartseffect onder de plausibele veronderstelling dat het inkomen van laaggeschoolden hoger wordt gewaardeerd dan het inkomen van hooggeschoolden.[4] Aan de andere kant zullen bedrijven minder geneigd zijn om laaggeschoolden aan te nemen vanwege toegenomen loonkosten. Dit leidt tot hogere werkloosheid, wat een negatief effect heeft op sociale welvaart, onder de veronderstelling dat onvrijwillig werklozen liever wel een baan hebben. De wenselijkheid van een minimumloon is dus ambigu en hangt af van het netto welvaartseffect van meer herverdeling en hogere werkloosheid.

De sociale kosten van een minimumloon nemen toe op het moment dat een inkomstenbelasting wordt geïntroduceerd. Terwijl de baten nog steeds bestaan uit eenzelfde herverdelingswinst, nemen de kosten van werkloosheid toe, aangezien een toename in werkloosheid nu ook leidt tot verlies van belastingopbrengsten. Evengoed kan een minimumloon de herverdeling via inkomstenbelasting complementeren op het moment dat inkomstenbelasting niet perfect toegespitst kan worden op

[4] Het gebruikelijke argument stelt dat een euro aan consumptie een groter positief effect heeft op het leven van een arm persoon dan op het leven van een rijk persoon.

loonvoeten of scholingsniveau. De reden hiervoor is dat een minimumloon direct het nettoloon van laaggeschoolden verhoogt en die van hooggeschoolden verlaagt. Met belastingen is een dergelijke herverdeling alleen mogelijk als deze geconditioneerd kunnen worden op het scholingsniveau. Als de overheid inderdaad de belastingtarieven af kan laten hangen van het scholingsniveau, kan ze een herverdeling teweegbrengen die identiek is aan die van een minimumloonsverhoging door het belastingtarief voor laaggeschoolden te verlagen en voor hooggeschoolden te verhogen. In vergelijking met een dergelijke belastinghervorming, leidt een minimumloonsverhoging echter nog altijd tot hogere arbeidskosten voor bedrijven, en dus tot hogere werkloosheid. Deze werkloosheid leidt vervolgens tot meer scholing, terwijl progressieve belasting juist tot een neerwaartse verstoring van de scholingsbeslissing leidt. De rol van het minimumloon bestaat in dit geval dus niet uit de herverdeling van inkomen, maar uit het tegengaan van scholingsverstoringen, veroorzaakt door een progressieve belastingstructuur. Een minimumloon is in dat geval wenselijk als de sociale baten van meer scholing opwegen tegen de kosten van hogere werkloosheid.

Optimale minimumlonen en verdeling van werkloosheid (hoofdstuk 4)

Hoofdstuk 4 bouwt voort op de resultaten van het voorgaande hoofdstuk, waarin we concludeerden dat een bindend minimumloon wenselijk is als de sociale baten van scholing groter zijn dan de sociale kosten van hogere werkloosheid. In hoofdstuk 4 laten we[5] zien dat deze conclusie geldig is onder zeer algemene aannamen. Zo nemen we algemene nutsfuncties aan en staan we toe dat individuen niet alleen beslissen over het aantal uren werk dat ze verrichten en of ze hoog- of laaggeschoold zijn, maar ook of ze participeren in de arbeidsmarkt of vrijwillig werkloos zijn. Een cruciale generalisatie in dit model betreft het zogenoemde rantsoeneringsschema. Dit schema is een theoretisch concept dat beschrijft hoe werkloosheid verdeeld wordt onder laaggeschoolden. Eerdere literatuur neemt veelal aan dat deze rantsoenering efficiënt is in de zin dat werkloosheid slechts laaggeschoolden treft die weinig welzijn ontlenen van werk. Hoofdstuk 3 neemt daarentegen aan dat de kans op werkloosheid gelijk is voor iedere laaggeschoolde. In hoofdstuk 4 stappen we af van dergelijke specifieke aannamen, en bepalen we de wenselijkheid van een minimumloon onder een algemeen rantsoeneringsschema. Dit is belangrijk omdat er geen sterke theoretische of empirische onderbouwingen bestaan voor één specifiek schema, terwijl dit schema wel een cruciale factor blijkt te zijn voor de

[5] Ook dit hoofdstuk is gezamenlijk werk met Bas Jacobs.

welvaartseffecten van een minimumloon.

Om te bepalen of een minimumloon wenselijk is in een dergelijke opzet, leiden we de gevolgen af van een zogenoemde netto-inkomensneutrale (NIN) minimumloonsverhoging. Als deel van deze hervorming verhoogt de overheid het minimumloon, terwijl ze de inkomstenbelasting zo aanpast dat het netto inkomen van werknemers constant blijft. De welvaartsgevolgen van een dergelijke NIN minimumloonsverhoging geven een duidelijk beeld van hoe een minimumloon zich onderscheidt van een soortgelijke herverdeling via de inkomstenbelasting. Aangezien het nettoloon van werknemers gelijk blijft, heeft de NIN hervorming geen effect op het aantal uren dat werknemers werken. Omdat de opbrengsten van arbeid constant blijven heeft de hervorming ook geen effect op de arbeidsparticipatiebeslissing van individuen. De hervorming verhoogt echter wel de kosten voor bedrijven om laaggeschoolden aan te nemen en leidt daarmee tot een lagere vraag naar laaggeschoold arbeid. Onder gelijkblijvend aanbod en dalende vraag zal de onvrijwillige werkloosheid onherroepelijk stijgen. Een hogere werkloosheid impliceert op haar beurt echter een lagere verwachte opbrengst van laaggeschoold werk, wat individuen een prikkel geeft om meer te investeren in scholing. De hervorming leidt dus tot zowel meer werklozen als meer hooggeschoolden.

De reden dat het rantsoeneringsschema zo belangrijk is voor de welvaartseffecten van een minimumloon is nu eenvoudig te begrijpen. Als werkloosheid voornamelijk plaats vindt onder laaggeschoolden die liever werkloos zijn dan investeren in scholing, zal een minimumloon voornamelijk zorgen voor hogere werkloosheid, zonder dat dit tot hogere scholing leidt. Als werkloosheid echter voornamelijk plaats vindt onder laaggeschoolden die net zo lief hooggeschoold zijn, zal een minimumloon juist een groot positief effect hebben op scholing. In het hoofdstuk leveren we het bewijs dat een minimumloon altijd optimaal 'gemaakt' kan worden door de juiste aanname te maken omtrent de verdeling van werkloosheid. Op dezelfde manier kan een minimumloon altijd suboptimaal gemaakt worden. Om deze reden is de wenselijkheid van een minimumloon vanuit theoretisch opzicht fundamenteel ambigu.

Om aan deze fundamentele ambiguïteit te ontsnappen herschrijven we de conditie voor de wenselijkheid van een minimumloon in termen van zogenaamde *sufficient statistics*. Dit zijn empirisch te bepalen grootheden die de wenselijkheid van een minimumloon weergeven zonder de noodzaak om diepere parameters uit het model (zoals het rantsoeneringsschema) empirisch te identificeren. We laten zien dat we genoeg hebben aan drie statistieken: de belastingopbrengsten van een extra hooggeschoolde, het verlies aan belastingopbrengsten vanwege een extra werkloze, en een elasticiteit die weergeeft hoe sterk het aantal hooggeschoolden reageert op werkloosheid.[6] Data over belastingopbrengsten van scholing en werkloosheid vinden we in publicaties van de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) voor een groot aantal OESO-landen. Op basis van deze cijfers kunnen we concluderen dat, voor alle OESO-landen behalve de Verenigde Staten, een verhoging van het minimumloon alleen wenselijk is als een procentpunt hogere werkloosheid zou leiden tot een toename van het scholingspercentage van meer dan 0,6 procentpunt. Voor de Verenigde Staten geldt een minimaal vereiste toename van het scholingspercentage van 0,4 procentpunt.[7] Als we ons richten tot de empirische literatuur die het scholingseffect van werkloosheid meet, vinden we dat een procentpunt toename van de werkloosheid typisch gepaard gaat met een toename in het scholingspercentage van tussen de 0,1 en 0,6 procentpunt. Hieruit concluderen we dat een NIN minimumloons*verlaging* wenselijk is voor alle landen binnen onze analyse – behalve mogelijkerwijs de Verenigde Staten. Een dergelijke hervorming zorgt voor zowel een toename in belastingopbrengsten, als voor een hoger welzijn van individuen die als gevolg van het lagere minimumloon een baan kunnen krijgen. Een verlaging van het minimumloon maakt dus deel uit van een zogeheten Pareto-verbeterende hervorming, als gevolg waarvan sommigen beter af zullen zijn en niemand slechter af. Voor landen die geen wettelijk minimumloon kennen – zoals Duitsland, Oostenrijk, Italië, of de Scandinavische landen – is het onwenselijk om deze alsnog in te voeren.

Optimale belastingen en werkloosheid (hoofdstuk 5)

In het laatste hoofdstuk van mijn proefschrift sta ik stil bij de consequenties van onvrijwillige werkloosheid voor optimale belastingen. Net als in het voorgaande hoofdstuk, ga ik er vanuit dat het *a priori* onbekend is hoe de werkloosheid verdeeld is over de beroepsbevolking. Ik laat zien dat de tegenwoordigheid van onvrijwillige werkloosheid leidt tot conclusies omtrent belastingen die volledig haaks staan op conventionele wijsheden uit de publieke economie, waarbij normaliter uitgegaan wordt van volledige werkgelegenheid. Eén van deze conventionele wijsheden stelt dat een hogere belastingvoet op inkomen aan de ene kant een hogere mate van

[6] Merk op dat we hiermee impliciet aannemen dat het directe welzijnsverlies van iemand die werkloos raakt verwaarloosbaar is. Daarmee schrijven we de conditie voor de wenselijkheid van een minimumloon dus sterk in het voordeel van een minimumloon.

[7] Zie kolom 5 in tabel 1 van hoofdstuk 4 voor specifieke waarden voor alle OESO-landen waarvoor voldoende data beschikbaar is.

herverdeling mogelijk maakt, maar aan de andere kant leidt tot een verstoring van arbeidsaanbod. Een hogere belastingvoet leidt immers tot lagere netto lonen en dus tot minder prikkels om te werken. Het lagere arbeidsaanbod zorgt op zijn beurt voor een afname in belastinginkomsten, wat een welvaartsverlies vertegenwoordigt. De optimale belastingvoet staat daarom in het teken van een *trade-off*: de herverdelingswaarde van hogere belastingopbrengsten versus de efficiëntiekosten van lager arbeidsaanbod.

Onvrijwillige werkloosheid impliceert echter dat het aanbod van arbeid groter is dan de vraag naar arbeid. Als daarbij ook nog sprake is van inefficiënte arbeidsrantsoenering wil dat zeggen dat sommige werklozen meer welzijn van een baan zouden ontlenen dan sommige werknemers die wél een baan hebben. In dat geval leidt een hogere belastingvoet voor iedere individuele werknemer nog altijd tot een lagere prikkel tot werken, waardoor sommige werknemers besluiten minder of geen arbeid meer aan te bieden. Maar omdat het aanbod van arbeid toch al groter was dan de vraag naar arbeid, leidt dit op geaggregeerd niveau niet tot lagere werkgelegenheid. Dat sommige mensen minder arbeid aanbieden door de belastingverhoging zorgt juist voor werkgelegenheid voor mensen die aanvankelijk onvrijwillig werkloos waren. Aangezien arbeidsaanbod op geaggregeerd niveau niet verandert is er geen sprake meer van een efficiëntieverlies door dalende belastingopbrengsten. In tegendeel: er is sprake van een efficiëntieverbetering als mensen die weinig welzijn ontlenen aan werk besluiten minder te werken en daarmee werk creëeren voor werklozen die daar meer welzijn aan ontlenen. Eenzelfde conclusie geldt voor uitkeringen, zoals de bijstandsuitkering. In afwezigheid van onvrijwillige werkloosheid leiden hogere uitkeringen tot minder arbeidsaanbod en dus tot een efficiëntieverlies. In de tegenwoordigheid van werkloosheid leiden hogere uitkeringen ertoe dat werknemers die weinig welzijn ontlenen aan een baan plaatsmaken voor werklozen die daar meer welzijn aan ontlenen. In andere woorden, door belastingen en uitkeringen te verhogen substitueert de overheid vrijwillige werkloosheid voor onvrijwillige werkloosheid, wat resulteert in een efficiëntiewinst.

In tegenstelling tot de conventionele trade-off tussen gelijkheid en efficiëntie leidt een belasting- of uitkeringsverhoging dus tot een hogere mate van zowel gelijkheid als efficiëntie. De beleidsimplicaties liggen dan ook voor de hand: verhoog de effectieve belastingvoet binnen arbeidsmarktsegmenten die gekenmerkt worden door hoge onvrijwillige werkloosheid, en gebruik de extra belastingopbrengsten om uitkeringen te verhogen. Vanuit een dynamisch perspectief zou de overheid belastingen en uitkeringen moeten verhogen als de economie zich in een recessie met hoge werkloosheid bevindt, en weer verlagen bij lage werkloosheid. Deze conclusie heeft ook consequenties voor overheidsbeleid dat gericht is op het bevorderen van arbeidsmarktparticipatie. Als de arbeidsmarkt gekenmerkt wordt door onvrijwillige werkloosheid heeft het geen zin om de participatie te bevorderen. Dergelijk beleid zorgt er alleen voor dat mensen die net zo lief niet werken toetreden tot de arbeidsmarkt en zodoende om dezelfde banen concurreren met mensen die wel graag een baan zouden willen hebben.

Verder laat ik zien onder welke voorwaarden deze beleidsimplicaties overeind blijven bij een endogene scholingsbeslissing. De resultaten zijn vergelijkbaar met het voorgaande hoofdstuk: als werkloosheid verdeeld is onder laaggeschoolden die net zo lief hooggeschoold zouden zijn, kan werkloosheid eventueel sociaal nut hebben door scholing te bevorderen. In dat geval zou de overheid eventueel niet alle onvrijwillige werkloosheid willen vervangen door vrijwillige werkloosheid. Daarnaast laat ik zien dat de beleidsimplicaties onveranderd blijven als de onvrijwillige werkloosheid veroorzaakt wordt door vakbonden. Als de overheid zich committeert aan een hogere belastingvoet wanneer de werkloosheid stijgt, dan zullen vakbonden minder geneigd zijn tot hogere looneisen omdat ze kunnen anticiperen op hogere belastingen. Als gevolg leidt een dergelijk beleid tot lagere onvrijwillige werkloosheid. Tot slot laat ik zien dat ook de conventionele theorie omtrent de weerslag van belastingen komt te vervallen bij de tegenwoordigheid van onvrijwillige werkloosheid. De conventionele theorie stelt dat het irrelevant is of de arbeidsinkomstenbelasting geheven wordt op werknemers of werkgevers. In beide gevallen leidt de belasting tot een lager netto loon voor de werknemer en hogere loonkosten voor de werkgever. Ik laat zien dat deze conventionele wijsheid voortkomt uit de aanname van een flexibele loonvoet, wat niet te rijmen is met onvrijwillige werkloosheid. Met onvrijwillige werkloosheid leidt een hogere belasting op werknemers tot herverdelings- en efficiëntiewinsten (zoals hierboven besproken), maar een hogere belasting op werkgevers leidt juist tot een nóg lagere vraag naar arbeid en daarmee tot een efficiëntieverlies door hogere werkloosheid.

Bibliography

Acemoglu, Daron, and Jörn-Steffen Pischke. 1999. "The structure of wages and investment in general training." *Journal of Political Economy*, 107(3): 539–572.

Allen, Stephen P. 1987. "Taxes, redistribution, and the minimum wage: A theoretical analysis." *Quarterly Journal of Economics*, 102(3): 477–489.

Aronsson, Thomas, and Erkki Koskela. 2008. "Optimal redistributive taxation and provision of public input goods in an economy with outsourcing and unemployment." *CESifo Working Paper No. 2481.*

Arulampalam, Wiji, Alison L Booth, and Mark L Bryan. 2004. "Training and the new minimum wage." *The Economic Journal*, 114(494): C87–C94.

Atkinson, Anthony B, and Joseph E Stiglitz. 1976. "The design of tax structure: Direct versus indirect taxation." *Journal of Public Economics*, 6(1-2): 55–75.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics.* New York/Singapore:McGraw-Hill.

Autor, David H, Alan Manning, and Christopher L. Smith. 2010. "The contribution of the minimum wage to U.S. wage inequality over three decades: a reassessment." *NBER Working Paper 16533.*

Backhouse, Roger E. 2013. "Scientific welfare analysis: The origins of Bergson-Samuelson welfare economics, 1936-1947." Mimeo, University of Birmingham.

Bentham, Jeremy. 1907. *An Introduction to the Principles of Morals and Legislation.* Oxford:Clarendon Press.

Bhaskar, Venkataraman, and Ted To. 1999. "Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition." *The Economic Journal*, 109(455): 190–203.

Bloemen, Hans G. 2008. "Job search, hours restrictions, and desired hours of work." *Journal of Labor Economics*, 26(1): 137–179.

Blomquist, Sören, and Luca Micheletto. 2006. "Optimal redistributive taxation when government's and agents' preferences differ." *Journal of Public Economics*, 90(6): 1215–1233.

Blundell, Richard, and Andrew Shephard. 2012. "Employment, hours of work and the optimal taxation of low-income families." *Review of Economic Studies*, 79(2): 481–510.

Blundell, Richard, and Thomas MaCurdy. 1999. "Labor supply: A review of alternative approaches." In *Handbook of Labor Economics.* Vol. 3, , ed. Orley Ashenfelter and David Card, 1559–1695. Amsterdam:Elsevier.

Boadway, Robin, and Katherine Cuff. 2001. "A minimum wage can be welfare-improving and employment-enhancing." *European Economic Review*, 45(3): 553–576.

Böheim, René, and Mark P Taylor. 2004. "Actual and preferred working hours." *British Journal of Industrial Relations*, 42(1): 149–166.

Booth, Alison L. 1995. *The Economics Of The Trade Union.* Cambridge:Cambridge University Press.

Booth, Alison L, and Jan C Van Ours. 2008. "Job satisfaction and family happiness: The part-time work puzzle." *The Economic Journal*, 118(526): F77–F99.

Brewer, Mike, Emmanuel Saez, and Andrew Shephard. 2010. "Means-testing and tax rates on earnings." In *Dimensions of Tax Design: The Mirrlees Review.* , ed. James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles and James Poterba, 90–173. Oxford:Oxford University Press.

Cahuc, Pierre, and Guy Laroque. 2013. "Optimal taxation and monopsonistic labor market: does monopsony justify the minimum wage?" *Journal of Public Economic Theory.*

Cahuc, Pierre, and Philippe Michel. 1996. "Minimum wage unemployment and growth." *European Economic Review*, 40(7): 1463–1482.

Cahuc, Pierre, Anne Saint-Martin, and André Zylberberg. 2001. "The consequences of the minimum wage when other wages are bargained over." *European Economic Review*, 45(2): 337–352.

Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. "Neuroeconomics: How neuroscience can inform economics." *Journal of Economic Literature*, 43(1): 9–64.

Card, David. 1992. "Using regional variation in wages to measure the effects of the federal minimum wage." *Industrial and Labor Relations Review*, 46(1): 22–37.

Card, David. 1999. "The causal effect of education on earnings." In *Handbook of Labor Economics*. Vol. 3, , ed. Orley Ashenfelter and David Card, 1801–1863. Amsterdam:Elsevier.

Card, David, and Alan B Krueger. 1995. *Myth and measurement: The new economics of the minimum wage.* Princeton:Princeton University Press.

Card, David, and Thomas Lemieux. 2001. "Dropout and enrollment trends in the postwar period: What went wrong in the 1970s?" In *Risky behavior among youths: An economic analysis.* , ed. Jonathan Gruber, 439–482. University of Chicago Press.

Chaplin, Duncan D, Mark D Turner, and Andreas D Pape. 2003. "Minimum wages and school enrollment of teenagers: A look at the 1990s." *Economics of Education Review*, 22(1): 11–21.

Clark, Andrew E., Paul Frijters, and Michael A. Shields. 2008. "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles." *Journal of Economic Literature*, 46(1): 95–144.

Clark, Damon. 2011. "Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrolment in Post-compulsory Education in England." *Economica*, 78(311): 523–545.

Couch, Kenneth A, and David C Wittenburg. 2001. "The response of hours of work to increases in the minimum wage." *Southern Economic Journal*, 68(1): 171–177.

Crofton, Stephanie O, William L Anderson, and Emily C Rawe. 2009. "Do higher real minimum wages lead to more high school dropouts? Evidence from Maryland across races, 1993–2004." *American Journal of Economics and Sociology*, 68(2): 445–464.

Cunningham, James. 1981. "The impact of minimum wages on youth employment, hours of work, and school attendance: Cross-sectional evidence from the 1960 and 1970 censuses." In *The Economics of Legal Minimum Wages.* , ed. Simon Rottenberg, 88–123. Washington, D.C.:American Enterprise Institute for Public Policy Research.

Danziger, Eliav, and Leif Danziger. 2010. "A graduated minimum wage with optimal taxation." Mimeo, Princeton University and Ben Gurion University.

Davis, Lucas W, and Lutz Kilian. 2011. "The allocative cost of price ceilings in the U.S. residential market for natural gas." *Journal of Political Economy*, 119(2): 212–241.

Deacon, Robert T, and Jon Sonstelie. 1989. "The welfare costs of rationing by waiting." *Economic Inquiry*, 27(2): 179–196.

Deltas, George. 2007. "Can a minimum wage increase employment and reduce prices in a neoclassical perfect information economy?" *Journal of Mathematical Economics*, 43(6): 657–674.

Diamond, Peter A. 1975. "A many-person Ramsey tax rule." *Journal of Public Economics*, 4(4): 335–342.

Dickens, William T, and Shelly J Lundberg. 1993. "Hours Restrictions and Labor Supply." *International Economic Review*, 34(1): 169–92.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux. 1996. "Labor market institutions and the distribution of wages, 1973–1992: a semparametric approach." *Econometrica*, 64(5): 1001–1044.

Dolado, Juan J, Florentino Felgueroso, and Juan F Jimeno. 2000. "The role of the minimum wage in the welfare state: An appraisal." *IZA Discussion Paper No. 152.*

Dustmann, Christian, and Uta Schönberg. 2009. "Training and union wages." *The Review of Economics and Statistics*, 91(2): 363–376.

Dworkin, Ronald. 2000. *Sovereign Virtue.* Cambridge:Harvard University Press.

Ehrenberg, Ronald G, and Alan J Marcus. 1982. "Minimum wages and teenagers' enrollment-employment outcomes: a multinomial logit model." *Journal of Human Resources*, 17(1): 39–58.

Fairris, David, and Roberto Pedace. 2004. "The impact of minimum wages on job training: an empirical exploration with establishment data." *Southern Economic Journal*, 70(3): 566–583.

Feldstein, Martin S. 1972. "Distributional equity and the optimal structure of public prices." *The American Economic Review*, 62(1/2): 32–36.

Flinn, Christopher J. 2006. "Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates." *Econometrica*, 74(4): 1013–1062.

Frech, Harry E, and William C Lee. 1987. "The welfare cost of rationing-by-queuing across markets: Theory and estimates from the U.S. gasoline crises." *The Quarterly Journal of Economics*, 102(1): 97–108.

Freeman, Richard B. 1996. "The minimum wage as a redistributive tool." *The Economic Journal*, 106(436): 639–649.

Gerritsen, Aart. 2013*a*. "Equity and efficiency in rationed labor markets." Mimeo, Erasmus University Rotterdam.

Gerritsen, Aart. 2013*b*. "Optimal taxation when people do not maximize well-being." Mimeo, Erasmus University Rotterdam.

Gerritsen, Aart, and Bas Jacobs. 2013*a*. "Is a minimum wage an appropriate instrument for redistribution?" Mimeo, Erasmus University Rotterdam.

Gerritsen, Aart, and Bas Jacobs. 2013*b*. "Minimum wages and taxation in competitive labor markets with endogenous skill formation." Mimeo, Erasmus University Rotterdam.

Glaeser, Edward L., and Erzo F.P. Luttmer. 2003. "The misallocation of housing under rent control." *The American Economic Review*, 93(4): 1027–1046.

Grossberg, Adam J, and Paul Sicilian. 1999. "Minimum wages, on-the-job training, and wage growth." *Southern Economic Journal*, 65(3): 539–556.

Guesnerie, Roger, and Kevin Roberts. 1987. "Minimum wage legislation as a second best policy." *European Economic Review*, 31(1-2): 490–498.

Gul, Faruk, and Wolfgang Pesendorfer. 2008. "The case for mindless economics." In *The Foundations of Positive and Normative Economics: A Handbook.* , ed. Andrew Caplin and Andrew Shotter, 3–42. New York:Oxford University Press.

Harsanyi, John C. 1982. "Morality and the theory of rational behavior." In *Utilitarianism and Beyond.* , ed. Amartya Sen and Bernard Williams, 39–62. Cambridge:Cambridge University Press.

Hashimoto, Masanori. 1982. "Minimum wage effects on training on the job." *American Economic Review*, 72(5): 1070–1087.

Hausman, Daniel M. 2011. *Preference, Value, Choice, and Welfare.* New York:Cambridge University Press.

Hosios, Arthur J. 1990. "On the efficiency of matching and related models of search and unemployment." *The Review of Economic Studies*, 57(2): 279–298.

Hungerbühler, Mathias, and Etienne Lehmann. 2009. "On the optimality of a minimum wage: New insights from optimal tax theory." *Journal of Public Economics*, 93(3): 464–481.

Immervoll, Herwig. 2013. "Minimum wages, minimum labour costs, and the tax treatment of low-wage employment." *IZA Discussion Paper No. 2555.*

Jacobs, Bas, and Ruud De Mooij. 2011. "Pigou meets Mirrlees: on the irrelevance of tax distortions for the second-best Pigouvian tax." *CESifo Working Paper 3342.*

Jacobs, Bas, Dirk Schindler, and Hongyan Yang. 2012. "Optimal Taxation of Risky Human Capital." *The Scandinavian Journal of Economics*, 114(3): 908–931.

Kahn, Shulamit, and Kevin Lang. 1991. "The effect of hours constraints on labor supply estimates." *The Review of Economics and Statistics*, 73(4): 605–611.

Kanbur, Ravi, Jukka Pirttilä, and Matti Tuomala. 2006. "Non-welfarist optimal taxation and behavioural public economics." *Journal of Economic Surveys*, 20(5): 849–868.

Kane, Thomas J. 1994. "College entry by blacks since 1970: The role of college costs, family background, and the returns to education." *Journal of Political Economy*, 102(5): 878–911.

Kaplow, Louis. 2008. *The Theory of Taxation and Public Economics.* Princeton University Press.

Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus Welfare.* Cambridge:Harvard University Press.

Katz, Lawrence F, and H David. 1999. "Changes in the wage structure and earnings inequality." In *Handbook of Labor Economics.* Vol. 3, , ed. Orley Ashenfelter and David Card, 1463–1555. Amsterdam:Elsevier.

Knabe, Andreas, and Steffen Rätzel. 2010. "Income, happiness, and the disutility of labour." *Economics Letters*, 107(1): 77–79.

Kőszegi, Botond, and Matthew Rabin. 2008. "Choices, situations, and happiness." *Journal of Public Economics*, 92(8-9): 1821–1832.

Landais, Camille, Pascal Michaillat, and Emmanuel Saez. 2013. "Optimal unemployment insurance over the business cycle." *National Bureau of Economic Research Working Paper No. 16526.*

Landon, Stuart. 1997. "High school enrollment, minimum wages and education spending." *Canadian Public Policy/Analyse de Politiques*, 23(2): 141–163.

Layard, Richard, Guy Mayraz, and Stephen Nickell. 2008. "The marginal utility of income." *Journal of Public Economics*, 92(8-9): 1846–1857.

Lee, David, and Emmanuel Saez. 2012. "Optimal minimum wage policy in competitive labor markets." *Journal of Public Economics*, 96(9–10): 739–749.

Lee, David S. 1999. "Wage inequality in the United States during the 1980s: rising dispersion or falling minimum wage?" *The Quarterly Journal of Economics*, 114(3): 977–1023.

Lehmann, Etienne, François Marical, and Laurence Rioux. 2013. "Labor income responds differently to income-tax and payroll-tax reforms." *Journal of Public Economics*, 99(4): 66–84.

Lokshin, Michael. 2006. "Semi-parametric difference-based estimation of partial linear regression models." *Stata Journal*, 6(3): 377–383.

Lott, John R. 1990. "Nontransferable rents and an unrecognized social cost of minimum wage laws." *Journal of Labor Research*, 11(4): 453–460.

Luttmer, Erzo F. P. 2007. "Does the Minimum Wage Cause Inefficient Rationing?" *BE Journal of Economic Analysis & Policy*, 7(1). (Contributions, Article 49).

Manning, Alan. 2003. *Monopsony in motion: Imperfect competition in labor markets.* Princeton:Princeton University Press.

Marceau, Nicolas, and Robin Boadway. 1994. "Minimum wage legislation and unemployment insurance as instruments for redistribution." *The Scandinavian Journal of Economics*, 96(1): 67–81.

Mattila, J Peter. 1981. "The impact of minimum wages on teenage schooling and on the part-time/full-time employment of youths." In *The Economics of Legal Minimum Wages.* , ed. Simon Rottenberg, 61–87. Washington, D.C.:American Enterprise Institute for Public Policy Research.

Mill, John Stuart. 1863. *Utilitarianism.* London:Parker, Son, and Bourn, West Strand.

Mincer, Jacob, and Linda S Leighton. 1981. "Effect of minimum wages on human capital formation." In *The Economics of Legal Minimum Wages.* , ed. Simon Rottenberg, 155–173. Washington, D.C.:American Enterprise Institute for Public Policy Research.

Mirrlees, James A. 1971. "An exploration in the theory of optimum income taxation." *Review of Economic Studies*, 38(2): 175–208.

Mirrlees, James A. 1976. "Optimal tax theory: A synthesis." *Journal of Public Economics*, 6(4): 327–358.

Montmarquette, Claude, Nathalie Viennot-Briot, and Marcel Dagenais. 2007. "Dropout, school performance, and working while in school." *The Review of Economics and Statistics*, 89(4): 752–760.

Neumark, David, and William Wascher. 1995*a*. "Minimum wage effects on employment and school enrollment." *Journal of Business & Economic Statistics*, 13(2): 199–206.

Neumark, David, and William Wascher. 1995*b*. "Minimum-wage effects on school and work transitions of teenagers." *The American Economic Review*, 85(2): 244–249.

Neumark, David, and William Wascher. 2001. "Minimum wages and training revisited." *Journal of Labor Economics*, 19(3): 563–595.

Neumark, David, and William Wascher. 2003. "Minimum wages and skill acquisition: Another look at schooling effects." *Economics of Education Review*, 22(1): 1–10.

Neumark, David, and William Wascher. 2006. "Minimum wages and employment: A review of evidence from the new minimum wage research." *National Bureau of Economic Research Working Paper No. 12663.*

Nickell, Stephen. 1979. "Education and lifetime patterns of unemployment." *Journal of Political Economy*, 87(5): S117–S131.

Nozick, Robert. 1974. *Anarchy, State, and Utopia.* New York:Basic Books.

OECD. 2011*a*. "Education at a Glance 2011: OECD Indicators." OECD Publishing, Paris.

OECD. 2011*b*. "Taxation and Employment." *OECD Tax Policy Studies No. 21*, OECD Publishing, Paris.

Pacheco, Gail A, and Amy A Cruickshank. 2007. "Minimum wage effects on educational enrollments in New Zealand." *Economics of Education Review*, 26(5): 574–587.

Palda, Filip. 2000. "Some deadweight losses from the minimum wage: The cases of full and partial compliance." *Labour Economics*, 7(6): 751–783.

Petrongolo, Barbara, and Marıa J. San Segundo. 2002. "Staying-on at school at 16: The impact of labor market conditions in Spain." *Economics of Education Review*, 21(4): 353–365.

Pouwels, Babette, Jacques Siegers, and Jan Dirk Vlasblom. 2008. "Income, working hours, and happiness." *Economics Letters*, 99(1): 72–74.

Rabin, Matthew. 1998. "Psychology and economics." *Journal of Economic Literature*, 36(1): 11–46.

Ragan, James F. 1977. "Minimum wages and the youth labor market." *The Review of Economics and Statistics*, 59(2): 129–136.

Rawls, John. 1971. *A Theory of Justice.* Cambridge:Harvard University Press.

Rebitzer, James B, and Lowell J Taylor. 1995. "The consequences of minimum wage laws some new theoretical ideas." *Journal of Public Economics*, 56(2): 245–255.

Robinson, Joan. 1933. *The economics of imperfect competition.* London:MacMillan.

Rosen, Sherwin. 1972. "Learning and experience in the labor market." *Journal of Human Resources*, 7(3): 326–342.

Saez, Emmanuel. 2001. "Using elasticities to derive optimal income tax rates." *Review of Economic Studies*, 68(1): 205–229.

Saint-Paul, Gilles. 1996. "Unemployment and increasing private returns to human capital." *Journal of Public Economics*, 61(1): 1–20.

Samuelson, Paul A. 1938. "A note on the pure theory of consumer's bahaviour." *Economica*, 5(17): 61–71.

Samuelson, Paul A. 1947. *Foundations of Economic Analysis.* Cambridge, MA:Harvard University Press.

Sandmo, Agnar. 1975. "Optimal taxation in the presence of externalities." *The Swedish Journal of Economics*, 77(1): 86–98.

Schiller, Bradley R. 1994. "Moving up: The training and wage gains of minimum-wage entrants." *Social Science Quarterly*, 75(3): 622–636.

Seade, Jesus. 1980. "Optimal non-linear policies for non-utilitarian motives." In *Income Distribution: The Limits to Redistribution.* , ed. David A. Collard, Richard Lecomber and Martin Slater. Bristol:Scientechnica.

Sen, Amartya. 1992. *Inequality Re-examined.* Cambridge:Harvard University Press.

Sidgwick, Henry. 1874. *The Methods of Ethics.* London:Macmillan.

Stewart, Mark B, and Joanna K Swaffield. 1997. "Constraints on the desired hours of work of British men." *The Economic Journal*, 107(441): 520–535.

Stigler, George J. 1946. "The economics of minimum wage legislation." *The American Economic Review*, 36(3): 358–365.

Stiglitz, Joseph E. 1982. "Self-selection and Pareto efficient taxation." *Journal of Public Economics*, 17(2): 213–240.

Teulings, Coen N. 2000. "Aggregation bias in elasticities of substitution and the minimum wage paradox." *International Economic Review*, 41(2): 359–398.

Teulings, Coen N. 2003. "The contribution of minimum wages to increasing wage inequality." *The Economic Journal*, 113(490): 801–833.

Tobin, James. 1952. "A Survey of the Theory of Rationing." *Econometrica*, 20(4): 521–553.

Tobin, James. 1977. "How dead is Keynes?" *Economic Inquiry*, 15(4): 459–468.

Tumino, Alberto, and Mark Taylor. 2013. "The impact of local labour market conditions on school leaving decisions." mimeo.

Tuomala, Matti. 1990. *Optimal Income Tax and Redistribution.* Oxford:Clarendon Press.

Zavodny, Madeline. 2000. "The effect of the minimum wage on employment and hours." *Labour Economics*, 7(6): 729–750.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

527. J. LINDE, *Experimenting with New Combinations of Old Ideas*
528. D. MASSARO, *Bounded Rationality and Heterogeneous Expectations in Macroeconomics*
529. J. GILLET, *Groups in Economics*
530. R. LEGERSTEE, *Evaluating Econometric Models and Expert Intuition*
531. M.R.C. BERSEM, *Essays on the Political Economy of Finance*
532. T. WILLEMS, *Essays on Optimal Experimentation*
533. Z. GAO, *Essays on Empirical Likelihood in Economics*
534. J. SWART, *Natural Resources and the Environment: Implications for Economic Development and International Relations*
535. A. KOTHIYAL, *Subjective Probability and Ambiguity*
536. B. VOOGT, *Essays on Consumer Search and Dynamic Committees*
537. T. DE HAAN, *Strategic Communication: Theory and Experiment*
538. T. BUSER, *Essays in Behavioural Economics*
539. J.A. ROSERO MONCAYO, *On the Importance of Families and Public Policies for Child Development Outcomes*
540. E. ERDOGAN CIFTCI, *Health Perceptions and Labor Force Participation of Older Workers*
541. T. WANG, *Essays on Empirical Market Microstructure*
542. T. BAO, *Experiments on Heterogeneous Expectations and Switching Behavior*
543. S.D. LANSDORP, *On Risks and Opportunities in Financial Markets*
544. N. MOES, *Cooperative Decision Making in River Water Allocation Problems*
545. P. STAKENAS, *Fractional Integration and Cointegration in Financial Time Series*
546. M. SCHARTH, *Essays on Monte Carlo Methods for State Space Models*
547. J. ZENHORST, *Macroeconomic Perspectives on the Equity Premium Puzzle*
548. B. PELLOUX, *On the Role of Emotions and Social Ties in Public Good Games: Behavioral and Neuroeconomic Studies*
549. N. YANG, *Markov-Perfect Industry Dynamics: Theory, Computation, and Applications*
550. R.R. VAN VELDHUIZEN, *Essays in Experimental Economics*
551. X. ZHANG, *Modeling Time Variation in Systemic Risk*

552. H.R.A. KOSTER, *The Internal Structure of Cities: The Economics of Agglomeration, Amenities and Accessibility*
553. S.P.T. GROOT, *Agglomeration, Globalization and Regional Labor Markets: Micro Evidence for the Netherlands*
554. J.L. MÖHLMANN, *Globalization and Productivity, Micro-Evidence on Heterogeneous Firms, Workers and Products*
555. S.M. HOOGENDOORN, *Diversity and Team Performance: A Series of Field Experiments*
556. C.L. BEHRENS, *Product Differentiation in Aviation Passenger Markets: The Impact of Demand Heterogeneity on Competition*
557. G. SMRKOLJ, *Dynamic Models of Research and Development*
558. S. PEER, *The Economics of Trip Scheduling, Travel Time Variability and Traffic Information*
559. V. SPINU, *Nonadditive Beliefs: From Measurement to Extensions*
560. S.P. KASTORYANO, *Essays in Applied Dynamic Microeconometrics*
561. M. VAN DUIJN, *Location, Choice, Cultural Heritage and House Prices*
562. T. SALIMANS, *Essays in Likelihood-Based Computational Econometrics*
563. P. SUN, *Tail Risk of Equity Returns*
564. C.G.J. KARSTEN, *The Law and Finance of M&A Contracts*
565. C. OZGEN, *Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth*
566. R.S. SCHOLTE, *The Interplay between Early-Life Conditions, Major Events and Health later in Life*
567. B.N. KRAMER, *Why Don't They Take a Card? Essays on the Demand for Micro Health Insurance*
568. M. KILIÇ, *Fundamental Insights in Power Futures Prices*
569. A.G.B. DE VRIES, *Venture Capital: Relations with the Economy and Intellectual Property*
570. E.M.F. VAN DEN BROEK, *Keeping up Appearances*
571. K.T. MOORE, *A Tale of Risk: Essays on Financial Extremes*
572. F.T. ZOUTMAN, *A Symphony of Redistributive Instruments*
573. M.J. GERRITSE, *Policy Competition and the Spatial Economy*
574. A. OPSCHOOR, *Understanding Financial Market Volatility*
575. R.R. VAN LOON, *Tourism and the Economic Valuation of Cultural Heritage*
576. I.L. LYUBIMOV, *Essays on Political Economy and Economic Development*