ESSAYS ON CATEGORICAL AND UNIVERSAL WELFARE PROVISION : DESIGN, OPTIMAL TAXATION AND ENFORCEMENT ISSUES

Sean Edward Slack

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2016

Essays on Categorical and Universal Welfare Provision:
Design, Optimal Taxation and Enforcement Issues.

Sean Edward Slack

University of St Andrews

This thesis is submitted in partial fulfilment for the degree of
PhD
at the
University of St Andrews

February 2016

Declaration

1. Candidate's declarations:

I, Sean Edward Slack, hereby certify that this thesis, which is approximately 61,010 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2012 and as a candidate for the degree of PhD in September 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2012 and 2015.

Date 26.02.16...... signature of candidate ....

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 26/2/16... signature of supervisor ....

3. Permission for publication:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

(b) Embargo on all of print copy for a period of ..2... years (maximum five) on the following ground(s):

- Publication would preclude future publication
- Publication would be in breach of laws or ethics

Supporting statement for printed embargo request: I request an embargo for reason b. First, I wish to avoid the risk of future papers based on the work in this thesis being rejected for journal publication on the grounds that the material is already published. Second, I have *already published* a paper based on some of the work in this thesis. To comply with open access requirements the journal allows an online version, but requires that it be embargoed for two years.

ELECTRONIC COPY

(b) Embargo on all of electronic copy for a period of ..2... years (maximum five) on the following ground(s):

- Publication would preclude future publication
- Publication would be in breach of laws or ethics

Supporting statement for electronic embargo request: I request an embargo for reason b. First, I wish to avoid the risk of future papers based on the work in this thesis being rejected for journal publication on the grounds that the material is already published. Second, I have *already published* a paper based on some of the work in this thesis. To comply with open access requirements the journal allows an online version, but requires that it be embargoed for two years.

Date ..26.02.16..... signature of candidate ....

Date ..26/2/16..... signature of supervisor ....

Abstract

Part I comprises three chapters (2-4) that analyse the optimal combination of a universal benefit ($B \geq 0$) and categorical benefit ($C \geq 0$) for an economy where individuals differ in both their ability to work and, if able to work, their productivity. C is *ex-ante* conditioned on applicants being unable to work, and *ex-post* conditioned on recipients not working.

In Chapter 2 the benefit budget is fixed but the test awarding C makes Type I and Type II errors. Type I errors guarantee $B > 0$ at the optimum to ensure all unable individuals have positive consumption. The analysis with Type II errors depends on the enforcement of the *ex-post* condition. Under *No Enforcement* $C > 0$ at the optimum conditional on the awards test having some discriminatory power; whilst maximum welfare falls with both error propensities. Under *Full Enforcement* $C > 0$ at the optimum always; and whilst maximum welfare falls with the Type I error propensity it may increase with the Type II error propensity.

Chapters 3 and 4 generalise the analysis to a linear-income tax framework. In Chapter 3 categorical status is perfectly observable. Optimal linear and piecewise-linear tax expressions are written more generally to capture cases where it is suboptimal to finance categorical transfers to eliminate inequality in the average social marginal value of income. Chapter 4 then derives the optimal linear income tax for the case with classification errors and *Full Enforcement*. Both equity and efficiency considerations capture the incentives an increase in the tax rate generates for able individuals to apply for C.

Part II (Chapter 5) focuses on the decisions of individuals to work when receiving C, given a risk of being detected and fined *proportional* to C. Under CARA preferences the risk premium associated with the variance in benefit income is *convex-increasing* in C, thus giving C a role in enforcement.

Acknowledgements

I would like to pay special thanks to David Ulph for his support throughout the Ph.D. programme and, indeed, throughout much of my academic career thus far. David is an excellent supervisor and to be able to learn from him is a privilege for which I will always be grateful. Thank you David.

I also warmly acknowledge the AXA Research Fund for financing this Doctoral thesis research. Further, AXA has also provided superb training on research dissemination through their 'Pop Days' event in Paris. I am honoured to have been awarded their Doctoral Research Fellowship.

Contents

Introduction

The extent to which cash benefits should be targeted at those in need, as opposed to being made more universally available, is a subject of much debate across welfare states. Targeted transfers play a prominent role and can take a number of forms, where eligibility may be conditioned on (i) a means test; (ii) belonging to some categorical group such as the disabled or involuntarily unemployed; or (iii) some combination of the two. A well-established result in economic theory is that 'tagging' individuals who belong to some categorical group that the policymaker wishes to assist is welfare improving because it provides the neediest individuals in society with greater support than they would receive under a simple negative income tax system and does so without resorting to high marginal tax rates (Akerlof, 1978).

There are, however, a number of caveats associated with targeting that the model generating this result does not account for.[1] First, categorical transfers can be both complex and costly to administer. Take the case of disability benefits: What is the appropriate definition of disability? How much weight should be placed on medically verifiable criteria as opposed to an applicant's alleged discomfort at work? Should the benefit be means tested? These complex issues lead us to a second point: wherever the eligibility 'line' is drawn, classification errors of both Type I (false rejection) and Type II (false award) are likely to be made. Broadly speaking, these errors may be made because (i) the 'line' is incorrect but the tests correctly classify individuals around the line; (ii) the 'line' is correct but the tests are imperfect and misclassify individuals around the line; or (iii) some combination of both.[2] However classification errors arise,

[1] Akerlof (1978) does discuss a number of these issues in his summary and conclusions.

[2] In line with the second case, consider an awards technology that has fixed propensities to misclassify individuals around an eligibility threshold that accurately distinguishes between needy and non-needy individuals. It follows that any adjustment in the eligibility threshold will involve a tradeoff between Type I and Type II errors. The relevant authority may therefore adjust the threshold to '*err on the side of harshness...or on the side of generosity*' (see Goodin, 1985, p.141).

they diminish the effectiveness of targeted programmes in providing support to those in need, in addition to generating incentives for non-needy individuals to masquerade as needy. Third, because targeted programmes focus on specific subgroups of society they may impose stigma on their intended recipients: this may arise due to accusatory eligibility tests or subsequent conditions placed on recipients (Moffitt, 1983). Fourth and finally, the complex and time-consuming nature of application forms present costs to those who apply, and these costs may greatest for those most in need (Currie, 2004). Both stigma and application costs give rise to non-take-up in targeted welfare programmes, as the very individuals that the programmes are designed to assist are deterred from applying.

At the other end of the spectrum, a universal benefit[3] provides all individuals in society with the same *unconditional*[4] cash benefit. Such schemes are often advocated on the grounds that they (i) provide all individuals in society with a guaranteed source of income that is independent of their employment status and financial circumstances; (ii) enhance work incentives because the benefit is not 'withdrawn' or 'phased-out' through working; (iii) are administratively simple; and (iv) are unlikely to suffer from the take-up related issues of targeted programmes (Van Parijs, 2004). The first point may be particularly important for individuals who are either ineligible for targeted welfare benefits or are incorrectly denied them by Type I classification error. If the universal benefit is financed by a linear (flat) income tax - as analysed by Atkinson (1995) - then both the second and third points may become stronger as flat taxes are argued to enhance work incentives whilst also lowering the administrative burden relative to more complex piecewise schedules (Paulus and Peichl, 2009; Peichl, 2014). In addition to these four points, Goodin (1992) advocates universal benefits on the grounds that they are 'minimally presumptuous'. Contrastingly, means tests often presume that families share earnings equitably among their members; whilst categorical benefit programmes may presume that certain conditions warrant support because they prohibit work, whilst other conditions do not. Whenever presumptions are made they

[3] A number of alternative names for the universal benefit are given in the literature, such as 'basic income', 'demogrant' or 'citizens income' (see Van Parijs, 2004).

[4] There is of course some eligibility condition: an individual must have citizenship to receive the benefit. In this sense a universal benefit is not entirely unconditional. Atkinson (2015, p.219) discusses how the citizenship condition may be too broad, in the sense that it can include individuals who live abroad and do not pay taxes. In part for this reason, Atkinson proposes replacing the citizenship condition with a participation condition; where participation may take the form of working, education, or some other type of social contribution.

may be either incorrect or become incorrect over time as society and the nature of employment changes.

A key issue surrounding universal benefit proposals is their affordability and, in turn, the level of taxation that would be required to finance payments to all individuals. This does, of course, depend on the nature of exactly what is being proposed. As discussed by both Atkinson (1995) and Van Parijs (2004), proposals typically take the form of a 'partial' universal system whereby a universal benefit replaces some but not all existing cash benefits.[5] There is reason to think that such proposals will garner more political support than the polar extremes of purely targeted or purely universal welfare provision.[6] Indeed, Skocpol (1991) refers to partial schemes as 'targeting within universalism' and, in examining the social policy history of the U.S., points to the relative support enjoyed by programmes that spread benefits over many individuals, whilst also leaving room for additional support to those most in need.

With these points in mind, this thesis is structured as follows. Part I contains three chapters (2,3 and 4) which model the optimal provision of a universal benefit (B) and categorical benefit (C) for an economy in which individuals differ in both their ability to work and, conditional on being able to work, their productivity (gross wage) when at work. The categorical benefit is targeted at unable individuals and is conditioned in two dimensions: *ex-ante* an applicant must be unable to work to be awarded the benefit; whilst *ex-post* a recipient must not work. Applications for the categorical benefit are taken to be costless in terms of money, stigma and time. If the benefit is only received by unable individuals then both dimensions of conditionality are automatically satisfied. However, the benefit may be administered with both Type I and Type II classification errors. A Type I error arises when an unable individual is incorrectly denied the categorical benefit, whilst a Type II error arises when an able individual is incorrectly awarded the categorical benefit. In the case of the latter error type, the *ex-post* no-work condition becomes relevant because individuals who are able to work are receiving the categorical benefit. Whether or not they will actually choose to work while receiving

[5] The state of Alaska provides the primary example of an existing partial universal benefit programme. In place since 1982, all residents of Alaska (including children) receive the same yearly cash benefit, the size of which is a function of the five year average interest on the Alaska Permanent Fund. This fund was set up in 1976 using oil revenue from the Trans-Alaska Pipeline System.

[6] To get a sense of the range of levels of taxation that may be considered, it is useful to note that Atkinson (1995, pp.114-129) simulates the distributional effect of two alternative partial schemes. The first scheme has a flat tax of 25 percent and a universal benefit of £10 per week; whilst the second scheme has a flat tax of 40 percent and a universal benefit of £35.60 per week.

it will depend on the enforcement of this condition.

A number of papers have analysed imperfectly targeted transfers in frameworks similar to that described above (see Parsons, 1996; Salanié, 2002). These papers make polar assumptions with regard to the enforcement of the *ex-post* 'no-work' condition, but the full implications of these differing assumptions are not directly comparable due to other differences in the modelling frameworks. First, Parsons (1996) adopts quasilinear preferences over consumption and leisure and assumes a homogenous able subpopulation where each individual has a marginal product of unity. In this setting able individuals who are tagged as unable by Type II error are allowed to work. From an analytical perspective this corresponds to the assumption that the *ex-post* 'no-work' condition is *not enforced*. Indeed, a 'dual negative income tax' system in which tagged able individuals are *incentivised* to work for less than their marginal product is found to be optimal. The greater the accuracy of the tagging technology the closer the optimum comes to the full insurance outcome.[7] Contrastingly, Salanié (2002) employs a more general framework with standard preferences and an able subpopulation where individuals differ over a productivity continuum. The *ex-post* condition that individuals do not work is *fully enforced* such that no tagged able individual will work. Given this, the surprising feature of the analysis is that there are no application decisions to be tagged. Instead, a fraction of the able subpopulation (corresponding to the Type II error propensity) are tagged and do not work. This must, however, be incentive incompatible for higher productivity able individuals who would choose not to be tagged because they are better off working and receiving the lower untagged/unconditional benefit.

Chapter 2 contributes to this literature by providing a systematic comparison of optimal welfare provision for the binary cases where the *ex-post* no-work condition is either (i) not enforced; or (ii) fully enforced. To pin down the intuition for how the optimal benefits change with classification errors, it is assumed that the government has a fixed budget for benefit expenditure (later chapters illustrate that many of the results generalise to an optimal tax analysis). In the baseline case where C is administered without error the standard result arises: if the benefit budget exceeds a critical level at which inequality in the average social marginal value of income (smvi)[8] between

[7]The full insurance outcome corresponds to the case where consumption is independent of ability status.

[8]Notice that we abstract from the term 'net' because the response of individual earnings to unearned

the unable and able subpopulations is eliminated through categorical transfers, it will be optimal to set both $C > 0$ and $B > 0$. This corresponds to a partial universal system because both targeted and universal benefits are provided. Contrastingly, if the budget falls at or below this critical level it will be optimal to set $C > 0$ but $B = 0$, thus corresponding to a purely targeted system. With perfect targeting, there is thus a clear ordering of priorities: spending should be purely categorical so long as the smvi of the unable exceeds the average smvi of the able (see also Beath et al., 1988; Diamond and Sheshinski, 1995; Parsons, 1996; Viard, 2001a,b).

In the main analysis C is administered with both Type I and Type II classification errors. Under No Enforcement of the *ex-post* condition benefit policy is one-dimensional in the sense that individuals receive a monetary benefit with no subsequent restriction on labour supply. Consequently, all able individuals will apply for C and, if awarded it, those of sufficiently high productivity will work when receiving it. Contrastingly, under Full Enforcement of the *ex-post* condition benefit policy is two-dimensional in the sense that individuals receive both a monetary benefit and a fully enforced zero quantity constraint on labour supply. In this case the only able individuals who will choose to apply for C will be those whose individual welfare is highest when receiving C and not working, and thus those of sufficiently low productivity (and therefore low opportunity cost). Application decisions in this latter setting will consequently be endogenous to the benefit levels.

The key results from Chapter 2 are as follows. Whenever Type I errors are made it will be optimal to set $B > 0$ so as to ensure that unable individuals who are incorrectly rejected C have some source of income to consume. Under No Enforcement it is optimal to set $C > 0$ only if the awards test has some discriminatory power.[9] In this case the optimal benefits are chosen so as to equate the average smvi of 'categorical recipients' with the average smvi of 'non-categorical recipients'[10], respectively, where classification errors mean that both groups may be composed of unable and able individuals. If however, the test has no discriminatory power, it is optimal to set $C = 0$ and spend the entire *per capita* budget on the universal benefit. It is shown that maximum welfare

income has no effect on the available budget for benefit expenditure. This assumption will be relaxed in the subsequent Chapters 3 and 4, where the benefit budget is determined endogenously by tax revenue.

[9] The awards test will have discriminatory power so long as it is more likely to award the categorical benefit to an unable applicant than to an able applicant.

[10] I.e. individuals who do not receive the categorical benefit.

falls with both Type I and Type II error propensities.

Contrastingly, under Full Enforcement it is optimal to set $C > 0$ at *all* levels of discriminatory power, therefore including the case of no discriminatory power. The Full Enforcement optimum is characterised by the condition that (i) the aggregate smvi of non-categorical recipients be equal to (ii) the aggregate smvi of categorical recipients *multiplied* by the increase in their total benefit income per unit reduction in the universal benefit.[11] In general, one cannot guarantee a unique solution to this condition because application decisions are endogenous to the benefit levels. The consequence of this endogeneity is that the two components of either aggregate smvi (i.e. individual smvi $\times$ number of individuals) may move in opposite directions. Consider the aggregate smvi of non-categorical recipients: an increase in the universal benefit lowers an individual's smvi *but* also reduces the number of individuals who apply for the categorical benefit, in turn increasing the number of non-recipients. Alternatively, consider the aggregate smvi of categorical recipients: an increase in the universal benefit may lower their total benefit income and thus increase their individual smvi, *but* also reduces the number of applicants and thus the number of categorical recipients. The overall effect on either aggregate smvi is ambiguous: it will depend on unspecified properties of both the utility function (e.g. third derivatives) and the distribution function (e.g. derivatives of the pdf).

Turning to the welfare effects of classification errors under Full Enforcement, it is shown that whilst welfare unambiguously falls with the Type I error propensity, there are conditions where it can increase with the Type II error propensity. The intuition here is that 'leakage' of the categorical benefit to lower productivity able individuals may play a redistributive role within the able subpopulation. This will depend in part on the proportion of able applicants who would be voluntarily unemployed when just receiving the universal benefit, and thus on the number of able applicants whose smvi is the same as that of unable applicants.

Finally, under both enforcement regimes between-group inequality in the average smvi will persist at the optimum whenever classification errors are made. Extensive numerical simulations with the CES preferences adopted by Stern (1976) illustrate how the optimal benefit levels change with the propensity to make classification errors.

[11] An implication of this is that at the Full Enforcement optimum the average smvi of categorical recipients will exceed that of non-categorical recipients due to the incentives that an increase in C generates for able individuals to apply.

Examples where welfare may increase with the Type II error propensity are also provided.

Chapters 3 and 4 proceed to generalise the analysis from Chapter 2 to an optimal linear income tax framework. Chapter 3 analyses the case of perfect discrimination, whilst Chapter 4 analyses the case with classification errors. In this richer setting, the benefit budget is determined endogenously by tax revenue (net of any revenue requirement for spending outside of welfare). In addition to choosing the optimum benefit levels, the optimum tax rate must therefore be chosen as well.

To the best of this author's knowledge, the analysis of categorical transfers within the standard optimal linear income tax framework[12] has thus far been restricted to the case of perfect discrimination. In this case optimal tax formulae are typically reported under the assumption that inequality in the average net smvi across categorical groups is eliminated through categorical transfers at the optimum (Viard, 2001a,b). This assumption allows the optimal tax expression to be written as in the uni-dimensional model where individuals differ only in productivity: the numerator (equity considerations) is the negative of the covariance between earnings and the net smvi; whilst the denominator (efficiency considerations) captures the response of (compensated) earnings to a change in the net-of-tax rate (see Atkinson and Stiglitz, 1980; Atkinson, 1995).

It is not immediately clear, however, that this will always be a valid assumption. There may be cases where it is suboptimal to finance categorical transfers to the point that inequality in the average net smvi across categorical groups is eliminated. For example, if a sufficiently large fraction of the population are dependent on categorical transfers for consumption then the level of taxation required to equate the average net smvi of the unable and able subpopulations may be too harmful to the latter group. This will also depend on the size of any revenue requirement in place for spending outside of welfare. Further, this argument is likely to hold beyond a simple linear income tax framework. For example, progressive piecewise linear tax systems[13] provide the government with additional tools to redistribute within categorical groups; but if shifting some of the tax burden away from lower earners in an able group (i) pushes the average net smvi of that group further below that of the unable group; and/or (ii) lowers tax revenue

[12] I.e. where in the productivity dimension there are a continuum of types.

[13] The term progressive is here used to refer to the case of increasing marginal tax rates.

relative to the flat tax case, this may limit further the cases where it is optimal to eliminate between-group inequality in the average net smvi.

Chapter 3[14] therefore derives more general expressions for (i) the optimal linear income tax rate; and (ii) the optimal two-bracket progressive piecewise linear tax rates, that allow for the persistence of between-group inequality at the optimum. The optimal linear income tax rate is derived under general preferences over consumption and leisure. The equity considerations in the numerator are now composed of both 'between-group' and 'within-group' terms. The first term is not found in standard tax expressions, whilst the latter term is now the negative of the covariance between *relative* earnings and the net smvi; where relative earnings refer to individual gross earnings as a fraction of aggregate gross earnings.[15] Numerical examples where between-group inequality in the average net smvi persists at the flat tax optimum are provided using a variant of framework employed by Stern (1976) (i.e. CES preferences, lognormal productivity distribution).

Turning to the piecewise analysis, individual utility is taken to be a concave transformation of preferences that are quasilinear in consumption. There is thus no income effect (see also Apps et al., 2014). Implicit expressions for the lower and upper tax rates are derived, in addition to that for the earnings threshold above which individuals face the upper marginal tax rate. Interestingly, the possibility that between-group inequality remains at the optimum enters into the equity considerations of the lower tax rate, but not the upper tax rate. The intuition here would seem to be that an increase in the lower tax rate is non-distortionary for those who choose to earn above the earnings threshold, and is thus an effective tool to help reduce between-group inequality. Numerical examples where between-group inequality persists at the piecewise optimum (and, for comparison, the flat tax optimum) are provided using preferences with a constant labour elasticity (see Atkinson, 1990; Saez, 2001).[16] Individual productivity is taken to be Pareto distributed as this is known to give rise to increasing marginal tax rates (Diamond, 1998). For cases where between-group inequality persists under both

[14]A shorter version of this chapter is published in the journal *Economics Letters*: See Slack (2015).

[15]If between-group inequality is eliminated at the optimum, the expression reduces to the well-documented optimal tax expression in the literature (see Atkinson and Stiglitz, 1980).

[16]Under these isoelastic preferences one can also establish analytically the cases where between-group inequality in the average smvi will persist at the flat tax optimum (for the special case where taxation is purely redistributive). Indeed, there is a critical level of the unable subpopulation size above which it will be suboptimal to eliminate this inequality.

flat tax and piecewise linear tax systems, the level of between-group inequality is typically higher under the piecewise system for any given unable subpopulation size and revenue requirement. Further, there are more cases where between-group inequality persists under the piecewise system than under the flat tax system.

The purpose of Chapter 4 is to tread new ground and derive an expression for the optimal linear tax rate when the categorical transfer is administered with both Type I and Type II classification errors. In this case it is assumed that the *ex-post* 'no-work' condition is fully enforced such that only able individuals of lower productivity will choose to apply.[17] Crucially, the application decisions of able individuals will be endogenous to the tax rate. Indeed, an increase in the tax rate will generate both direct and indirect behavioural responses that affect the government budget constraint. The direct effect enters only the tax revenue side of the budget constraint and is found in all conventional analyses: it is simply the response of gross earnings in the intensive margin to a marginal increase in the tax rate. The indirect effect, meanwhile, captures the fact that a marginal increase in the tax rate increases the critical productivity at or below which an able individual chooses to apply for the categorical benefit. Intuitively, this affects both the tax revenue and benefit expenditure sides of the budget constraint.

To simplify the analysis, individual utility is - as in the piecewise analysis of Chapter 3 - taken to be a concave transformation of preferences that are quasilinear in consumption. There are thus no income effects associated with a working individual's smvi and, further, the size of the universal benefit does not directly influence an able individual's decision to apply for the categorical benefit. This assumption is of great assistance in interpreting the optimal tax expression because a precise relationship between the average smvi of able applicants and the shadow price of public expenditure can be established.

In the resulting expression that characterises the optimal tax rate, an important component of both equity (numerator) and efficiency (denominator) considerations is the elasticity of the distribution function with respect to individual productivity, evaluated at the critical productivity at or below which able individuals choose to apply for the categorical benefit. Type II errors generate conflicting effects in both the equity and

[17] An analysis of the less realistic No Enforcement regime is documented in the appendix to this chapter.

efficiency considerations. In the equity dimension, Type II errors (i) mean that some able individuals of lower productivity - who the government would not wish to tax highly - receive the categorical benefit and do not work, which acts to raise the tax rate; but also (ii) redistribute 'within' the able subpopulation through 'leaking' categorical transfers to these lower productivity individuals, which acts to lower the tax rate because there may be less need to redistribute through the universal benefit. In the efficiency dimension, Type II errors (i) mean that some able individuals of lower productivity are awarded the categorical benefit and thus do not respond in the intensive margin to the tax rate, which acts to increase the tax rate; but (ii) result in both foregone tax revenue and additional benefit expenditure as able individuals depart the labour force because they are awarded the categorical benefit, both of which act to reduce the tax rate because the number of applicants for the categorical benefit are, *ceteris paribus*, an increasing function of the tax rate. The efficiency considerations are thus composed of both the direct and indirect effects discussed above.

Numerical simulations using the same isoelastic preferences employed in Chapter 3 suggest that an increase in *either* error propensity (i) increases the optimal marginal tax rate; (ii) increases the optimal universal benefit; but (iii) decreases the optimal categorical benefit. Whilst optimal welfare provision always includes some targeting, it becomes overall more universal as the propensity to make classification errors increases.

In Part I of this thesis the enforcement of the *ex-post* 'no-work' condition is binary: it is either not enforced or fully enforced. Both are strong assumptions. In particular, without explicitly modelling the enforcement parameters we cannot state all the conditions under which full enforcement can be achieved. The size of the categorical benefit will surely play a role in this enforcement. In reality, an individual who works (and pays taxes[18]) will certainly face some risk of detection, but it is unlikely that they will be detected with certainty. This would require perfect information sharing between the tax authority, benefit authority and local government. There is ample evidence that this is not the case (see Fuller et al., 2015).

With these enforcement issues in mind, Part II of this thesis abstracts from optimal welfare design and instead focuses on the decisions of individuals to violate *ex-post* conditionality, given a risk of being detected and sanctioned. In this regard, a small

[18]This thesis abstracts from 'cash-in-hand' work in the shadow economy.

existing literature analyses individual decisions to fraudulently claim unemployment benefits (Yaniv, 1986, 1997). Fraud is here taken to be the act of working whilst receiving benefits. Within this literature, relatively little focus is placed on the role that the benefit level plays in enforcement when penalties are made proportional to the benefit level itself. The most work in this area comes from Yaniv (1986), who analyses two alternative penalty structures from the tax evasion literature. Individual preferences are taken to exhibit decreasing absolute risk aversion. The author finds that an increase in the benefit level (i) increases incentives to fraudulently claim when the fine is proportional to the number of claiming days (i.e. independent of the benefit size); but (ii) has an *ambiguous* effect on incentives when the fine is proportional to the benefit size. In this latter case, an increase in the benefit level increases the expected fine. The model abstracts from both extensive and intensive labour supply decisions: in particular, there is no voluntary unemployment due to the disability benefit.

Chapter 5 aims to draw out more explicitly the role that the benefit level may play in exposing fraudulent recipients to risk. Once more, there is a continuum of productivity differentiated able individuals who may choose to apply for a categorical benefit that is *ex-ante* conditional on an applicant being unable to work; and *ex-post* conditional on a recipient not working. Recipients may also be required to spend a fraction of the 'working day' at the benefit office. There are no checks or penalties in place for an able individual who is incorrectly awarded the categorical benefit and does not work. This assumption is made for two reasons. First, such behaviour is extremely difficult to detect because the able recipient exactly mimics the unable recipient (Yaniv, 1986). Second, on the grounds of legal uncertainty applicants may be unaware of their true eligibility. However, the act of working is detectable and provides the benefit authority with the 'smoking gun' that a recipient is able to work and knowingly breaking the rules. Any recipient detected working is made to repay the categorial benefit and also fined at a rate proportional to the benefit size. This form of financial sanction mirrors that 'offered' by benefit authorities in reality (see, for example, Department for Work and Pensions, 2015).[19]

Drawing on the seminal work of Arrow (1970) and Pratt (1964), the analysis captures

[19]In the U.K., for example, a benefit recipient who is found to be fraudulent may be offered this type of financial sanction as an alternative to prosecution. In addition to reclaiming the benefits incorrectly received, the benefit authority imposes a fine corresponding to 50% of the overpaid benefits (up to a maximum of £2000) (Department for Work and Pensions, 2015).

the expected utility of a working recipient via the utility of their expected income net of the risk premium associated with the variance in benefit income. In general, the risk premium is approximated as the coefficient of absolute risk aversion multiplied by half the variance in benefit income. This variance is *convex-increasing* in the benefit level. If preferences exhibit constant absolute risk aversion the risk premium is also convex-increasing in the benefit level, and independent of individual productivity. An able individual who would choose to work conditional on receiving the categorical benefit will only apply for it if the expected benefit income exceeds the risk premium. Given that (i) the expected benefit is linearly increasing in the benefit size; but (ii) the risk premium is convex increasing in the benefit size, there is a critical benefit level above which full enforcement is attained. Further, the more lenient the standard enforcement parameters (detection probability, penalty rate), the higher this critical benefit level.

An unsatisfactory implication of this result is that for any benefit set below the critical level, *all* individuals who would choose to work conditional on receiving the benefit will apply. If, however, recipients are required to spend a fraction of the working day at the benefit office, those of higher productivities will not apply because the opportunity cost of foregone earnings is too large. In this analysis the categorical benefit can still be set to attain full enforcement, *but* even if it is not the range of wages for which fraud occurs will be bounded above by a critical wage at which the opportunity cost of foregone earnings offsets the expected benefit net of risk premium.

The remainder of this thesis is structured as follows. Chapter 1 precedes the main analyses in Parts I and II through providing an overview of the existing economics literature on tax/benefit programmes. Where possible, a common notation is adopted to facilitate comparison across the various works. Part I then contains Chapters 2, 3 and 4. These chapters focus on the optimal choice of categorical and universal benefits, given differing assumptions on both the benefit budget and tagging technology available to the government. In Chapter 2 the benefit budget is exogenously fixed, whereas in Chapters 3 and 4 it is endogenously determined by tax revenue. Part II contains Chapter 5, which analyses the decision of ineligible individuals to violate an *ex-post* 'no-work' condition, given a probability of being detected and fined proportional to the benefit size. Finally, Chapter 6 concludes the thesis.

Chapter 1

Literature Review

The design and behavioural implications of welfare programmes have received much attention in the economics literature, both theoretically and empirically.[1] This literature can be partitioned into three broad strands. First, a large theoretical literature analyses the optimal design of welfare benefits under the alternate assumptions that the benefit authority has (i) no formal awards technology to determine eligibility, but instead chooses consumption bundles that satisfy incentive compatibility constraints and induce self revelation; (ii) an awards technology that uses categorical information to 'tag' only eligible individuals, but may make Type I (false rejection) errors; or (iii) an awards technology that makes *both* Type I and Type II (false award) errors in tagging. The analyses that fall under the second of these informational assumptions are firmly grounded in the optimal income tax framework originating from Mirrlees (1971).

A small second strand of literature largely abstracts from design issues and instead focuses on the decisions of truly ineligible individuals to fraudulently claim welfare benefits. This literature draws from the economics of crime (Becker, 1968) and, relatedly, the tax-evasion literature. It does, however, recognise that there are *ex-post* conditions associated with receiving benefits; such as spending a fraction of the working day 'signing on' at the benefit office. These will influence individual behaviour.

Related to the second strand, a third empirical literature assesses both (i) the disincentives provided by; and (ii) the scope for classification errors, in the U.S. Social Security

[1] Design issues and behavioural incentives are of course intrinsically related.

Table 1.1: Some Common Notation

Notation	*Interpretation*
θ	Population proportion (typically unable or low ability)
x	Individual consumption
H	Labour supply
n	Individual productivity/ gross wage
n_l, n_h	Low productivity, High productivity (Unskilled, Skilled)
$f(n)$, $F(n)$	Productivity density function, productivity distribution function
y	Individual gross earnings
$\bar{y}$	Aggregate/Average gross earnings
T	Tax/Transfer
t	Linear income tax rate
M;B, C	Unearned income (general); Welfare benefit, Categorical benefit

Disability awards process[2]. In the latter case, a number of studies indicate there are non-negligible propensities to make both Type I and Type II classification errors.

This review discusses some of the key contributions in each of these three bodies of literature. Where possible, a common notation has been adopted to facilitate comparison across the various works (see Table 1.1).

1.1 The Design of Welfare Benefits

1.1.1 No Awards Technology: Mechanism Design

A body of literature focuses on the design of welfare benefits when the benefit authority has no formal discriminatory test to determine eligibility. Instead, it chooses consumption bundles/transfer levels which satisfy incentive compatibility constraints, and thus prevent the non-needy from masquerading as the needy.

Diamond and Mirrlees (1978) model the optimal provision of social insurance when a population of size 1 *ex-ante* identical individuals face an exogenous probability θ of becoming unable to work. The government cannot observe whether an individual is

[2]This literature analyses both the insurance and assistance components of disability welfare provision.

unemployed voluntarily or due to disability: it is therefore restricted to choosing consumption levels for workers and non-workers. An able individual who works has utility $v^e(x^e)$; where x^e denotes the consumption level a worker receives. Contrastingly, an able individual who does not work has utility $v^o(x^o)$; where x^o denotes the consumption level a non-worker receives. Finally, an unable individual also receives the consumption bundle x^o but faces utility $u(x^o)$. An able individual who chooses to work produces an output of 1 such that, were there no disability or voluntary unemployment, total resources in the economy would be 1.

The government problem is therefore described by[3]:

$$\begin{aligned} &\max_{x^e,x^o} W = (1-\theta)v^e(x^e) + \theta u(x^o) \\ \text{s.t.}\quad &\theta x^o + (1-\theta)x^e = (1-\theta) \quad \text{(Budget Constraint)} \\ &v^e(x^e) \geq v^o(x^o) \quad \text{(Incentive Compatibility constraint)} \end{aligned} \tag{1.1}$$

The assumption is made that an able individual will choose to work if $v^e(x^e) = v^o(x^o)$: This guarantees positive resources in the economy even if the incentive compatibility constraint is binding. Given an indifference curve between the two consumption states of $(1-\theta)v^e(x^e) + \theta u(x^o) = k$, where k is a constant, it is straightforward to establish that:

$$\frac{dx^o}{dx^e} = -\underbrace{\left(\frac{1-\theta}{\theta}\right)}_{\text{budget slope}} \frac{(v^e)'}{u'} \tag{1.2}$$

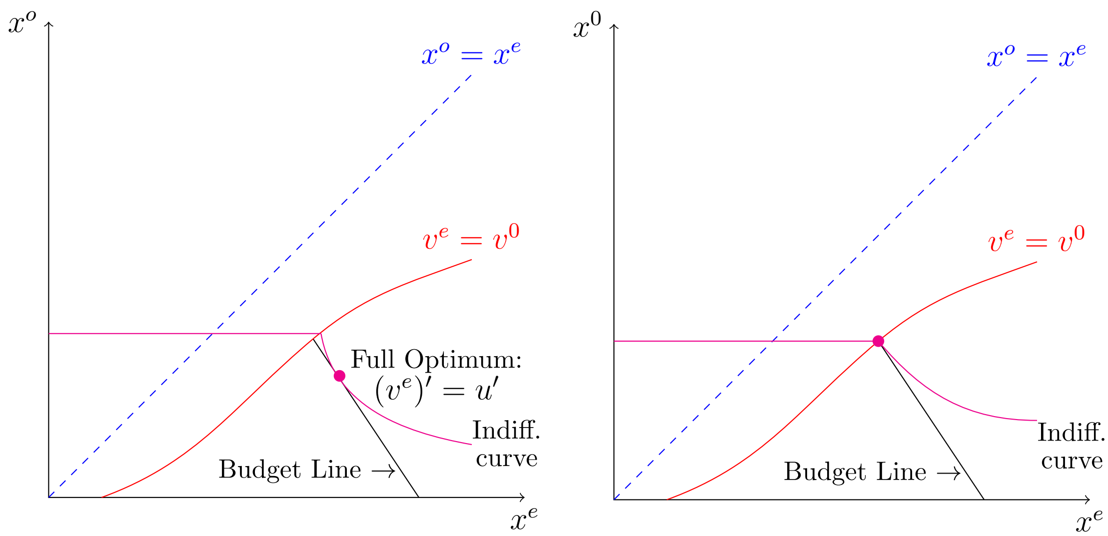
One of two types of optimum can emerge depending on whether $(v^e)' > u'$ or $(v^e)' \leq u'$ when evaluated at the point of indifference $v^e = v^o$. This is illustrated in Figure 1: Panel (a) shows that if $(v^e)' > u'$ at this point the incentive compatibility is non-binding and the 'Full Optimum' where $(v^e)' = u'$ can be obtained. Contrastingly, if $(v^e)' \leq u'$ the incentive compatibility constraint will bind.

Nichols and Zeckhauser (1982) argue that optimal transfer programmes should place restrictions on intended transfer beneficiaries. For example, restrictions may be placed

[3]Note that, because individuals are identical, the objective function here is the expected utility of an individual who faces an exogenous probability of becoming unable to work.

Figure 1.1: Optima from Diamond and Mirrlees (1978)

(a) **Full Optimum**: $MRS = (1-\theta)/\theta$ (b) **Binding IC constraint optimum**



Notes. See Diamond and Mirrlees (1978, pp.333-334)

on earnings or the consumption of goods 'indicating' low ability. In a simple two-person setting, individuals differ in their wage and possibly preferences. Let n_l denote the productivity (wage) of a low ability individual, and n_h the productivity of a high ability individual, where $n_l < n_h$. Individuals have preferences over consumption (x) and labour (H), as given by $u^i(x_i, H_i)$; $i \in \{l, h\}$. Suppose that the government only has income information and operates a tax/transfer system giving rise to the individual budget constraint:

$$x = \begin{cases} n_i H_i + T & : n_i H_i \leq Z \\ n_i H_i - T & : n_i H_i > Z \end{cases}$$

where T is a tax/transfer depending on where earnings lie relative to the threshold Z.

Letting H_i^* denote an individual's optimal labour supply and $y_i^* = n_i H_i^*$ their resulting gross earnings, the authors show that the optimal choice of Z is slightly *smaller* than y_l^*. To qualify for the welfare payment then, a low ability individual has to restrict their labour supply to $Z/n_l < H_l^*$. This, however, only imposes a *second-order* welfare loss on them because they are still in the neighbourhood their optimal choice of H:

$$u^l(Z+T, Z/n_l) \approx u^l(y_l^*+T, y_l^*/n_l) \quad \text{(Second-order welfare loss)} \tag{1.3}$$

However, under the assumption that $u^h(y_l^*+T, y_l^*/n_h) = u^h(y_h^*-T, y_h^*/n_h)$ - such that a high skilled individual would be indifferent between masquerading or working their optimal amount were $Z = y_l^*$ - it must hold that setting $Z < y_l^*$ imposes a *first-order* welfare loss on a masquerader because they are pushed even further from their optimum choice of H (conditional on *receiving* T). Formally:

$$u^h(Z+T, Z/n_h) < u^h(y_l^*+T, y_l^*/n_h) = u(y_h^*-T, y_h^*/n_h) \quad \text{(First-order welfare loss)} \tag{1.4}$$

Such a restriction rules out masquerading and further, allows the government to finance a higher (compensatory) transfer to the low type through increasing the tax rate. Developing the argument further, should low ability individuals also consume more of an 'indicator' good such as medical care than a high ability person, there may be a role for in-kind transfers. In this case, converting part of the cash transfer into a quantity of the indicator good set slightly higher than the low income individual would optimally consume generates, analogous to the first case, (i) a second-order welfare loss for the low income type; but (ii) a first-order welfare loss for the high type whose consumption of the indicator good is far in excess of their optimum (Nichols and Zeckhauser, 1982).

Drawing on the work of Nichols and Zeckhauser (1982) and others, Blackorby and Donaldson (1988) are also concerned with the role non-cash transfers can play in the prevention of non-needy individuals masquerading as needy. Their analysis takes place in a simple two-person (able, infirm/ill), two-good (yams, medical care) model. Whilst both the able and infirm desire yams, only the infirm desires medical care. Specifically, the utility of an able person is given by the quantity of yams they consume, whilst the utility of an infirm is a general strictly-quasiconcave function of yams and medical care. For simplicity, an exogenous linear production possibility frontier constrains the quantity of yams and medical care in the economy. The authors establish, and compare, the first-best; second-best; and third-best Pareto optima which can arise. In a first-best setting, the government observes individual preferences and agents know this. The first-best frontier can be attained as a decentralised Walrasian equilibrium where the price of medical care relative to yams is unity (with yams normalised to price one). Next, in a second-best setting the government only knows the population proportion of individual types and cannot observe preferences. It designs (rations) consumption bundles which satisfy self-selection constraints for the able and infirm[4]. Any second-best optima which do not coincide with first-best allocations are characterised by overprovision of medical care because this does not violate incentive compatibility, whilst an increase in yams rationed to the infirms may. Turning to a third-best setting, the government faces the same informational asymmetry but now chooses (i) a price for medical care; and (ii) an equal income for both population types, so as to maximize the indirect utility function of the infirm subject to the production possibility frontier. Any deviation in the price of medical care relative to that in first-best corresponds to a tax or subsidy on medical care and generates optima that are Pareto dominated by some second-best optima. So with informational constraints, in many cases rationing (second-best) may be superior to a market-solution (third-best).

Building further on the Nichols and Zeckhauser (1982) analysis of restrictions on welfare recipients, a number of papers explore the role of conditioning transfers on unproductive work requirements (Cuff, 2000; Kreiner and Tranaes, 2005). Most recently, Kreiner and Tranaes (2005) explore the use of workfare requirements to restrict unemployment insurance to individuals in the labour force who become involuntarily unemployed, as opposed to those out of the labour force who are voluntarily unemployed. In particular, they are concerned with whether an unemployment insurance system with unproductive workfare requirements can be Pareto improving relative to a purely pecuniary unemployment insurance system. Individual preferences over consumption (x) and labour (H) are represented by $u^i(x_i, H_i)$ $i \in \{f, v\}$, where f denotes an individual in the labour force (working or involuntarily unemployed) whilst v denotes a voluntarily unemployed individual. Letting n_i denote individual productivity, a type v individual will never choose to work because $n_v < u^v_H(0,0)/u^v_x(0,0)$ [5]. Contrastingly, a working type f individual optimises labour supply, H^*_f , such that:

[4] So for the able to choose their intended consumption bundle they must receive more yams than an infirm. Meanwhile, for an infirm to choose their intended bundle their utility from the yams and medical care rationed to them must exceed that for consuming the yams allocated to the able.

[5] The right side is simply the reservation wage rate above which an individual would choose to supply labour, as derived by setting $H = 0$ in the first order condition for labour supply.

$$n_f = \frac{u_H^f(n_f H_f^* - T, H_f^*)}{u_x^f(n_f H_f^* - T, H_f^*)} \equiv MRS^f(n_f H_f^* - T, H_f^*)$$

where T denotes a tax on workers necessary to finance an unemployment insurance scheme (and other transfers) and MRS is the marginal rate of substitution. In the population a fraction $\theta \in (0,1)$ of individuals are of type f; whilst the remainder are of type v. For those of type f; a fraction $\rho \in (0,1)$ receive a job offer.

An unemployment insurance scheme provides a consumption-workfare bundle $\{B, H^R\}$ to involuntarily unemployed individuals, where B denotes the pecuniary benefit and H^R the work requirement. For type v individuals there is an unconditional income support $B_0 \leq B$. With knowledge of the population proportion θ, the government's problem is described by:

$$\max_{B,B_0,T,H^R} E(u^f) = \rho \underbrace{u^f(n_f H_f^* - T, H_f^*)}_{\text{utility from working}} + (1-\rho)\underbrace{u^f(B, H^R)}_{\text{utility from UI}}$$

$$\begin{aligned}
\text{s.t.} \quad u^f(n_f H_f^* - T, H_f^*) &\geq u^f(B, H^R) && \text{(I1: Type } f \text{ prefers work to UI)} \\
u^v(B_0, 0) &\geq u^v(B, H^R) && \text{(I2: Type } v \text{ prefers } B_0 \text{ to UI)} \\
\theta(1-\rho)B + (1-\theta)B_0 &\leq \theta\rho T && \text{(I3: Benefit Budget Constraint)} \\
\underline{B} &\leq B_0 && \text{(I4: Minimum Type N support)}
\end{aligned} \tag{1.5}$$

Under a pure pecuniary unemployment insurance scheme, $\{B, 0\}$, it is straightforward to see from I2 that $B_0 = B$, whilst from I3 we have $T = B(1-\theta\rho)/\theta\rho$ such that, provided I1 holds:[6]

$$\underbrace{1 - \frac{u_x^f(n_f H_f^* - T, H_f^*)}{u_x^f(B, 0)}}_{\text{MB of increasing B}} = \underbrace{\frac{1-\theta}{1-\theta\rho}}_{\text{MC of increasing B}} \leq 1 \tag{1.6}$$

[6]With no work requirements, the only choice variable is B, and thus the first order condition resulting from (1.5) is:

$$-u_x^f(n_f H_f^* - T, H_f^*) \cdot \left\{\frac{1-\theta\rho}{\theta}\right\} + (1-\rho)u_x^f(B, 0) = 0 :$$

Rearranging and adding 1 to both sides gives the result in the text.

The left-side captures how much a type f individual values, in monetary terms, a marginal increase in B. The right side, meanwhile, captures the extent to which an increase in taxes necessary to increase B are misspent 'leaking' unemployment insurance to individuals of type v.

The relevant question is now whether a workfare scheme, $\{B, H^R\}$, can improve upon the pure pecuniary system, for a given B_0. Totally differentiating $E(u^f)$ and evaluating at $H^R = 0$ yields the below condition for $dE(u^f) > 0$: [7]

$$\underbrace{1 - \frac{u_x^f(n_f H_f^* - T, H_f^*)}{u_x^f(B,0)}}_{\text{MB of increasing B}} \geq -\frac{u_H^f(B,0)}{u_x^f(B,0)} \cdot \frac{u_x^v(B,0)}{u_H^v(B,0)} \equiv \frac{MRS^f(B,0)}{MRS^v(B,0)} \tag{1.7}$$

Notice that whilst the marginal benefit (left side) is identical to the optimality condition for the pecuniary unemployment insurance scheme $\{B, 0\}$, the marginal cost (right side) now captures the reduction in welfare from increased workfare. If the marginal utility of leisure of Type f individuals is sufficiently different (lower) from that of Type v individuals, then the introduction of workfare into a unemployment insurance scheme can be Pareto improving (Kreiner and Tranaes, 2005).

There has also been some work exploring the role of workfare when the objective is poverty alleviation. Besley and Coate (1992) analyse the standard screening argument for the imposition of workfare, in addition to a deterrent argument whereby ability enhancing effort is encouraged earlier in life to avoid welfare (and thus the requirement to engage in workfare). In a population of N individuals, a fraction $\theta \in (0,1)$ are of low earning ability, n_l, whilst the remaining proportion $(1-\theta)$ are of high earning ability, n_h. Quasilinear individual preferences are given by $u(x,H) = x - g(H)$; $g'(\cdot) > 0$, $g''(\cdot) > 0$, where x is consumption and H is labour supply. Let H_i^* ; $i = \{l, h\}$

[7] Totally differentiating (1.5) given the choice variables T, B and H^R gives:

$$dE(u_f) = \rho u_x^f(n_f H_f^* - T, H_f^*) \cdot -dT + (1-\rho)\left\{u_x^f(B, H^R)dB + u_H^f(B, H^R)dH^R\right\}$$

Noting that (i) from I2: $0 = u_x^v dB + u_H^v dH^R \Rightarrow dH^R = -dB u_x^v(B, H^R)/u_H^v(B, H^R)$; whilst (ii) from I3: $\theta(1-\rho)dB = \theta\rho dT \Rightarrow dT = dB(1-\rho)/\rho$; and so:

$$= (1-\rho)dB\left\{-u_x^f(n_f H_f^* - T, H_f^*) + u_x^f(B, H^R) - u_H^f(B, H^R) \cdot \frac{u_x^v(B,0)}{u_H^v(B, H^R)}\right\}$$

Evaluating at $H^R = 0$, it is straightforward to obtain (1.7).

denote optimal labour supply, where at an interior solution H_i^* satisfies $n_i = g'(H_i^*)$. The resulting indirect utility function for a type i individual is $v(n_i) = n_i H_i^* - g(H_i^*)$. Letting Z denote an arbitrary poverty line, it is assumed that $n_l H_l^* \leq Z < n_h H_h^*$, thereby giving the government a reason to implement a poverty reduction programme to assist the low types. A welfare consumption bundle, (B_i, H_i^R), consists of a welfare payment, B, and an unproductive workfare requirement, H^R. Notably, an individual of either ability type receiving welfare may still work in the private sector. Intuitively, they will choose to do so if the work requirement does not exceed their optimal labour supply, and thus if $H_i^R \leq H_i^*$.

Turning to information structure, the government knows the distribution parameter θ but cannot identify the type of a given individual, thereby giving rise to the problem of incentive compatibility. The indirect utility of an individual receiving their intended welfare package is $v(n_i, B_i, H_i^R)$, whilst if they receive that intended for the other type it is $v(n_i, B_{-i}, H_{-i}^R)$. Putting this all together, the objective of the government is to choose expenditure minimising consumption bundles that eliminate poverty and satisfy incentive compatibility constraints. The incentive compatibility constraints depend on whether individual earnings can be observed. Should earnings be unobservable then we have the following incentive compatibility(IC) constraints :

$$\begin{aligned} v(n_h, B_l, H_l^R) &\leq v(n_h, B_h, H_h^R) \\ v(n_l, B_h, H_h^R) &\leq v(n_l, B_l, H_l^R) \end{aligned} \quad \text{(Earnings Unobservable IC constraints)} \tag{1.8}$$

Contrastingly, should earnings be observable then there is now a cost to high types masquerading, as they must restrict their labour supply to $\bar{H}_h$ satisfying $n_H \bar{H}_h = n_l(MAX[0, H_l^* - H_l^R])$. The IC constraint is thus:

$$B_l + n_h \bar{H}_h - g(\bar{H}_h + H_l^R) \leq v(B_h, H_h^R, N_h) \quad \text{(Earnings Observable IC constraint)} \tag{1.9}$$

In both cases, there exists a work requirement level that separates the two types. Necessarily, this requirement is lower with observable earnings because masquerading individuals are already harmed through restricting their labour supply. The optimality of workfare relative to a purely pecuniary welfare programme rests on the extent to which workfare reduces, or 'crowds out', private earnings, thereby exacerbating the poverty gap. Intuitively, the greater n_l and the lower θ, the more likely workfare is to be optimal. Further, workfare is more likely to be optimal where earnings are unobservable. It may therefore be more relevant in a developing country context, as opposed to a developed one. Where poverty is partly endogenous, in the sense that ability is determined partly by effort earlier in life, it may be optimal to impose a 'maximal' work requirement that leaves a low type indifferent to welfare programme participation, thereby incentivising effort earlier in life (Besley and Coate, 1992).

1.1.2 Awards Technology that 'Tags' only Needy Individuals.

In each of the frameworks considered so far there have been two types of individual and the aim has been to design a mechanism that prevents the non-needy from claiming benefits. In this section we now discuss optimal transfers when the government can use categorical information to 'tag' individuals as needy and award them categorical transfers. It will be assumed that only needy individuals are tagged, though not necessarily all needy individuals. Issues surrounding masquerading are therefore abstracted from and the awards technology does not make false award errors.

These analyses are firmly grounded in the optimal income tax framework originating from Mirrlees (1971). It is therefore natural to first discuss the standard uni-dimensional model of optimal taxation; where individuals differ over a continuum of productivities. This has the benefit of allowing us to explore optimal universal welfare provision. Following the seminal contribution of Akerlof (1978), we then introduce categorical heterogeneity into the optimal tax framework. The government will have access to technology that allows it to partition society into categorical groups and design tax/transfer schedules using this information.

The Uni-Dimensional Optimal Tax Framework

The standard optimal nonlinear tax framework originates from Mirrlees (1971). It also provides the framework for optimal linear taxation (Sheshinski, 1972). Individuals have identical preferences over consumption, $x \geq 0$, and labour, $H \in [0, 1]$, as represented by the utility function $u(x, H)$. The standard assumptions apply: u is continuous, increasing in x ($u_x > 0$), decreasing in H ($u_H < 0$) and strictly concave ($u_{xx} < 0$,

$u_{HH} < 0$, $u_{xx}u_{HH} - u_{xH}^2 > 0$). Individuals differ in the sole dimension of productivity (n); where n is continuously distributed with density function $f(n)$ and associated distribution function $F(n)$.

Social welfare takes the (generalised) utilitarian form $W = \int G(u)f(n)dn$, where G is a monotonic concave (or linear) transformation of individual utility.[8] Note that if $G(u) = u$ (or equivalently if $G'(u) = 1$) the government is strictly utilitarian and a desire for redistribution arises solely from the concavity of individual utility. For simplicity, we throughout focus on this case. Given this objective, the government wishes to design a redistributive tax schedule but is constrained by the fact that it only observes earnings (which are endogenous to the tax system) and not the two drivers of these earnings: productivity and labour supply (effort). This gives rise to the fundamental tradeoff between redistributive gains (equity) and behavioural responses (efficiency). Redistribution from high earners to low earners is desirable on the grounds of diminishing marginal utility of income, but taxes have a distortionary affect on individual behaviour.

Benchmark: First-Best (Lump-Sum) Taxation

As a benchmark case it is useful to first outline the first-best tax optimum which would arise if the government could in fact observe individual productivity. This will also prove useful for the later discussion of categorical transfers, which share a number of similarities with first-best taxation (Viard, 2001a). Note that a tax on individual productivity is non-distortionary in the sense that individuals cannot alter the amount of tax they pay through adjusting their labour supply.

Let $M(n)$ be a lump-sum tax/transfer that is conditioned on individual productivity. Individual labour supply (H^*) and the resulting indirect utility function (v) are:

$$\begin{aligned} H^*[n, M(n)] \equiv& \arg\max_{H\in[0,1]} u[nH + M(n), H] \\ v[n, M(n)] \equiv& u[nH^* + M(n), H^*] \end{aligned} \tag{1.10}$$

[8] This welfare function is the dominant approach in much of public finance. For useful discussions see Beath et al. (1988); Lewis and Ulph (1988); Mankiw and Weinzierl (2010); Piketty and Saez (2012); Saez and Stantcheva (2013).

where at an interior solution $H^* > 0$ satisfies $n = -u_H/u_x$. By the envelope theorem we have:

$$\begin{aligned}\frac{\partial v}{\partial n} &= u_x H^* + \frac{\partial H^*}{\partial n}\underbrace{(nu_x + u_H)}_{=0} = u_x H^* \\ \frac{\partial v}{\partial M} &= u_x + \frac{\partial H^*}{\partial M}\underbrace{(nu_x + u_H)}_{=0} = u_x\end{aligned} \tag{1.11}$$

from which we obtain Roy's identity $v_n = v_M H^*$ and thus $v_{nM} = v_{MM}H^* + v_M H^*_M$. By the concavity of utility and the normality of leisure it is straightforward to establish that $H^*_M < 0$ and thus $v_{nM} < 0$: i.e. the marginal indirect utility of unearned income is decreasing in productivity.

From (1.11) we can readily establish that:

$$\frac{dv}{dn} = \frac{\partial v}{\partial n} + \frac{\partial v}{\partial M}M'(n) = u_x\left(H^* + M'\right) \tag{1.12}$$

$$\frac{dH^*}{dn} = \frac{\partial H^*}{\partial n} + \frac{\partial H^*}{\partial M}M'(n) = \frac{\partial H^c}{\partial n} + \frac{\partial H^*}{\partial M}\left(H^* + M'\right) \tag{1.13}$$

where H^c is compensated labour supply satisfying the Slutsky-Hicks equation:[9]

$$\frac{\partial H^c}{\partial n} = \frac{\partial H^*}{\partial n} + H^*\frac{\partial H^*}{\partial M} \tag{1.14}$$

With this by way of background, the ability-dependent tax transfer schedule $M(n)$ is

[9] The expenditure minimisation problem is:

$$\min_{x,H} x - nH \quad \text{s.t.} \quad u(x,H) \geq \bar{u}$$

Let $x^c(n,u)$, $H^c(n,u)$ and the lagrange multiplier $\mu^c(n,u)$ denote the optimal choices. The FOC's characterising these choices are:

$$(x): \quad 1 - \mu^c \mu_x = 0 \;\; , \;\; (H): \quad n + \mu^c u_H = 0 \;\; , \;\; (\mu): \quad u - \bar{u} = 0$$

Let $E(n,u) = x^c - nH^c$ denote the resulting expenditure function. By the envelope theorem:

$$\frac{\partial E}{\partial n} = \frac{\partial x^c}{\partial n}\underbrace{(1 - \mu^c u_x)}_{=0} - \frac{\partial H^c}{\partial n}\underbrace{(n + \mu^c u_H)}_{=0} - \frac{\partial \mu^c}{\partial n}\underbrace{(u - \bar{u})}_{=0} - H^c = -H^c$$

Differentiating the identity $H^c(n,u) = H^*(n, E(n,u))$ w.r.t. n then gives (1.14) in the main text.

chosen so as to:

$$\max_{M(n)} W = \int_0^\infty v[n, M(n)] f(n) dn \tag{1.15}$$
$$\text{s.t.} \quad \int_0^\infty M(n) f(n) dn - R = 0$$

Letting $\hat{M}(n)$ denote the resulting optima, we have:

$$v_M[n, \hat{M}(n)] = \lambda \ \ \forall\, n \tag{1.16}$$

The left side of (1.16) is the social marginal value of income (smvi) of a productivity n individual, whilst the right side is the shadow price of public expenditure. Intuitively, the condition states that the tax/transfer schedule M should be chosen so as to equate the smvi of all ability types (Viard, 2001a,b).

Differentiating (1.16) with respect to n gives:

$$M'(n) = -H^* - \underbrace{\left\{ \frac{v_M}{v_{MM}} \right\}}_{\leq 0} H_M \leq 0 \tag{1.17}$$

The optimal lump-sum transfer is therefore decreasing in productivity: i.e. higher ability individuals will pay taxes ($M < 0$) to finance transfers ($M > 0$) to those of lower ability. Substituting (1.17) into (1.12) and (1.13) clearly illustrates that at the optimum we have $dv/dn < 0$ and $dH^*/dn > 0$ for all those who work. This arises because it is efficient to encourage the most productive individuals in society to work (see Hellwig, 1986; Helpman and Sadka, 1978; Viard, 2001a,b).[10]

[10] Suppose preferences are separable in consumption and labour and take the form $u(x, H) = v(x) - \gamma H$, where $v' > 0$ and $v'' < 0$. The first-best problem is now:

$$\max_{M(n)} \int_0^\infty v[nH + M(n)] - \gamma H f(n) dn \quad \text{s.t.} \quad \int_0^\infty M(n) f(n) dn - R = 0$$

The resulting FOC is $v'[nH + M(n)] = \lambda \ \forall\, n$. This states that the marginal utility of consumption should be equated across productivity types, and thus so too should be consumption. At the optimum then, those with above average productivity will be taxed, whilst those with below average productivity will receive transfers (Mankiw and Weinzierl, 2010). More eloquently: '*the solution to this problem in the abstract is that the richer should be taxed for the benefit of the poorer up to the point at which complete equality of fortunes is attained*' (Edgeworth, 1897, p.553).

Optimal Linear Income Tax Framework

In his seminal contribution on nonlinear taxation, Mirrlees (1971, p.208) discusses the desirability of approximately linear income taxation and the administrative advantages this would bring. Indeed, proponents of linear (flat) income taxes cite their administrative simplicity and enhanced work incentives (Paulus and Peichl, 2009; Peichl, 2014). Analytically, the analysis of linear income taxes captures the equity-efficiency tradeoff inherent in income taxation more tractably than nonlinear taxation.

Individuals face the budget constraint $x = n(1-t)H + B$; where $t \in (0, 1)$ is a constant marginal tax rate on earnings and B is a tax-free universal benefit. B receives many names in the literature, such as 'demogrant' or 'basic income' (see Van Parijs, 2004). Atkinson (1995) employed this framework in his analysis of 'the basic income/flat tax proposal'.

An implication of providing a universal benefit is that there will be a reservation productivity $\bar{n}$ satisfying $\bar{n}(1-t) = -u_H(B, 0)/u_x(B, 0)$ at or below which an individual will choose not to work. So for all $n \leq \bar{n}$: $H^* = 0$; whilst for all $n > \bar{n}$: $H^* > 0$. To save on notation, let $y^*(n, 1-t, B) = nH^*[n(1-t), B]$ be individual gross earnings; whilst $\bar{y}^* = \int y f(n) dn$ is the average earnings in the economy.

In a population of size 1, the government's problem is now described by:

$$\max_{t,B} W(t, B) = \int_0^\infty v[n(1-t), B] f(n) dn \tag{1.18}$$
$$\text{s.t.} \quad B = t \int_0^\infty y^* f(n) dn - R$$

where R is an exogenous revenue requirement for spending outside of welfare.

For the purpose discussing the results which follow, let the *net* smvi of a productivity individual be given by (Atkinson and Stiglitz, 1980; Atkinson, 1995; Viard, 2001a,b):

$$s(n, t, M, \lambda) = \begin{cases} u_x(M, 0) & : n \leq \bar{n} \\ v_M[n(1-t), M] + \lambda t y^*_M(n, 1-t, M) & : n > \bar{n} \end{cases} \tag{1.19}$$

where λ is the shadow price of public expenditure. For individuals who do not work s

is simply the social marginal utility of income. Contrastingly, for individuals who do work s also captures - in welfare units - the fact that an increase in unearned income induces an individual to reduce their labour supply and thus lowers tax revenue. We let $\bar{s}(t, M, \lambda) = \int s f(n) dn$.

Let $\hat{B}$ and $\hat{t}$ denote the optimal choices which result from (1.18). The first order conditions (FOCs) characterising these optima are:

$$(B): \quad \bar{s}\left(\hat{t}, \hat{B}, \hat{\lambda}\right) = \hat{\lambda} \tag{1.20}$$

$$(t): \quad \int_0^\infty \left\{-nv_\omega + \hat{\lambda}\left[y^* - \hat{t}\frac{\partial y^*}{\partial(1-t)}\right]\right\} f(n) dn = 0 \tag{1.21}$$

where $\hat{\lambda}$ is the shadow price of public expenditure at the optimum.

Condition (1.20) states that at the optimum the average net smvi associated with an increase in the universal benefit must equate with, in welfare units, the marginal cost of increasing the benefit (given a population of size 1 this enters as unity). Next, following (Atkinson and Stiglitz, 1980) we can use Roy's identity ($v_\omega = v_M H^* = v_M y^*/n$) and the Slutsky-Hicks equation[11] to write (1.21) as:

$$\begin{aligned} \int_0^\infty y^*\left(1 - \frac{v_M}{\hat{\lambda}} - \hat{t}\frac{\partial y^*}{\partial M}\right) f(n) dn &= \hat{t}\int \frac{\partial y^c}{\partial(1-t)} f(n) dn \\ \Rightarrow \quad \int_0^\infty y^*\left(\hat{\lambda} - s\right) f(n) dn &= \hat{t}\hat{\lambda}\int \frac{\partial y^c}{\partial(1-t)} f(n) dn \end{aligned} \tag{1.22}$$

From (1.20) we know that $\bar{s} = \hat{\lambda}$ and so the left side of (1.22) can be written as $\bar{y}^*\bar{s} - \int y^* s f(n) dn = -\text{Cov}(y^*, s)$. Finally, dividing both sides of (1.22) by the net-of-tax rate $(1-t)$ thus gives the well-documented expression:

$$\frac{\hat{t}}{1-\hat{t}} = \frac{-\text{Cov}(y, s)}{\hat{\lambda}\int y\mathcal{E}^c f(n) dn} \quad \text{i.e.} \left(\frac{\text{Equity}}{\text{Efficiency}}\right) \tag{1.23}$$

[11] In terms of earnings, the Slutsky-Hicks equation can be written as:

$$\frac{\partial y^c}{\partial(1-t)} = \frac{\partial y^*}{\partial(1-t)} - y^*\frac{\partial y^*}{\partial M}$$

where:

$$\mathcal{E}^c = \frac{1-t}{y} \cdot \frac{\partial y^c}{\partial(1-t)}$$

is the compensated elasticity of earnings with respect to the net-of-tax rate.

The optimal tax expression in (1.23) has equity considerations in the numerator and efficiency considerations in the denominator. The numerator is the negative of the covariance between individual gross earnings and the net smvi. Under the frequently employed assumption of agent monotonicity earnings will rise with productivity. Contrastingly, by *ordinal* properties the net smvi will fall with productivity. The intuition is that the government would like to redistribute from those of high productivity (and thus low smvi) towards those of low productivity (and thus high smvi).

This redistribution is constrained, however, by the efficiency considerations in the denominator. *Ceteris paribus*, larger compensated elasticities of earnings favour lower tax rates; with emphasis placed on both high productivities and productivities where the population is most dense (Kaplow, 2008).

Moving Beyond the Uni-Dimensional Framework

Suppose that we modify the linear income tax framework so that a fraction $\theta \in (0,1)$ of the population faces a zero quantity constraint on labour supply and is thus unable to work. The model remains otherwise the same.

In this case the government's problem is:

$$\begin{aligned}\max_{t,B} W(t,B;\theta) &= \theta u(B,1) + (1-\theta)\int_0^\infty v[n(1-t),B]f(n)dn \\ \text{s.t. } \; B &= t\bar{y}^* - R\end{aligned} \tag{1.24}$$

The FOCs characterising the optimal universal benefit and tax rate are now:

$$(B): \quad \theta u_x(\hat{B},1) + (1-\theta)\bar{s}\left(\hat{t},\hat{B},\hat{\lambda}\right)f(n)dn = \hat{\lambda} \tag{1.25}$$

$$(t): \quad \frac{\hat{t}}{1-\hat{t}} = \frac{\int_0^\infty y^*(\hat{\lambda}-s)f(n)dn}{\hat{\lambda}\int \mathcal{E}^c f(n)dn} \tag{1.26}$$

The presence of a dependent population means that we can *no longer* use the FOC characterising $\hat{B}$ to write the numerator of the optimal tax expression as the negative of the covariance between earnings and the net smvi. Further, at the optimum we can see that $u_x(\hat{B},1) > \hat{\lambda} > \bar{s}(\hat{t},\hat{B},\hat{\lambda})$ because the universal benefit does not remove inequality in the average net smvi *between* the unable and able subpopulations. If the government could condition transfers on disability, it would certainly seem desirable to provide a categorical transfer to the unable (Atkinson, 1995).

Tagging Transfers: Categorical Heterogeneity

In the standard optimal tax model individuals differ in the sole dimension of productivity. However, the value of introducing additional dimensions of heterogeneity into tax/transfer models - and exploiting this through categorical information - has been known since Akerlof (1978). Akerlof considers a simple model in which there is a population composed of an equal share of high ability and low ability individuals. Low ability individuals may only work in an easy job and earn (pre-tax/transfer) n_e. Contrastingly, high ability individuals may choose to work in either (i) a difficult job and earn (pre-tax/transfer) n_d ; or (ii) an easy job and earn n_e.

In the absence of any categorical information on ability, the government has in place a tax/transfer system whereby those in difficult jobs pay a tax $t_d \geq 0$, whilst those in easy jobs receive a transfer $t_e \geq 0$. The utility of an individual working in an easy job is $u(n_e + t_e)$, where $u'(\cdot) > 0$ and $u''(\cdot) < 0$. Meanwhile, the utility of an individual working in a difficult job is $u(n_d - t_d) - \gamma$, where γ captures the disutility of working in the harder job. The assumption is made that $u(n_d) - \gamma > u(n_e)$ such that, in absence of any taxes or transfers, a high ability individual would always choose a difficult job. Putting this all together, the government's problem is described by:

$$\begin{aligned}
&\max_{t_e,t_d} W(t_d,t_e) = \frac{1}{2}\max[u(n_d - t_d) - \gamma, u(n_e + t_e)] + \frac{1}{2}u(n_e + t_e) \\
\text{s.t.} \quad & u(n_d - t_d) - \gamma < u(n_e + t_e) \Rightarrow t_e = 0 \quad \text{(Incentive compatibility)} \\
& u(n_d - t_d) - \gamma \geq u(n_e + t_e) \Rightarrow t_d = t_e \quad \text{(Taxes = Transfers)}
\end{aligned} \tag{1.27}$$

The first condition states that if the tax system is such that high ability individuals prefer easy jobs to difficult jobs then no tax revenue will be generated and consequently $t_e = 0$. The second condition, meanwhile, states that tax revenue is used entirely to finance transfers and, given the equal population shares, taxes equate with transfers. Letting $\hat{t}_d$ and $\hat{t}_e$ denote the optimum transfers, it is straightforward to establish that:[12]

$$\hat{t}_d = \hat{t}_e \quad , \quad u(n_d - \hat{t}_d) - \gamma = u(n_e + \hat{t}_e) \quad \text{(No Tagging Optimum)} \tag{1.28}$$

Next, when the government has can identify the fraction $\alpha \in (0, 1]$[13] of low ability individuals as low ability, it chooses t'_d, t'_e (which may now be a tax or transfer) in addition to a tagged transfer t_l so as to:

$$\begin{aligned} \max_{t'_d, t'_e, t_l} W^{\mathrm{TAG}}(t'_d, t'_e, t_l) = & \frac{1}{2} \max \left\{ u(n_d - t'_d) - \gamma, u(n_e + t'_e) \right\} \\ & + \frac{1}{2} \left\{ \alpha u(n_e + t_l) + (1 - \alpha) u(n_e + t'_e) \right\} \end{aligned}$$

$$\begin{aligned} \text{s.t.} \quad & u(n_d - t'_d) - \gamma < u(n_e + t'_e) \Rightarrow \alpha t_l = -(2 - \alpha) t'_e \quad \text{(Incentive compatibility)} \\ & u(n_d - t'_d) - \gamma \geq u(n_e + t'_e) \Rightarrow t'_d = \alpha t_l + (1 - \alpha) t'_e \quad \text{(Taxes = Transfers)} \end{aligned} \tag{1.29}$$

The incentive compatibility condition simply states that if everyone chooses easy jobs then any tax on easy jobs must finance transfers to tagged individuals. Akerlof (1978)

[12] Drawing on Akerlof (1978), the proof tests two hypotheses.

(i) First, suppose that $u(n_d - \hat{t}_d) - \gamma > u(n_e + \hat{t}_e)$ and let $\tilde{t}_d = \hat{t}_d + \epsilon$. From (1.27), a Taylor approximation of $W(\tilde{t}_d, \tilde{t}_e)$ around $\epsilon = 0$ yields $W(\hat{t}_d, \hat{t}_e) + \frac{1}{2}[u'(n_e - \hat{t}_e) - u'(n_d - \hat{t}_d)] \cdot \epsilon + \sigma_\epsilon$, where $\lim_{\epsilon \to 0} \sigma_\epsilon = 0$. Next, the assumption that $u(n_d - \hat{t}_d) - \gamma > u(n_e - \hat{t}_e)$ implies, via the fact that $u'' < 0$, that $u'(n_d - \hat{t}_d) < u'(n_e + t_e)$. From the Taylor approximation, this implies that $W(\tilde{t}_d, \tilde{t}_e) > W(\hat{t}_d, \hat{t}_e)$ for small ϵ. But this contradicts the fact that $\hat{t}_d$ and $\hat{t}_e$ are optimum. The assertion that $u(n_d - \hat{t}_d) - \gamma > u(n_e + \hat{t}_e)$ must therefore be false. (ii) Next, suppose that $u(n_d - t_d) - \gamma < u(n_e + t_e)$ and thus $t^*_e = 0$ because no high skilled individuals works in a difficult job. Total welfare thus becomes $u(n_e)$. But by the concavity of welfare this can be easily improved upon $\frac{1}{2}u(n_d) - \gamma + \frac{1}{2}u(n_e)$. Putting this all together, it must hold that $u(n_d - t_d) - \gamma = u(n_e + t_e)$.

[13] Note that if $\alpha < 1$ some low ability individuals are not tagged. This is analogous to a Type I (false rejection) error.

demonstrates that, at the optimum:

$$\hat{t}_l > \hat{t}_e \quad , \quad u(n_d - t_d') - \gamma = u(n_e + t_e') \tag{1.30}$$

So it is welfare enhancing to award tagged individuals higher transfers than untagged workers in easy jobs. An increase in the transfer size provides less of an incentive for skilled workers to move to an easy job than in the no tagging case, where all those in an easy job receive a non-negative transfer.

Following Akerlof (1978) , a number of papers introduce categorical transfers into the optimal tax framework (both linear and nonlinear). These papers assume that the government can perfectly partition society into categorical groups and thus has two tools available to it: categorical transfers and income based taxes/transfers.

Categorical Transfers in the Linear Income Tax Framework

Viard (2001a,b) analyses categorical transfers in the the optimal linear income tax framework. Using categorical information on exogenous characteristics which are correlated with ability, the planner can perfectly classify individuals into J disjoint groups, each with its own ability distribution, $f_j(n)$. Each group represents the proportion θ_j ; $j = 1, ..., J$ of the total population, where $\sum_j \theta_j = 1$ and $\sum_j \theta_j f_j(n) = f(n)$. In terms of information structure, the planner cannot observe ability but does know the distribution of abilities for each group. However, because it knows the categorical group to which an individual belongs, it can provide a categorical transfer C_j to individuals in each group.

Putting this all together, the planner's optimisation problem is described by:

$$\begin{aligned} &\max_{C_j,t} W = \sum_j \theta_j \int_0^\infty v[n(1-t), C_j] f_j(n) dn \\ \text{s.t.} \quad &\sum_j \theta_j C_j \leq \sum_j \theta_j \cdot t \int_0^\infty n H^*[n(1-t), C_j] f_j(n) dn - R \end{aligned} \tag{1.31}$$

where R is an exogenous revenue requirement[14].

[14] Viard (2001a, p.489) notes in his numerical analysis that $C_j \geq 0$: a negative transfer would harm too much the very low productivity individuals in a categorical group. Categorical transfers are therefore financed entirely through tax revenue.

Letting $\hat{C}_j$ and $\hat{t}$ denote the optimal choices, the FOC characterising optimal categorical transfers is:

$$\int_0^{\infty} s\left(n, \hat{t}, \hat{C}_j, \hat{\lambda}\right) f_j(n) dn = \hat{\lambda} \;\; \forall\, j \tag{1.32}$$

At the optimum then, the role of categorical transfers is to eliminate inequality in the average net smvi between categorical groups. Viard (2001a) notes the 'fundamental' similarity between categorical transfers and first-best lump-sum transfers, as can be seen from comparing (1.32) with (1.16). Whilst lump-sum transfers eliminate inequality in the smvi between *indivdiuals*, categorical transfers aim to (if feasible) eliminate inequality in the average net smvi between categorical groups. In the first case, the planner has information on each individual, whilst in the second it only has information on categorical groups. In the extreme case where the planner has categorical information on characteristics which are perfectly correlated with earnings ability, the two types of transfers coincide and the planner can eliminate inequality in the smvi across all individuals. Analogous to first-best, optimal categorical transfers are shown to be decreasing in group average ability under the assumption that $H_{MM} < 0$ and $H_{M\omega} < 0$. This assumption guarantees that the net smvi is decreasing in both ω and M. Further, categorical transfers can induce some quantiles of higher ability groups to have greater labour supply and lower utility levels than their counterparts in lower ability groups.

In reporting the optimal tax formula Viard (2001a,b) implicitly *assumes* that (1.32) holds: this allows the optimal tax expression to be written analogously to that in the unidimensional model where individuals differ only in productivity (i.e. equation 1.23). Chapter 3 of this thesis will return to discuss the appropriateness of this assumption.

Viard (2001b) notes that categorical transfers and the income tax are imperfect substitutes: The former remedies between-group inequality but not within-group inequality, whilst the latter plays some role in addressing both. Numerical simulations using a log-normal ability distribution and CES utility function are used to provide some support to propositions which follow from linear approximations of first-order conditions. Whenever there is low within-group inequality or labour supply is highly elastic, the importance of categorical transfers relative to the income tax as a redistributive tool is increased. Because the income tax is used to a lesser extent, a greater proportion of the between-group inequality is remedied by the perfectly administered categorical transfers, as captured by larger differences in transfers between-groups.

Categorical Transfers in the NonLinear Income Tax Framework

In a nonlinear income tax framework, Immonen et al. (1998) explore a similar issue to Viard (2001a,b). Specifically, categorical information is used identify a proportion $\theta \in (0,1)$ of the population as belonging to categorical group a, whilst the remaining $(1-\theta)$ of the population fall into categorical group b. The distribution of ability within each of the two groups is given by the density function $f_j(n)$; $j \in \{a, b\}$, with corresponding distribution function $F_j(n)$. There may also be heterogeneity in individual preferences, as given by $u^j(x, H)$, where x denotes consumption and H labour time. An individual from group j faces the budget constraint $x = nH - T_j(nH)$, where T_j denotes the group specific nonlinear tax schedule. An individual's resulting indirect utility function is $v^j(n)$. Given this, the government's overall problem is to choose group specific tax schedules and revenue requirements, given an overall exogenous revenue requirement of R, so as to:

$$\begin{aligned} \max_{T_j, R_j} W &= \int_0^\infty [\theta v^a(n) f_a(n) + (1-\theta) v^b(n) f_b(n)] dn \\ \text{s.t.} \quad R_a + R_b = R &= \int_0^\infty [\theta T_a f_a(n) + (1-\theta) T_b f_b(n)] dn \end{aligned} \tag{1.33}$$

This overall problem can, however, be decomposed into a two-stage problem. In the first stage, the government solves the within-group standard Mirrlees problem, taking as arbitrarily fixed R_j:

$$\max_{T_j} W^j = \int_0^\infty v^j(n) f_j(n) dn \quad \text{s.t.} \quad \int_0^\infty T_j f_j(n) dn = R_j \tag{1.34}$$

Given the optimal tax schedules from the first stage, the government then chooses the optimal group specific revenue requirements so as to:

$$\max_{R_a} W = W^a(R_a) + W^b(R - R_a) \tag{1.35}$$

The authors reconcile tensions between the Dilnot et al. (1984) argument of increasing (decreasing) marginal tax rates in the poor (rich) groups and the conventional optimal income tax literature result of decreasing marginal rates in income. Through numerical simulations employing CES preferences[15] and lognormal ability distributions, Immonen et al. (1998) argue that the intuition relating the two-dimensional and one-dimensional problems respectively, lies in group-specific revenue requirements and intergroup redistribution. Specifically, the revenue requirement for the poor group will differ in magnitude and potentially sign, thereby capturing intergroup transfers.

Next, in a setting where individual ability is revealed through successful job search, Boone and Bovenberg (2006) model the optimal combination of (i) an out-of-work (welfare) benefit; and (ii) in-work benefits , with the latter delivered directly through the nonlinear tax-schedule individuals face. There are two sources of inequity in the model: disparities in ability and involuntary unemployment. Formally, ability, n, is distributed continuously in the interval $[\underline{n}, \bar{n}]$ with density function $f(n)$. Individuals have quaslinear preferences $u(x, H) = v(x) - H$; $v' > 0$, $v'' < 0$, where x is consumption and H is labour supplied[16]. An individual who does not work therefore has utility $v(B)$, where B is the out-of-work welfare benefit. Contrastingly, an individual who does work has utility $u(n) = v[x(n)] - y^*(n)/n$, where $y^* = nH^*$ is gross earnings and $x(n) = y^*(n) - T(n)$ is net earnings.

In this static model individuals search for a job and the probability of finding a job is given directly be the search effort $e \in [0, 1]$. The costs involved in searching with intensity e are given by the cost function $k(e)$.[17] Individuals choose search intensity e so as to:

$$\tilde{u} = \max_{e} \; e \cdot u(n) + (1 - e)v(B) - \gamma \cdot e \tag{1.36}$$

Note that through finding a job an individual's ability is revealed to the government.

[15]Specifically, they use functions where the elasticity of substitution between consumption and leisure is 1 (i.e. Cobb-Douglas) and 0.5, respectively.

[16]As Boone and Bovenberg (2006) state, the assumption of quaslinear preferences that are concave in consumption and linear in labour renders a utilitarian government concerned with the distribution of consumption, but not the distribution of labour effort.

[17]For simplicity, we omit the precise properties of the cost function, which are stated in detail in Boone and Bovenberg (2006, p.169).

The utilitarian government's problem is formally described by:

$$\max_{s,y,x} \int_{\underline{n}}^{\bar{n}} \tilde{u}(n) f(n) dn \tag{1.37}$$

$$\text{s.t.} \quad \underbrace{B \int_{\underline{n}}^{\bar{n}} [1 - s(n)] f(n) dn}_{\substack{\text{Benefit expenditure on those} \\ \text{who do not find a job}}} = \underbrace{\int_{\underline{n}}^{\bar{n}} s(n) T(n) f(n) dn - R}_{\substack{\text{Tax Revenue net} \\ \text{of Revenue requirement}}}$$

The conclusions reached in the paper depend - among other things - on whether the government can optimise with respect to both the welfare benefit level and the nonlinear tax-schedule, or just the latter (taking the welfare benefit level as exogenously given). In the first case, the in-work benefits may exceed the welfare benefit due to the effect that the welfare benefit has on tax rates. Whilst high welfare benefits assist the most needy (the unemployed), they require the tax rate to be lowered on high productivity workers to prevent these workers departing the labour force in favour of receiving the welfare benefit. In contrast, in-work benefits have no such disincentive effects because the government can observe the ability of workers. Meanwhile, in the second case the authors describe a U-shaped relationship between welfare benefits and in-work benefits: A marginal increase in the welfare benefit from low levels substitutes for in-work benefits because the labour force participation constraint of high ability workers is non-binding (i.e. they are not incentivised to depart the labour force). However, higher levels of the welfare benefit require higher in-work benefits because the participation constraint becomes binding.

1.1.3 Awards Technology with Two-Sided Classification errors

A third strand of literature explores optimal transfers when the test, or monitoring technology, administering a given benefit makes two-sided classification errors. The propensity to make classification errors is largely taken to be exogenous, though a few papers do endogenize the propensities by making the test accuracy a function of the resources dedicated to it.

Exogenous Awards Technology

Diamond and Sheshinski (1995) model the optimal provision of (i) a disability benefit; and (ii) a retirement benefit, when the test determining disability makes both Type I and Type II classification errors. All individuals have the same marginal product of unity but differ over a continuum of labour disutilities, $\gamma \geq 0$, which are distributed by the continuous density function $f(\gamma)$ and corresponding distribution function $F(\gamma)$. An individual who works derives utility $\tilde{u}(x_w) - \gamma$, where $\tilde{u}'(\cdot) > 0$, $\tilde{u}''(\cdot) < 0$ and x_w is the consumption level associated with working. Contrastingly, an individual who does not work will have utility $u(x_i)$, $i \in \{d, r\}$, where $u'(\cdot) > 0$, $u''(\cdot) < 0$; x_d is the consumption when receiving the disability benefit; and x_r is the consumption when receiving the retirement benefit. The retirement benefit is received automatically by anyone who is (a) not working and (b) not receiving the disability benefit. The disability benefit, however, must be applied for. The assumption is made that $x_d \geq x_r$. The probability that a given individual with imperfectly observable labour disutility γ is awarded the disability benefit is given by $\rho(\gamma)$, where $\rho'(\gamma) > 0$. Drawing this all together, we can establish that:

1. An individual will choose to work over all other options (apply for disability, retire) if $\tilde{u}(x_w) - \gamma > u(x_d) \geq u(x_r)$. This implies that there is a critical disutility level, $\bar{\gamma}_d$ above which an individual will apply for disability benefits, and which satisfies:
$$\tilde{u}(x_w) - \bar{\gamma}_d \equiv u(x_d) \tag{1.38}$$

2. Conditional on applying for the disability benefit (i.e. $\gamma > \bar{\gamma}_d$) and being rejected it, an individual will work if $\bar{\gamma}_d \leq \gamma \leq \bar{\gamma}_r$, where $\bar{\gamma}_r$ satisfies:
$$\tilde{u}(x_w) - \bar{\gamma}_r \equiv u(x_r) \tag{1.39}$$

3. Conditional on applying for the disability benefit and being rejected it, an individual will choose to retire if $\bar{\gamma}_r \leq \gamma$.

Given a *fixed* budget of size β, the government chooses the consumption levels x_w, x_d

and x_r so as to maximize the utilitarian social welfare function:

$$\max_{x_w,x_d,x_r} W = \underbrace{\int_0^{\bar{\gamma}_d}[\tilde{u}(x_w)-\gamma]dF(\gamma)}_{\text{Those who always work}} + \underbrace{\int_{\bar{\gamma}_d}^{\bar{\gamma}_r}[\rho(\gamma)u(x_d)+(1-\rho(\gamma))(\tilde{u}(x_w)-\gamma)]dF(\gamma)}_{\text{Those who apply for the disability benefit but work if rejected it}}$$
$$+ \underbrace{\int_{\bar{\gamma}_R}^{\infty}[\rho(\gamma)u(x_d)+(1-\rho(\gamma))u(x_r)]dF(\gamma)}_{\text{Those who apply for the disability benefit and are retired}}$$

$$\text{s.t.} \quad \beta = \int_0^{\bar{\gamma}_d}[x_w-1]dF(\gamma) + \int_{\bar{\gamma}_d}^{\bar{\gamma}_r}[\rho(\gamma)x_w+(1-\rho(\gamma))(x_w-1)]dF(\gamma)$$
$$+ \int_{\bar{\gamma}_r}^{\infty}[\rho(\gamma)x_d+(1-\rho(\gamma))x_r]dF(\gamma) \tag{1.40}$$

Diamond and Sheshinski (1995) demonstrate that, provided $\tilde{u}'(x_w) < u'(x_d) \leq u'(x_r)$, a sufficient condition to have $\hat{x}_d > \hat{x}_r$ is that $\rho'(\gamma) > 0$, as assumed. In choosing the optimal benefit levels there is a tradeoff between (i) the reduction in labour supply (in the extensive margin) that a marginal increase in benefits brings; and (ii) the aim to efficiently allocate consumption across individual types. The optimal consumption levels are chosen so as to balance the labour supply reduction associated with an increase in benefits. Were labour disutility perfectly observable then it would be optimal to provide a single benefit which equates the marginal utility of workers with that of non-workers. However, because labour disutility is in fact imperfectly observed, the optimal disability and retirement benefits are optimally set at lower levels to encourage labour supply.

Within this literature, a number of papers alternatively model disability as a (0,1) phenomenon and impose a zero quantity constraint on labour supply for the disabled. When the benefit authority makes two-sided classification errors in tagging individuals as disabled, Parsons (1996) shows, under a number of specific assumptions, that a 'dual negative income tax' system providing incentives for tagged able individuals to work is optimal. Individuals are *ex-ante* identical but face a probability $\theta \in (0,1)$ of being disabled and unable to work. Contrastingly, with probability $(1-\theta)$ an individual is able to work. Any able individual who works has a marginal product of unity. Conditional on ability status, the probability that an individual is tagged is given by:

$$\begin{aligned}\text{Prob. (Tagged when Disabled)} &= (1-p_I)\ ;\ 0 \le p_I \le 1 \\ \text{Prob. (Tagged when Able)} &= p_{II}\ ;\ 0 \le p_{II} \le 1\end{aligned} \tag{1.41}$$

A Type I error (false rejection) is thus made with probability p_I, whilst a Type II error (false award) is made with probability p_{II}. The assumption is made that $(1 - p_I) > p_{II} \Leftrightarrow p_I + p_{II} \le 1$ and thus that the test awards the disability benefit to a disabled individual with at least as high a probability as it awards it to a truly able individual.

Individual preferences are given by $u(x) - \zeta \cdot \gamma$, where $u' > 0$; $u'' < 0$; x is consumption; γ is labour disutility; and $\zeta \in \{0, 1\}$ is a binary variable taking the value 1 if the individual works. By definition, $\zeta = 0$ for a disabled individual.

A 'dual negative income tax' system is characterised by the consumption levels x_o^t , x_o , x_w^t and x_w, where the subscripts w and o capture the working and non-working states, whilst the superscript t, or lack thereof, indicates whether or not an individual is tagged. Under the assumption that all able individuals do in fact work, the government chooses consumption levels so as to maximise a utilitarian social welfare function (which in this case is equivalent to an individual's expected utility):

$$\begin{aligned}\max_{x_n^t, x_n, x_w^t, x_w} W &= (1-\theta)\left\{(1-p_{II})[u(x_w)-\gamma] + p_{II}[u(x_w^t)-\gamma]\right\} \\ &\quad + \theta\left\{p_I u(x_o) + (1-p_I)u(x_o^t)\right\} \\ \text{s.t.}\quad (1-\theta) &= \theta\left\{(1-p_{II})x_w + p_{II}x_w^t\right\} + (1-\theta)\left\{p_I x_o + (1-p_I)x_o^t\right\}, \\ u(x_o) &\le u(x_w) - \gamma \quad \text{(Incentive compatibility constraint for } \textit{untagged} \text{ able)}, \\ u(x_o^t) &\le u(x_w^t) - \gamma \quad \text{(Incentive compatibility constraint for } \textit{tagged} \text{ able)}.\end{aligned} \tag{1.42}$$

Note that resources in the economy are given by $(1 - \theta)$ because all individuals who work produce a marginal product of unity. The analysis in Parsons proceeds as follows.

It is first shown that the 'dual negative income tax system' $\{x_o^t, x_o, x_w^t, x_w\}$ must be welfare superior to the single negative income tax system $\{x_o, x_w\}$ with no tagging because the latter is a available under the former (i.e. through setting $x_o^t = x_o$ and $x_w^t = x_w$), but is not chosen whenever $p_I + p_{II} < 1$ - and thus whenever tagging brings some useful information. Second, a dual negative income tax system can be welfare improving over a 'three price' system $\{x_o, x_o^t, x_w\}$ if tagged able individuals - who are voluntarily unemployed under the three price system because $x_o < x_w < x_o^t$ - can be induced to work for less than their marginal product. Formally then, there are gains to be had from a dual negative income tax system relative to a three price system if $u(x_o^t + 1) - \gamma > u(x_o^t)$.

In a more general exposition than Parsons (1996) - and with the additional restriction that tagged individuals are *not* allowed to work - Salanié (2002) shows that a higher benefit (demogrant) should be awarded to those who are tagged relative to those who are not. Once more, the proportion θ (resp. $(1-\theta)$) of the population are unable (resp. able) to work, and the terms p_I and p_{II} denote the propensity of the test awarding the disability benefit to make Type I and Type II errors respectively, where $p_I + p_{II} \leq 1$. However, able individuals now differ over a continuum of productivities, $n \geq 0$, with distribution function $F(n)$. All individuals who are not tagged as disabled receive a demogrant B and are taxed $T(nH)$ on any earnings, where the function $T(\cdot)$ may be nonlinear and is treated as exogenous and fixed. Meanwhile, any individual who is tagged as unable to work is (i) not allowed to work; and (ii) receives the tagged demogrant B^t.

The disabled have preferences $u^d(x, l) = u^d(M, 1)$ over consumption, $x \geq 0$, and leisure, $l \in [0, 1]$, where $M \in \{B, B^t\}$ depending on whether the individual is tagged or not. Meanwhile, the able have utility:

$$v^a(n, M) \equiv \begin{cases} u^a[nH^* - T(nH^*) + B, 1] & \text{if untagged} \\ u^a(B^t, 1) & \text{it tagged} \end{cases} \tag{1.43}$$

The social utility of an unable individual is given by $s^d(M) = g^d[u^d(M, 1)]$, whilst that of an able individual is given by $s^a(n, M) = g^a[v^a(n, M)]$, where the function $g^i; \quad i \in \{a, d\}$ satisfies $(g^i)'(\cdot) > 0$ and $(g^i)''(\cdot) < 0$. To proceed in the analysis, Salanié (2002) makes the following *standard* assumptions:

$$
\begin{aligned}
T'(nH) &\geq 0 && \text{(A1)} \\
H^*_M(n, M) &\leq 0 && \text{(A2)} \\
s^a_{nM}(n, M) &\leq 0 && \text{(A3)} \\
s^d_{MM}(M) &\leq 0 && \text{(A4a)} \\
s^a_{MM}(M) &\leq 0 && \text{(A4b)} \\
s^a_M(0, M) &\leq s^d_M(M) \ \forall M && \text{(A5)}
\end{aligned}
\tag{1.44}
$$

Assumption 1 (i.e. A1) simply states that the marginal tax rate is non-negative; A2 is the normality of leisure; A3 states that social marginal utility is decreasing in productivity; A4 states that social utility is concave in M; whilst A5 assumes that the social marginal utility of an able individual with zero productivity is no higher than that of an unable individual. Putting this all together, the government's objective is to choose B and B^t so as to:

$$
\begin{aligned}
\max_{B,B^t} W = \theta \left\{(1-p_I)s^d(B^t) + p_I s^d(B)\right\} \\
+ (1-\theta)\left\{p_{II}s^a(0, B^t) + (1-p_{II})\int_0^\infty s^a(n, B)dF(n)\right\} \qquad (1.45) \\
\text{s.t.} \quad B\chi + B^t(1-\chi)t \leq (1-\theta)(1-p_{II})\int_0^\infty tnH^*(n, M)dF(n) - R
\end{aligned}
$$

where $\chi = \theta p_I + (1-\theta)(1-p_{II})$.

Under assumptions A1 - A5, Salanié (2002) demonstrates that $B \leq B^t$ (i.e. tagged individuals should receive a higher benefit). Given that tagged individuals are not allowed to work, the *unusual* feature of this model is the absence of application decisions. Indeed, it can be seen from the second line of (1.45) that the proportion p_{II} of *all* able individuals are tagged. This must however, be incentive *incompatible*: all those with productivity exceeding some critical level would rather receive the lower benefit B and be *allowed* to work, as opposed to receiving the higher benefit B^t with the imposed zero quantity constraint on labour supply.

In the Salanié (2002) analysis the tax schedule is exogenous and consequently no focus is placed on how the propensity to make classification errors may affect the optimal tax schedule (and the benefit levels). To gain some insights into this question we turn to Stern (1982), who compares (i) imperfect lump-sum taxation and linear income taxation; with (ii) optimal nonlinear income taxation. For the purpose of this discussion, we are concerned with the first of these. There are two ability types in the economy: unskilled and skilled. Let n_l denote the gross wage of an unskilled individual, whilst $n_h > n_l$ denotes the gross wage of a skilled individual. These wages are endogenous to the model and are determined as the marginal product of a given production function. Both types of individual have identical preferences $u(x, H)$ over consumption and labour. In terms of the population, there are θ unskilled individuals and $(2-\theta)$ skilled individuals.[18]

A lump-sum tax system provides a tax/transfer $C_i; i \in \{l, h\}$ to each individual that is conditioned on ability type. Earned income is taxed at the linear (flat) rate $t \in (0, 1)$. The consumption of an individual of type i is therefore given by $x_i = (1-t)n_i H_i + C_i$. Let H_i^* and v denote optimal labour supply and the resulting indirect utility function, respectively. Importantly, the lump sum grants are administered with classification errors such that (i) an unskilled individual is incorrectly classified as skilled with probability p_l; whilst (ii) a skilled individual is incorrectly classified as unskilled with probability p_h.

Given these classification error propensities, it is straightforward to see that the average labour supply of the unskilled and skilled ability types is given by:

$$\begin{aligned} \bar{H}_l &= (1-p_l)H_l^*[(n_l(1-t), C_l] + p_l H_l^*[n_l(1-t), C_H] \\ \bar{H}_h &= (1-p_h)H_h^*[(n_h(1-t), C_h] + p_h H_h^*[n_h(1-t), C_l] \end{aligned} \tag{1.46}$$

The wages n_l and n_h are driven endogenously by the production function $Y = F[\theta \bar{H}^l, (2-\theta)\bar{H}^h]$, where $n_l = F_1$ and $n_h = F_2$:

Putting this all together, the government chooses the parameters t and C_l so as to:

[18] For much of the analysis Stern (1982) restricts attention to the case where $\theta = 1$ such that there are any equal number of low and high ability individuals.

$$\max_{t,C_l} W = \theta \left\{(1-p_l)v^{\rho}[n_l(1-t),C_l] + p_l v^{\rho}[n_l(1-t),C_h]\right\} + (2-\theta)\left\{(1-p_h)v^{\rho}[n_h(1-t),C_h] + p_h v^{\rho}[n_h(1-t),C_l]\right\} \tag{1.47}$$

s.t. $\chi_l C_l + \chi_h C_h = tY - R$

where $\chi_l = [(2-\theta)(1-p_l) + \theta p_h]$ denotes the number of individuals classified as unskilled; whilst $\chi_h = [(2-\theta)p_l + \theta(1-p_h)]$ denotes the fraction classified as skilled. The parameter ρ captures the government's concern for equity (i.e. a concern for equity beyond that of deceasing marginal utility of income).[19] The term R is an exogenous revenue requirement.

Adopting Constant Elasticity of Substitution (CES) preferences and a Cobb-Douglas production function, Stern numerically simulates the optima (t, C_l, C_h) under various parameter assumptions. With no classification errors it may be optimal to impose a lump-sum tax on skilled individuals to finance transfers to the unskilled, with the optimum income tax rate set at zero. So in this case $C_l = -C_h > 0$. However, a positive propensity to misclassify the unskilled and skilled exposes the former to the risk of receiving a smaller and even negative transfer. This is welfare reducing and, consequently, the linear income tax may rise to ensure that transfers to the skilled are positive. In summary, the simulations illustrate that the optimal linear tax rate tends to increase with the propensity to make either error type.[20]

Endogenous Awards Technology Accuracy

A number of papers explore an endogenous eligibility test (monitoring technology) that makes two-sided classification errors. First, Kleven and Kopczuk (2011) explicitly

[19] So if $\rho = 1$ we have the standard utilitarian case whereby the concern for equity arises solely from individual risk aversion (concavity).

[20] In terms of the comparison between (i) imperfect lump sum taxation and (ii) nonlinear taxation, the numerical results in Stern (1982) illustrate that the critical error propensity at which lump-sum taxation and optimum income taxation become equally desirable is highly dependent on the distributional values that the government has, as embodied in ρ. In particular, the greater the concern the government exhibits for equality, the lower the size of classification errors that will be tolerated before optimum taxation becomes preferable to an imperfect lump-sum system. In general, whenever classification errors are sufficiently small, it is optimal to rely on lump-sum taxation.

model the eligibility test adopted by the benefit authority and capture the interplay between screening intensity - or test complexity - and benefit applications. Whilst screening intensity improves the estimation of individual ability, it also imposes a cognitive cost on applicants. This cognitive cost may deter some truly eligible individuals from applying for a given benefit. The paper therefore goes beyond the standard definitions of Type I (false rejection) and Type II (false award) errors around a given eligibility threshold - both of which are conditional on applying for a benefit - to also include 'Type Ia' errors arising from incomplete take-up by eligibles.

Individuals differ in both (i) their innate ability (n); and (ii) the 'precision' with which this innate ability can be measured (σ). Any individual who applies for a benefit B is subject to tests which provide the unbiased *estimate* of their true ability $n^e = n + \epsilon/\pi$; where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ and π is the screening intensity (e.g. number of tests). Given an eligibility threshold of $\tilde{n}$, an applicant will only be awarded the benefit B if $n^e \leq \tilde{n}$; and thus if $\epsilon < (\tilde{n} - n)\pi$. Let ϵ/σ be distributed with distribution function $A(\epsilon/\sigma)$. The probability that an individual with characteristics n and σ will, conditional on applying, be awarded benefits is therefore:

$$\text{Award Probability} = A\left\{\frac{(\tilde{n} - n)\pi}{\sigma}\right\} \tag{1.48}$$

Individual preferences over consumption (x) and application costs ($K(\pi)$) are given by, $u[x - \mu K(\pi)]$, where $u' > 0$; $u'' \leq 0$; $K' > 0$; and $\mu \in \{0, 1\}$ denotes whether an application has been made. Application costs are therefore an increasing function of complexity. An individual will thus apply for the benefit B if;

$$A\left\{\frac{(\tilde{n} - n)\pi}{\sigma}\right\} \cdot \{u[n + B - K(\pi)] - u[n - K(\pi)]\} \geq u(n) - u[n - K(\pi)]$$

and therefore if:

$$A\left\{\frac{(\tilde{n} - n)\pi}{\sigma}\right\} > \bar{A}\left\{\frac{(\tilde{n} - n)\pi}{\sigma}\right\} \equiv \frac{u(n) - u[n - K(\pi)]}{u[n + B - K(\pi)] - u[n - K(\pi)]} \tag{1.49}$$

To simplify the analysis, Kleven and Kopczuk (2011) impose constant absolute risk aversion (CARA) preferences on the utility function. This eliminates the dependence of the reservation probability, $\bar{A}$, on individual ability. The function $\bar{A}$ therefore depends only on σ and π. Accordingly, the threshold precision level , $\bar{\sigma}_n$ at which $A = \bar{A}$ (i.e. at which the award probability renders an individual indifferent between applying or not) is given by[21]:

$$\bar{\sigma}_n = \frac{(\tilde{n} - n)\pi}{A^{-1}\bar{A}} \tag{1.50}$$

Turning to the population, there are two ability types: low (n_l) and high (n_h). Both ability groups has a distribution of precision parameters, $\sigma \in [0, \infty)$, as given by $F_i(\sigma) \ \ ; i = \in \{l, h\}$. Now, given an exogenously *fixed* budget size, β, and an income maintenance objective of providing a minimum benefit $\underline{B}$ to those of low ability, the government chooses (i) the screening intensity; (ii) the eligibility threshold; and (iii) the benefit level parameters so as to:

$$\begin{aligned}
&\max_{\pi, \bar{n}, B} \int_0^{\bar{\sigma}_l} A\left[\frac{(n^e - n)\pi}{\sigma}\right] dF_l(\sigma) \\
&\text{s.t.} \quad \left\{\int_0^{\bar{\sigma}_l} A\left[\frac{(n^e - n)\pi}{\sigma}\right] dF_l(\sigma) + \int_0^{\bar{\sigma}_h} A\left[\frac{(n^e - n)\pi}{\sigma}\right] dF_l(\sigma) dF_h(\sigma)\right\} B \leq \beta \ , \\
&\qquad B \geq \underline{B}
\end{aligned} \tag{1.51}$$

Kleven and Kopczuk (2011) go on to show that (i) optimal targeting programmes feature high screening intensity for the purposes of restricting the conventional Type I and Type II errors *and* (ii) to go some way to combatting the negative effect of high screening intensity on take-up by eligibles, the government can use its two remaining instruments.

Jacquet (2014) models the optimal provision of disability and income-based (welfare) benefits when the monitoring technology administering the disability benefit is costly and classification errors are a function of the resources dedicated to the monitoring technology. Individual applications are, however, taken to be costless. Individuals differ along the two dimensions of productivity at work and labour disutility. First, an

[21] Setting $A = \bar{A}$ gives $\frac{1}{\sigma}(n^e - n)\pi = A^{-1}\bar{A} \Rightarrow \bar{\sigma}_n = \frac{(n^e - n)\pi}{A^{-1}\bar{A}}$.

individual may be disabled or able bodied and disability status is perfectly correlated with productivity. That is, a disabled individual commands a low productivity, n_l, whilst an able individual commands a high productivity, n_h, where of course $n_h > n_l > 0$. Whilst a disabled person can only engage in low productivity work, an able person may choose to engage in either low or high productivity work[22]. Turning to the second dimension of heterogeneity, individuals differ continuously in the disutility they experience at work - where work is modelled in the extensive margin. Notice that this approach to modelling disability is similar to that in Diamond and Sheshinski (1995).

However, the source of this disutility depends on whether an individual is disabled or able. The disutility a disabled individual endures working - denoted by γ_d - arises due to physical or mental pain. Contrastingly, the disutility an able individual experiences when working - as denoted by γ_a - is due to distaste for work. These parameters are distributed with density functions f_d and f_a and corresponding distribution functions F_d and F_a respectively. The benefit authority is taken to know these distributions but cannot observe an individual's γ. Finally, individual utility takes the quasilinear form $u(x) - \gamma$, where x denotes consumption and $u'(x) > 0,\ u''(x) \leq 0$.

The awards technology administering the disability benefit makes Type I errors with probability $p_I(K)$ and Type II errors with probability $p_{II}(K)$, where $K(p_I, p_{II})$ is the cost per applicant of running the test. It is assumed that $\partial K/\partial p_I < 0$ and $\partial K/\partial p_{II} < 0$. Intuitively, this captures the idea that a reduction in the resources devoted to the test lower its accuracy . Let $x_h \geq x_l > x_d \geq x_w$ denote consumption levels when (i) working in a high productivity job; (ii) working in a low productivity job; (iii) receiving disability benefits; and (iv) receiving welfare benefits, respectively. Applications for the disability benefit are costless. It is therefore straightforward to see that application decisions are driven by labour disutilities and are independent of error propensities. Specifically, a disabled individual with disutility γ_d who has been rejected the disability benefit will only work if $\gamma \leq \bar{\gamma}_d$, where:

$$u(x_l) - \bar{\gamma}_d \equiv u(x_w) \tag{1.52}$$

In terms of application decisions, a disabled individual will only apply for x_d in the

[22] In the model Jacquet (2014) demonstrates that it is never optimal for an able individual to choose low productivity employment.

first place if $\gamma > \bar{\bar{\gamma}}_d$, where:

$$u(x_l) - \bar{\bar{\gamma}}_d \equiv u(x_d) \tag{1.53}$$

where $\bar{\bar{\gamma}}_d \leq \bar{\gamma}_d$ because $u(x_d) \geq u(x_w) \Leftrightarrow x_d \geq x_w$. Analogous definitions of $\bar{\gamma}_a$ and $\bar{\bar{\gamma}}_a$ hold for able individiuals.

Jacquet (2014) adopts a 'paternal' objective function which gives no weight to the disutility of able individuals (i.e. that arising from distaste for leisure). Where $\theta \in (0,1)$ denotes the proportion of disabled individuals in the population, paternal welfare is:

$$\begin{aligned} W^P \\ = \theta & \left\{ \begin{array}{c} \underbrace{\int_0^{\bar{\bar{\gamma}}_d} [u(x_l) - \gamma_d] dF_d(\gamma_d)}_{\text{Non-take-up}} + p_I \left\langle \underbrace{\int_{\bar{\bar{\gamma}}_d}^{\bar{\gamma}_d} [u(x_l) - \gamma_d] dF_d(\gamma_d)}_{\text{Rejected disabled who work}} + \underbrace{\int_{\bar{\gamma}_d}^{\infty} u(x_w) dF_d(\gamma_d)}_{\text{Rejected disabled not working}} \right\rangle \\ + \underbrace{[1 - F(\bar{\bar{\gamma}}_d)] u(x_d)}_{\text{Disabled awarded the benefit}} \end{array} \right\} \\ + (1-\theta) & \left\{ \begin{array}{c} \underbrace{F_a(\bar{\bar{\gamma}}_a) u(x_h)}_{\text{Non Applicants}} + (1 - p_{II}) \left\langle \underbrace{[F_a(\bar{\gamma}_a) - F_a(\bar{\bar{\gamma}}_a)] u(x_h)}_{\text{Rejected able who work}} + \underbrace{[1 - F(\bar{\gamma}_a)] u(x_w)}_{\text{Rejected able not working}} \right\rangle \\ + \underbrace{p_{II}[1 - F_a(\bar{\bar{\gamma}}_a)] u(x_d)}_{\text{Able awarded the benefit}} \end{array} \right\} \end{aligned} \tag{1.54}$$

The problem is then to:

$$\begin{aligned} \max_{\{x_h, x_l, x_d, x_w, p_I, p_{II}\}} W^P \quad & \text{s.t.} [1 - F_d(\bar{\bar{\gamma}}_d)] \left\{ [(1 - p_I) x_d] + K(p_I, p_{II}) \right\} + (1 - F(\bar{\gamma}_d) x_l \\ & + [1 - F_a(\bar{\bar{\gamma}}_a)] \left\{ p_{II} x_d + K(p_I, p_{II}) \right\} + (1 - F(\bar{\gamma}_d)) x_w \\ & = \left\{ F_d(\bar{\bar{\gamma}}_d) + p_I [F_d(\bar{\gamma}_d) - F_d(\bar{\bar{\gamma}}_d)] \right\} (n_l - x_l) \\ & + \left\{ F_a(\bar{\bar{\gamma}}_a) + (1 - p_{II}) [F_a(\bar{\gamma}_a) - F_a(\bar{\bar{\gamma}}_a)] \right\} (n_h - x_h) \end{aligned} \tag{1.55}$$

The author demonstrates that optimum Type I errors balance the trade-off between (i) the increased tax revenue due to some disabled individuals working and lower resources devoted to monitoring; and (ii) the welfare loss associated with Type I errors. Optimum Type II errors, meanwhile, balance the trade-off between (i) the benefit that more accurate monitoring induces able individuals who would have previously received the benefit to work, thus increasing tax revenue; and (ii) the monitoring costs of reducing Type II errors (Jacquet, 2014).

In an alternative approach, Boadway et al. (1999) endogenise classification errors through the effort that social workers exert in applying a monitoring technology (awards test) on applicants. The population consists of (i) N high ability individuals, (ii) N low ability individuals and (iii) N disabled individuals, where the latter are unable to work. The high ability types pay taxes, whilst the low ability and disabled types receive transfers of some fashion. For high and low ability individuals preferences are described by $u^i(x, y)$; $i \in \{l, h\}$; $u_x > 0$, $u_y < 0$ where x is consumption and y earnings. Regarding earnings, there is a minimal earning requirement of y^{min} which may or may not bind. The disabled, meanwhile, have preferences $u^d(x)$; $(u^d)'(x) > 0$, $(u^d)''(x) < 0$ over consumption. The government chooses (x, y) bundles for individuals and may operate a two-tier welfare system composed of a general welfare benefit - where recipients may or may not be allowed to work - and a targeted welfare benefit for the disabled. In either case the high ability are offered a bundle (x_h, y_h) and those tagged as disabled receive a bundle $(x^t, 0)$. Under a system where general welfare recipients are allowed to work (system A), both the bundles (x_l, y_l) and $(x_d, 0)$ are offered. Untagged low ability individuals choose between these, whilst untagged disabled must receive the second bundle. For their to be incentive compatability, the bundles must satisfy $u^h(x_h, y_h) \geq u^l(x_l, y_l)$ (to induce the high ability to work) and $u^l(x_l, y_l) \geq u^l(x_d, 0)$ (to induce the low ability to work). Contrastingly, under a system where general welfare recipients are not allowed to work (system B), only the bundle $(x_d, 0)$ is offered. Under this system the only incentive compatibility constraint to satisfy is $u^h(x_h, y_h) \geq u^h(x_d, 0)$.

Regarding the awards test, let p_i denote the probability that an individual of type i is tagged, where $p_h = 0$ and $0 \leq p_l \leq 1/2 \leq p_d \leq 1$, where the latter two propensities will depend on the effort exerted by social workers. Boadway et al. (1999) further assume that:

$$p \equiv \underbrace{(1 - p_d)}_{\text{Pr(Type I error)}} = \underbrace{p_l}_{\text{Pr(Type II error)}} \tag{1.56}$$

So the term p captures the effectiveness of the awards process - where $p = 1/2$ corresponds to the case where a disabled individual is tagged with the same probability as a low skilled individual. This arises when the social worker exerts no effort. Social workers are risk neutral and paid n_s per client, which may be a low skilled or disabled individual with equal probability.[23] However, if the social worker is found to have made an award error (of either Type I or Type II), an absolute fine of n_s is invoked upon them. There are two ways in which errors may be detected. First, under system A where recipients are allowed to work, any untagged welfare recipients who are not working must be disabled and untagged by Type I error. So the identification of Type I errors is costless under system A. Second, the government may audit welfare recipients with probability e and, in doing so, ascertain their true ability. Auditing, however, is costly. Tagging therefore induces administrative costs, both in terms of paying social workers to administer the awards test and in terms of auditing welfare recipients. The costs of targeting a benefit may outweigh the benefits of discerning the disabled from the low skilled, and lead the government to instead operate simple negative income tax system.

Finally, the government wishes to maximise the welfare of the disabled, and so adopts the social welfare function:

$$W = p_l u^d(x^t, 0) + (1 - p_d)u^d(x_d, 0) \tag{1.57}$$

Briefly, a number of conclusions emerge from the Boadway et al. (1999) analysis. First, if social worker effort is observable and the minimum income constraint is non-binding, system A will always be preferable to system B because low ability individuals can be incentivised to work even with a positive marginal tax rate. Second, if the minimum income constraint is binding it may be difficult to incentivise untagged low ability individuals to work without resorting to a negative income tax scheme. Accordingly system B may be preferred to system A depending on the cost of monitoring social workers. Third, if the earnings constraint is not binding but effort is not observable, the balance of considerations becomes more complicated. On the one hand, system A may allow Type I errors to be identified through separation, but this implies that social workers should simply tag everyone. It may therefore be optimal to adopt system B.

[23] Because there are N low income and N disabled individuals who would like to be tagged.

1.2 Individual Behaviour and Risk-Taking.

A small literature explores the fraudulent receipt of welfare benefits when individuals risk being detected and fined. This literature predominantly focuses on unemployment insurance schemes and the specific costs that benefit conditionality places on recipients.

The most notable contribution in this literature is that of Yaniv (1986), who considers the decision of employed individuals to claim unemployment insurance benefits for a duration of their time endowment. Individual preferences over consumption (x) are given by $u(x)$; where $u' > 0$ and $u'' < 0$. Further, for $r(x) = -u''(x)/u'(x)$ preferences satisfy $r'(x) < 0$ and thus exhibit decreasing absolute risk aversion. Let ω denote an individual's wage rate and D their full time allocation in days, giving rise to an earned income of ωD. The worker may choose to fraudulently claim an unemployment insurance benefit of $B < \omega$ per day, by *sacrificing* a proportion k of each claiming day to labour exchange duties. The model does not feature voluntary unemployment, such that any departure from providing labour for the full time endowment arises only due to the decision to claim unemployment insurance and the resulting labour exchange commitments. There is also a waiting period of q days before benefits can be received such that, for z days of claiming, benefits are only received for $z - q$ days.

Of course, claiming unemployment insurance for days spent working comes with the risk of being detected and fined. Let ρ denote the detection probability and F_i the fine imposed if detected. In this simple setting, Yaniv (1986) explores two alternative fines from the tax evasion literature: the first is proportional to the duration of fraudulent claims but independent of the benefit size (à la Allingham and Sandmo, 1972); whilst the second is proportional to the total benefit income fraudulently claimed (à la Yitzhaki, 1974). Formally;

$$F_i = \begin{cases} \pi z \; ; \; \pi > B & : i = 1 \\ \phi B(z-q) \; ; \; \phi > 1 & : i = 2 \end{cases} \tag{1.58}$$

There are two states of the world: detected and not detected. Consumption when not detected is given by $x^N = \omega(D - kz) + B(z - q)$. Meanwhile, consumption if detected will depend on the fine structure in place and is given by:

$$x^{Di} = \begin{cases} \omega D - Bq + z(B - \omega k - \pi) & : \text{if } i = 1 \\ \omega D - Bq(1-\phi) + z[B(1-\phi) - \omega k] & : \text{if } i = 2 \end{cases} \tag{1.59}$$

The individual problem is given by:

$$\max_{z} \rho u(x^{Di}) + (1-\rho)u(x^N) \;\; ; \; i \in \{1, 2\} \tag{1.60}$$

Assuming that the conditions for an interior solution ($z > 0$) are satisfied under both penalty structures, the optimal number of claiming days under the respective schemes - denoted by z_1 and z_2 - are characterised by:

$$\begin{gathered} \rho u'(x^{D1})[B - \omega k - \pi] + (1-\rho)u'(x^N)(B - \omega k) = 0 \\ \rho u'(x^{D2})[B(1-\phi) - \omega k] + (1-\rho)u'(x^N)(B - \omega k) = 0 \end{gathered} \tag{1.61}$$

Through comparative statics exercises on these optimality conditions, Yaniv (1986) generates the following results. First, dz_i/dB differs with the fine structure due to differences in income and substitution effects. Unsurprisingly, we have $dz_1/dB > 0$ because (i) the expected fine remains unchanged such that the gains to fraud increase with the benefit (positive substitution effect); and (ii) because preferences exhibit decreasing absolute risk aversion the greater expected income induces more fraud (via a positive income effect). Contrastingly, the sign of dz_2/dB is ambiguous because the expected fine is increasing with the benefit level, thus generating income and substitution effects of opposing signs. Second, the sign of $dz_i/d\omega$ is ambiguous under both fine structures because an increase in the wage increases the opportunity cost of labour exchange time (negative substitution effect), but encourages risk taking (positive income effect)[24]. Turning to the effect of the enforcement parameters on the extent of fraud, we have $dz_i/d\rho < 0 \; \forall \; i$, $dz_1/d\pi < 0$, $dz_2/d\phi < 0$ and $dz_i/dk < 0$. However, the affect

[24] Yaniv (1986) shows how $dz_i/d\omega$ would change if the unemployment insurance benefit was proportional to earnings, and thus given by $B = \beta\omega D \;\; ; \; 0 < \beta < 1$. Under fine structure F_1, an increase in the wage increases in the benefit level but has no affect on the penalty, thereby generating only a positive substitution effect. Contrastingly, under fine structure F_2, an increase in the wage increase both the benefit level and the fine level.

of an increase in the waiting period differs between the two schemes, with $dz_1/dq < 0$ and $dz_2/dq > 0$. Under the second fine structure, an increase in q serves to, *ceteris paribus*, reduce the expected penalty which generates a substitution effect which more than offsets the income effect.

Next, Wolf and Greenberg (1986) analyse the binary choice of a benefit recipient to either (i) fully report or (ii) not report any earned income to the welfare authority, given an earnings stream of n per month for D months and zero thereafter. Individuals know that the formula adopted by the benefit authority for determining monthly welfare payments is $\max(0, B - tn)$, where B is the maximum monthly benefit and t is the 'tax' rate on earnings. So notice that if $B = tn$ then the individual has no welfare entitlement for the D months each earning n. Individual utility is linear in consumption and, for those who do not report income, includes a linear disutility/stigma term, $\gamma > 0$, which enters negatively. Further, an individual who does not report income faces a constant probability ρ of being detected in any period j and being fined at a rate ϕ_j proportional to the welfare income fraudulently obtained, as given by $j \min(B, tn)$. Whilst it is beyond the scope of this discussion to go into further detail, note that ϕ_j will depend on (a) whether the benefit income incorrectly obtained exceeds a fraud prosecution threshold; and (b) what the recipient's true entitlement is.

Putting this all together, should an individual fully comply ($\mathcal{C}$) with earnings reporting then, given a monthly discount rate of δ_j, their present discounted value of utility is simply given by:

$$u^{\mathcal{C}} = \delta_1[n + \max(0, B - tn)] + ... + \delta_D[n + \max(0, B - tn) + \delta_{D+1}B + ... \tag{1.62}$$

Contrastingly, should non-compliance ($\mathcal{N}$) be chosen the present discounted value of utility takes into account that an individual may be detected in any period and, once detected, receives only their true entitlement thereafter. Formally, we have:

$$u^{\mathcal{N}} = \rho \left\{ \begin{array}{c} \delta_1[n + B - \phi_1 \min(B, tn) - \gamma] + \delta_2[n + \max(0, B - tn)] + \\ ... + \delta_D[n + \max(0, B - tn)] + d_{D+1}B + .. \end{array} \right\}$$

$$+(1-\rho)\rho \left\{ \begin{array}{c} \delta_1[n + B - \gamma] + \delta_2[n + B - \phi_2 \min(B, tn) - \gamma] + \delta_3[n + \max(0, B - tn)] \\ +... + \delta_D[n + \max(0, B - tn)] + d_{D+1}B + ... \end{array} \right\}$$

$$\vdots$$

$$+ (1-\rho)^{D-1}\rho \left\{ \begin{array}{c} \delta_1[n + B - \gamma] + ... + \delta_{D-1}[n + B - \gamma] + \\ \delta_D[n + B - \phi_D \min(B, tn) - \gamma] + \delta_{D+1}B + .. \end{array} \right\}$$

$$+ (1-\rho)^D \{\delta_1[n + B - \gamma] + ... + \delta_D[n + B - \gamma] + \delta_{D+1}B + ...\} \quad (1.63)$$

Note that the welfare recipient makes the compliance decision once at the start of earnings stream. Non-compliance will thus only be chosen over full compliance if $u^{\mathcal{N}} > u^{\mathcal{C}}$. Unsurprisingly, the incentive to commit fraud is decreasing in the penalty imposed if detected, ϕ_j, the detection probability, ρ, and on the disutility, γ, an individual endures through committing fraud.

In the literature discussed so far the fraudulent activity involves either working or not reporting earnings. However, given that benefits benefits may be conditioned on spending time at the benefit office and/or providing evidence of job search, there is another compliance decision to be made. Whilst Yaniv (1986) assumes that individuals fully comply with this dimension of *ex-post* conditionality, other authors do not. We proceed to discuss this below.

Burgess (1992) develops and empirically tests a simple theoretical model capturing the eligibility compliance decision of unemployment insurance recipients, where maintained benefit receipt is conditional on job-search and filing requirements. Risk-neutral recipients make their compliance decision by comparing the net gains of compliance ($\mathcal{C}$) with those of non-compliance (N). In both cases, the net gains can be decomposed into (i) benefit size, (ii) search and (iii) cost components. Compliant individuals receive a benefit of B per week with certainty, whilst non-compliant individuals face an expected benefit of $B(1-\rho\phi)$, where ρ is the detection probability and $(1-\phi)$ is the proportion of the weekly benefit payment that can be retained. The present discounted value of engaging in job search is given by $\delta^{\mathcal{C}}[F(s^{\mathcal{C}},e)]$ and $\delta^{\mathcal{N}}[F(s^{\mathcal{N}},e)]$, respectively, where $F(\cdot)$ denotes the distribution of wage offers given a chosen search intensity, $s^{\mathcal{C}}$or $s^{\mathcal{N}}$, and e is the employment state of the labour market. Unsurprisingly, $\delta^{\mathcal{C}}[f(s^{\mathcal{C}},e)] > \delta^{\mathcal{N}}[F(s^{\mathcal{N}},e)]$. Let an individual's weekly time opportunity costs be k. A compliant individual who spends the fraction $s^{\mathcal{C}}$ of the week searching for employment and the fraction a 'filing' at the benefit office has a weekly time cost $k(s^{\mathcal{C}}+a)$. Contrastingly, a non-compliant individual does not incur a and has weekly time cost $k(s^{\mathcal{N}})$.

Putting this all together, the probability that an individual will choose to comply is given by:

$$\text{Prob}\left\{\underbrace{B\rho\phi}_{\text{Benefit}} + \underbrace{\delta^{\mathcal{C}}[F(s^{\mathcal{C}},e)] - \delta^{\mathcal{N}}[F(s^{\mathcal{N}},e)]}_{\text{Search}} - \underbrace{k(a+s^{\mathcal{C}}-s^{\mathcal{N}})}_{\text{Cost}} > 0\right\}$$

Burgess (1992) notes that even when $\rho = 0$ compliance may still be chosen should the gains in the present discounted values of job search exceed the cost of compliance. It is straightforward to see that the returns to benefit compliance are increasing in ρ and ϕ (i.e. decreasing in $(1-\phi)$). Notice that the conditionality parameter a - corresponding to time spent filing - enters only through the cost term and plays no productive role through search. Accordingly, the tightening of this eligibility requirement actually serves to reduce compliance.

1.3 Empirical Studies: Classification Error Estimates and Labour Force Participation Disincentives.

In the models discussed in Section 1.1.3 it is assumed that the technology awarding benefits has a propensity to make classification errors of both Type I (false rejection) and Type II (false award). These error propensities are either exogenously given or endogenously determined by the resources dedicated to the monitoring technology (and the effort exerted by employees). A natural question to ask is therefore; are there estimates of error propensities in real-world welfare programmes? Further, how is the scope for these errors related to the design of welfare programmes?

Both the Social Security Disability Insurance (DI) and Supplemental Security Income (SSI) programmes in the U.S. provide rich case studies to answer these questions. Both benefits are administered by the U.S. Social Security Administration (SSA) and the disability criterion for awards is common to both programmes. Under the Social Security Act, an individual is considered disabled if their medically verifiable physical or mental condition prevents them from engaging in 'substantial gainful activity' (see SSA, 2012, p.2)[25]. Further, this condition should be expected to be terminal or last for no less than one year.

1.3.1 Estimating Classification Error Propensities

A number of studies have attempted to estimate the propensity of the SSA to make classification errors of both Type I and Type II in awarding both DI and SSI. The first is Nagi (1969), who used external audit data on the independent health assessment of DI applicants by professionals (such as doctors and psychologists) and compared these with the ultimate award decision of the SSA. Based on these comparisons, the author inferred a Type I error propensity of approximately 48% and a Type II error propensity of approximately 19%.

More recently, Benitez-Silva et al. (2004) treat (i) the self-declared disability status of DI/SSI applicants in the Health and Retirement Study; and (ii) the ultimate SSA awards decision (post any appeals), as noisy but unbiased indicators of true disability status. They estimate the Type I error propensity to be approximately 60%, whilst the Type II error propensity is estimated at approximately 22%. Their results are thus quantitatively similar to those in Nagi (1969).

1.3.2 Labour Force Participation Disincentives

A growing empirical literature focuses on the labour force participation (LFP) disincentives generated by both Social Security Disability Insurance (DI) and Supplemental

[25]Substantial gainful activity was in 2011 a monthly income of $1,000 for a non-blind individual (SSA, 2012).

Security Income (SSI) programmes in the United States. There are a number of motivating factors behind the research in this literature. First, legislation changes in 1984 required the SSA to place greater emphasis on an applicant's ability to function in the workplace and any pain experienced in doing so, as opposed to basing assessments solely on strict medical criteria. With regard to difficult to verify conditions such as mental illness and musculoskeletal disease (i.e. back pain), these changes are argued to correspond to a reduction in the screening intensity of applicants (Autor and Duggan, 2003, 2006; Von Wachter et al., 2011). Second, the earnings replacement rate[26] has also risen markedly since 1984. The concern is that both of these changes account for (i) the large increase in DI recipients following the 1984 legislation change; (ii) mental illness and musculoskeletal disease being the two most prevalent claimant categories (SSA, 2012); and (iii) the increased propensity of younger individuals to exit the labour force and claim benefits.

The impact of the DI programme on LFP is far from clear in the literature. The basis for this discussion stems from the method developed by Bound (1989), who uses the labour force participation of rejected DI applicants as a control, or upper bound, for that which could be expected from awarded applicants. The assumption is therefore made that the tests run by the benefit authority have some discriminatory power, and thus that rejected applicants are more capable of work than awarded applicants. Using data on the labour market performance of rejected applicants between 1972 and 1978, Bound shows that less than half of rejected prime age (45-64) male applicants provide labour for a sustained period.[27] He thus argues that the majority of DI recipients are in fact disabled.

Parsons (1991), however, argues that Bounds' analysis is flawed because it does not factor in the persistent role that the DI programme plays in the lives of applicants beyond an initial rejection. Owing in part to the difficulties in precisely defining the eligibility criterion, the programme features an extensive appeals system whereby a number of successive appeals through different bodies can be made prior to an ultimate rejection. Given the lags between an applicant filing an appeal and the response of the relevant authority, coupled with the numerous appeals that can be made, the appeals process is a lengthy one. Parsons notes three reasons - that are acknowledged,

[26] I.e. Disability income relative to earnings.

[27] Bound (1989) does not consider those below 45 because less than 20% of recipients in the studied period were younger than this age.

but given little weight in Bounds' analysis - as to why a rejected applicant may not return to work. First, rejected applicants may be filing an appeal. Second, they may be enduring a period of unemployment to strengthen a future reapplication. Third and finally, rejected applicants may face difficulty in finding employment due to their sustained period out of work.[28]

More recent studies provide little consensus. On the one had, the analysis of Chen and van der Klaauw (2008) points to relatively low LFP disincentives induced by the DI programme and thus supports Bound (1989). These authors adopt Bound's comparator group approach on a data-set which features a number of important differences from that used in Bound (1989). First, merged survey and administrative data from the 1990s is used, whilst Bound only used survey data which can be unreliable given potential misreporting. Second, data is not restricted to the DI programme, but also includes the SSI programme. Finally, both male and female applicants are included in the data set. The analysis suggests that, in absence of any disability benefit provision, the LFP of recipients would be at most twenty percent higher.

On the other hand, however, the most recent study of Von Wachter et al. (2011) finds substantial LFP for rejected younger applicants and rejected applicants who claimed for musculoskeletal diseases and/or mental health conditions. The authors use a database containing (i) DI administrative information on applications and awards (from 1981-1999); in addition to (ii) information on earnings pre- and post-applications (from 1978-2006). The results depend on the age group of applicants. Whilst the conclusions of Bound (1989) continue to hold for those in the older age category of 45-64; it is shown that younger rejected applicants aged 30-44 exhibit substantial post-rejection employment. As discussed in the opening of this section, younger applicants form a much larger proportion of total applicants in the data used by Von Wachter et al. (2011) than in that from the 70's for Bound (1989).

Finally, Autor and Duggan (2003) estimate that the combined effects of (i) the 1984

[28] In a reply, Bound (1991) readdresses each of these three issues. Concerning the appeals process, he notes that most rejected applicants in his data-set had applied at least 18 months prior to the survey, and very few appeals processes take this period of time. Turning to reapplications, it is not clear that enduring a sustained period of unemployment can enhance an application. In fact, it may have the opposite effect. Whilst some rejected applicants may behave this way, it is unlikely to play a large role. Finally, concerning the impact of processing lags lowering the employment prospects of rejected applicants, Bound notes that, on average, the applicants studied had been unemployed for seven months prior to making an application.

changes to disability assessment; (ii) increasing replacement rates; and (iii) a reduction in the demand for low-skilled labour, have served to double the propensity of high-school dropouts to exit the labour force over the period 1984-2001. The authors identify instrumental variables to capture exogenous variation in both the 'supply' and 'demand' for DI. On the supply side, it is noted that the benefit formula is a function of the average wage in the U.S. - which has risen relative to low-skilled wages - but does not account for state level differences in wages. As a consequence, the replacement rate is higher in some states than others. The supply effect of programme expansions and contractions may thus have different effects in different states. Next, to capture exogenous variation in the demand for DI, the authors project national industry employment changes onto state level industry composition. In summary, the analysis suggests that state level contractions in DI supply generate large increases in LFP among high-school dropouts; whilst DI applications are much more responsive to state level demand shocks following the 1984 legislation changes to eligibility.

1.4 Concluding Remarks

The purpose of this chapter has been to review some of the important contributions in the economics literature relating to cash welfare programmes. The emphasis has been largely placed on targeted categorical programmes and the scope for classification errors in administering these benefits. Along the way, a number of interesting questions have emerged.

First, the discussion in Section 1.1.2 illustrated the important contributions on categorical transfers in the optimal income tax framework. These analyses typically assume that categorical status can be perfectly observed.[29] In this case, a well-established result is that categorical transfers should be set so as to eliminate inequality in the average net social marginal value of income across categorical groups (Viard, 2001a,b). For the purpose of writing the optimal linear tax expression, it is assumed that categorical transfers do indeed eliminate this between-group inequality at the optimum. This allows the tax expression to be written as in unidimensional model where individuals differ only in productivity. However, is there a more general way to write the optimal

[29] Indeed, Immonen et al. (1998, p.181) state that '*individuals are unable to alter or disguise the group to which they belong, which is observed costlessly by the government*'.

linear tax tax expression to allow for cases where between-group inequality persists at the optimum? Under what cases is this likely to arise?

Second, the discussion in Section 1.1.3 illustrated that a number of the analyses of imperfectly targeted categorical transfers make different assumptions as to whether recipients of an incapacity benefit are 'allowed' to work, but the full implications of these differing assumptions are not directly comparable due to other differences in the modelling frameworks. In particular, able individuals who are incorrectly tagged as unable are allowed to work in Parsons (1996); but not allowed to work in Salanié (2002). These polar assumptions can be interpreted as no enforcement and full enforcement, respectively, of an *ex-post* 'no-work' condition. Whilst the framework in Salanié is more general (continuum of productivities, standard preferences) than in Parsons, it abstracts from individual application decisions. Instead, a fraction of the able subpopulation are simply tagged as unable and not allowed to work. This must, however, be incentive incompatible for higher productivity individuals who would rather not be tagged and allowed to work. With these points in mind, it would be useful to analyse imperfectly targeted categorical transfers in a framework that both endogenises application decisions and allows for a systematic comparison across enforcement regimes. We could then ask the following questions: (i) How do the optimal combinations of targeted (categorial) and non-targeted (universal) benefits differ across enforcement regimes? (ii) What are the welfare effects of classification errors and how do these differ across enforcement regimes?

Third, to the best of this author's knowledge there has been no analysis of imperfectly targeted categorical transfers in the optimal income tax framework where individuals differ in both some categorical dimension *and* over a continuum of productivities. Indeed, we saw in Section 1.1.2 that the analysis of categorical transfers in the optimal linear income tax framework assumes perfect targeting. How would Type I and Type II classification errors affect the equity and efficiency considerations in the optimal tax expression? Suppose that there is an unable subpopulation who cannot work and an able subpopulation of individuals who can work and differ continuously in productivity. Further, suppose that the categorical benefit targeted at the unable has a fully enforced 'no-work' condition *and* individuals must choose to apply for the categorical benefit. In this setting the additional efficiency considerations that we need to account for seem quite clear: an increase in the tax rate may incentivise able individuals to apply for the categorical benefit and, if awarded it, stop working.

Part I

Design Issues: Optimal Benefits and Taxation.

Chapter 2

Optimal Universal and Categorical Benefit Provision:

Classification Errors and Imperfect Enforcement.[1]

2.1 Introduction

Partial universal welfare programmes can be defined as those which (i) provide all members of society with an unconditional universal cash benefit; but also (ii) allow for additional, *targeted*, assistance to those considered by the policymaker to be most in need. The concept of a universal benefit has a rich history and is known under various alternative guises in the economic and social policy literature, such as 'demogrant' or 'basic income' (Van Parijs, 2004). The state of Alaska provides the primary example of a real-world universal benefit programme.[2] Targeted benefits take centre stage in modern welfare systems. Of these, categorical transfers (e.g. unemployment benefits, disability benefits) play a prominent role. The optimal balance between these two types of support (universal, targeted) is likely to depend on the effectiveness of targeted

[1] This chapter represents a significant extension that I have made to the discussion paper of Slack and Ulph (2014).

[2] In place since 1982, all residents of Alaska receive the same yearly cash benefit, whose size is a function of the five year interest on the Alaska Permanent Fund. This fund was set-up in 1976 using oil revenue from the Trans-Alaska Pipeline System.

benefits at reaching those in need. In particular, the less that needy individuals are assisted by targeted benefits, the more important universal provision may become in ensuring those in need in have some form of financial support. Both types of transfers have been extensively analysed in the economics literature (Atkinson and Sutherland, 1989; Atkinson, 1995; Callan et al., 1999; Diamond and Sheshinski, 1995; Parsons, 1996; Salanié, 2002; Viard, 2001a,b). Further, the merits of schemes which feature both targeted and universal dimensions have been widely discussed in the social policy and political science literature (Mkandawire, 2005; Skocpol, 1991).[3]

Two important and related features of any benefit system that involves targeting are (i) the conditions attached to targeted benefits; and (ii) the degree to which these conditions are enforced. We discuss both in turn.

Double Conditionality. Targeted benefits are typically conditioned in two dimensions: *ex-ante* an applicant must satisfy certain eligibility conditions to be awarded a given benefit; whilst *ex-post* a recipient must comply with certain behavioural requirements or restrictions. For example, disability benefits may be *ex-ante* conditioned on an applicant having a disability that substantially affects their ability to work; but also *ex-post* conditioned on, among other things, a recipient not working or only working low permitted amounts.[4] In many cases these two dimensions of conditionality will be linked under the presumption that an individual who satisfies the former will automatically satisfy the latter.[5]

Enforcement Errors. In large and complex welfare programmes both *ex-ante* and *ex-post* dimensions of conditionality may be imperfectly enforced.

[3]Skocpol (1991, p.414) refers to such programmes as 'targeting within universalism'. Drawing on the social policy history of the United States, she notes that whilst targeted programmes in isolation have been politically unsustainable, those that are more universal and spread benefits across groups have received broad political support and have been effective at targeting benefits to the needy.

[4]Whether or not a no-work condition is imposed will depend on the nature of the disability benefit. For example, the Employment and Support Allowance benefit in the U.K. allows recipients to undertake certain permitted work activities, such as earning very low amounts per week; potentially as part of a rehabilitation programme.

[5]Consider the following two examples, one more clear-cut than the other. First, an individual who is unable to work and awarded a disability benefit on this basis will automatically satisfy any *ex-post* 'no-work' condition. Second, an involuntarily unemployed individual who is actively searching for work, and awarded unemployment benefits on this basis, will *likely* satisfy any *ex-post* requirement to provide evidence of job search (this will of course depend on the stringency of the job search requirement).

- In the case of *ex-ante* conditionality, classification errors of Type I (false rejection) and Type II (false award) may be made in the awards process. A number of papers analyse the scope for classification errors in the U.S. Social Security Disability Insurance programme. Benitez-Silva et al. (2004) estimate a Type I error propensity of approximately 60% and a Type II error propensity of approximately 20%. These results are quantitatively similar to the earlier study of Nagi (1969). Classification errors may arise for two broad reasons: (i) the awards technology is imperfect and misclassifies individuals around a given eligibility threshold; and/or (ii) the eligibility threshold itself may be incorrect (e.g. too harsh or too lenient).[6]

- In the case of *ex-post* conditionality, the enforcement mechanisms put in place by the benefit authority may be insufficient to fully deter recipients from breaking the conditions placed on benefit receipt. This may arise because authorities (i) fail to detect all recipients who break the requirements; and, further, (ii) the sanctions imposed if detected are too lenient.

To the extent that *ex-ante* eligible individuals automatically satisfy *ex-post* conditions, Type II errors and the subsequent behaviour of ineligible recipients will be the source of *ex-post* enforcement issues.

In the context of partial universal benefit programmes three central questions therefore arise:

1. How does the propensity of the benefit system to make Type I and Type II errors affect (a) the decision of whether or not to provide a targeted benefit; and (b) conditional on providing *both* targeted and universal benefits, the respective levels of each?

2. How do these error propensities affect the resulting level of social welfare?

3. How do the answers to both of these questions depend on how well the *ex-post* conditionality is enforced?

[6]Disability benefits once more provide a good example where these sources of error can arise because (i) certain medical conditions are difficult to diagnose and verify (such as musculoskeletal illness and mental disorders) and (ii) the eligibility definition, or 'threshold', will determine the weight given to both strict medical criteria and the subjective assessment of ability to function in the workplace. Indeed, there has been much work exploring Type II errors and the work capability of recipients in the U.S. Social Security Disability Insurance programme (Autor and Duggan, 2006; Bound, 1989, 1991; Chen and van der Klaauw, 2008; Parsons, 1991, 1980; Von Wachter et al., 2011).

The existing literature on the targeting of benefits has tended to focus on part (a) of the first question. In doing so a particular enforcement structure in relation to *ex-post* conditionality has been assumed, but there has been no comparison across enforcement regimes. Moreover, while different papers have made different assumptions about the way *ex-post* conditionality is enforced, other differences in the modelling frameworks have not facilitated direct comparison. This is illustrated in the two related contributions of Parsons (1996) and Salanié (2002), where both papers explore the optimal provision of targeted benefits that are administered with classification errors. Parsons uses a framework where individuals are *ex-ante* identical[7] but face an exogenous probability of becoming unable to work. Individuals who are incorrectly tagged as unable by Type II error are allowed to work. Indeed, the author demonstrates that a 'dual negative income tax system' in which tagged able individuals are incentivised to work will be optimal, provided these individuals can be incentivised to work for less than their marginal product.[8] However, in a more general framework where able individuals differ over a productivity continuum, Salanié imposes the opposite restriction that tagged able individuals are not allowed to work, and shows that it is optimal to award a higher benefit to tagged individuals than those who are not tagged. Notably, Salanié does not model the decision of individuals to apply for the targeted benefit: instead, a fixed proportion - corresponding to the Type II error probability - of able individuals are tagged and do not work. Yet, this must be incentive incompatible for higher productivity able individuals who would rather receive the lower unconditional benefit and be allowed to work. Were application decisions modelled, these higher productivity individuals would not choose to apply.

The contribution of this chapter is to address the three questions raised above within a framework that allows for a systematic comparison of how the answer to the first two questions depends on how well the *ex-post* conditionality is enforced. We consider a framework where individuals differ in two dimensions. First, there is a categorical dimension: Individuals are either able to work or unable to work - modelled as a zero-hours quantity constraint on labour supply. Second, individuals who are able to work differ over a continuum of productivities. The government operates a tax/benefit system that comprises four elements: (i) a constant marginal tax rate on earned income;

[7]In particular, Parsons (1996) assumes that all individuals have the same productivity.

[8]From a *technical* perspective, allowing tagged individuals to work corresponds to not enforcing an *ex-post* 'no-work' requirement.

(ii) a tax-free universal benefit (B) which is received automatically by all individuals in society; (iii) a tax-free categorical benefit (C) which is targeted at those who are unable to work; and, for simplicity, (iv) a fixed budget for benefit expenditure.[9]

The categorical benefit is *ex-ante* conditional on an applicant being unable to work; and *ex-post* conditional on a recipient not working. Upon receiving an application, the benefit authority applies a test to determine whether or not an applicant is eligible to receive the benefit[10]. However, this test may make Type I and Type II classification errors with fixed (exogenous) probabilities that are independent of productivity (Salanié, 2002). The awards test will be said to have (i) *perfect discriminatory power* if the propensities to make Type I and Type II errors are both zero; (ii) *no discriminatory power* if unable and able applicants face the same probability of being awarded the benefit; and (iii) *some discriminatory power* for all intermediate cases where an unable applicant is more likely to be awarded the benefit than an able applicant, but there is a positive propensity to make at least one type of error.

Applications for the categorical benefit are taken to be costless in terms of money, stigma and time. This frequently employed assumption eliminates the direct dependence of application decisions on the propensity of the awards test to make errors (Jacquet, 2006, 2014). It is assumed that no checks or penalties are in place for an able individual who is incorrectly awarded the categorical benefit but does not work when receiving it. There are two reasons for this: (i) such behaviour is highly difficult to detect because the recipient does not reveal their true type through working (Yaniv, 1986); and, further, (ii) it is not immediately clear that such behaviour is 'fraudulent' as the applicant may be unsure of their own eligibility upon applying. The implication is that some able individuals with low productivities may always choose to apply for the benefit and, if awarded it, not work.

However, the application decisions of individuals with higher productivities will intuitively depend on how effectively the *ex-post* no-work condition is enforced. As

[9]Note that some countries such as the U.K. now operate a welfare cap on (forecasted) benefit expenditures. Programmes which fall under the cap include, to name a few, incapacity benefits, income support and child benefits. One of the motivations for such a policy is that it increases the monitoring of welfare spending and elicits policy decisions on the 'appropriate' level of welfare spending (see Crawford et al., 2014; Rhodes and Mcinnes, 2014).

[10]No penalty attaches to an individual whose application is rejected because they are deemed to be able-bodied. Indeed, on the grounds of legal uncertainty an applicant may be unsure of their own eligibility upon applying.

indicated, a number of different enforcement assumptions are made in the literature on categorical transfers administered with classification errors.[11] In this chapter we consider two discrete alternative enforcement regimes: *No Enforcement* and *Full Enforcement*.

- Under the *No Enforcement* regime there are no effective mechanisms in place to deter able recipients of C from subsequently working. The categorical benefit policy is therefore one dimensional: individuals receive a monetary amount C with no enforced restriction on labour supply. In this case all able individuals will apply: both those who would choose *not* to work when receiving C, and those who *would* choose to work when receiving C.

- Under the *Full Enforcement* regime it is assumed that the probability of detection and the penalty regime are sufficiently tough that no able individual who is wrongly awarded the categorical benefit will choose to subsequently work. The categorical benefit policy is therefore two-dimensional: individuals receive a monetary amount C and also a fully enforced zero quantity constraint on labour supply. Accordingly, the only able individuals who will choose to apply for the categorical benefit will be those of lower productivity, and thus those for whom the opportunity cost of not working is low. The exact range of productivities for which individuals choose to apply will be endogenous to the benefit size. Moreover, some of those who do apply would have chosen to work under the No Enforcement set-up and so are constrained by the fully enforced no-work requirement.

The government's optimisation problem is to choose the levels of categorical (C) and

[11] The literature on optimal benefits/transfers can be partitioned into three related strands. The first strand focuses on the design of welfare benefits when the benefit authority has no formal discriminatory test to determine eligibility. It instead chooses consumption bundles/transfer levels such that the non-needy opt against masquerading as the needy (see Besley and Coate, 1992; Blackorby and Donaldson, 1988; Cuff, 2000; Diamond and Mirrlees, 1978; Kreiner and Tranaes, 2005; Nichols and Zeckhauser, 1982). An important second strand analyses transfers within the optimal income taxation framework (Mirrlees, 1971; Sheshinski, 1972). In the standard model where individuals differ only through unobservable ability, Atkinson (1995) models a universal benefit financed by a linear income tax and allows for an unable subpopulation that cannot work. Following Akerlof (1978), a number of papers have modelled taxes and transfers when categorical information can be perfectly observed (see Immonen et al., 1998; Mankiw and Weinzierl, 2010; Viard, 2001a,b). The third important strand - and most related to this chapter - models categorical transfers administered with classification errors (see Diamond and Sheshinski, 1995; Jacquet, 2006, 2014; Kleven and Kopczuk, 2011; Parsons, 1996; Salanié, 2002; Stern, 1982).

universal (B) benefits that maximise a strictly utilitarian social welfare function subject to its budget constraint. Both the welfare function and expenditure component of the budget constraint will depend on the enforcement regime in place and, in turn, on the propensity of the benefit authority to make classification errors. The assumption of a fixed benefit budget simplifies the exposition because labour supply responses to unearned income do not affect the budget size. The purpose of making this assumption is that it allows us to pin down more clearly the intuition for the conditions under which it is optimal to provide a categorical benefit. Many of the key results do, however, generalise to the cases where (a) the benefit budget is driven by tax revenue; and (b) the government also optimises with respect to the tax rate. Indeed, characterising the optimal linear tax rate - both with and without classification errors - deserves more space than can be afforded here and is the subject matter of the following two chapters.

In the absence of any form of welfare provision, there are two types of inequality in this model.

1. *Within-group* inequality in the able subpopulation: Those with higher productivity have higher absolute utility but lower marginal indirect utility of income, or social marginal value of income (smvi).[12]

2. *Between-group* inequality: Those who are able to work will have a higher average level of utility but lower average smvi than those who are unable able to work.

As is well established from the literature on categorical transfers (Viard, 2001a,b), the purpose of the categorical benefit in this model is to reduce and, if possible, eliminate, inequality in the average smvi between the two subpopulations. When the benefit authority can perfectly discriminate between unable and able applicants this chapter demonstrates (as a baseline case) that it is always optimal to set $C > 0$; but $B > 0$ *only* if the benefit budget exceeds a critical level which, if spent entirely on categorical transfers, would equate the smvi of the unable with the average smvi of the able. There is thus an ordering of priorities: the first aim is to eliminate between-group inequality in the average smvi through categorical spending, whilst the second is to reduce inequality in the smvi throughout the population through universal spending.[13] Contrastingly,

[12]Under a strictly utilitarian objective function the social marginal value of income is simply the marginal (indirect) utility of income.

[13]Intuitively, social welfare will be raised more through categorical spending than universal spending so long as the smvi of an unable individual exceeds the average smvi of the able subpopulation.

whenever the test administering the categorical benefit makes Type I and/or Type II errors, this chapter demonstrates that between-group inequality the average smvi will never be eliminated at the optimum.

The major conclusions of this chapter are as follows:

- Under a *No Enforcement* regime:

1. It is optimal to set $C > 0$ whenever the test administering C has *some* discriminatory power. In this case the optimal benefits are chosen to equate - if budget feasible - the average smvi of categorical recipients with the average smvi of those not receiving the categorical benefit. A positive propensity to make Type I errors guarantees $B > 0$ at the optimum to ensure rejected unable individuals have some source of income to consume. However, if Type I errors are never made and the benefit budget is insufficiently large for categorical spending to equate the average smvi of categorical recipients with that of non-recipients, it will be optimal to set $B = 0$. Finally, if the test administering C has *no* discriminatory power it is optimal to set $C = 0$ and spend the entire benefit budget on B.[14] The intuition is that whenever the test has no discriminatory power (i) a targeted system does no better *between-group* than a pure universal system because the unable and able receive, on average, the same in benefit income; and further (ii) a targeted system does worse *within* group than a pure universal system because classification errors introduce horizontal inequities.

2. The associated value function of social welfare (i.e. maximum social welfare) is decreasing in the propensity to make both Type I and Type II errors, respectively.

- Under a *Full Enforcement* regime:

1. It is optimal to set $C > 0$ for *all* levels of discriminatory power, thus including the case of no discriminatory power. The intuition rests on the fact that, even if the test administering the categorical benefit has no discriminatory power, a targeted system (i) provides the unable with more in benefit income, on average, than the able; and (ii) redistributes within the able subpopulation through leak-

[14] An exception arises in the extreme case where the test has a zero propensity to make Type I errors but always makes Type II errors. In this case of no discriminatory power there is no optimisation problem to solve because all individuals in the economy receive the same benefit income. Consequently, any budget-feasible combination of B and C generates the same level of welfare.

age of the categorical benefit to lower productivity individuals. If Type I errors occur with positive propensity then it is always optimal to set $B > 0$. So under Full Enforcement it is always optimal to adopt a system with targeting: a *pure universal system* is never chosen.

2. Whilst maximum welfare is unambiguously decreasing in the propensity to make Type I errors, there are conditions under which it can be increasing in the propensity to make Type II errors. In particular, this is more likely to occur the larger the proportion of able applicants who are voluntarily unemployed (due to the universal benefit) and who thus have the same smvi as the unable.

One cannot in general guarantee a unique solution to the Full Enforcement optimisation problem. The optimal benefits are characterised by the condition that (i) the aggregate smvi of those not receiving the categorical benefit be equal to (ii) the aggregate smvi of categorical recipients *multiplied* by the increase in their total benefit income per unit reduction in the universal benefit. However, because the number of individuals who apply for the categorical benefit is endogenous to the benefit levels an increase in the universal benefit may generate conflicting effects on the two components of either aggregate smvi (i.e. individual smvi $\times$ number of individuals). For example, an increase in the universal benefit lowers the smvi of each individual who does not receive the categorical benefit, *but* also reduces the number of individuals who apply for - and thus receive - the categorical benefit. The overall effect is ambiguous: it will depend on unspecified properties of both the utility function (e.g. third derivatives) and the distribution function (e.g. derivatives of the pdf). As will be discussed below, numerical simulations suggest that the assumption of a unique optimum is valid in most cases - and identify the cases where it is not.

In line with much of the literature on optimal tax/benefits, we turn to numerical simulation methods to gain further comparative statics insights (see, for example, Immonen et al., 1998; Viard, 2001a,b). The purpose of the simulations is twofold: (i) to establish how the optimal benefit levels change with the propensity to make Type I and Type II classification errors, respectively; and (ii) to provide examples where maximum welfare increases with the Type II error propensity under the Full Enforcement regime. For this, we employ CES preferences and take individual productivity to be exponentially distributed. The numerical results we obtain are consistent with those established in the theory. The analysis of the Full Enforcement case is particularly sensitive to the value of the elasticity of substitution between leisure and consumption, which we systematically vary between 0.5 and 0.99. At lower values of the elasticity the welfare function may not be concave in the benefit levels and, consequently, care has to be taken in searching for a global optima.

The remainder of this chapter is structured as follows. Section 2.2 sets out the model and discusses how the different enforcement assumptions affect individual application decisions for the categorical benefit. Section 2.3 then presents the main analysis: here the optimum benefit levels are characterised and the effects of errors on maximum social welfare are determined under both enforcement assumptions. Section 2.4 then numerically simulates the optimal benefits under both enforcement regimes. This provides insights into how the optimal benefit levels change with the propensity to make classification errors. Finally, Section 2.5 concludes the chapter.

2.2 The Model

2.2.1 Background: Individuals

Individuals in the economy have identical preferences given by the utility function $u(x, l)$, where $x \geq 0$ denotes consumption and $l \in [0, 1]$ denotes leisure respectively. The standard assumptions apply: u is continuous, differentiable, strictly increasing in both arguments ($u_x > 0$, $u_l > 0$) and strictly concave ($u_{xx} < 0$, $u_{ll} < 0$, $u_{xx}u_{ll} - u_{xl}^2 > 0$); with both goods normal ($u_l u_{xx} - u_x u_{xl} < 0$).

We also assume:

$$\lim_{x \to 0} u_x(x, l) = +\infty \tag{2.1}$$

An individual with net wage $\omega \geq 0$ and unearned income $M \geq 0$ chooses labour supply, $H \in [0, 1]$, so as to maximise utility. The resulting optimal labour supply (H^*) and indirect utility (v) functions are given by:

$$\begin{aligned} H^*(\omega, M) &\equiv \arg\max_{H\in[0,1]} u(\omega H + M, 1 - H) \\ v(\omega, M) &\equiv \max_{H\in[0,1]} u(\omega H + M, 1 - H) = u(\omega H^* + M, 1 - H^*) \end{aligned} \tag{2.2}$$

Formally, H^* satisfies:

$$\omega \leq \frac{u_l(\omega H^* + M, 1 - H^*)}{u_x(\omega H^* + M, 1 - H^*)} \;\; ; \; H^* \geq 0$$

where the pair of inequalities hold with complementary slackness.

Let $\bar{\omega}(M) \equiv u_l(M, 1)/u_x(M, 1)$ denote the reservation wage at or below which an individual with lump sum unearned income M will choose not to work. Formally, $\bar{\omega}$ satisfies $H^*[\bar{\omega}(M), M] = 0$. Furthermore, given that leisure is normal $\bar{\omega}'(M) > 0$.[15]

For $\omega \leq \bar{\omega}(M)$ we have $H^*(\omega, M) \equiv 0$ which implies $v(\omega, M) \equiv u(M, 1)$, from which it follows that $v_M(\omega, M) \equiv u_x(M, 1)$ and $v_{MM}(\omega, M) = u_{xx}(M, 1) < 0$. So, over this range of net wage rates the marginal indirect utility of unearned income is constant and independent of the wage rate.

For $\omega > \bar{\omega}(M)$ we have $H^*(\omega, M) > 0$. Further, by the normality of leisure:

$$H_M^* = \frac{\omega u_{xx} - u_{xl}}{2\omega u_{xl} - \omega^2 u_{xx} - u_{ll}} = \frac{(u_l u_{xx} - u_x u_{xl})/u_l}{[2u_{xl} - (u_l/u_x)u_{xx} - (u_x/u_l)u_{ll}]} < 0 \tag{2.3}$$

[16]By the envelope theorem the marginal indirect utility of unearned income is $v_M(\omega, M) = u_x(\omega H^* + M, 1 - H^*)$, whilst from (2.3) it also follows that:

$$v_{MM} = u_{xx} + H_M^*(\omega u_{xx} - u_{xl}) = \frac{u_x^2(u_{xl}^2 - u_{xx}u_{ll})}{u_x u_l[2u_{xl} - (u_l/u_x)u_{xx} - (u_x/u_l)u_{ll}]} < 0 \tag{2.4}$$

[15]Formally, $\bar{\omega}'(M) = (u_x u_{xl} - u_l u_{xx})/u_x^2 > 0$.

[16]The denominator of (2.3) is positive by the concavity - and thus quasiconcavity - of the utility function. See, for example, Simon and Blume (1994).

Roy's identity (i.e. $v_\omega = v_M H^*$) and the assumption that leisure is normal together imply that $v_{\omega M} = v_{M\omega} = v_{MM}H^* + v_M H^*_M < 0$. So over this range of net wages the marginal indirect utility of unearned income is a strictly decreasing function of the net wage. Formally:

$$\forall\ \omega > \bar{\omega}(M) : v_{\omega M} < 0 \Rightarrow v_M(\omega, M) < v_M(\bar{\omega}(M), M) = u_x(M, 1) \tag{2.5}$$

2.2.2 The Population and Tax-Benefit System

In a population of size 1, there are two distributional issues. First, the fraction $\theta \in (0, 1)$ of individuals face a zero quantity constraint on labour supply and are thus unable to work. Absent any provision of state financial support, these individuals would have no source of income to consume. Second, the remaining fraction $(1 - \theta)$ of individuals are able to work, but differ continuously in their productivity, thereby giving rise to earned income inequality.

There is a tax-benefit system comprising four elements:

1. A constant marginal tax rate, $t \in (0, 1)$, on all earned income;
2. A tax-free universal benefit, $B \geq 0$, paid automatically to everyone[17] ;
3. A tax-free categorical benefit, $C \geq 0$, that is targeted - potentially imperfectly - at those who are unable to work and is received in addition to B ;
4. A fixed benefit budget, $\beta > 0$, to be spent on the universal and/or the categorical benefits.

In all that follows we take the tax rate as exogenously fixed. Assuming a constant marginal tax rate allows productivity and productivity differences to be captured through the net wage, $\omega \geq 0$. Net wages are distributed with density $f(\omega)$, where $f(\omega) > 0\ \forall\ \omega \geq 0$ and $\int_0^\infty f(\omega)d\omega = 1$. The associated distribution function is $F(\omega) = \int_0^\omega f(z)dz$, where $0 \leq F(\omega) \leq 1$. Note from (2.5) that:

$$\int_0^\infty v(\omega, M)f(\omega)d\omega > u(M, 1)\ \ ;\ \ \int_0^\infty v_M(\omega, M)f(\omega)d\omega < u_x(M, 1) \tag{2.6}$$

[17] With just these first two elements, we would have effectively a simple negative income tax system.

Conditional on everyone receiving the same unearned income, this simply states that the average utility of the able exceeds that of the unable and, by diminishing marginal utility of income, the average marginal utility of income for the able must be lower than that of the unable. This latter inequality can be reduced through targeting additional resources at the unable subpopulation (we have of course yet to specify the objective function of the government).

Whilst B is automatically received by everyone, C must be applied for and is subject to the following *double conditionality*:

- *Ex-ante* conditionality: An applicant for C must be unable to work to be awarded it.
- *Ex-post* conditionality: A recipient of C must not subsequently work.[18]

An individual who has applied for C is subject to a test to determine whether they are indeed unable to work. If deemed unable to work, they are awarded the benefit. However, this test may be imperfect and, in statistical parlance, subject to Type I (false rejection) and/or Type II (false award) classification errors. A Type I error occurs when an individual who is unable to work is incorrectly classified as being able to work and consequently rejected C. We denote the probability of a Type I error occurring by $p_I \in [0,1]$. Contrastingly, a Type II error arises when an individual who is able to work is incorrectly awarded C. For simplicity, we assume that the probability of Type II error is independent of productivity and denote this by $p_{II} \in [0,1]$. In what follows, we confine our attention to error propensities satisfying:

$$p_I + p_{II} \leq 1 \tag{2.7}$$

The level of *discriminatory power* that the test has can be simply characterised by:

$$p_I + p_{II} \begin{cases} = 0 & : \textit{Perfect}\text{ Discriminatory Power} \\ \in (0,1) & : \textit{Some}\text{ Discriminatory power} \\ = 1 & : \textit{No}\text{ Discriminatory power} \end{cases}$$

So the test has *perfect* discriminatory power if no classification errors are made. In this

[18]Indeed, since C is targeted at the unable, it is *de facto* conditional on not subsequently working.

setting the categorical benefit is only awarded to unable applicants and, further, all unable applicants are awarded it. Contrastingly, the test has *no* discriminatory power if it awards the categorical benefit to an able applicant with the same propensity that it does an unable applicant - i.e. the test brings no useful information in distinguishing between the two types of applicant. Finally, for all intermediate cases where $p_1 + p_{II} \in (0, 1)$, we say that the test has *some*, but not perfect, discriminatory power; because it is more likely to award the categorical benefit to an unable applicant than an able one.

Trivially, if $p_{II} = 0$ *ex-post* conditionality is automatically satisfied because all recipients of the categorical benefit are unable to work. Contrastingly, $p_{II} > 0$ generates a host of enforcement issues: whether or not an able recipient of C will choose to work will depend on - in addition to their underlying productivity - the strength of the enforcement regime in place. We turn to discuss this in the following section.

2.2.3 Benefit Applications and Enforcement.

Applications for C are taken to be costless in terms of money, stigma and time. This frequently employed assumption eliminates the direct dependence of application decisions on p_I and p_{II} because individual welfare in the rejected state coincides with that from having not applied (Jacquet, 2006, 2014).[19]

It follows that an unable individual will always choose to apply for C. Contrastingly, the application decision of an able individual will depend on the enforcement mechanisms in place to detect ineligible recipients and, conditioning on this, their underlying productivity. First, it is assumed that there are no checks or penalties in place for an able individual who applies for C and, if awarded it, subsequently complies with *ex-post* conditionality through not working. There are two reasons for making this assumption: (i) such behaviour is highly difficult to detect because the recipient does not reveal their true type through working (Yaniv, 1986); and, further, (ii) it is not immediately clear that such behaviour is 'fraudulent' as the applicant may be unsure of their own eligibility upon applying. Given this assumption, all able individuals with

[19] As will be discussed, which able individuals will apply for the categorical benefit will depend on the size of the benefit itself, which - at the optimum - will in turn be an implicit function of p_I and p_{II}. There will therefore be an indirect dependency of application decisions of the able on p_I and p_{II}.

$\omega \leq \bar{\omega}(B+C)$ will apply for C for sure, where $\bar{\omega}(B+C) > \bar{\omega}(B)\ \forall\ C > 0$.[20]

However, which individuals commanding the higher productivities $\bar{\omega}(B+C) < \omega$ will choose to apply for C will depend on how effectively the no-work condition is enforced. In this regard, we analyse two binary enforcement assumptions where the no-work requirement is either (i) not enforced at all; or is (ii) fully enforced. Both alternative assumptions are detailed below.

No Enforcement

In this first case we assume that there are no effective mechanisms in place to deter an able recipient of C from working should they wish to. In this sense, benefit policy is one-dimensional because recipients receive only the monetary amount C, with no subsequent restrictions on labour supply. Accordingly, all able individuals along the net wage continuum will choose to apply for C. Of the proportion $(1 - p_{II})$ who are correctly denied the benefit, all those with $\omega \leq \bar{\omega}(B)$ will choose not to work, whilst those with $\bar{\omega}(B) < \omega$ will work. Turning to the proportion p_{II} who are awarded C, all those with $\omega \leq \bar{\omega}(B+C)$ will not work, whilst those with $\bar{\omega}(B+C) < \omega$ will work.

Full Enforcement

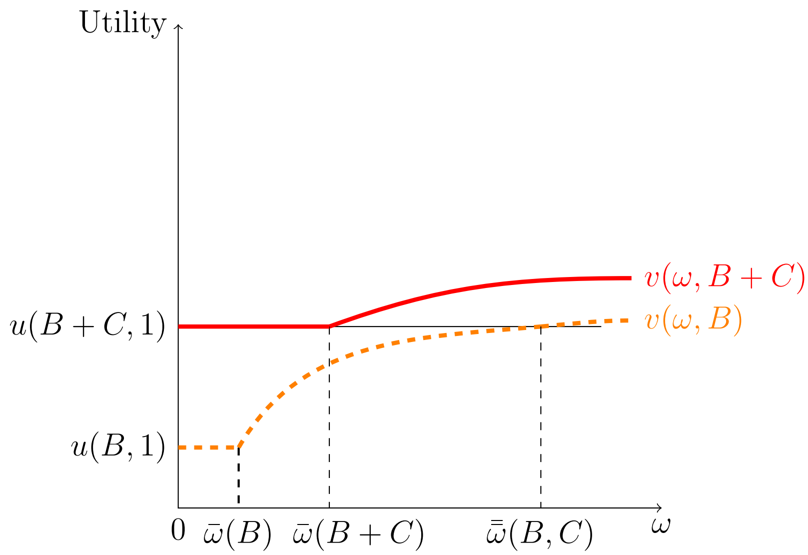
In this second case, we alternatively assume that there are totally effective mechanisms in place that fully deter any able individual from working whilst receiving C. The benefit policy now has two dimensions: individuals who are awarded the benefit receive both the monetary amount C and a fully enforced zero quantity constraint on labour supply. As under No Enforcement, all those with $\omega \leq \bar{\omega}(B+C)$ will apply for C because they would choose not to work when receiving it even if there were no restrictions on labour supply. However, those with $\omega > \bar{\omega}(B+C)$ will only choose to apply for C if $\omega \leq \bar{\bar{\omega}}(B, C)$, where:

$$v[\bar{\bar{\omega}}(B,C), B] \equiv u(B+C, 1) \tag{2.8}$$

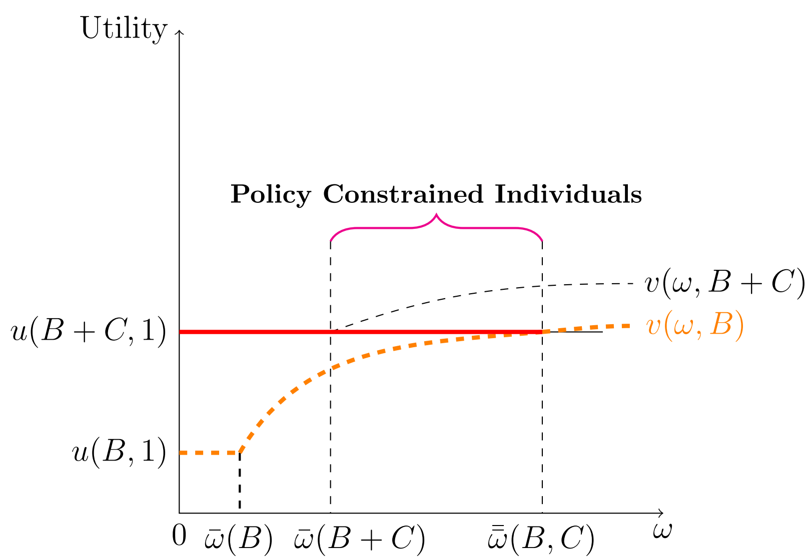
[20] Recall that $\bar{\omega}(M) = u_l(M,1)/u_x(M,1)$ is the reservation wage at or below which an individual chooses voluntary unemployment.

Figure 2.1: Work Decision and Utility under Alternative Enforcement Regimes

(a) No Enforcement

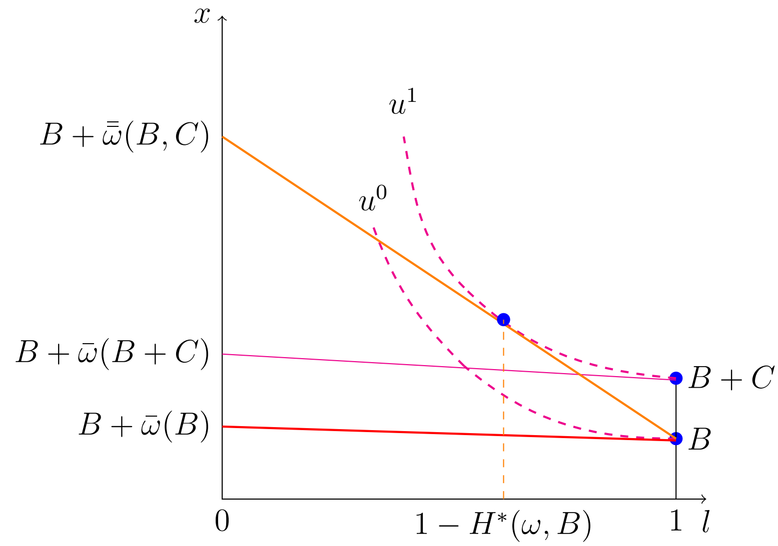


(b) Full Enforcement



Notes. The bold unbroken lines denote recipients of the categorical benefit; whilst the bold broken lines denote non-recipients of the categorical benefit. Under No Enforcement (panel (a)) *all* individuals choose to apply for C. Of those awarded it all those with $\omega \leq \bar{\omega}(B+C)$ choose not to work; whilst all those with $\omega > \bar{\omega}(B+C)$ instead work. Alternatively, under Full Enforcement (panel(b)) only those with $\omega \leq \bar{\bar{\omega}}(B,C)$ apply for C. Of the individuals awarded it, notice that those with $\bar{\omega}(B+C) < \omega \leq \bar{\bar{\omega}}$ would have chosen to work under the No Enforcement regime.

Figure 2.2: The critical wages $\bar{\omega}(B)$, $\bar{\omega}(B+C)$ and $\bar{\bar{\omega}}(B,C)$



Notes. The figure provides graphical intuition for the critical net wages $\bar{\omega}$ and $\bar{\bar{\omega}}$. u^0 is the indifference curve associated with the utility level $u(B,1)$; whilst u^1 is the indifference curve associated with the utility level $u(B+C,1)$. The function $\bar{\omega}(M)$ captures the corner solution case where $H^*[\bar{\omega}(M),M]=0$ and thus $v(\omega,M)=u(M,1)$. Contrastingly, $\bar{\bar{\omega}}(B,C)$ captures the interior solution case where $H^*[\bar{\omega}(B+C),B]>0$ and thus $v(\omega,B)>u(B,1)$ *but* $v(\bar{\bar{\omega}},B)=u(B+C,1)$.

So $\bar{\bar{\omega}}(B,C)$ is the net wage at which an able individual is just indifferent between (i) not working and receiving total benefit income $B+C$, and (ii) working as much as desired and receiving only B in benefit income. It is straightforward to show that:

$$\bar{\bar{\omega}}_C=\frac{u_x(B+C,1)}{H^*(\bar{\bar{\omega}},B)\cdot v_M(\bar{\bar{\omega}},B)}>\bar{\bar{\omega}}_B=\frac{u_x(B+C,1)-v_M(\bar{\bar{\omega}},B)}{H^*(\bar{\bar{\omega}},B)\cdot v_M(\bar{\bar{\omega}},B)}>0 \tag{2.9}$$

and thus $\bar{\bar{\omega}}_C-\bar{\bar{\omega}}_B=1/H^*(\bar{\bar{\omega}},B)$. Note that $\bar{\bar{\omega}}_B>0$ by the normality of leisure; which implies that the marginal utility of consumption is increasing in leisure along an indifference curve.[21]

Finally, given that $v[\bar{\omega}(B+C),B+C]\equiv v[\bar{\bar{\omega}}(B,C),B]\equiv u(B+C,1)$, it must hold

[21] Let $x(l)$ satisfy $u[x(l),l]=k$, where k is a constant. Differentiating w.r.t. l thus gives $x'(l)=-u_l/u_x$. From this it follows that:

$$\frac{du_x[x(l),l]}{dl}=u_{xx}x'(l)+u_{xl}=\frac{u_xu_{xl}-u_lu_{xx}}{u_x}>0$$

by the normality of leisure.

that:

$$\bar{\omega}(B + C) < \bar{\bar{\omega}}(B, C) \ \ \forall\, C > 0 \tag{2.10}$$

where $\lim_{C\to 0} \bar{\bar{\omega}}(B, C) = \bar{\omega}(B)$.

Policy Constrained Individuals

Under Full Enforcement, the only recipients of C for whom the enforced zero quantity constraint on labour actually binds are those with $\bar{\omega}(B+C) < \omega \leq \bar{\bar{\omega}}(B, C)$. To see this, note that under the No Enforcement regime these individuals would optimally choose to work when receiving C. They are therefore constrained by the fully enforced benefit policy. Figure 2.1 illustrates this: the figure depicts the work decision and resulting utility of individuals over the net wage continuum under both the No Enforcement and Full Enforcement regimes. The bold unbroken lines denote utility when receiving the categorical benefit, whilst the bold broken lines denote utility when not receiving it. Figure 2.2, meanwhile, illustrates that the budget constraints with slope $\bar{\omega}(B+C)$ and $\bar{\bar{\omega}}(B, C)$, respectively, are tangent to the same indifference curve, but the tangency in the former case occurs at a corner solution ($H = 0$), whilst the latter occurs at an interior solution ($H > 0$).

2.3 Analysis

We now turn to the main theoretical analysis. When the test administering the categorical benefit makes Type I and/or Type II classification errors with positive propensity the key questions we wish to answer are:

1. For what levels of discriminatory power would it be optimal to adopt: (i) a pure universal system ($B > 0, C = 0$); (ii) a partial universal system ($B > 0, C > 0$); or (iii) a pure targeted system ($B = 0, C > 0$)?
2. How does maximum social welfare change with the propensity to make classification errors?
3. How do the answers to these questions differ between the No Enforcement and Full Enforcement regimes?

In all cases, the social welfare (objective) function that the government seeks to maximise is strictly utilitarian, and thus given by the sum of individual (expected) utilities. A desire for redistribution therefore arises solely from the concavity of individual utility. We will throughout adopt the notational convention of letting the superscripts P, N and F denote Perfect Discrimination, No Enforcement and Full Enforcement, respectively.

To proceed, we initially analyse the first design question under the baseline case of Perfect Discrimination.

2.3.1 Perfect Discrimination

When the test administering C can perfectly discern unable applicants from able applicants, social welfare is given by:

$$W^P(B,C;\theta) = \theta u(B+C,1) + (1-\theta)\int_0^\infty v(\omega,B)f(\omega)d\omega \tag{2.11}$$

Given the fixed budget size β available for expenditure on B and C, the problem is to choose B and C so as to maximise social welfare. Formally:

$$\max_{B,C} W^P(B,C;\theta) \quad \text{s.t.} \;\; B+\theta C=\beta \;\;, B\geq 0 \;\;, C\geq 0 \tag{2.12}$$

Notice that the budget constraint must hold with equality because, if underspent, welfare could always be raised through paying a higher B to all individuals. If we denote the optimum benefit levels by $\hat{B}^P(\beta,\theta)$ and $\hat{C}^P(\beta,\theta)$ respectively, we have:

Proposition 1. *$\hat{C}^P > 0$ and $\hat{B}^P \geq 0$ satisfy:*

$$\int_0^\infty v_M(\omega,\hat{B}^P)f(\omega)d\omega \leq u_x(\hat{B}^P+\hat{C}^P,1) \;\; ; \;\; \hat{B}^P \geq 0 \tag{2.13}$$

where the pair of inequalities hold with complementary slackness.

Proof: See Appendix.

Corollary 1.

$$\hat{B}^P \left\{\begin{matrix} > \\ = \end{matrix}\right\} 0 \text{ if } \beta \left\{\begin{matrix} > \\ \leq \end{matrix}\right\} \bar{\beta}^P \tag{2.14}$$

where $\bar{\beta}^P$ is defined by:

$$\int_0^\infty v_M(\omega, 0) f(\omega) d\omega \equiv u_x\left(\frac{\bar{\beta}^P}{\theta}, 1\right) \tag{2.15}$$

Proposition 1 states that (i) it is always optimal to provide a categorical benefit; but (ii) it will only be optimal to provide a universal benefit if categorical transfers have been financed up the point where inequality in the average smvi between the unable and able subpopulations is eliminated. It thus follows directly from Corollary 1 that there will be a critical budget ($\bar{\beta}^P$) which, if spent solely on categorical transfers, will equate the smvi of the unable with the average smvi of the able. For any budget exceeding this level it will be optimal to provide a universal benefit.[22] Finally, notice that in any optimum $\hat{B}^P + \hat{C}^P > \beta > \hat{B}^P$.[23]

The result that the categorical benefit should be set so as to equate the average smvi across categorical groups (in our case the unable and able) is well established in the optimal tax-benefit literature (Atkinson, 1995; Beath et al., 1988; Immonen et al., 1998; Viard, 2001a,b)

To provide some intuition for Proposition 1, let $V^P(\beta, \theta) \equiv W^P(\hat{B}^P, \hat{C}^P; \theta)$ denote maximum social welfare and let W^U denote welfare under a pure universal system (i.e.

[22] There is therefore an ordering or priorities. The first priority is to eliminate, budget allowing, between-group inequality in the average smvi - i.e. support unable individuals because they are the most needy in society from the perspective of smvi. Conditional on the benefit budget being sufficiently large for this to be achieved and, further, there being money left over, the second priority is to spend the remainder of the benefit budget on the universal benefit. This in turn reduces inequality in the smvi *within* the able subpopulation.

[23] **Perfect Discrimination One Dimensional Problem.** From the budget constraint we could alternatively define $C^P(B; \beta, \theta) = (\beta - B)/\theta$ and solve the unconstrained one-dimensional problem $\max_{B \in [0,\beta]} W^P(B, C^P; \theta)$. This yields the first order condition:

$$\theta u_x(\hat{B}^P + C^P, 1) \cdot (1 + \partial C^P/\partial B) + (1-\theta) \int_0^\infty v_M(\omega, \hat{B}^P) f(\omega) d\omega \leq 0 \; ; \;\; \hat{B}^P \geq 0$$

If we substitute in $\partial C^P/\partial B = -1/\theta$ the results in the main text directly follow.

$B = \beta, C = 0$), where:

$$W^U(\beta, \theta) \equiv \theta u(\beta, 1) + (1 - \theta) \int_0^\infty v(\omega, \beta) f(\omega) d\omega \tag{2.16}$$

Since the pure universal system is always a feasible choice, but is rejected in favour of a system which targets a categorical benefit at the unable, it must follow that $V^P(\beta, \theta) - W^U(\beta, \theta) > 0$. To see this, note that a first-order Taylor approximation around β gives:

$$\begin{aligned} V^P - W^U &\approx \theta u_x(\beta, 1) \cdot \left(\hat{B}^P + \hat{C}^P - \beta\right) + (1 - \theta) \int_0^\infty v_M(\omega, \beta) f(\omega) d\omega \cdot \left(\hat{B}^P - \beta\right) \\ &= \theta(1 - \theta) \left[u_x(\beta, 1) - \int_0^\infty v_M(\omega, \beta) f(\omega) d\omega\right] \hat{C}^P > 0 \end{aligned}$$

By diminishing marginal utility of income, social welfare can always be raised above W^U by transferring some benefit income away from those who are able to work towards the unable. Of course, the relative amounts that we would choose to transfer will depend on the proportion of the population who are unable.

With this section serving as a baseline case, we now turn to analyse optimal welfare provision when the test administering the categorical benefit makes classification errors. We first study the No Enforcement regime, and then subsequently turn to the Full Enforcement regime.

2.3.2 Imperfect Discrimination: No Enforcement

With no enforcement mechanisms in place to restrict the work behaviour of able individuals who receive C, we know that all able individuals will apply for C. Social welfare is therefore now given by:

$$
\begin{aligned}
&W^N(B, C; \theta, p_I, p_{II}) \\
&= \theta \left\{ (1 - p_I)u(B + C, 1) + p_I u(B, 1) \right\} \\
&+ (1 - \theta) \left\{ \begin{array}{l} p_{II} \left\langle F[\bar{\omega}(B + C)]u(B + C, 1) + \int_{\bar{\omega}(B+C)}^{\infty} v(\omega, B + C) f(\omega) d\omega \right\rangle \\ \qquad +(1 - p_{II}) \left\langle F[\bar{\omega}(B)]u(B, 1) + \int_{\bar{\omega}(B)}^{\infty} v(\omega, B) f(\omega) d\omega \right\rangle \end{array} \right\}
\end{aligned} \tag{2.17}
$$

Notice that $W^N(B, C; \theta, 0, 0) = W^P(B, C; \theta)$ - i.e. when there are no classification errors in the awards process the welfare function in (2.17) reduces to its Perfect Discrimination counterpart in (2.11).[24] The first line on the right side of (2.17) illustrates that horizontal inequity in utility levels is introduced into the unable subpopulation through Type I errors. These individuals derive consumption solely from unearned income and so any disparities in their utility arise due to unequal treatment by the benefit system. The proportion p_I are incorrectly denied C, whilst the proportion $(1 - p_I)$ are correctly awarded it. The second line, meanwhile, concerns the able subpopulation. Again, classification errors introduce inequalities into this subpopulation as the individual welfare rankings now differ from the case where $p_{II} = 0$. Consider, for example, those with $\omega \leq \bar{\omega}(B)$: these individuals are voluntarily unemployed and share the same welfare level $u(B, 1)$ when $p_{II} = 0$, but when $p_{II} > 0$ some have welfare $u(B + C, 1)$ whilst others have welfare $u(B, 1)$.

The government budget constraint is given by:

$$B + [\theta(1 - p_I) + (1 - \theta)p_{II}]C \leq \beta \tag{2.18}$$

[24]Writing social welfare as in (2.17) illustrates the voluntary unemployment induced by Type II classification errors. We could instead write social welfare more compactly as:

$$W^N = \theta \left\{ (1 - p_I)u(B + C, 1) + p_I u(B, 1) \right\} + (1 - \theta) \left\{ \int_0^{\infty} \langle p_{II} v(\omega, B + C) + (1 - p_{II}) v(\omega, B) \rangle f(\omega) d\omega \right\}$$

Properties of the budget constraint. Let $C^N(B;\beta,\theta,p_I,p_{II})$ denote the level of categorical benefit that exhausts the budget constraint for any $B \in [0,\beta]$. Formally:

$$C^N(B;\beta,\theta,p_I,p_{II}) = \frac{\beta - B}{[\theta(1-p_I)+(1-\theta)p_{II}]} \tag{2.19}$$

It is straightforward to see that $\partial C^N/\partial B < -1$ and thus $d[B+C^N]/dB = 1 + \partial C^N/\partial B < 0$. An increase in the universal benefit therefore reduces the total benefit income of categorical recipients.[25]

Optimisation problem. The optimisation problem of the government is given by:

$$\begin{aligned} &\max_{B,C} W^N(B,C;\theta,p_I,p_{II}) \\ \text{s.t. } & B + [\theta(1-p_I)+(1-\theta)p_{II}]C = \beta\,, \\ & B \geq 0\,, C \geq 0. \end{aligned} \tag{2.20}$$

To proceed, let us define (i) the aggregate smvi of those not receiving the categorical benefit (henceforth 'non-categorical recipients') and (ii) the aggregate smvi of categorical recipients, respectively, by:

$$\begin{aligned} \sigma^N_{NR}(B;\theta,p_I,p_{II}) &\equiv \theta p_I u_x(B,1) + (1-\theta)(1-p_{II})\int_0^\infty v_M(\omega,B)dF(\omega) \\ \sigma^N_R(B,C;\theta,p_I,p_{II}) &\equiv \theta(1-p_I)u_x(B+C,1) + (1-\theta)p_{II}\int_0^\infty v_M(\omega,B+C)dF(\omega) \end{aligned} \tag{2.21}$$

The subscript NR denotes 'non-categorical recipients', whilst the subscript R denotes categorical recipients. We let the corresponding averages be given by $\bar{\sigma}^N_{NR} \equiv \sigma^N_{NR}/[\theta p_I + (1-\theta)(1-p_{II})]$ and $\bar{\sigma}^N_R \equiv \sigma^N_R/[\theta(1-p_I)+(1-\theta)p_{II}]$, respectively. It is useful to note that $\partial\bar{\sigma}^N_{NR}/\partial p_I > 0$ and $\partial\bar{\sigma}^N_{NR}/\partial p_{II} > 0$. Intuitively, a *ceteris paribus* increase in

[25] The choice set of B and C is clearly convex because $\partial C^N/\partial B < 0$ whilst $\partial^2 C^N/\partial B^2 = 0$.

either p_I or p_{II} acts to increase the proportion of non-categorical recipients who are in fact unable, and thus acts to increase the average. By parallel argument, it also holds that both $\partial\bar{\sigma}_R^N/\partial p_I < 0$ and $\partial\bar{\sigma}_R^N/\partial p_{II} < 0$.

If we now denote the optimal benefit levels by $\hat{B}^N(\beta,\theta,p_I,p_{II})$ and $\hat{C}^N(\beta,\theta,p_I,p_{II})$ respectively, we have:

Proposition 2a:

(i) $\forall\ p_I + p_{II} < 1\ \ \hat{C}^N > 0$ *and* $\hat{B}^N \geq 0$ *satisfy:*

$$\bar{\sigma}_{NR}^N\left(\hat{B}^N;\theta,p_I,p_{II}\right) \leq \bar{\sigma}_R^N\left(\hat{B}^N,\hat{C}^N;\theta,p_I,p_{II}\right)\ ;\ \ \hat{B}^N \geq 0 \tag{2.22}$$

where the pair of inequalities in (2.22) hold with complementary slackness.

(ii) $\forall\ p_I + p_{II} = 1$ *(a)* $\hat{C}^N = 0$ *and* $\hat{B}^N = \beta$ *if* $p_I > 0$*; but (b) any* (B,C) *satisfying* $B + C = \beta$ *will be optimal in the extreme case where* $p_I = 0$ *(and* $p_{II} = 1$*).*

Proof: See Appendix[26]

Corollary 2: $\hat{B}^N > 0$ *if* $p_I > 0$ *; otherwise* $\hat{B}^N \begin{Bmatrix} > \\ = \end{Bmatrix} 0$ *as* $\beta \begin{Bmatrix} > \\ \leq \end{Bmatrix} \bar{\beta}^N$, *where* $\bar{\beta}^N$ *is the critical budget level satisfying*

$$\bar{\sigma}_{NR}^N(0;\theta,0,p_{II}) = \int_0^\infty v_M(\omega,0)dF(\omega) = \bar{\sigma}_R^N\left(0,\frac{\bar{\beta}^N}{\theta+(1-\theta)p_{II}};\theta,p_I,p_{II}\right) \tag{2.23}$$

Proof: See Appendix

To discuss Proposition 2a, it is useful to distinguish between the cases where $p_I > 0$ and $p_I = 0$. Suppose first that $p_I > 0$: Proposition 2a states that a necessary and sufficient condition to provide a categorical benefit is that the test administering it has

[26]**Unconstrained uni-dimensional problem.** We could alternatively substitute the function C^N into the welfare function in (2.17) and solve the unconstrained uni-dimensional problem $\max_{B\in[0,\beta]} W^N(B,C^N;\theta,p_I,p_{II})$. This yields the first order condition:

$$\theta p_I u_x(\hat{B}^N,1) + (1-\theta)(1-p_{II})\int v_M(\omega,\hat{B}^N)f(\omega)d\omega$$
$$\leq \left\{\theta(1-p_I)u_x(\hat{B}^N+C^N,1) + (1-\theta)p_{II}\int v_M(\omega,\hat{B}^N+C^N)f(\omega)d\omega\right\}\cdot -(1+\partial C^N/\partial B)\ ;\ \hat{B}^N \geq 0$$

Substituting in $(1+\partial C^N/\partial B) = -[\theta p_I + (1-\theta)(1-p_{II})]/[\theta(1-p_I)+(1-\theta)p_{II}]$ and rearranging then gives the same expression as in (2.22).

some positive discriminatory power. If this holds then the benefit levels are chosen in accordance with (2.22). The term to the left of the first inequality in (2.22) is the average smvi of non-categorical recipients, whilst the term to the right is the average smvi of categorical recipients. The complementary slackness condition implies that spending should be exclusively categorical up to the point where the average smvi of categorical recipients is equated with that of non-categorical recipients. Corollary 2 immediately follows from (2.22). Given our assumption in (2.1) that $\lim_{x\to 0} u_x = +\infty$ a universal benefit will always be provided to ensure rejected unable individuals have some source of income to consume. At the optimum then, the average smvi of non-categorical recipients is equated with the average smvi of categorical recipients. However, if the awards test has no discriminatory power it will be suboptimal to provide a categorical benefit. Instead, the entire benefit budget should be spent on the universal benefit.

Suppose alternatively that $p_I = 0$: in this case the story in Proposition 2a is a little more nuanced, but the main message remains. If the awards test has positive discriminatory power it will certainly be optimal to provide a categorical benefit. Given that all able individuals receive the categorical benefit, there is no guarantee that it will be optimal to provide a universal benefit. Indeed, it follows from (2.23) and Corollary 2 that it will only be optimal to provide a universal benefit if the benefit budget exceeds a critical level ($\bar{\beta}^N$) which - if spent entirely on categorical transfers - would equate the average smvi of non-categorical recipients with the average smvi of categorical recipients. However, if the awards test has no discriminatory power then the social welfare function to be 'maximised' is simply $W^N(B, C; \theta, 0, 1) = \theta u(B+C, 1) + (1-\theta) \int v(\omega, B+C) f(\omega) d\omega$, such that *any* combination of B and C satisfying $B + C = \beta$ is 'optimal'. Any budget feasible combination of the benefit levels yields the same level of welfare and there is no optimisation problem to solve. Note that within this framework this is equivalent to a pure universal system.[27]

For the ease of exposition, we assume that $p_I > 0$ for the remainder of the No Enforcement analysis.

[27]That is, we abstract from any differences in costs there may be between administering a universal benefit and administering a categorical benefit.

Intuition for Proposition 2a To provide the intuition for Proposition 2a we compare welfare under an arbitrary budget feasible targeted system ($B \in [0, \beta), C^N > 0$) with that under a pure universal system ($B = \beta, C^N = 0$):

$$
\begin{aligned}
&W^N - W^U \\
&= \theta \left\{ \begin{array}{c} \langle (1-p_I)u(B+C^N,1) + p_I u(B,1)\rangle - u[B+(1-p_I)C^N,1] \\ \\ +\langle u[B+(1-p_I)C^N,1] - u(\beta,1)\rangle \end{array} \right\} \\
&+ (1-\theta) \left\{ \begin{array}{c} \int_0^\infty \left\langle p_{II} v(\omega, B+C^N) + (1-p_{II}) v(\omega, B)\right\rangle - v[\omega, B+p_{II}C^N] f(\omega) d\omega \\ \\ +\int_0^\infty \left\langle v[\omega, B+p_{II}C^N] - v(\omega,\beta)\right\rangle f(\omega) d\omega \end{array} \right\}
\end{aligned} \tag{2.24}
$$

The two curly braces multiplied by θ and $(1-\theta)$ concern the unable and able subpopulations, respectively. Within both pairs of braces the first line captures *within-group* considerations, whilst the second line captures *between-group* considerations. By the concavity of individual utility the first line must be negative in both cases: i.e. the utility of consuming the average benefit must exceed the expected utility; where $B + (1-p_I)C^N$ is the average benefit income for the unable and $B + p_{II}C^N$ is the average benefit income for the able. Within-group, therefore, an imperfectly targeted programme does worse than a pure universal programme. Next, the sign of the second line will differ between the two pairs of braces whenever $p_I + p_{II} < 1$. In this case the unable receive a higher benefit income on average through targeting than under a pure universal system, whilst the able receive a lower benefit income on average through targeting than under a pure universal system. However, both second lines are instead zero if $p_I + p_{II} = 1$. This arises because a targeted system provides no more benefit income on average to the unable than the able.[28] Consequently, if $p_I + p_{II} = 1$ a

[28] Note that we can write the benefit budget as:

$$
\beta = B + [\theta(1-p_I) + (1-\theta)p_{II}]C = \theta \underbrace{[B+(1-p_I)C]}_{\substack{\text{Average benefit} \\ \text{(unable)}}} + (1-\theta) \underbrace{[B+p_{II}C]}_{\substack{\text{Average benefit} \\ \text{(able)}}}
$$

targeted system does worse than a pure universal system *within-group* and no better *between-group*. This explains why $\hat{C}^N = 0$ when $p_I + p_{II} = 1$. Finally, one can also do a Taylor Approximation of $W^N - W^U$ around β and make use of the budget constraint to obtain:

$$\begin{aligned} W^N - W^U &\approx \theta u_x(\beta, 1) \cdot [B + (1 - p_I)C - \beta] \\ &+ (1-\theta)\int_0^{\infty} v_M(\omega, \beta) f(\omega) d\omega \cdot [B + p_{II}C - \beta] \\ &= \theta(1-\theta)[1 - p_I - p_{II}] \left\{ u_x(\beta, 1) - \int_0^{\infty} v_M(\omega, \beta) f(\omega) d\omega \right\} C \end{aligned}$$

This will be positive so long as $p_I + p_{II} < 1$, which further illustrates that the pure universal welfare level can be improved upon through targeting a categorical benefit at the unable whenever the test administering it has some discriminatory power.

The intuition for Corollary 2 is twofold. First, if Type I errors are made but no universal benefit is provided then unable individuals who have been wrongly denied the categorical benefit would have no source of income to provide consumption. Accordingly, a universal benefit must be provided whenever Type I errors are made. Second, even if there are no Type I errors, the first priority of a benefit system is to support the most needy - in this case those who are unable to work - and a universal benefit should be used only if the budget is above a critical level.

Recall that at the Perfect Discrimination optimum inequality in the average smvi between the unable and able subpopulations will be eliminated *if* the benefit budget is sufficiently large. The following proposition asserts that this no longer holds when classification errors are made.

Proposition 2b: *Let inequality in the average smvi between the unable and able sub-*

So: $B + (1-p_I)C > \beta > B + p_{II}C \; \forall \; p_I + p_{II} < 1$, but $B + (1-p_I)C = \beta = B + p_{II}C \; \forall \; p_I + p_{II} = 1$.

To see that welfare is improved through the unable receiving more on average than the per capita budget (but consequently the able receiving less), note that a Taylor Approximation of $\theta\langle u[B + (1 - p_I)C] - u(\beta, 1)\rangle + (1-\theta)\int\langle v[\omega, B + p_{II}C] - v(\omega, \beta)\rangle f(\omega) d\omega$ around β yields:

$$\theta[B + (1-p_I)C - \beta]\left[u_x(\beta, 1) - \int_0^{\infty} v_M(\omega, \beta) f(\omega) d\omega\right] \geq 0$$

populations be given by:

$$\delta^N \equiv \underbrace{[p_I u_x(B,1) + (1-p_I)u_x(B+C,1)]}_{\text{Average smvi (unable)}} - \underbrace{\int_0^\infty [(1-p_{II})v_M(\omega,B) + p_{II}v_M(\omega,B+C)]f(\omega)d\omega}_{\text{Average smvi (able)}} \tag{2.25}$$

Then if $p_I > 0$ and/or $p_{II} > 0$ it will hold that $\delta^N > 0$ at the optimum because:

$$u_x(\hat{B}^N,1) > u_x(\hat{B}^N+\hat{C}^N,1) \geq \hat{\lambda}^N \geq \int v_M(\omega,\hat{B}^N)f(\omega)d\omega > \int v_M(\omega,\hat{B}^N+\hat{C}^N)f(\omega)d\omega \tag{2.26}$$

where (i) $u_x(\hat{B}^N + \hat{C}^N, 1) = \hat{\lambda}^N$ only if $p_{II} = 0$; whilst (ii) $\int v_M(\omega,\hat{B}^N)f(\omega)d\omega = \hat{\lambda}^N$ only if $p_I = 0$ and β is sufficiently large.

Proof: See Appendix

We now turn our attention to how classification errors affect maximum social welfare, as defined by the value function $V^N(\beta,\theta,p_I,p_{II}) \equiv W^N(\hat{B}^N,\hat{C}^N;\theta,p_I,p_{II})$. Using (2.26) we can establish the following result:

Proposition 2c: *Maximum social welfare is decreasing in the propensity to make classification errors of either Type I or Type II. Formally*:

$$\frac{\partial V^N}{\partial p_I} < 0 \;,\; \frac{\partial V^N}{\partial p_{II}} < 0 \;;\; \forall\, 0 \leq p_I + p_{II} < 1 \tag{2.27}$$

Proof: See Appendix

Proposition 2c implies that improvements in the discriminatory power of the awards technology - corresponding to a *ceteris paribus* reduction in p_I and/or p_{II} - will be unambiguously welfare improving. Suppose, alternatively, that genuine improvements in discriminatory power are not possible, but instead a reduction in p_I (p_{II}) can only be brought about through making the awards test more lenient (stringent), which in turn implies an increase in p_{II} (p_I). There would therefore be a tradeoff between error propensities, discriminatory power held fixed (see Goodin, 1985). The welfare effects of a change in the propensity to make either error type would then depend on the relative size of $|\partial V^N/\partial p_I|$ and $|\partial V^N/\partial p_{II}|$, respectively. In particular, if $|\partial V^N/\partial p_I| > |\partial V^N/\partial p_{II}|$ then an increase in the leniency of the awards test may be welfare improving. Contrastingly, if $|\partial V^N/\partial p_I| < |\partial V^N/\partial p_{II}|$ then an increase in the stringency of the awards test may be welfare improving.

Without adopting restrictive assumptions on the third derivatives u_{xxx} and v_{MMM}, comparative statics results concerning the effect of p_I and p_{II} on the optimal benefit levels are scarce. The exception is the effect of an increase in the Type I error propensity on the optimal universal benefit:

Proposition 2d: $\partial\hat{B}^N/\partial p_I > 0$

Proof: See Appendix

The intuition for this result follows directly from the optimality condition in (2.22). Indeed, consider the effect of an increase in p_I on both $\bar{\sigma}^N_{NR}$ and $\bar{\sigma}^N_R$, when evaluated at $(B, C) = (\hat{B}^N, C^N)$:

$$\frac{d\bar{\sigma}^N_{NR}(\hat{B}^N, C^N; \theta, p_I, p_{II})}{dp_I} = \underbrace{\frac{\partial\bar{\sigma}^N_{NR}}{\partial p_I}}_{>0} + \underbrace{\frac{\partial\bar{\sigma}^N_{NR}}{\partial B}}_{<0} \cdot \frac{\partial\hat{B}^N}{\partial p_I} \tag{2.28}$$

$$\frac{d\bar{\sigma}^N_R(\hat{B}^N, C^N; \theta, p_I, p_{II})}{dp_I} = \underbrace{\frac{\partial\bar{\sigma}^N_R}{\partial p_I}}_{<0} + \underbrace{\frac{\partial\bar{\sigma}^N_R}{\partial B}}_{<0} \cdot \left\{ \frac{\partial\hat{B}^N}{\partial p_I} \underbrace{\left[1 + \frac{\partial C^N}{\partial B}\right]}_{<0} + \underbrace{\frac{\partial C^N}{\partial p_I}}_{>0} \right\} \tag{2.29}$$

where in writing (2.29) we have used the property that $\partial\bar{\sigma}^N_R/\partial B = \partial\bar{\sigma}^N_R/\partial C$. Suppose that $\partial\hat{B}^N/\partial p_I = 0$. Then one can readily establish from (2.28) and (2.29) that $d\bar{\sigma}^N_{NR}/dp_I > 0$ whilst $d\bar{\sigma}^N_R/dp_I < 0$. The average smvi of the two groups thus move in opposite directions and we are no longer at an optimum. Similarly, if we suppose $\partial\hat{B}^N/\partial p_I < 0$ it will also be the case that $d\bar{\sigma}^N_{NR}/dp_I > 0$ whilst $d\bar{\sigma}^N_R/dp_I < 0$. Accordingly, the only way to restore balance (i.e. $\bar{\sigma}^N_{NR} = \bar{\sigma}^N_R$) following an increase in p_I is to lower the average smvi of those non-categorical recipients through increasing B, which explains why $\partial\hat{B}^N/\partial p_I > 0$. Indeed, setting (2.28) equal to (2.29) we can readily establish that:

$$\frac{\partial \hat{B}^N}{\partial p_I} = \frac{\dfrac{\partial \bar{\sigma}^N_{NR}}{\partial p_I} - \left[\dfrac{\partial \bar{\sigma}^N_R}{\partial p_I} + \dfrac{\partial \bar{\sigma}^N_R}{\partial B} \cdot \dfrac{\partial C^N}{\partial p_I}\right]}{\dfrac{\partial \bar{\sigma}^N_R}{\partial B}\left[1 + \dfrac{\partial C^N}{\partial B}\right] - \dfrac{\partial \bar{\sigma}^N_{NR}}{\partial B}} > 0 \tag{2.30}$$

Notice that the result $\partial \hat{B}^N / \partial p_I > 0$ does *not* guarantee that $\partial \hat{C}^N / \partial p_I < 0$. Indeed, because an increase in p_I means that fewer unable individuals receive the categorical benefit we can establish from the budget constraint that:

$$\frac{\partial \hat{C}^N}{\partial p_I} \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \text{ if } \frac{\partial \hat{B}^N}{\partial p_I} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \frac{\theta(\beta - \hat{B}^N)}{[\theta(1-p_I) + (1-\theta)p_{II}]} = \theta \hat{C}^N$$

Remark: *If we make the strong assumption that $v_{MM} = -k \;\forall\; \omega \leq \bar{\omega}(M)$; but $v_{MM} = 0 \;\forall\; \omega > \bar{\omega}(M)$, where $k > 0$ is a constant, then:*

$$\frac{\partial \hat{C}^N}{\partial p_I} < 0$$

This would arise under quadratic preferences $u(x, l) = [x - \alpha x^2] - \gamma(1 - l)$; where α and γ are constants and $k = 2\alpha$. A well documented feature of these preferences is that $u_x > 0$ only if $x < 1/(2\alpha)$. Viard (2001a) similarly assumes that v_{MM} is a negative constant to obtain comparative statics predictions.

The effects of an increase in p_I or p_{II} on the optimal benefit levels are discussed further in the numerical analysis of Section 2.4.

This concludes our discussion of the No Enforcement regime. We now proceed to analyse the Full Enforcement regime.

2.3.3 Imperfect Discrimination: Full Enforcement

When the condition that recipients of C do not work is fully enforced, we know from (2.8) that the only able individuals who apply for C will be those with $\omega \leq \bar{\bar{\omega}}$. For individuals with $\omega > \bar{\bar{\omega}}$ the opportunity cost of not working is simply too high. Given that the unable always apply for C, we can therefore write social welfare under the

Full Enforcement regime as:

$$
\begin{aligned}
W^F(B,C;\theta,p_I,p_{II}) &= \theta\left\{(1-p_I)u(B+C,1)+p_I u(B,1)\right\} \\
&\quad + (1-\theta)\left\{\begin{array}{c} \int_0^\infty v(\omega,B)f(\omega)d\omega \\ \\ +p_{II}\int_0^{\bar{\bar{\omega}}}\langle u(B+C,1)-v(\omega,B)\rangle f(\omega)d\omega \end{array}\right\}
\end{aligned} \tag{2.31}
$$

The first line is the aggregate welfare of unable individuals and is written exactly as under the No Enforcement analysis. The second line, meanwhile, is the aggregate welfare of able individuals. Within the second pair of curly braces the first term captures the average welfare over the able subpopulation were no able individuals to be awarded the categorical benefit, whilst the second term captures the welfare *gain* to able applicants who are awarded the categorical benefit by Type II error.[29]

The government budget constraint is given by:

$$
B + \chi(B,C;\theta,p_I,p_{II})C \leq \beta \tag{2.32}
$$

where $\chi \equiv \theta(1-p_I)+(1-\theta)p_{II}F(\bar{\bar{\omega}})$ denotes the number of categorical recipients in the economy. Conversely, $1-\chi = \theta p_I + (1-\theta)[1-F(\bar{\bar{\omega}})p_{II}]$ denotes the number of individuals not receiving the categorical benefit.

Properties of the budget constraint. Let $C^F(B;\beta,\theta,p_I,p_{II})$ denote the level of the categorical benefit that exhausts the budget constraint for any given $B \in [0,\beta]$. Formally:[30]

[29]To arrive at the expression for welfare in (2.31), note that average welfare over the able subpopulation when the fraction $F(\bar{\bar{\omega}})$ apply for the categorical benefit is:

$$
\begin{aligned}
&\int_0^{\bar{\bar{\omega}}}\left[p_{II}u(B+C,1)+(1-p_{II})v(\omega,B)\right]f(\omega)d\omega + \int_{\bar{\bar{\omega}}}^\infty v(\omega,B)f(\omega)d\omega \\
&= p_{II}\int_0^{\bar{\bar{\omega}}}\left[u(B+C,1)-v(\omega,B)\right]f(\omega)d\omega + \int_0^\infty v(\omega,B)f(\omega)d\omega
\end{aligned}
$$

[30]There must be a unique root $C = C^F$ to the condition $\chi(B,C;\theta,p_I,p_{II})\cdot C = \beta - B$ because the left side is unambiguously increasing in C (and zero when $C=0$) whilst the right side is constant.

$$B + \chi(B, C^F; \theta, p_I, p_{II}) \cdot C^F = \beta \tag{2.33}$$

Differentiating (2.33) w.r.t B thus gives:

$$\frac{\partial C^F}{\partial B} = -\frac{\left[1 + (1-\theta) f(\bar{\bar{\omega}}) p_{II} \bar{\bar{\omega}}_B C^F\right]}{\chi + (1-\theta) f(\bar{\bar{\omega}}) p_{II} \bar{\bar{\omega}}_C C^F} \leq 0 \tag{2.34}$$

From this we can directly establish that:

$$\frac{d\bar{\bar{\omega}}(B, C^F)}{dB} = \bar{\bar{\omega}}_B + \bar{\bar{\omega}}_C \cdot \left(\frac{\partial C^F}{\partial B}\right) = \frac{\chi \bar{\bar{\omega}}_B - \bar{\bar{\omega}}_C}{\chi + (1-\theta) f(\bar{\bar{\omega}}) p_{II} \bar{\bar{\omega}}_C C^F} < 0 \tag{2.35}$$

The total effect of an increase in the universal benefit is therefore to reduce the number of individuals who apply for the categorical benefit. Note that this does not necessarily imply that the total benefit income of categorical recipients falls. This requires $d[B + C^F]/dB = 1 + \partial C^F/\partial B < 0$ and thus $\partial C^F/\partial B < -1$, which in turn corresponds to the condition that:

$$(1-\theta) f(\bar{\bar{\omega}}) p_{II} C^F < (1-\chi) H^*(\bar{\bar{\omega}}, B)$$

For simplicity, we will throughout assume this to be the case.

Finally, to guarantee convexity of the feasible set of choices of B and C it is throughout assumed that $\partial^2 C^F/\partial B^2 \leq 0$, which requires:

$$2 f(\bar{\bar{\omega}}) \frac{d\bar{\bar{\omega}}(B, C^F)}{dB} \frac{\partial C^F}{\partial B} \geq C^F \left\{ \begin{array}{c} f'(\bar{\bar{\omega}}) \left[\dfrac{d\bar{\bar{\omega}}(B, C^F)}{dB}\right]^2 \\ + f(\bar{\bar{\omega}})[\bar{\bar{\omega}}_{BB} + 2\bar{\bar{\omega}}_{BC} C_F' + \bar{\bar{\omega}}_{CC}(\partial C^F/\partial B)^2] \end{array} \right\} \tag{2.36}$$

The left side is unambiguously positive, but the sign of the terms within curly braces on the right side depend on as yet unspecified properties of (i) the distribution function (necessary to sign $f'(\omega)$) and (ii) the third derivatives of individual preferences (necessary to sign $\bar{\bar{\omega}}_{BB}$, $\bar{\bar{\omega}}_{CC}$ and $\bar{\bar{\omega}}_{BC}$). For this reason one cannot conclusively show that $\partial^2 C^F/\partial B^2 \leq 0$. However, the subsequent numerical analysis in Section 2.4 will illustrate that the convexity assumption is readily satisfied in all considered cases.

Optimisation Problem. The optimisation problem of the government is thus now given by:

$$\begin{aligned} &\max_{B,C} W^F(B,C;\theta,p_I,p_{II}) \\ \text{s.t. } \; & B + \chi(B,C;\theta,p_I,p_{II}) \cdot C = \beta \;, \\ & B \geq 0 \;, \quad C \geq 0 \end{aligned} \tag{2.37}$$

We will assume that there is a unique solution to the above optimisation problem. However, the concavity of the welfare function W^F with respect to the choice variables is not guaranteed under Full Enforcement due to the endogeneity of application decisions. We will return to discuss this both below and in the numerical analysis in Section 2.4. The latter will illustrate that this assumption is appropriate in a large range of cases - and identify cases where it is not.

Analogous to the analysis of No Enforcement, let us define the aggregate smvi of non-categorical recipients and categorical recipients by, respectively:

$$\begin{aligned} &\sigma^F_{NR}(B,C;\theta,p_I,p_{II}) = \\ &\theta p_I u_x(B,1) + (1-\theta)\left\{\int_0^\infty v_M(\omega,B)dF(\omega) - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,B)dF(\omega)\right\} \;, \\ &\sigma^F_R(B,C;\theta,p_I,p_{II}) = u_x(B+C,1)\cdot\chi \end{aligned} \tag{2.38}$$

Once more the subscript NR denotes non-categorical recipients, whilst the subscript R denotes categorical recipients.

If we now denote the optimal benefit levels by $\hat{B}^F(\beta,\theta,p_I,p_{II})$ and $\hat{C}^F(\beta,\theta,p_I,p_{II})$, we have:

Proposition 3a: *$\hat{C}^F > 0 \,\forall\, p_I + p_{II} \leq 1$ and $\hat{B}^F \geq 0$ satisfy:*

$$\sigma^F_{NR}\left(\hat{B}^F, \hat{C}^F; \theta, p_I, p_{II}\right)$$

$$\leq \sigma^F_R\left(\hat{B}^F, \hat{C}^F; \theta, p_I, p_{II}\right) \cdot \left\{ \frac{(1-\chi) + (1-\theta)p_{II} f(\bar{\bar{\omega}})(\bar{\bar{\omega}}_B - \bar{\bar{\omega}}_C)\hat{C}^F}{\chi + (1-\theta)p_{II} f(\bar{\bar{\omega}})\bar{\bar{\omega}}_C \hat{C}^F} \right\} \quad , \; \hat{B}^F \geq 0 \tag{2.39}$$

where the pair of inequalities hold with complementary slackness.

Proof: See Appendix.

Corollary 3: *$\hat{B}^F > 0$ if $p_I > 0$, but otherwise $\hat{B}^F > 0$ if:*

$$\sigma^F_{NR}\left(0, C^F; \theta, 0, p_{II}\right)$$

$$> u_x\left(C^F, 1\right) \cdot \chi \left\{ \frac{(1-\chi) + (1-\theta)p_{II} f(\bar{\bar{\omega}})(\bar{\bar{\omega}}_B - \bar{\bar{\omega}}_C)C^F}{\chi + (1-\theta)p_{II} f(\bar{\bar{\omega}})\bar{\bar{\omega}}_C C^F} \right\}$$

The principal message from Proposition 3a is that it is optimal to provide a categorical benefit at *all* levels of discriminatory power, therefore including the case of no discriminatory power. Condition (2.39) characterises the optimum. The left side is the aggregate smvi of individuals *not* receiving the categorical benefit. Contrastingly, the right side is the aggregate smvi of categorical recipients *multiplied* by the increase in their *total* benefit income per unit reduction in the universal benefit. This is made more transparent through the below remark.

Remark: *From (2.34) we can write the terms within curly braces on the right side of (2.39) more compactly as $\cdot - \left(1 + \partial C^F/\partial B\right) > 0$.* [31]

[31] **Unconstrained uni-dimensional problem (Full Enforcement).** Analogous to footnote 26 we could alternatively substitute the function C^F into the welfare function in (2.31) and solve the unconstrained uni-dimensional problem $\max_{B\in[0,\beta]} W^F(B, C^F; \theta, p_I, p_{II})$. The resulting first order condition is:

$$\theta p_I u_x(\hat{B}^F, 1) + (1-\theta)\left\{\int_0^\infty v_M(\omega, \hat{B}^F) f(\omega) d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega, \hat{B}^F) f(\omega) d\omega\right\}$$
$$\leq u_x(\hat{B}^F + C^F, 1) \cdot -(1 + \partial C^F/\partial B)\chi \; ; \; \hat{B}^F \geq 0$$

The increase in total benefit income associated with a reduction in the universal benefit is a function of the terms $\bar{\bar{\omega}}_B$ and $\bar{\bar{\omega}}_C$, capturing the fact that a *ceteris paribus* increase in either benefit generates incentives to apply for the categorical benefit. However, because $\bar{\bar{\omega}}_B < \bar{\bar{\omega}}_C$ a direct implication of (2.39) is that at the optimum the smvi of categorical recipients will exceed the average smvi of non-categorical recipients.

To see why we cannot in general guarantee a unique solution to the optimality condition in (2.39), it is useful to substitute the function C^F into both σ^F_{NR} and σ^F_R and establish how these functions change with the universal benefit. Formally:

$$\begin{aligned}\frac{d\sigma^F_{NR}(B, C^F; \theta, p_I, p_{II})}{dB} &= \underbrace{\theta p_I u_{xx} + (1-\theta)\left\{\int_0^\infty v_{MM} dF(\omega) - p_{II}\int_0^{\bar{\bar{\omega}}} v_{MM} dF(\omega)\right\}}_{<0} \\ &\quad + \underbrace{(1-\theta)p_{II} f(\bar{\bar{\omega}}) \cdot -\frac{d\bar{\bar{\omega}}(B, C^F)}{dB} v_M(\bar{\bar{\omega}}, B)}_{>0} \end{aligned} \tag{2.40}$$

whilst:

$$\frac{d\sigma^F_R(B, C^F; \theta, p_I, p_{II})}{dB} = \chi \cdot \underbrace{\left[u_{xx}(1 + \partial C^F/\partial B)\right]}_{>0} + u_x \cdot \underbrace{\left[(1-\theta)p_{II} f(\bar{\bar{\omega}}) \frac{d\bar{\bar{\omega}}(B, C^F)}{dB}\right]}_{<0} \tag{2.41}$$

The total effect of an increase in the universal benefit on σ^F_{NR} and σ^F_R is in both cases composed of two *conflicting* effects. In the case of σ^F_{NR} an increase in the universal benefit (i) increases the benefit income of each individual and thus lowers their individual smvi; *but* also (ii) increases the number of non-categorical recipients through reducing the number of individuals that apply for the categorical benefit. Alternatively, in the case of σ^F_R an increase in the universal benefit (i) lowers the total benefit income of each individual and consequently increases their smvi; *but* also (ii) reduces the number of categorical recipients. Notice that in either case the size of the second effect will depend on the responsiveness of $\bar{\bar{\omega}}$ to changes in benefit income. *Ceteris paribus*, the less responsive this critical net wage is the lower the absolute size of this effect. In

Substituting in (2.34) for $\partial C^F/\partial B$ then gives the expression in (2.39).

both cases the overall effect is ambiguous and will depend on unspecified properties of both the utility function (e.g. third derivatives) and the distribution function (e.g. the derivative of the pdf). These conflicting effects can result in non-concavity of the welfare function and multiple stationary points.[32] We will discuss this in greater depth in the numerical analysis of Section 2.4.

The conditions under which a universal benefit should be provided follow directly from (2.39) and are given in Corollary 3. As under the No Enforcement analysis, it is optimal to provide a universal benefit whenever the test administering the categorical benefit has a positive propensity to make Type I errors. This ensures that unable individuals who are incorrectly denied the categorical benefit have some source of income to consume. Contrastingly, if Type I errors never occur it is only optimal to provide a universal benefit if, when evaluated at $(B, C) = (0, C^F)$, the total welfare gain of a marginal increase in the universal benefit to non-categorical recipients exceeds the total welfare loss to categorical recipients.

Intuition for Proposition 3a. To provide the intuition for the main result in Proposition 3a, which states that a categorical benefit should be provided no matter what the discriminatory power of the test administering it, we take a Taylor approximation of $W^F - W^U$ around β to obtain:

$$W^F - W^U \approx (1-\theta)C \left\{ \begin{array}{l} \theta[1 - p_I - F(\bar{\bar{\omega}})p_{II}] \left\langle u_x(\beta, 1) - \int_0^\infty v_M(\omega, \beta) f(\omega) d\omega \right\rangle \\ \\ + p_{II} F(\bar{\bar{\omega}}) \left\langle \dfrac{F[\bar{\omega}(\beta)]}{F(\bar{\bar{\omega}})} u_x(\beta, 1) - \int_0^\infty v_M(\omega, \beta) f(\omega) d\omega \right\rangle \end{array} \right\} \tag{2.42}$$

The first line within curly braces is always positive because a targeted system awards, on average, the unable with more benefit income than the able for *all* levels of discrim-

[32] It is also useful to note that:

$$\frac{d}{dB}\left\{\sigma_R^F(B, C^F; \theta, p_I, p_{II}) \cdot -(1 + \partial C^F/\partial B)\right\}$$
$$= \chi \cdot \underbrace{\left\{u_{xx} \cdot -(1 + C^F)^2 - u_x \partial^2 C^F/\partial B^2\right\}}_{>0} + u_x \cdot -(1 + \partial C^F/\partial B) \underbrace{\left\{(1-\theta) f(\bar{\bar{\omega}}) p_{II} \frac{d\bar{\bar{\omega}}(B, C^F)}{dB}\right\}}_{<0}$$

The assumption that $\partial^2 C^F/\partial B^2 < 0$ thus acts to compound the effect of a reduction in total benefit income on a categorical recipient's smvi.

inatory power. This arises because able individuals of high productivity (i.e. $\omega > \bar{\bar{\omega}}$) choose not to apply for C. Next, the second line within curly braces is likely to be positive for a small categorical benefit where $\bar{\bar{\omega}} \sim \bar{\omega}(\beta) \Rightarrow F(\bar{\omega}) \sim F(\bar{\bar{\omega}})$. This line captures the fact that Type II errors actually redistribute *within* the able subpopulation because only those with $\omega \leq \bar{\bar{\omega}}$ will apply for the benefit. Indeed, the closer $F(\bar{\omega})/F(\bar{\bar{\omega}})$ to unity the larger this effect is and this arises precisely because able individuals with $\omega < \bar{\omega}$ are formally equivalent to unable individuals from the perspective of smvi.

In terms of benefit design then, the principal difference between the No Enforcement and Full Enforcement regimes is that under the latter a targeted system (pure or partial) is chosen in all cases, whereas under the former a targeted system (pure or partial) is only chosen if the test administering C has some discriminatory power.

From the first order conditions characterising the optimal benefit levels (see Appendix) we can establish the following result.

Proposition 3b: *Let inequality in the average smvi between the unable and able subpopulations be written as:*

$$\delta^F = \underbrace{[p_I u_x(B,1) + (1-p_I)u_x(B+C,1)]}_{\text{Average smvi (unable)}} - \underbrace{\left\{\int_0^\infty v_M(\omega,B)f(\omega)d\omega + p_{II}\int_0^{\bar{\bar{\omega}}} \left[u_x(B+C,1) - v_M(\omega,B)\right] f(\omega)d\omega\right\}}_{\text{Average smvi (able)}} \tag{2.43}$$

Then for $p_I > 0$ and/or $p_{II} > 0$ it will hold that $\delta^F > 0$ at the optimum because:

$$\frac{1}{F(\bar{\bar{\omega}})}\int_0^{\bar{\bar{\omega}}} v_M(\omega,\hat{B}^F)f(\omega)d\omega > u_x(\hat{B}^F+\hat{C}^F,1) \geq \hat{\lambda}^F > \frac{1}{1-F(\bar{\bar{\omega}})}\int_{\bar{\bar{\omega}}}^\infty v_M(\omega,\hat{B}^F)f(\omega)d\omega \tag{2.44}$$

where $u_x(\hat{B}^F + \hat{C}^F, 1) = \hat{\lambda}^F$ only if $p_{II} = 0$.

Proof: See Appendix

Proposition 3b states that classification errors of either type prevent the elimination of between-group inequality in the average smvi, even if the benefit budget would be sufficiently large to eliminate the inequality under Perfect Discrimination.

We now turn to discuss how classification errors affect maximum social welfare, as defined by the value function $V^F(\beta, \theta, p_I, p_{II}) \equiv W^F(\hat{B}^F, \hat{C}^F; \theta, p_I, p_{II})$. From (2.44) we can write the following result.

Proposition 3c: *Maximum social welfare is decreasing in the propensity to make Type I classification errors, but can be increasing in the propensity to make Type II classification errors. Formally:*

$$\frac{\partial V^F}{\partial p_I} < 0 \ , \ \ \frac{\partial V^F}{\partial p_{II}} \begin{Bmatrix} \leq \\ > \end{Bmatrix} 0 \tag{2.45}$$

where: (i)

$$\frac{\partial V^F}{\partial p_{II}} = (1-\theta) F(\bar{\bar{\omega}}) \cdot$$
$$\left\{ \underbrace{\left[u(\hat{B}^F + \hat{C}^F, 1) - u(\hat{B}^F, 1) - \hat{\lambda}^F \hat{C}^F \right]}_{\substack{\text{Benefit to individual of receiving } C \\ \text{conditional on facing zero labour} \\ \text{constraint (net of exchequer cost)}}} - \underbrace{\int_{\bar{\omega}(B)}^{\bar{\bar{\omega}}} \left[v(\omega, \hat{B}^F) - u(\hat{B}^F, 1) \right] \frac{f(\omega)}{F(\bar{\bar{\omega}})} d\omega}_{\substack{\text{Cost of facing zero labour constraint} \\ \text{conditional on receiving } B}} \right\}$$

and (ii) a sufficient condition for welfare to be increasing in Type II errors is that:

$$\frac{F(\bar{\omega}(B))}{\dfrac{\theta(1-p_I)}{(1-\theta)p_{II}} + F(\bar{\bar{\omega}})} \cdot \underbrace{\frac{\bar{\bar{\omega}} f(\bar{\bar{\omega}})}{F(\bar{\bar{\omega}})}}_{\substack{\text{elasticity of } F \text{ w.r.t. } \omega, \\ \text{evaluated at } \bar{\bar{\omega}}}} \cdot \underbrace{\frac{\hat{C}^F \bar{\bar{\omega}}_C}{\bar{\bar{\omega}}}}_{\substack{\text{elasticity of } \bar{\bar{\omega}} \text{ w.r.t. } C \\ \text{evaluated at } \hat{C}^F}} - \left[1 - \frac{F(\bar{\omega}(B))}{F(\bar{\bar{\omega}})} \right] > 0 \tag{2.46}$$

Proof: See Appendix

We can make a number of observations from (2.46). First, the closer $F(\bar{\omega}(B))/F(\bar{\bar{\omega}})$ is to unity the more likely that $\partial V^F/\partial p_{II} > 0$. This arises precisely because the categorical benefit is being 'leaked' predominantly to able applicants who have the same smvi as unable applicants. Second, the higher p_I and/or p_{II} the larger the first term in (2.46) and thus the more likely that $\partial V^F/\partial p_{II} > 0$. We will see from the numerical examples which follow that an increase in the propensity to make either error type is likely to lower the categorical benefit, which in turn will increase $\bar{\omega}$ relative to $\bar{\bar{\omega}}$. Finally, the smaller θ is - and thus the fewer unable individuals in society - the more likely that $\partial V^F / \partial p_{II} > 0$.

2.3.4 Welfare Comparison: No Enforcement vs. Full Enforcement

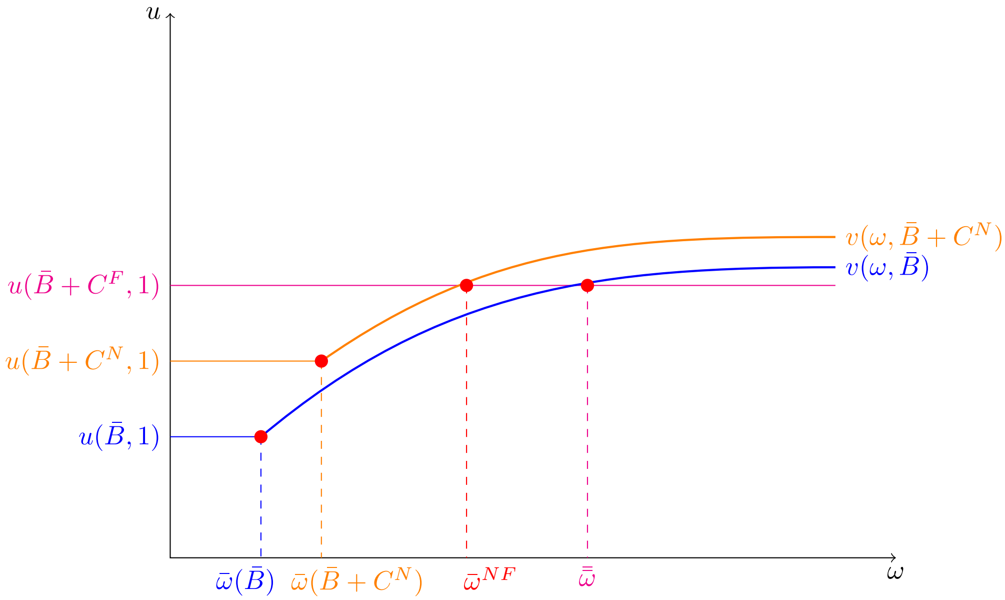
Given our analysis of optimal welfare provision under the alternative enforcement regimes of No Enforcement and Full Enforcement, it is natural to ask under which regime is social welfare highest? Clearly, if $p_{II} = 0$ then there are zero enforcement issues and welfare under both regimes will coincide. Contrastingly, if $p_I + p_{II} = 1$ then it must hold that $V^F > V^N = W^U$ because the pure universal outcome is not chosen under the Full Enforcement regime. We are therefore concerned with the intermediate cases where $0 < p_I + p_{II} < 1$ *and* $p_{II} > 0$.

Suppose that we fix B at any budget feasible value $\bar{B} \in (0, \beta)$. From (2.19) and (2.33) the resulting categorical benefit sizes under the alternative enforcement regimes are then C^N and C^F, respectively. For a finite benefit budget it must hold that $F(\bar{\bar{\omega}}) < 1$ and so $C^N < C^F$.

For a given $\bar{B}$ there will be a critical productivity at which the welfare of a categorical recipient under No Enforcement equates with the welfare of a categorical recipient under Full Enforcement. Formally, let $\bar{\omega}^{NF}$ satisfy $u(\bar{B} + C^F, 1) \equiv v\left(\bar{\omega}^{NF}, \bar{B} + C^N\right)$.[33] This critical net wage is illustrated graphically in Figure 2.3. A comparison of the two welfare regimes $(\bar{B}, C^N)$ and $(\bar{B}, C^F)$ thus yields:

[33] Note that $\forall\, \omega < \bar{\omega}(\bar{B} + C^N)$ we have $v(\omega, \bar{B} + C^N) = u(\bar{B} + C^N, 1) < u(\bar{B} + C^F, 1)$. There will therefore be a critical productivity $\bar{\omega}^{NF}$ as asserted in the main text.

Figure 2.3: Welfare comparison and the critical productivity $\bar{\omega}^{NF}$.



Notes: From the relationship $v(\omega,\bar{B}+C^N)=u(\bar{B}+C^F,1)=v(\bar{\bar{\omega}},\bar{B})$ it must hold that $\bar{\omega}^{NF}<\bar{\bar{\omega}}$

$$
\begin{aligned}
& W^F(\bar{B},C^F;\theta,p_I,p_{II}) - W^N(\bar{B},C^N;\theta,p_I,p_{II}) \\
&= \theta(1-p_I)\big[\underbrace{u(\bar{B}+C^F,1)-u(\bar{B}+C^N,1)}_{>0}\big] \\
&+ (1-\theta)p_{II}\left\{\begin{array}{c}
\int_0^{\bar{\omega}^{NF}}\underbrace{\big[u(\bar{B}+C^F,1)-v(\omega,\bar{B}+C^N)\big]}_{>0}f(\omega)d\omega \\
+\int_{\bar{\omega}^{NF}}^{\bar{\bar{\omega}}}\underbrace{\big[u(\bar{B}+C^F,1)-v(\omega,\bar{B}+C^N)\big]}_{<0}f(\omega)d\omega \\
+\int_{\bar{\bar{\omega}}}^{\infty}\underbrace{\big[v(\omega,\bar{B})-v(\omega,\bar{B}+C^N)\big]}_{<0}f(\omega)d\omega
\end{array}\right\}
\end{aligned} \tag{2.47}
$$

[34]The first line concerns the unable subpopulation and is unambiguously positive. Un-

[34]It is important to stress that we are not comparing the value functions for No Enforcement and Full Enforcement here. However, we can always (and will below) set $\bar{B}=\hat{B}^N$ and compare an arbitrary Full Enforcement system with the optimum No Enforcement system (or vice versa). If there are

able individuals are on aggregate better of under Full Enforcement than No Enforcement because (i) those awarded the categorical benefit receive $C^F > C^N$; whilst (ii) those not awarded the categorical benefit receive the same as they would under No Enforcement (i.e. $\bar{B}$) and are thus no worse off. The second line concerns the able subpopulation: as can be seen from the terms within curly braces there are three considerations. First, categorical recipients with $\omega \leq \bar{\omega}^{NF}$ are by definition better off under Full Enforcement than under No Enforcement. Within this range of wages notice that recipients with $\bar{\omega}(\bar{B} + C^F) < \omega \leq \bar{\bar{\omega}}$ would like to work under the Full Enforcement regime, but they would not be willing to accept a reduction of $(C^F - C^N)$ in their benefit income in return for a full relaxation of the no-work requirement. Second, however, categorical recipients with $\bar{\omega}^{NF} < \omega \leq \bar{\bar{\omega}}$ are worse off under Full Enforcement than under No Enforcement. These individuals would be willing to accept the lower categorical benefit size in return for a full relaxation of the no-work requirement. Third, individuals with $\omega > \bar{\bar{\omega}}$ do not apply for the categorical benefit under Full Enforcement because they are better off working and receiving only the universal benefit. They are therefore worse off than their counterparts under No Enforcement who work *and* receive the categorical benefit.

Suppose that we set $\bar{B} = \hat{B}^N$ (and consequently $C^N = \hat{C}^N$). In this case (2.47) becomes a comparison of (i) welfare under an arbitrary Full Enforcement system where $B = \hat{B}^N$ and $C > \hat{C}^N$; with (ii) *maximum* welfare under No Enforcement. An arbitrary Full Enforcement system thus does better than an *optimal* No Enforcement system for both the unable subpopulation and, within the able subpopulation, individuals of lower productivity (i.e. $\omega \leq \bar{\omega}^{NF}$). If the welfare losses to those with higher productivities (i.e. $\omega > \bar{\omega}^{NF}$) do not offset the redistribute gains from awarding higher transfers to the unable and those of lower productivity, it will certainly be the case that a Full Enforcement scheme with *optimally* chosen welfare benefits can improve upon a No Enforcement scheme.

We shall revisit this question in the numerical analysis which follows.

---

conditions where the former exceeds the latter then there will certainly be conditions where optimal Full Enforcement welfare exceeds optimal No Enforcement welfare.

2.4 Numerical Simulations

The purpose of this section is twofold: (i) to obtain insights into how the optimal benefit levels change with the propensity to make classification errors of either type; and (ii) to provide examples where welfare under the Full Enforcement regime can indeed be increasing in the propensity to make Type II classification errors. To satisfy both objectives, we turn to numerical methods.

In line with much of the optimal tax/benefit literature (Immonen et al., 1998; Mirrlees, 1971; Stern, 1982; Viard, 2001a,b), we take individual preferences to be of the constant elasticity of substitution (CES) form:

$$u(x,l) = \begin{cases} \left[\alpha x^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha) l^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right]^{\frac{\mathcal{E}}{\mathcal{E}-1}} & : \mathcal{E} \neq 1 \\ x^{\alpha} l^{1-\alpha} & : \mathcal{E} = 1 \end{cases} \tag{2.48}$$

where $\mathcal{E}$ is elasticity of substitution between leisure and consumption. The properties of this function are well documented in the literature (see Appendix for the derivation of labour supply and indirect utility.).

The choice of productivity distribution is influenced by the sufficient condition in (2.46), which is more likely to hold the closer $F(\bar{\omega})/F(\bar{\bar{\omega}}) < 1$ is to unity. This suggests that a productivity distribution with a thick lower tail may give rise to conditions where welfare can be increasing in the propensity to make Type II errors. There are a number of candidates that satisfy this criterion; such as the Exponential Distribution or the Pareto Distribution. The former does not require the imposition of a minimum productivity and so we choose this. Individual productivities are therefore exponentially distributed with density function:

$$f(n) = \mu e^{-\mu n} \tag{2.49}$$

and distribution function

$$F(n) = 1 - e^{-\mu n} \tag{2.50}$$

where $\mu > 0$ is a scale parameter and $n = \omega/(1-t)$ is individual productivity.

We note at the outset that many of the results obtained in this section concerning the effect of classification errors on optimal benefits continue to hold under alternative distributional assumptions.[35]

In the simulations that follow the choices of parameter values are: $\mu = 3$; $\mathcal{E} \in \{0.5, 0.6, 0.7\ 0.8, 0.9, 0.99\}$; $\alpha = 0.6$; $\theta \in \{0.05, 0.1\}$; $t = 0.3$; and $\beta = 0.1$. For $\mu = 3$ the average productivity (maximum earnings) in the economy is 1/3. Analogous to Stern (1976), the parameter $\alpha = 0.6$ is chosen such that the average productivity individual works roughly two-thirds of their time endowment (unity) when $\mathcal{E} = 0.5$. The choices of the unable subpopulation size (θ) seems sensible in light of real-world statistics. For example, Mcinnes (2012, p.4) reports that 6.7% of those in the United Kingdom aged 16-64 (i.e. working age) claimed Incapacity Benefit/ Earnings and Support Allowance (ESA) in 2011. Note that higher values of θ would also be permissible[36]: Atkinson (1995, Ch.2) - who refers to those who are unable to work as the 'sick and retired' - considers $\theta = 0.15$ (and higher). Finally, the chosen budget size of $\beta = 0.1$ corresponds to 30% of maximum earnings. This value is slightly larger than tax revenue at $t = 0.3$ (when able individuals receive no unearned income), but the results do not qualitatively change at lower budget sizes. Indeed, examples at the lower budget size $\beta = 0.05$ are available from the author upon request.

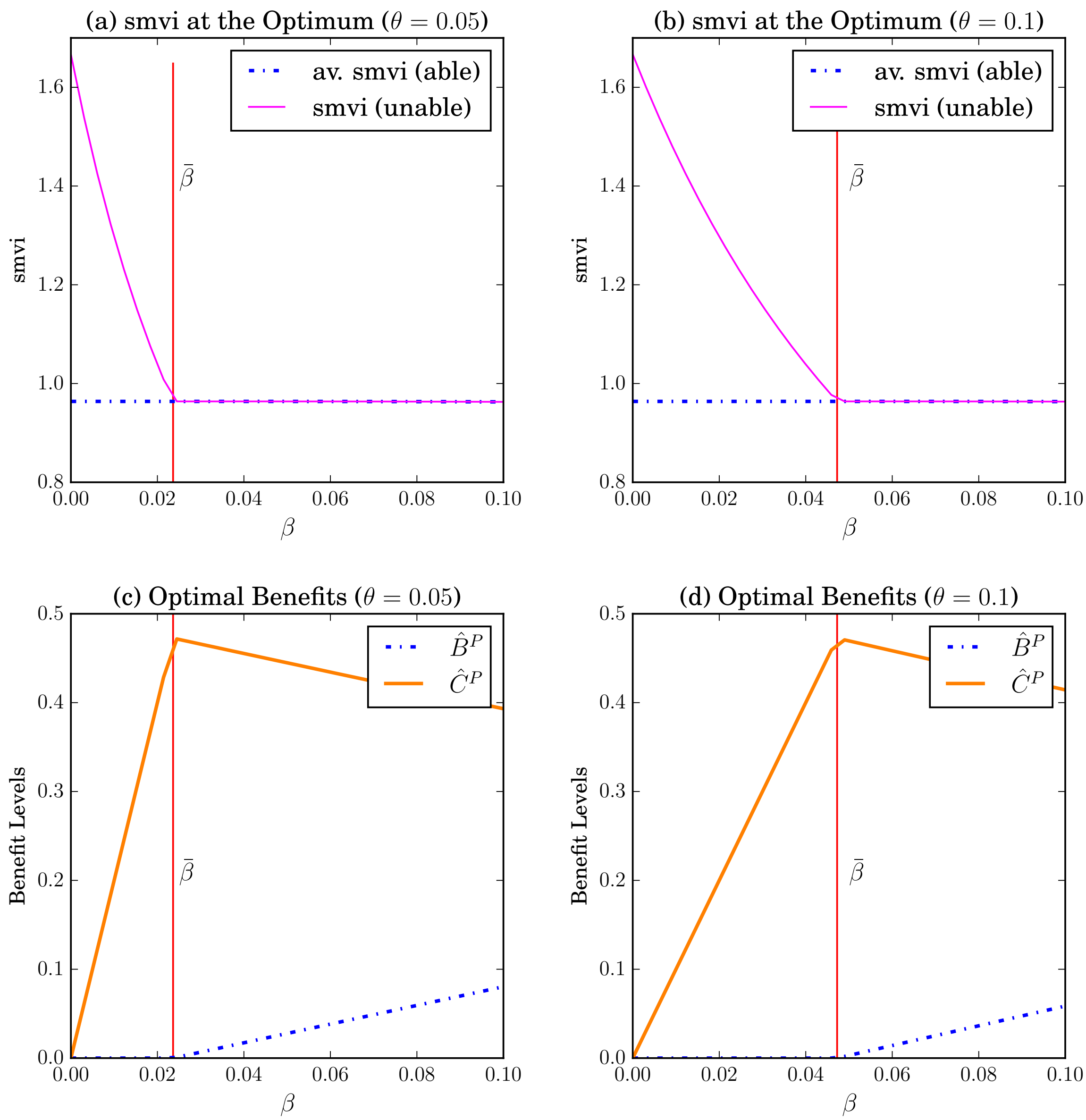
2.4.1 Baseline Case: Perfect Discrimination Simulations

To provide baseline numerical results, Figures 2.4 and 2.5 display the Perfect Discrimination optima for the cases where $\mathcal{E} = 0.5$ and $\mathcal{E} = 0.99$, respectively. Both figures contain four subplots. The subplots on the left side are drawn for $\theta = 0.05$, whilst the subplots on the right side are drawn for $\theta = 0.10$. Each subplot has the budget size β on the horizontal axis. Subplots (a) and (b) illustrate how the average smvi of the two subpopulations change with β, where for each value of β the optimal benefit levels are chosen. Subplots (c) and (d) illustrate how the optimal benefit levels change with β.

[35]Simulation results under the frequently employed assumption that productivity is lognormally distributed are available from the author upon request.

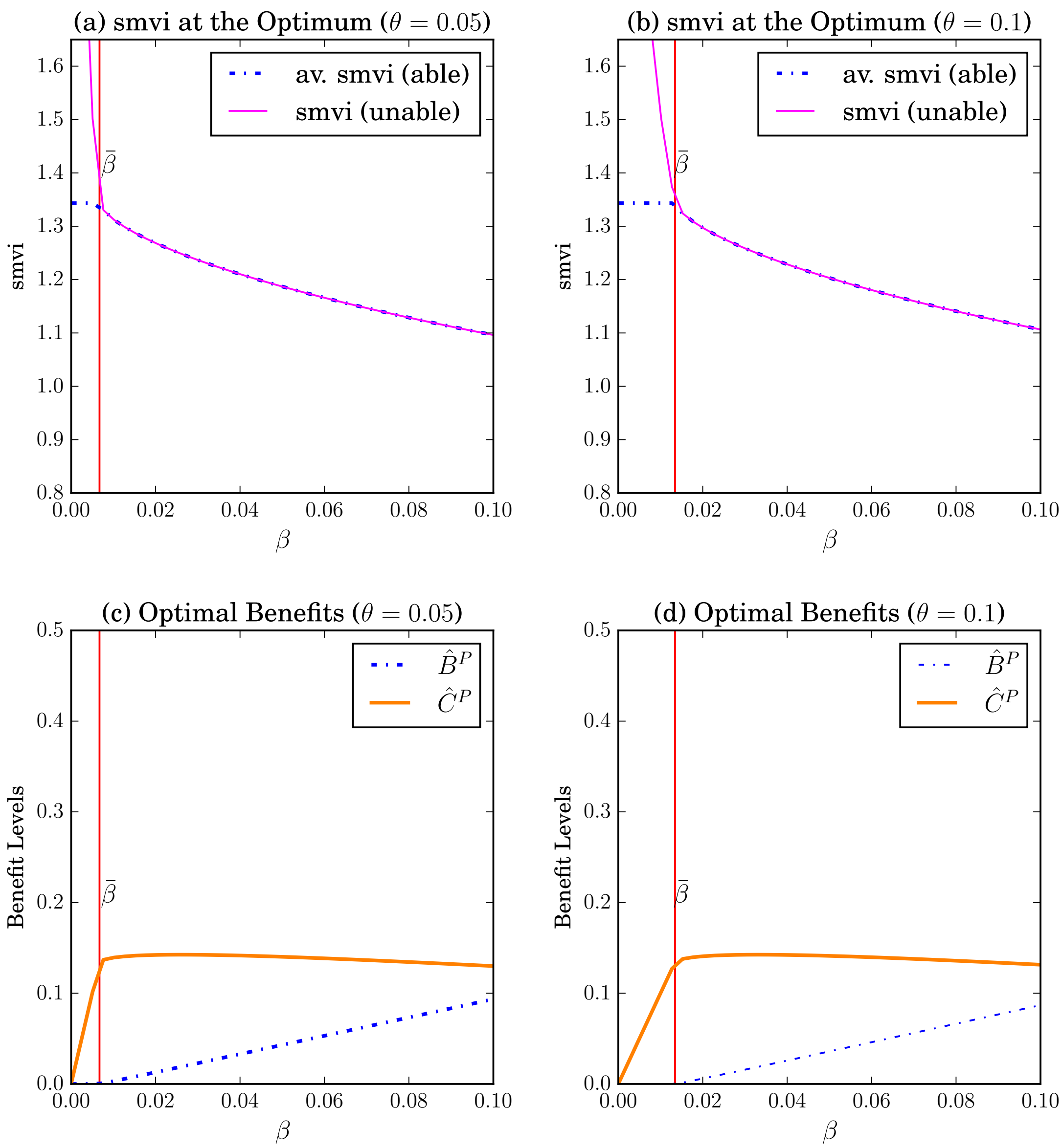
[36]To put these values further in context, the UK population size was estimated at 63.2 million in 2011/12, with 66 percent (42.1 million) of working age. Of these, approximately 13 percent (5.7 million) were estimated to be disabled, where the disability definition adopted includes, non-exhaustively, those with a long term illness or substantial difficulties engaging in daily activities (see ONS, 2012; ODI, 2014).

Figure 2.4: Perfect discrimination optima and average smvi ($\mathcal{E} = 0.5$)



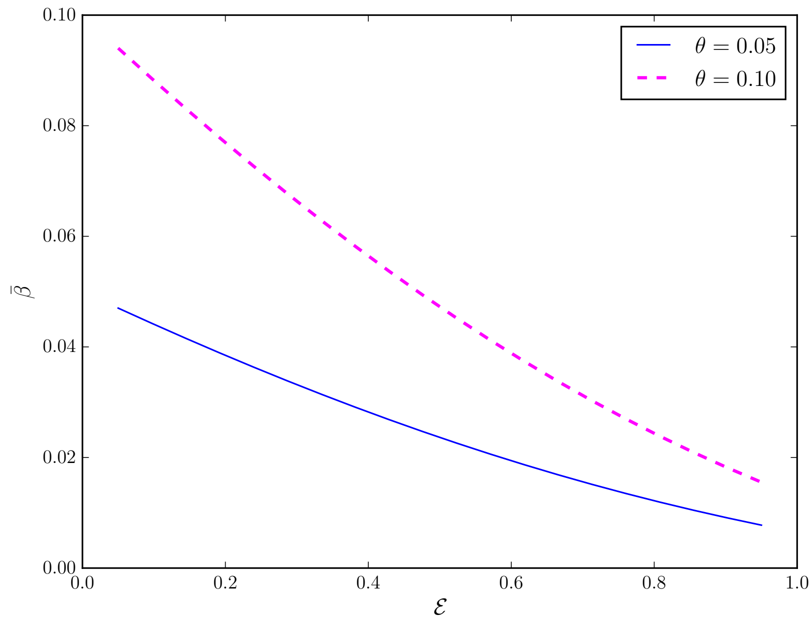
Notes. This figure captures how the average smvi of the unable and able subpopulations changes with the budget size when the benefits are optimally chosen. The subplots on the left side are drawn for an unable subpopulation size of $\theta = 0.05$; whilst the subplots on the right side are drawn for $\theta = 0.10$. $\bar{\beta}$ is the critical budget level satisfying $u_x\left(\bar{\beta}/\theta, 1\right) \equiv \int v_M(\omega, 0) f(\omega) d\omega$. Whenever $\beta < \bar{\beta}$ we have $\hat{B}^P = 0$ whilst $\hat{C}^P > 0$ increases with β. At $\beta = \bar{\beta}$ - and thus $\hat{C}^P = \bar{\beta}/\theta$ - categorical spending eliminates inequality in the average smvi between the unable and able subpopulations. For $\beta > \bar{\beta}$ we have $\hat{B}^P > 0$ and $\hat{C}^P > 0$, such that a combination of the universal and categorical transfers are used to keep the average smvi of the two subpopulations equal.

Figure 2.5: Perfect discrimination optima and the average smvi ($\mathcal{E} = 0.99$)



Notes. A comparison with Figure 2.4 illustrates that $\bar{\beta}$ is lower when $\mathcal{E} = 0.99$ than when $\mathcal{E} = 0.5$.

Figure 2.6: Variation of $\bar{\beta}$ with $\mathcal{E}$.



Notes: On the horizontal axis is the elasticity of substitution between leisure and consumption ($\mathcal{E}$), whilst on the vertical axis is the budget size ($\bar{\beta}$) required to eliminate inequality in the average smvi between the unable and able subpopulations, when spent entirely on categorical transfers.

The vertical line labelled $\bar{\beta}$ corresponds to $\bar{\beta}^P$ in (2.15). It is thus the critical budget size at which there are just enough resources to eliminate inequality in the average smvi between the unable and able subpopulations through categorical transfers.

Both figures illustrate that whenever $\beta < \bar{\beta}$ the smvi of the unable exceeds the average smvi of the able at the optimum. In this case the optimal universal benefit is zero and the optimal categorical benefit is increasing in the budget size. Increasing categorical transfers lower the smvi of the unable towards the average smvi of the able. Consequently, the smvi of the unable is falling with the budget size. At $\beta = \bar{\beta}$ the average smvi of the two subpopulations is equated. For $\beta > \bar{\beta}$ it is optimal to provide a universal benefit, where *both* the universal and categorical benefits are chosen to keep the average smvi of the two subpopulations equated.

Comparing the left and right columns within both figures illustrates, unsurprisingly, that an increase in the unable subpopulation size increases the critical budget $\bar{\beta}$. This is easily verified: differentiating the identity $u_x(\bar{\beta}/\theta, 1) \equiv \int v_M(\omega, 0) f(\omega) d\omega$ with respect to θ gives $\partial\bar{\beta}/\partial\theta = \bar{\beta}/\theta > 0$. Comparing Figure 2.4 with Figure 2.5 illustrates that an increase in the elasticity of substitution between leisure and consumption *increases* the universal benefit and *lowers* the categorical transfer. This arises because an increase in $\mathcal{E}$ lowers inequality in the average smvi between the two subpopulations. This can be seen from the fact that, *ceteris paribus*, $\bar{\beta}$ falls with $\mathcal{E}$, as illustrated in Figure 2.6.

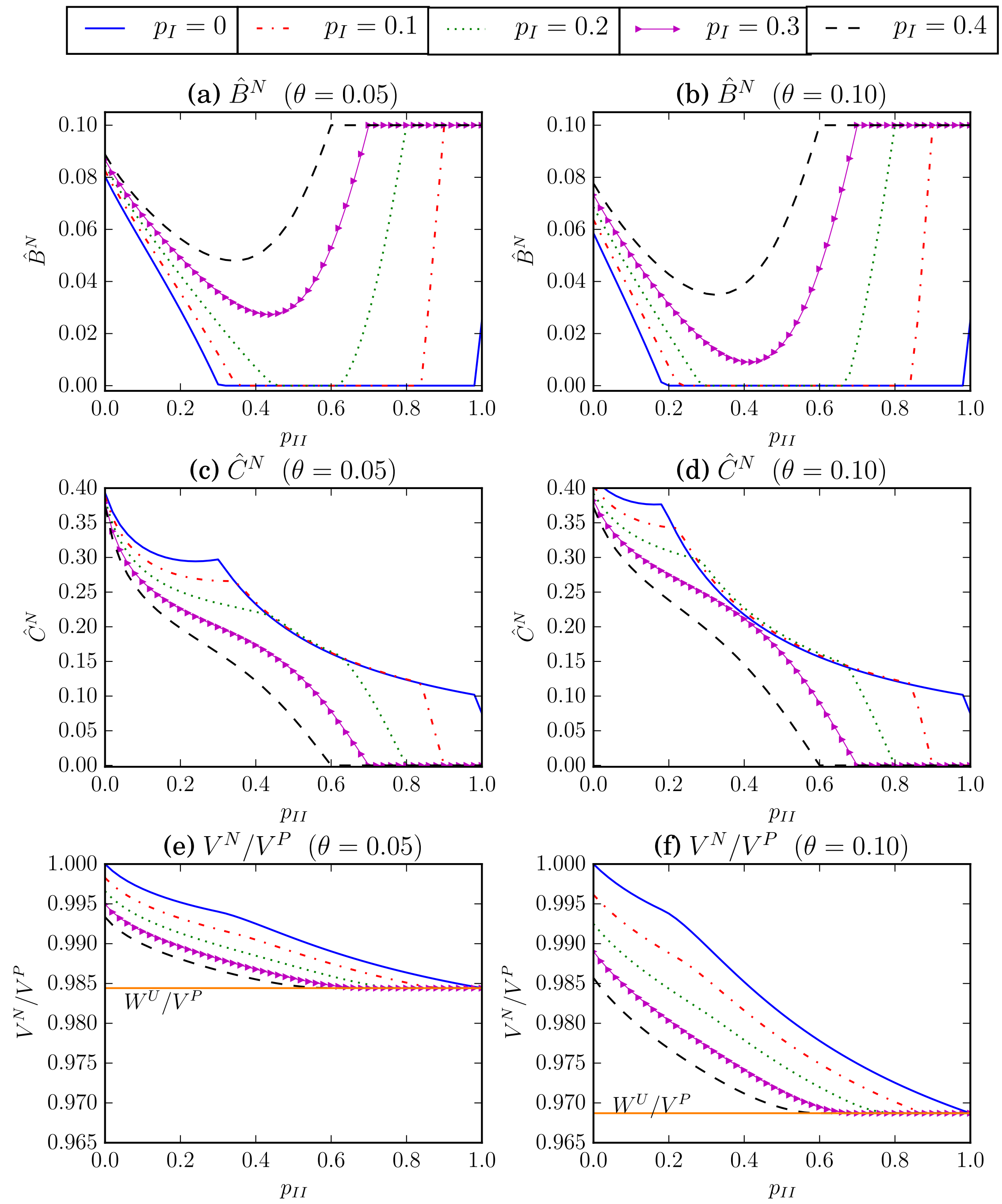
2.4.2 No Enforcement Simulations

Figures 2.7 to 2.12 display how the No Enforcement optimum benefit levels and maximum social welfare change with the propensity to make both Type I and Type II errors. Each figure is drawn for a different $\mathcal{E} \in \{0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$. Within each figure there are six subplots. The subplots on the left side (a,c, and e) are drawn for $\theta = 0.05$; whilst those on the right side (b,d and f) are drawn for $\theta = 0.10$. Subplots (a) and (b) display the optimal universal benefit; subplots (c) and (d) display the optimal categorical benefit; whist subplots (e) and (f) display maximum welfare under these choices as a proportion of welfare under the Perfect Discrimination optimum (this ensures values are between 0 and 1). On the horizontal axis in each subplot is the propensity to make Type II errors; varied in between 0 and 1. The different curves within a given subplot are drawn for different Type I error propensities, where $p_I \in \{0, 0.1, 0.2, 0.3, 0.4\}$. To generate each subplot p_{II} was increased from 0 to 1 in discrete intervals of 0.02 and at each stage the optimal benefit levels were simulated.

The most immediate observation from each figure is that the simulation results are consistent with the underlying theory in Proposition 2a. At the point of no discriminatory power (i.e. $p_I + p_{II} = 1$) we observe that: (i) if $p_I > 0$ then $\hat{C}^N$ is set at zero whilst $\hat{B}^N$ is set at the *per capita* budget size β; but (ii) if $p_I = 0$ any combination of C and B that satisfy the budget constraint will be optimal. The intuition is clear: when Type I errors are never made but Type II errors are always made the categorical transfer is effectively administered as a universal transfer because it is received by *all* individuals in society. It is useful to note, however, that setting p_I at a very low value such as 0.0001 will guarantee that $\hat{B}^N = \beta$ and $\hat{C}^N = 0$ at the point of no discriminatory power.

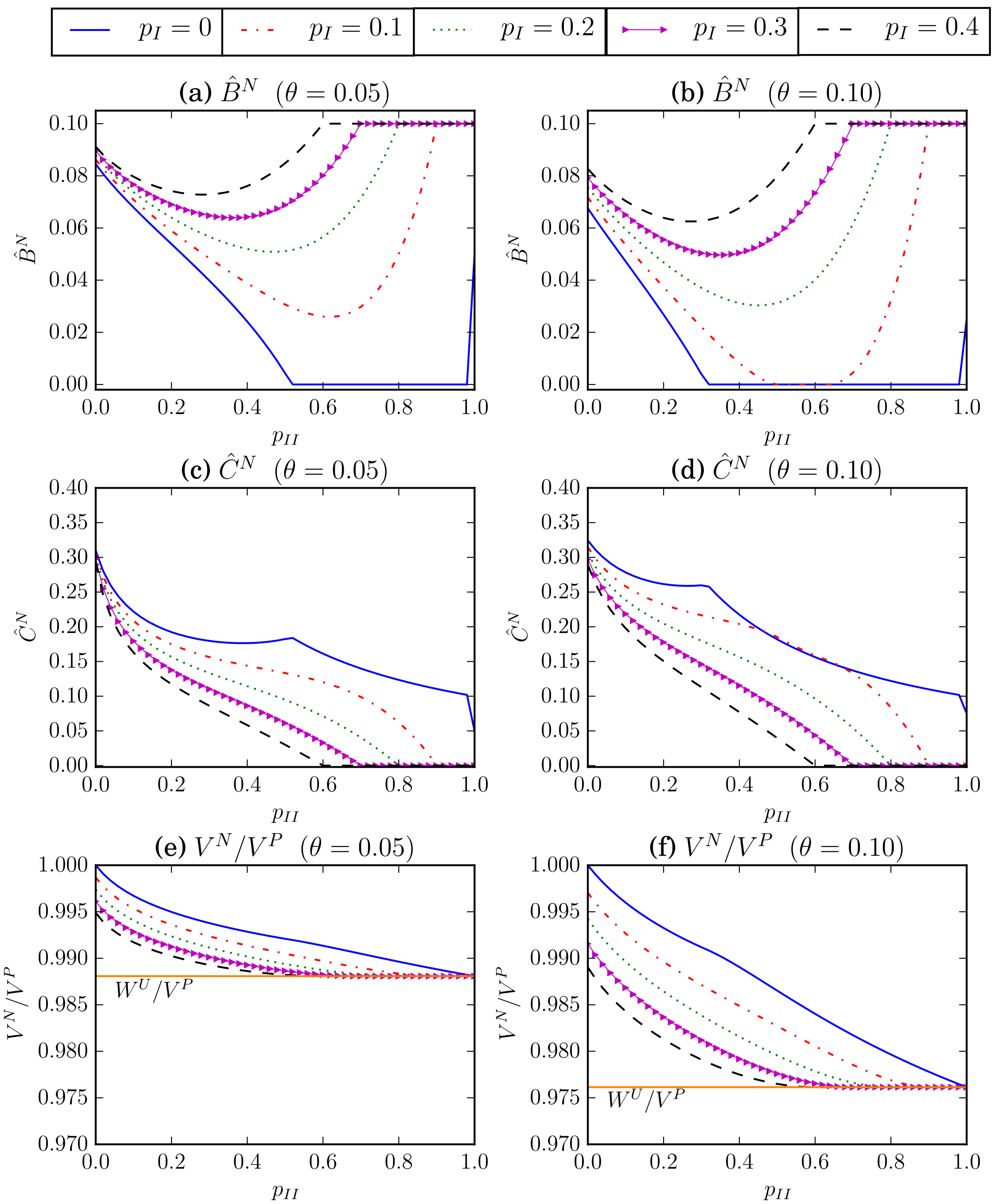
Within each figure, subplots (a) to (d) illustrate that a *ceteris paribus* increase in p_I increases $\hat{B}^N$ and tends to decrease $\hat{C}^N$. The observation that $\hat{B}^N$ increases with p_I

Figure 2.7: No Enforcement Optima for $\mathcal{E} = 0.5$



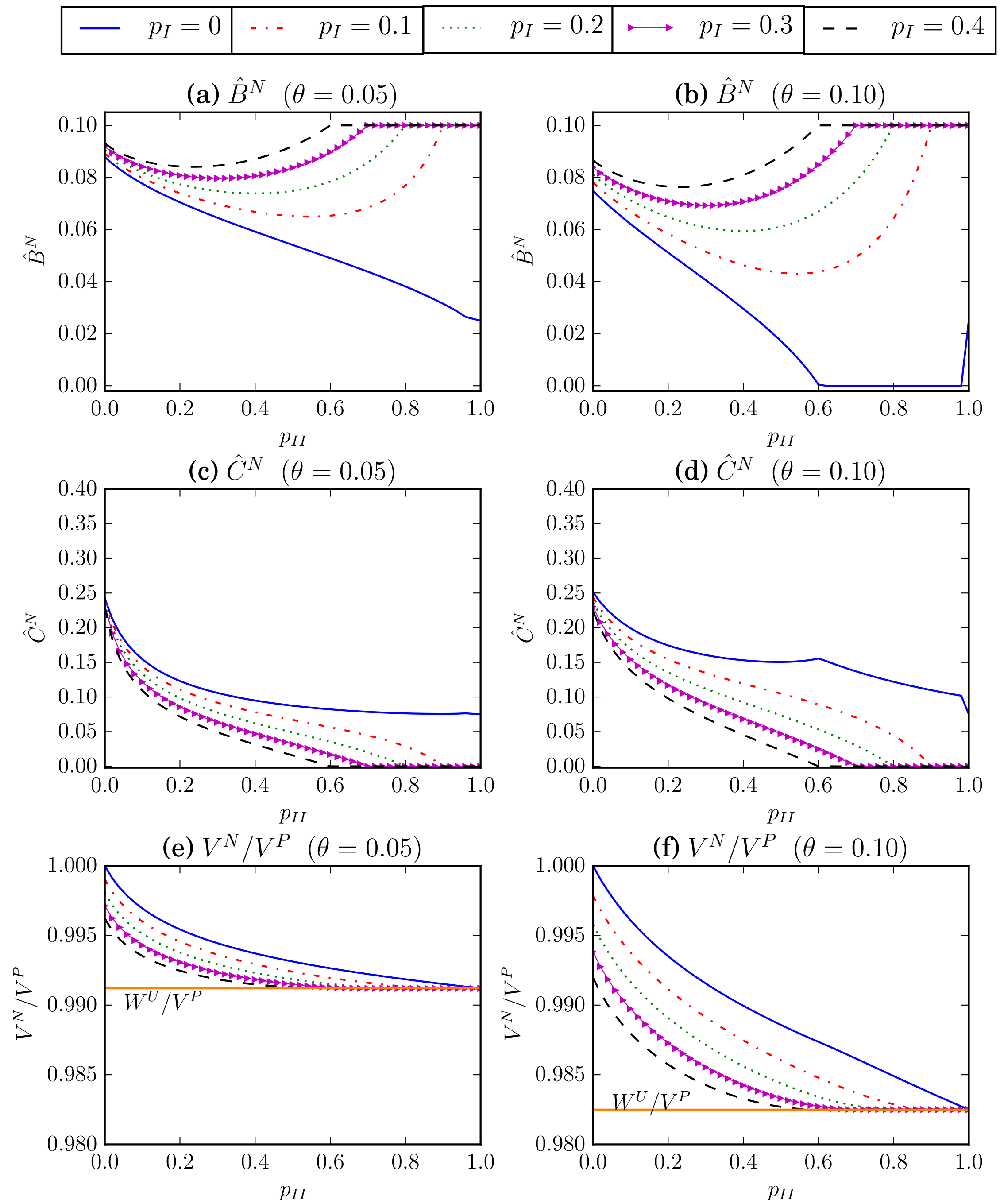
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.8: No Enforcement Optima for $\mathcal{E} = 0.6$



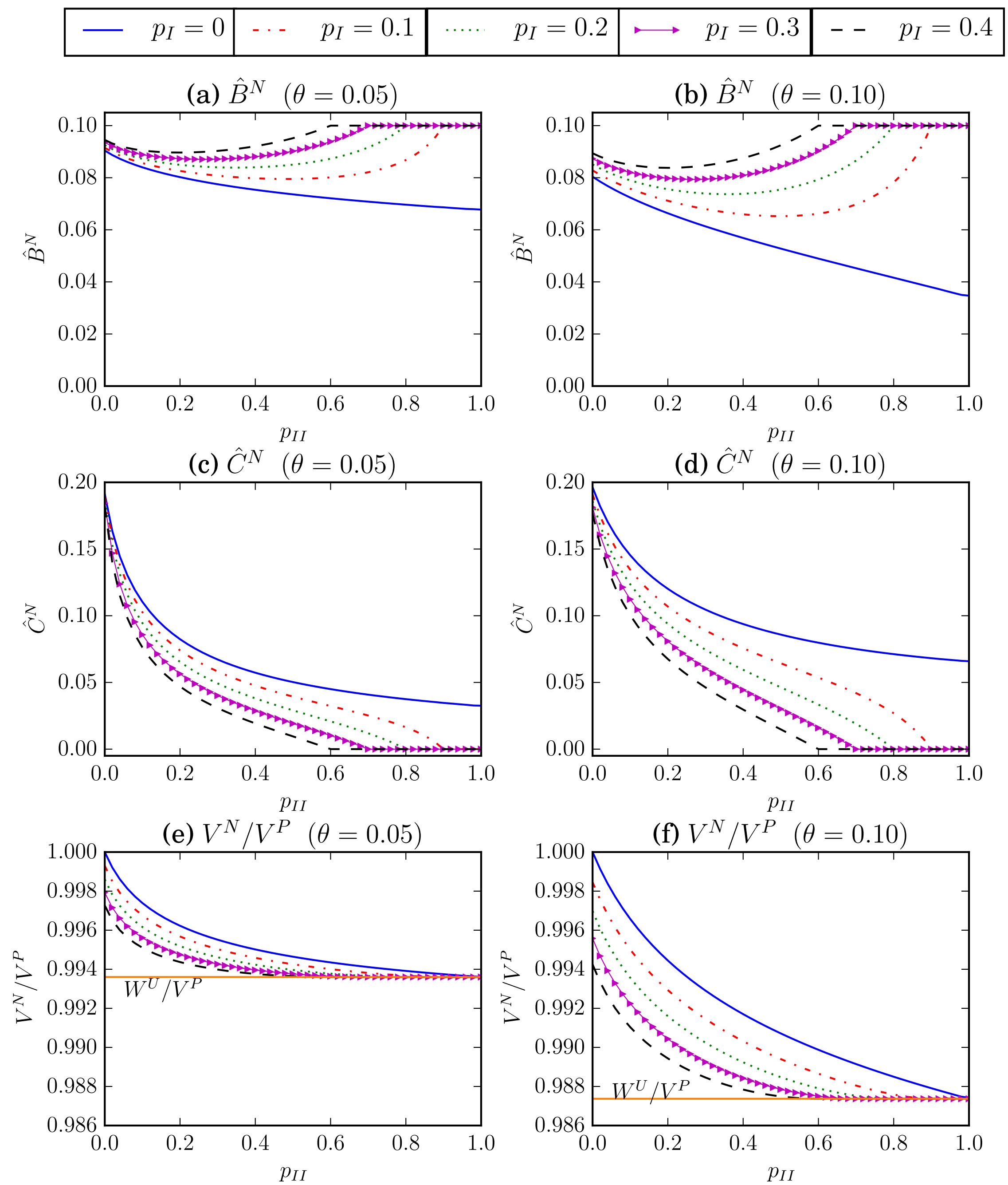
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.9: No Enforcement Optima for $\mathcal{E} = 0.7$



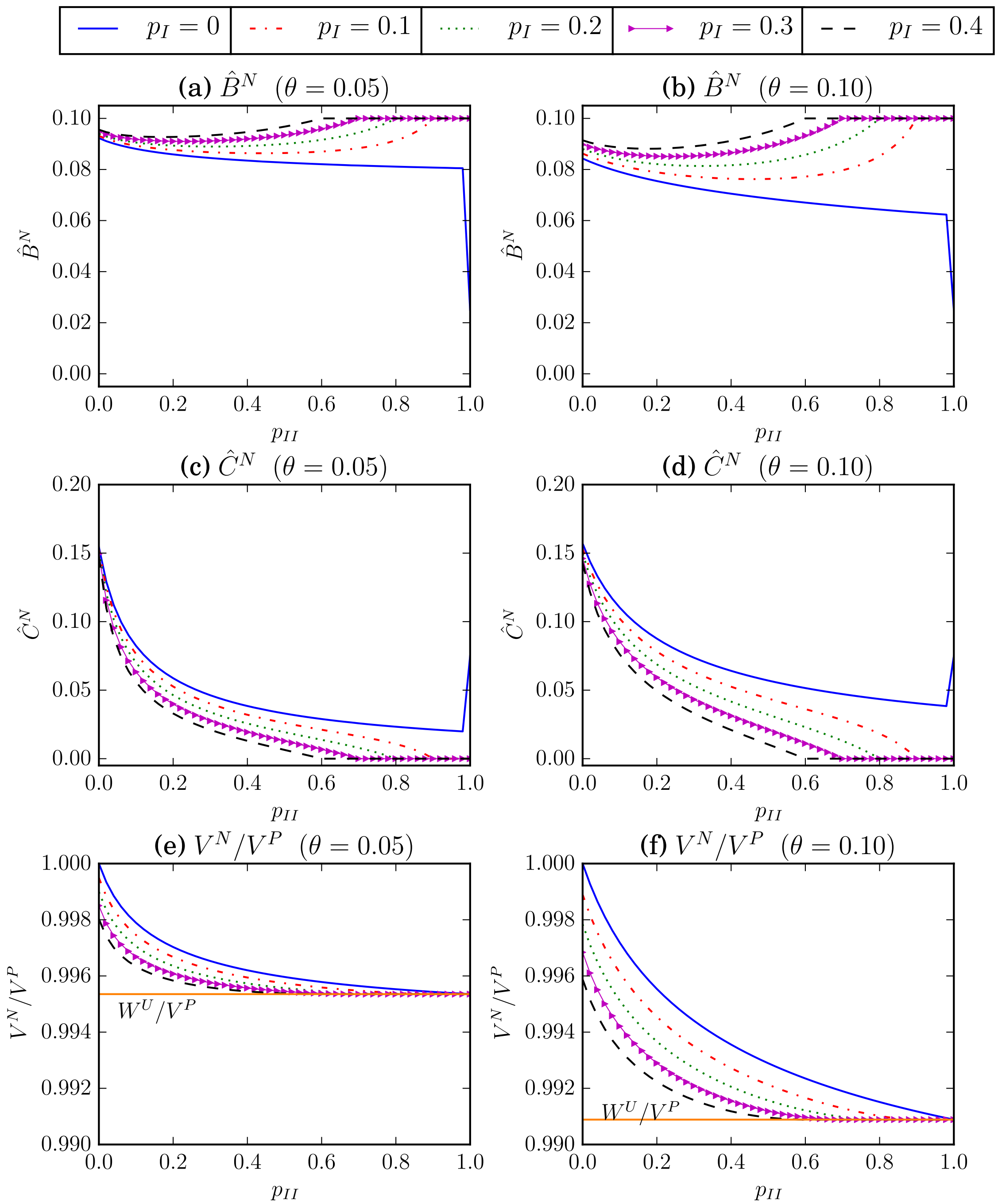
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.10: No Enforcement Optima for $\mathcal{E} = 0.8$



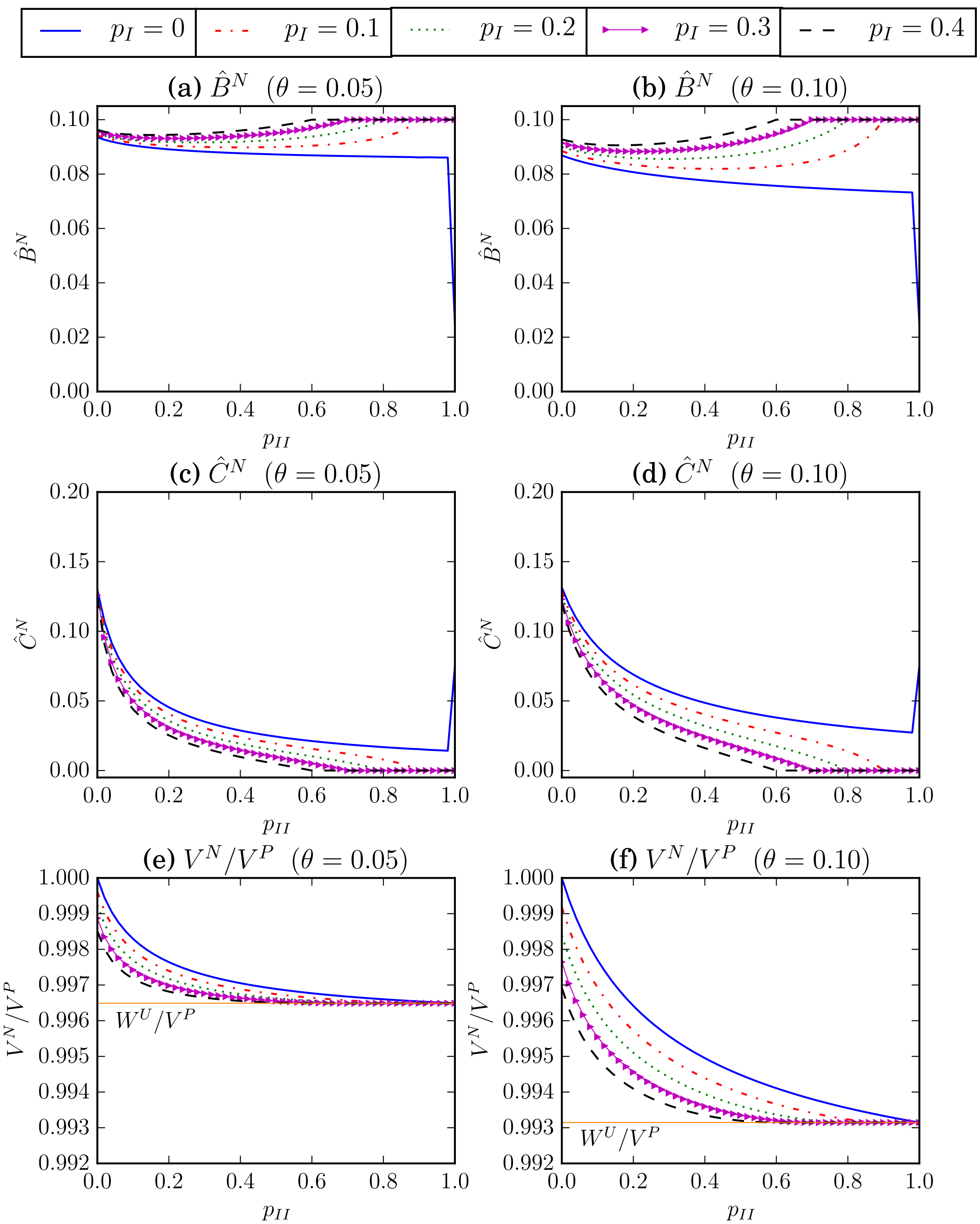
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.11: No Enforcement Optima for $\mathcal{E} = 0.9$



Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.12: No Enforcement Optima for $\mathcal{E} = 0.99$



Notes. W^U/V^P is the horizontal line in panels (e) and (f).

is thus consistent with the theoretical result of $\partial\hat{B}^N/\partial p_I > 0$ in Proposition 2d. The intuition is as described in that section.

Next, the effect of Type II errors on the optimal benefit levels is somewhat more complicated. In particular, each figure illustrates that $\hat{B}^N$ is non-monotonic in p_{II} whenever $p_I > 0$. $\hat{B}^N$ tends to initially fall with p_{II} when p_{II} is low and discriminatory power is reasonably high; but then rises with p_{II} at higher values of p_{II} and thus lower levels of discriminatory power. As discriminatory power approaches zero $\hat{B}^N$ converges to the *per capita* budget size β. Overall, $\hat{C}^N$ tends to fall with p_{II}; but falls less steeply in the region where $\hat{B}^N$ is falling.

To provide the intuition for how $\hat{B}^N$ and $\hat{C}^N$ change with p_{II} we return to the optimality condition in (2.22). Recall this states that at the optimum the average smvi of non-categorical recipients (i.e. $\bar{\sigma}^N_{NR}$) should equate with the average smvi of categorical recipients (i.e. $\bar{\sigma}^N_R$). Consider the effect of an increase in p_{II} on both $\bar{\sigma}^N_{NR}$ and $\bar{\sigma}^N_R$, when evaluated at $(B, C) = (\hat{B}^N, C^N)$:[37]

$$\frac{d\bar{\sigma}^N_{NR}(\hat{B}^N, C^N; \theta, p_I, p_{II})}{dp_{II}} = \underbrace{\frac{\partial\bar{\sigma}^N_{NR}}{\partial p_{II}}}_{>0} + \underbrace{\frac{\partial\bar{\sigma}^N_{NR}}{\partial B}}_{<0} \cdot \frac{\partial\hat{B}^N}{\partial p_{II}} \tag{2.51}$$

$$\frac{d\bar{\sigma}^N_{R}(\hat{B}^N, C^N; \theta, p_I, p_{II})}{dp_{II}} = \underbrace{\frac{\partial\bar{\sigma}^N_{R}}{\partial p_{II}}}_{<0} + \underbrace{\frac{\partial\bar{\sigma}^N_{R}}{\partial B}}_{<0} \cdot \left\{ \frac{\partial\hat{B}^N}{\partial p_{II}} \underbrace{\left[1 + \frac{\partial C^N}{\partial B}\right]}_{<0} + \underbrace{\frac{\partial C^N}{\partial p_{II}}}_{<0} \right\} \tag{2.52}$$

At an interior optimum it must hold that $d\bar{\sigma}^N_{NR}/dp_{II} = d\bar{\sigma}^N_R/dp_{II}$. We can therefore establish that:

$$\frac{\partial\hat{B}^N}{\partial p_{II}} = \frac{\dfrac{\partial\bar{\sigma}^N_{NR}}{\partial p_{II}} - \left[\dfrac{\partial\bar{\sigma}^N_R}{\partial p_{II}} + \dfrac{\partial\bar{\sigma}^N_R}{\partial B} \cdot \dfrac{\partial C^N}{\partial p_{II}}\right]}{\dfrac{\partial\bar{\sigma}^N_R}{\partial B}\left[1 + \dfrac{\partial C^N}{\partial B}\right] - \dfrac{\partial\bar{\sigma}^N_{NR}}{\partial B}} \tag{2.53}$$

The denominator is unambiguously positive and so the the sign of $\partial\hat{B}^N/\partial p_{II}$ will

[37]Recall that, as defined in (2.19), C^N is the level of categorical benefit that exhausts the benefit budget for any $B \in [0, \beta]$.

depend on the relative sizes of the terms in the numerator. In this regard it is useful to note that $\partial^2\bar{\sigma}^N_{NR}/\partial p^2_{II} > 0$ *but* both $\partial^2\bar{\sigma}^N_R/\partial p^2_{II} < 0$ and $\partial^2 C^N/\partial p^2_{II} < 0$. This helps explain why we observe that $\partial\hat{B}^N/\partial p_{II} < 0$ when p_{II} is low, but $\partial\hat{B}^N/\partial p_{II} > 0$ at higher values of p_{II}.

To complete this part of the discussion, each figure illustrates that the extent to which $\hat{B}^N$ falls with p_{II} (when p_{II} is low) is decreasing in the size of p_I. The intuition for this observation follows directly from the theoretical result that $\partial\hat{B}^N/\partial p_I > 0$.

Notice that in both Figure 2.7 ($\mathcal{E} = 0.5$) and Figure 2.8 ($\mathcal{E} = 0.6$) there are cases where $\hat{B}^N = 0$ *despite* the fact that $p_I > 0$. This implies that at the corner solution $B = 0$ the average smvi of categorical recipients must (weakly) *exceed* that of non-recipients. At higher values of $\mathcal{E}$ (i.e. Figures 2.9-2.12) this does not arise and $\hat{B}^N > 0$ whenever $p_I > 0$. Recall that in the theory there are no corner solutions when $p_I > 0$ due to the assumption that $\lim_{x\to 0} u_x = +\infty$. To shed some light as to why corner solutions can occur under CES preferences it is useful to note that we can write the smvi of an unable individual/non-worker as:

$$\begin{aligned}
u_x &= \alpha M^{-\frac{1}{\mathcal{E}}}\left[\alpha M^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]^{\frac{1}{\mathcal{E}-1}} \\
&= \alpha\left\{M^{\frac{1-\mathcal{E}}{\mathcal{E}}}\left[\alpha M^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]\right\}^{\frac{1}{\mathcal{E}-1}} \\
&= \alpha\left[\alpha + (1-\alpha)M^{\frac{1-\mathcal{E}}{\mathcal{E}}}\right]^{\frac{1}{\mathcal{E}-1}}
\end{aligned}$$

and thus:

$$\lim_{x\to 0} u_x = \alpha^{\frac{\mathcal{E}}{\mathcal{E}-1}} \tag{2.54}$$

which is finite and *increasing* in $\mathcal{E}$. In particular, this limiting case takes the values: 1.67 when $\mathcal{E} = 0.5$; 2.15 when $\mathcal{E} = 0.6$; 3.29 when $\mathcal{E} = 0.7$; 7.72 when $\mathcal{E} = 0.8$; 99.23 when $\mathcal{E} = 0.9$ and 9.18×10^{21} when $\mathcal{E} = 0.99$. It is the lower values when $\mathcal{E} = 0.5$ and $\mathcal{E} = 0.6$ that result in their being corner solutions when $\hat{B}^N = 0$ despite $p_I > 0$. Consider, for example, subplot (a) in Figure 2.7. We can see that $\hat{B}^N$ is zero - or very close to identically zero - when $p_I = 0.1$ and $p_{II} = 0.4$. As stated, this can only arise if the average smvi of categorical recipients weakly exceeds that of non-recipients under a purely targeted system. This can be easily verified: setting $B = 0$, $C = \beta/[\theta(1 - p_I) + (1 - \theta)p_{II}]$ and choosing parameters $p_I = 0.1$, $p_{II} = 0.4$, $\beta = 0.1$, $\theta = 0.05$ and $\mu = 3$ gives:

$$\bar{\sigma}^N_{NR} = 0.96989 < \bar{\sigma}^N_R = 0.97663$$

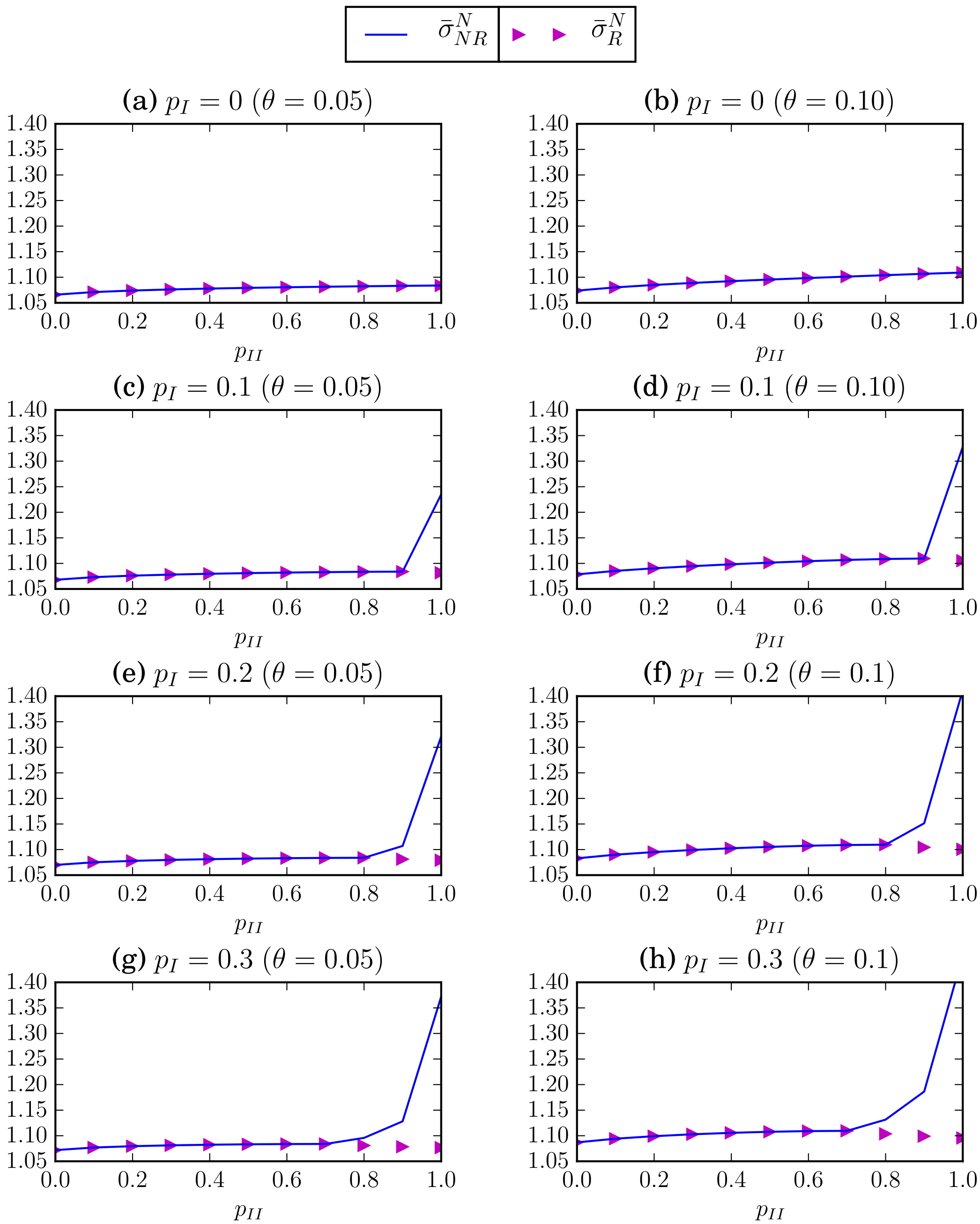
which explains why $\hat{B}^N = 0$ at the optimum despite $p_I > 0$.

The effects of changes in the values of the parameters $\mathcal{E}$ and θ are analogous to the Perfect Discrimination simulations. First, a comparison across Figures 2.7 to 2.12 illustrates that an increase in $\mathcal{E}$ tends to increase the universal benefit and lower the categorical benefit. Conversely, within each figure it can be seen that a *ceteris paribus* increase in θ from 0.05 to 0.10 serves to increase the resources devoted to the categorical transfer and, consequently, lower the resources dedicated to the universal transfer. Unsurprisingly, an increase in the unable subpopulation size raises the importance of categorical transfers due to the greater need for redistribution between the two subpopulations.

Subplots (e) and (f) of each figure plot maximum social welfare under No Enforcement as a proportion of that under Perfect Discrimination. This has the attractive property that the scale on the vertical axis is between 0 and 1, with 1 capturing the Perfect Discrimination outcome. Each figure illustrates that social welfare under the optimal choices $\hat{B}^N$ and $\hat{C}^N$ is unambiguously *decreasing* in the propensity to make Type I and/or Type II classification errors. The results are thus consistent with the theory in Proposition 2c. The horizontal line in both subplots is labelled W^U/V^P and therefore captures the Pure Universal level of welfare as a fraction of the Perfect Discrimination welfare level. The purpose of including this line in the subplots is to illustrate that *whenever* $p_I + p_{II} = 1$ (including the case where $p_I = 0$) a Pure Universal system is chosen and, consequently, $V^N = W^U$.

Finally, Figures 2.13 to 2.15 illustrate how the average smvi of categorical recipients and non-recipients change with p_{II} for various values of p_I. Figure 2.13 is drawn for $\mathcal{E} = 0.8$, Figure 2.14 is drawn for $\mathcal{E} = 0.9$ and Figure 2.15 is drawn for $\mathcal{E} = 0.99$. Within each figure, the subplots on the left column are drawn for $\theta = 0.05$, whilst those on the right column are drawn for $\theta = 0.10$. Each figure clearly illustrates that for all cases where $p_I + p_{II} < 1$ (i.e. some discriminatory power) the average smvi of categorical

Figure 2.13: $\bar{\sigma}^N_{NR}$ and $\bar{\sigma}^N_R$ evaluated at the Optimum. ($\mathcal{E} = 0.8$, $\beta = 0.1$)



Notes. The purpose of this figure is to illustrate that the numerical results satisfy the optimality condition in Proposition 2a. On the horizontal axis in each subplot p_{II} is varied between 0 and 1 in discrete increments of 0.1. For each value of p_{II} (and other parameters) the optimal benefits are calculated and substituted into the functions $\bar{\sigma}^N_{NR}$ and $\bar{\sigma}^N_R$. In all cases where $p_I + p_{II} < 1$ we observe $\bar{\sigma}^N_{NR} = \bar{\sigma}^N_R$, thus satisfying the optimality condition in Proposition 2a.

Figure 2.14: $\bar{\sigma}^N_{NR}$ and $\bar{\sigma}^N_R$ evaluated at the Optimum. ($\mathcal{E} = 0.9$, $\beta = 0.1$)

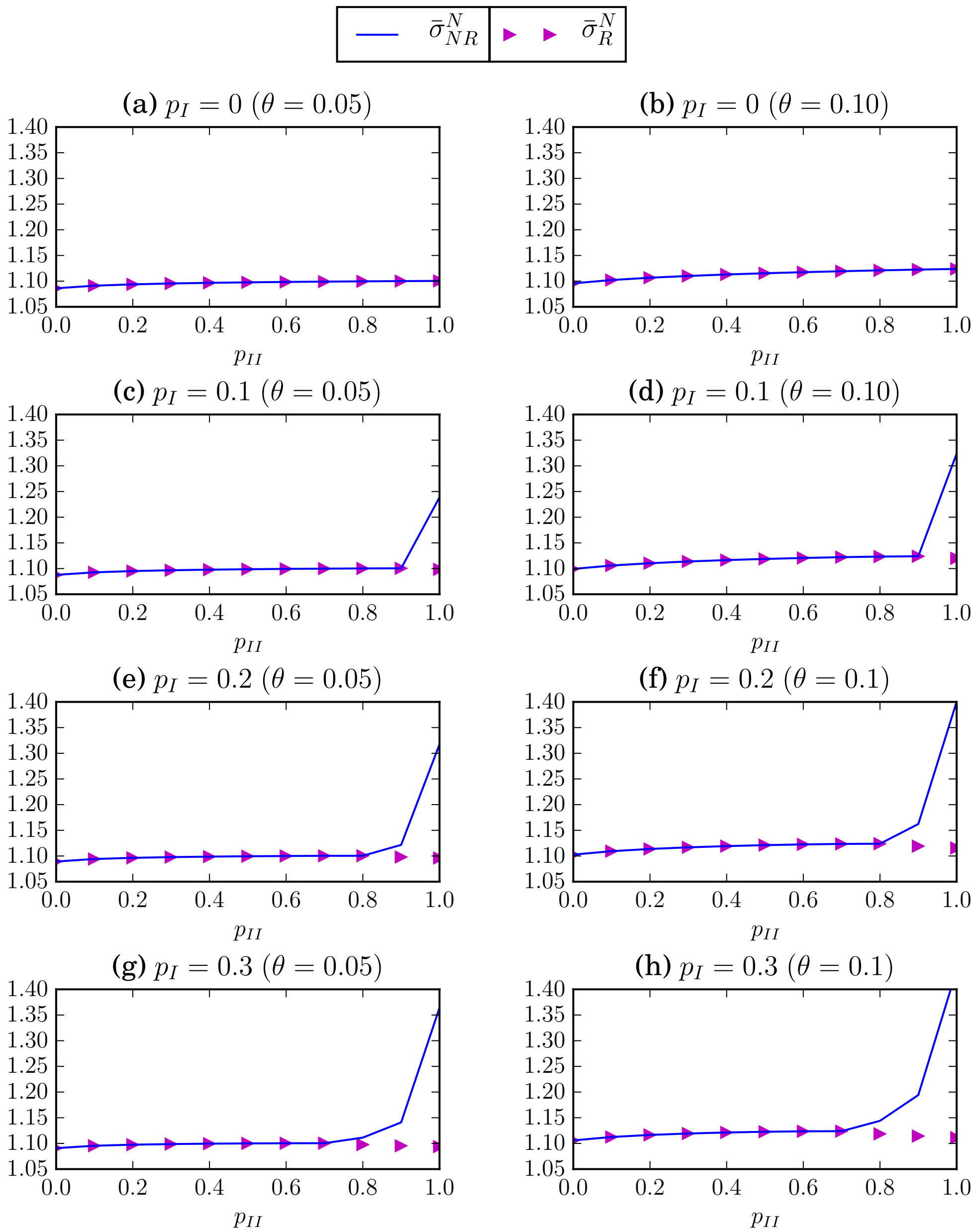
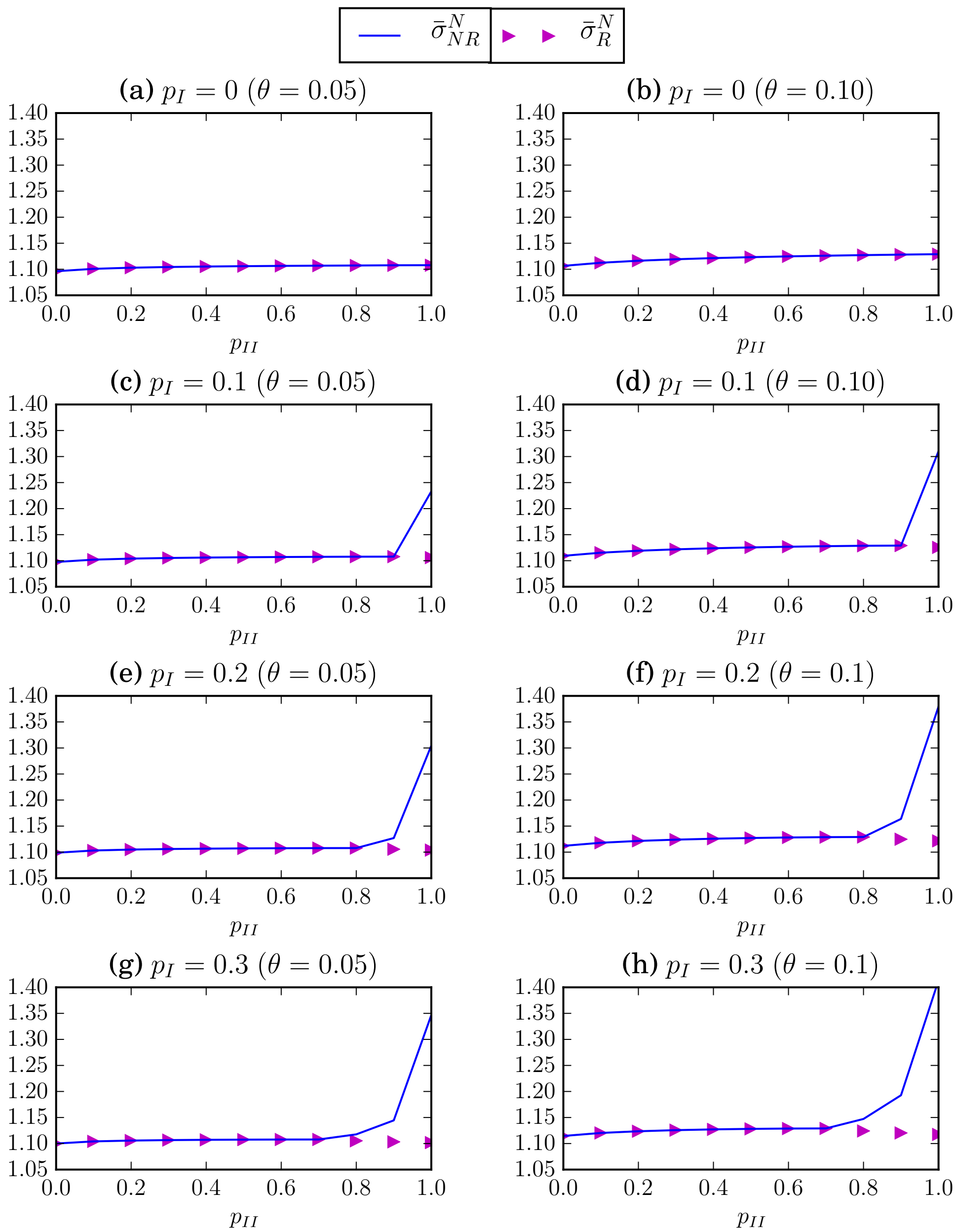


Figure 2.15: $\bar{\sigma}^N_{NR}$ and $\bar{\sigma}^N_R$ evaluated at the Optimum. ($\mathcal{E} = 0.99$, $\beta = 0.1$)



recipients and non-recipients are equated, thus satisfying the optimality condition in Proposition 2a. These figures therefore provide a useful validation of the optimal benefits simulated in Figures 2.10 - 2.12 and this is the reason for including them. Of course, when $p_I + p_{II} = 1$ only a universal benefit is provided and, consequently, the average smvi of 'recipients' and 'non-recipients' departs.

This completes our discussion of the No Enforcement simulation results. We now proceed to discuss the case of Full Enforcement.

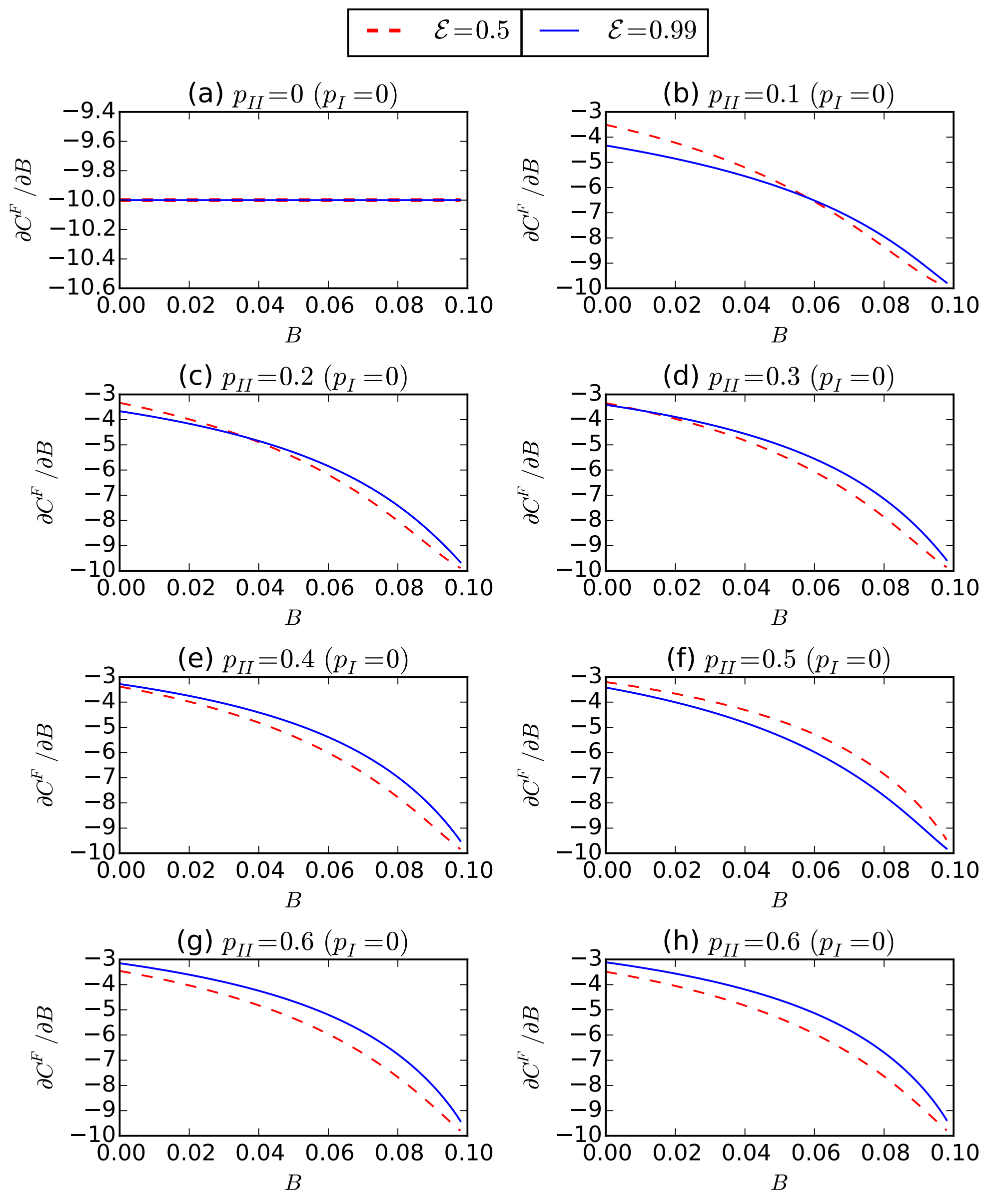
2.4.3 Full Enforcement Simulations

Under No Enforcement the constraint set is unambiguously convex and the objective function is unambiguously concave. This rendered the search for an optimum in the numerical analysis relatively straightforward given the single stationary point corresponding to a global (interior) optimum. Turning to the Full Enforcement case, however, we know from Section 2.3.3 that - without specific assumptions on both the productivity distribution and third derivatives of individual preferences - we cannot guarantee either the convexity of the constraint set or the concavity of the objective function.

To proceed with the Full Enforcement numerical analysis we must therefore exercise a little caution and first examine both the convexity of the constraint set and how the welfare function changes with the choice variables. It is below demonstrated that the constraint set is indeed convex in *all* considered cases. However, social welfare is not always concave in the choice variables. Whilst this is not a problem in itself provided a sufficient number of starting points are employed in the search for a global optima, there are a small number of cases in which multiple optima emerge. These cases are restricted to lower values of $\mathcal{E}$ and are detailed below.

Convexity of the constraint set For the constraint set to be convex it must hold that $\partial^2 C^F/\partial B^2 \leq 0 \; \forall B \in [0, \beta]$, where C^F is as defined in (2.33). Figure 2.16 plots how the function $\partial C^F/\partial B < 0$ changes with $B \in [0, \beta]$ for both $\mathcal{E} = 0.5$ and $\mathcal{E} = 0.99$. The figure is composed of eight subplots, where each successive subplot has a higher

Figure 2.16: The Function $\partial C^F/\partial B$ over $B \in [0, \beta]$.



Notes: C^F is the level of the categorical benefit that exhausts the budget constraint for any given level of $B \in [0, \beta]$. $\partial C^F/\partial B < 0$ therefore captures how C^F changes with B. For the budget constraint to be convex we require that C^F either decreases linearly (such that $\partial C^F/\partial B < 0$ is constant and $\partial^2 C^F/\partial B^2 = 0$) or C^F decreases more rapidly as B increases (such that $\partial C^F/\partial B < 0$ is decreasing and $\partial^2 C^F/\partial B^2 < 0$). In subplot (a) $\partial C^F/\partial B$ is a negative constant because $p_{II} = 0$ and, consequently, the budget constraint is linear. In the remaining subplots $\partial C^F/\partial B < 0$ becomes more negative with B and thus $\partial^2 C^F/\partial B^2 < 0$.

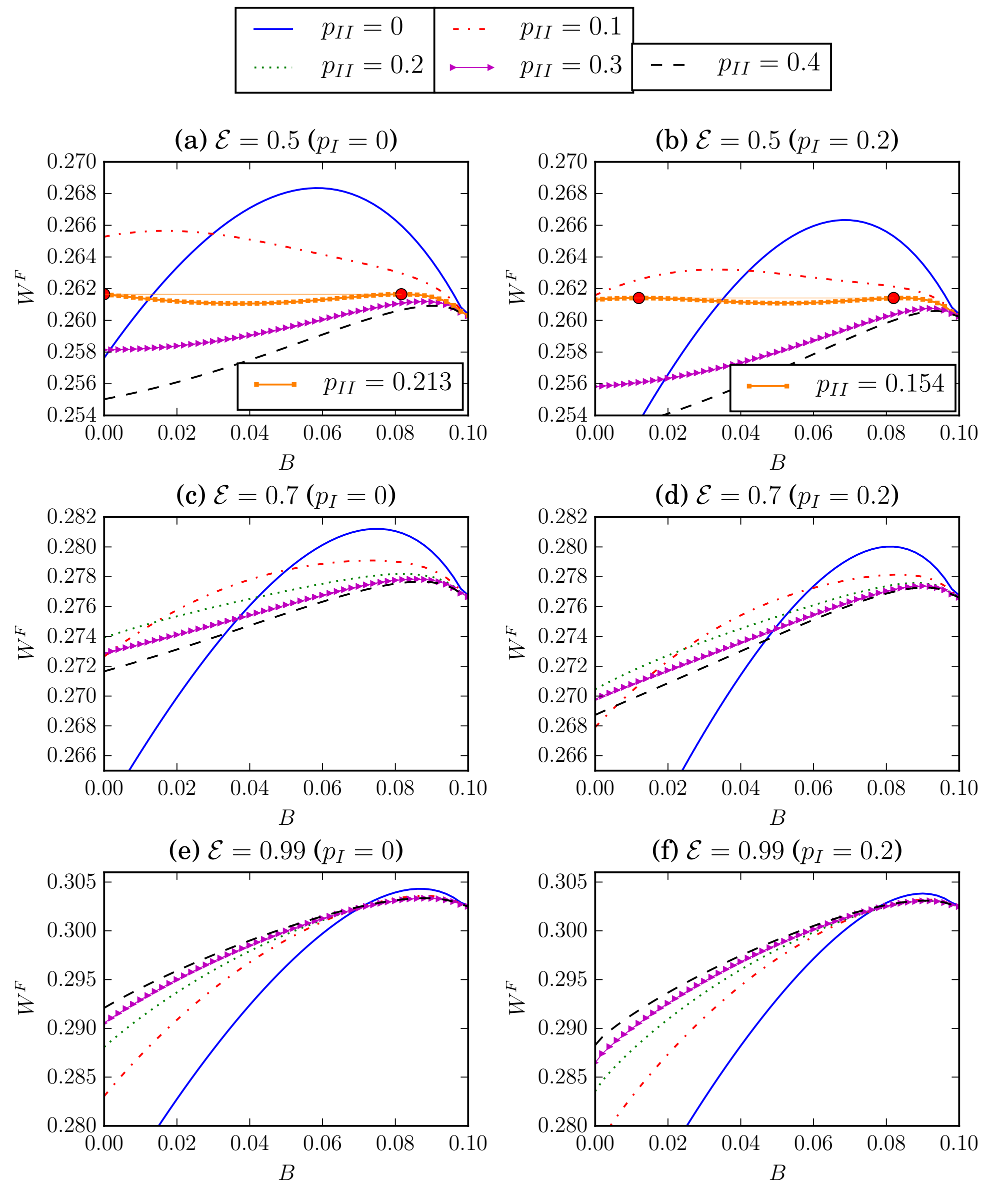
value of p_{II}. In all cases $p_I = 0$ and $\theta = 0.1$[38]. Each subplot in which $p_{II} > 0$ clearly illustrates that $\partial C^F/\partial B$ is negative and *falling* with B, from which we can directly infer that $\partial^2 C^F/\partial B^2 < 0$. Notice that in all cases $\partial C^F/\partial B < -1$ such that an increase in B unambiguously lowers the total benefit income of categorical recipients.

The objective function Figure 2.17 plots how the welfare function $W^F(B, C^F; \theta, p_I, p_{II})$ changes with B. The figure is generated as follows: for a given value of $B \in [0, \beta]$ we calculate the resulting level of C that satisfies the budget constraint (i.e. C^F) and substitute both benefit levels into the Full Enforcement welfare function. The subplots on the first row are generated for $\mathcal{E} = 0.5$, those on the second row are generated for $\mathcal{E} = 0.7$, whilst those on the third row are generated for $\mathcal{E} = 0.99$. The subplots on the left column are drawn for $p_I = 0$, whilst those on the right column are drawn for $p_I = 0.2$. Finally, the different curves within each subplot are drawn for different values of p_{II}. Subplots (a) and (b) clearly illustrate that there is a critical value of p_{II} at which there are two global optima. In subplot (a) this critical value is given by $p_{II} \approx 0.213$. The first of these optima is a purely targeted system (corner solution), whilst the second is a partial system providing a small categorical benefit. In subplot (b) the critical value is $p_{II} \approx 0.154$ and both equilibria are partial systems. For values of p_{II} at or just above these two critical values welfare provision will therefore shift between the two equilibria. In the simulation results which follow we will thus observe a 'jump' in the benefit levels around these critical values of p_{II}. Importantly, this behaviour of the objective function is *not* observed at higher values of $\mathcal{E}$ - as illustrated through subplots (c) to (f) - and there are thus no further problems of multiple equilibria.

The intuition for the multiple equilibria observed in subplots (a) and (b) of Figure 2.17 follows from our discussion in Section 2.3.3 concerning how the functions σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ change with the universal benefit, when evaluated at (B, C^F). Recall from (2.40) and (2.41) that in both cases an increase in B generates two conflicting effects. In terms of σ^F_{NR} an increase in B (i) lowers the smvi of each individual; but also (ii) increases the number of non-categorical recipients through reducing the number that apply for the categorical benefit. In terms of $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ an increase in B (i) increases a categorical recipient's smvi because their total benefit income falls; but also (ii) reduces the number of categorical recipients. Notice that the

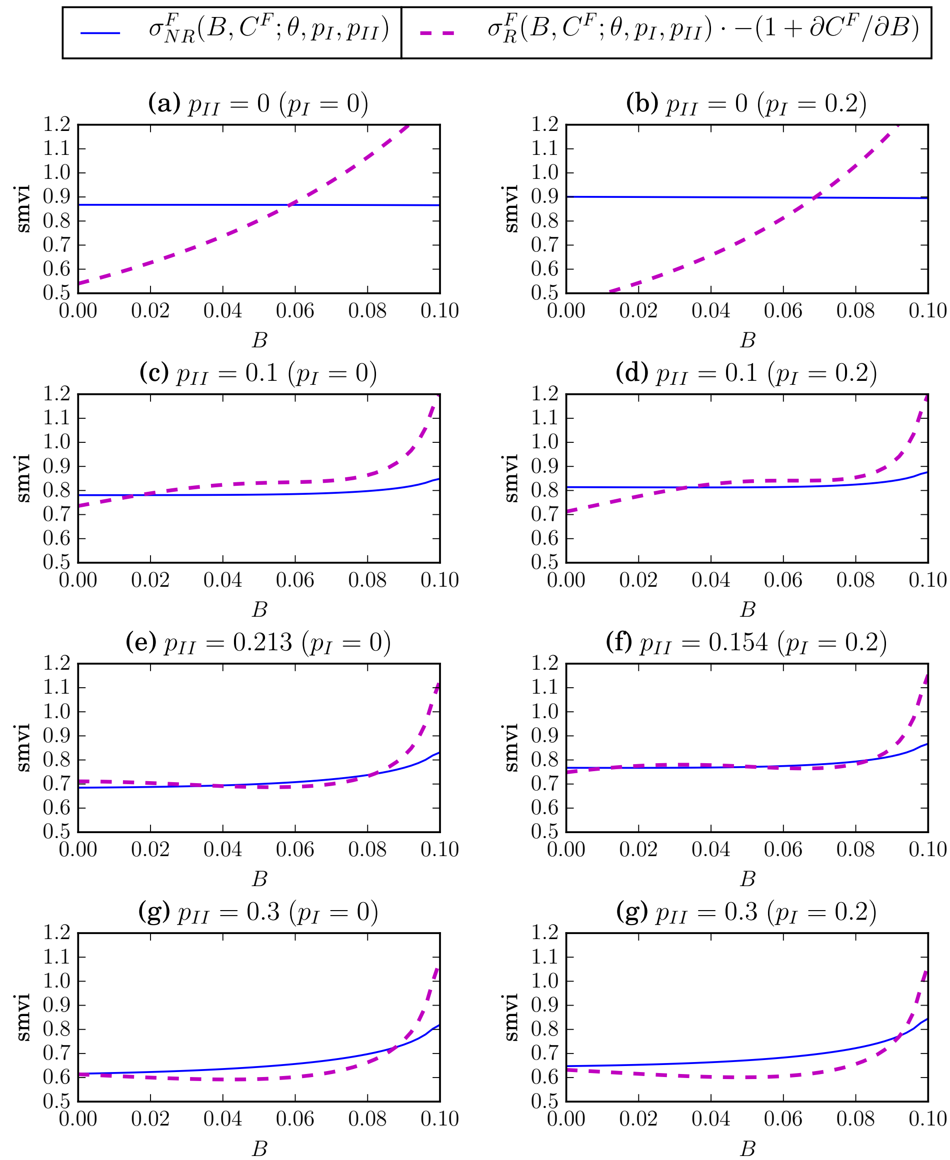
[38]Extensive examples with variation in both θ and p_I are available from the author upon request but are omitted to save on space.

Figure 2.17: Full Enforcement Objective Function



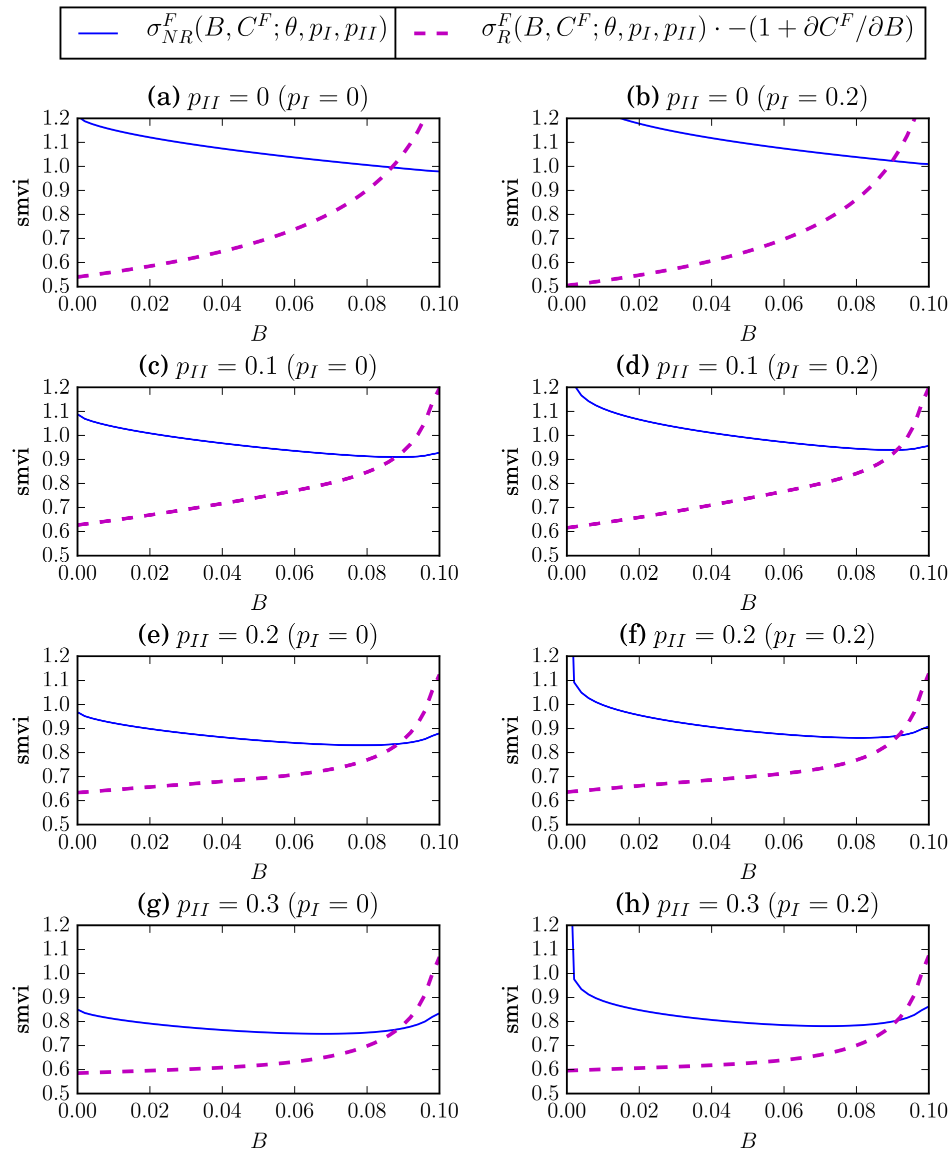
Notes: In each subplot $\theta = 0.1$. Subplots (a) and (b) are drawn for $\mathcal{E} = 0.5$. In subplot (a) the parameter value $p_{II} \approx 0.213$ gives rise to two global optima. The first of these optima is a purely targeted system (corner solution) whilst the second is a partial system with a small value of the categorical benefit. In subplot (b) the parameter value $p_{II} \approx 0.154$ gives rises to two global optima. Both optima are this time partial systems (i.e. interior solutions). Contrastingly, for the cases where $\mathcal{E} = 0.7$ (subplots (c) and (d)) and $\mathcal{E} = 0.99$ (subplots (e) and (f)) no problems of multiple equilibria arise at any value of p_{II}.

Figure 2.18: σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ evaluated over (B, C^F). $(\mathcal{E} = 0.5)$



Notes: In each subplot $\mathcal{E} = 0.5$ and $\theta = 0.1$. The subplots in the left column are generated for $p_I = 0$, whilst those in the right column are generated for $p_I = 0.2$. The parameter p_{II} is increased as we move down the columns. Note that in subplots (e) and (f) p_{II} is set at the critical values established in Figure 2.17. Intersections between the functions σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ correspond to stationary points on the welfare function.

Figure 2.19: σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ evaluated over (B, C^F). ($\mathcal{E} = 0.99$)



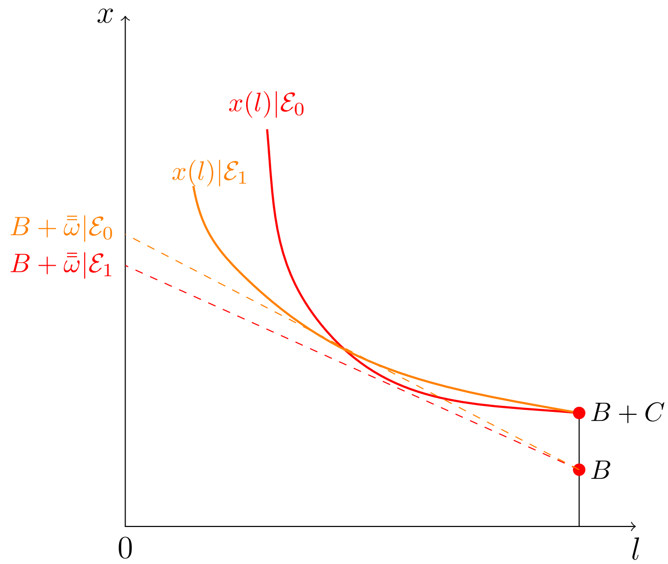
Notes: In each subplot $\mathcal{E} = 0.99$ and $\theta = 0.1$. The structure of the figure is as in Figure 2.17. In all cases $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ monotonically increases with B and there is unique intersection - corresponding to a stationary point on the welfare function - between σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$. The value of B (and resulting value of C) at each unique intersection gives rise to a global maximum of the welfare function.

first effect is compounded by the fact that $\partial^2 C^F/\partial B^2 < 0$. These conflicting effects can given rise to non-monotonicity in either function. This non-monotonicity can lead to multiple stationary points and, in the special cases illustrated in Figure 2.17, multiple optima.

To illustrate this, Figure 2.18 plots how both σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ change with the benefit levels. The purpose of this figure is to explain subplots (a) and (b) in Figure 2.17 and so the parameter values are chosen accordingly (i.e. $\mathcal{E} = 0.5$, $\theta = 0.1$). The subplots on the left column are generated for $p_I = 0$, whilst those on the right column are generated for $p_I = 0.2$. The value of p_{II} is increased as we move down the columns. In the special case where $p_{II} = 0$ the effect of an increase in B on the number of individuals who apply for the categorical benefit necessarily disappears and so $\sigma^F_R \cdot -(1+\partial C^F/\partial B)$ unambiguously increases with B; whilst σ^F_{NR} unambiguously falls with B, thus resulting in a unique intersection between the two functions. However, as p_{II} is increased from zero the two conflicting effects associated with an increase in B result in a point of inflexion in $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$. This in turn can result in multiple values of B for which $\sigma^F_{NR} = \sigma^F_R \cdot -(1 + \partial C^F/\partial B)$. Indeed, in subplots (e) and (f) the parameter p_{II} is set at the critical values illustrated in subplots (a) and (b) of Figure 2.17, respectively. In subplot (e) there are two intersections between the the two functions: the benefit levels at the first of these corresponds to a global minima, whilst those at the second correspond to maxima that generate the same welfare level as the corner solution where $B = 0$. Contrastingly, in subplot (f) we observe three intersections: the benefit levels at the first and third are both interior optima, whilst those at the second correspond to minima. Finally, as p_{II} is increased further (subplots (g) and (h)) the effect of an increase in B on the number of benefit applicants (i.e. reduced applications) initially dominates the increase in the smvi of categorical recipients, thus causing $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ to initially fall. In these cases there is a unique value of B at which $\sigma^F_{NR} = \sigma^F_R \cdot -(1 + \partial C^F/\partial B)$.

That we do not observe multiple optima at higher values of $\mathcal{E}$ would seem to stem from the fact that $\bar{\bar{\omega}}$ is increasing in $\mathcal{E}$ when $M < 1$ (see Figure 2.20). If we re-examine equation (2.41) we can note that higher values of $\bar{\bar{\omega}}$ (i) increase χ and thus place more weight on the effect of a reduction in total benefit income on the smvi of categorical recipients; but also (ii) will imply a lower value of $f(\bar{\bar{\omega}})$ under the Exponential distribution, thus lessening the effect of a reduction in the number categorical recipients. Further, at higher values of $\mathcal{E}$ the number of applicants for the categorical benefit tends

Figure 2.20: $\bar{\bar{\omega}}$ and $\mathcal{E}$.

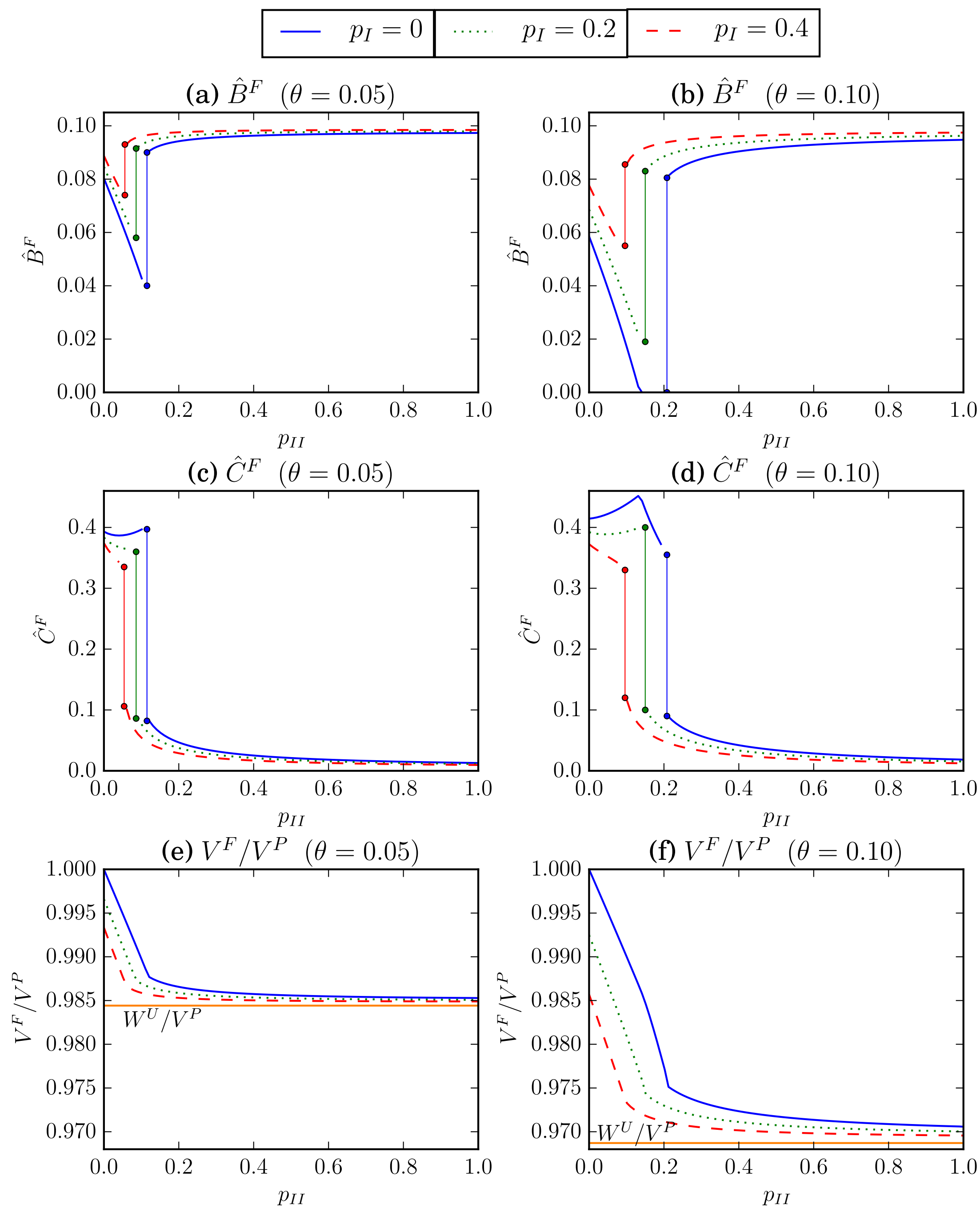


Notes: The figure provides graphical intuition for how $\bar{\bar{\omega}}$ changes with $\mathcal{E}$, where $\mathcal{E}_0 < \mathcal{E}_1$ denote different levels of the elasticity of substitution. Notice that the indifference curve $x(l)|\mathcal{E}_0$ is shallower than $x(l)|\mathcal{E}_1$ at high values of l. This arises because the reservation wage, $\bar{\omega} = (\frac{1-\alpha}{\alpha})M^{\frac{1}{\mathcal{E}}}$, is increasing in $\mathcal{E}$ when $M < 1$.

to be less responsive to changes in the level of B (and in turn the total benefit income of categorical recipients). The implication of both of these observations is that, at higher values of $\mathcal{E}$, the effect of a reduction in the total benefit income of categorical recipients and their corresponding increase in smvi will tend to dominate the effect of a reduction in the number of applicants, thus causing $\sigma_R^F \cdot -(1 + \partial C^F/\partial B)$ to monotonically increase with B (see Figure 2.19).

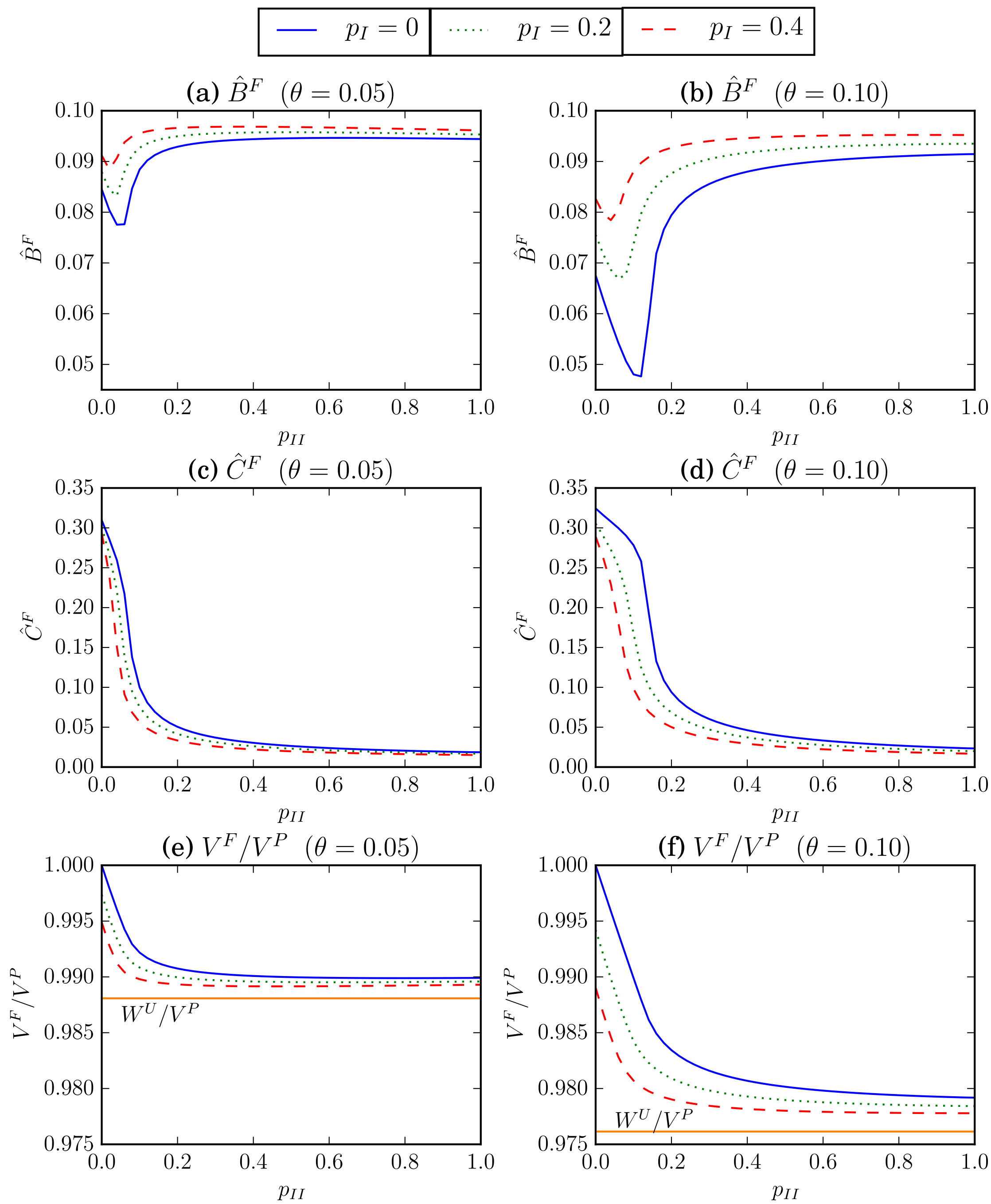
Discussion of the numerical results We now proceed to discuss the numerical results. Figures 2.21 to 2.26 illustrate how the Full Enforcement optimal benefit levels and resulting maximum welfare change with the propensity to make classification errors. Each figure is generated for a different $\mathcal{E} \in \{0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$. The format of these figures parallels Figures 2.7 to 2.12 from the No Enforcement analysis. The subplots in the left column of each figure are generated for $\theta = 0.05$, whilst those in the right column are generated for $\theta = 0.10$. The parameter p_{II} is increased in discrete increments of 0.02 from 0 to 1 on the horizontal axis. The different curves within each

Figure 2.21: Full Enforcement Optima for $\mathcal{E} = 0.5$



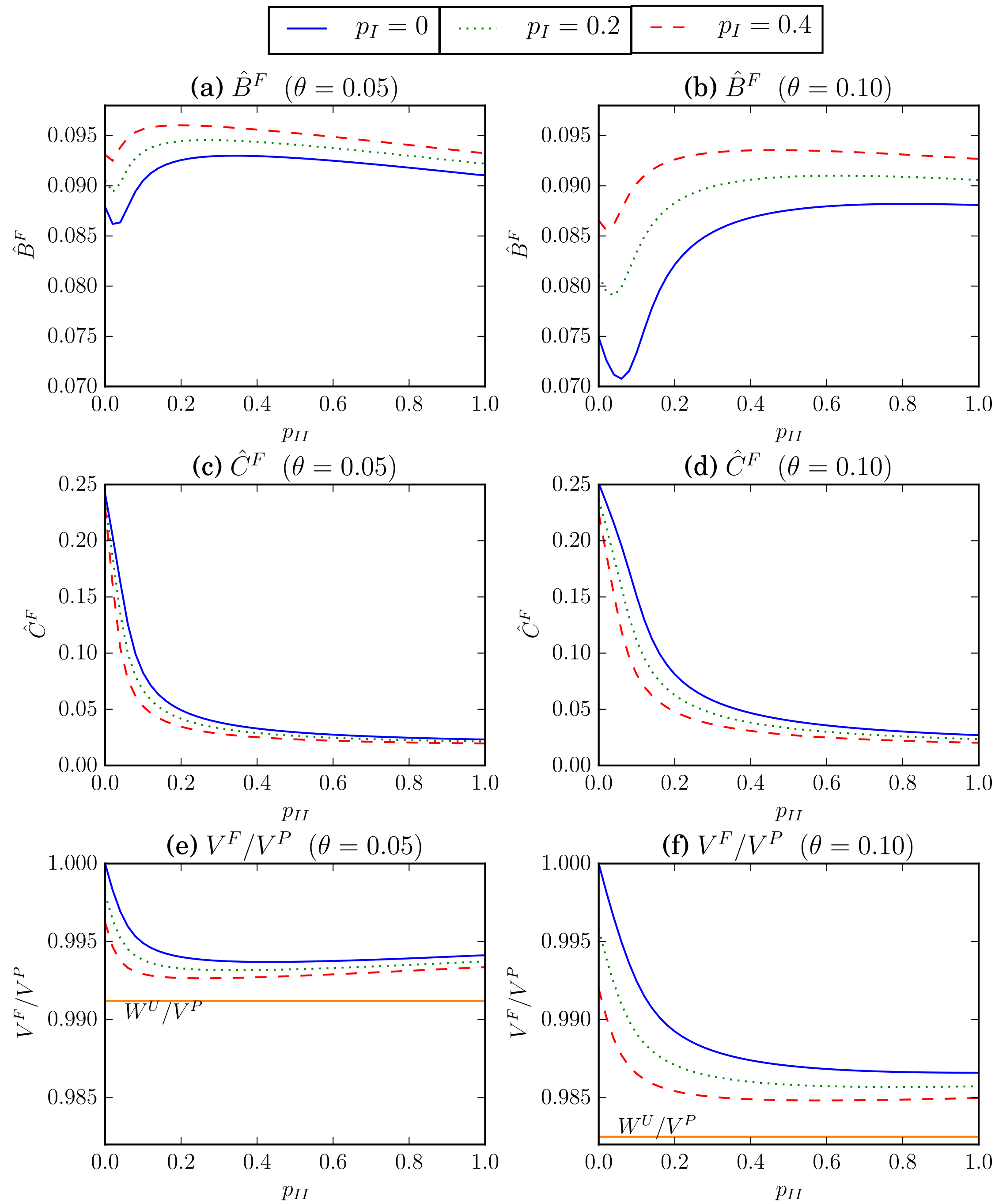
Notes. The vertical lines in subplots (a) to (d) capture discontinuities in the optimal benefit functions. Note that the results in subplots (b) and (d) are consistent with the discussion surrounding Figure 2.17: i.e. we observe discontinuities at the parameters $(p_I, p_{II}) = (0, 0.213)$ and $(p_I, p_{II}) = (0.2, 0.154)$, respectively. The horizontal line W^U/W^P in subplots (e) and (f) gives the Pure Universal welfare level as a fraction of the Perfect Discrimination welfare.

Figure 2.22: Full Enforcement Optima for $\mathcal{E} = 0.6$



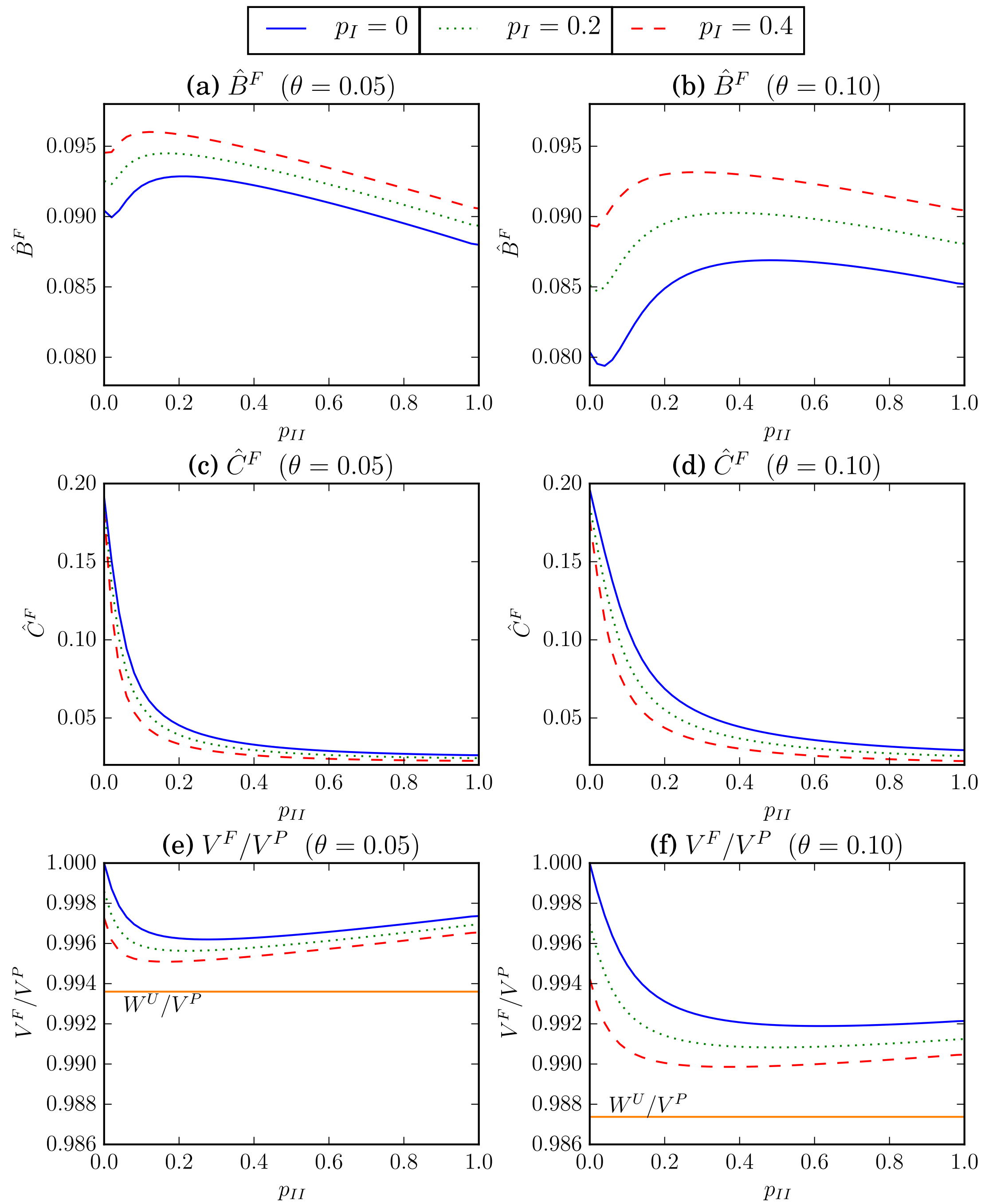
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.23: Full Enforcement Optima for $\mathcal{E} = 0.7$



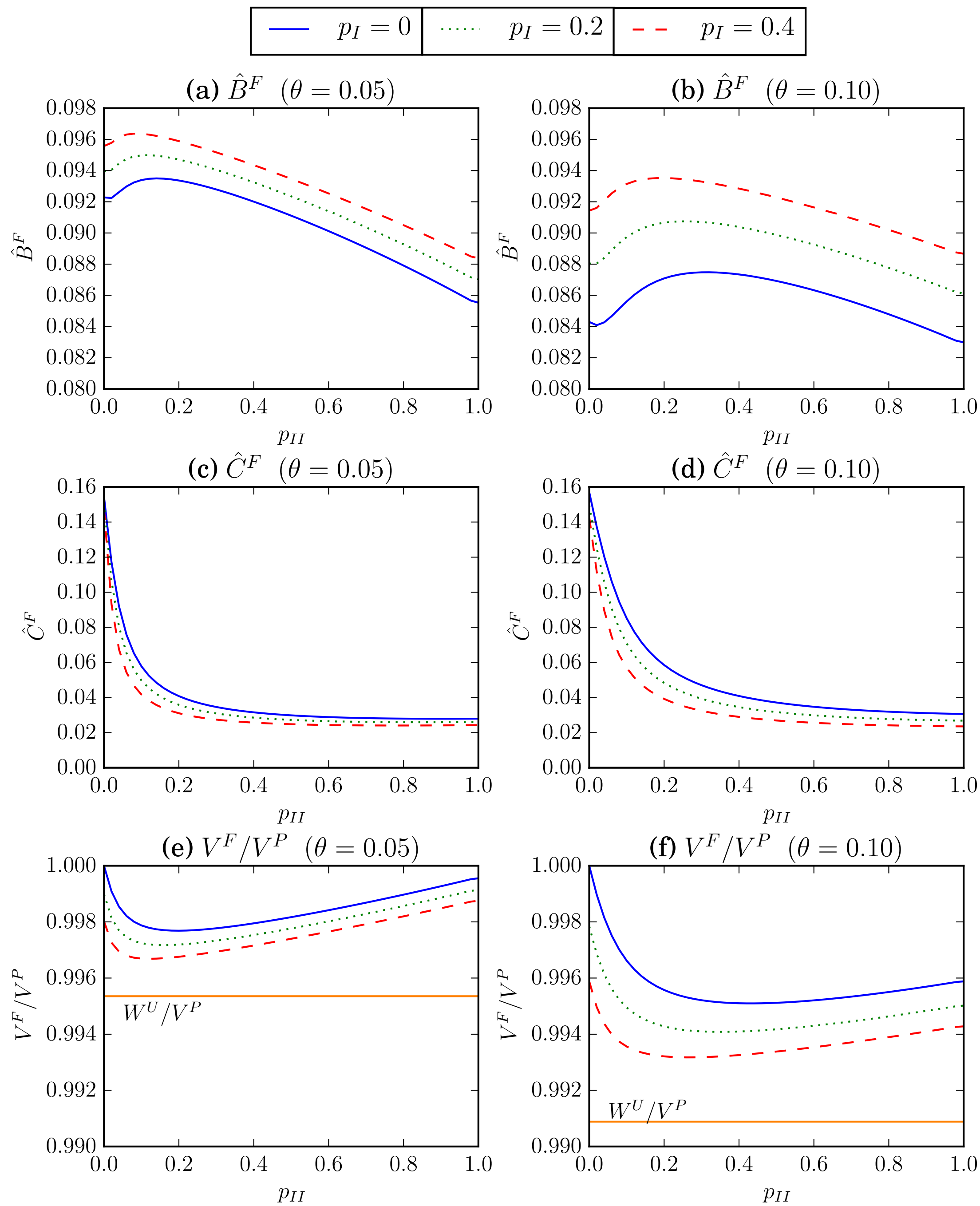
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.24: Full Enforcement Optima for $\mathcal{E} = 0.8$



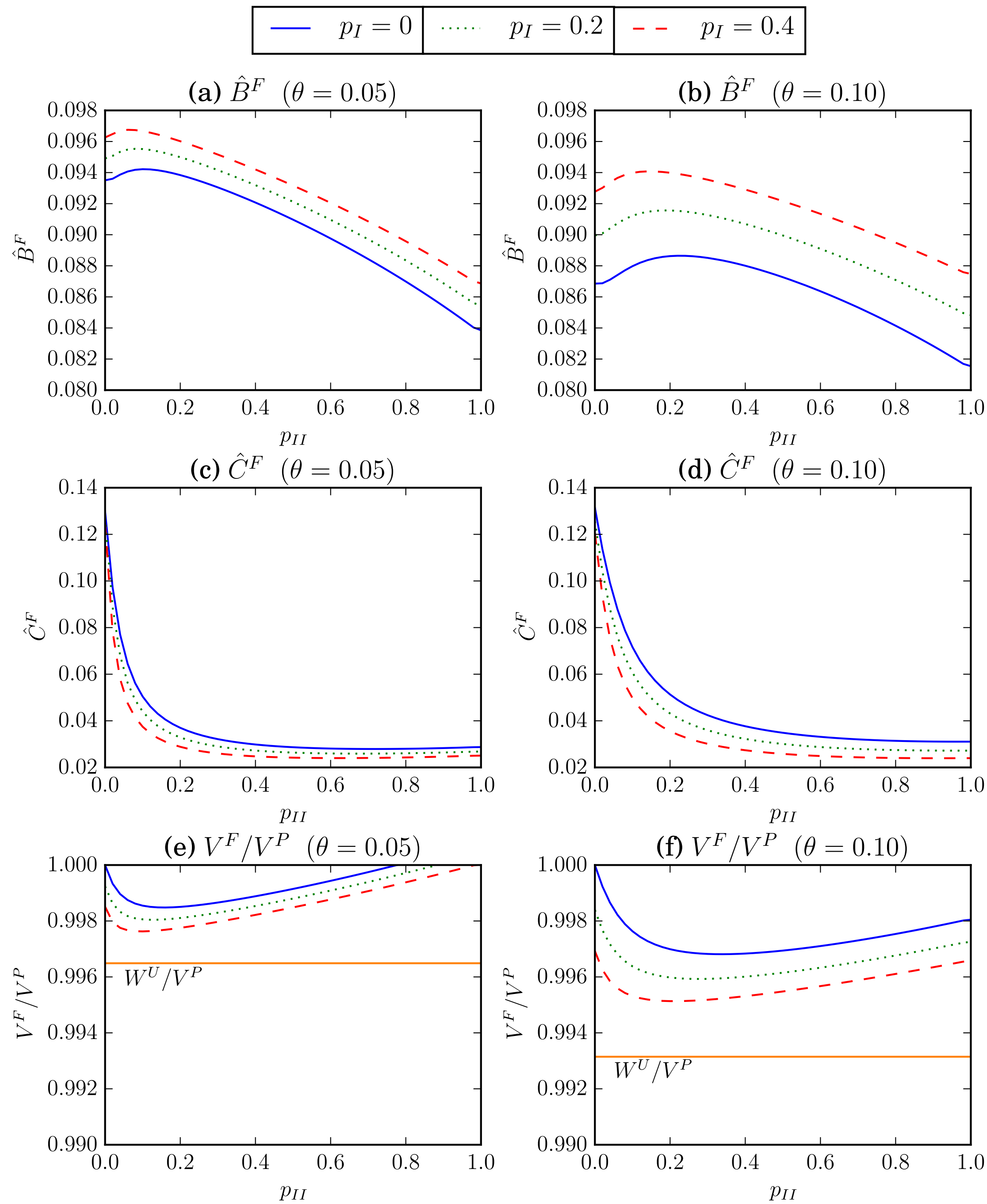
Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.25: Full Enforcement Optima for $\mathcal{E} = 0.9$



Notes. W^U/V^P is the horizontal line in panels (e) and (f).

Figure 2.26: Full Enforcement Optima for $\mathcal{E} = 0.99$



Notes. W^U/V^P is the horizontal line in panels (e) and (f).

subplot are generated for a different value of $p_I \in \{0, 0.2, 0.4\}$.[39] The subplots on the first row of each figure illustrate how the optimal universal benefit changes with p_{II}. Those on the second row illustrate how the optimal categorical benefit changes with p_{II}. Finally, those on the third row illustrate how optimal welfare as a fraction of the Perfect Discrimination welfare level changes with p_{II}.

Given our pre-emptive discussion of the multiple equilibria that arise when $\mathcal{E} = 0.5$, we will primarily focus the present discussion on Figures 2.22 to 2.26 (i.e. $\mathcal{E} = 0.6$ to $\mathcal{E} = 0.99$). In these figures the optimal benefit functions are smooth and continuous. With respect to Figure 2.21 ($\mathcal{E} = 0.5$), it is comforting to note that the optimal benefit functions exhibit discontinuities are the critical parameters (p_I, p_{II}) established in Figure 2.17. This is discussed further in the caption immediately immediately below the figure.

The first point to make is that the simulation results are consistent with the theory in Proposition 3a. That is, the optimal categorical benefit is *positive* for all levels of discriminatory power; including the case of no discriminatory power (i.e. $p_I + p_{II} = 1$).

Subplots (a)-(d) in each figure illustrate that an increase in p_I serves to increase $\hat{B}^F$ and lower $\hat{C}^F$. The intuition remains the same as before: because some unable individuals are incorrectly rejected the categorical benefit more resources are devoted towards the unconditional transfer to prevent the consumption of these individuals becoming too low (and consequently their smvi becoming too large).

The effect of an increase in p_{II} on the optimal benefit levels is more complicated. Figures 2.21 to 2.26 illustrate that $\hat{B}^F$ is non-monotonic in p_{II}, whilst $\hat{C}^F$ tends to fall with p_{II} throughout much of the interval $p_{II} \in (0, 1]$.[40] In more detail, $\hat{B}^F$ tends to (i) fall with p_{II} when p_{II} is small; (ii) rise with p_{II} as p_{II} is increased further; and may also (iii) fall with p_{II} at higher values of p_{II}. The initial fall in $\hat{B}^F$ limits the extent to which $\hat{C}^F$ falls with p_{II} at low values of p_{II}. The intuition would seem to be that when p_{II} is small few able individuals receive the categorical benefit and so a reduction in the universal benefit can be used to keep the categorical support near the level when $p_{II} = 0$. The extent of the initial fall in $\hat{B}^F$ will depend negatively on the size of p_I; and thus

[39] To ensure that the differences between each curve are visible we have restricted p_I to these values (i.e. we have omitted $p_I = 0.1$ and $p_I = 0.2$).

[40] At higher values of p_{II} the optimal categorical benefit may rise.

on the number of unable individuals who are incorrectly denied the categorical benefit and rely on the universal benefit for consumption. The increase in $\hat{B}^F$ as p_{II} increases further is accompanied by a necessary large decrease in $\hat{C}^F$. Here, the intuition would seem to be that it is optimal to lower the categorical benefit so as to restrict the number of able individuals who apply down to those of lower productivities.[41] Indeed, Table 2.1 illustrates that for each $\mathcal{E} \in \{0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$; $F(\bar{\bar{\omega}})$ falls with both p_I and p_{II}.

We now turn to discuss the effect of Type I and Type II classification errors on maximum welfare. In line with Proposition 3c subplots (e) and (f) in each figure illustrate that an increase in p_I lowers maximum welfare for all values of p_{II}. The intuition remains the same as under No Enforcement: Type I errors introduce horizontal inequity into the unble subpopulation and consequently raise their average smvi. Unlike the No Enforcement regime, however, the simulations illustrate that an increase in p_{II} may not be welfare decreasing. In particular, Figures 2.23 through to 2.26 illustrate that welfare is increasing in p_{II} for p_{II} sufficiently high. The intuition rests on the fact that 'leakage' of the categorical benefit into the able subpopulation still plays a redistributive role because it is received only by lower productivity able individuals. As p_{II} increases the welfare gain to this portion of the able subpopulation may become sufficiently large so as to offset any *further* welfare losses to the unable who are harmed by receiving lower transfers than they would under Perfect Discrimination.

This effect of Type II errors on maximum welfare is more pronounced the larger is $\mathcal{E}$ and the smaller is θ. The intuition for why welfare is more likely to be increasing in p_{II} at higher values of $\mathcal{E}$ can be seen from the sufficient condition in (2.46) and Table 2.1. Expression (2.46) implies higher values of $F(\bar{\omega})/F(\bar{\bar{\omega}})$ are more likely to give rise to the result that welfare is increasing in p_{II}; precisely because higher values of this ratio imply that able applicants for the categorical benefit are very similar to unable applicants in terms of their smvi. Table 2.1, meanwhile, clearly illustrates that at the optimum $F(\bar{\omega})/F(\bar{\bar{\omega}})$ is increasing $\mathcal{E}$. This effect is predominantly driven by the fact that the reservation productivity $\bar{\omega}$ is unambiguously increasing in $\mathcal{E}$ when $M \in (0, 1)$ (recall our discussion of Figure 2.20).

Comparing the subplots in the left and right columns of Figures 2.21 - 2.26 once

[41]Note that the effect of a reduction in the categorical benefit on the number of able individuals who choose to apply for the categorical benefit was not available under the No Enforcement regime because there all able individuals apply irrespective of the benefit size.

Table 2.1: Proportion of voluntarily unemployed and able applicants at the optimum.

	$F[\bar{\omega}(\hat{B}^F)]$			$F[\bar{\bar{\omega}}(\hat{B}^F, \hat{C}^F)]$			$F(\bar{\omega})/F(\bar{\bar{\omega}})$		
p_{II}	$\mathcal{E}=0.6$	$\mathcal{E}=0.8$	$\mathcal{E}=0.99$	$\mathcal{E}=0.6$	$\mathcal{E}=0.8$	$\mathcal{E}=0.99$	$\mathcal{E}=0.6$	$\mathcal{E}=0.8$	$\mathcal{E}=0.99$
(a) $p_I = 0$									
0.0	0.045	0.132	0.230	0.983	0.936	0.891	0.046	0.141	0.258
0.2	0.053	0.136	0.230	0.477	0.566	0.613	0.111	0.241	0.375
0.4	0.054	0.135	0.226	0.353	0.490	0.568	0.154	0.276	0.399
0.6	0.055	0.133	0.221	0.307	0.460	0.552	0.178	0.289	0.401
(b) $p_I = 0.1$									
0.0	0.047	0.134	0.231	0.982	0.935	0.890	0.048	0.143	0.260
0.2	0.054	0.138	0.232	0.452	0.550	0.603	0.119	0.250	0.384
0.4	0.055	0.136	0.228	0.340	0.479	0.560	0.162	0.284	0.406
0.6	0.055	0.134	0.223	0.297	0.452	0.546	0.186	0.297	0.408
(c) $p_I = 0.2$									
0.0	0.048	0.136	0.233	0.981	0.933	0.888	0.049	0.145	0.262
0.2	0.055	0.139	0.233	0.428	0.533	0.592	0.128	0.261	0.393
0.4	0.056	0.137	0.229	0.326	0.469	0.553	0.171	0.293	0.414
0.6	0.056	0.135	0.224	0.287	0.444	0.540	0.194	0.305	0.415
(d) $p_I = 0.3$									
0.0	0.05	0.137	0.234	0.98	0.931	0.887	0.051	0.148	0.264
0.2	0.056	0.14	0.234	0.403	0.517	0.58	0.138	0.271	0.403
0.4	0.056	0.139	0.23	0.312	0.458	0.545	0.18	0.302	0.422
0.6	0.056	0.136	0.225	0.277	0.436	0.534	0.203	0.313	0.422

Notes: These results were generated for parameter values $\theta = 0.05$ and $\beta = 0.1$.

more illustrates that an increase in θ tends to lower the universal benefit as more resources are devoted to categorical transfers. Note that this does *not* imply that the categorical benefit size necessarily increases, because targeted resources are spread over more unable individuals. Next, comparing Figures 2.21 - 2.26 illustrates that an increase in $\mathcal{E}$ serves to increase the universal transfer and lower the categorical transfer.

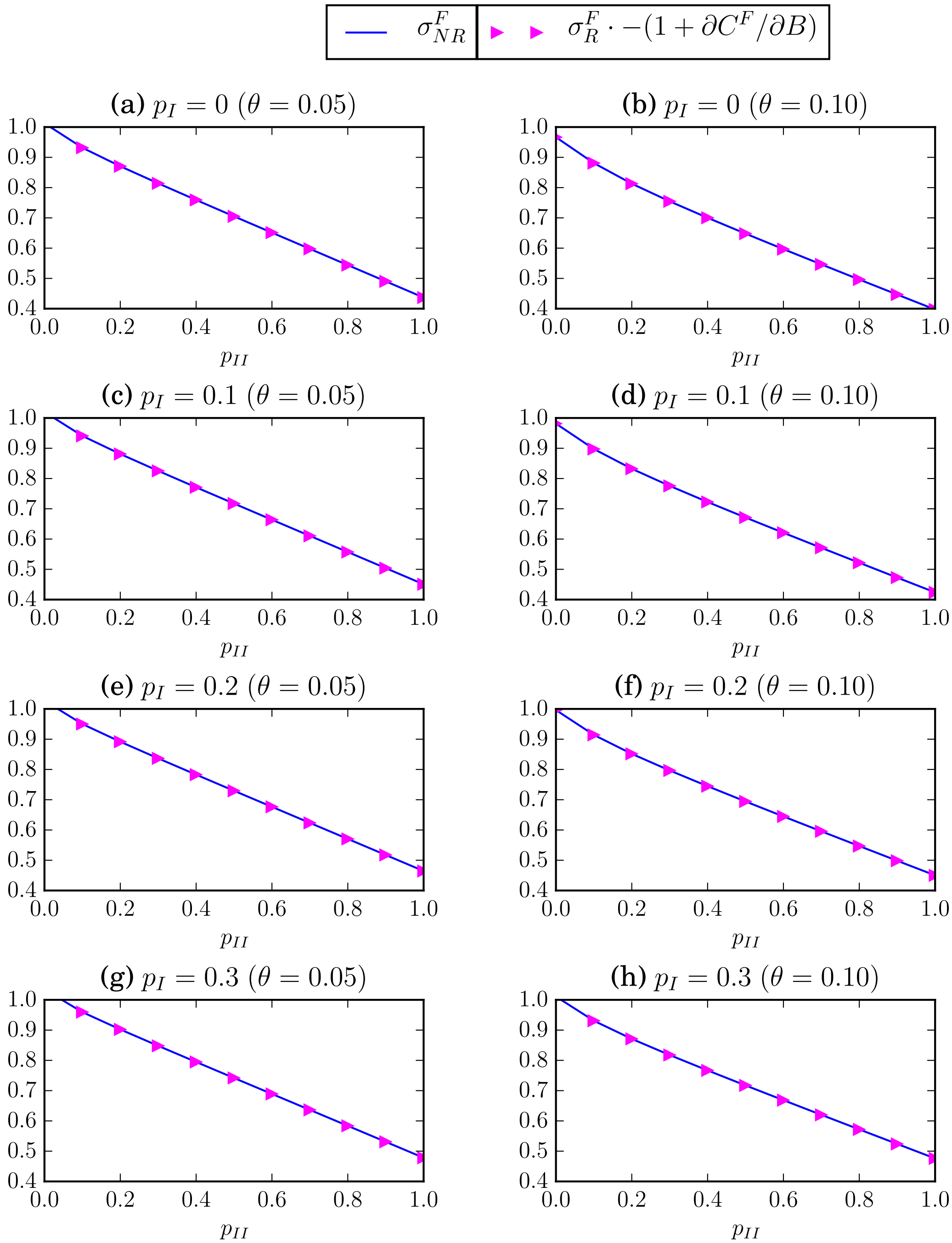
Finally, Figures 2.27 to 2.29 provide useful confirmation that the simulated Full Enforcement optima satisfy with *equality* the optimality condition in equation (2.39) of Proposition 3a. Figure 2.27 is drawn for $\mathcal{E} = 0.8$, Figure 2.28 is drawn for $\mathcal{E} = 0.9$, whilst Figure 2.29 is drawn for $\mathcal{E} = 0.99$. Each figure plots how both σ_{NR}^F and $\sigma_R^F \cdot -(1 + \partial C^F/\partial B)$ change with p_{II}, when evaluated at the optimal benefit levels. In all cases the two functions are equated.

2.4.4 Welfare Comparison Simulations

This section provides numerical examples of how maximum welfare under the No Enforcement regime compares with that under the Full Enforcement regime. Figures 2.30 to 2.32 display maximum welfare under the two regimes as a fraction of the Perfect Discrimination welfare level. Figure 2.30 is drawn for $\mathcal{E} = 0.8$; Figure 2.31 is drawn for $\mathcal{E} = 0.9$; whilst Figure 2.32 is drawn for $\mathcal{E} = 0.99$. On the horizontal axis in the each subplot p_{II} is varied from 0 to 1. The subplots on the left column are generated for $\theta = 0.05$, whilst those in the right column are generated for $\theta = 0.10$. The parameter p_I is increased as we move down through the subplots.

At higher levels of discriminatory power we observe cases where maximum welfare under No Enforcement exceeds that under Full Enforcement. In these cases the Full Enforcement system provides a lower benefit income to categorical recipients so as to reduce the number of able individuals who apply for the categorical benefit and, in turn, the number of quantity constrained able individuals in the economy. The welfare losses relative to No Enforcement for the unable are not offset by the redistributive welfare gains within the able subpopulation and, consequently, welfare under Full Enforcement is lower than under No Enforcement. Contrastingly, at lower levels of discriminatory power we observe that the Full Enforcement system outperforms the No Enforcement system in all cases. Note that in some of these cases the benefit income for categorical

Figure 2.27: σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ evaluated at the Optimum ($\mathcal{E} = 0.8$)



Notes. The purpose of this figure is to illustrate that the numerical results satisfy the optimality condition in Proposition 3a. That is, in each subplot (i) the aggregate smvi of non-categorical recipients is equated with (ii) the aggregate smvi of categorical recipients multiplied by the reduction in their total benefit income associated with an increase in the universal benefit.

Figure 2.28: σ^F_{NR} and $\sigma^F_R \cdot -(1 + \partial C^F/\partial B)$ evaluated at the Optimum ($\mathcal{E} = 0.9$)

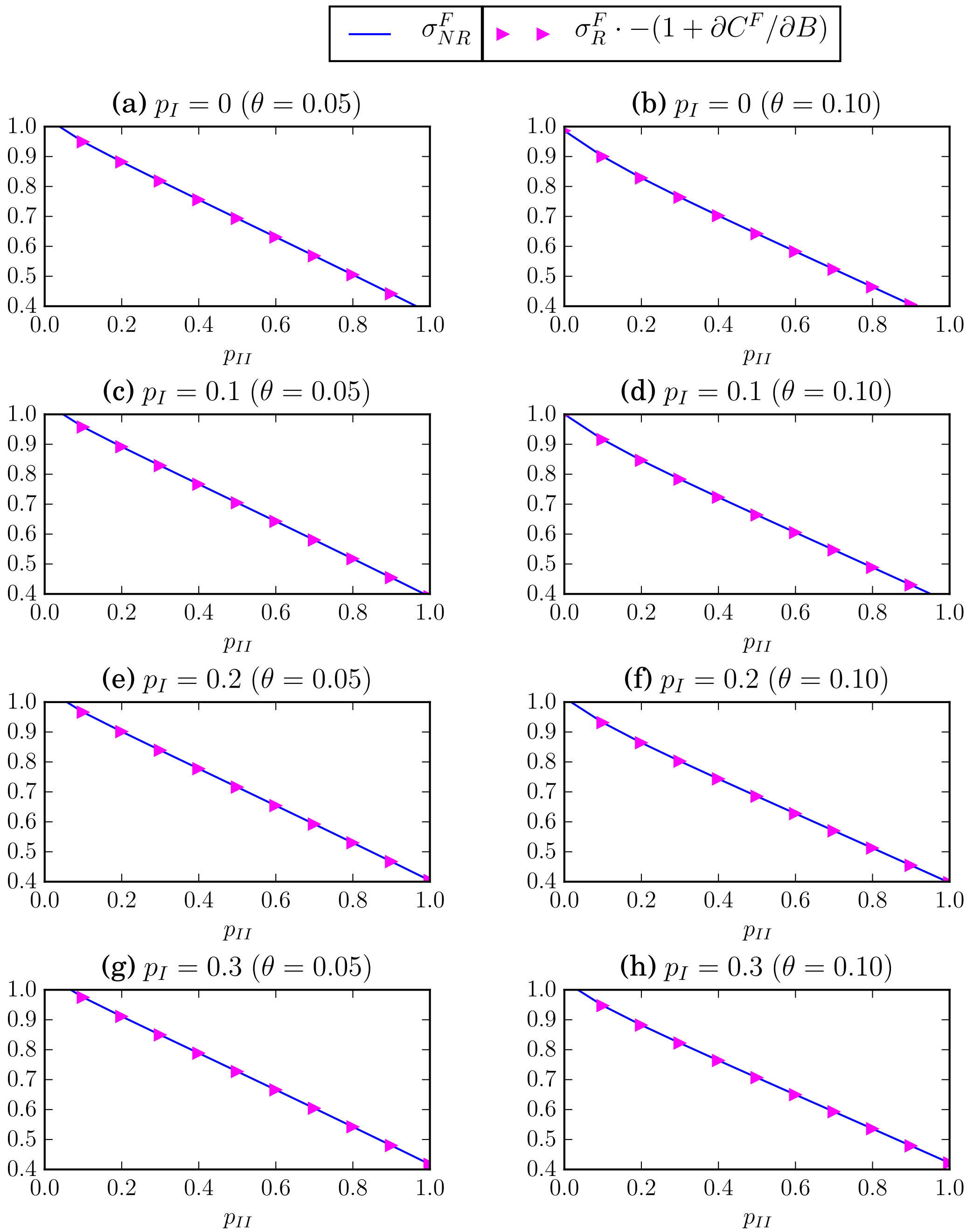


Figure 2.29: σ_{NR}^F and $\sigma_R^F \cdot -(1 + \partial C^F/\partial B)$ evaluated at the Optimum ($\mathcal{E} = 0.99$)

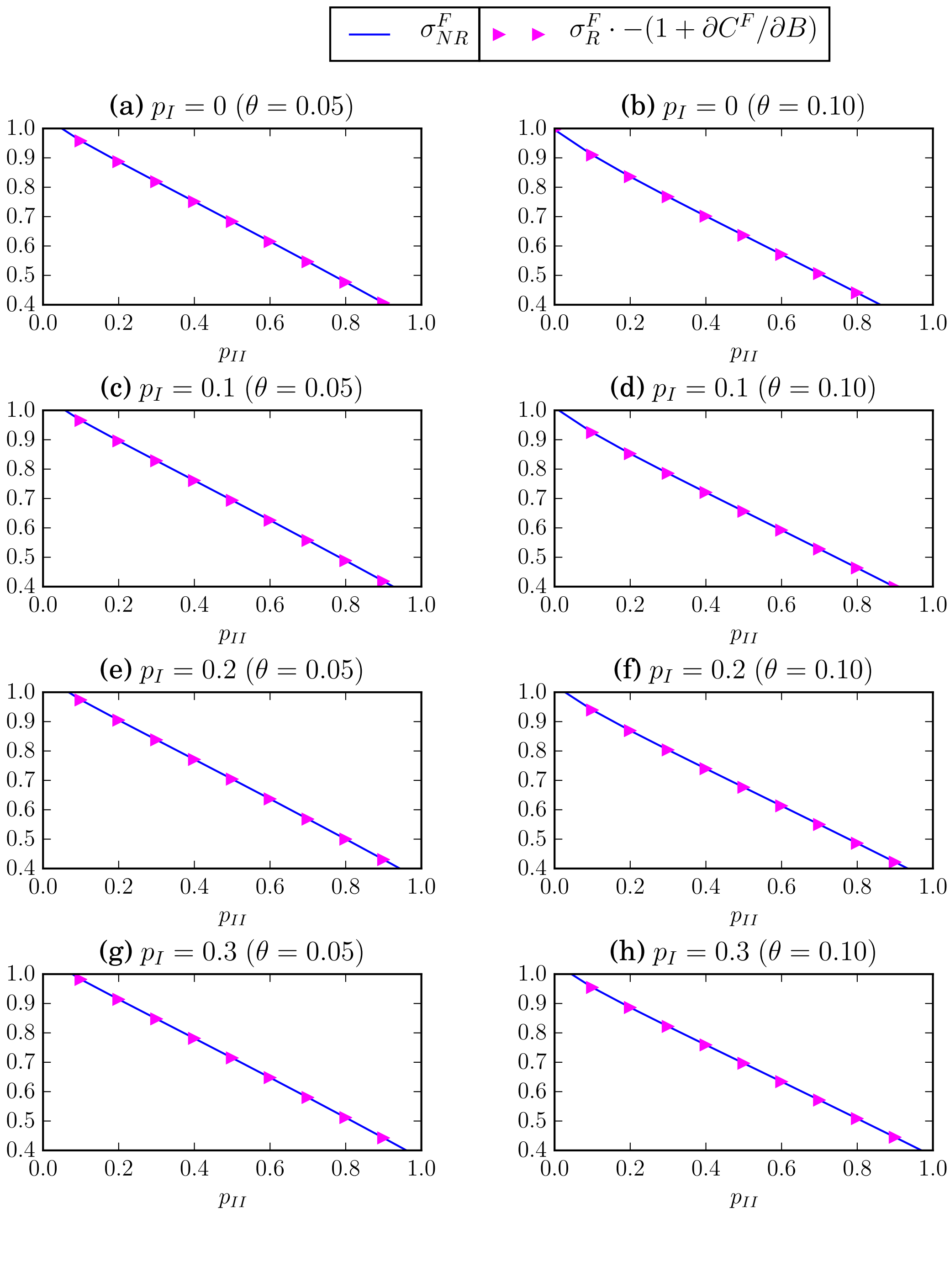


Figure 2.30: Welfare Comparision $\mathcal{E} = 0.8$, $\beta = 0.1$.

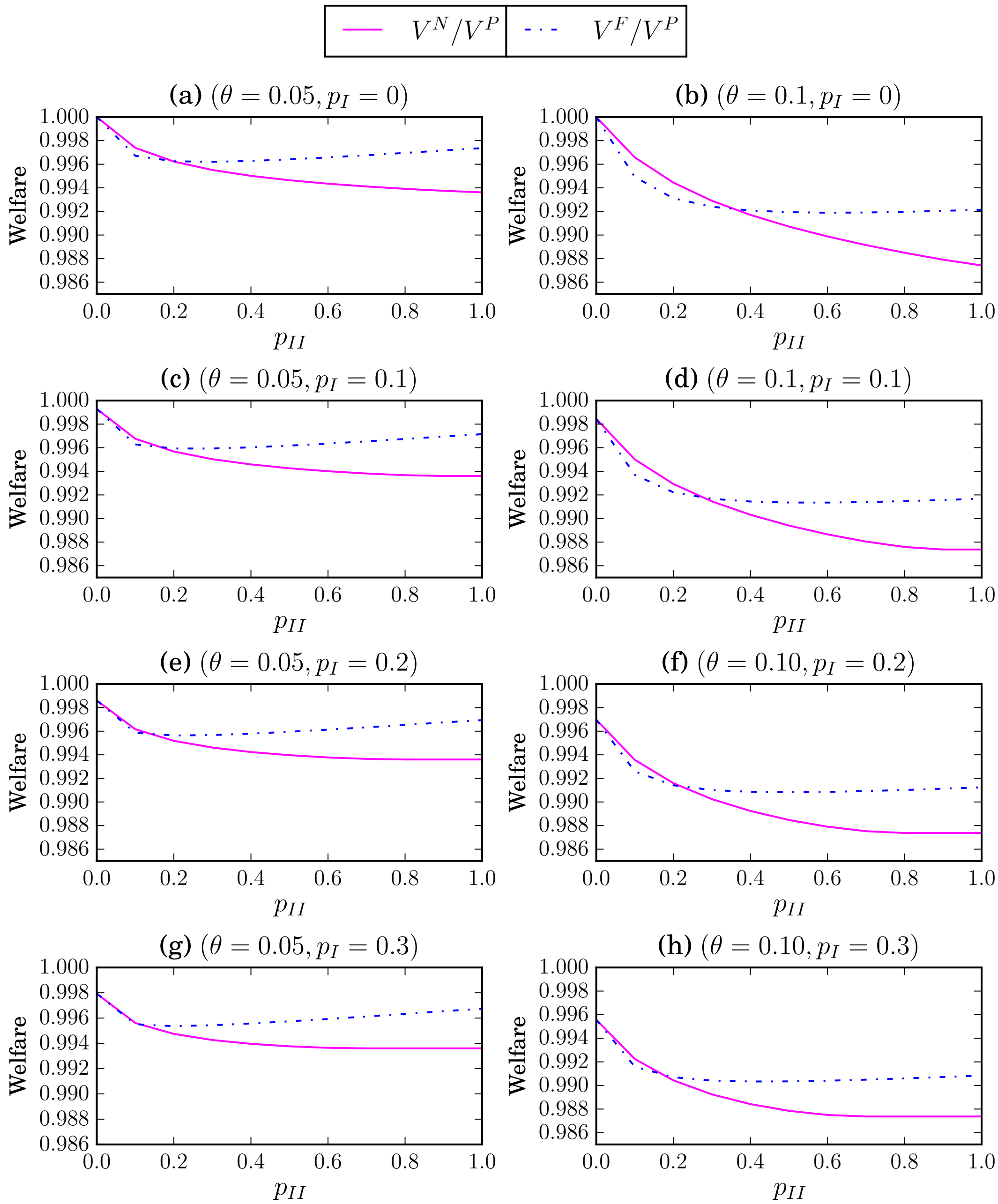


Figure 2.31: Welfare Comparision $\mathcal{E} = 0.9$, $\beta = 0.1$.

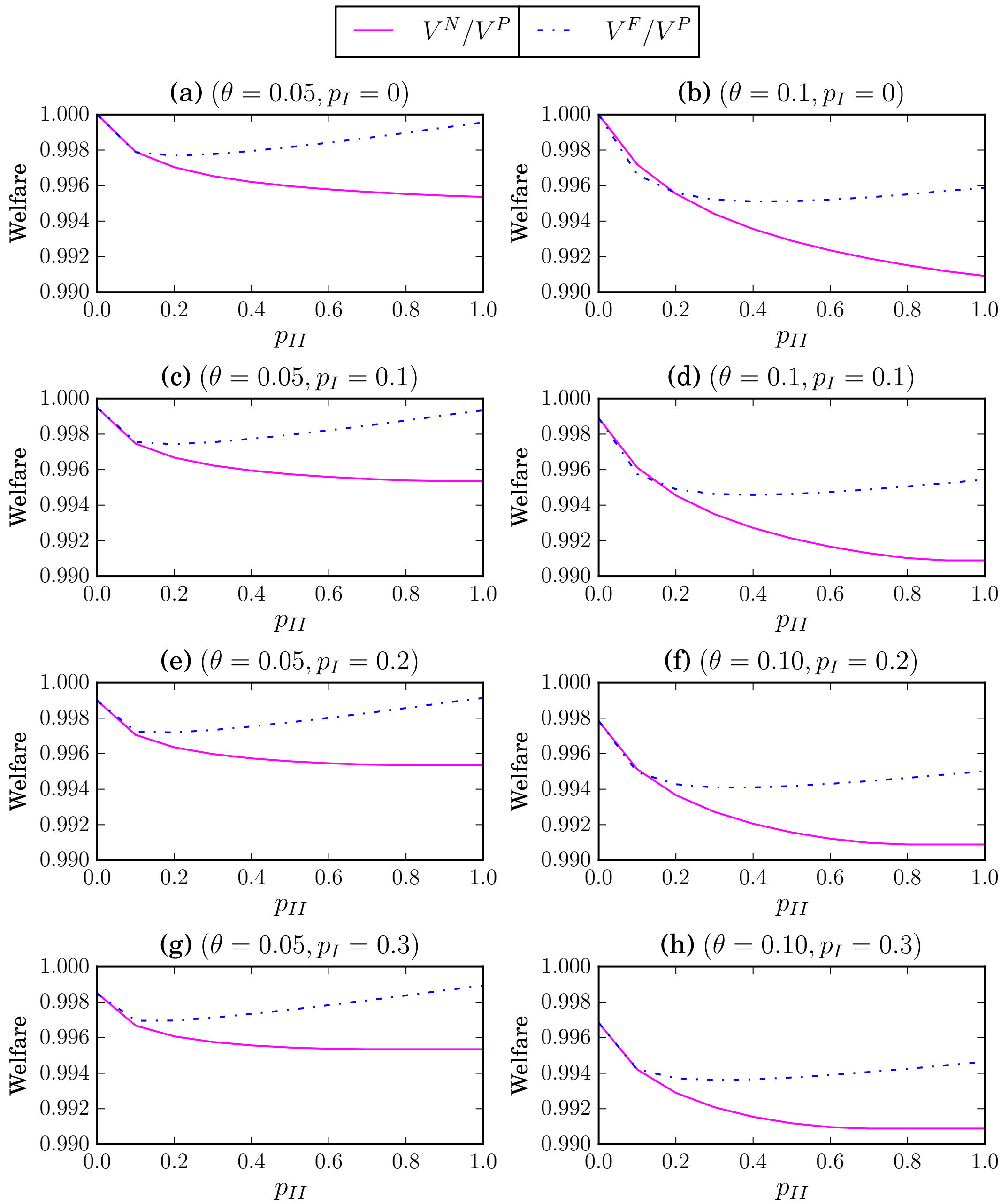
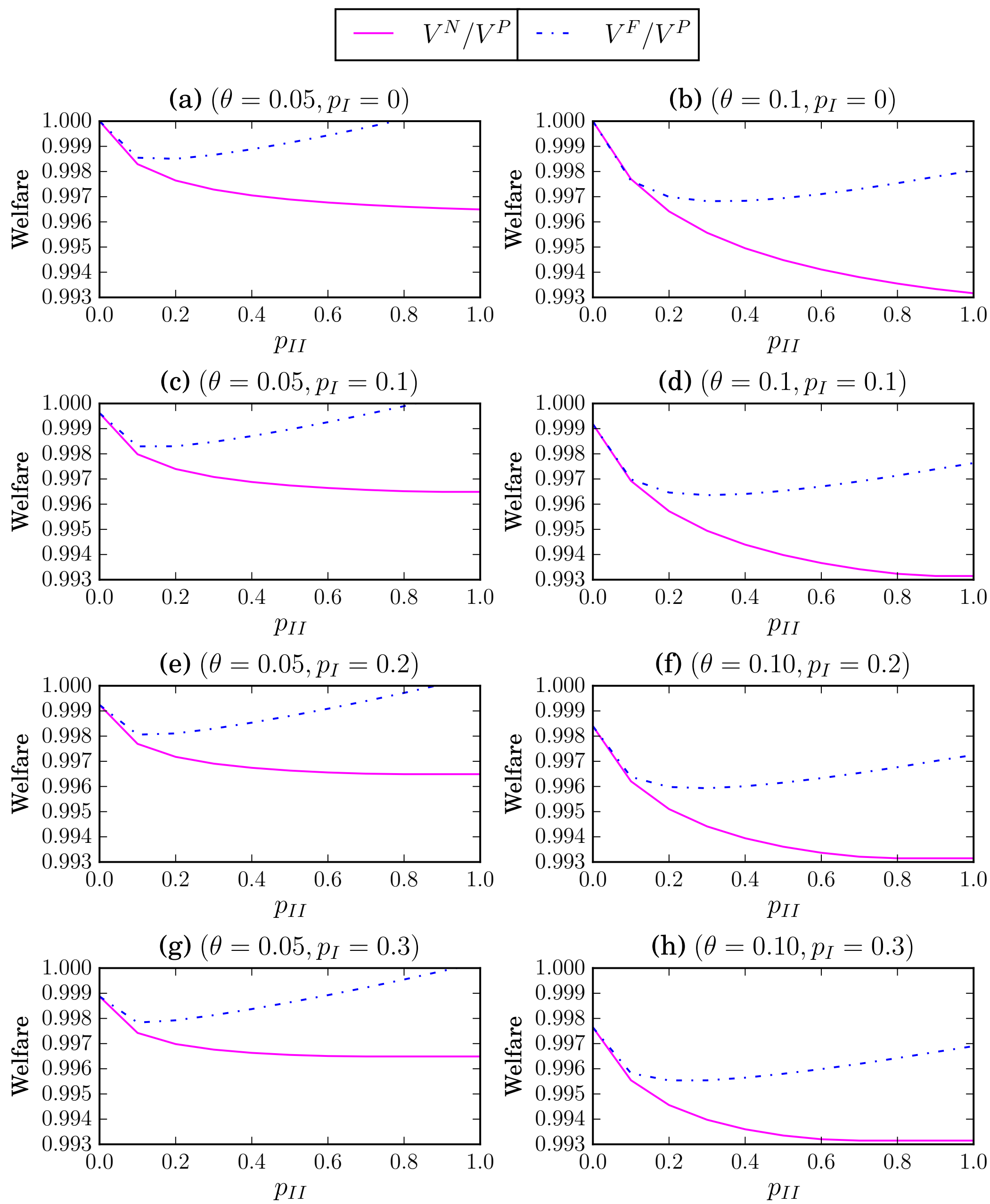


Figure 2.32: Welfare Comparision $\mathcal{E} = 0.99$, $\beta = 0.1$.



recipients is still lower than under No Enforcement, but the redistributive gains within the able subpopulation are sufficiently large.

2.5 Concluding Remarks

Real world targeted benefits typically feature what this chapter terms *double conditionality*: *ex-ante* the award of a benefit is conditioned on an applicant meeting some initial eligibility conditions, whilst *ex-post* a recipient must typically conform with certain behavioural requirements. For example, an *ex-ante* condition for being awarded a disability benefit may be that an individual faces significant physical or mental difficulties in working; whilst an *ex-post* condition may be that a recipient does not work when receiving the disability benefit.

In large and complex welfare systems where *ex-ante* eligibility may be difficult to ascertain, the test awarding benefits is likely to be susceptible to, in statistical parlance, classification errors of both Type I (false rejection) and Type II (false award). The latter error type gives rise to enforcement issues because benefits are awarded to individuals for whom they are not intended, and thus individuals who may choose not to comply with the *ex-post* requirements in place. Both the propensity of an awards test to make classification errors and the effectiveness of the enforcement of *ex-post* requirements will influence the design of welfare programmes, in particular with regard to the weight placed on targeted transfers over those that are made universally available.

The literature on optimal transfers administered with classification errors typically assumes specific enforcement structures in relation to *ex-post* conditionality: recipients of a disability benefit that is targeted at those unable to work are either allowed to work (as in Parsons, 1996); or not allowed to work (as in Salanié, 2002). Moreover, the models employed differ sufficiently so as not to allow for a systematic comparison across enforcement regimes. There has also been little attention devoted to understanding how the propensity of the awards test to make classification errors of either type affects maximum social welfare and, further, how this differs depending on the enforcement regime in place.

In relation to these gaps in the literature, this chapter posed three questions at the outset. First, how does the propensity of an awards test to make classification errors of Type I and/or Type II affect the choice between (i) a pure universal system; (ii) a partial universal system where *both* targeted and universal benefits are provided; and (iii) a purely targeted welfare system? Further, conditioning on providing both benefits, how does the propensity to make these errors affect the optimal levels of the universal and targeted benefits, respectively? Second, how does an increase in the propensity to make either type of error affect maximum social welfare? Third, how do the answers to both of these questions depend on how effectively the *ex-post* conditionality is enforced?

To answer these questions this chapter has considered a framework where individuals differ in both their ability to work - modelled as zero quantity constraint on labour supply - and, conditional on being able to work, their productivity when at work. Given a fixed benefit budget the government chooses the optimal combination of (i) a universal benefit received unconditionally by all; and (ii) a categorical benefit that is imperfectly targeted at the unable and has an *ex-post* no-work requirement. We analyse the cases where the *ex-post* no-work requirement is either not enforced at all, or is fully enforced.

The principal messages are as follows:

- Under a *No Enforcement* regime:

    1. It is optimal to provide a categorical benefit whenever the test administering it has some discriminatory power. In this case the optimal benefits are chosen so as to equate - if budget feasible - the average smvi of categorical recipients with the average smvi of those not receiving the categorical benefit. A positive propensity to make Type I errors guarantees the provision of a universal benefit, such that a *partial universal system* is optimal. However, if Type I errors are never made and the benefit budget is not large enough to finance categorical transfers to the level that equates the average smvi of categorical recipients with that of non-recipients, a *purely targeted system* will be optimal. Finally, if the awards test has no discriminatory power a *pure universal system* is chosen.

    2. Maximum social welfare is decreasing in the propensity to make both Type I and Type II classification errors.

- Under a *Full Enforcement* regime:

1. It is optimal to provide a categorical benefit at *all* levels of discriminatory power, thus including the case of no discriminatory power. It is therefore never optimal to adopt a pure universal system. The optimal benefits are chosen to equate (i) the aggregate smvi of those not receiving the categorical benefit with (ii) the aggregate smvi of categorical recipients *multiplied* by the increase in their total benefit income per unit reduction in the universal benefit. A positive propensity to make Type I errors once more guarantees the provision of a universal benefit, such that a *partial universal system* is optimal.

2. Whilst maximum welfare is unambiguously decreasing in the propensity to make Type I errors, there are conditions under which it can be *increasing* in the propensity to make Type II errors. In particular, this is more likely to arise the larger is the fraction of able applicants who are voluntarily unemployed (due to the universal benefit), as these individuals have the same smvi as unable applicants.

Appendix A Proofs

Proof of Proposition 1

Solving the maximisation problem in (2.12) yields the following FOCs characterising $\hat{B}^P$ and $\hat{C}^P$:

$$\begin{aligned} (B):&\quad \theta u_x(\hat{B}^P + \hat{C}^P, 1) + (1-\theta)\int_0^\infty v_M(\omega, \hat{B}^P) f(\omega) d\omega \leq \hat{\lambda}^P \quad, \hat{B}^P \geq 0 \\ (C):&\quad u_x(\hat{B}^P + \hat{C}^P, 1) \leq \hat{\lambda}^P \quad, \hat{C}^P \geq 0 \end{aligned} \tag{A.1}$$

The pairs of inequalities hold with complementary slackness and λ denotes the shadow price of public expenditure. Given that the budget constraint must be exhausted (i.e. $\hat{B}^P + \theta\hat{C}^P = \beta$), we now test the following two hypotheses:

(i) $\hat{B}^P = \beta, \hat{C}^P = 0$ **(Pure universal system)**

If $\hat{B}^P = \beta$ and $\hat{C}^P = 0$ the FOCs in (A.1) become:

$$\begin{aligned} \theta u_x(\beta, 1) + (1-\theta)\int_0^\infty v_M(\omega, \beta) f(\omega) d\omega = \lambda \\ u_x(\beta, 1) \leq \lambda \end{aligned}$$

Combining both equations gives the contradictory statement:

$$\int_0^\infty v_M(\omega, \beta) f(\omega) d\omega \geq u_x(\beta, 1)$$

It cannot hold that the average smvi of the able weakly exceeds that of the unable when both subpopulations receive the same benefit income. The assertion that $\hat{C}^P = 0$ must therefore be false. Instead, we must have $\hat{C}^P > 0$. The shadow price of public expenditure is therefore equal to the smvi for the unable at the optimum.

(ii) $\hat{B}^P = 0, \hat{C}^P = \beta/\theta$ **(Pure targeted system)**

If $\hat{B}^P = 0$ and $\hat{C}^P = \beta/\theta$ the FOCs in (A.1) become:

$$\theta u_x\left(\frac{\beta}{\theta}, 1\right) + (1-\theta)\int_0^\infty v_M(\omega, 0) f(\omega) d\omega \leq \lambda$$
$$u_x\left(\frac{\beta}{\theta}, 1\right) = \lambda$$

Combining both equations gives:

$$\int_0^\infty v_M(\omega, 0) f(\omega) d\omega \leq u_x\left(\frac{\beta}{\theta}, 1\right)$$

The right side is unambiguously decreasing in β. Further, given that $\lim_{x\to 0} u(x,l) = +\infty$ it must hold that there is a critical budget level $\bar{\beta}^P$ satisfying:

$$\int_0^\infty v_M(\omega, 0) f(\omega) d\omega \equiv u_x\left(\frac{\bar{\beta}^P}{\theta}, 1\right)$$

Proof of Proposition 2a

Solving the maximisation problem in (2.20) yields the FOCs:

$$\begin{aligned}(B): \quad & \theta\left[(1-p_I)u_x(\hat{B}^N + \hat{C}^N, 1) + p_I u_x(\hat{B}^N, 1)\right] \\ & + (1-\theta)\int_0^\infty \left[p_{II} v_M(\omega, \hat{B}^N + \hat{C}^N) + (1-p_{II}) v_M(\omega, \hat{B}^N)\right] f(\omega) d\omega \\ & \leq \hat{\lambda}^N \;\; ; \;\; \hat{B}^N \geq 0\end{aligned} \tag{A.2}$$

$$\begin{aligned}(C): \quad & \frac{\theta(1-p_I)u_x(\hat{B}^N + \hat{C}^N, 1) + (1-\theta)p_{II}\int_0^\infty v_M(\omega, \hat{B}^N + \hat{C}^N) f(\omega) d\omega}{\theta(1-p_I) + (1-\theta)p_{II}} \\ & \leq \hat{\lambda}^N \;\; ; \;\; \hat{C}^N \geq 0\end{aligned} \tag{A.3}$$

where the pairs of inequalities in (A.2) and (A.3) hold with complementary slackness. Note that in deriving these FOCs we have used property that $u(M, 1) \equiv v[\bar{\omega}(M), M]$ in differentiating the integral limits (Leibniz rule).

The left side of (A.2) is the average smvi over the entire population. This captures how much, on average, welfare will increase with a marginal increase in the universal benefit. On the right side is the shadow price of public expenditure ($\hat{\lambda}^N$) multiplied by 1, which is the marginal expenditure cost of increasing each individual's universal transfer (in an economy of size 1).

The left side of (A.3) is the average smvi of categorical recipients, as composed of both unable and able individuals. Notice that this is an average because the denominator - i.e. $\theta(1-p_I)+(1-\theta)p_{II}$ - is the total number of categorical recipients in the economy. This captures how much, on average, welfare will increase with an increase in the categorical transfer size. On the right side is the shadow price of public expenditure multiplied by unity (the *per capita* marginal cost associated with increasing the categorical benefit).

With respect to the pairs of inequalities in (A.2) and (A.3), we proceed to test the following two hypotheses: (i) ($\hat{B}^N > 0$, $\hat{C}^N = 0$) ; and (ii) ($\hat{B}^N = 0$, $\hat{C}^N > 0$).

(i) $\hat{B}^N = \beta, \hat{C}^N = 0$ (Pure universal system)

If $\hat{C}^N = 0$ then it must hold from the budget constraint $(B+[\theta(1-p_I)+(1-\theta)p_{II}]C = \beta)$ that $\hat{B}^N = \beta$. In this case, the FOCs for B and C reduce to:

$$\theta u_x(\beta,1)+(1-\theta)\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega = \lambda$$
$$\frac{\theta(1-p_I)u_x(\beta,1)+(1-\theta)p_{II}\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega}{[\theta(1-p_I)+(1-\theta)p_{II}]} \leq \lambda$$

Writing the first of these equations as $\theta u_x(\beta,1) = \lambda-(1-\theta)\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega$ and substituting into the second gives:

$$\frac{(1-p_I)\left[\lambda-(1-\theta)\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega\right]+(1-\theta)p_{II}\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega}{[\theta(1-p_I)+(1-\theta)p_{II}]} \leq \lambda$$
$$\Rightarrow \frac{(1-p_I)(1-\theta)\left[\lambda-\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega\right]}{p_{II}(1-\theta)\left[\lambda-\int_0^\infty v_M(\omega,\beta)f(\omega)d\omega\right]} \leq 1$$
$$\Rightarrow 1-p_I \leq p_{II}$$

Given our discriminatory power assumption (i.e. $p_I+p_{II} \leq 1$), this condition can only hold with equality, and thus when $p_I+p_{II} = 1$. It follows that $\hat{C}^N = 0$ only if the test awarding C has no discriminatory power. Otherwise, $\hat{C}^N > 0 \;\forall\; p_I+p_{II} < 1$.

It follows that at any optimum where $p_I+p_{II} < 1$ the average smvi of categorical recipients should equate with the shadow price of public expenditure:

$$\frac{\theta(1-p_I)u_x(\hat{B}^N+\hat{C}^N,1)+(1-\theta)p_{II}\int_0^\infty v_M(\omega,\hat{B}^N+\hat{C}^N)f(\omega)d\omega}{\theta(1-p_I)+(1-\theta)p_{II}}=\hat{\lambda}^N \tag{A.4}$$

from which we can immediately ascertain that, by diminishing marginal utility of income:

$$u_x(\hat{B}^N+\hat{C}^N,1)\geq\hat{\lambda}^N>\int_0^\infty v_M(\omega,\hat{B}^N+\hat{C}^N)f(\omega)d\omega \tag{A.5}$$

where $u_x(\hat{B}^N+\hat{C}^N,1)=\hat{\lambda}^N$ only if $p_{II}=0$.

Substituting (A.4) into (A.2) then gives:

$$\begin{aligned}&\frac{\theta p_I u_x(\hat{B}^N,1)+(1-\theta)(1-p_{II})\int_0^\infty v_M(\omega,\hat{B}^N)f(\omega)d\omega}{\theta p_I+(1-\theta)(1-p_{II})}\\ &\leq\frac{\theta(1-p_I)u_x(\hat{B}^N+\hat{C}^N,1)+(1-\theta)p_{II}\int_0^\infty v_M(\omega,\hat{B}^N+\hat{C}^N)f(\omega)d\omega}{\theta(1-p_I)+(1-\theta)p_{II}}\quad;\quad\hat{B}^N\geq 0\\ &=\hat{\lambda}^N\end{aligned} \tag{A.6}$$

which in turn implies that, by diminishing marginal utility of income:

$$u_x(\hat{B}^N,1)>\hat{\lambda}^N\geq\int_0^\infty v_M(\omega,\hat{B}^N)f(\omega)d\omega \tag{A.7}$$

where $\int_0^\infty v_M(\omega,\hat{B}^N)f(\omega)d\omega=\hat{\lambda}^N$ *only* if $p_I=0$ and $\hat{B}^N>0$.

(ii) $\hat{B}^N=0,\hat{C}^N=\beta/[\theta(1-p_I)+(1-\theta)p_{II}]$ **(Pure targeted system)**

From the budget constraint ($B+[\theta(1-p_I)+(1-\theta)p_{II}]C=\beta$) it follows that if $B=0$ then $C=\beta/[\theta(1-p_I)+(1-\theta)p_{II}]$. First off, it must hold that $\hat{B}^N>0$ whenever $p_I>0$ because $\lim_{x\to 0}u_x(x,l)=+\infty$. Suppose otherwise, then the left side of (A.2) or (A.6) blows up to infinity whenever some unable individuals have zero income to consume. Next, if $p_I = 0$ but $p_{II} \geq 0$ then (A.6) becomes:

$$\int_0^\infty v_M(\omega, 0) f(\omega) d\omega$$
$$\leq \frac{\theta u_x \left(\frac{\beta}{\theta + (1-\theta)p_{II}}, 1 \right) + (1-\theta)p_{II} \int_0^\infty v_M \left(\omega, \frac{\beta}{\theta + (1-\theta)p_{II}} \right) f(\omega) d\omega}{\theta + (1-\theta)p_{II}}$$

The left side is independent of β, whilst the right side is unambiguously decreasing in β. Suppose that $\beta \to 0$, then $\lim_{\beta \to 0} u_x(\beta/[\theta + (1-\theta)p_{II}], 1) = +\infty$ such that the right side approaches $+\infty$ and the condition must hold with strict inequality. There must therefore be a critical budget level $\bar{\beta}^N$ for which the condition holds with equality. Using the definition of $\bar{\beta}^P$ from (2.15), we can implicitly define $\bar{\beta}^N$ by:

$$u_x \left(\frac{\bar{\beta}^P}{\theta}, 1 \right) \equiv \frac{\theta u_x \left(\frac{\bar{\beta}^N}{\theta + (1-\theta)p_{II}}, 1 \right) + (1-\theta)p_{II} \int_0^\infty v_M \left(\omega, \frac{\bar{\beta}^N}{\theta + (1-\theta)p_{II}} \right) f(\omega) d\omega}{\theta + (1-\theta)p_{II}}$$

It is straightforward to see that this implies $u_x(\bar{\beta}^N/[\theta + (1-\theta)p_{II}], 1) > u_x(\bar{\beta}^P/\theta, 1) \Rightarrow \bar{\beta}^N < \bar{\beta}^P \cdot [\theta + (1-\theta)p_{II}]/\theta$. Note that $p_{II} = 0 \Rightarrow \bar{\beta}^P = \bar{\beta}^N$.

Proof of Proposition 2b.

To establish that $\delta^N > 0$ at the optimum whenever $p_I > 0$ and/or $p_{II} > 0$, we proceed in three cases: (i) ($p_I > 0$, $p_{II} = 0$) ; (ii) ($p_I = 0$, $p_{II} > 0$) ; and (iii) ($p_I > 0$, $p_{II} > 0$).

(i) ($p_I > 0$**,** $p_{II} = 0$)

In this case $\delta^N = [p_I u_x(B, 1) + (1-p_I) u_x(B+C, 1)] - \int_0^\infty v_M(\omega, B) f(\omega) d\omega$. From (A.5) and (A.7) we have $u_x(\hat{B}^N + \hat{C}^N, 1) = \hat{\lambda}^N > \int_0^\infty v_M(\omega, \hat{B}^N) f(\omega) d\omega$. Further, by diminishing marginal utility of income $u_x(\hat{B}^N, 1) > u_x(\hat{B}^N + \hat{C}^N, 1)$ $\forall p_I < 1$ and so $\delta^N > 0$ for $p_I < 1$. If $p_I = 1$ then $\hat{B}^N = \beta$ and $\hat{C}^N = 0$ such that $u_x(\beta, 1) > \int_0^\infty v_M(\omega, \beta) f(\omega) d\omega$. So $\delta^N > 0$ when $p_I = 1$.

(ii) ($p_I = 0$ **,** $p_{II} > 0$)

In this case $\delta^N = u_x(B+C,1) - \int_0^\infty [p_{II} v_M(\omega, B+C) + (1-p_{II}) v_M(\omega, B)] f(\omega) d\omega$. From (A.5) and (A.7):

$$u_x(\hat{B}^N + \hat{C}^N, 1) > \hat{\lambda}^N \geq \int_0^\infty v_M(\omega, \hat{B}^N) f(\omega) d\omega > \int_0^\infty v_M(\omega, \hat{B}^N + \hat{C}^N) f(\omega) d\omega$$

So for all $p_{II} < 1 : \delta^N > 0$. Once more, if $p_{II} = 1$ then welfare provision effectively takes a pure universal form and $\delta^N > 0$ because $u_x(\beta, 1) > \int v_M(\omega, \beta) f(\omega) d\omega$.

(iii) $(p_I > 0,\ p_{II} > 0)$

In this case $\delta^N = [p_I u_x(B,1) + (1-p_I) u_x(B+C,1)] - \int_0^\infty [p_{II} v_M(\omega, B+C) + (1 - p_{II}) v_M(\omega, B)] f(\omega) d\omega$. From (A.5) and (A.7):

$$u_x(\hat{B}^N, 1) > u_x(\hat{B}^N + \hat{C}^N, 1) > \hat{\lambda}^N > \int_0^\infty v_M(\omega, \hat{B}^N) f(\omega) d\omega > \int_0^\infty v_M(\omega, \hat{B}^N + \hat{C}^N) f(\omega) d\omega \tag{A.8}$$

and thus $\delta^N > 0$ at the optimum. If $p_I + p_{II} = 1$ then a pure universal system is chosen and $\delta^N > 0$ because $u_x(\beta, 1) > \int v_M(\omega, \beta) f(\omega) d\omega$.

In summary then, we have $\delta^N > 0$ whenever $p_I > 0$ and/or $p_{II} > 0$. Q.E.D.

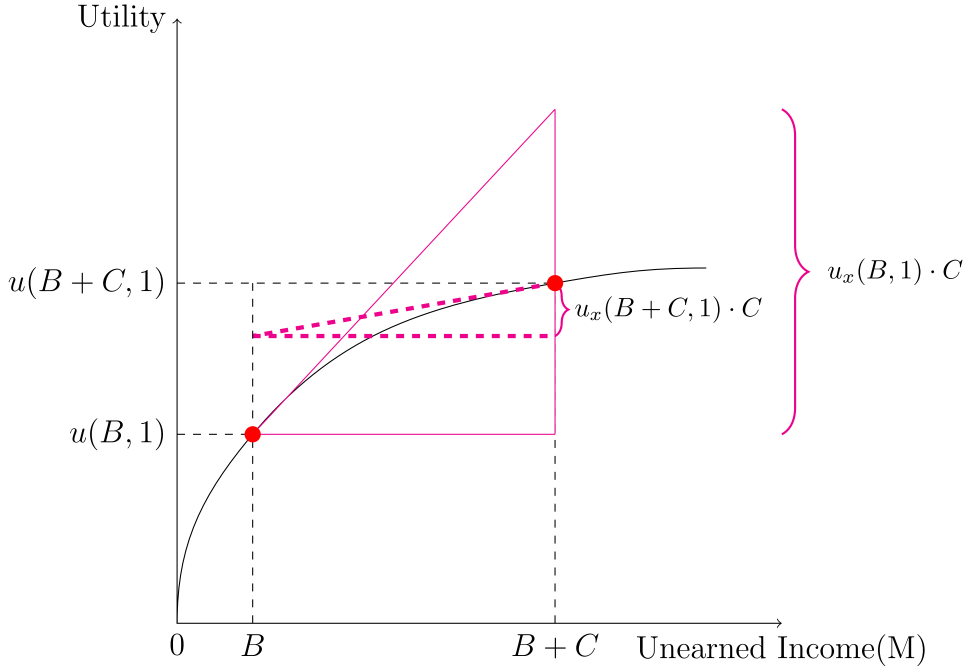
Proof of Proposition 2c

To establish the effect of classification errors on social welfare, we make use of the following standard property of concave functions:

$$u_x(B,1) \cdot C > u(B+C,1) - u(B,1) > u_x(B+C,1) \cdot C \tag{A.9}$$

Figure 2.33 illustrates this. From (2.26) in the main text and the above property in

Figure 2.33: Property of Concave Utility



(A.9), it must therefore hold that:

$$\frac{\partial V^N}{\partial p_I} = \theta \left\{ \left[u(\hat{B}^N, 1) - u(\hat{B}^N + \hat{C}^N, 1) \right] + \hat{\lambda}^N \hat{C}^N \right\} < \theta \hat{C}^N \left[\hat{\lambda}^N - u_x(\hat{B}^N + \hat{C}^N, 1) \right] \leq 0$$

$$\begin{aligned} \frac{\partial V^N}{\partial p_{II}} &= (1-\theta) \left\{ \int_0^\infty \left[v(\omega, \hat{B}^N + \hat{C}^N) - v(\omega, \hat{B}^N) \right] f(\omega) d\omega - \hat{\lambda}^N \hat{C}^N \right\} \\ &< \theta \hat{C}^N \left[\int_0^\infty v_M(\omega, \hat{B}^N) f(\omega) d\omega - \hat{\lambda}^N \right] \leq 0 \end{aligned} \tag{A.10}$$

Q.E.D.

Proof of Proposition 2d (Comparative Statics)

To save on notation, let $u^B = u(B,1)$; $u^C = u(B+C,1)$; $v^B = v(\omega, B)$ and $v^C = v(\omega, B+C)$. The notation for the income derivatives also follows this convention.

Differentiating the system of FOCs in (A.2) and (A.3) (and the budget constraint) w.r.t. p_I gives:

$$(B): \; -\frac{\partial\hat{\lambda}^N}{\partial p_I} + \left\{\theta[(1-p_I)u_{xx}^C + p_I u_{xx}^B] + (1-\theta)\int \left[p_{II}v_{MM}^C + (1-p_{II})v_{MM}^B\right] f(\omega)d\omega\right\}\frac{\partial\hat{B}^N}{\partial p_I}$$
$$+\left\{\theta(1-p_I)u_{xx}^C + (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\right\}\frac{\partial\hat{C}^N}{\partial p_I} = \theta\left[u_x^C - u_x^B\right]$$

$$(C): \; -[\theta(1-p_I)+(1-\theta)p_{II}]\frac{\partial\hat{\lambda}^N}{\partial p_I}$$
$$+\left\{\theta(1-p_I)u_{xx}^C + (1-\theta)p_{II}\int v_{MM}^C f(\omega)d\omega\right\}\left[\frac{\partial\hat{B}^N}{\partial p_I} + \frac{\partial\hat{C}^N}{\partial p_I}\right] = \theta\left[u_x^C - \hat{\lambda}^N\right]$$

$$(\lambda): \; \frac{\partial\hat{B}^N}{\partial p_I} + [\theta(1-p_I)+(1-\theta)p_{II}]\frac{\partial\hat{C}^N}{\partial p_I} = \theta\hat{C}^N$$

In matrix form this system of equations is:

$$\begin{bmatrix} 0 & -1 & -[\theta(1-p_I)+(1-\theta)p_{II}] \\ -1 & \begin{matrix}\theta[(1-p_I)u_{xx}^C + p_I u_{xx}^B]+ \\ (1-\theta)\int \left[p_{II}v_{MM}^C + (1-p_{II})v_{MM}^B\right] f(\omega)d\omega\end{matrix} & \begin{matrix}\theta(1-p_I)u_{xx}^C+ \\ (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\end{matrix} \\ -[\theta(1-p_I)+(1-\theta)p_{II}] & \begin{matrix}\theta(1-p_I)u_{xx}^C+ \\ (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\end{matrix} & \begin{matrix}\theta(1-p_I)u_{xx}^C+ \\ (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\end{matrix} \end{bmatrix}$$
$$\cdot \begin{bmatrix} \frac{\partial\hat{\lambda}^N}{\partial p_I} \\ \frac{\partial\hat{B}^N}{\partial p_I} \\ \frac{\partial\hat{C}^N}{\partial p_I} \end{bmatrix} = \begin{bmatrix} -\theta\hat{C}^N \\ \theta\left(u_x^C - u_x^B\right) \\ \theta\left(u_x^C - \hat{\lambda}^N\right) \end{bmatrix}$$

The determinant of the bordered Hessian is thus:

$$= -\left\{\theta(1-p_I)u_{xx}^C + (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\right\}[\theta p_I + (1-\theta)(1-p_{II})]$$
$$+\left\{\theta(1-p_I)u_{xx}^C + (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\right\}[\theta(1-p_I)+(1-\theta)p_{II}][1-\theta(1-p_I)-(1-\theta)p_{II}]$$
$$-\left\{\theta p_I u_{xx}^B + (1-\theta)\int (1-p_{II})v_{MM}^B f(\omega)d\omega\right\}[\theta(1-p_I)+(1-\theta)p_{II}]^2$$
$$= -\left\{\theta(1-p_I)u_{xx}^C + (1-\theta)\int p_{II}v_{MM}^C f(\omega)d\omega\right\}[\theta p_I + (1-\theta)(1-p_{II})][1-\theta(1-p_I)-(1-\theta)p_{II}]$$
$$-\left\{\theta p_I u_{xx}^B + (1-\theta)\int (1-p_{II})v_{MM}^B f(\omega)d\omega\right\}[\theta(1-p_I)+(1-\theta)p_{II}]^2$$

and thus

$$
\begin{aligned}
=&-\left\{\theta(1-p_I)u^C_{xx}+(1-\theta)\int p_{II}v^C_{MM}f(\omega)d\omega\right\}[\theta p_I+(1-\theta)(1-p_{II})]^2\\
&-\left\{\theta p_I u^B_{xx}+(1-\theta)\int(1-p_{II})v^B_{MM}f(\omega)d\omega\right\}[\theta(1-p_I)+(1-\theta)p_{II}]^2>0
\end{aligned}
$$

By Cramer's rule, the sign of $\partial\hat{B}^N/\partial p_I$ is thus given by (because the determinant of the bordered Hessian is positive):

$$
\det\begin{bmatrix}
0 & -\theta\hat{C}^N & -[\theta(1-p_I)+(1-\theta)p_{II}]\\
-1 & \theta(u^C_x-u^B_x) & \begin{array}{c}\theta(1-p_I)u^C_{xx}+\\(1-\theta)\int p_{II}v^C_{MM}f(\omega)d\omega\end{array}\\
-[\theta(1-p_I)+(1-\theta)p_{II}] & \theta(u^C_x-\hat{\lambda}^N) & \begin{array}{c}\theta(1-p_I)u^C_{xx}+\\(1-\theta)\int p_{II}v^C_{MM}f(\omega)d\omega\end{array}
\end{bmatrix}
$$

$$
\begin{aligned}
=&\theta\hat{C}^N\left\{-\left[\theta(1-p_I)u^C_{xx}+(1-\theta)\int p_{II}v^C_{MM}f(\omega)d\omega\right][\theta p_I+(1-\theta)(1-p_{II})]\right\}\\
&-[\theta(1-p_I)+(1-\theta)p_{II}]\left\{-\theta(u^C_x-\hat{\lambda}^N)+\theta(u^C_x-u^B_x)[\theta(1-p_I)+(1-\theta)p_{II}]\right\}>0
\end{aligned}
$$

where we have used the fact that $u^C_x>\hat{\lambda}^N$ when $p_I>0$ and $p_{II}>0$. It follows that $\partial\hat{B}^N/\partial p_I>0$.

Next, the sign of $\partial\hat{C}^N/\partial p_1$ is given by:

$$
\det\begin{bmatrix}
0 & -1 & -\theta\hat{C}^N\\
-1 & \begin{array}{c}\theta[(1-p_I)u^C_{xx}+p_I u^B_{xx}]+\\(1-\theta)\int\left[p_{II}v^C_{MM}+(1-p_{II})v^B_{MM}\right]f(\omega)d\omega\end{array} & \theta(u^C_x-u^B_x)\\
-[\theta(1-p_I)+(1-\theta)p_{II}] & \begin{array}{c}\theta(1-p_I)u^C_{xx}+\\(1-\theta)\int p_{II}v^C_{MM}f(\omega)d\omega\end{array} & \theta(u^C_x-\hat{\lambda}^N)
\end{bmatrix}
$$

$$
\begin{aligned}
=&\theta\left\{\underbrace{(\hat{\lambda}^N-u^C_x)}_{<0}+\underbrace{(u^C_x-u^B_x)[\theta(1-p_I)+(1-\theta)p_{II}]}_{<0}\right\}\\
&-\theta\hat{C}^N\left\{\begin{array}{l}-\left\langle\theta(1-p_I)u^C_{xx}+(1-\theta)p_{II}\int v^C_{MM}f(\omega)d\omega\right\rangle[\theta p_I+(1-\theta)(1-p_{II})]\\+\left\langle\theta p_I u^B_{xx}+(1-\theta)(1-p_{II})\int v^B_{MM}f(\omega)d\omega\right\rangle[\theta(1-p_I)+(1-\theta)p_{II}]\end{array}\right\}
\end{aligned}
$$

This can be written as:

$$
\begin{aligned}
=&\theta\left\{(\hat{\lambda}^N - u_x^C) + (u_x^C - u_x^B)[\theta(1-p_I) + (1-\theta)p_{II}]\right\} \\
&- \theta\hat{C}^N \left\{ \begin{array}{c} \theta^2 p_I(1-p_I)\left(u_{xx}^B - u_{xx}^C\right) \\ +\theta(1-\theta)p_I p_{II}\left(u_{xx}^B - \int v_{MM}^C f(\omega)d\omega\right) \\ +\theta(1-\theta)(1-p_I)(1-p_{II})\left(\int v_{MM}^B f(\omega)d\omega - u_{xx}^C\right) \\ +(1-\theta)^2 p_{II}(1-p_{II})\int\left(v_{MM}^B - v_{MM}^C\right)f(\omega)d\omega \end{array} \right\}
\end{aligned}
$$

A sufficient condition for $\partial\hat{C}^N/\partial p_I < 0$ will therefore be that the terms within curly braces in the second line be positive. To make further progress, suppose that $v_{MM} = -k\ \forall\ \omega \leq \bar{\omega}$; but $v_{MM} = 0\ \forall\omega > \bar{\omega}$ (where $k > 0$ is a constant)[42]. Then the terms within the second pair of braces become:

$$
k\left\{ \begin{array}{c} \theta^2 p_I(1-p_I)(1-1) \\ +\theta(1-\theta)p_I p_{II}\left[F(\bar{\omega}^C) - 1\right] \\ +\theta(1-\theta)(1-p_I)(1-p_{II})\left[1 - F(\bar{\omega}^B)\right] \\ +(1-\theta)^2 p_{II}(1-p_{II})\left[F(\bar{\omega}^C) - F(\bar{\omega}^B)\right] \end{array} \right\} = k\left\{ \begin{array}{c} \theta(1-\theta)p_I p_{II}\left[F(\bar{\omega}^C) - F(\bar{\omega}^B)\right] \\ +\theta(1-\theta)[1-p_I-p_{II}][1-F(\bar{\omega}^B)] \\ +(1-\theta)^2 p_{II}(1-p_{II})[F(\bar{\omega}^C) - F(\bar{\omega}^B)] \end{array} \right\} > 0
$$

and thus $\partial\hat{C}^N/\partial p_I < 0$

Proof of Proposition 3a

Solving the optimisation problem described in (2.37) yields the following FOCs:

$$
\begin{aligned}
(B):\quad & \theta\left[(1-p_I)u_x(\hat{B}^F + \hat{C}^F, 1) + p_I u_x(\hat{B}^F, 1)\right] \\
& + (1-\theta)\left\{\int v_M(\omega, \hat{B}^F)f(\omega)d\omega + p_{II}\int^{\bar{\bar{\omega}}}\left[u_x(\hat{B}^F + \hat{C}^F, 1) - v_M(\omega, \hat{B}^F)\right]f(\omega)d\omega\right\} \\
& \leq \hat{\lambda}^F\left[1 + (1-\theta)f(\bar{\bar{\omega}})\bar{\bar{\omega}}_B p_{II}\hat{C}^F\right] \quad ;\ \hat{B}^F \geq 0 \qquad \text{(A.11)}
\end{aligned}
$$

and:

$$
(C):\quad u_x(\hat{B}^F + \hat{C}^F, 1) \leq \hat{\lambda}^F\left\{1 + \frac{(1-\theta)f(\bar{\bar{\omega}})\bar{\bar{\omega}}_C p_{II}\hat{C}^F}{\theta(1-p_I) + (1-\theta)F(\bar{\bar{\omega}})p_{II}}\right\} \quad ;\ \hat{C}^F \geq 0 \qquad \text{(A.12)}
$$

[42]Details of preferences satisfying these (stringent) properties are given below.

where the pairs of inequalities hold with complementary slackness. Note that in writing these FOCs we have made use of the fact that $v(\bar{\bar{\omega}}, B) \equiv u(B + C, 1)$.

The FOCs have the following interpretations. The left side of (A.11) is the marginal welfare gain associated with a marginal increase in B. Notice that because all individuals in the population receive B, the marginal gain is the sum of each individual's smvi. Next, the right side captures in welfare units (i.e. multiplied by the shadow price of public expenditure, λ) the marginal cost associated with a marginal increase in B. Once more, because the government budget is exogenously fixed there are no tax revenue effects but simply an expenditure effect. The expenditure effect can be decomposed into two sub-effects: first, a marginal increase in B increases the amount of money spent on *all* individuals ; second, a marginal increase in B increases the number of individuals who would choose to apply for C and not work if awarded it.

The left side of (A.12) is the average marginal welfare gain associated with a marginal increase in C. On the right side is the marginal cost in welfare units of an increase in C, divided by the number of categorical recipients. The first term within curly braces captures the fact that a marginal increase in C results in a higher categorical payment to an existing recipient. The second term within curly braces captures the fact that an increase in C induces more individuals to apply for C.

We proceed to test the following two hypotheses: (i) $(\hat{B}^F > 0, \hat{C}^F = 0)$; and (ii) $(\hat{B}^F = 0,\, \hat{C}^F > 0)$.

(i) $\hat{B}^F = \beta\ ,\ \hat{C}^F = 0$ **(Pure universal system)**:

Suppose that $\hat{C}^F = 0$ - and thus $\hat{B}^F = \beta$ - such that $\bar{\bar{\omega}}(\beta, 0) = \bar{\omega}(\beta)$. It follows $\forall\ \omega \leq \bar{\bar{\omega}}(\beta, 0) = \bar{\omega}(\beta)\ :\ H^*(\omega, \beta) = 0 \rightarrow u(\beta, 1) = v(\omega, \beta) \rightarrow v_M(\omega, \beta) = u_x(\beta, 1)$. The FOCs (A.11) and (A.12) therefore become:

$$[\theta + (1 - \theta)F(\bar{\omega})]u_x(\beta, 1) + (1 - \theta) \int_{\bar{\omega}}^{\infty} v_M(\omega, \beta) f(\omega) d\omega = \hat{\lambda}^F,$$

$$u_x(\beta, 1) \leq \hat{\lambda}^F.$$

Combining these equations implies the contradictory statement:

$$\frac{1}{1 - F(\bar{\omega})} \int_{\bar{\omega}}^{\infty} v_M(\omega, \beta) f(\omega) d\omega \geq u_x(\beta, 1)$$

The average smvi for those with $\bar{\bar{\omega}} = \bar{\omega} < \omega$ cannot exceed that of the unable and so the assertion that $\hat{C}^F = 0$ must therefore be false. Instead, it must hold that $\hat{C}^F > 0 \; \forall \; p_I + p_{II} \leq 1$.

An immediate implication of this result is that the FOC for C in (A.12) must hold with equality and thus:

$$u_x(\hat{B}^F + \hat{C}^F, 1) = \hat{\lambda}^F \left\{ 1 + \frac{(1-\theta) f(\bar{\bar{\omega}}) \bar{\bar{\omega}}_C p_{II} \hat{C}^F}{[\theta(1-p_I) + (1-\theta)F(\bar{\bar{\omega}})p_{II}]} \right\} \geq \hat{\lambda}^F \tag{A.13}$$

From this we can directly ascertain that:

$$u_x(\hat{B}^F, 1) > \frac{1}{F(\bar{\bar{\omega}})} \int_0^{\bar{\bar{\omega}}} v_M(\omega, \hat{B}^F) f(\omega) d\omega > u_x(\hat{B}^F + \hat{C}^F, 1)$$
$$\geq \hat{\lambda}^F > \frac{1}{1 - F(\bar{\bar{\omega}})} \int_{\bar{\bar{\omega}}}^{\infty} v_M(\omega, \hat{B}^F) f(\omega) d\omega$$

where $u_x(\hat{B}^F + \hat{C}^F, 1) = \hat{\lambda}^F$ *only* if $p_{II} = 0$.

Prior to addressing the second hypothesis, it is useful to substitute (A.13) into (A.11) to obtain:

$$\begin{aligned} &\theta p_I u_x(\hat{B}^F, 1) + (1-\theta) \left\{ \int_0^{\infty} v_M(\omega, \hat{B}^F) f(\omega) d\omega - p_{II} \int_0^{\bar{\bar{\omega}}} v_M(\omega, \hat{B}^F) f(\omega) d\omega \right\} \\ &\leq u_x(\hat{B}^F + \hat{C}^F, 1) \chi \left\{ \frac{(1-\chi) + (1-\theta) f(\bar{\bar{\omega}}) p_{II} (\bar{\bar{\omega}}_B - \bar{\bar{\omega}}_C) \hat{C}^F}{\chi + (1-\theta) f(\bar{\bar{\omega}}) \bar{\bar{\omega}}_C p_{II} \hat{C}^F} \right\} \\ &\leq u_x(\hat{B}^F + \hat{C}^F, 1) \end{aligned} \tag{A.14}$$

where the final inequality arises because $\bar{\bar{\omega}}_B < \bar{\bar{\omega}}_C$. Recall that $\chi = [\theta(1-p_I) + (1-\theta)F(\bar{\bar{\omega}})p_{II}]$ denotes the number of individuals receiving the categorical transfer. Conversely, $(1-\chi) = [\theta p_I + (1-\theta)(1 - F(\bar{\bar{n}})p_{II}]$ denotes the number of individuals not receiving the categorical transfer.

We are now in a position to address the second hypothesis.

(ii) $\hat{B}^F = 0, \hat{C}^F = \beta / [\theta + (1-p_{II})F(\bar{\bar{\omega}})]$ **(Pure targeted system)**:

Suppose now that $\hat{B}^F = 0$. If $p_I > 0$ then some unable individuals have no income to consume and so the left side of (A.14) blows up because $\lim_{x\to} u_x(x,l) = +\infty$. The weak inequality cannot hold and so the assertion that $\hat{B}^F = 0$ must be false.

However, if $p_I = 0$ then we can write (A.14) with $B = 0$ as:

$$\frac{\int_0^\infty v_M(\omega,0)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,0)f(\omega)d\omega}{1-F(\bar{\bar{\omega}})p_{II}} \\ \leq u_x\left(C^F,1\right)\chi\cdot\left\{\frac{(1-\chi)+(1-\theta)f(\bar{\bar{\omega}})p_{II}(\bar{\bar{\omega}}_B-\bar{\bar{\omega}}_C)C^F}{\chi+(1-\theta)f(\bar{\bar{\omega}})\bar{\bar{\omega}}_C p_{II}C^F}\right\} \tag{A.15}$$

Proof of Proposition 3b.

To establish that $\delta^F > 0$ at the optimum whenever $p_I > 0$ and/or $p_{II} > 0$, we proceed to check three cases: (i) ($p_I > 0$, $p_{II} = 0$) ; (ii) ($p_I = 0$, $p_{II} > 0$) ; and (iii) ($p_I > 0$, $p_{II} > 0$).

(i) ($p_I > 0$ **,** $p_{II} = 0$)

In this case $\delta^F = [p_I u_x(B,1) + (1-p_I)u_x(B+C,1)] - \int_0^\infty v_M(\omega,B)f(\omega)d\omega$. Given no Type II errors are made the enforcement structure does not matter and the proof that $\delta^F > 0$ simply follows that under the No Enforcement case.

(ii) ($p_I = 0$ **,** $p_{II} > 0$)

In this case inequality in the average smvi between the unable and able subpopulations can be written as:

$$\begin{aligned}\delta^F =& u_x(B+C,1)[1-F(\bar{\bar{\omega}})p_{II}] - \left\{\int_0^\infty v_M(\omega,B)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,B)f(\omega)d\omega\right\} \\ =& [1-F(\bar{\bar{\omega}})p_{II}]\left\{u_x(B+C,1) - \frac{\int_0^\infty v_M(\omega,B)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,B)f(\omega)d\omega}{1-F(\bar{\bar{\omega}})p_{II}}\right\}\end{aligned} \tag{A.16}$$

The first term inside the curly braces is the smvi of categorical benefit recipients, whilst the second term is the average smvi of able individuals who do not receive the categorical benefit.

Substituting $p_I = 0$ into the equation defining optimal transfers in (A.14) we can immediately see that the term within curly braces will be positive and thus $\delta^F >$

0.

(iii) $(p_I > 0\ ,\ p_{II} > 0)$

In this case we can write δ^F as:

$$
\begin{aligned}
\delta^F =& u_x(B+C,1)[1-F(\bar{\bar{\omega}})p_{II}-p_I] \\
&-\left\{\int_0^\infty v_M(\omega,B)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,B)f(\omega)d\omega - p_I u_x(B,1)\right\} \\
=&[1-F(\bar{\bar{\omega}})p_{II}-p_I]\cdot \\
&\left\{u_x(B+C,1) - \frac{\int_0^\infty v_M(\omega,B)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,B)f(\omega)d\omega - p_I u_x(B,1)}{1-F(\bar{\bar{\omega}})p_{II}-p_I}\right\}
\end{aligned}
$$

From (A.14) we know that the smvi of categorical recipients (i.e. $u_x(B+C,1)$) will exceed the average smvi of non-categorical recipients. It follows that if the second term within curly braces is less than the average smvi of non-categorical recipients (i.e. the left side of (A.14)) then for sure $\delta^F > 0$. We check this below:

$$
\begin{aligned}
&\frac{\theta p_I u_x(\hat{B}^F,1) + (1-\theta)\left\{\int_0^\infty v_M(\omega,\hat{B}^F)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,\hat{B}^F)f(\omega)d\omega\right\}}{\theta p_I + (1-\theta)[1-F(\bar{\bar{\omega}})p_{II}]} \\
&-\left\{\frac{\int_0^\infty v_M(\omega,B)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,B)f(\omega)d\omega - p_I u_x(B,1)}{1-F(\bar{\bar{\omega}})p_{II}-p_I}\right\} \\
&=\frac{p_I[1-F(\bar{\bar{\omega}})p_{II}]\left\{u_x(\hat{B}^F,1) - \frac{\int_0^\infty v_M(\omega,\hat{B}^F)f(\omega)d\omega - p_{II}\int_0^{\bar{\bar{\omega}}} v_M(\omega,\hat{B}^F)f(\omega)d\omega}{1-F(\bar{\bar{\omega}})p_{II}}\right\}}{\{\theta p_I + (1-\theta)[1-F(\bar{\bar{\omega}})p_{II}]\}\{1-F(\bar{\bar{\omega}})p_{II}-p_I\}} > 0
\end{aligned}
\tag{A.17}
$$

because the smvi of unable individuals receiving only the universal benefit will always exceed the average smvi of able individuals also receiving only the universal benefit. It must therefore hold that $\delta^F > 0$ at the optimum.

In summary, whenever $p_I > 0$ and/or $p_{II} > 0$ it must hold that $\delta^F > 0$. Q.E.D.

Proof of Proposition 3c.

To establish the sign of $\partial V^F/\partial p_I$ we use (A.9) (i.e. $u(B,1)-u(B+C,1) < u_x(B+C,1)C$) to obtain:

$$\frac{\partial V^F}{\partial p_I} = \theta\langle[u(\hat{B}^F,1) - u(\hat{B}^F + \hat{C}^F,1)] + \hat{\lambda}^F\hat{C}^F\rangle < \theta\hat{C}^F[\hat{\lambda}^F - u_x(\hat{B}^F + \hat{C}^F,1)] < 0$$

Turning to the affect of Type II errors on maximum welfare, we have:

$$\begin{aligned}
&\frac{\partial V^F}{\partial p_{II}} \\
&= (1-\theta)F(\bar{\bar{\omega}})\left\{\left[u(\hat{B}^F + \hat{C}^F,1) - \frac{1}{F(\bar{\bar{\omega}})}\int_0^{\bar{\bar{\omega}}} v(\omega,\hat{B}^F)f(\omega)d\omega\right] - \hat{\lambda}^F\hat{C}^F\right\} \\
&= (1-\theta)F(\bar{\bar{\omega}})\left\{\begin{array}{c}\left[u(\hat{B}^F + \hat{C}^F,1) - u(\hat{B}^F,1) - \hat{\lambda}^F\hat{C}^F\right] \\ -\frac{1}{F(\bar{\bar{\omega}})}\int_0^{\bar{\bar{\omega}}}\left[v(\omega,\hat{B}^F) - u(\hat{B}^F,1)\right]f(\omega)d\omega\end{array}\right\} \\
&= (1-\theta)F(\bar{\bar{\omega}})\left\{\begin{array}{c}\left[u(\hat{B}^F + \hat{C}^F,1) - u(\hat{B}^F,1) - \hat{\lambda}^F\hat{C}^F\right] \\ -\frac{1}{F(\bar{\bar{\omega}})}\int_{\bar{\omega}(B)}^{\bar{\bar{\omega}}}\left[v(\omega,\hat{B}^F) - u(\hat{B}^F,1)\right]f(\omega)d\omega\end{array}\right\}
\end{aligned}$$

To transition from the first line to the second line we add and subtract $u(\hat{B}^F,1)$. In the second line there are two effects in square braces: the first is the benefit to an individual - net of exchequer costs - of receiving the categorical benefit conditional on facing a zero quantity constraint on labour supply; the second measures the cost of facing a zero quantity constraint on labour conditional on only receiving the universal benefit. The transition from the second to third line is made through recognising that $v_M(\omega,B) = u_x(B,1)\ \forall\ \omega \le \bar{\omega}(B)$.

But since $v(\omega, B) < v(\bar{\bar{\omega}}, B) = u(B + C, 1) \;\; \forall \, \bar{\omega}(B) \leq \omega < \bar{\bar{\omega}}$ it must hold that:

$$\begin{aligned}
\frac{\partial V^F}{\partial p_{II}} &> (1-\theta)F(\bar{\bar{\omega}}) \left\{ \left\langle 1 - \frac{F(\bar{\bar{\omega}}) - F(\bar{\omega}(\hat{B}^F))}{F(\bar{\bar{\omega}})} \right\rangle \left[u(\hat{B}^F + \hat{C}^F, 1) - u(\hat{B}^F, 1) \right] - \hat{\lambda}^F \hat{C}^F \right\} \\
&= (1-\theta)F(\bar{\bar{\omega}}) \left\{ \frac{F[\bar{\omega}(\hat{B}^F)]}{F(\bar{\bar{\omega}})} \left[u(\hat{B}^F + \hat{C}^F, 1) - u(\hat{B}^F, 1) \right] - \hat{\lambda}^F \hat{C}^F \right\} \\
&> (1-\theta)F(\bar{\bar{\omega}})\hat{C}^F \left\{ \frac{F[\bar{\omega}(\hat{B}^F)]}{F(\bar{\bar{\omega}})} u_x(\hat{B}^F + \hat{C}^F, 1) - \hat{\lambda}^F \right\} \\
&= (1-\theta)F(\bar{\bar{\omega}})\hat{C}^F \left\{ \frac{F[\bar{\omega}(\hat{B}^F)]}{F(\bar{\bar{\omega}})} \hat{\lambda}^F \left\langle 1 + \frac{(1-\theta)f(\bar{\bar{\omega}})\bar{\bar{\omega}}_C p_{II} \hat{C}^F}{\theta(1-p_I) + (1-\theta)F(\bar{\bar{\omega}})p_{II}} \right\rangle - \hat{\lambda}^F \right\} \\
&= (1-\theta)F(\bar{\bar{\omega}})\hat{C}^F \left\{ \begin{array}{c} \frac{F[\bar{\omega}(\hat{B}^F)]}{F(\bar{\bar{\omega}})} \hat{\lambda}^F \left\langle \frac{(1-\theta)f(\bar{\bar{\omega}})\bar{\bar{\omega}}_C p_{II} \hat{C}^F}{\theta(1-p_I) + (1-\theta)F(\bar{\bar{\omega}})p_{II}} \right\rangle \\ -\hat{\lambda}^F \left[1 - \frac{F[\bar{\omega}(\hat{B}^F)]}{F(\bar{\bar{\omega}})} \right] \end{array} \right\}
\end{aligned}$$

The transition from the first line to the second follows from simple manipulation of fractions, whilst to transition from the second to third line we use the property that $u(B + C, 1) - u(B, 1) > u_x(B + C, 1)C$. Next, the transition from the third through to fifth line uses the definition of $u_x(\hat{B}^F + \hat{C}^F, 1)$ from (A.13). Rewriting the terms in curly braces in the final line in terms of elasticities then gives:

$$\frac{\partial V^F}{\partial p_{II}} > (1-\theta)F(\bar{\bar{\omega}})\hat{C}^F E \tag{A.18}$$

where:

$$E \equiv \frac{F[\bar{\omega}(\hat{B}^F)]}{\frac{\theta(1-p_I)}{(1-\theta)p_{II}} + F(\bar{\bar{\omega}})} \cdot \frac{\bar{\bar{\omega}} f(\bar{\bar{\omega}})}{F(\bar{\bar{\omega}})} \cdot \frac{\hat{C}^F \bar{\bar{\omega}}_C}{\bar{\bar{\omega}}} - \left[1 - \frac{F[\bar{\omega}(\hat{B}^F)]}{F(\bar{\bar{\omega}})} \right] \tag{A.19}$$

The sufficient condition is therefore: $E > 0 \Rightarrow \partial V^F / \partial p_{II} > 0$. If $(1-\theta)p_{II} \approx 0$ then the first term on the right side is approximately zero such that $E < 0$. Yet, given that $\partial V^F / \partial p_{II} > (1-\theta)F(\bar{\bar{\omega}})\hat{C}^F E$ this is insufficient to sign $\partial V^F / \partial p_{II}$. However, if $\theta(1 - p_I) \approx 0$ then:

$$E \approx \frac{F(\bar{\omega}(B))}{F(\bar{\bar{\omega}})} \cdot \frac{\bar{\bar{\omega}} f(\bar{\bar{\omega}})}{F(\bar{\bar{\omega}})} \cdot \frac{\hat{C}^F \bar{\bar{\omega}}_C}{\bar{\bar{\omega}}} - \left[1 - \frac{F(\bar{\omega}(B))}{F(\bar{\bar{\omega}})} \right] \tag{A.20}$$

If the product of (i) the elasticity of F with respect to ω - evaluated at $\bar{\bar{\omega}}$ - and (ii) the elasticity of $\bar{\bar{\omega}}$ with respect to C - evaluated at $\hat{C}^F$ - are sufficiently high, then we can have $E > 0$ and thus $\partial V^F/\partial p_{II} > 0$. Since these elasticities depend on as yet unspecified properties of the distribution and utility functions respectively, we have enough degrees of freedom to choose parameters so that $E > 0$.

Appendix B Numerical Simulations

This section (i) derives some key properties of the CES utility function in (2.48) in the main text; and (ii) presents the numerical code generating the simulation results in the main text.

The CES Utility Function

Given preferences $u(x,l) = \left[\alpha x^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha) l^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right]^{\frac{\mathcal{E}}{\mathcal{E}-1}}$ (where $\mathcal{E} \neq 1$) and a budget constraint $x = \omega H + M$, an individual's optimal choice of $H \in (0,1]$ satisfies [43]

$$H^* = \operatorname{Arg} \max_{H \in (0,1]} \left[\alpha(\omega H + M)^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)(1-H)^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right]^{\frac{\mathcal{E}}{\mathcal{E}-1}}$$

[43] It is straightforward to show that the elasticity of substitution of the CES utility function is given by $\mathcal{E}$. Writing the individual budget constraint as $x + \omega l = \omega + M$ and noting that at any interior optimum $\omega = u_l/u_x$, we can define the elasticity of substitution between l and x by (where x^* and l^* denote optimal choices):

$$\mathcal{E} = \frac{d(x^*/l^*)/(x^*/l^*)}{d(p_l/p_x)/(p_l/p_x)} = \frac{p_l/p_x}{x^*/l^*} \cdot \frac{d(x^*/l^*)}{d(p_l/p_x)} = \frac{\omega}{x^*/l^*} \cdot \frac{d(x^*/l^*)}{d(p_l/p_x)} = \frac{u_l/u_x}{x^*/l^*} \cdot \frac{d(x^*/l^*)}{d(u_l/u_x)}$$

From the CES preferences above we have:

$$\frac{u_l}{u_x} = \frac{(1-\alpha)l^{-\frac{1}{\mathcal{E}}}}{\alpha x^{-\frac{1}{\mathcal{E}}}} = \frac{(1-\alpha)}{\alpha} \cdot \left(\frac{x}{l}\right)^{\frac{1}{\mathcal{E}}} \Rightarrow \frac{x}{l} = \left(\frac{\alpha}{1-\alpha}\right)^{\mathcal{E}} \left(\frac{u_l}{u_x}\right)^{\mathcal{E}}$$

Substituting this into the definition of $\mathcal{E}$ gives:

$$\mathcal{E} = \frac{\mathcal{E}\left(\frac{\alpha}{1-\alpha}\right)^{\mathcal{E}} \left(\frac{u_l}{u_x}\right)^{\mathcal{E}-1}}{\left(\frac{\alpha}{1-\alpha}\right)^{\mathcal{E}} \left(\frac{u_l}{u_x}\right)^{\mathcal{E}-1}} = \mathcal{E}$$

Thus yielding the FOC:

$$\alpha\omega\left(\frac{1}{\omega H^{*}+M}\right)^{\frac{1}{\mathcal{E}}}-(1-\alpha)\cdot\left(\frac{1}{1-H^{*}}\right)^{\frac{1}{\mathcal{E}}}\leq 0\ ;\ H^{*}\geq 0$$
$$\Rightarrow \alpha\omega-(1-\alpha)\left(\frac{\omega H^{*}+M}{1-H^{*}}\right)^{\frac{1}{\mathcal{E}}}\leq 0\ ;\ H^{*}\geq 0 \tag{B.1}$$

where the pair of inequalities hold with complementary slackness.

Setting $H^{*}=0$ it is straightforward to verify that the reservation wage $\bar{\omega}(M)$ at or below which an individual chooses not to work is given by:

$$\bar{\omega}(M)=\left(\frac{1-\alpha}{\alpha}\right)M^{\frac{1}{\mathcal{E}}} \tag{B.2}$$

It is useful to note that:

$$\frac{\partial\bar{\omega}}{\partial\mathcal{E}}=-\frac{(1-\alpha)M^{\frac{1}{\mathcal{E}}}\log(M)}{\alpha\mathcal{E}^{2}}\begin{Bmatrix}>\\=\\<\end{Bmatrix}0\text{ if }\ M\begin{Bmatrix}>\\=\\<\end{Bmatrix}1$$

Next, for all $\omega>\bar{\omega}(M)$ it must hold that $H^{*}>0$ such that the FOC in (B.1) holds with equality (i.e. $\alpha^{\mathcal{E}}\omega^{\mathcal{E}}(1-H^{*})-(1-\alpha)^{\mathcal{E}}(\omega H^{*}+M)=0$). Optimal labour supply in this case is therefore:

$$H^{*}=\frac{\alpha^{\mathcal{E}}\omega^{\mathcal{E}}-(1-\alpha)^{\mathcal{E}}M}{\alpha^{\mathcal{E}}\omega^{\mathcal{E}}+(1-\alpha)^{\mathcal{E}}\omega}=\frac{\alpha^{\mathcal{E}}\omega^{\mathcal{E}}\left[1-\left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}M\omega^{-\mathcal{E}}\right]}{\alpha^{\mathcal{E}}\omega^{\mathcal{E}}+(1-\alpha)^{\mathcal{E}}\omega}=\frac{1-\left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}M\omega^{-\mathcal{E}}}{1+\left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\omega^{1-\mathcal{E}}} \tag{B.3}$$

which is analogous to that in Stern (1976). Notice that if $M=0$ then labour supply will for sure be falling in the net wage rate.[44]

[44] Note that if $\mathcal{E}=1$ preferences are Cobb-Douglas and given by $u(x,l)=x^{\alpha}l^{1-\alpha}$. The FOC for optimal labour supply is:

$$\alpha\omega-(1-\alpha)\left(\frac{\omega H^{*}+M}{1-H^{*}}\right)\leq 0\ ;\ \ H^{*}\geq 0$$

Setting H^{*} gives rise to the reservation wage $\bar{\omega}(M)=(\frac{1-\alpha}{\alpha})M$ (i.e. as in (B.2)). Similarly, for $\omega>\bar{\omega}$

Indirect Utility Function

If $\omega \leq \bar{\omega}(M)$ then $H^* = 0$ and so indirect utility is simply $v(\omega, M) = u(M, 1) = \left[\alpha M^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]^{\frac{\mathcal{E}}{\mathcal{E}-1}}$; with partial derivatives:

$$\begin{aligned} v_M &= \alpha\left[\alpha + (1-\alpha)M^{\frac{1-\mathcal{E}}{\mathcal{E}}}\right]^{\frac{1}{\mathcal{E}-1}} > 0 \\ v_{MM} &= -\left(\frac{1}{\mathcal{E}}\right)\alpha(1-\alpha)M^{\frac{1-2\mathcal{E}}{\mathcal{E}}}\left[\alpha + (1-\alpha)M^{\frac{1-\mathcal{E}}{\mathcal{E}}}\right]^{\frac{2-\mathcal{E}}{\mathcal{E}-1}} < 0 \end{aligned}$$

and

$$v_{MMM} = -\left(\frac{1}{\mathcal{E}^2}\right)\alpha(1-\alpha)\left\{\begin{array}{c} (1-2\mathcal{E})M^{\frac{1-3\mathcal{E}}{\mathcal{E}}}[\alpha + (1-\alpha)M^{\frac{1-\mathcal{E}}{\mathcal{E}}}]^{\frac{2-\mathcal{E}}{\mathcal{E}-1}} \\ -(2-\mathcal{E})(1-\alpha)M^{\frac{2-4\mathcal{E}}{\mathcal{E}}}[\alpha + (1-\alpha)M^{\frac{1-\mathcal{E}}{\mathcal{E}}}]^{\frac{3-2\mathcal{E}}{\mathcal{E}-1}} \end{array}\right\}$$

Note that it will certainly hold that $v_{MMM} > 0$ whenever $\mathcal{E} \in [1/2, 1)$.

Meanwhile, for $\omega > \bar{\omega}(M)$ we substitute (B.3) into individual preferences to obtain:

$$\begin{aligned} v &= \left\{\alpha\left[\frac{\omega - \left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}} M\omega^{1-\mathcal{E}}}{1 + \left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\omega^{1-\mathcal{E}}} + M\right]^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\left\langle 1 - \left[\frac{1 - \left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}} M\omega^{-\mathcal{E}}}{1 + \left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\omega^{1-\mathcal{E}}}\right]\right\rangle^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right\}^{\frac{\mathcal{E}}{\mathcal{E}-1}} \\ &= \left\{\alpha\left[\frac{\omega + M}{1 + \left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\omega^{1-\mathcal{E}}}\right]^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\left[\frac{\left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\omega^{-\mathcal{E}}(\omega + M)}{1 + \left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\omega^{1-\mathcal{E}}}\right]^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right\}^{\frac{\mathcal{E}}{\mathcal{E}-1}} \\ &= (\omega + M)\left\{\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\omega^{1-\mathcal{E}}\right\}^{\frac{1}{\mathcal{E}-1}} \end{aligned} \tag{B.4}$$

where $v_M = [\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\omega^{1-\mathcal{E}}]^{\frac{1}{\mathcal{E}-1}}$. As v_M is not a function of M we necessarily have $v_{MM} = v_{MMM} = 0$.[45]

we obtain:

$$H^* = \frac{\alpha\omega - (1-\alpha)M}{\omega} = \alpha - (1-\alpha)M/\omega$$

[45]Under Cobb Douglas preferences ($\mathcal{E} = 1$) indirect utility when $\omega \leq \bar{\omega}(M)$ is simply $v = M^{\alpha}$; with partial derivatives $v_M = \alpha M^{-(1-\alpha)} > 0$ and $v_{MM} = \alpha(\alpha - 1)M^{\alpha-2} < 0$. Next, for $\omega > \bar{\omega}(M)$ we

Derivation of $\bar{\bar{\omega}}_B$ and $\bar{\bar{\omega}}_C$

For the CES utility function the condition defining the critical net wage $\bar{\bar{\omega}}$ is:

$$[\bar{\bar{\omega}} + B]\left\{\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\right\}^{\frac{1}{\mathcal{E}-1}} \equiv \left[\alpha(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]^{\frac{\mathcal{E}}{\mathcal{E}-1}}$$

Differentiating this identity with respect to B gives:

$$\begin{aligned}
&\frac{\partial\bar{\bar{\omega}}}{\partial B}\left[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\right]^{\frac{1}{\mathcal{E}-1}} \cdot \left\{1 - \frac{(1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{-\mathcal{E}}(\bar{\bar{\omega}}+B)}{[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}]}\right\} \\
&= \alpha(B+C)^{-\frac{1}{\mathcal{E}}}\left[\alpha(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]^{\frac{2-\mathcal{E}}{\mathcal{E}-1}} - \left[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\right]^{\frac{1}{\mathcal{E}-1}}
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow\frac{\partial\bar{\bar{\omega}}}{\partial B}\left[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\right]^{\frac{2-\mathcal{E}}{\mathcal{E}-1}}\left\{\alpha^{\mathcal{E}} - (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{-\mathcal{E}}B\right\} \\
&= \alpha\left\{(B+C)^{\frac{1-\mathcal{E}}{\mathcal{E}}}\left[\alpha(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]\right\}^{\frac{1}{\mathcal{E}-1}} - \left[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\right]^{\frac{1}{\mathcal{E}-1}}
\end{aligned}$$

From which it is straightforward to establish that:

$$\frac{\partial\bar{\bar{\omega}}}{\partial B} = \frac{\alpha\left[\alpha + (1-\alpha)(B+C)^{\frac{1-\mathcal{E}}{\mathcal{E}}}\right]^{\frac{1}{\mathcal{E}-1}} - \left[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\right]^{\frac{1}{\mathcal{E}-1}}}{[\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}]^{\frac{2-\mathcal{E}}{\mathcal{E}-1}} \cdot [\alpha^{\mathcal{E}} - (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{-\mathcal{E}}B]} \tag{B.6}$$

Similarly, we can readily establish that:

$$\frac{\partial\bar{\bar{\omega}}}{\partial C} = \frac{\alpha\left\{\alpha + (1-\alpha)(B+C)^{\frac{1-\mathcal{E}}{\mathcal{E}}}\right\}^{\frac{1}{\mathcal{E}-1}}}{\{\alpha^{\mathcal{E}} + (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{1-\mathcal{E}}\}^{\frac{2-\mathcal{E}}{\mathcal{E}-1}}\,[\alpha^{\mathcal{E}} - (1-\alpha)^{\mathcal{E}}\bar{\bar{\omega}}^{-\mathcal{E}}B]} \tag{B.7}$$

Upper and Lower Bounds of $\bar{\bar{\omega}}(B, C)$

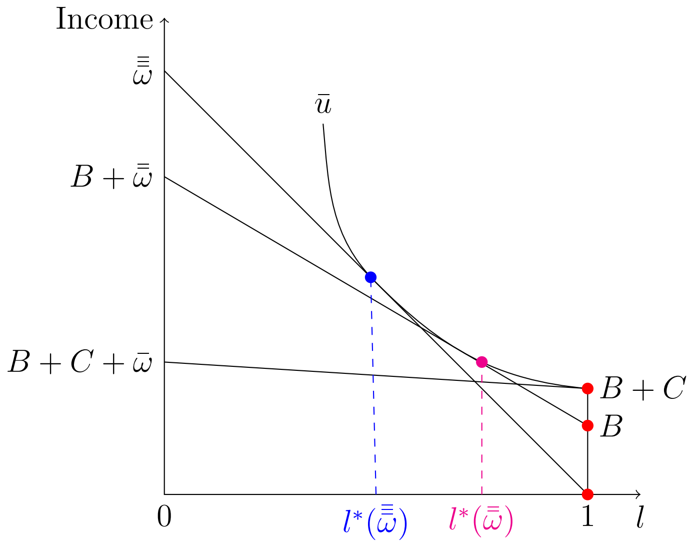
For the purpose of numerically simulating $\bar{\bar{\omega}}$ we want to put lower and upper bounds on the search region. Given that $v[\bar{\omega}(B+C), B+C] = u(B+C, 1) = v(\bar{\bar{\omega}}, B)$ and thus

substitute (B.3) into $u = (\omega H + M)^{\alpha}(1-H)^{1-\alpha}$ to obtain:

$$v = \left\{\frac{\omega + M}{1 + \left(\frac{1-\alpha}{\alpha}\right)}\right\}^{\alpha}\left\{\frac{\left(\frac{1-\alpha}{\alpha}\right)(1 + M/\omega)}{1 + \left(\frac{1-\alpha}{\alpha}\right)}\right\}^{1-\alpha} = \frac{\left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha}}{1 + \left(\frac{1-\alpha}{\alpha}\right)} \cdot \frac{\omega + M}{\omega^{1-\alpha}} = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\left(\frac{\omega + M}{\omega^{1-\alpha}}\right) \tag{B.5}$$

with partial derivative $v_M = \alpha^{\alpha}(1-\alpha)^{1-\alpha}/\omega^{1-\alpha}$.

Figure 2.34: Upper and Lower Bounds of $\bar{\omega}$



Notes: At the critical wage $\bar{\bar{\bar{\omega}}}$ an individual's optimal labour/leisure choice is such that they are indifferent between: (i) working whilst receiving no unearned income; and (ii) not working and receiving $B+C$ in unearned income.

$\bar{\omega}(B+C) < \bar{\bar{\omega}}$, the function $\bar{\omega}(B+C)$ serves as a suitable lower bound for $\bar{\bar{\omega}}$. Under the CES preferences in (2.48) we already have an explicit expression for $\bar{\omega}$, as given by (B.2). To establish an expression for an upper bound, let the function $\bar{\bar{\bar{\omega}}}(B,C)$ satisfy $v(\bar{\bar{\bar{\omega}}},0) \equiv u(B+C,1)$, where clearly $\bar{\bar{\bar{\omega}}}_B > 0$ and $\bar{\bar{\bar{\omega}}}_C > 0$. Given that $v(\bar{\bar{\bar{\omega}}},0) = v(\bar{\bar{\omega}},B)$ it must hold that $\bar{\bar{\bar{\omega}}} > \bar{\bar{\omega}}$ whenever $B > 0$, whilst $\bar{\bar{\bar{\omega}}} = \bar{\bar{\omega}}$ if $B = 0$. The function $\bar{\bar{\bar{\omega}}}$ is therefore an *upper bound* for $\bar{\bar{\omega}}$. Figure 2.34 graphically illustrates the upper ($\bar{\bar{\bar{\omega}}}$) and lower ($\bar{\omega}$) bounds, respectively.

For the CES preferences in (2.48) we can derive an explicit expression for $\bar{\bar{\bar{\omega}}}$. Formally, $\bar{\bar{\bar{\omega}}}$ satisfies:

$$\bar{\bar{\bar{\omega}}}\left[\alpha^{\mathcal{E}} + (1-\alpha)\bar{\bar{\bar{\omega}}}^{1-\mathcal{E}}\right]^{\frac{1}{\mathcal{E}-1}} = \left[\alpha(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha)\right]^{\frac{\mathcal{E}}{\mathcal{E}-1}}$$

which can be written more conveniently as:

$$
\begin{aligned}
\left\{\bar{\bar{\bar{\omega}}}^{\mathcal{E}-1}\left[\alpha^{\mathcal{E}-1}+(1-\alpha)^{\mathcal{E}}\bar{\bar{\bar{\omega}}}^{1-\mathcal{E}}\right]\right\}^{\frac{1}{\mathcal{E}-1}} &= \left\{\alpha(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}}+(1-\alpha)\right\}^{\frac{\mathcal{E}}{\mathcal{E}-1}} \\
\Rightarrow \alpha^{\mathcal{E}}\bar{\bar{\bar{\omega}}}^{\mathcal{E}-1}+(1-\alpha)^{\mathcal{E}} &= \left\{\alpha(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}}+(1-\alpha)\right\}^{\mathcal{E}} \\
\Rightarrow \bar{\bar{\bar{\omega}}}^{\mathcal{E}-1} &= \left\{\frac{\alpha M^{\frac{\mathcal{E}-1}{\mathcal{E}}}+(1-\alpha)}{\alpha}\right\}^{\mathcal{E}}-\left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}
\end{aligned}
$$

and thus, finally:

$$
\bar{\bar{\bar{\omega}}} = \left\{\left[(B+C)^{\frac{\mathcal{E}-1}{\mathcal{E}}}+\left(\frac{1-\alpha}{\alpha}\right)\right]^{\mathcal{E}}-\left(\frac{1-\alpha}{\alpha}\right)^{\mathcal{E}}\right\}^{\frac{1}{\mathcal{E}-1}}
$$

Numerical Code

```
from __future__ import division
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize
from numpy import vectorize
from scipy.stats import lognorm
from scipy.stats import expon
from scipy.integrate import quad
from scipy.optimize import brentq
from scipy.optimize import newton

#Python Code run using Enthought Canopy 1.5.1 (64bit) on MacBook Pro 2.3GHz Intel Core i7

#FIXED BUDGET ANALYSIS CODE
#---------------------------------------------------------

#1) PRELIMINARIES
#.........................................................

#In what follows: (i) 'n' is individual productivity; (ii) 't' is the tax rate; (iii) m is unearned income; (iv) 'e' is the elasticity of substitution between leisure and consumption; (v) 'a' (alpha) is the consumption factor share in CES preferences; (vi) 'theta' is the unable subpopulation size; and (vii) 'beta' is the fixed benefit budget size.

#In addition, we will throughout let 'b' denote the universal benefit, whilst 'c' denotes the categorical benefit.

#The reservation productivity is:
def nbar(t,m,a,e):
    if m<=0:
        return 0
    else:
        return ((1-a)/a)*(m**(1/e))/(1-t)

#Individual labour supply is defined by:
def h(n,t,m,a,e):
    if n <= nbar(t,m,a,e):
        return 0
    else:
        return ((1-(((1-a)/a)**e)*m*((n*(1-t))**(-e)))/
        (1+(((1-a)/a)**e)*((n*(1-t))**(1-e))))

#The partial derivative of h with respect to m is:
def hm(n,t,m,a,e):
    if n <= nbar(t,m,a,e):
        return 0
    else:
        return -((((1-a)/a)**e)*((n*(1-t))**(-e))/
        (1+(((1-a)/a)**e)*((n*(1-t))**(1-e))))
```

```

#Individual earnings are therefore:
def y(n,t,m,a,e):
    return n*h(n,t,m,a,e)

#Multiplying y by the distribution pdf gives:
def ypdf(n,t,m,a,e,mu):
    return n*h(n,t,m,a,e)*expon.pdf(n,scale=1/mu)

#Average earnings in the economy are:
def ybar(t,b,a,e,mu):
    return quad(ypdf,0,3,args=(t,b,a,e,mu))[0]

#The indirect utility function is thus:
def v(n,t,m,a,e):
    if n <= nbar(t,m,a,e):
        if m<=0:
            return 0
        else:
            return (a*(m**((e-1)/e))+(1-a))**(e/(e-1))
    else:
        return ((n*(1-t)+m)*(a**e+((1-a)**e)*
        (n*(1-t))**(1-e))**(1/(e-1)))

#Individual utility for the unable is simply
    v(0,0,m,a,e). To check this simply define:
def u(m,a,e):
    return (a*(m**((e-1)/e))+(1-a))**(e/(e-1))

# and then
#print u(0.3,0.6,0.5)
#0.416666666667
#print v(0,0,0.3,0.6,0.5)
#0.416666666667

#To construct the average welfare over the able
    subpopulation who receive unearned income m, we
    first define a function (vpdf) which is v
    multiplied by the productivity pdf. We then define
    another function (wa) which integrates vpdf over n.
def vpdf(n,t,m,a,e,mu):
    return v(n,t,m,a,e)*expon.pdf(n,scale=1/mu)

def wa(t,m,a,e,mu):
    return quad(vpdf,0,3,args=(t,m,a,e,mu))[0]

#NB.In the simulations mu=3 and so at an upper
    integration limit of n=3 we can see that:
#print expon.cdf(3,scale=1/3)
#0.999876590196

#1.1) SOCIAL MARGINAL VALUE OF INCOME(smvi)
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

#The partial derivative of v with respect to m is:
def vm(n,t,m,a,e):
    if n <= nbar(t,m,a,e):
        return a*((a+(1-a)*(m**((1-e)/e)))**(1/(e-1)))
```

```
    else:
        return ((a**e+((1-a)**e)*(n*(1-t))**(1-e))
        **(1/(e-1)))

#Note:For cases where n<nbar writing vm=ux as above -
    as opposed to
    a*m**(-1/e)(a*m**((e-1)/e)+(1-a))**(1/(e-1)) - has
    the convenient property that m is not raised to a
    negative power, thus avoiding issues of raising
    zero to a negative power (in the limiting case
    where m=x=0). Indeed, we can check that:

#print vm(0,0,0,0.6,0.5)
#1.66666666667
#print vm(0,0,0,0.6,0.6)
#2.15165741456
#print vm(0,0,0,0.6,0.7)
#3.29341972638
#print vm(0,0,0,0.6,0.8)
#7.71604938272
#print vm(0,0,0,0.6,0.9)
#99.2290301275
#print vm(0,0,0,0.6,0.99)
#9.18388024492e+21

#Multiplying vm by the distribution pdf gives:
def vmpdf(n,t,m,a,e,mu):
    return vm(n,t,m,a,e)*expon.pdf(n,scale=1/mu)

#Integrating over n then gives the average social
    marginal value of income (smvi) of the able
    subpopulation (with unearned income m).
def wam(t,m,a,e,mu):
    return quad(vmpdf,0,3,args=(t,m,a,e,mu))[0]

#The critical value of beta that eliminates
    between-group inequality in the average smvi
    between the unable and able subpopulations is given
    as follows. We first define the function capturing
    between-group inequality (inequalityins) and then
    define a function which finds the root of
    'inequalityins':
def inequalityins(beta,t,a,e,mu,theta):
    return (vm(0,0,beta/theta,a,e)-wam(t,0,a,e,mu))

def betastar(t,a,e,mu,theta):
    return brentq(inequalityins,0.001,1,
    args=(t,a,e,mu,theta))

#2) PERFECT DISCRIMINATION OPTIMISATION PROBLEM
#.................................................

#Let the vector x=(x[0],x[1])=(b,c) denote the
    variables we are optimising over. The optimisation
    problem is now described by:
def resultspd(t,a,e,mu,theta,beta):
    conspd=({'type':'eq',
'fun':lambda x:np.array([x[0]+theta*x[1]-beta])},
```

```
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])})
    def objpd(x,t,a,e,mu,theta):
        return (-theta*v(0,0,x[0]+x[1],a,e)-
        (1-theta)*wa(t,x[0],a,e,mu))
    res=minimize(objpd,[0.1,0.2],args=(t,a,e,mu,theta),
    constraints=conspd,method='SLSQP',
    options={'ftol':1e-10,'disp':False})
    return ('b=',res.x[0],'c=',res.x[1],
    'V^P=',objpd(res.x,t,a,e,mu,theta),
    'sbar=',wam(t,res.x[0],a,e,mu),
    'lambda=',vm(0,0,res.x[0]+res.x[1],a,e))

#Notice that because we are using the minimisation
    algorithm 'minimize' we necessarily multiply the
    objective function by -1.The dictionary of
    constraints above specify the following. The first
    constraint is an equality constraint and states
    that expenditure must equal the fixed budget size.
    The second and third constraints are non-negativty
    inequality constraints on b(x[0]) and c(x[1]),
    respectively.

#Note that the Pure Universal welfare level is:
def wpu(t,a,e,mu,theta,beta):
    return
        theta*v(0,t,beta,a,e)+(1-theta)*wa(t,beta,a,e,mu)

#The Pure Universal welfare level as a proportion of
    the Perfect Discrimination welfare level is thus:
def wuvp(t,a,e,mu,theta,beta):
    return (wpu(t,a,e,mu,theta,beta)/
    (-resultspd(t,a,e,mu,theta,beta)[5]))

#3)IMPERFECT DISCRIMINATION
#.....................................................

#We now turn to the case where the categorical benefit
    may be administered with Type I and Type II
    classification errors. Further, the condition that
    a recipient of the categorical benefit does not
    work may be either not enforced or fully enforced.

#3)(i) NO ENFORCEMENT
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

#Let us define the average smvi of categorical
    recipients and non-recipients by the following two
    functions:
def avsmvir(t,b,c,a,e,mu,theta,p1,p2):
    return ((theta*(1-p1)*vm(0,0,b+c,a,e)+(1-theta)*
    p2*wam(t,b+c,a,e,mu))/
    (theta*(1-p1)+(1-theta)*p2))

def avsmvnr(t,b,a,e,mu,theta,p1,p2):
    return ((theta*p1*vm(0,0,b,a,e)+(1-theta)*(1-p2)*
```

```
    wam(t,b,a,e,mu))/(theta*p1+(1-theta)*(1-p2)))

#Once more, let the vector x=(x[0],x[1])=(b,c) denote
    the variables we are optimising over.

#Benefit expenditure (left side of the budget
    constraint) is now defined by the following
    function:
def expenditurene(x,t,theta,p1,p2):
    return x[0]+(theta*(1-p1)+(1-theta)*p2)*x[1]

#The optimisation problem is now described by:
def resultsne(t,a,e,mu,theta,p1,p2,beta,s1,s2):
    consne=({'type':'eq','fun':lambda x:
    np.array([expenditurene(x,t,theta,p1,p2)-beta])},
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])})
    def objne(x,t,a,e,mu,theta,p1,p2):
        return (-theta*((1-p1)*v(0,0,x[0]+x[1],a,e)+
        p1*v(0,0,x[0],a,e))-(1-theta)*
        (p2*wa(t,x[0]+x[1],a,e,mu)+
        (1-p2)*wa(t,x[0],a,e,mu)))
    res=minimize(objne,[s1,s2],
    args=(t,a,e,mu,theta,p1,p2),constraints=consne,
    method='SLSQP', options={'ftol':1e-10,'disp':False})
    return ('b=',res.x[0], 'c=',res.x[1],
    'V^N/V^P=',((-objne(res.x,t,a,e,mu,theta,p1,p2))/
    (-resultspd(t,a,e,mu,theta,beta)[5])),
    'avsmvir=',
    avsmvir(t,res.x[0],res.x[1],a,e,mu,theta,p1,p2),
    'avsvinr=',
    avsmvnr(t,res.x[0],a,e,mu,theta,p1,p2))

#Note that s1 and s2 are starting points in the search
    for the optimal x[0](b) and x[1](c) - we can vary
    s1 and s2 to check the robustness of the results.

#The value of c that exhausts the budget constraint for
    any given level of b is:
def cnebudget(b,t,a,e,mu,theta,p1,p2,beta):
   return ((beta-b)/(theta*(1-p1)+(1-theta)*p2))

#The below functons loops 'cnebudget' over b and prints
    out the the resulting array. This allows us to
    check the budget set.
def necbudget(t,a,e,mu,theta,p1,p2,beta):
    print "(p1=",p1,")","[",cnebudget(0.00001,
    t,a,e,mu,theta,p1,
    p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print cnebudget(b,t,a,e,mu,theta,p1,p2,beta),",",
    print cnebudget(0.099,t,a,e,mu,theta,p1,p2,beta),"]"

#To ensure that both starting points s1 and s2 are
```

```
    budget feasible, we can simply let
    s2=cnebudget(s1,t,a,e,mu,theta,p1,p2,beta).

#The below function extracts optima from 'resultsne'
    and prints the results for successive values of p2
    (for a given p1). For x[0] set i=1; for x[1] set
    i=3; for V^N/W^U set i=5, for 'avsmvir' set i=7;
    and for 'avsmvinr' set i=9
def optne(t,a,e,mu,theta,p1,beta,s1,i):
    p2=-0.02
    while p2<0.98:
        p2=p2+0.02
        print resultsne(t,a,e,mu,theta,p1,p2,beta,
        s1,cnebudget(s1,t,a,e,mu,theta,p1,p2,beta))[i]

#This function prints out the above results in array
    form (which can be then used for plotting)
def optnearray(t,a,e,mu,theta,p1,beta,s1,i):
    print "(",i,")(p1=",p1,")","[",
        resultsne(t,a,e,mu,theta,p1,
    0,beta,s1,cnebudget(s1,t,a,e,mu,theta,p1,0,beta))[i],",",
    p2=0
    while p2<0.98:
        p2=p2+0.02
        print resultsne(t,a,e,mu,theta,p1,p2,beta
        ,s1,cnebudget(s1,t,a,e,mu,theta,p1,p2,beta))[i],",",
    print resultsne(t,a,e,mu,theta,p1,0.99,beta,
    s1,cnebudget(s1,t,a,e,mu,theta,p1,p2,beta))[i],"]"

#For the purpose of establishing how the welfare
    function changes with the universal benefit level -
    and resulting categorical benefit size that
    exhausts the budget constraint - we define:
def objnenooptimisation(b,t,a,e,mu,theta,p1,p2,beta):
  return
      (theta*((1-p1)*v(0,0,b+cnebudget(b,t,a,e,mu,theta,
  p1,p2,beta),a,e)+
      p1*v(0,0,b,a,e))+(1-theta)*(p2*wa(t,
      b+cnebudget(b,t,a,e,
      mu,theta,p1,p2,beta),a,e,mu)+
      (1-p2)*wa(t,b,a,e,mu)))

#The below function loops 'objnenooptimisation' over b
    and at each stage prints out the resulting level of
    welfare. We can use the resulting array to plot how
    welfare changes with the budget feasible benefit
    levels.
def objnenooptimisationarray(t,a,e,mu,theta,p1,p2,beta):
    print
        "(p1=",p1,"p2=",p2,")","[",objnenooptimisation(
    0.00001,
    t,a,e,mu,theta,p1,p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print objnenooptimisation(b,t,a,e,mu,
        theta,p1,p2,beta),",",
    print objnenooptimisation(0.099,t,a,e,mu,
```

```
    theta,p1,p2,beta),"]"

#3.2 FULL ENFORCEMENT
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

#When the condition that recipients of the categorical
    benefit do not work is fully enforced,the critical
    gross wage at or below which an able individual
    chooses to apply for the categorical benefit is
    given by the root to the below function ('vdiff'):
def vdiff(n,t,mun,ma,a,e):
    return v(0,t,mun,a,e)-v(n,t,ma,a,e)

#where 'mun' is the unearned income intended for the
    unable (i.e. b+c); whilst 'ma' is the unearned
    income intended for the able (i.e. b).

#We wish to provide a search region [a,b] in which the
    root to vdiff is contained, whereby
    vdiff(a,t,mun,ma,a,e)>0 but
    vdiff(b,t,mun,ma,a,e)<0. The lower bound (a) is
    given by the function nbar, whilst the upper bound
    (b) is given by specifying the function:
def ntbar(t,mun,a,e):
    return
        ((((a*((mun)**((e-1)/e))+(1-a))**e)*(a**(-e))-
    (((1-a)/a)**e))**(1/(e-1)))/(1-t)

#The below function solves for the critical
    productivity:
def ndbar(t,mun,ma,a,e):
    if (v(ntbar(t,mun,a,e),t,ma,a,e)>v(0,0,mun,a,e)>
    v(nbar(t,mun,a,e),t,ma,a,e)):
        return
            brentq(vdiff,nbar(t,ma,a,e),ntbar(t,mun,a,e),
        args=(t,mun,ma,a,e))
    else:
        return newton(vdiff,nbar(t,mun,a,e),
        args=(t,mun,ma,a,e))

#The proportion of able individuals who choose to apply
    is thus:
def approportion(t,mun,ma,a,e,mu):
    return expon.cdf(ndbar(t,mun,ma,a,e),scale=1/mu)

#For the purposes of later constructing the objective
    (social welfare) function, we define the following
    two functions: the first ('vpdfdiff') is the
    welfare gain to an applicant who is awarded the
    categorical benefit (multiplied by the distribution
    pdf), whilst the second ('wadiff) integrates this
    over all applicants.
def vpdfdiff(n,t,b,c,a,e,mu):
    return ((v(0,t,b+c,a,e)-v(n,t,b,a,e))*
    expon.pdf(n,scale=1/mu))

def wadiff(t,b,c,a,e,mu):
    return (quad(vpdfdiff,0,ndbar(t,b+c,b,a,e),
```

```
    args=(t,b,c,a,e,mu))[0])

#By the implicit function theorem, we can write the
    partial derivatives of ndbar with respect to b and
    c as the functions:
def ndbarb(t,b,c,a,e,mu,theta):
    return ((a*((a+(1-a)*((b+c)**((1-e)/e)))**(1/(e-1)))
    -(((a**e)+((1-a)**e)*
    ((ndbar(t,b+c,b,a,e)*(1-t))**(1-e)))**(1/(e-1))))/
    ((1-t)*(((a**e)+((1-a)**e)*
    ((ndbar(t,b+c,b,a,e)*(1-t))**(1-e)))**((2-e)/(e-1)))
    *((a**e)-((1-a)**e)*((ndbar(t,b+c,b,a,e)*(1-t))**(-e))
    *b)))

def ndbarc(t,b,c,a,e,mu,theta):
    return
        (a*((a+(1-a)*(((b+c)**((1-e)/e))))**(1/(e-1)))/
    ((1-t)*(((a**e)+((1-a)**e)*
    ((ndbar(t,b+c,b,a,e)*(1-t))**(1-e)))**((2-e)/(e-1)))
    *((a**e)-((1-a)**e)*((ndbar(t,b+c,b,a,e)*(1-t))**(-e))
    *b)))

#From Proposition 3a in the main text, we can write the
    left and right sides of (32) - evaluated at any
    benefit levels - by:
def fetsmvinr(t,b,c,a,e,mu,theta,p1,p2):
    return (theta*p1*vm(0,0,b,a,e)+(1-theta)*
    (wam(t,b,a,e,mu)-p2*quad(vmpdf,0,
    ndbar(t,b+c,b,a,e),args=(t,b,a,e,mu))[0]))

def psi(t,b,c,a,e,mu,theta,p1,p2):
    return (vm(0,0,b+c,a,e)*
    (theta*(1-p1)+(1-theta)*
    expon.cdf(ndbar(t,b+c,b,a,e),
    scale=1/mu)*p2)*(((1-theta*(1-p1)-(1-theta)*
    expon.cdf(ndbar(t,b+c,b,a,e),scale=1/mu)*p2)
    +(1-theta)*expon.pdf(ndbar(t,b+c,b
    ,a,e),scale=1/mu)*p2*(ndbarb(t,b,c,a,e,mu,
    theta)-ndbarc(t,b,c,a,e,mu,theta))*c)/(
    (theta*(1-p1)+(1-theta)*expon.cdf(ndbar(t,b+c,b,a,e),
    scale=1/mu)*p2)+(1-theta)*expon.pdf(ndbar(t,b+c,b,a,e),
    scale=1/mu)*p2*ndbarc(t,b,c,a,e,mu,theta)*c)))

#We once more let the vector x=(x[0],x[1])=(b,c) denote
    the variables we are optimising over.

#Benefit expenditure (left side of the budget
    constraint) is now given by:
def expenditurefe(x,t,a,e,mu,theta,p1,p2):
    return (x[0]+(theta*(1-p1)+(1-theta)*
    p2*approportion(t,x[0]+x[1],x[0],a,e,mu))*x[1])

#The optimisation problem is:
def resultsfe(t,a,e,mu,theta,p1,p2,beta,s1,s2):
    consfe=({'type':'eq',
'fun':lambda x:np.array([expenditurefe(x,t,a,e,mu
,theta,p1,p2)-beta])},
{'type':'ineq',
```

```
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])})
    def objfe(x,t,a,e,mu,theta,p1,p2):
        return (-theta*((1-p1)*u(x[0]+x[1],a,e)+
        p1*u(x[0],a,e))-
        (1-theta)*(wa(t,x[0],a,e,mu)+
        p2*wadiff(t,x[0],x[1],a,e,mu)))
    res=minimize(objfe,[s1,s2],args=(t,a,e,mu,theta,p1,p2)
    ,bounds=((0,None),(0,None)),
    constraints=consfe,method='SLSQP',
    options={'ftol':1e-10,'disp':False})
    return ('b=',res.x[0], 'c=', res.x[1],
    'V^F/V^P=',((-objfe(res.x,t,a,e,mu,theta,p1,p2))/
    (-resultspd(t,a,e,mu,theta,beta)[5])),
    'fetsmvinr=',
    fetsmvinr(t,res.x[0],res.x[1],a,e,mu,theta,p1,p2)
    ,'psi=',
    psi(t,res.x[0],res.x[1],a,e,mu,theta,p1,p2),
    'Fnbar=',expon.cdf(nbar(t,res.x[0],a,e),scale=1/mu),
    'Fndbar=',
    approportion(t,res.x[0]+res.x[1],res.x[0],a,e,mu),
    'Fnbar/Fndbar=',
    (expon.cdf(nbar(t,res.x[0],a,e),scale=1/mu)/
    approportion(t,res.x[0]+res.x[1],res.x[0],a,e,mu)))

#To establish the value of c that exhausts the budget
    constraint we first define the budget constraint
    for any value of b and c:
def febudgetcondition(c,b,t,a,e,mu,theta,p1,p2,beta):
    return (b+(theta*(1-p1)+(1-theta)*
    p2*approportion(t,b+c,b,a,e,mu))*c)-beta

#and then find the root to the above condition:
def cbudget(b,t,a,e,mu,theta,p1,p2,beta):
    return brentq(febudgetcondition,0.001,2,
    args=(b,t,a,e,mu,theta,p1,p2,beta))

#The below functon loops 'cbudget' over b and prints
    out the the resulting array. This allows us to
    check the budget set.
def fecbudget(t,a,e,mu,theta,p1,p2,beta):
    print "(p1=",p1,"p2=",p2,")","[",cbudget(0.00001,
    t,a,e,mu, theta,p1,p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print cbudget(b,t,a,e,mu,theta,p1,p2,beta),",",
    print cbudget(0.099,t,a,e,mu,theta,p1,p2,beta),"]"

#The derivative of 'cbudget' with respect to b is:
def cbudgetprime(b,t,a,e,mu,theta,p1,p2,beta):
    return -((1+(1-theta)*p2*expon.cdf(ndbar(t,b+
    cbudget(b,t,a,e,mu,theta,p1,p2,beta),b,a,e),
    scale=1/mu)*ndbarb(t,b,cbudget(b,t,a,e,mu,theta,p1
    ,p2,beta),a,e,mu,theta)
    *(1-t)*cbudget(b,t,a,e,mu,theta,p1,p2,beta))/
    (theta*(1-p1)+(1-theta)*p2*expon.cdf(ndbar(t,b+
```

```
    cbudget(b,t,a,e,mu,theta,p1,p2,beta),b,a,e),
    scale=1/mu)*ndbarc(t,b,cbudget(b,t,a,e,mu,
    theta,p1,p2,beta),a,e,mu,theta)
    *(1-t)*cbudget(b,t,a,e,mu,theta,p1,p2,beta)))

#Instead of picking starting points (s1,s2) in the
    search for the optimal benefit levels, we can use
    function 'cbudget' so that the choice of s1
    automatically determines the budget exhausting
    value of s2. This ensures our search always starts
    in the feasible set.

#The below function extracts optima from 'resultsfe'
    and prints the results for successive values of p2
    (for a given p1). For x[0] set i=1, for x[1] set i=3
def optfe(t,a,e,mu,theta,p1,beta,s1,i):
    p2=-0.1
    while p2<0.9:
        p2=p2+0.1
        print resultsfe(t,a,e,mu,theta,p1,p2,beta
        ,s1,cbudget(s1,t,a,e,mu,theta,p1,p2,beta))[i]

#To print the results out in an array for plotting use:
def optfearray(t,a,e,mu,theta,p1,beta,s1,i):
    print "(p1=",p1,")","[",resultsfe(t,a,e,mu,theta,p1,
    0,beta,s1,
    cbudget(s1,t,a,e,mu,theta,p1,0,beta))[i],",",
    p2=0
    while p2<0.98:
        p2=p2+0.02
        print resultsfe(t,a,e,mu,theta,p1,p2,beta,s1,
        cbudget(s1,t,a,e,mu,theta,p1,p2,beta))[i],",",
    print resultsfe(t,a,e,mu,theta,p1,1,beta,s1,
    cbudget(s1,t,a,e,mu,theta,p1,p2,beta))[i],"]"

#For the purpose of establishing how the welfare
    function changes with the universal benefit level -
    and resulting categorical benefit size that
    exhausts the budget constraint - we define:
def objfenooptimisation(b,t,a,e,mu,theta,p1,p2,beta):
        return (theta*((1-p1)*u(b+cbudget(b,t,a,e,mu,
        theta,p1,p2,beta),a,e)+p1*u(b,a,e))+
        (1-theta)*(wa(t,b,a,e,mu)+
    p2*wadiff(t,b,cbudget(b,t,a,e,mu,
    theta,p1,p2,beta),a,e,mu)))

#The below function loops 'objfenooptimisation' over b
    and prints out the resulting array. This allows us
    to plot how welfare changes with the budget
    feasible benefit levels.
def objfenooptimisationarray(t,a,e,mu,theta,p1,p2,beta):
    print
        "(p1=",p1,"p2=",p2,")","[",objfenooptimisation(
    0.00001,t,a,e,mu,theta,p1,p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print objfenooptimisation(b,t,a,e,mu,theta,
```

```
        p1,p2,beta),",",
    print objfenooptimisation(0.099,t,a,e,mu,theta,
    p1,p2,beta),"]"

#It is of interest to establish how the number of able
    individuals who choose to apply for the categorical
    benefit changes over the feasible set. We thus
    define the below function that substitutes
    'cbudget' into 'approportion':
def feappnooptimisation(b,t,a,e,mu,theta,p1,p2,beta):
    return approportion(t,b+cbudget(b,t,a,e,mu,theta,p1,
    p2,beta),b,a,e,mu)

#We then define the below function to loop
    'feappnooptimisation' over b:
def feappnooptimisationarray(t,a,e,mu,theta,p1,p2,beta):
    print
        "(p1=",p1,"p2=",p2,")","[",feappnooptimisation(
    0.00001, t,a,e,mu,theta,p1,p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print feappnooptimisation(b,t,a,e,mu,theta,p1
        ,p2,beta),",",
    print feappnooptimisation(0.099,t,a,e,mu,theta,p1,
    p2,beta),"]"

#Recall that the Full Enforcement optimality condition
    (Proposition 3a in the main text) equates (i) the
    total smvi of non-categorical recipients; with (ii)
    with the total smvi of categorical recipients
    multiplied by the reduction in their benefit income
    associated with an increase in the universal
    benefit. To check how thee two terms change with
    the universal benefit size (and resulting
    categorical benefit that exhausts the budget) we
    define the below two functions:
def fetsmvinrnoopt(b,t,a,e,mu,theta,p1,p2,beta):
    return
        fetsmvinr(t,b,cbudget(b,t,a,e,mu,theta,p1,p2,beta)
    ,a,e,mu,theta,p1,p2)

def psinoopt(b,t,a,e,mu,theta,p1,p2,beta):
    return psi(t,b,cbudget(b,t,a,e,mu,theta,p1,p2,beta)
    ,a,e,mu,theta,p1,p2)

#To plot how these functions change with the universal
    benefit we define two functions which loop them
    over b:
def fetsmvinrnooptarray(t,a,e,mu,theta,p1,p2,beta):
    print "(p1=",p1,"p2=",p2,")","[",fetsmvinrnoopt(
    0.00001,t,a,e,mu,theta,p1,p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print
            fetsmvinrnoopt(b,t,a,e,mu,theta,p1,p2,beta),",",
    print
```

```
        fetsmvinrnoopt(0.099,t,a,e,mu,theta,p1,p2,beta),"]"

def psinooptarray(t,a,e,mu,theta,p1,p2,beta):
    print "(p1=",p1,"p2=",p2,")","[",psinoopt(
    0.00001,t,a,e,mu,theta,p1,p2,beta),",",
    b=0
    while b<0.098:
        b=b+0.002
        print psinoopt(b,t,a,e,mu,theta,p1,p2,beta),",",
    print psinoopt(0.099,t,a,e,mu,theta,p1,p2,beta),"]"

vcbudget=np.vectorize(cbudget)
```

2

Chapter 3

Revisiting the Optimal Linear Income Tax with Categorical Transfers.[1]

3.1 Introduction

When individuals differ in both their productivity and some categorical dimension such as disability, a well-established result is that categorical transfers should be set so as to eliminate inequality in the average social marginal value of income (smvi) between categorical groups (Diamond and Sheshinski, 1995; Parsons, 1996). The linear income tax framework has played an important role in the analysis of categorical transfers: proponents of flat tax schedules cite their administrative simplicity and enhanced work incentives; whilst analytically a flat tax captures the equity-efficiency tradeoff of income taxation more tractably than nonlinear taxation (Atkinson, 1995; Paulus and Peichl, 2009; Peichl, 2014).[2] The resulting optimal tax formulae are typically reported under the assumption that inequality in the average *net* smvi is indeed eliminated at the optimum (Viard, 2001a). This assumption allows the optimal tax expression to be written as in the uni-dimensional model where individuals differ only in their productivity: the numerator (equity considerations) is the negative of the covariance between gross

[1] A shorter version of this chapter is published in *Economics Letters*: see Slack (2015). This publication presents some of the key results derived in this chapter.

[2] Mirrlees (1971, p.208) discusses the desirability of approximately linear tax schedules.

earnings and the *net* smvi; whilst the denominator (efficiency considerations) captures the response of compensated gross earnings to a change in the net wage rate.

However, it is not immediately clear that this between-group inequality will always be eliminated at the optimum. Indeed, where categorical transfers are financed by tax revenue there may be cases where it is suboptimal to do so. For example, if a sufficiently large fraction of the population are dependent on categorical transfers for consumption then the level of taxation required to equate the average net smvi of dependent and non-dependent groups may be too harmful to the latter group. This will also depend on the size of any revenue requirement in place for spending outside welfare.

Moreover, this is likely to hold beyond a simple flat tax framework. For example, progressive piecewise linear tax systems provide the government with additional tools to redistribute *within* categorical groups; but if shifting some of the tax burden away from lower earners in an able group: (i) pushes the average net smvi of that group further below that of a dependent group; and/or (ii) lowers tax revenue relative to the flat tax case, this may limit further the cases where it is optimal to eliminate between-group inequality in the average net smvi.

This chapter addresses this issue in both linear and piecewise linear income tax frameworks. It demonstrates that the optimal tax expressions can be written more generally to allow for cases where the average net smvi of categorical groups are not equated at the optimum. In these cases welfare provision is purely categorical, such that no universal benefit is provided. Alternatively, if between-group inequality is eliminated *and* there are resources left over a universal benefit is also provided. Extensive numerical simulations provide examples where between-group inequality is not eliminated at the optimum. Further, they indicate that it is more likely to arise under a progressive piecewise system for the reasons outlined above.

The remainder of this chapter is structured as follows. Section 3.2 sets up the model and analyses the flat tax case. Within this section, numerical examples where between-group inequality persists at the optimum are provided using a variant of the framework employed by Stern (1976) (i.e. CES preferences, lognormal productivity distribution). Section 3.3 then extends the analysis to the less restrictive case of piecewise linear taxation with two tax brackets and increasing marginal tax rates. To simplify the exposition preferences are taken to be quasilinear in consumption (see also Apps et al., 2014). As agent monotonicity[3] is readily satisfied under these preferences there will be a bunching of earners at the earnings threshold that separates the two tax brackets. Numerical examples where between-group inequality persists under the piecewise optima (and, for comparison, the flat tax optima) are obtained using preferences with a constant labour elasticity (see Atkinson, 1990; Saez, 2001).[4] Individual productivity is taken to be Pareto distributed, as this is known to give rise to increasing marginal tax rates (Diamond, 1998). Finally, Section 3.4 concludes the chapter.

3.2 The Model

3.2.1 Background

Individual preferences over consumption, $x \geq 0$, and leisure, $l \in [0, 1]$, are represented by the utility function $u(x, l)$. The standard assumptions apply: u is continuous; differentiable; increasing in both arguments ($u_x > 0, u_l > 0$) and concave ($u_{xx} < 0$, $u_{ll} < 0$, $u_{xx}u_{ll} - u_{xl}^2 > 0$); with both goods normal ($u_l u_{xx} - u_x u_{xl} < 0$).

For an individual with net wage $\omega \geq 0$ and unearned income $M \geq 0$, optimal labour supply (H^*) and the resulting indirect utility function (v) are defined by:

$$H^*(\omega, M) \equiv \arg \max_{H \in (0,1)} u(\omega H + M, 1 - H) \ ,$$
$$v(\omega, M) \equiv u(\omega H^* + M, 1 - H^*) \ .$$

Let $\bar{\omega}(M) = u_l(M, 1)/u_x(M, 1)$ be the reservation wage satisfying: $H^* = 0 \ \forall \ \omega \leq \bar{\omega}$ and $H^* > 0 \ \forall \ \omega > \bar{\omega}$; where $\bar{\omega}' > 0$. It follows that $\forall \ \omega \leq \bar{\omega} : \ v(\omega, M) = u(M, 1)$ and thus $v_M(\omega, M) = u_x(M, 1)$. Contrastingly, Roy's identity ($v_\omega = v_M H^*$) and the normality of leisure ($H_M^* < 0$) imply that $\forall \ \omega > \bar{\omega} : \ v_{\omega M} = v_{MM} H^* + v_M H_M^* < 0$. So for $\omega > \bar{\omega}$ the marginal indirect utility of unearned income is strictly decreasing in the

[3] Agent monotonicity requires that indifference curves in gross income - consumption space become flatter in individual productivity (see, for example, Seade, 1982).

[4] These isoelastic preferences also allow one to derive analytical results for the conditions under which between-group inequality will persist under the flat tax optimum. Specifically, for the special case where taxation is purely redistributive (i.e. no revenue requirement) this chapter demonstrates that the optimal flat tax will fall below that required to eliminate inequality in the average smvi if the unable subpopulation is too large.

net wage rate.

3.2.2 The Tax-Benefit System

Consider a population of size 1, where a fraction $\theta \in (0,1)$ of individuals face a zero quantity constraint on labour supply and are thus unable to work. Absent any form of state financial provision these individuals would have zero income to consume. The remaining $(1-\theta)$ individuals are able to work but differ in their underlying productivity $n \geq 0$, where n is distributed with density function $f(n)$ and associated distribution function $F(n)$.

The government operates a tax-benefit system comprising (i) a constant marginal income tax rate $t \in (0,1)$; (ii) a tax-free universal benefit $B \geq 0$ received unconditionally by all individuals in society; and (iii) a tax-free categorical benefit $C \geq 0$ that is perfectly targeted at unable individuals.

To save on notation, let $y = nH$ and $\bar{y} = \int_0^\infty y f(n)dn$. Evaluated at the optimal labour choices of individuals, we thus define the gross earnings of a productivity n individual by $y^*(n, 1-t, M) \equiv nH^*[n(1-t), M]$; whilst the average gross earnings over able individuals are $\bar{y}^*(1-t, M) \equiv \int_0^\infty y^* f(n)dn$.

Under a strictly utilitarian criterion, social welfare is:

$$W(t, B, C; \theta) = \theta u(B + C, 1) + (1-\theta) \int_0^\infty v[n(1-t), B] f(n)dn \tag{3.1}$$

The first term is the welfare of unable individuals multiplied by their population share; whilst the second term is the average welfare of able individuals multiplied by their population share.

The government's optimisation problem is thus described by:

$$\begin{aligned} & \max_{t,B,C} W(t, B, C; \theta) \\ \text{s.t.} \quad & B + \theta C = (1-\theta)t \cdot \bar{y}^*(t, B) - R \ , \\ & t \in (0,1) \ , \ B \geq 0 \ , \ C \geq 0 \ . \end{aligned} \tag{3.2}$$

where $R \geq 0$ is an exogenous revenue requirement.

To discuss the results which follow, let the net smvi of a productivity n individual be (Viard, 2001a):

$$s(n,t,M,\lambda) = \begin{cases} u_x(M,1) & : n \leq \bar{n}(t,M) \\ v_M[n(1-t),M] + \lambda t y_M^*(n,1-t,M) & : n > \bar{n}(t,M) \end{cases} \tag{3.3}$$

where $\bar{n} \equiv \bar{\omega}/(1-t)$ and λ is the shadow price of public expenditure.[5] For the voluntarily unemployed, s is simply the social marginal utility of income. However, for working individuals s also captures - in welfare units - the fact that an increase in unearned income induces a worker to reduce their labour supply and, consequently, lowers tax revenue.

Let $\hat{t}$, $\hat{B}$ and $\hat{C}$ denote the optimal choices resulting from the maximisation problem in (3.2). We state the following result:

Result 1:

(i) $\hat{C} > 0$ *and* $\hat{B} \geq 0$ *satisfy*:

$$\bar{s}\left(\hat{t},\hat{B},\hat{\lambda}\right) \leq u_x\left(\hat{B}+\hat{C},1\right) = \hat{\lambda}\ ;\ \hat{B} \geq 0 \tag{3.4}$$

where the pair of inequalities hold with complementary slackness and $\hat{\lambda}$ *is the shadow price of public expenditure at the optimum.*

(ii) *For* $\delta \equiv (\lambda - \bar{s})$ *and* $r = y^*/\bar{y}^*$; $\hat{t}$ *is implicitly characterised by:*

$$\frac{\hat{t}}{1-\hat{t}} = \begin{cases} \dfrac{\delta - \mathrm{Cov}(r,s)}{\hat{\lambda}\int r\mathcal{E}^c f(n)dn} & : \delta > 0 \\ \\ \dfrac{-\mathrm{Cov}(y^*,s)}{\hat{\lambda}\int y\mathcal{E}^c f(n)dn} & : \delta = 0 \end{cases} \tag{3.5}$$

where $\mathcal{E}^c$ *is the compensated elasticity of individual gross earnings with respect to the net of tax rate.*

Proof: See Appendix

[5] Atkinson and Stiglitz (1980, p.387) normalise s by λ and define it as the net social marginal valuation. In terms of deriving the optimal tax expression the two approaches are equivalent.

Result 1(i) states that, budget allowing, C should be set so as to eliminate inequality in the average net smvi between the unable and able subpopulations (see Viard, 2001a). Further, so long as there is inequality in this dimension it is optimal to set $B = 0$ because social welfare can be increased more through targeting resources at the unable.

Next, Result 1(ii) provides a general expression for the optimal linear tax rate that captures the equity (numerator) - efficiency (denominator) tradeoff inherent in income taxation. This *differs* from standard linear tax expressions whenever $\delta > 0$, and thus whenever there is inequality in the average net smvi between the unable and able subpopulations. In this case, the numerator is composed of *two* terms. The first is δ itself, which will be larger the greater is the between-group disparity at the optimum. The second is the covariance between *relative* earnings (r) and the net smvi. Notice that $\mathrm{Cov}(r, s) \cdot \bar{y} = \mathrm{Cov}(y, s)$, where the latter term is found in all linear tax formulae and captures a desire to redistribute from those of high productivity to those of lower productivity.[6] The intuition for these two terms is that the presence of a dependent subpopulation shifts the equity focus away from disparities in earnings ability *within* the able subpopulation, and towards the *between*-group disparity in the average net smvi.

The denominator in (3.5) captures the efficiency considerations involved in setting the optimal tax rate and is unambiguously positive. *Ceteris paribus*, higher compensated elasticities of labour supply imply lower tax rates, with emphasis placed on both very high productivities and productivities at which the population is most dense.

Finally, if $\delta = 0$ then between-group inequality is eliminated at the optimum and the tax formula reduces to the standard representation in the literature (Atkinson and Stiglitz, 1980; Atkinson, 1995; Viard, 2001a,b).

[6] This covariance will be negative if: (i) $y_n \geq 0$; and (ii) $s_n < 0$. The requirement that the net smvi falls with productivity is formally given by:

$$\frac{\partial s}{\partial n} = (1-t)v_\omega + \lambda t y_{nM} = (1-t)v_\omega + \lambda t \left\{ H_M^* + n(1-t)H_{\omega M}^* \right\}$$

As stated in the main text, we have $v_{\omega M} < 0 \; \forall \; n > \bar{n}$ and similarly $H_M^* < 0 \; \forall \; n > \bar{n}$. However, the sign of $H_{\omega M}^*$ is unclear and requires assumptions on third derivatives.

3.2.3 Numerical Results: Flat Tax

The purpose of this section is to provide examples where between-group inequality in the average net smvi persists at the flat tax optimum. In line with the key numerical studies on linear income taxation - both those with and without categorical transfers - we take preferences over consumption and leisure to be of the constant elasticity of substitution form (see Immonen et al., 1998; Stern, 1976; Viard, 2001a,b):

$$u(x,l) = \left[\alpha x^{\frac{\mathcal{E}-1}{\mathcal{E}}} + (1-\alpha) l^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right] \quad ; \; \mathcal{E} \neq 1 \tag{3.6}$$

where $\mathcal{E}$ is the elasticity of substitution between leisure and consumption.

The social welfare function is given by:

$$W = \left(\frac{1}{1-\eta}\right)\left\{\theta u(B+C,1)^{1-\eta} + (1-\theta)\int_0^\infty v[n(1-t),B]^{1-\eta} f(n)dn\right\} \tag{3.7}$$

Notice that the parameter η determines the degree of concavity of the social welfare function: setting $\eta > 0$ allows for express concern with regard to the distribution of utilities. We will consider $\eta \in \{0,2\}$. The smvi of an unable or voluntarily unemployed individual is thus $u^{-\eta} \cdot u_x$; whilst the net smvi of a worker is $v^{-\eta} \cdot v_M + \lambda t y_M^*$.

Whilst the main properties of (3.6) are well established in the literature, it is useful to note that the optimal earnings function is given by:

$$y^* = \frac{n - (\frac{1-\alpha}{\alpha})^{\mathcal{E}} n^{1-\mathcal{E}}(1-t)^{-\mathcal{E}} M}{1 + (\frac{1-\alpha}{\alpha})^{\mathcal{E}}[n(1-t)]^{1-\mathcal{E}}} \tag{3.8}$$

from which it follows that $\partial y^*/\partial M < 0$ whilst $\partial^2 y^*/\partial M \partial(1-t) > 0$.

If $M = 0$ for workers - which from Result 1 would arise if between-group inequality in the average (net) smvi is not eliminated through categorical transfers - then (3.8) illustrates that earnings will be falling in the net-of-tax rate. Conversely, earnings will be rising in the tax rate, as will be tax revenue.

In the simulations which follow we consider the following parameter values: $n \sim \ln\mathcal{N}(\mu = -1, \sigma = 0.39)$; $\alpha = 0.614$; $\mathcal{E} \in \{0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$; $\theta = 0.1$ and $R \in \{0, 0.05, 0.10\}$. The key studies of categorical transfers within an optimal tax

Table 3.1: Numerical results: optimal linear income tax and between-group inequality.

		$\eta = 0$					$\eta = 2$				
$\mathcal{E}$	t_R	$\hat{t}$	$\hat{B}$	$\hat{C}$	δ	t_δ	$\hat{t}$	$\hat{B}$	$\hat{C}$	δ	t_δ
(a) $R = 0$											
0.5	0	0.304	0	0.755	**0.014**	0.311	0.458	0.087	0.139	0	0.087
0.6	0	0.289	0	0.700	**0.022**	0.302	0.416	0.076	0.134	0	0.083
0.7	0	0.274	0	0.647	**0.029**	0.293	0.382	0.067	0.129	0	0.079
0.8	0	0.259	0	0.597	**0.035**	0.285	0.353	0.060	0.123	0	0.075
0.9	0	0.245	0	0.551	**0.040**	0.276	0.329	0.054	0.117	0	0.072
0.99	0	0.233	0	0.512	**0.044**	0.268	0.310	0.049	0.112	0	0.069
(b) $R = 0.05$											
0.5	0.205	0.473	0	0.724	**0.055**	0.499	0.526	0.060	0.130	0	0.274
0.6	0.209	0.460	0	0.654	**0.078**	0.501	0.484	0.048	0.128	0	0.275
0.7	0.213	0.446	0	0.583	**0.102**	0.506	0.449	0.038	0.124	0	0.277
0.8	0.218	0.432	0	0.515	**0.126**	0.513	0.419	0.030	0.120	0	0.279
0.9	0.223	0.419	0	0.451	**0.153**	0.524	0.393	0.023	0.115	0	0.281
0.99	0.228	0.408	0	0.397	**0.177**	0.540	0.373	0.017	0.111	0	0.284
(c) $R = 0.10$											
0.5	0.394	0.629	0	0.703	**0.114**	0.674	0.601	0.036	0.118	0	0.446
0.6	0.403	0.619	0	0.615	**0.165**	0.696	0.562	0.025	0.117	0	0.453
0.7	0.414	0.608	0	0.524	**0.223**	0.736	0.529	0.015	0.114	0	0.462
0.8	0.426	0.597	0	0.435	**0.293**	*	0.500	0.006	0.111	0	0.472
0.9	0.440	0.588	0	0.352	**0.376**	*	0.485	0	0.105	**0.571**	0.485
0.99	0.454	0.583	0	0.283	**0.470**	*	0.497	0	0.094	**2.730**	0.499

Notes: t_R denotes the tax rate that generates just enough tax revenue to satisfy the revenue requirement. $\hat{t}, \hat{B}$ and $\hat{C}$ denote the optimal choices. t_δ is the critical tax rate that generates just enough tax revenue - net of any revenue requirement - to eliminate between-group inequality in the average net smvi through categorical spending. Cases where t_δ does not exist are denoted by an $*$: this arises because the average net smvi may fall with the tax rate at high tax rates. Values $\leq 10^{-7}$ are reported as 0.

framework take individual productivity to be lognormally distributed, and we follow convention here (see Immonen et al., 1998; Viard, 2001a,b). We set the mean of $\log n$ equal to -1 and the standard deviation of $\log n$ equal to 0.39; as has been the standard since Mirrlees (1971). The choice of α follows from Stern (1976): at $\alpha = 0.614$ the average productivity individual works exactly 2/3 of their time endowment when $\mathcal{E} = 0.5$ and $t = M = 0$. Turning to the proportion of unable individuals in society, the choice of $\theta = 0.1$ seems sensible in light of statistics on real-world benefit programmes (Mcinnes, 2012). Finally, the choices of positive revenue requirement can be written as a proportion of maximum earnings when $t = 0$. When $\mathcal{E} = 0.5$ aggregate earnings in the economy are 0.235. A value of $R = 0.05$ thus corresponds to roughly 21% of maximum earnings, whilst $R = 0.10$ corresponds to roughly 43% of maximum earnings.

The numerical results are presented in Table 3.1.[7] The column labelled t_R denotes the tax rate that generates just enough tax revenue to satisfy the revenue requirement, with benefit expenditure set at zero. This is independent of η. The columns labelled $\hat{t}$, $\hat{B}$ and $\hat{C}$ denote the optimal choices of the tax rate and benefit levels. The columns labelled δ denote the level of between-group inequality in the average (net) smvi at the optimal choices. Given the welfare function in (3.7), δ is formally defined as:

$$\delta = u^{-\eta} u_x(B + C, 1) - \int_0^\infty \left\{ v^{-\eta} v_M[n(1-t), B] + \lambda t y_M^* \right\} dF(n)$$

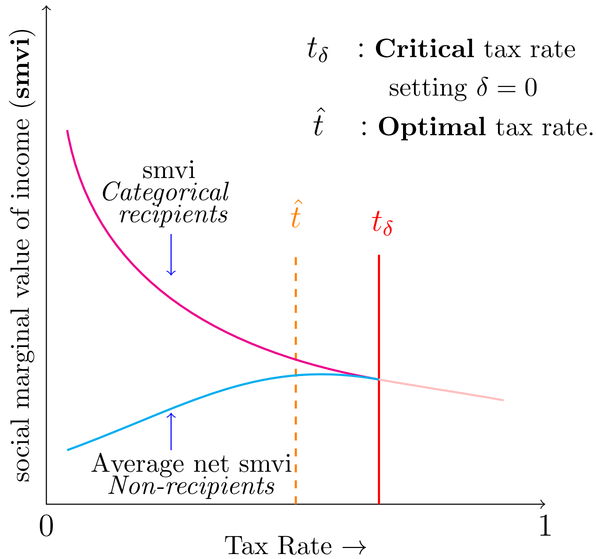
Relatedly, the columns labelled t_δ denote the critical tax rate that generates just enough tax revenue - net of any revenue requirement - to eliminate between-group inequality through categorical spending and thus set $\delta = 0$. Formally, t_δ satisfies:

$$u^{-\eta} u_x \left[\frac{(1-\theta) t_\delta \bar{y}(1 - t_\delta, 0) - R}{\theta}, 1 \right] = \int_0^\infty \left\{ v^{-\eta} v_M[n(1 - t_\delta), 0] + \lambda t_\delta y_M^* \right\} dF(n)$$

The implication is that $t < t_\delta \Rightarrow \delta > 0$, but $t \geq t_\delta \Rightarrow \delta = 0$. Figure 3.1 provides the graphical intuition for the critical flat tax t_δ.

[7] Note that if we set $\theta = 0$ - such that there is no need for categorical transfers - the numerical results in Table 3.1 collapse down to those in Stern (1976). This exercise provides a useful *additional* check on the accuracy of the simulation results.

Figure 3.1: The Critical Tax Rate t_δ



The most immediate observation from Table 3.1 is that when $\eta = 0$ we observe $\delta > 0$ and $t < t_\delta$ in *all* cases. Consequently, optimal welfare provision is purely targeted and the universal benefit is set at zero in all cases. Note that when $R = 0.10$ there is no t_δ that eliminates between-group inequality. The intuition is as follows: because high levels of R imply high levels of taxation (see t_R) without any form of benefit expenditure, the term $\lambda t y_M^*$ from the net smvi of an able individual may become sufficiently negative that the average net smvi over the able subpopulation falls with the tax rate (recall from our discussion of the earnings function in (3.8) that $y_M^* < 0$ becomes more negative with the tax rate). If this fall is sufficiently large there will be no way to eliminate inequality in the average net smvi.

Turning to the case where $\eta = 2$, we only observe cases where $t < t_\delta$ when $R = 0.10$. Higher levels of the revenue requirement thus still render it suboptimal to eliminate inequality in the average net smvi between the unable and able subpopulations. Notice that - in contrast to the case where $\eta = 0$ - the critical tax rate t_δ exists in all cases. The reason for this is that when $\eta = 2$ more weight is placed on the marginal indirect utility of lower productivity workers who are being taxed highly but receive no benefit income. Consequently, the average net smvi of the able subpopulation tends to rise with the tax rate.

Turning to the remaining observations from Table 3.1, it is interesting to note that t_δ

is (i) falling in $\mathcal{E}$ when $R = 0$; but (ii) increasing in $\mathcal{E}$ when $R > 0$. The intuition for this follows from the well-established result (see Stern, 1976) that t_R rises with $\mathcal{E}$. Finally, we observe that the optimal tax rate, $\hat{t}$, falls with $\mathcal{E}$ in almost all considered cases, as do the the optimal benefit levels.

3.3 A Progressive Piecewise Linear Income Tax System

Whilst a number of countries adopt a flat income tax, the majority employ progressive[8] piecewise linear tax schedules (see Paulus and Peichl, 2009; Peichl, 2014). This section first motivates the idea that the above discussion is likely to hold in the more general setting of piecewise taxation, where the set of instruments available to the government are less restrictive than in the simple flat tax case. We then proceed to demonstrate this numerically.

3.3.1 A two-bracket progressive piecewise linear tax schedule

Consider a simple progressive piecewise linear tax system with two tax brackets and an earnings threshold Y. Individuals are taxed at the rate t_1 on all earnings $y \leq Y$; but are taxed at the rate $t_2 \geq t_1$ on any *additional* earnings $(y - Y)$ in excess of the threshold. Formally, an individual's budget constraint is given by:

$$x = \begin{cases} (1 - t_1)y + B & : y \leq Y \\ (t_2 - t_1)Y + (1 - t_2)y + B & : y > Y \end{cases} \tag{3.9}$$

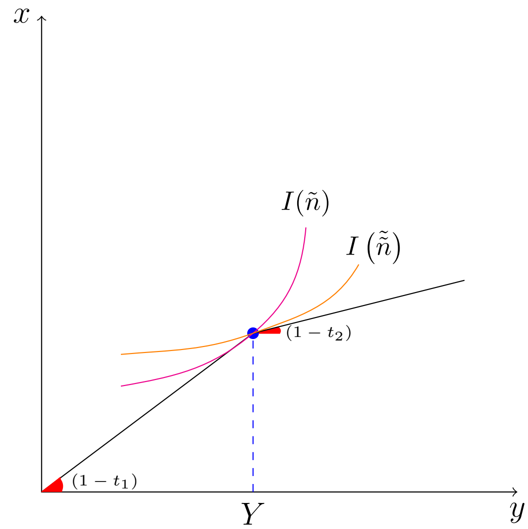
The tax-benefit system remains otherwise as in Section 3.2.2.

To progress it is helpful to abstract from income effects[9] through assuming preferences take the form:

[8] The term progressive is used here to refer to increasing marginal tax rates.

[9] This ensures that the first order condition for earnings $y > Y$ is unaffected by t_1, which simplifies the analysis (see Apps et al., 2014).

Figure 3.2: 'Bunching' of earnings at the bracket threshold Y



Note: This figure illustrates that under the preferences in (3.10) indifference curves in (y, x) space become flatter in n, resulting in the productivity interval $n \in [\tilde{n}, \tilde{\tilde{n}}]$ where individuals choose to earn *exactly* Y (see the Appendix for a more detailed discussion). Apps et al. (2014) refer to this progressive piecewise tax system as the 'convex case' because the budget set in (y, x) space is convex. Notice that in drawing the figure we have abstracted from unearned income (if $M > 0$ an individual would still have positive consumption when $y = 0$).

$$U(x, H) = u\left(x - g(H)\right) \tag{3.10}$$

where $u' > 0$; $u'' < 0$; $g' > 0$; and $g'' > 0$. In addition, we assume that $g(0) = g'(0) = 0$. The implication of this latter assumption is simply that there is no reservation wage at or below which an able individual chooses voluntary unemployment.

As is well documented (see Apps et al., 2014), there will be a 'bunching' of earnings at Y for individuals who *would* earn more than Y if additional earnings were still taxed at rate t_1; but choose not to because they are in fact taxed at rate t_2. Formally, this bunching occurs for $\tilde{n}(1-t_1, Y) < n \le \tilde{\tilde{n}}(1-t_2, Y)$; where:

$$\tilde{n}(1-t_1) \equiv g'\left(Y/\tilde{n}\right) \quad , \quad \tilde{\tilde{n}}(1-t_2) \equiv g'\left(Y/\tilde{\tilde{n}}\right) \tag{3.11}$$

It follows directly from (3.11) that both $\partial\tilde{n}/\partial(1-t_1) < 0$ and $\partial\tilde{\tilde{n}}/\partial(1-t_2) < 0$, whilst both $\partial\tilde{n}/\partial Y > 0$ and $\partial\tilde{\tilde{n}}/\partial Y > 0$. Figure 3.2 provides some intuition for the two critical productivities $\tilde{n}$ and $\tilde{\tilde{n}}$.

Taking into account the bunching of earnings at the 'kink' point in the budget constraint, we define the optimal earnings function of a productivity n individual by:

$$y^*(n, 1-t_1, 1-t_2, Y)$$

$$\equiv \begin{cases} \operatorname{Arg\,max}_{y \in (0,Y)} u\left[(1-t_1) + M - g(y/n)\right] & : \forall\ n \in (0, \tilde{n}) \\ Y & : \forall\ n \in [\tilde{n}, \tilde{\tilde{n}}] \\ \operatorname{Arg\,max}_{y \in (Y,\infty)} u\left[Y(t_2 - t_1) + y(1-t_2) + M - g(y/n)\right] & : \forall\ n \in (\tilde{\tilde{n}}, \infty) \end{cases} \tag{3.12}$$

From (3.11) and (3.12) one can readily verify that y^* satisfies:

$$\begin{aligned} n(1-t_1) = g'(y^*/n) &\quad : \forall\ n \in (0, \tilde{n}] \\ (1-t_1) > g'(Y/n)/n > (1-t_2) &\quad : \forall\ n \in (\tilde{n}, \tilde{\tilde{n}}) \\ n(1-t_2) = g'(y^*/n) &\quad : \forall\ n \in [\tilde{n}, \infty) \end{aligned} \tag{3.13}$$

where $\forall\ n \in (0, \tilde{n}) : \partial y^*/\partial n > 0$ and $\partial y^*/\partial(1-t_1) > 0$; whilst $\forall\ n \in (\tilde{\tilde{n}}, \infty)$: $\partial y^*/\partial n > 0$ and $\partial y/\partial(1-t_2) > 0$.

If we let $v(n, 1-t_1, 1-t_2, Y, M)$ denote the resulting indirect utility function we can establish that:

$$\frac{\partial v}{\partial(1-t_1)} = \begin{cases} v_M \cdot y^* & : \forall\ n \in (0, \tilde{n}] \\ v_M \cdot Y & : \forall n \in (\tilde{n}, \infty) \end{cases}, \quad \frac{\partial v}{\partial(1-t_2)} = \begin{cases} 0 & : \forall n \in (0, \tilde{\tilde{n}}] \\ v_M(y^* - Y) & : \forall n \in (\tilde{\tilde{n}}, \infty) \end{cases} \tag{3.14}$$

whilst:

$$\frac{\partial V}{\partial Y} = \begin{cases} 0 & : \forall n \in (0, \tilde{n}] \\ v_M \cdot \left[(1-t_1) - g'(Y/n)/n\right] & : \forall n \in (0, \tilde{\tilde{n}}] \\ v_M \cdot (t_2 - t_1) & : \forall n \in (\tilde{\tilde{n}}, \infty) \end{cases} \tag{3.15}$$

where $v_M = u'$. *Ceteris paribus*, a reduction in the lower net-of-tax rate (i.e. $1 - t_1$) benefits all individuals in the economy; whilst a reduction in the upper net-of-tax rate (i.e. $1 - t_2$) benefits only those for whom $y > Y$ and thus those with $n \in (\tilde{n}, \infty)$. Meanwhile, a *ceteris paribus* increase in the earnings threshold benefits both those bunched at the earnings threshold and those who earn above it. In particular, individuals with $n \in (\tilde{n}, \tilde{\tilde{n}}]$ can optimally increase their earnings; whilst individuals with $n \in (\tilde{\tilde{n}}, \infty)$ have less of their earnings taxed at the higher rate t_2.

Optimisation Problem. The government's optimisation problem is now described by:

$$\max_{t_1,t_2,Y,B,C} W = \theta u(B + C) + (1 - \theta) \int_0^\infty v(n, 1 - t_1, 1 - t_2, Y, B) f(n) dn$$

$$\begin{aligned} \text{s.t. } & B + \theta C \\ & = (1 - \theta) \left\{ t_1 \left\langle \int_0^{\tilde{n}} y^* f(n) dn + Y[1 - F(\tilde{n})] \right\rangle + t_2 \int_{\tilde{\tilde{n}}}^\infty (y^* - Y) f(n) dn \right\} - R \, , \\ & t_1 \in (0,1) \, , \; t_2 \in (0,1) \, , \; t_1 \leq t_2 \, , \; Y \geq 0 \, , \; B \geq 0 \, , \; C \geq 0 \, . \end{aligned} \tag{3.16}$$

As we are abstracting from income effects, an able individual's smvi is simply their marginal indirect utility of income. However, to be consistent with the notation from Section 3.2.2 we let $s = v_M$ be a productivity n individual's smvi and $\bar{s} = \int s f(n) dn$ be the average smvi over the able subpopulation.

Letting $\hat{t}_1$,$\hat{t}_2$,$\hat{Y}$,$\hat{B}$ and $\hat{C}$ denote the resulting optima, we can state the following:

Result 2:

(i) $\hat{C} > 0$ *and* $\hat{B} \geq 0$ *satisfy:*

$$\bar{s}\left(\hat{t}_1, \hat{t}_2, \hat{Y}, \hat{B}\right) \leq u'\left(\hat{B} + \hat{C}\right) = \hat{\lambda} \;\; ; \;\; \hat{B} \geq 0 \tag{3.17}$$

where the pair of inequalities hold with complementary slackness.

(ii) *For* $\mathcal{E}_i = \frac{(1-t_i)}{y^*}\frac{\partial y^*}{\partial(1-t_i)}$; $i \in \{1,2\}$, $\hat{t}_1$, $\hat{t}_2$ *and* $\hat{Y}$ *are characterised by:*

$$(t_1)\ :\ \frac{\hat{t}_1}{1-\hat{t}_1} = \frac{\delta\hat{Y} + \int_0^{\tilde{n}} (Y-y)\,(s-\hat{\lambda})f(n)dn}{\hat{\lambda}\int_0^{\tilde{n}} y\mathcal{E}_1 f(n)dn} \tag{3.18}$$

$$(t_2)\ :\ \frac{\hat{t}_2}{1-\hat{t}_2} = \frac{\int_{\tilde{\tilde{n}}}^{\infty}\left(y-\hat{Y}\right)(\hat{\lambda}-s)f(n)dn}{\hat{\lambda}\int_{\tilde{\tilde{n}}}^{\infty} y\mathcal{E}_2 f(n)dn} \tag{3.19}$$

$$(Y)\ :\ \int_{\tilde{n}}^{\tilde{\tilde{n}}}\left(\frac{\partial v}{\partial Y} + \hat{\lambda}\hat{t}_1\right)f(n)dn = \left(\hat{t}_2 - \hat{t}_1\right)\int_{\tilde{\tilde{n}}}^{\infty}\left(\hat{\lambda}-s\right)f(n)dn \tag{3.20}$$

Proof: See Appendix.

Result 2(i) parallels Result 1(i): a universal benefit will only be provided conditional on categorical transfers eliminating inequality in the average smvi; and there being resources left over. Result 2(ii) characterises the optimal tax parameters and is analogous to Apps et al. (2014). The important difference with these authors is that the presence of a dependent population changes how we write the expression for $\hat{t}_1$. In particular, whilst both the expressions for $\hat{t}_1$ and $\hat{t}_2$ have equity concerns in the numerator and efficiency concerns in the denominator, it is only the numerator of the former that contains δ; entering as δY. The intuition is seemingly that an increase in t_1 has no distortionary effect on the gross earnings of those with $n \in (\tilde{n}, \infty)$ and is therefore an effective tool to help reduce δ. This does, of course, come at the cost of imposing a higher tax rate on those with $n \in (0, \tilde{n})$ and thus those with lower productivities. Finally, (3.20) illustrates that Y should be set so as to equate (i) the marginal benefit of allowing individuals with $n \in (\tilde{n}, \tilde{\tilde{n}})$ to work more at rate t_1 (and in welfare units the associated increase in tax revenue); with (ii) the marginal cost of foregone tax revenue from those with $n \in (\tilde{\tilde{n}}, \infty)$, weighted by the positive term $\int_{\tilde{\tilde{n}}}^{\infty}(\lambda - s)f(n)dn$.

3.3.2 Numerical Results: Piecewise Linear Taxation

To provide examples where $\delta > 0$ at the piecewise taxation optimum, we once more turn to numerical methods.

In line with the preferences adopted in (3.10), let preferences take the frequently employed isoelastic form (see Atkinson, 1990; Saez, 2001):

$$u(x, H) = \log\left(x - \alpha\frac{H^{1+k}}{1+k}\right) \tag{3.21}$$

where in this setting $1/k$ is the constant elasticity of labour supply with respect to the net wage rate and α is a constant.[10] Low labour elasticities are observed empirically and we follow convention by setting $1/k = 0.25$ and thus $k = 4$. The literature typically adopts values of the labour elasticity between 0.1 and 1.

Productivities are Pareto distributed where $f(n) = \mu\underline{n}^{\mu}/n^{\mu+1}\ \forall\ n \geq \underline{n}$. The Pareto distribution captures well the upper tail of observed income distributions and its adoption in the more recent optimal tax literature has supported increasing marginal tax rates on higher earners (see Diamond, 1998). It would therefore seem appropriate for simulating progressive piecewise tax schedules.[11] To capture how the spread of abilities affects the results, we consider two alternative distributions: (i) $\underline{n} = 1$, $\mu = 4$; and (ii) $\underline{n} = 1.067$, $\mu = 5$, where $\underline{n}$ is adjusted so that the average productivity is 1.333 in both cases. The second distribution has a smaller spread of abilities than the first.

The remaining parameter choices are $\alpha = 8$; $\theta \in \{0.10, 0.15\}$ and $R \in \{0, 0.10, 0.15, 0.20\}$. Analogous to the numerical analysis in Section 3.2.3, the leisure preference parameter α is set so that the average worker has a labour supply of roughly 2/3. Again, the values of the unable subpopulation size θ seem sensible following statistics on real-world welfare programmes (Mcinnes, 2012). Finally, the choices of revenue requirement fall well within maximum tax revenue under the two considered distributions. The revenue maximising tax rate is given by $t = k/(1+k)$ and is thus 80% when $k = 4$.[12] Setting

---

[10]Under the preferences in (3.21) optimal labour supply is given by $H^* = [n(1-t)/\alpha]^{\frac{1}{k}}$. Letting $\omega = n(1-t)$ denote the net wage, the elasticity of labour supply with respect to the net wage rate is thus:

$$\frac{\omega}{H^*} \cdot \frac{\partial H^*}{\partial \omega} = \left(\omega^{\frac{k-1}{k}}\alpha^{\frac{1}{k}}\right) \cdot \left(\left(\frac{1}{k}\right)\omega^{\frac{1-k}{k}}\alpha^{-\frac{1}{k}}\right) = \frac{1}{k}$$

where the term within the first pair of braces is ω/H^*, whilst the term within the second pair of braces is $\partial H^*/\partial\omega$.

[11]Slemrod et al. (1994) demonstrate that extending the framework of Stern (1976) (a variant of which we applied in Section 3.2.3) to a two bracket piecewise setting gives rise to $t_1 > t_2$ in all cases. Apps et al. (2014) discuss how the assumption that productivities are lognormally distributed may bias results in favour of decreasing marginal tax rates.

[12]Individual earnings are given by $y^* = nH^* = (1-t)^{\frac{1}{k}} n^{\frac{1+k}{k}} \alpha^{-\frac{1}{k}}$. The revenue maximising tax rate is thus defined as:

$$t_L = \operatorname{Arg}\max_t t(1-t)^{\frac{1}{k}}\alpha^{-\frac{1}{k}} \int_0^{\infty} n^{\frac{1+k}{k}} dF(n) = \operatorname{Arg}\max_t t(1-t)^{\frac{1}{k}}$$

$\theta = 0.15$ we find that maximum revenue under the first distribution ($\mu = 4, \underline{n} = 1$) is 0.393; whilst maximum revenue under the second distribution ($\mu = 5, \underline{n} = 1.067$) is 0.391.

In the numerical results to be presented below it is useful to compare the piecewise optima with the flat tax optima. Intuitively, the flat tax is always available under the piecewise system (i.e. through setting $t_1 = t_2$) and so where it is not chosen welfare must be higher under the piecewise system. We once more let t_δ denote the critical flat tax that generates just enough tax revenue (net of any revenue requirement) to set $\delta = 0$ through categorical transfers. In this regard, the isoelastic preferences in (3.21) simplify the flat tax problem sufficiently that we can establish conditions where between-group inequality will persist at the optimum for the purely redistributive case with no revenue requirement. We discuss this in the below remark and subsequent result.

Remark: *Under the isoelastic preferences in (3.21) the explicit solution for t_δ when $R = 0$ is:*

$$t_\delta = \frac{\theta}{(1-\theta)(\frac{1+k}{k}) \int_0^\infty n^{\frac{1+k}{k}} dF(n) \cdot \int_0^\infty n^{-\frac{1+k}{k}} dF(n) + \theta} \tag{3.22}$$

which is unambiguously increasing in the unable subpopulation size θ.

Result 3: *Consider the pure targeting problem when $R = 0$:*

$$\begin{aligned} \max_t W &= \theta u(C(t), 1) + (1-\theta) \int_0^\infty v[n(1-t), 0] dF(n) \\ \text{where } C(t) &= \left(\frac{1-\theta}{\theta}\right) t \int_0^\infty y^*(n, 1-t) dF(n) \end{aligned} \tag{3.23}$$

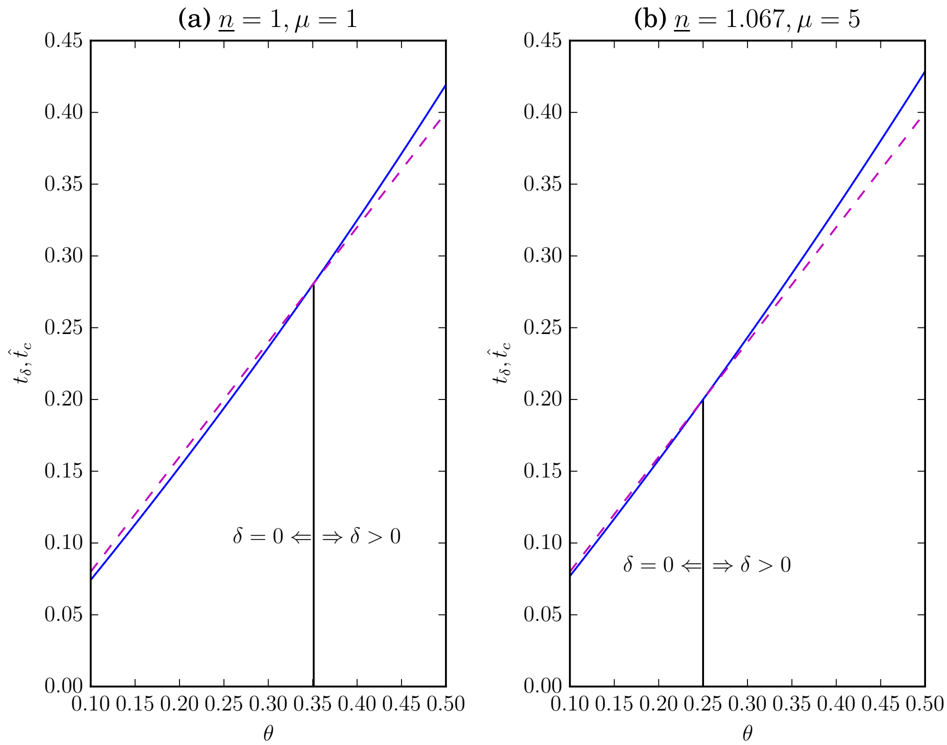
Then under the isoelastic preferences in (3.21) the optimal tax rate, denoted by $\hat{t}_c$, is thus given by:

$$\hat{t}_c = \theta \left(\frac{k}{1+k}\right) \tag{3.24}$$

which is increasing in both the size of the dependent population and in the inverse labour elasticity. A simple comparison between (3.22) and (3.24) illustrates that for θ

which yields the first order condition $1 - (1/k)t_L(1 - t_L)^{-1} = 0$.

Figure 3.3: Between-group inequality and θ when $R = 0$.



Notes: This figure plots the functions t_δ and $\hat{t}_c$ (as defined in (3.22) and (3.23)) over θ. As discussed in Result 3, there is a critical unable subpopulation size above which it is suboptimal to eliminate inequality in the average smvi between the unable and able subpopulations. In subplot (a) the critical value of θ is 0.35, whilst in subplot (b) the critical value of θ is 0.25.

large enough it will be suboptimal to eliminate inequality in the average smvi between the unable and able subpopulations. This is illustrated in Figure 3.3.

Proof: See Appendix

Table 3.2 displays the numerical results and its structure is described in the caption immediately below the table. The most immediate observation from Table 3.2 is that there are indeed cases under *both* the flat tax and piecewise linear tax schedules where $\delta > 0$ at the optimum. Moreover, there are a number of cases where $\delta = 0$ under the flat tax schedule, *but* $\delta > 0$ under the corresponding piecewise schedule. The intuition here is that the additional tools available to the government under the piecewise system allow it to lower the tax burden on individuals of lower productivity, which in turn acts to lower the average smvi over the able subpopulation (as illustrated by the observation

Table 3.2: Numerical Results: Piecewise Linear Income Taxation

	Flat Tax System							*Piecewise Linear Tax System*							
R	$\hat{t}$	$\hat{B}$	$\hat{C}$	δ	$\bar{s}$	G	t_δ	$\hat{t}_1$	$\hat{t}_2$	$\hat{Y}$	$\hat{B}$	$\hat{C}$	δ	$\bar{s}$	G
(a) $\underline{n} = 1$, $\mu = 4$, $\theta = 0.1$															
0	0.251	0.137	0.442	0	1.726	0.182	0.074	0.127	0.368	0.823	0.060	0.522	0	1.717	0.113
0.10	0.268	0.051	0.425	0	2.101	0.093	0.200	0.183	0.394	0.849	0	0.476	**0.017**	2.085	0.048
0.15	0.279	0.008	0.416	0	2.359	0.050	0.267	0.257	0.422	0.935	0	0.418	**0.060**	2.334	0.042
0.20	0.336	0	0.362	**0.089**	2.673	0.036	0.338	0.331	0.452	1.073	0	0.357	**0.143**	2.654	0.036
(b) $\underline{n} = 1$, $\mu = 4$, $\theta = 0.15$															
0	0.255	0.109	0.437	0	1.830	0.175	0.113	0.129	0.372	0.814	0.030	0.519	0	1.821	0.108
0.10	0.275	0.024	0.419	0	2.259	0.086	0.242	0.230	0.411	0.900	0	0.439	**0.041**	2.236	0.066
0.15	0.310	0	0.385	**0.043**	2.548	0.058	0.312	0.303	0.440	1.012	0	0.380	**0.105**	2.523	0.057
0.20	0.381	0	0.322	**0.193**	2.915	0.048	0.386	0.378	0.471	1.216	0	0.319	**0.232**	2.902	0.048
(c) $\underline{n} = 1.067$, $\mu = 5$, $\theta = 0.10$															
0	0.185	0.085	0.503	0	1.701	0.136	0.077	0.093	0.292	0.842	0.024	0.566	0	1.695	0.081
0.10	0.203	0	0.485	**0.003**	2.061	0.048	0.203	0.197	0.341	0.957	0	0.479	**0.041**	2.045	0.048
0.15	0.269	0	0.422	**0.074**	2.295	0.042	0.271	0.266	0.374	1.092	0	0.419	**0.098**	2.287	0.042
0.20	0.338	0	0.358	**0.194**	2.599	0.036	0.342	0.337	0.410	1.346	0	0.357	**0.206**	2.596	0.036
(d) $\underline{n} = 1.067$, $\mu = 5$, $\theta = 0.15$															
0	0.189	0.057	0.498	0	1.804	0.131	0.117	0.104	0.300	0.846	0	0.556	**0.002**	1.797	0.083
0.10	0.245	0	0.445	**0.045**	2.204	0.067	0.247	0.241	0.362	1.035	0	0.441	**0.073**	2.194	0.066
0.15	0.312	0	0.382	**0.141**	2.475	0.057	0.316	0.310	0.396	1.227	0	0.381	**0.157**	2.470	0.057
0.20	0.383	0	0.318	**0.310**	2.835	0.048	0.391	0.383	0.433	1.662	0	0.318	**0.315**	2.834	0.048

Notes: Column 1 is the revenue requirement; columns 2-4 are the flat tax optima; columns 5-7 give the resulting values of δ, $\bar{s}$ and G (tax revenue net of R); column 8 is the critical tax rate t_δ generating just enough tax revenue (net of R) to set $\delta = 0$; columns 9-13 are the piecewise optima; whilst columns 14-16 give the resulting values of δ, $\bar{s}$ and G. Values $\leq 10^{-7}$ reported as 0.

that $\bar{s}$ takes a lower value at the piecewise optimum than at the flat tax optimum in all cases). Even though the unable receive a similar benefit income to the flat tax case (and thus have a similar smvi) δ is necessarily higher due to the fall in $\bar{s}$. In these cases net tax revenue is also lower under the piecewise schedule. Note that where $\delta > 0$ under both types of system, the unable tend to receive a slightly lower level of categorical support under the piecewise system than they do under the flat tax system. Finally, given that the flat tax is always available but not chosen, welfare is higher under the piecewise system (to save on space this is omitted from the table).

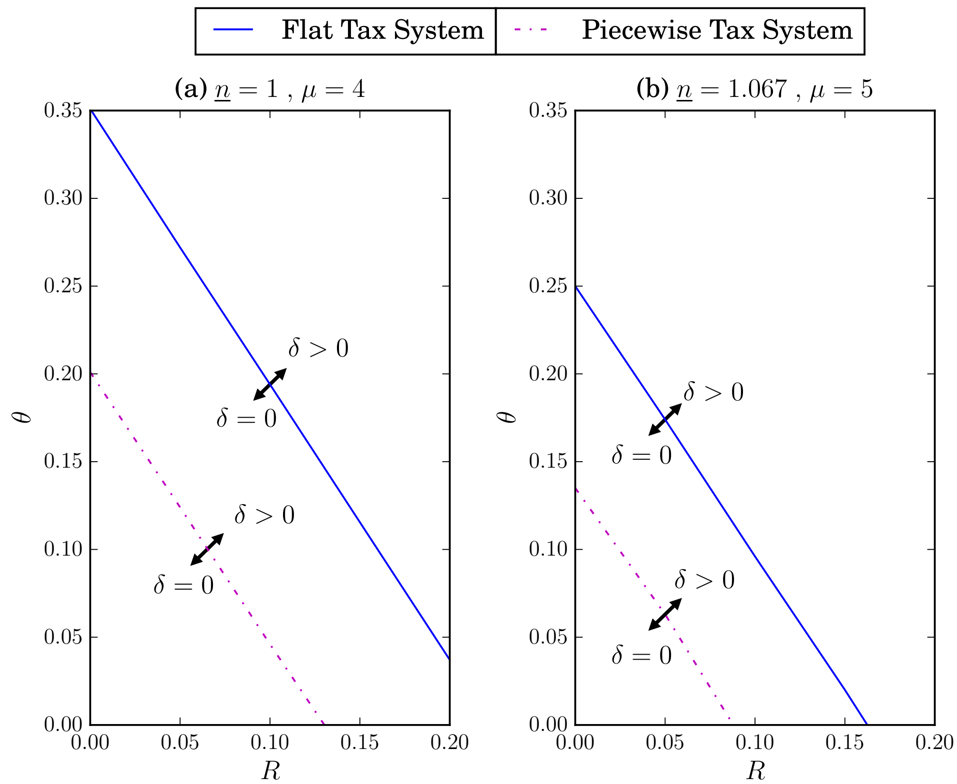
Table 3.2 illustrates that an increase in θ or R increases the number of cases where $\delta > 0$. Further, for cases where $\delta > 0$ prior to the increase, we observe that an increase in either parameter increases the size of δ. With respect to the piecewise system, we note that an increase in R (i) raises the average smvi of the able (because the tax parameters $\hat{t}_1, \hat{t}_2$ and $\hat{Y}$ rise whilst $\hat{B}$ either falls or remains at zero); and (ii) raises the smvi of the unable (because $\hat{C}$ falls). Given that δ rises in all cases the latter effect must dominate the first. A parallel argument holds in the flat tax case. Interestingly, a mean-preserving reduction in the spread of abilities *increases* (i) the number of cases where $\delta > 0$; and (ii) the magnitude of δ for cases where $\delta > 0$ already. This is best explained in terms of the flat tax system. For any given value of t, both $\bar{s}$ and tax revenue fall. Consequently, the critical tax rate t_δ at which between-group inequality is eliminated rises.[13] Figure 3.4 summarises the conditions where $\delta > 0$ in (R, θ) space.

3.4 Concluding Remarks

The analysis of categorical transfers in optimal linear and piecewise-linear income tax frameworks will depend on whether categorical transfers are financed so as to eliminate between-group inequality in the average net smvi at the optimum. If one simply assumes this to hold then the resulting optimal tax expressions can be written as in the uni-dimensional model where individuals differ only in productivity. The purpose of this chapter has been to demonstrate that this may not always be a good assumption. For example, the presence of a sufficiently large unable subpopulation dependent on

[13]Note that t_δ would still rise even if we did not adjust $\underline{n}$ so as to keep mean productivity constant (for any given t: $\bar{s}$ may increase but tax revenue falls).

Figure 3.4: Between-Group Inequality at the Optimum in (R, θ) Space



Notes: For optimum flat tax and optimum piecewise tax schedules, this figure illustrates the cases in (R, θ) space where between group inequality in the average smvi is not eliminated. Notice that when $R = 0$ the critical values of θ above which between-group inequality will persist in the flat tax optimum correspond to those established in Figure 3.3.

categorical transfers, coupled with spending commitments outside welfare, may render it suboptimal to spend on categorical transfers up to this point because doing so would be too harmful to the non-dependent working subpopulation. Numerical examples have illustrated cases where this arises.

This chapter has demonstrated that optimal income tax expressions can be written more generally to allow for cases where it is indeed suboptimal to eliminate between-group inequality in the average net smvi. In these cases the equity considerations in the numerator of optimal tax expressions are composed of both between-group and within-group terms. Further, if one employs preferences with a constant labour elasticity (see Atkinson, 1990; Saez, 2001) it is also possible to derive analytical results for the conditions under which between-group inequality will persist under the flat tax optimum. Specifically, in the special case where taxation is purely redistributive the optimal flat tax will fall below that required to eliminate inequality in the average smvi whenever the unable subpopulation exceeds a critical level. How these optimal tax expressions change when categorical transfers are administered with classification errors - and thus where it may not be possible to eliminate inequality in the average net smvi - warrants investigation.

Appendix A Linear Income Taxation

Proof of Result 1

From the optimisation problem described in (3.2), the first-order conditions (henceforth FOCs) characterising the optimal benefits $\hat{B}$ and $\hat{C}$ are:

$$(B) : \theta u_x\left(\hat{B}+\hat{C},1\right)+(1-\theta)\bar{s}\left(\hat{t},\hat{B},\hat{\lambda}\right)\leq\hat{\lambda}\ ;\ \hat{B}\geq 0 \tag{A.1}$$

$$(C) : u_x\left(\hat{B}+\hat{C},1\right)\leq\hat{\lambda}\ ;\ \hat{C}\geq 0 \tag{A.2}$$

where the pairs of inequalities hold with complementary slackness.

We test the following two hyotheses:

(i) $\hat{B}>0$ **,** $\hat{C}=0$ **(Pure Universal System)**

Setting $\hat{C}=0$ in both (A.1) and (A.2) gives:

$$\begin{aligned}\theta u_x(B,1)+(1-\theta)\bar{s}(t,B,\lambda)&=\lambda\\ u_x(C,1)&\leq\lambda\end{aligned}$$

Taken together, these equations imply that $\bar{s}(t,B,\lambda)\geq u_x(B,1)$. This is clearly a contradiction because the average net social marginal value of income (smvi) of the able cannot exceed that of the unable when receiving the same benefit income. The assertion that $\hat{C}=0$ is therefore false and it must instead hold that $\hat{C}>0$.

(ii) $\hat{B}=0$**,** $\hat{C}>0$ **(Pure Targeted System)**

Alternatively, setting $\hat{B}=0$ in the FOCs in (A.1) and (A.2) gives:

$$\begin{aligned}\theta u_x(C,1)+(1-\theta)\bar{s}(t,0,\lambda)&\leq\lambda\\ u_x(C,1)&=\lambda\end{aligned}$$

Combining these equations gives the condition:

$$\bar{s}(t,0,\lambda)\leq u_x(C,1) \tag{A.3}$$

This simply states that it will not be optimal to provide a universal benefit if, at the optimum, categorical spending does not eliminate between-group inequality in the average net smvi. Intuitively, it is optimal to expend resources on the most needy in society, and this is the unable so long as the aforementioned between-group inequality persists.

The FOC for the interior tax rate optimum is given by:

$$(t): \int_0^\infty \left\{ -nv_\omega + \hat{\lambda} \left\langle y^* - \hat{t} \cdot \frac{\partial y^*}{\partial(1-t)} \right\rangle \right\} f(n)dn = 0 \tag{A.4}$$

where $\partial y^*/\partial(1-t) = n^2 H_\omega^*$. Note that because $y^* = 0 \ \forall \ n \leq \bar{n}$ we could equivalently use $\bar{n}$ as the lower integral limit.

By standard methods (Atkinson and Stiglitz, 1980) we use Roy's identity ($v_\omega = v_M(y^*/n)$) and the Slutsky-Hicks equation [14] ($\frac{\partial y^*}{\partial(1-t)} = \frac{\partial y^c}{\partial(1-t)} + y_M^* y^*$) to write (A.4) as:

$$\int_0^\infty y^* \left(1 - \frac{s}{\hat{\lambda}}\right) f(n)dn = \hat{t} \int_0^\infty \frac{\partial y^c}{\partial(1-t)} f(n)dn$$

Letting $\mathcal{E}^c = \dfrac{(1-t)}{y} \dfrac{\partial y^c}{\partial(1-t)}$ be the compensated elasticity of earnings with respect to

[14] The expenditure minimisation problem is:

$$\min_{x,y} x - y(1-t) \ \text{ s.t. } \ u(x, 1 - y/n)$$

yielding the FOCs:

$$\begin{aligned} (x)&: \ 1 - \gamma u_x = 0 \\ (y)&: \ -(1-t) + \gamma u_l/n = 0 \\ (\gamma)&: \ u - \bar{u} = 0 \end{aligned}$$

Let $x^c(n, 1-t, \bar{u})$, $y^c(n, 1-t, u)$ and $\gamma^c(n, 1-t, \gamma)$ denote the optimal 'compensated' choices and further, let the expenditure function be $E(n, 1-t, \bar{u}) = x^c - y^c(1-t)$. By the envelope theorem:

$$\frac{\partial E}{\partial(1-t)} = \frac{\partial x^c}{\partial(1-t)}(1 - \gamma^c u_x) + \frac{\partial y^c}{\partial(1-t)}\left[-(1-t) + \gamma^c \frac{u_l}{n}\right] - y^c - \frac{\partial \gamma^c}{\partial(1-t)}(u - \bar{u}) = -y^c$$

Next, differentiating the identity $y^c(n, 1-t, \bar{u}) \equiv y^*(n, 1-t, E(n, 1-t, \bar{u}))$ w.r.t. $(1-t)$ gives:

$$\frac{\partial y^c}{\partial(1-t)} = \frac{\partial y^*}{\partial(1-t)} + \frac{\partial y^*}{\partial M}\frac{\partial E}{\partial(1-t)} = \frac{\partial y^*}{\partial(1-t)} - y^* \frac{\partial y^*}{\partial M}$$

the net of tax rate we obtain:

$$\frac{\hat{t}}{1-\hat{t}} = \frac{\int_0^\infty y^* \left(\hat{\lambda} - s\right) f(n)dn}{\hat{\lambda} \int_0^\infty y\mathcal{E}^c f(n)dn} \tag{A.5}$$

The numerator will reduce to the negative of the covariance between gross earnings (y^*) and the net smvi (s) **only if** $\hat{\lambda} = \bar{s}$, and thus if between-group inequality in the average net smvi is eliminated through categorical spending at the optimum. Indeed, to see this note that if $\hat{\lambda} = \bar{s}$ then $\int y^*(\hat{\lambda} - s)f(n)dn = \int y^* f(n)dn \cdot \bar{s} - \int y^* s f(n)dn = \bar{y}^*\bar{s} - \int y^* s f(n)dn = -\text{Cov}(y^*, s)$.

Letting $\delta = (\lambda - \bar{s})$ denote between-group inequality in the average net smvi and $r = y^*/\bar{y}^*$ relative income; the numerator of (A.5) can be written as:

$$\begin{aligned}
\int_0^\infty y^*(\hat{\lambda} - s)f(n)dn &= \int_0^\infty y^* \left[(\hat{\lambda} - \bar{s}) + \bar{s}\left(1 - \frac{s}{\bar{s}}\right)\right] f(n)dn \\
&= \delta\bar{y}^* + \bar{s}\int_0^\infty y^* \left(1 - \frac{s}{\bar{s}}\right) f(n)dn \\
&= \delta\bar{y}^* + \bar{s}\left\{\int_0^\infty y^* f(n)dn \int_0^\infty \left(\frac{s}{\bar{s}}\right) f(n)dn - \int_0^\infty y^* \left(\frac{s}{\bar{s}}\right) f(n)dn\right\} \\
&= \delta\bar{y}^* + \left\{\int_0^\infty y^* f(n)dn \int_0^\infty s f(n)dn - \int_0^\infty y^* s f(n)dn\right\} \\
&= \delta\bar{y} - \text{Cov}(y^*, s) \qquad \text{(A.6)}
\end{aligned}$$

Substituting (A.6) into the optimal tax expression in (A.5) and then dividing both the numerator and denominator by $\bar{y}$ gives the optimal tax expression in the main text.

Appendix B Piecewise Linear Income Taxation

Properties of Preferences

Suppose preferences are of the form $U(x, y/n) = u[x - g(y/n)]$; where $u' > 0$, $u'' < 0$; $g' > 0$ and $g'' > 0$. This is simply a concave transformation of quasi-linear preferences.[15] Let the function $x(y)$ satisfy the indifference condition $U(x(y), y/n) = \bar{u}$. Differentiating this expression with respect to y yields $x'(y) = g'(y/n)/n > 0$ and $x''(y) = g''(y/n)/n^2 > 0$, such that indifference curves in (y, x) space are convex-increasing in y. From this it immediately follows that:

$$\frac{dx'(y)}{dn} = \frac{-\left(\frac{y}{n}\right) g''\left(\frac{y}{n}\right) - g'\left(\frac{y}{n}\right)}{n^2} < 0 \tag{B.1}$$

such the slope of indifference curves in (y, x)-space is decreasing in n. This is illustrated in Figure (3.2) in the main text.

Given the piecewise tax system described in (3.9) in the main text, it follows from (B.1) that there will be (i) a critical productivity $\tilde{n}(1 - t_1, Y)$ at which an individual facing the marginal tax rate t_1 has an optimal *unconstrained* gross earnings of Y; and (ii) a critical productivity $\tilde{\tilde{n}}(1 - t_2, Y)$ at which an individual facing the marginal tax rate t_2 has optimal *unconstrained earnings* of Y. Formally:

$$\tilde{n}(1 - t_1) = g'\left(\frac{Y}{\tilde{n}}\right) \quad ; \quad \tilde{\tilde{n}}(1 - t_2) = g'\left(\frac{Y}{\tilde{\tilde{n}}}\right)$$

Differentiating w.r.t. the net of tax rates gives:

$$\begin{aligned} \frac{\partial \tilde{n}}{\partial(1 - t_1)} &= \frac{-(\tilde{n})^2}{(1 - t_1)\tilde{n} + g''(Y/\tilde{n}) \cdot (Y/\tilde{n})} < 0 \\ \frac{\partial \tilde{\tilde{n}}}{\partial(1 - t_2)} &= \frac{-(\tilde{\tilde{n}})^2}{(1 - t_2)\tilde{\tilde{n}} + g''(Y/\tilde{\tilde{n}}) \cdot (Y/\tilde{\tilde{n}})} < 0 \end{aligned} \tag{B.2}$$

Alternatively, differentiating w.r.t. the earnings threshold Y gives:

$$\begin{aligned} \frac{\partial \tilde{n}}{\partial Y} &= \frac{g''(Y/\tilde{n})}{(1 - t_1)\tilde{n} + g''(Y/\tilde{n}) \cdot (Y/\tilde{n})} > 0 \\ \frac{\partial \tilde{\tilde{n}}}{\partial Y} &= \frac{g''(Y/\tilde{\tilde{n}})}{(1 - t_2)\tilde{\tilde{n}} + g''(Y/\tilde{\tilde{n}}) \cdot (Y/\tilde{\tilde{n}})} > 0 \end{aligned} \tag{B.3}$$

[15] Alternatively, we could write $U(x, l) = u[x - g(1 - l)]$. It is straightforward to verify that:

$$U_l = u' \cdot \alpha g' > 0 \quad , \quad U_{ll} = u'' \cdot [\alpha g'(1 - l)]^2 - u' \alpha g'(1 - l) < 0$$

Properties of indirect utility. Substituting the optimal earnings function y^* (as defined in (3.12) in the main text) into preferences gives the indirect utility function $v(n, 1-t_1, 1-t_2, Y, M)$. Differentiating v w.r.t. the net of tax rates and the earnings threshold thus gives:

$$\frac{\partial v}{\partial(1-t_1)} = \begin{cases} u' \cdot \left\langle y^* + \frac{\partial y^*}{\partial(1-t_1)} \underbrace{[(1-t_1) - g'(y^*/n)/n]}_{=0} \right\rangle = u'y^* & : n \in (0, \tilde{n}] \\ u' \cdot Y & : n \in (\tilde{n}, \infty) \end{cases}$$

whilst

$$\frac{\partial v}{\partial(1-t_2)} = \begin{cases} 0 & : n \in (0, \tilde{\tilde{n}}] \\ u' \cdot \left\langle (y^* - Y) + \frac{\partial y^*}{\partial(1-t_2)} \underbrace{[(1-t_2) - g'/n]}_{=0} \right\rangle = u'(y^* - Y) & : n \in (\tilde{\tilde{n}}, \infty) \end{cases}$$

and finally:

$$\frac{\partial v}{\partial Y} = \begin{cases} 0 & : n \in (0, \tilde{n}) \\ u' \cdot [(1-t_1) - g'(Y/n)/n] > 0 & : n \in (\tilde{n}, \tilde{\tilde{n}}) \\ u' \left\langle (t_2 - t_1) + \frac{\partial y}{\partial Y} \underbrace{[(1-t_2) - g'/n]}_{=0} \right\rangle = u'(t_2 - t_1) & : n \in (\tilde{\tilde{n}}, \infty) \end{cases}$$

Given that $v_M = u'$ it follows that $\partial v/\partial(1-t_1) = v_M \min(y^*, Y)\ \forall n$; whilst $\partial v/\partial(1-t_2) = v_M(y^* - Y)\ \forall\ n \in (\tilde{\tilde{n}}, \infty)$.

Proof of Result 2

Optimal benefits $(\hat{B}, \hat{C})$. The FOCs characterising the optimal benefit levels are:

$$(B) \ : \quad \theta u' \left(\hat{B} + \hat{C}\right) + (1-\theta)\bar{s}\left(\hat{t}_1, \hat{t}_2, \hat{Y}, \hat{B}\right) \leq \hat{\lambda}\ ;\ \hat{B} \geq 0 \tag{B.4}$$

$$(C) \ : \quad u' \left(\hat{B} + \hat{C}\right) \leq \hat{\lambda}\ ;\ \hat{C} \geq 0 \tag{B.5}$$

where the pairs of inequalities hold with complementary slackness. Suppose that $\hat{C} = 0$ but $\hat{B} > 0$: then (B.4) and (B.5) together imply that $u'(B) = \bar{s}(t_1, t_2, Y, B)$ which is a contradiction given that the smvi is falling in n. It must therefore hold that $\hat{C} > 0$ at the optimum.

Optimal lower tax rate ($\hat{t_1}$). The FOC characterising the optimal lower tax rate (t_1) is[16]:

$$(t_1)\ :\ \int_0^{\infty} -\frac{\partial v}{\partial(1-t_1)} f(n)dn + \hat{\lambda}\left\langle \int_0^{\tilde{n}} \left[y^* - \hat{t}_1 \frac{\partial y^*}{\partial(1-t_1)} \right] f(n)dn + \hat{Y}[1 - F(\tilde{n})] \right\rangle = 0$$

where the only individuals who adjust their labour supply in response to an increase t_1 are those with $n \in (0, \tilde{n}]$. Taking this into account and using the fact that $\partial v / \partial(1-t_1) = v_M \min(y, Y)\ \forall\ n$ we can write this as:

$$\int_0^{\tilde{n}} y \left(1 - \frac{v_M}{\lambda}\right) f(n)dn + \int_{\tilde{n}}^{\infty} \hat{Y} \left(1 - \frac{v_M}{\hat{\lambda}}\right) f(n)dn = \hat{t}_1 \int^{\tilde{n}} \frac{\partial y^*}{\partial(1-t_1)} f(n)dn$$

Letting $\mathcal{E}_i = \dfrac{(1-t_i)}{y^*} \dfrac{\partial y^*}{\partial(1-t_i)}$ $i \in \{1, 2\}$ we obtain:

$$\frac{\hat{t}_1}{1-\hat{t}_1} = \frac{\int^{\tilde{n}} y^* \left(\hat{\lambda} - v_M\right) f(n)dn + \int_{\tilde{n}}^{\infty} Y \left(\hat{\lambda} - v_M\right) f(n)dn}{\hat{\lambda} \int^{\tilde{n}} y^* \mathcal{E}_1 f(n)dn} \tag{B.6}$$

$$= \frac{\int^{\tilde{n}} (y^* - Y) \left(\hat{\lambda} - v_M\right) f(n)dn + \int Y \left(\hat{\lambda} - v_M\right) f(n)dn}{\hat{\lambda} \int^{\tilde{n}} y^* \mathcal{E}_1 f(n)dn}$$

$$= \frac{Y(\lambda - \bar{s}) + \int^{\tilde{n}} (y^* - Y)(\lambda - s) f(n)dn)}{\hat{\lambda} \int^{\tilde{n}} y^* \mathcal{E}_1 f(n)dn}$$

$$= \frac{\delta Y + \int^{\tilde{n}} (Y - y^*) \left(s - \hat{\lambda}\right) f(n)dn}{\hat{\lambda} \int^{\tilde{n}} y^* \mathcal{E}_1 f(n)dn} \tag{B.7}$$

where to progress from (B.6) to (B.7) we simply added and subtracted $\int^{\tilde{n}} Y(\hat{\lambda} - v_M) f(n)dn$ and then subsequently used the definitions $s = v_M$ and $\bar{s} = \int s f(n)dn$.

[16] In the unidimensional model where individuals differ only in their underlying productivity, Apps et al (2013) add and subtract $\int^{\tilde{n}} Y(\lambda - s)$ and use the property that $\lambda = \bar{s}$ at the optimum. Of course, in our setting with categorical transfers this may not be the case.

Optimal upper tax rate ($\hat{t}_2$). Next, the FOC charactering the upper bracket tax rate (t_2) is:

$$(t_2) \ : \quad \int_{\tilde{\tilde{n}}}^{\infty} -\frac{\partial v}{\partial(1-t_2)} + \hat{\lambda}\left\langle \left(y^* - \hat{Y}\right) - \hat{t}_2\frac{\partial y^*}{\partial(1-t_2)} \right\rangle f(n)dn = 0 \tag{B.8}$$

where of course $\partial y^*/\partial(1-t_2) = 0 \ \forall n \in (0, \tilde{\tilde{n}})$. Using the fact that $\partial v/\partial(1-t_2) = v_M(y-Y)$ this becomes;

$$\int_{\tilde{\tilde{n}}}^{\infty} \left(y^* - \hat{Y}\right)\left(1 - \frac{v_M}{\hat{\lambda}}\right) f(n)dn = \hat{t}_2 \int_{\tilde{\tilde{n}}}^{\infty} \frac{\partial y}{\partial(1-t_2)} f(n)dn \tag{B.9}$$

From which we obtain:

$$\frac{\hat{t}_2}{1-\hat{t}_2} = \frac{\int_{\tilde{\tilde{n}}}^{\infty} \left(y - \hat{Y}\right)\left(\hat{\lambda} - v_M\right) f(n)dn}{\hat{\lambda}\int_{\tilde{\tilde{n}}}^{\infty} y\mathcal{E}_2 f(n)dn} \tag{B.10}$$

Optimal earnings threshold ($\hat{Y}$). Finally, the FOC with respect to the earnings threshold (Y) is:

$$(Y) \ : \quad \int_{\tilde{n}}^{\infty} \frac{\partial v}{\partial Y} f(n)dn + \hat{\lambda}\left\langle \hat{t}_1[1-F(\tilde{n})] + \hat{t}_2[1-F(\tilde{\tilde{n}})] \right\rangle \tag{B.11}$$

Using the fact that $\partial v/\partial Y = -(t_2 - t_1)v_M \ \forall n \in [\tilde{\tilde{n}}, \infty)$ this becomes:

$$\begin{aligned}
&\int_{\tilde{n}}^{\tilde{\tilde{n}}} \frac{\partial v}{\partial Y} f(n)dn + (\hat{t}_2 - \hat{t}_1)\int_{\tilde{\tilde{n}}}^{\infty} v_M f(n)dn = \hat{\lambda}\left\{(\hat{t}_2 - \hat{t}_1) + \hat{t}_1 F(\tilde{n}) - \hat{t}_2 F(\tilde{\tilde{n}})\right\} \\
\Rightarrow &\int_{\tilde{n}}^{\tilde{\tilde{n}}} \frac{\partial v}{\partial Y} f(n)dn + (\hat{t}_2 - \hat{t}_1)\int_{\tilde{\tilde{n}}}^{\infty} (v_M - \hat{\lambda}) f(n)dn = \hat{\lambda}\hat{t}_1\left[F(\tilde{n}) - F(\tilde{\tilde{n}})\right]
\end{aligned}$$

and thus:

$$\int_{\tilde{n}}^{\tilde{\tilde{n}}} \left(\frac{\partial v}{\partial Y} + \hat{\lambda}\hat{t}_1\right) f(n)dn = (\hat{t}_2 - \hat{t}_1)\int_{\tilde{\tilde{n}}}^{\infty} (\hat{\lambda} - v_M) f(n)dn \tag{B.12}$$

Appendix C Numerical Simulations with Isoelastic Preferences

Properties of preferences with constant labour elasticity

Under the isoelastic preferences $u(x, H) = \log[x - \alpha H^{1+k}/(1+k)]$ specified in (3.21) in the main text, the optimal labour supply function is defined as:

$$H^* = \arg\max_{H \in [0,1]} \log\left(nH(1-t) + M - \alpha\frac{H^{1+k}}{1+k}\right)$$

This yields the explicit labour supply and optimal earnings functions:

$$\begin{aligned} H^*[n(1-t)] &= \left[\frac{n(1-t)}{\alpha}\right]^{\frac{1}{k}} \\ y^*(n, 1-t) &= nH^* = \alpha^{-\frac{1}{k}}(1-t)^{\frac{1}{k}} n^{\frac{1+k}{k}} \end{aligned} \tag{C.1}$$

from which we can immediately see that $\dfrac{\partial y^*}{\partial n} > 0$ and $\dfrac{\partial y^*}{\partial(1-t)} > 0$. We can also directly establish from (C.1) that the indirect utility function is given by:

$$\begin{aligned} v(n, 1-t, M) &= \log\left[(1-t)y^* + M - \alpha\frac{(y^*/n)^{1+k}}{1+k}\right] \\ &= \log\left[n^{\frac{1+k}{k}}(1-t)^{\frac{1}{k}}\alpha^{-\frac{1}{k}}\left(\frac{k}{1+k}\right) + M\right] \end{aligned} \tag{C.2}$$

Agent Monotonicity. Let the function $x(y)$ satisfy the indifference condition $u\left(x(y), y/n\right) = \bar{u}$ and thus:

$$\begin{aligned} x(y) &= e^{\bar{u}} + \alpha\left(\frac{(y/n)^{1+k}}{1+k}\right) \\ \Rightarrow\ x'(y) &= \alpha\left(\frac{y}{n}\right)^k/n \\ \Rightarrow\ \frac{dx'(y)}{dn} &= -\alpha\left(\frac{y}{n}\right)^k(k+1)/n < 0 \end{aligned} \tag{C.3}$$

So indifference curves in (y, x) space are (i) convex-increasing in y *but* (ii) flatter in n (i.e. agent monotonicity).

Piecewise Tax System. Consider a piecewise system with two tax rates $t_1 \leq t_2$ and earnings threshold Y generating an individual budget constraint:

$$x = \begin{cases} y(1-t_1) + M & : y \leq Y \\ Y(t_2 - t_1) + y(1-t_2) + M & : y > Y \end{cases}$$

Consider an individual facing the portion of the budget constraint for which $y \leq Y$. From (C.1) and (C.3) it is straightforward to establish that there will be critical productivity $\tilde{n}$ at which an individual will choose to earn *exactly* Y. Formally:

$$\tilde{n} = \left[Y^k \left(\frac{\alpha}{1-t_1} \right) \right]^{\frac{1}{1+k}} \tag{C.4}$$

Next, consider an individual facing the upper part of the budget constraint with marginal tax rate t_2. Once more, from (C.1) and (C.3) we can establish that there is a critical productivity $\tilde{\tilde{n}}$ at which an individual will choose to earn *exactly* Y. Formally:

$$\tilde{\tilde{n}} = \left[Y^k \left(\frac{\alpha}{1-t_2} \right) \right]^{\frac{1}{1+k}} \tag{C.5}$$

where for $t_1 < t_2 : \tilde{n} < \tilde{\tilde{n}}$.

Note that:

$$\frac{\partial \tilde{n}}{\partial t_1} > 0 \; ; \; \frac{\partial \tilde{n}}{\partial Y} > 0 \; ; \; \frac{\partial \tilde{\tilde{n}}}{\partial t_2} > 0 \; ; \; \frac{\partial \tilde{\tilde{n}}}{\partial Y} > 0$$

Proof of Result 3

Optimal Tax Rate under pure targeting scheme (when $R = 0$). Under the isoelastic preferences in (3.21) the optimisation problem in (3.23) can be written as:

$$\begin{aligned} \max_{t \in (0,1)} \quad & \theta \log \left[\left(\frac{1-\theta}{\theta} \right) t(1-t)^{\frac{1}{k}} \alpha^{-\frac{1}{k}} \int_0^\infty n^{\frac{1+k}{k}} dF(n) \right] \\ & + (1-\theta) \int_0^\infty \log \left[(1-t)^{\frac{1+k}{k}} \alpha^{-\frac{1}{k}} \left(\frac{k}{1+k} \right) n^{\frac{1+k}{k}} \right] dF(n) \end{aligned} \tag{C.6}$$

This yields the FOC for the optimal tax rate ($\hat{t}_c$):

$$\frac{\theta\left[(1-\hat{t}_c)^{\frac{1}{k}}-(\frac{1}{k})\hat{t}_c(1-\hat{t}_c)^{\frac{1-k}{k}}\right]}{\hat{t}_c(1-\hat{t}_c)^{\frac{1}{k}}}=(1-\theta)\frac{(\frac{1+k}{k})}{(1-\hat{t}_c)}$$

$$\Rightarrow\theta\left[1-\left(\frac{1}{k}\right)\hat{t}_c(1-\hat{t}_c)^{-1}\right]=(1-\theta)\left(\frac{1+k}{k}\right)\hat{t}_c(1-\hat{t}_c)^{-1}$$

$$\Rightarrow\theta\left[(1-\hat{t}_c)-\left(\frac{1}{k}\right)t_c\right]=(1-\theta)\left(\frac{1+k}{k}\right)\hat{t}_c$$

and thus $\hat{t}_c=\theta(k/(1+k))$.

The critical tax rate t_δ (when $R=0$). When $R=0$ the elimination of between-group inequality in the average smvi requires:

$$\left[\left(\frac{1-\theta}{\theta}\right)t_\delta(1-t_\delta)^{\frac{1}{k}}\alpha^{-\frac{1}{k}}\int_0^\infty n^{\frac{1+k}{k}}dF(n)\right]^{-1}$$

$$=\left[(1-t_\delta)^{\frac{1+k}{k}}\alpha^{-\frac{1}{k}}\left(\frac{k}{1+k}\right)\right]^{-1}\cdot\int_0^\infty n^{-\frac{1+k}{k}}dF(n)$$

and thus:

$$\theta\left(\frac{k}{1+k}\right)=t_\delta\left[(1-\theta)\int_0^\infty n^{-\frac{1+k}{k}}dF(n)\int_0^\infty n^{\frac{1+k}{k}}dF(n)+\theta\left(\frac{k}{1+k}\right)\right]$$

Simple manipulation then yields:

$$t_\delta=\theta\left[(1-\theta)\left(\frac{1+k}{k}\right)\int_0^\infty n^{\frac{1+k}{k}}dF(n)\int_0^\infty n^{-\frac{1+k}{k}}dF(n)+\theta\right]^{-1}$$

Numerical Code for Flat Tax/Piecewise Analysis with Isoelastic Preferences

```
from __future__ import division
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize
from numpy import vectorize
from scipy.stats import pareto
from scipy.integrate import quad
from scipy.optimize import newton

#Python Code run using Enthought Canopy 1.5.1 (64bit)
    on MacBook Pro 2.3GHz Intel Core i7

#LINEAR AND PIECEWISE LINEAR TAXATION WITH ISOELASTIC
    PREFERENCES
#----------------------------------------------------------

#In what follows: 'n' is individual productivity; 't'
    is the flat tax rate; 't1' and 't2' are the lower
    and upper tax rates from the piecewise system;
    'ycheck' is the earnings threshold from the
    piecewise system; 'k' is the inverse of the
    elasticity of labour supply with respect to the net
    wage rate; 'a' is a constant from the isoelastic
    preferences.

#1) PURE FLAT TAX ANALYSIS
#..........................................................

#The optimal earnings function of a productivity n
    individual is:
def yopt(n,t,a,k):
    return ((n*(1-t)/a)**(1/k))*n

#The resulting indirect utility function is thus:
def vopt(n,t,m,a,k):
    return np.log(yopt(n,t,a,k)*(1-t)+m-
    a*(((yopt(n,t,a,k)/n)**(1+k))/(1+k)))

#Multiplying the optimal earnings function by the
    productivity distribution pdf gives:
def yoptpdf(n,t,a,k,mu,minn):
    return
        yopt(n,t,a,k)*pareto.pdf(n,mu,loc=0,scale=minn)

#Similarly, multiplying the indirect utility utility
    function by the productivity distribution pdf gives:
def voptpdf(n,t,m,a,k,mu,minn):
    return
        vopt(n,t,m,a,k)*pareto.pdf(n,mu,loc=0,scale=minn)

#The average indirect utility over the able
    subpopulation is:
def wabaseline(t,m,a,k,mu,minn):
    return
```

```
        quad(voptpdf,minn,100,args=(t,m,a,k,mu,minn))[0]

#The marginal indirect utility of income multiplied by
    the productivity distribution pdf is:
def vmoptpdf(n,t,m,a,k,mu,minn):
    return ((1/(yopt(n,t,a,k)*(1-t)+m-
    a*(((yopt(n,t,a,k)/n)**(1+k))/(1+k))))
    *pareto.pdf(n,mu,loc=0,scale=minn))

#The utility of an unable individual is simply given by:
def u(m):
    return np.log(m)

#Their marginal utility of consumption is thus:
def ux(m):
    return 1/m

#The average marginal indirect utility of income over
    the able subpopulation is:
def sbaropt(t,m,a,k,mu,minn):
    return
        quad(vmoptpdf,minn,100,args=(t,m,a,k,mu,minn))[0]

#Under a strictly utilitarian welfare function (to be
    defined below), inequality in the average social
    marginal value of income (smvi) between the unable
    and able subpopulations is given by:
def betaopt(t,ma,mun,a,k,mu,minn):
    return ux(mun)-sbaropt(t,ma,a,k,mu,minn)
#where 'ma' denotes the unearned income received by the
    able; whilst 'mun' denote denotes the benefit
    income received by the unable.

#The critical tax rate - defined as t* in the main text
    - that eliminates the above between-group
    inequality can be solved for as follows. We first
    define net tax revenue below by the function:
def ntr(t,a,k,theta,mu,minn,r):
    return ((1-theta)*t*quad(yoptpdf,minn,100,
    args=(t,a,k,mu,minn))[0]-r)

#Next, we define the level of between-group inequality
    in the average smvi when net tax revenue is spent
    entirely on categorical transfers for the unable by:
def betapc(t,a,k,theta,mu,minn,r):
    return betaopt(t,0,ntr(t,a,k,theta,mu,minn,r)/theta
    ,a,k,mu,minn)

#Finally, the critical tax rate that generates just
    enough (net) tax revenue to eliminate between-group
    inequality in the average smvi is given by:
def tcrit(a,k,theta,mu,minn,r):
    return newton(betapc,0.6,args=(a,k,theta,mu,minn,r))

#1.1) FLAT TAX OPTIMISATION PROBLEM
#.................................................

#Let the vector x=(x[0],x[1],x[2])=(t,b,c) be variables
```

```
    we are maximising with respect to:

#Written in terms of these choice variables, net tax
    revenue in the economy is:
def budgetflat(x,a,k,theta,mu,minn,r):
   return ((1-theta)*x[0]*quad(yoptpdf,minn,100,
   args=(x[0],a,k,mu,minn))[0]-r )

#The optimisation problem is described by the below
    function:
def resultsbaseline(a,k,theta,mu,minn,r):
   consbaseline=({'type':'eq',
'fun':lambda x:np.array([x[1]+theta*x[2]-
budgetflat(x,a,k,theta,mu,minn,r)])},
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([1-x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])},
{'type':'ineq',
'fun':lambda x: np.array([x[2]])})
   def objbaseline(x,a,k,theta,mu,minn):
      return (-theta*u(x[1]+x[2])-
      (1-theta)*wabaseline(x[0],x[1],a,k,mu,minn))
   res=minimize(objbaseline,[0.1,0.1,0.1],
   args=(a,k,theta,mu,minn),constraints=consbaseline,
   method='SLSQP',options={'ftol':1e-10,'disp':False})
   return ('(t,b,c)=',res.x,'beta=',
   betaopt(res.x[0],res.x[1],res.x[1]+res.x[2],a,k,
   mu,minn),
   'sbar=',sbaropt(res.x[0],res.x[1],a,k,mu,minn),
   'G=', budgetflat(res.x,a,k,theta,mu,minn,r),
   'tstar=',tcrit(a,k,theta,mu,minn,r),
   'W=',-objbaseline(res.x,a,k,theta,mu,minn))

#The dictionary of constraints is given by
    'consbaseline'. The first of these is an equality
    constraint specifying that benefit expenditure must
    equate with net tax revenue. The remaining
    constraints are inequality constraints specifying
    that the optimal benefit levels must be
    non-negative; whilst the optimal tax rate must be
    interior and fall between 0 and 1.

#-------------------------------------------------------
# 2) TWO-BRACKET PIECEWISE ANALYSIS
#.......................................................

#The critical productivity at which an individual
    facing the lower marginal tax rate (t1) chooses to
    earn exactly the earnings threshold 'ycheck' is
    given by:
def ntilde(t1,ycheck,a,k):
   return ((a*((ycheck)**k))/(1-t1))**(1/(1+k))

#Similarly, the critical productivity at which an
    individual facing the upper marginal tax rate (t2)
```

```python
    chooses to earn exactly the earnings threshold
    'ycheck' is given by:
def ntilde2(t1,t2,ycheck,a,k):
    return ((a*((ycheck)**k))/(1-t2))**(1/(1+k))

#Putting this all together, the optimal earnings
    function for a productivity n individual is given
    by:
def y(n,t1,t2,ycheck,a,k):
    if n<ntilde(t1,ycheck,a,k):
        return ((n*(1-t1)/a)**(1/k))*n
    elif (ntilde(t1,ycheck,a,k)<=n<=
    ntilde2(t1,t2,ycheck,a,k)):
        return ycheck
    else:
        return ((n*(1-t2)/a)**(1/k))*n

#Multiplying the optimal earnings function by the
    productivity distribution pdf gives:
def ypdf(n,t1,t2,ycheck,a,k,mu,minn):
    return y(n,t1,t2,ycheck,a,k)*pareto.pdf(n,mu,loc=0,
    scale=minn)

#Integrating this over all productivities thus gives
    the average earnings over the able subpopulation.
def ybar(t1,t2,ycheck,a,k,mu,minn):
    return quad(ypdf,minn,100,args=(t1,t2,ycheck,a,
    k,mu,minn))[0]

#The differences between the earnings of an individual
    with n>ntilde2 and the earnings threshold,
    multiplied by the productivity distribution pdf are
    given by:
def ydiffpdf(n,t1,t2,ycheck,a,k,mu,minn):
    return ((((n*(1-t2)/a)**(1/k))*n-ycheck)*
    pareto.pdf(n,mu,loc=0
    ,scale=minn))

#The number of able individuals earning above the
    threshold 'ycheck' is given by:
def propoycheck(t1,ycheck,a,k,mu,minn):
    return 1-pareto.cdf(ntilde(t1,ycheck,a,k),
    mu,loc=0,scale=minn)

#The indirect utility of a productivity n individual is
    given by:
def v(n,t1,t2,ycheck,m,a,k):
    if n<ntilde(t1,ycheck,a,k):
        return np.log(((n*(1-t1)/a)**(1/k))*n*(1-t1)+m-
        a*(((((n*(1-t1)/a)**(1/k)))**(1+k))/(1+k)))
    elif ntilde(t1,ycheck,a,k)<=n<=
    ntilde2(t1,t2,ycheck,a,k):
        return np.log(ycheck*(1-t1)+m-a*
        (((ycheck/n)**(1+k))/(1+k)))
    else:
        return
            np.log(ycheck*(t2-t1)+((n*(1-t2)/a)**(1/k))
        *n*(1-t2)+m-a*(((((n*(1-t2)/a)**
```

```
        (1/k)))**(1+k))/(1+k)))

#Multiplying the indirect utility function by the
    productivity pdf gives:
def vpdf(n,t1,t2,ycheck,m,a,k,mu,minn):
    return (v(n,t1,t2,ycheck,m,a,k)*
    pareto.pdf(n,mu,loc=0,scale=minn))

#The average indirect utility over the able
    subpopulation is:
def wa(t1,t2,ycheck,m,a,k,mu,minn):
    return quad(vpdf,minn,100,
    args=(t1,t2,ycheck,m,a,k,mu,minn))[0]

#The marginal indirect utility of an able individual is:
def vmpw(n,t1,t2,ycheck,m,a,k):
    if n<ntilde(t1,ycheck,a,k):
        return 1/(((n*(1-t1)/a)**(1/k))*n*(1-t1)+m-
        a*(((((n*(1-t1)/a)**(1/k)))**(1+k))/(1+k)))
    elif (ntilde(t1,ycheck,a,k)<=n<=
    ntilde2(t1,t2,ycheck,a,k)):
        return 1/(ycheck*(1-t1)+m-
        a*(((ycheck/n)**(1+k))/(1+k)))
    else:
        return 1/(ycheck*(t2-t1)+((n*(1-t2)/a)**(1/k))*n*
        (1-t2)+
        m-a*(((((n*(1-t2)/a)**(1/k)))**(1+k))/(1+k)))
#Multiplying this by the productivity distribution pdf
    then gives:
def vmpwpdf(n,t1,t2,ycheck,m,a,k,mu,minn):
    return (vmpw(n,t1,t2,ycheck,m,a,k)*
    pareto.pdf(n,mu,loc=0,scale=minn))

#The average smvi over the able subpopulation is thus:
def sbarpw(t1,t2,ycheck,m,a,k,mu,minn):
    return quad(vmpwpdf,minn,100,args=
    (t1,t2,ycheck,m,a,k,mu,minn))[0]

#Between-group inequality in the avergae smvi is
    defined by:
def betapw(t1,t2,ycheck,ma,mun,a,k,mu,minn):
    return ux(mun)-sbarpw(t1,t2,ycheck,ma,a,k,mu,minn)

#2.1) PIECEWISE OPTIMISATION PROBLEM
#.................................................

# Let x=(x[0],x[1],x[2],x[3],x[4])=(t1,t2,ycheck,b,c)
    denote the variables we are optimising over.
#Written in terms of the choice vector, net tax revenue
    in the economy is given by:
def budget(x,a,k,theta,mu,minn,r):
    return
        (1-theta)*(x[0]*(quad(ypdf,minn,ntilde(x[0],x[2]
    ,a,k) ,args=(x[0],x[1],x[2],a,k,mu,minn))[0]+
    x[2]*propoycheck(x[0],x[2],a,k,mu,minn))+
    x[1]*quad(ydiffpdf,ntilde2(x[0],x[1],x[2],a,k),
    100,args=(x[0],x[1],x[2],a,k,mu,minn))[0])-r

```

```
#The optimisation problem is given by the below
    function.
def resultpw(a,k,theta,mu,minn,r,s):
    conspw=({'type':'eq',
'fun':lambda x:np.array([x[3]+theta*x[4]-
budget(x,a,k,theta,mu,minn,r)])},
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([1-x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])},
{'type':'ineq',
'fun':lambda x: np.array([1-x[1]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]-x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[2]])},
{'type':'ineq',
'fun':lambda x: np.array([x[3]])},
{'type':'ineq',
'fun':lambda x: np.array([x[4]])})
    def objpw(x,a,k,theta,mu,minn):
        return (-theta*u(x[3]+x[4])-
        (1-theta)*wa(x[0],x[1],x[2],x[3],a,k,mu,minn))
    res=minimize(objpw,[0.1,0.1,s,0.1,0.1],
    args=(a,k,theta,mu,minn),constraints=conspw,
    method='SLSQP',
    options={'ftol':1e-10,'disp':False})
    return ('(t1,t2,Y,b,c)=',res.x,'beta=',
    betapw(res.x[0],res.x[1],res.x[2],res.x[3],res.x[3]+
    res.x[4],a,k,mu,minn),
    'sbar=',
        sbarpw(res.x[0],res.x[1],res.x[2],res.x[3],a,k,mu
    ,minn),'G=',budget(res.x,a,k,theta,mu,minn,r),
    'ybar=',ybar(res.x[0],res.x[1],res.x[2],a,k,mu,minn),
    'W=',-objpw(res.x,a,k,theta,mu,minn))

#The dictionary of constraints is given by 'conspw'.
    The first constraint is an equality constraint and
    specifies that benefit expenditure must equate with
    net tax revenue at the optimum.The remaining
    constraints are inequality constraints specifying
    that: (i) the upper and lower tax rates must be
    interior and fall between zero and one; (ii) the
    upper tax rate must be at least as great as the
    lower tax rate; and (iii) the optimal benefit
    levels must be non-negative.

#The term 's' is a starting point in the search for the
    optimal earnings threshold 'ycheck'. The starting
    point can be (and is) varied systematically in the
    search for the optimal piecewise system.

#2.2. OPTIMISATION PROBLEM FOR A GIVEN VALUE OF
    'ycheck'.
#.................................................
#As a further check on the results generated by
```

```
    'resultpw',we also consider the two-stage
    optimisation problem. This involves first taking
    'ycheck' as given and solving the optimisation for
    the optimal tax rates and and benefit levels. We
    then vary the fixed value of 'ycheck' and determine
    the value of 'ycheck' at which welfare is highest.

#To proceed, we now let the vector
    x=(x[0],x[1],x[2],x[3])=(t1,t2,b,c).

#Net tax revenue written in terms of these choices
    variables is thus:
def budget1(x,ycheck,a,k,theta,mu,r):
    return
        (1-theta)*(x[0]*(quad(ypdf,1,ntilde(x[0],ycheck,a,k),
    args=(x[0],x[1],ycheck,a,k,mu))[0]+
    ycheck*propoycheck(x[0],ycheck,a,k,mu))+
    x[1]*quad(ydiffpdf,ntilde2(x[0],x[1],ycheck,a,k),
    100,args=(x[0],x[1],ycheck,a,k,mu))[0])-r

#The optimisation problem is now described by:
def results1(ycheck,a,k,theta,mu,r):
    cons1=({'type':'eq',
'fun':lambda x:np.array([x[2]+theta*x[3]-
budget1(x,ycheck,a,k,theta,mu,r)])},
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([1-x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])},
{'type':'ineq',
'fun':lambda x: np.array([1-x[1]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]-x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[2]])},
{'type':'ineq',
'fun':lambda x: np.array([x[3]])})
    def obj1(x,ycheck,a,k,theta,mu):
        return (-theta*u(x[2]+x[3])-
        (1-theta)*wa(x[0],x[1],ycheck,x[2],a,k,mu))
    res=minimize(obj1,[0.1,0.1,0.1,0.1],
    args=(ycheck,a,k,theta,mu)
    ,constraints=cons1,method='SLSQP',
    options={'ftol':1e-10,'disp':False})
    return res.x, -obj1(res.x,ycheck,a,k,theta,mu)
```

2

Chapter 4

The Optimal Linear Income Tax with Imperfectly Administered Categorical Transfers

4.1 Introduction

When the linear income tax framework is augmented with categorical heterogeneity that the planner can *perfectly* observe, there is a well-defined ordering of priorities between tax financed categorical transfers and a universal benefit (demogrant): net tax revenue should be spent solely on categorical transfers up to the point that inequality in the average social marginal value (smvi) across categorical groups is eliminated.[1] Conditional on it being optimal to raise enough tax revenue to achieve this, *and* there being tax revenue left over, a universal benefit (demogrant) will be provided to all individuals in society. The resulting optimal tax expression captures the equity-efficiency tradeoff inherent in income taxation; with equity terms in the numerator and those of efficiency in the denominator. In general, the equity considerations will be a function of both (i) any between-group disparity in the average smvi; and (ii) the 'within-group' negative of the covariance between individual earnings and their smvi. The efficiency considerations, meanwhile, capture the aggregate (compensated) responsiveness of individual

[1] The result that categorical transfers should be set to eliminate inequality in the average smvi between categorical groups is well-established (see Diamond and Sheshinski, 1995; Viard, 2001a)

gross earnings to the tax rate.

In reality, however, the complexities of determining an individual's categorical status (e.g. disability or involuntary unemployment) mean that categorical benefits are likely to be administered with both Type I (false rejection) and Type II (false award) classification errors.[2] The extent to which these errors are made will affect the ordering of priorities between categorical and universal benefits and, in turn, have implications for the optimal tax rate. This chapter is concerned with the additional equity and efficiency considerations that classification errors introduce into the optimal tax expression. For example, if an increase in the tax rate induces an additional ineligible individual to apply for - and ultimately be awarded - a categorical benefit, this may give rise to both (i) losses in tax revenue from this individual and (ii) additional welfare expenditure costs.

This chapter considers an economy where a fraction of the population is unable to work, whilst the remaining fraction is composed of individuals who are able to work but differ continuously in their productivity. The government operates a three-part tax-benefit system comprising (i) a constant marginal tax rate on all earned income; (ii) a tax-free universal benefit received unconditionally by all individuals in society; and (iii) a tax-free categorical benefit that is *ex-ante* conditional on an applicant being unable to work and *ex-post* conditional on a recipient not working. The test awarding the categorical benefit makes Type I and Type II classification errors. However, it is assumed that the *ex-post* 'no-work' condition is fully enforced such that no able individual who is awarded the benefit by Type II error will subsequently work. This therefore restricts the able individuals who apply for the categorical benefit to those of lower productivities.[3] Moreover, the critical productivity at or below which an able individual chooses to apply for the categorical benefit will be an *increasing* function of the tax rate.

An increase in the tax rate will therefore generate both *direct* and *indirect* behavioural responses that affect the government budget constraint. The direct effect is simply that found in all conventional analyses: an increase in the tax rate induces individuals

[2] A growing empirical literature analyses the scope for Type II errors in the U.S. Social Security Disability Insurance programme (see Autor and Duggan, 2003, 2006; Benitez-Silva et al., 2004; Chen and van der Klaauw, 2008; Von Wachter et al., 2011).

[3] Applications for the categorical benefit are taken to be costless in terms of money, stigma and time.

to adjust their labour supply in the intensive margin. The indirect effect, meanwhile, captures the fact that an increase in the tax rate will induce additional able individuals to apply for the categorical benefit and, if awarded it, stop working and generating tax revenue. This will also affect the expenditure side of the budget constraint because more individuals receive the categorical benefit.

The key contributions of this chapter are (i) to provide an expression for the optimal linear tax rate that captures the additional equity and efficiency considerations of income taxation that arise with classification errors; and (ii) to numerically simulate how the optimal tax rate (and benefit levels) changes with the propensity to make classification errors of either type.

With regard to the optimal tax expression, an important consideration in both the numerator (equity terms) and denominator (efficiency terms) becomes the elasticity of the distribution function with respect to individual productivity, evaluated at the critical productivity at or below which able individuals choose to apply for the categorical benefit. Type II errors generate conflicting effects in both the equity and efficiency dimensions. In the equity dimension, Type II errors (i) mean that some able individuals of low productivity - who the government would not wish to tax highly - receive the categorical benefit and do not work, which acts to raise the tax rate; but also (ii) redistribute within the able subpopulation through 'leaking' categorical transfers to lower productivity individuals, which acts to lower the tax rate because there may be less need to redistribute via the universal benefit. In the efficiency dimension (i) Type II errors mean that some individuals of lower productivity do not work and thus their response to taxation can be disregarded, therefore acting to increase the tax rate; but (ii) the number of applications for the categorical benefit is, *ceteris paribus*, an increasing function of the tax rate, and because the government must pay each new recipient the categorical benefit in addition to losing tax revenue from this individual, this acts to lower the tax rate.

Individual utility is throughout taken to be a concave transformation of quasilinear preferences (linear in consumption and convex-decreasing in labour). The implications of this assumption are that (i) there are are no income effects associated with a working individual's smvi; and (ii) the size of the universal benefit does not influence an able individual's decision to apply for the categorical benefit. This allows us to establish a precise relationship between the average smvi of an able applicant for the categorical benefit and the shadow price of public expenditure, which in turn greatly assists in interpreting the equity considerations in the optimal tax expression.

To gain comparative statics insights into how the optimal tax rate changes with the propensity to make classification errors, we turn to numerical methods. Preferences take a frequently employed isoelastic form (see Atkinson, 1990; Saez, 2001) that is consistent with those adopted in the theoretical section of the chapter. Individual productivity is lognormally distributed. This has been the distribution of choice for the key studies of categorical transfers within the optimal income tax framework (Immonen et al., 1998; Viard, 2001a). For varying values of the constant labour supply elasticity the results all suggest that an increase in the propensity to make *either* error type (i) increases the optimal tax rate; (ii) increases the optimal universal benefit; but (iii) *decreases* the optimal categorical benefit. Welfare provision thus becomes increasingly universal as the the discriminatory power of the test awarding the categorical benefit decreases. As in the theory, however, the categorical benefit always remains positive such that some form of targeting is always desirable.

Whilst a number of papers have analysed perfectly administered categorical transfers in a variant of the standard linear income tax framework where individuals differ over *both* a productivity continuum and some categorical attribute[4], there has been little work on imperfectly administered transfers in this setting. Indeed, the literature typically restricts individual heterogeneity in the productivity dimension to two types (Jacquet, 2014; Stern, 1982). In this regard, the closest precursor to this chapter dates back to Stern (1982), who compares social welfare under (i) imperfect lump-sum taxation[5] and a proportional income tax; with (ii) optimal non-linear income taxation. It is the first case that is related to this chapter: individual earnings are taxed linearly at source and the government can classify, albeit with error, individuals into 'skilled' and 'unskilled' groups. With no classification errors it may be optimal to impose a lump-sum tax on skilled individuals to finance transfers to the unskilled, with the income tax set at zero. However, a positive propensity to incorrectly classify the unskilled as skilled exposes

[4]See, for example, Viard (2001a,b).

[5]Whilst there is a clear distinction between categorical transfers and lump-sum transfers in the general case where the categorical attribute (e.g. disability) is imperfectly correlated with ability, it is also important to stress that the two converge in the extreme case where there is a perfect correlation. This would arise, for example, in a simple two-type setting where disability status is perfectly correlated with low ability. For a further discussion on the similarity between categorical transfers and lump-sum transfers see Viard (2001a,b).

the former to the risk of receiving a smaller and even negative transfer, which is welfare reducing. Consequently, the linear income tax - which is independent of 'categorical' status and thus the same for both types - may rise to ensure transfers to the skilled are sufficiently large. As in the current chapter, the numerical simulations in Stern illustrate that the optimal linear income tax rate increases with the propensities to misclassify either type of individual.

The remainder of this chapter is structured as follows. Section 4.2 presents the model and the main theoretical analysis. Section 4.3 then presents the numerical analysis. Finally, Section 4.4 concludes the chapter.

4.2 The Model

4.2.1 Individuals

Let individual preferences over consumption ($x \geq 0$) and labour ($H \geq 0$) be represented by the utility function:

$$U(x, H) = u\left(x - g(H)\right) \tag{4.1}$$

where $u' > 0$; $u'' < 0$ and $\lim_{x \to 0} u' = +\infty$; whilst $g' > 0$; $g'' > 0$ and $\lim_{H \to 0} g = \lim_{H \to 0} g' = 0$. This last assumption simply ensures that there will always be an interior optimum for H.[6]

An individual with net wage $\omega \geq 0$ and unearned income $M \geq 0$ has an optimal labour supply function $H^*(\omega)$ that satisfies $\omega \equiv g'(H^*)$. Notice that this is independent of unearned income. By the assumptions placed on g, it follows directly that H^* is increasing in ω.[7]

We denote an individual's indirect utility function by $v(\omega, M) \equiv u[\omega H^* + M - g(H^*)]$. It follows follows from the envelope theorem that $v_\omega = u'H^*$ and $v_M = u'$, which in turn implies $v_\omega = v_M H^*$ (i.e. Roy's identity). In addition, $v_{\omega M} = u''H^* < 0$ such that the marginal indirect utility of income is decreasing in the net wage.

[6]The implication is therefore that there will be no reservation wage.

[7]Differentiating both sides of $\omega = g'(H^*)$ w.r.t. ω gives $dH^*(\omega)/d\omega = [g''(H^*)]^{-1} > 0$.

With this by way of background we proceed to define the population and tax-benefit system in place.

4.2.2 The Population and Tax-Benefit System

Consider an economy of size 1 where a fraction $\theta \in (0,1)$ of the population face a zero quantity constraint on labour supply and are thus unable to work. The remaining $(1-\theta)$ individuals are able to work but differ in their underlying productivity, n ; where $n \in (0,\infty)$ is continuously distributed with density function $f(n)$ and distribution function $F(n)$.

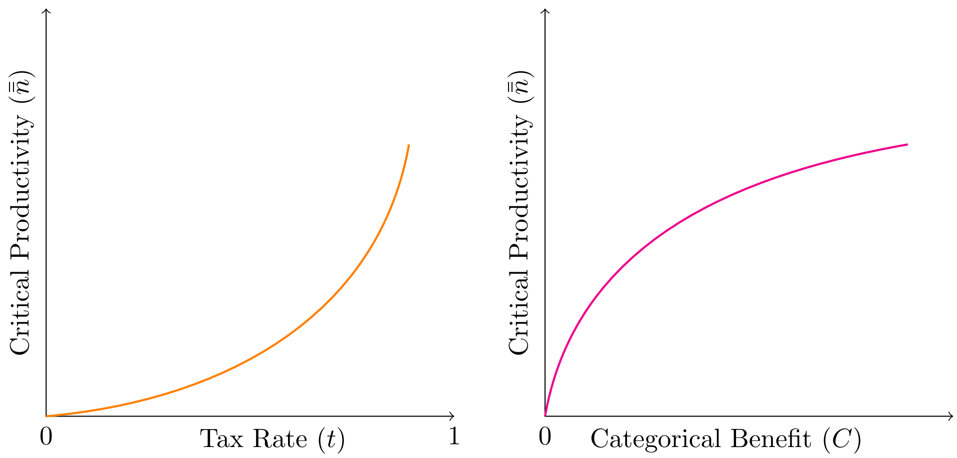
The government operates a three-part tax-benefit system comprising (i) a constant marginal tax rate $t \in (0,1)$ on all earned income; (ii) a tax-free universal benefit $B \geq 0$ that is received unconditionally by all individuals in society; and (iii) a tax-free categorical benefit $C \geq 0$ that is *ex-ante* conditioned on applicants being unable to work and *ex-post* conditioned on recipients not working. Applications for the categorical benefit are taken to be costless in terms of money, time and stigma. The test awarding the benefit makes Type I (false rejection) and Type II (false award) classification errors with probabilities p_I and p_{II}, respectively. We assume that $p_I + p_{II} \leq 1$, which guarantees that the test is never more likely to award the categorical benefit to an able applicant than an unable applicant.

Enforcement and the critical application productivity. In what follows, we assume that the *ex-post* no-work condition is fully enforced such that no able individual who is incorrectly awarded C by Type II error ever chooses to work. The implication is that only those of sufficiently low productivity will choose to apply for C. With this in mind, let the critical productivity $\bar{\bar{n}}(1-t, C)$ satisfy $u(B+C) \equiv v[\bar{\bar{n}} \cdot (1-t), B]$ and thus:

$$C = \bar{\bar{n}}(1-t)H^*\left(\bar{\bar{n}}(1-t)\right) - g\left[H^*\left(\bar{\bar{n}}(1-t)\right)\right] \tag{4.2}$$

It follows that all able individuals with $n \in [0, \bar{\bar{n}}]$ will choose to apply for C; whilst those with $n \in (\bar{\bar{n}}, \infty)$ will not apply because the opportunity cost of not working is

Figure 4.1: The Critical Productivity $\bar{\bar{n}}(1-t, C)$



Notes. This figure illustrates how the critical productivity $\bar{\bar{n}}$ changes with both the tax rate (left subplot) and the categorical benefit size (right subplot).

too high[8]. Differentiating (4.2) with respect to $(1-t)$ and C, respectively, yields:

$$\frac{\partial \bar{\bar{n}}}{\partial(1-t)} = -\frac{\bar{\bar{n}}}{1-t} < 0 \ ; \ \frac{\partial^2 \bar{\bar{n}}}{\partial(1-t)^2} = \frac{2\bar{\bar{n}}}{(1-t)} > 0 \ . \tag{4.3}$$

and

$$\frac{\partial \bar{\bar{n}}}{\partial C} = \frac{1}{(1-t)H^*} > 0 \ ; \quad \frac{\partial^2 \bar{\bar{n}}}{\partial C^2} = -\frac{1}{(1-t)g''(H^*) \cdot (H^*)^3} < 0 \tag{4.4}$$

A marginal increase in either t or C thus increases the critical productivity at or below which an able individual chooses to apply for C and, if awarded it, stop working (see Figure 4.1).

Aggregate gross earnings and the tax tate. Let aggregate gross earnings over the able subpopulation be given by the composite function:

$$Z\left(1-t, \bar{\bar{n}}; \theta, p_{II}\right) = (1-\theta)\left\{\int_0^{\infty} y(n, 1-t)dF(n) - p_{II}\int_0^{\bar{\bar{n}}} y(n, 1-t)dF(n)\right\} \tag{4.5}$$

[8]We assume that at the point of indifference an able individual will choose to apply for C.

where $y(n, 1-t)$ denotes individual optimal gross earnings. The *direct* and *indirect* effects of an increase in the net-of-tax rate are given by, respectively:

$$\frac{\partial Z}{\partial(1-t)} = (1-\theta)\left(\int_0^\infty \frac{\partial y}{\partial(1-t)} dF(n) - p_{II}\int_0^{\bar{\bar{n}}} \frac{\partial y}{\partial(1-t)} dF(n)\right) > 0 \;, \quad (4.6)$$

$$\frac{\partial Z}{\partial \bar{\bar{n}}} \cdot \frac{\partial \bar{\bar{n}}}{\partial(1-t)} = (1-\theta)F(\bar{\bar{n}})p_{II}\mathcal{E}_f(\bar{\bar{n}})y|_{n=\bar{\bar{n}}}(1-t)^{-1} > 0 \;. \quad (4.7)$$

where $\mathcal{E}_f(n) \equiv nf(n)/F(n)$ is the elasticity of the distribution function with respect to individual productivity.

The direct effect of an increase in the net-of-tax rate is simply that found in all conventional analyses: it captures the aggregate intensive margin response of working individuals to the increased work incentives, holding constant the critical productivity $\bar{\bar{n}}$. Meanwhile, the indirect effect of an increase in the net-of-tax rate is captured by the fact that $\bar{\bar{n}}$ falls with the net-of-tax rate, and thus so too do the number of applicants for the categorical benefit. This increases the number of working individuals in the economy. Notice that the size of the indirect effect depends on the elasticity of the distribution function with respect to individual productivity, evaluated at the critical productivity $\bar{\bar{n}}$. Both direct and indirect effects act to increase aggregate earnings.

Government Budget Constraint. The government budget constraint is given by:

$$B + \chi\left(t, C; \theta, p_I, p_{II}\right) \cdot C \leq t \cdot Z\left(1-t, \bar{\bar{n}}; \theta, p_{II}\right) - R \quad (4.8)$$

where $\chi \equiv \theta(1-p_I) + (1-\theta)F(\bar{\bar{n}})p_{II}$ denotes the number of categorical recipients in the economy, whilst $R \geq 0$ is an exogenous revenue requirement for spending outside of welfare. One can readily establish that:

$$\frac{\partial \chi}{\partial t} = (1-\theta)F(\bar{\bar{n}})p_{II}\mathcal{E}_f(\bar{\bar{n}})(1-t)^{-1} > 0 \quad (4.9)$$

Intuitively, a *ceteris paribus* increase in the tax rate reduces the gain to working and, consequently, incentivises additional able individuals to apply for the categorical benefit. The size of this effect is increasing in the elasticity of the distribution function with respect to individual productivity, evaluated at the critical productivity $\bar{\bar{n}}$.

Let $C_F(t, B; \theta, p_I, p_{II}, R)$ be the value of the categorical benefit that exhausts the budget constraint for any pair (t, B) and exogenous parameters (θ, p_I, p_{II}, R). Formally, C_F is defined by:

$$t \cdot Z\left[1 - t, \bar{\bar{n}}(1 - t, C_F); \theta, p_{II}\right] - \chi(t, C_F; \theta, p_I, p_{II}) \cdot C_F \equiv (B + R) \tag{4.10}$$

Written in this way, we can see that the left side is unambiguously decreasing in C (because $\partial Z/\partial C < 0$, $\partial \chi/\partial C > 0$), whilst the right side is independent of C. Differentiating (4.10) with respect to t gives:

$$\begin{aligned}
\frac{\partial C_F}{\partial t} &= \frac{Z - t\left(\dfrac{\partial Z}{\partial(1-t)} + \dfrac{\partial Z}{\partial \bar{\bar{n}}}\dfrac{\partial \bar{\bar{n}}}{\partial(1-t)}\right) - \dfrac{\partial \chi}{\partial t}C_F}{\dfrac{\partial \chi}{\partial C}C_F + \chi - t\dfrac{\partial Z}{\partial \bar{\bar{n}}}\dfrac{\partial \bar{\bar{n}}}{\partial C}} \\
&= \frac{Z - t\dfrac{\partial Z}{\partial(1-t)} - (1-\theta)F(\bar{\bar{n}})p_{II}\mathcal{E}_f(\bar{\bar{n}})(ty|_{n=\bar{\bar{n}}} + C_F)/(1-t)}{\chi + (1-\theta)p_{II}f(\bar{\bar{n}})\dfrac{\partial \bar{\bar{n}}}{\partial C}\left(C_F + ty|_{n=\bar{\bar{n}}}\right)}
\end{aligned} \tag{4.11}$$

The numerator in (4.11) captures the effect of an increase in the tax rate on the benefit budget, holding constant the categorical benefit size. The first term, Z, captures the effect of an increase in the tax rate on tax revenue in the absence of any behavioural responses. Were this the only term it would unambiguously be the case that $\partial C_F/\partial t > 0$. There are of course behavioural responses and these are captured by the remaining terms. The second term is the negative of the direct effect in (4.6). It thus corresponds to the reduction in aggregate earnings caused by individuals reducing their labour time in the intensive margin. The third term captures the fact that $\bar{\bar{n}}$ increases with the tax rate, and thus so too do the number of categorical applicants (and recipients) from the able subpopulation. As discussed in both (4.7) and (4.9), this results in both (i) a loss in tax revenue as individuals who previously worked now do not; and (ii) the expenditure cost of paying the categorical benefit to the new recipients. Overall, the sign of the numerator - and in turn the sign of $\partial C_F/\partial t$ - will depend on whether or not any rise in tax revenue generated by the increase in t is offset by the additional benefit expenditure costs it induces.

The denominator in (4.11) captures the effect of an increase in the categorical benefit on the benefit budget. It is unambiguously positive. An increase in the categorical benefit requires paying all existing categorical recipients a higher benefit level. This in turn incentivises additional individuals to apply, which generates both (i) a loss in tax revenue as individuals who previously worked now do not; and (ii) the expenditure cost of paying new recipients the categorical benefit. The extent to which C_F changes with the tax rate (for a given B) will be decreasing in both the number of existing recipients; and the size of tax revenue effects and expenditure effects associated with new applicants.

Notice that differentiating (4.10) with respect to B yields:

$$\frac{\partial C_F}{\partial B} = -\left[\chi + (1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial \bar{\bar{n}}}{\partial C}\left(ty|_{n=\bar{\bar{n}}} + C_F\right)\right]^{-1} < 0 \tag{4.12}$$

This is simply the negative of the denominator in (4.11). Unsurprisingly, an increase in B necessarily lowers C_F because there are less resources to spend on categorical transfers.

In all that follows we assume that $\partial C_F/\partial B < -1$, which requires:

$$(1-\chi) > (1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial \bar{\bar{n}}}{\partial C}(ty|_{n=\bar{\bar{n}}} + C_F)$$

The implication is that the total benefit income of categorical recipients will fall with an increase in the universal benefit.

Laffer Rate. The revenue maximising (or 'Laffer') tax rate is defined as:

$$t_L(B;\theta,p_I,p_{II},R) = \operatorname{Arg}\max_{t\in(0,1)} t\cdot Z\left[1-t,\bar{\bar{n}}\left(1-t,C_F\right);\theta,p_{II}\right] \tag{4.13}$$

Notice that whilst individual earnings are not directly affected by the level of B, the categorical benefit size C_F is. Because a *ceteris* paribus increase in B lowers C_F it will increase the number of able individuals who work (through lowering the number that apply for the categorical benefit). The revenue maximising tax rate will therefore depend on B.

The resulting first-order-condition (henceforth FOC) characterising t_L yields:

$$t_L = Z\left[\frac{\partial Z}{\partial(1-t)} - \frac{\partial Z}{\partial\bar{\bar{n}}}\left(\frac{\partial\bar{\bar{n}}}{\partial C}\cdot\frac{\partial C_F}{\partial t} - \frac{\partial\bar{\bar{n}}}{\partial(1-t)}\right)\right]^{-1} \tag{4.14}$$

Dividing both sides $(1-t)$ allows us to write the expression in terms of elasticities.

$$\frac{t_L}{1-t_L} = \left\{\left[\frac{(1-t)}{Z}\cdot\frac{\partial Z}{\partial(1-t)}\right] + \left[\left(\frac{\bar{\bar{n}}}{Z}\cdot\frac{\partial Z}{\partial\bar{\bar{n}}}\right)\cdot\left(\frac{(1-t)}{\bar{\bar{n}}}\cdot\frac{d\bar{\bar{n}}(1-t,C_F)}{d(1-t)}\right)\right]\right\}^{-1} \tag{4.15}$$

The first term within curly braces is the partial elasticity of aggregate earnings with respect to the net-of-tax rate. It captures the aggregate responsiveness of earnings in the intensive margin to an increase in the net-of-tax rate. In all that subsequently follows we will denote this by $\mathcal{E}_Z \equiv (1-t)Z^{-1}\partial Z/\partial(1-t)$.[9] The second term within curly braces, meanwhile, is the product of two elasticities. The first of these is the elasticity of aggregate earnings with respect to the critical productivity $\bar{\bar{n}}$. The second, meanwhile, is the elasticity of this critical productivity with respect to the net-of-tax rate. Intuitively, larger values of the aforementioned elasticities imply a lower revenue maximising tax rate.

4.2.3 The Optimisation Problem

Let the social welfare function be strictly utilitarian and given by:

[9]Notice that if $p_{II} = 0$ the first-order-condition characterising the Laffer Rate would simply be $Z - t_L\partial Z/\partial(1-t) = 0$, thus giving rise to the familiar expression (Saez, 2001):

$$\frac{t_L}{1-t_L} = \frac{1}{\mathcal{E}_Z}$$

.

$$W(t,B,C;\theta,p_I,p_{II}) = \theta\left[(1-p_I)u(B+C)+p_I u(B)\right] + (1-\theta)\left\{\begin{array}{c}\int_0^\infty v[n(1-t),B]f(n)dn \\ +p_{II}\int_0^{\bar{\bar{n}}}\langle u(B+C)-v[n(1-t),B]\rangle f(n)dn\end{array}\right\} \tag{4.16}$$

The first line is the average welfare of an unable individual - accounting for the fact that with probability p_I they are incorrectly denied the categorical benefit - weighted by their population share. The second line, meanwhile, is the average welfare of unable individuals multiplied by their population share. The first term within curly braces is the average welfare of able individuals in the case that all receive the universal benefit, whilst the second term captures the welfare gain to able individuals who are awarded the categorical benefit with probability p_{II}.

Given the government's budget constraint described in (4.8), the optimisation problem is therefore:

$$\begin{aligned}&\max_{\{t,B,C\}} W(t,B,C;\theta,p_I,p_{II})\\ &\text{s.t.}\quad B+\chi(t,C;\theta,p_I,p_{II})\cdot C = t\cdot Z\left[t,\bar{\bar{n}}(1-t,C);\theta,p_{II}\right]-R\ ,\\ &\qquad t\in(0,1)\ ,\ B\geq 0\ ,\ C\geq 0\end{aligned} \tag{4.17}$$

In what follows we let $\hat{t}$, $\hat{B}$ and $\hat{C}$ denote the optimal choices resulting from (4.17). In addition, $\hat{\lambda}$ will denote the shadow price of public expenditure at the optimum. To save on notation, we will henceforth denote the aggregate smvi of non-categorical recipients by:

$$\sigma(t,B,C) = \theta p_I u'(B) + (1-\theta)\left[\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} v_M dF(n)\right] \tag{4.18}$$

The solution to the optimisation problem in (4.17) is characterised in the below result.[10]

Result 1:

(i) *$\hat{C} > 0 \;\forall\; p_I + p_{II} \leq 1$ and $\hat{B} \geq 0$ are characterised by:*

$$\sigma\left(\hat{t}, \hat{B}, \hat{C}\right) \leq u'\left(\hat{B} + \hat{C}\right)\chi \cdot - \underbrace{\left[1 + \frac{\partial C_F(\hat{t}, \hat{B}; \cdot)}{\partial B}\right]}_{<0} \;\; ; \; \hat{B} \geq 0 \tag{4.19}$$

where the pair of inequalities hold with complementary slackness.

(ii) *$\hat{\lambda} < u'(\hat{B} + \hat{C}) = v_M[\bar{\bar{n}}(1 - \hat{t}), \hat{B}]$ and so:*

$$v_M[n(1 - \hat{t}), \hat{B}] > \hat{\lambda} \;\forall\; n \in (0, \bar{\bar{n}}] \tag{4.20}$$

(iii) *$\hat{t} \in (0, 1)$ is implicitly defined by:*

$$\frac{\hat{t}}{1 - \hat{t}} = \frac{\left[\int_0^\infty y\left(\hat{\lambda} - v_M\right) dF(n) - p_{II}\int_0^{\bar{\bar{n}}} y\left(\hat{\lambda} - v_M\right) dF(n)\right] - F(\bar{\bar{n}})p_{II}\hat{\lambda}\mathcal{E}_f(\bar{\bar{n}})\hat{C}}{\hat{\lambda}\left[\bar{Z}\mathcal{E}_Z + F(\bar{\bar{n}})p_{II}\mathcal{E}_f\left(y|_{n=\bar{\bar{n}}} + \hat{C}\right)\right]} \tag{4.21}$$

where $\bar{Z} = Z/(1 - \theta)$ denotes average gross earnings over the able subpopulation.

Proof: See Appendix

Corollary 1: *If $p_I > 0$ then $\hat{B} > 0$ by the property $\lim_{x\to 0} u' = +\infty$ and, consequently, (4.19) holds with equality.*

Result 1 is composed of three related parts, culminating with the optimal tax expression in part (iii). We proceed to discuss each in turn.

Result 1(i) states that a categorical benefit should be provided at all levels of discriminatory power[11], whilst the optimality condition in (4.19) characterises the conditions under which it is optimal to provide a universal benefit. The left side of (4.19) is the aggregate smvi of individuals who do *not* receive the categorical benefit, as composed of both unable and able individuals. This captures the aggregate welfare gain to these

[10]In line with the discussion of the Full Enforcement regime in Chapter 2, we cannot in general guarantee a unique solution to this optimisation problem. However, in the numerical analysis which follows welfare is always concave and, consequently, no problems of multiple optima arise.

[11]Therefore including the case of no discriminatory power where $p_I + p_{II} = 1$.

individuals from an increase in the universal benefit. The right side, meanwhile is the aggregate smvi of categorical recipients multiplied by the *negative of* the total change in their benefit income associated with an increase in the universal benefit. Given our assumption that $\partial C_F/\partial B < -1$, an increase in the universal benefit acts to reduce the total benefit income of categorical recipients. The right side therefore captures the aggregate welfare gain to categorical recipients associated with a reduction in the universal benefit. The complementary slackness condition implies that it will only be optimal to provide a positive universal benefit if the left side of (4.19) exceeds the right side, when evaluated at $B = 0$. As stated in Corollary 1, this is guaranteed to hold whenever the propensity to make Type I errors is positive because unable individuals would otherwise have zero income to consume.

A direct implication of (4.19) is that:

$$u'\left(\hat{B} + \hat{C}\right) > \sigma\left(\hat{t}, \hat{B}, \hat{C}\right)/(1-\chi) \tag{4.22}$$

I.e. at the optimum the smvi of categorical recipients will exceed the average smvi of non-categorical recipients. This arises because:

$$\chi \cdot - \left[1 + \frac{\partial C_F(\hat{t}, \hat{B}; \cdot)}{\partial B}\right] = (1-\chi) \cdot \left[\frac{1 - (1-\theta)p_{II} f(\bar{\bar{n}}) \frac{\partial \bar{\bar{n}}}{\partial C}\left(ty|_{n=\bar{\bar{n}}} + C_F\right)/(1-\chi)}{1 + (1-\theta)p_{II} f(\bar{\bar{n}}) \frac{\partial \bar{\bar{n}}}{\partial C}\left(ty|_{n=\bar{\bar{n}}} + C_F\right)/\chi}\right] < 1$$

Result 1(ii) states that at the optimum the smvi of an able applicant for the categorical benefit *exceeds* the shadow price of public expenditure. The reason for including this as part of the main result is that it assists us in interpreting the optimal tax expression in Result 1(iii), which we turn to discuss now.

Result 1(iii) provides an implicit expression for the optimal tax rate, as given by (4.21). As is standard, the implicit expression characterising the optimal tax rate has equity considerations in the numerator and efficiency considerations in the denominator. We start by discussing the numerator. The two terms inside square braces are aggregates of $y(\hat{\lambda} - v_M)$: i.e. the difference between the shadow price of public expenditure and an individual's smvi, weighted by gross earnings. The first term is the aggregate (or in this case average) of $y(\hat{\lambda} - v_M)$ over the entire able subpopulation; whilst the second term - which enters negatively - is the aggregate of $y(\hat{\lambda} - v_M)$ over able applicants. From (4.20) we know that this aggregate must itself be must be negative because the smvi of each able applicant exceeds the shadow price of public expenditure at the optimum. This second term therefore acts to increase the tax rate because those who the government would not wish to tax highly are - due to Type II error - receiving the categorical transfer and not working. However, a further equity implication of Type II errors on the tax rate is that 'leaked' categorical transfers do play a redistributive role within the able subpopulation, which may lessen the need to employ higher tax rates to redistribute through the more expensive universal transfer. This is captured by the third term outside square brackets, which acts to lower the tax rate. Notice that the size of this third term is an increasing function of the elasticity of the distribution function with respect to individual productivity, evaluated at $\bar{\bar{n}}$ (i.e. $\mathcal{E}_f$).

We now turn to discuss the efficiency considerations in the denominator of (4.21). The first term within square brackets (i.e. $\bar{Z}\mathcal{E}_Z$) is the *partial* elasticity of aggregate gross earnings with respect to the net-of-tax rate, weighted by average gross earnings over the able subpopulation. Recall that this elasticity captures the responsiveness of individual earnings in the intensive margin to the net-of-tax rate, holding constant the number of applicants for the categorical benefit. This consideration is present in all standard optimal tax expressions.[12] *Ceteris paribus*, higher values of this elasticity act to lower the optimal tax rate. Notice also that a *ceteris paribus* increase in p_{II} will act to reduce this term because fewer lower productivity individuals work and respond to the tax rate. The second and final term within square brackets (i.e. $F(\bar{\bar{n}})p_{II}\mathcal{E}_f(y|_{n=\bar{\bar{n}}} + C)$) captures the fact that a marginal increase in the tax rate increases $\bar{\bar{n}}$ and thus the number of applicants for the categorical benefit. This in turn generates the dual efficiency considerations of (i) foregone tax revenue from individuals who previously worked but

[12]Note that there a number of equivalent notational conventions in the literature to capture efficiency considerations. Specifically, the first term in the denominator of (4.21) will often appear in terms of *individual* elasticities. To see that this is equivalent to the formulation here, which draws from Piketty and Saez (2012), note that:

$$\bar{Z}\mathcal{E}_Z = \frac{Z}{(1-\theta)}\left(\frac{1-t}{Z}\cdot\frac{\partial Z}{\partial(1-t)}\right) = (1-t)\left[\int_0^\infty \frac{\partial y}{\partial(1-t)}dF(n) - p_{II}\int_0^{\bar{\bar{n}}} \frac{\partial y}{\partial(1-t)}dF(n)\right]$$
$$= \int_0^\infty y\mathcal{E}_y dF(n) - p_{II}\int_0^{\bar{\bar{n}}} y\mathcal{E}_y dF(n)$$

where $\mathcal{E}_y \equiv (1-t)y^{-1}\partial y\partial(1-t)$ is the elasticity of individual gross earnings with respect to the net-of-tax rate.

now receive the categorical benefit by Type II error; and (ii) the expenditure effect of paying each new recipient the categorical benefit. Notice that the size of this effect is increasing in the gross earnings of the marginal $\bar{\bar{n}}$ individual; the categorical benefit size; and the elasticity of the distribution function with respect to individual productivity, evaluated at $\bar{\bar{n}}$. Higher values of any of these these considerations act to lower the optimal tax rate.

4.2.4 Between-group inequality in the average smvi

By way of background, in the case where categorical transfers are perfectly targeted at the unable it will only be optimal to provide a universal benefit if inequality in the average smvi between the unable and able subpopulations is eliminated through categorical spending *and* there are resources left over. However, as illustrated in Result 1(i) and Corollary 1, the conditions under which it is optimal to provide a universal benefit change with classification errors. In particular, Type I errors guarantee the provision of a universal benefit. The purpose of this section is to also demonstrate that whenever classification errors are made inequality in the average smvi will remain at the optimum.

Formally, let inequality in the average smvi between the unable and able subpopulations be given by:

$$\begin{aligned}\delta = & \underbrace{[(1-p_I)u'(B+C)+p_I u'(B)]}_{\text{average smvi (unable)}} - \\ & - \underbrace{\left\{\int_0^\infty v_M dF(n) + p_{II}\int_0^{\bar{\bar{n}}} [u'(B+C)-v_M]\, dF(n)\right\}}_{\text{average smvi (able)}}\end{aligned} \tag{4.23}$$

Result 2: *Whenever $p_I > 0$ and/or $p_{II} > 0$ we have $\delta > 0$ at the optimum.*

Proof: We proceed to check that $\delta > 0$ at the optimum in three cases.

(i) ($p_I > 0$, $p_{II} = 0$)

Given that $p_I > 0$ the expression in (4.19) must hold with equality such that:[13]

$$\frac{\theta p_I u'(\hat{B}) + (1-\theta)\int_0^\infty v_M dF(n)}{\theta p_I + (1-\theta)} = u'(\hat{B}+\hat{C})$$

which directly implies $u'(B) > u'(\hat{B}+\hat{C}) > \int_0^\infty v_M dF(n)$, in turn implying $\delta > 0$.

(ii) ($p_I = 0$, $p_{II} > 0$)

In this case $\delta = u'(B+C)[1 - F(\bar{\bar{n}})p_{II}] - [\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} v_M dF(n)]$ and consequently $\delta > 0$ at the optimum if:

$$u'(\hat{B}+\hat{C}) > \frac{\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} v_M dF(n)}{1 - F(\bar{\bar{n}})p_{II}}$$

which is guaranteed to hold by expression (4.19).

(iii) ($p_I > 0$, $p_{II} > 0$)

In this case $\delta > 0$ at the optimum if:

$$u'(\hat{B}+\hat{C}) > \frac{\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} v_M dF(n) - p_I u'(B)}{1 - F(\bar{\bar{n}})p_{II} - p_I}$$

From (4.19) it follows that a sufficient condition for $\delta > 0$ is:

$$\frac{\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} v_M dF(n) - p_I u'(B)}{1 - F(\bar{\bar{n}})p_{II} - p_I} < \frac{\theta p_I u'(B) + (1-\theta)[\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} dF(n)]}{\theta p_I + (1-\theta)[1 - F(\bar{\bar{n}})p_{II}]}$$

which requires:

$$u'(\hat{B}) > \frac{\int_0^\infty v_M dF(n) - p_{II}\int_0^{\bar{\bar{n}}} v_M dF(n)}{1 - F(\bar{\bar{n}})p_{II}}$$

When receiving the same benefit income (i.e. B), the smvi of an unable individual will always be at least as great as that of an able individual and so this condition must

[13] Note that when $p_{II} = 0$ we have $-(1 + \partial C_F/\partial B) = (1-\chi)/\chi$ and $(1-\chi) = \theta p_I + (1-\theta)$.

hold. Q.E.D.

4.3 Numerical Simulations

Recall that (4.21) is an implicit expression for the optimal tax rate: gross earnings, indirect utility and, importantly, the critical productivity $\bar{\bar{n}}$ are all functions of the (net-of) tax rate. We therefore turn to numerical methods to gain insights into how the propensities to make Type I and Type II errors affect the optimal tax rate.

To do so, let individual preferences over consumption and labour be given by the frequently employed isoelastic form:

$$u(x, H) = \log\left(x - \alpha\frac{H^{1+k}}{1+k}\right) \tag{4.24}$$

where $1/k$ is a constant elasticity of labour supply with respect to the net wage rate (see also Atkinson, 1990, 1995; Saez, 2001). In line with the more general preferences in (4.1), the resulting optimal labour supply function $H^* = [n(1-t)/\alpha]^{1/k}$ is independent of unearned income and gives rise to the indirect utility function:

$$v[n(1-t), M] = \log\left\{[n(1-t)]^{\frac{1+k}{k}}\alpha^{-\frac{1}{k}}\left(\frac{k}{1+k}\right) + M\right\}$$

From this it can be readily established that the critical productivity $\bar{\bar{n}}$ at or below which able individuals apply for the categorical benefit is:

$$\bar{\bar{n}} \equiv \left\{\left(\frac{1+k}{k}\right)\alpha^{\frac{1}{k}}C\right\}^{\frac{k}{1+k}} /(1-t) \tag{4.25}$$

This is clearly increasing in both t and C.

In the simulations which follow we assume: $n \sim \ln\mathcal{N}(\mu = -1, \sigma = 0.39)$; $\alpha = 1$; $k \in \{1, 2, 3\}$; $\theta = 0.1$. The assumption that productivities are lognormally distributed with mean of $\log n$ set at -1 and standard deviation of $\log n$ set at 0.39 is frequently adopted in the literature (see Immonen et al., 1998; Mirrlees, 1971; Stern, 1976; Viard, 2001a,b). Further, the key analyses of categorical transfers within the optimal income

Table 4.1: Numerical Results

	$k = 1$					$k = 2$					$k = 3$				
p_{II}	$\hat{t}$	$\hat{B}$	$\hat{C}$	$F(\bar{\bar{n}})$	V	$\hat{t}$	$\hat{B}$	$\hat{C}$	$F(\bar{\bar{n}})$	V	$\hat{t}$	$\hat{B}$	$\hat{C}$	$F(\bar{\bar{n}})$	V
(a) $p_I = 0.0$															
0	0.231	0.026	0.038	0.473	-2.667	0.281	0.048	0.087	0.469	-1.944	0.320	0.065	0.114	0.473	-1.672
0.1	0.236	0.028	0.018	0.152	-2.676	0.291	0.054	0.042	0.099	-1.953	0.335	0.073	0.055	0.076	-1.680
0.2	0.239	0.028	0.013	0.083	-2.678	0.297	0.055	0.034	0.052	-1.955	0.341	0.076	0.045	0.039	-1.682
0.3	0.241	0.029	0.011	0.056	-2.680	0.300	0.056	0.030	0.035	-1.956	0.345	0.077	0.041	0.026	-1.683
0.4	0.243	0.029	0.010	0.042	-2.680	0.302	0.057	0.028	0.026	-1.957	0.347	0.078	0.038	0.019	-1.684
0.5	0.244	0.029	0.009	0.034	-2.681	0.304	0.057	0.026	0.021	-1.957	0.349	0.078	0.036	0.015	-1.684
(b) $p_I = 0.1$															
0	0.235	0.026	0.037	0.462	-2.670	0.291	0.051	0.083	0.453	-1.947	0.334	0.069	0.109	0.455	-1.676
0.1	0.239	0.028	0.016	0.134	-2.678	0.300	0.056	0.039	0.084	-1.955	0.346	0.076	0.051	0.064	-1.683
0.2	0.242	0.029	0.012	0.072	-2.680	0.304	0.057	0.032	0.044	-1.957	0.351	0.078	0.043	0.033	-1.684
0.3	0.244	0.029	0.011	0.049	-2.681	0.306	0.058	0.029	0.030	-1.958	0.353	0.079	0.039	0.022	-1.685
0.4	0.245	0.030	0.010	0.037	-2.682	0.308	0.058	0.026	0.022	-1.958	0.355	0.080	0.036	0.016	-1.685
0.5	0.246	0.030	0.009	0.029	-2.682	0.309	0.059	0.025	0.018	-1.959	0.357	0.080	0.035	0.013	-1.686
(c) $p_I = 0.2$															
0	0.239	0.027	0.036	0.452	-2.673	0.300	0.053	0.080	0.438	-1.951	0.345	0.073	0.104	0.439	-1.679
0.1	0.243	0.029	0.015	0.117	-2.680	0.307	0.058	0.036	0.071	-1.958	0.355	0.079	0.048	0.054	-1.685
0.2	0.245	0.029	0.012	0.063	-2.682	0.310	0.059	0.030	0.037	-1.959	0.358	0.080	0.040	0.028	-1.686
0.3	0.247	0.030	0.010	0.042	-2.682	0.312	0.059	0.027	0.025	-1.960	0.361	0.081	0.037	0.019	-1.687
0.4	0.248	0.030	0.009	0.032	-2.683	0.313	0.060	0.025	0.019	-1.960	0.362	0.082	0.035	0.014	-1.687
0.5	0.248	0.030	0.008	0.025	-2.683	0.314	0.060	0.024	0.015	-1.961	0.363	0.082	0.033	0.011	-1.687

Notes: V denotes welfare evaluated at the optimum choices. The numerical results were generated in Python 2.7 (Enthought Canopy 1.5.1 (64 bit)) and also replicated in Mathematica 10 Student Edition. The numerical code is situated in the Appendix.

Table 4.1: Numerical Results *Continued*

	$k=1$					$k=2$					$k=3$				
p_{II}	$\hat{t}$	$\hat{B}$	$\hat{C}$	$F(\bar{\bar{n}})$	V	$\hat{t}$	$\hat{B}$	$\hat{C}$	$F(\bar{\bar{n}})$	V	$\hat{t}$	$\hat{B}$	$\hat{C}$	$F(\bar{\bar{n}})$	V
(d) $p_I = 0.3$															
0	0.242	0.028	0.035	0.442	-2.675	0.307	0.055	0.077	0.424	-1.954	0.354	0.076	0.100	0.424	-1.682
0.1	0.246	0.029	0.014	0.100	-2.682	0.313	0.059	0.034	0.060	-1.960	0.362	0.081	0.045	0.045	-1.687
0.2	0.248	0.030	0.011	0.053	-2.683	0.316	0.060	0.028	0.031	-1.961	0.365	0.082	0.038	0.023	-1.688
0.3	0.249	0.030	0.009	0.036	-2.684	0.317	0.060	0.026	0.021	-1.961	0.367	0.083	0.035	0.015	-1.688
0.4	0.250	0.030	0.008	0.027	-2.684	0.318	0.061	0.024	0.016	-1.962	0.368	0.083	0.033	0.012	-1.689
0.5	0.250	0.030	0.008	0.021	-2.684	0.319	0.061	0.023	0.013	-1.962	0.369	0.084	0.031	0.009	-1.689
(e) $p_I = 0.4$															
0	0.245	0.029	0.034	0.432	-2.678	0.313	0.057	0.074	0.411	-1.957	0.362	0.079	0.096	0.410	-1.684
0.1	0.249	0.030	0.013	0.085	-2.683	0.318	0.060	0.032	0.050	-1.961	0.369	0.083	0.042	0.037	-1.689
0.2	0.250	0.030	0.010	0.045	-2.684	0.320	0.061	0.027	0.026	-1.962	0.371	0.084	0.036	0.019	-1.689
0.3	0.251	0.030	0.009	0.030	-2.685	0.322	0.061	0.024	0.017	-1.963	0.372	0.084	0.033	0.013	-1.690
0.4	0.252	0.031	0.008	0.022	-2.685	0.322	0.062	0.023	0.013	-1.963	0.373	0.085	0.031	0.010	-1.690
0.5	0.252	0.031	0.007	0.018	-2.685	0.323	0.062	0.021	0.010	-1.963	0.374	0.085	0.030	0.008	-1.690
(f) $p_I = 0.5$															
0	0.248	0.029	0.033	0.423	-2.680	0.319	0.059	0.072	0.399	-1.959	0.369	0.081	0.093	0.397	-1.687
0.1	0.251	0.030	0.012	0.070	-2.685	0.323	0.062	0.030	0.040	-1.963	0.375	0.085	0.040	0.030	-1.690
0.2	0.252	0.031	0.009	0.036	-2.685	0.325	0.062	0.025	0.021	-1.964	0.377	0.085	0.034	0.015	-1.691
0.3	0.253	0.031	0.008	0.024	-2.686	0.326	0.062	0.023	0.014	-1.964	0.378	0.086	0.031	0.010	-1.691
0.4	0.253	0.031	0.007	0.018	-2.686	0.326	0.063	0.021	0.011	-1.964	0.378	0.086	0.030	0.008	-1.691
0.5	0.254	0.031	0.007	0.015	-2.686	0.327	0.063	0.020	0.008	-1.965	0.379	0.086	0.028	0.006	-1.691

tax framework typically employ the lognormal distribution.[14] The choice of $\alpha = 1$ gives rise to $H^* = 0.63$ for the average productivity ($n = 0.3969$) individual when $k = 2$. Finally, the choices of k correspond to labour elasticities of 1,1/2 and 1/3 respectively: these are broadly consistent with those adopted by other authors (see, for example Atkinson, 1995; Saez, 2001). For simplicity, it is assumed that the government has no revenue generating commitments and thus $R = 0$. Taxation is thus purely redistributive.

The simulation procedure is as follows: for a given value of $p_I \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ we systematically increase p_{II} in intervals of 0.025 from 0 to 1; at each stage calculating the optimal tax rate and benefit levels. Table 4.1 reports the resulting optima over the stated values of k and p_I, for the cases where $p_{II} \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$. We discuss the results below.

A *ceteris paribus* increase in p_I tends to increase the optimal tax rate; increase the universal benefit; but lower the categorical benefit. Whilst the increase in the tax rate will act to increase $\bar{\bar{n}}$, the fall in the categorical benefit acts in the opposite direction and is sufficiently large that $\bar{\bar{n}}$ falls with p_I in all cases. Next, the effect of an increase in p_{II} on the optimal choices is more pronounced but the directions of movement remain the same. Indeed, a *ceteris paribus* increase in p_{II} tends to increase the optimal tax rate; increase the universal benefit; but lower the categorical benefit. In particular, notice that an increase in p_{II} from 0 to 0.1 (for any p_I) results in all cases in a reduction of over 50% in the categorical benefit, which in turn reduces the proportion of able individuals who would choose to apply for the categorical benefit (were $p_{II} > 0$) from around 50% to less than 20%. For all the examples given in Table 4.1 the value function of social welfare (i.e. welfare at the optimum choices given exogenous error propensities) is falling in the propensity to make either error type.

Finally, notice that $\hat{C}$ is increasing in k in all cases: this arises because an increase in k corresponds to a reduction in the elasticity of labour supply, which in turn generates a reduction in the number of able individuals who would choose to apply for C. Formally, we have:

[14] Detailed numerical results under the alternative assumption that productivity is exponentially distributed are available from the author upon request.

$$\frac{\partial \bar{\bar{n}}}{\partial k} = \left[\frac{\alpha^{\frac{1}{k}}(1+k)C}{k(1-t)^{\frac{1+k}{k}}}\right]^{\frac{k}{1+k}} \left\{\frac{[k^2-(1+k)k-(1+k)\log a]}{(1+k)^2 k} + \frac{1}{(1+k)^2}\log\left[\frac{\alpha^{\frac{1}{k}}(1+k)C}{k}\right]\right\}$$

When evaluated at $\alpha = 1$ the term $\log(\alpha)$ drops out such that it will unambiguously hold that $\partial\bar{\bar{n}}/\partial k < 0$ so long as $C(1+k)/k < 1$; which is certainly the case from Table 4.1.

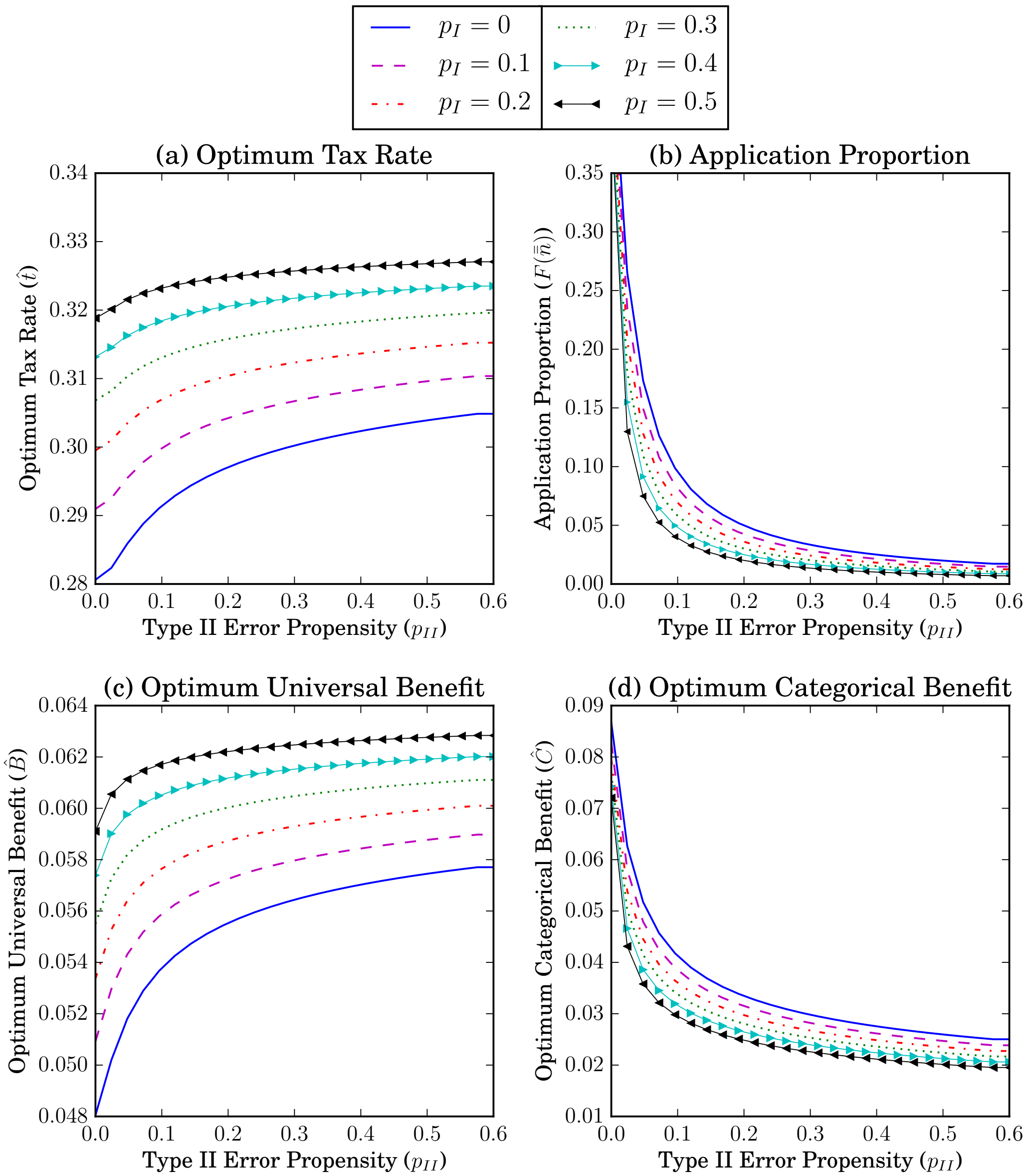
As discussed, a key observation from Table 4.1 is that an increase in p_{II} above zero induces a large proportional fall in the optimal categorical benefit and, in turn, in the number of able individuals who choose to apply for the categorical benefit. It is therefore important to note that in all considered cases the optimum benefit functions are smooth and continuous. Figure 4.2 provides useful examples for the case where $k = 2$. The figure graphically illustrates how $\hat{t}$, $\hat{B}$, $\hat{C}$ and $F(\bar{\bar{n}})$ change with the propensity to make classification errors. On the horizontal axis in each subplot p_{II} is varied from 0 to 0.6 in discrete increments of 0.025, whilst the different curves within each figure are generated for a different value of p_I.

Finally, Figure 4.3 provides a useful check that the simulated optimal tax rates are in fact global optima and that the welfare function is 'well-behaved' with classification errors. The figure illustrates how welfare changes with the tax rate, for different values of p_I and p_{II}. In each of the six subplots the tax rate is varied on the horizontal axis and welfare is given on the vertical axis. The different subplots are generated for successively higher values of p_I, whilst the different curves within each subplot are generated for a different value of p_{II}. For any given tax rate the optimal benefit levels were chosen and then substituted into the welfare function. As can be seen, in each case there is a unique optimal choice of t.

4.4 Concluding Remarks

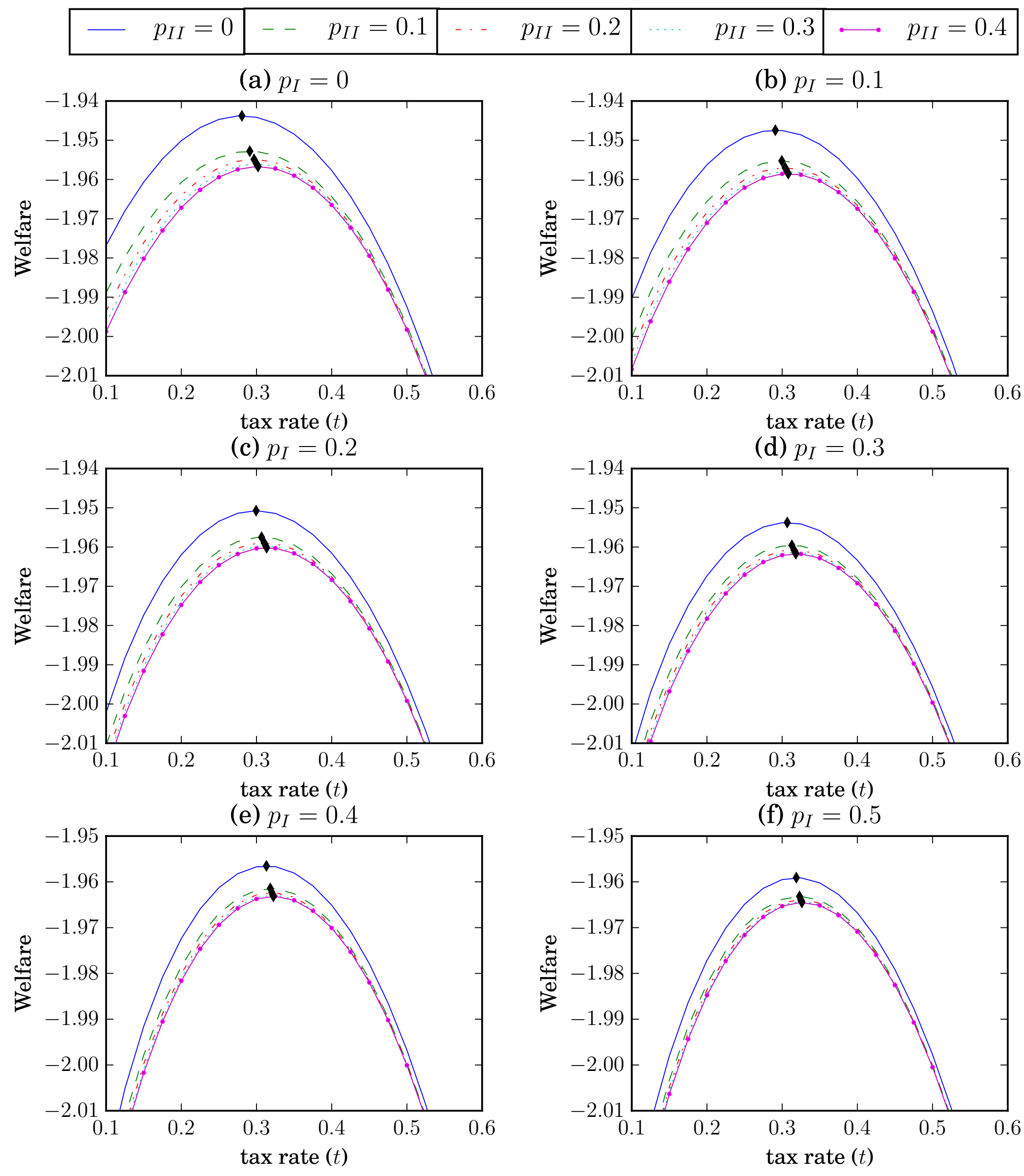
As outlined in the introduction to this chapter, many of the key results from the analysis of categorical transfers in a linear income tax framework are derived under the strong assumption that the government can perfectly identify the categorical group to which an individual belongs. In this case there is a well-defined ordering of priorities

Figure 4.2: Graphical Illustration of Optima for $k = 2$



Notes: Subplots (a), (c) and (d) illustrate how the optimal choices $(\hat{t}, \hat{B}, \hat{C})$ change with the propensity to make classification errors; whilst subplot (b) illustrates how these changes in the choice variables affect the number of individuals from the able subpopulation who choose to apply for the categorical benefit. The different curves within each figure are generated for a different value of $p_I \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$.

Figure 4.3: Welfare and the Tax Rate ($k = 2$)



Notes: This figure serves as a useful check that the simulated optimal tax rates correspond to global optima. Within each subplot the diamond marker (i.e. $\diamond$) denotes welfare at the optimal tax rate, $\hat{t}$.

between categorical (targeted) and universal benefits. Welfare provision should be purely categorical up to the point that inequality in the average smvi across categorical groups is eliminated. If it is optimal to generate enough tax revenue to achieve this, *and* there are resources left over, an unconditional universal benefit will be provided to all individuals in society.

In reality, however, an individual's true categorical status (e.g. disability, involuntary unemployment) may be difficult to identify and this will inevitably lead to misclassifications; in turn generating both Type I (false rejection) and Type II (false award) errors. These errors are likely to change the ordering of priorities between targeted and universal welfare provision, and in turn have implications for the equity and efficiency considerations that characterise the optimal tax rate. The implications for the optimal tax rate have been the subject matter of this chapter.

This chapter considered a framework where a fraction of the population is unable to work, whilst the remaining fraction are able to work but differ continuously in their productivity. The government operates a tax-benefit system comprising (i) a linear income tax on all earned income; (ii) a tax-free universal benefit received unconditionally by all individuals in society; and (iii) a tax-free categorical benefit that is targeted at the unable, albeit imperfectly. The categorical benefit is administered with both Type I and Type II classification errors. Any able individual who applies for the categorical benefit and is incorrectly awarded it is not allowed to work. This condition is fully enforced, such that the only able individuals who choose to apply for the categorical benefit are those of lower productivities, and thus those for whom the opportunity cost of not working is insufficiently high. An important implication is that the critical productivity at or below which individuals choose to apply for the categorical benefit will be, *ceteris paribus*, an increasing function of the tax rate.

The key contribution of this chapter has been to derive an expression for the optimal linear tax rate when the categorical benefit is administered with two-sided classification errors. In setting the tax rate, the government must now factor in the additional equity and efficiency considerations that arise through classification errors. In particular, Type II errors generate conflicting effects in both the equity and efficiency dimensions. In the equity dimension, Type II errors (i) mean that some able individuals of low productivity - who the government would not wish to tax highly - are not working, and this acts to raise the tax rate; but also (ii) play a redistributive role through 'leaking' the categorical benefit to lower productivity individuals, which may in turn reduce the need to redistribute through the universal benefit and thus act to lower the tax rate. In the efficiency dimension (i) individuals who would have worked and adjusted their earnings in response to the tax rate now do not work due to Type II errors, which acts to raise the tax rate; but (ii) the number of individuals who choose to apply for the categorical benefit is, *ceteris paribus*, increasing in the tax rate, which acts to lower the tax rate so as to avoid the dual efficiency concerns of reducing the tax base whilst also paying more individuals the categorical benefit.

Numerical simulations suggest that an increase in the propensity to make either error type (i) increases the optimal tax rate; (ii) increases the optimal universal benefit; but (iii) lowers the optimal categorical benefit. For any given propensity to make Type I errors, an increase from zero in the propensity to make Type II errors generates a large proportional fall in the categorical benefit size, thus reducing the proportion of the able subpopulation that would choose to apply for it.

Appendix A Derivations and Proofs

Derivatives of the function $\bar{\bar{n}}$.

Differentiating both sides of (4.2) in the main text with respect to $(1-t)$ and C, respectively, yields:

$$(1-t): \quad \left[\frac{\partial \bar{\bar{n}}}{\partial(1-t)} \cdot (1-t) + \bar{\bar{n}}\right] \left\{ H^* + \frac{dH^*}{d\omega} \underbrace{\left[\bar{\bar{n}}(1-t) - g'(H^*)\right]}_{=0} \right\} = 0$$

$$(C): \quad \frac{\partial \bar{\bar{n}}}{\partial C} \cdot (1-t) \left\{ H^* + \frac{dH^*}{d\omega} \underbrace{\left[\bar{\bar{n}}(1-t) - g'(H^*)\right]}_{=0} \right\} - 1 = 0$$

from which it directly follows that $\partial\bar{\bar{n}}/\partial(1-t) = -\bar{\bar{n}}/(1-t) < 0$, whilst $\partial\bar{\bar{n}}/\partial C = 1/[(1-t)H^*] > 0$. Taking second derivatives then yields:

$$\frac{\partial^2 \bar{\bar{n}}}{\partial(1-t)^2} = -\left[\frac{(1-t)\frac{\partial \bar{\bar{n}}}{\partial(1-t)} - \bar{\bar{n}}}{(1-t)^2}\right] = \frac{2\bar{\bar{n}}}{(1-t)^2} > 0$$

$$\frac{\partial^2 \bar{\bar{n}}}{\partial C^2} = -\frac{1}{(1-t)(H^*)^3} \cdot \frac{dH^*}{d\omega} = -\frac{1}{(1-t)(H^*)^3 g''(H^*)} < 0$$

Proof of Result 1(i).

From the optimisation problem in (4.17) the first-order-conditions (henceforth FOCs) characterising the optimal benefits $\hat{B}$ and $\hat{C}$ are:

$$(B): \quad \chi u'\left(\hat{B} + \hat{C}\right) + \sigma(\hat{t}, \hat{B}, \hat{C}) \leq \hat{\lambda} \quad ; \quad \hat{B} \geq 0 \tag{A.1}$$

$$(C): \ \chi u'\left(\hat{B} + \hat{C}\right) \leq \hat{\lambda}\left(\chi + \frac{\partial \chi}{\partial C}\hat{C} - \hat{t} \cdot \frac{\partial Z}{\partial \bar{\bar{n}}}\frac{\partial \bar{\bar{n}}}{\partial C}\right) \ ; \ \hat{C} \geq 0$$

$$\Rightarrow u'\left(\hat{B} + \hat{C}\right) \leq \hat{\lambda}\left\{1 + \frac{(1-\theta) f(\bar{\bar{n}}) p_{II} \frac{\partial \bar{\bar{n}}}{\partial C}\left[\hat{C} + \hat{t}y|_{n=\bar{\bar{n}}}\right]}{\chi}\right\} \ ; \ \hat{C} \geq 0 \tag{A.2}$$

where the pairs of inequalities hold with complementary slackness. Notice that in writing (A.2) we have substituted in both $\frac{\partial \chi}{\partial C}C = (1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial \bar{\bar{n}}}{\partial C}C$ and $\frac{\partial Z}{\partial \bar{\bar{n}}}\frac{\partial \bar{\bar{n}}}{\partial C} = -(1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial \bar{\bar{n}}}{\partial C}y|_{n=\bar{\bar{n}}}$.

We briefly discuss both of these FOCs in turn. Given that incentives to apply for the categorical benefit are unaffected by the level of universal benefit, the FOC characterising $\hat{B}$ is simple. The left side of (A.1) is the average smvi over the entire population and thus corresponds to the marginal benefit of increasing B. The right side of (A.1) is the marginal cost - in welfare units - of increasing B: this cost simply enters as unity (i.e. the shadow price multiplied by 1) because *every* individual is paid a higher universal benefit. As illustrated by the complementary slackness condition, whether or not it will be optimal to provide a universal benefit will depend on whether or not the average smvi over the population equates with or exceeds the shadow price of public expenditure (λ) when $B = 0$.

The shadow price is in turn pinned down by the FOC characterising $\hat{C}$. The left side of (A.2) is the average smvi of a categorial recipient.[15] The right side captures - in welfare units - the average marginal cost associated with increasing the categorical benefit. This marginal cost is composed of two effects. First, a marginal increase in the categorical benefit means that the government pays each *existing* recipient a higher benefit and so this effect simply enters as unity. Second, a marginal increase in the categorical benefit induces an *additional* able individual to apply for the categorical benefit, where they did not prior to the increase. This second effect generates both an expenditure cost and a loss in tax revenue, because: (i) the government pays an additional individual the categorical benefit where it did not before (this effect enters as $\hat{C}$); and (ii) the new recipient stops working as a result of the fully enforced no-work requirement (this effect enters as $\hat{t}y|_{n=\bar{\bar{n}}}$). As this is an average marginal cost, these second effects enter as a fraction of the number of existing recipients (i.e. χ).

With respect to the FOCs for B and C in (A.1) and (A.2) we test two hypotheses: (i) $\hat{B} > 0$, $\hat{C} = 0$; and (ii) $\hat{B} = 0$, $\hat{C} > 0$.

(i) ($\hat{B} > 0$, $\hat{C} = 0$)

[15] All recipients of the categorical benefit of course have the same marginal welfare because they do not work.

If we set $C = 0$ then $\bar{\bar{n}} = 0$ and so:

$$\chi = \theta(1 - p_I) \;\; ; \quad \sigma = \theta p_I u'(B, 1) + (1 - \theta) \int_0^\infty v_M dF(n)$$

The FOCs in (A.1) and (A.2) thus become:

$$\theta u'(B) + (1 - \theta) \int_0^\infty v_M[n(1 - t), B] dF(n) = \lambda \;\; ,$$
$$u'(B) \leq \lambda \;\; .$$

Taken together, however, this implies that:

$$\theta u'(B) + (1 - \theta) \int_0^\infty v_M dF(n) \geq u'(B) \Rightarrow \int_0^\infty v_M[n(1 - t), B] dF(n) \geq u'(B)$$

This must be a contradicton because for any interior tax rate able individuals are better-off than unable individuals and thus their average smvi must be lower than that of unable individuals. The assertion that $\hat{C} = 0$ is therefore **false** and, instead, at all levels of discriminatory power we have $\hat{C} > 0$.

Given that $\hat{C} > 0$ the FOC in (A.2) must hold with equality such that:

$$\hat{\lambda} = u'(\hat{B} + \hat{C}) \cdot \left\{ \frac{\chi}{\chi + (1 - \theta) f(\bar{\bar{n}}) p_{II} \frac{\partial \bar{\bar{n}}}{\partial C} [\hat{C} + \hat{t} y|_{n = \bar{\bar{n}}}]} \right\} \tag{A.3}$$

Substituting (A.3) into (A.1) then gives:

$$\begin{aligned} \sigma(\hat{t}, \hat{B}, \hat{C}) &\leq u'(B + C) \cdot \chi \left\{ \frac{1}{\chi + (1 - \theta) f(\bar{\bar{n}}) p_{II} \frac{\partial \bar{\bar{n}}}{\partial C} [\hat{C} + \hat{t} y|_{n = \bar{\bar{n}}}]} - 1 \right\} \quad ; \;\; \hat{B} \geq 0 \\ &= u'(B + C) \cdot \chi \left\{ \frac{(1 - \chi) - (1 - \theta) f(\bar{\bar{n}}) p_{II} \frac{\partial \bar{\bar{n}}}{\partial C} [\hat{C} + \hat{t} y|_{n = \bar{\bar{n}}}]}{\chi + (1 - \theta) f(\bar{\bar{n}}) p_{II} \frac{\partial \bar{\bar{n}}}{\partial C} [\hat{C} + \hat{t} y|_{n = \bar{\bar{n}}}]} \right\} \;\; ; \;\; \hat{B} \geq 0 \end{aligned} \tag{A.4}$$

Finally, it is straightforward to establish that the term within curly braces corresponds to $-(1 + \frac{\partial C_F}{\partial B})$, where C_F is as defined in (4.10) in the main text.

(ii) **(**$\hat{B} = 0$ **,** $\hat{C} > 0$**)**

Given that $\lim_{x \to 0} u'(x) = +\infty$ it is straightforward to see from (4.19) in the main text that $\hat{B} > 0$ whenever $p_I > 0$. Otherwise, it will only be optimal to provide a universal benefit if (i) the aggregate smvi of non-recipients exceeds (ii) the aggregate smvi of categorical recipients multiplied by the (negative of) the change in their total benefit income associated with an increase in the universal benefit, when evaluated at $B = 0$.

Proof of Result 1(ii)

It is straightforward to establish from (A.3) that $u'(\hat{B}+\hat{C}) > \hat{\lambda}$ at the optimum. This arises because the term within curly braces is less than unity. Now, from the definition of $\bar{\bar{n}}$ in (4.2) an able individual with $n \in [0, \bar{\bar{n}})$ has $u[n(1-t)H^* + B - g(H^*)] < u(B+C)$ and so:

$$v_M[n(1-t), B] = u'[n(1-t)H^* + B - g(H^*)] > u'(B+C) \ ; \forall \, n \in [0, \bar{\bar{n}}) \tag{A.5}$$

Putting $u'(\hat{B}+\hat{C}) > \hat{\lambda}$ and (A.5) together thus gives:

$$\begin{aligned} v_M[n(1-\hat{t}), \hat{B}] &> u'\left(\hat{B}+\hat{C}\right) > \hat{\lambda} \ ; \ \forall \, n \in [0, \bar{\bar{n}}) \ , \\ v_M[\bar{\bar{n}}(1-\hat{t}), \hat{B}] &= u'\left(\hat{B}+\hat{C}\right) > \hat{\lambda} \ ; \ \text{if } n = \bar{\bar{n}} \end{aligned} \tag{A.6}$$

This assists us in interpreting the optimal tax expression derived below.

Proof of Result 1(iii) (Optimal Tax Expression)

From the optimisation problem in (4.17), the FOC characterising an interior solution for the optimal tax rate is given by:

$$\begin{aligned} (t): \quad & (1-\theta)\left\{\int_0^{\infty} -nv_\omega dF(n) + p_{II}\int_0^{\bar{\bar{n}}} nv_\omega dF(n)\right\} \\ &= \hat{\lambda}\left\{\frac{\partial \chi}{\partial t}\hat{C} - Z + \hat{t}\left[\frac{\partial Z}{\partial(1-t)} + \frac{\partial Z}{\partial \bar{\bar{n}}}\cdot\frac{\partial \bar{\bar{n}}}{\partial(1-t)}\right]\right\} \end{aligned}$$

Note that in writing this FOC we have used the identity $u(B+C) \equiv v[\bar{\bar{n}}(1-t), B]$ when differentiating the integral limits in the welfare function. Substituting in both Roy's identity (i.e. $v_\omega = v_M(y/n)$ and the definition of aggregate earnings Z then gives:

$$\begin{aligned}&\hat{\lambda}(1-\theta)\left\{\int_0^\infty y\left(1-\frac{v_M}{\hat{\lambda}}\right)dF(n) - p_{II}\int_0^{\bar{\bar{n}}} y\left(1-\frac{v_M}{\hat{\lambda}}\right)dF(n)\right\}\\ =&\hat{\lambda}\left\{\frac{\partial\chi}{\partial t}\hat{C} + \hat{t}\left[\frac{\partial Z}{\partial(1-t)} + \frac{\partial Z}{\partial\bar{\bar{n}}}\cdot\frac{\partial\bar{\bar{n}}}{\partial(1-t)}\right]\right\}\end{aligned}$$

To progress we recall that $\frac{\partial\chi}{\partial t}\hat{C} = -(1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial\bar{\bar{n}}}{\partial(1-t)}\hat{C}$ whilst $\frac{\partial Z}{\partial\bar{\bar{n}}}\frac{\partial\bar{\bar{n}}}{\partial(1-t)} = -(1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial\bar{\bar{n}}}{\partial(1-t)}y|_{n=\bar{\bar{n}}}$. Substituting in these terms and using the property that $\frac{\partial\bar{\bar{n}}}{\partial(1-t)}\cdot(1-t) = -\bar{\bar{n}} \Rightarrow \frac{\partial\bar{\bar{n}}}{\partial(1-t)} = t\cdot\frac{\partial\bar{\bar{n}}}{\partial(1-t)} - \bar{\bar{n}}$, gives:

$$\begin{aligned}&\frac{1}{\hat{\lambda}}(1-\theta)\left\{\int_0^\infty y(\hat{\lambda}-v_M)dF(n) - p_{II}\int_0^{\bar{\bar{n}}} y(\hat{\lambda}-v_M)dF(n) - \hat{\lambda}p_{II}\bar{\bar{n}}f(\bar{\bar{n}})\hat{C}\right\}\\ &= \hat{t}\left\{\frac{\partial Z}{\partial(1-t)} - (1-\theta)p_{II}f(\bar{\bar{n}})\frac{\partial\bar{\bar{n}}}{\partial(1-t)}\left[y|_{n=\bar{\bar{n}}} + \hat{C}\right]\right\}\end{aligned} \tag{A.7}$$

Substituting in the definition $\frac{\partial\bar{\bar{n}}}{\partial(1-t)} = -\bar{\bar{n}}/(1-t)$ and dividing both sides by $(1-t)$ yields:

$$\frac{\hat{t}}{1-\hat{t}} = \frac{\int_0^\infty y\left(\hat{\lambda}-v_M\right)dF(n) - p_{II}\int_0^{\bar{\bar{n}}} y\left(\hat{\lambda}-v_M\right)dF(n) - \hat{\lambda}F(\bar{\bar{n}})p_{II}\left(\frac{\bar{\bar{n}}}{F(\bar{\bar{n}})}f(\bar{\bar{n}})\right)\hat{C}}{\hat{\lambda}\left\{\bar{Z}\cdot\left(\frac{(1-t)}{Z}\frac{\partial\bar{Z}}{\partial(1-t)}\right) + p_{II}F(\bar{\bar{n}})\cdot\left(\frac{\bar{\bar{n}}}{F(\bar{\bar{n}})}f(\bar{\bar{n}})\right)\left[y|_{n=\bar{\bar{n}}} + \hat{C}\right]\right\}} \tag{A.8}$$

where $\bar{Z} = Z/(1-\theta)$ is simply the average gross earnings over the able subpopulation. Substituting in $\mathcal{E}_{\bar{Z}} \equiv \frac{(1-t)}{\bar{Z}}\frac{\partial\bar{Z}}{\partial(1-t)}$ and $\mathcal{E}^f \equiv nf(n)/F(n)$ then gives the expression in the main text.

Appendix B Numerical Code

```python

from __future__ import division
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize
from numpy import vectorize
from scipy.stats import lognorm
from scipy.integrate import quad
from scipy.optimize import brentq
from scipy.optimize import newton

plt.rc('text',usetex=True)
plt.rc('font',family='serif')

#Python Code run using Enthought Canopy 1.5.1 (64bit)
    on MacBook
#Pro 2.3GHz Intel Core i7

#OPTIMAL TAXATION WITH CLASSIFICATION ERRORS CODE
#------------------------------------------------------

#1) PRELIMINARIES
#......................................................

#In what follows: 'n' denotes individual productivity;
    't' is the tax rate on earned income; 'm' is
    unearned income; k is the constant elasticity of
    labour supply with respect to the net wage rate;
    'a' is a constant; 'p1' is the Type I error
    propensity; and 'p2' is the Type II error
    propensity. Throughout the lower case variables 'b'
    and 'c' denote the universal and categorical
    benefits, respectively.

#Optimal labour supply is given by
def hopt(n,t,a,k):
    return (n*(1-t)/a)**(1/k)

#For the purposes of aggregating individual earnings,
    we define the below function which multiplies
    optimal labour supply by the distribution function.
def ypdf(n,t,a,k,mu,sigma):
    return n*hopt(n,t,a,k)*lognorm.pdf(n,sigma,
    scale=np.exp(mu))

#Average gross earnings over workers in the case where
    all able individuals work are thus given by:
def ybar(t,a,k,mu,sigma):
    return quad(ypdf,0,100,args=(t,a,k,mu,sigma))[0]

#The critical productivity at or below which an able
    individual chooses to apply for the categorical
    benefit is defined by:
def ndbar(t,c,a,k):
```

```
    if c<=0:
        return 0
    else:
        return ((((1+k)/k)*(a**(1/k))*c)**(k/(1+k)))/(1-t)

#The utility of an individual who does not work (unable,voluntary unemployed) is given by:
def u(m):
    if m<=0:
        return -1e100
    else:
        return np.log(m)
#Notice that we specify a value for the limiting case where m=0, so as to avoid the program trying to evaluate log(0)=-inf.

#The indirect utility of a worker is given by:
def v(n,t,m,a,k):
    return np.log(((n*(1-t))**((1+k)/k))*
    (a**(-1/k))*(k/(1+k))+m)

#For the purposes of aggregating individual indirect utility over the productivity continuum we define the below function which multiplies indirect utility by the distribution pdf.
def vpdf(n,t,m,a,k,mu,sigma):
    return v(n,t,m,a,k)*lognorm.pdf(n,sigma,scale=np.exp(mu))

#Aggregate welfare over workers - for the counterfactual case where no able individual applies for the categorical benefit - is thus given by:
def wa(t,m,a,k,mu,sigma):
    return quad(vpdf,0,100,args=(t,m,a,k,mu,sigma))[0]

#The welfare gain to an able individual who is awarded the categorical benefit by Type II error (multiplied by the productivity pdf) is given by:
def vpdfdiff(n,t,b,c,a,k,mu,sigma):
    return (u(b+c)-v(n,t,b,a,k))*lognorm.pdf(n,sigma,scale=
    np.exp(mu))

#Aggregating up to the critical productivity 'ndbar' then gives:
def wadiff(t,b,c,a,k,mu,sigma):
    return quad(vpdfdiff,0,ndbar(t,c,a,k),
    args=(t,b,c,a,k,mu,sigma))[0]

#The below function integrates over the earnings of workers who would choose to apply for the categorical benefit.
def yapp(t,c,a,k,mu,sigma):
    return quad(ypdf,0,ndbar(t,c,a,k),args=(t,a,k,
    mu,sigma))[0]
```

```
#Aggregate earnings in the economy are thus given by:
def z(t,c,a,k,mu,sigma,theta,p2):
    return (1-theta)*(ybar(t,a,k,mu,sigma)-
    p2*yapp(t,c,a,k,mu,sigma))

#The fraction of the able subpopulation that chooses to
    apply for the categorical benefit is given by:
def appprop(t,c,a,k,mu,sigma):
    return
        lognorm.cdf(ndbar(t,c,a,k),sigma,scale=np.exp(mu))


#2)REVENUE MAXIMISING (OR 'LAFFER') TAX RATE
#======================================================

#The benefit budget is given by:
def benefitbudget(c,t,b,a,k,mu,sigma,theta,p1,p2,r):
    return (b+(theta*(1-p1)+(1-theta)*p2*
    appprop(t,c,a,k,mu,sigma))*c
    -t*z(t,c,a,k,mu,sigma,theta,p2)-r)

#For any pair (t,b) and exogenous parameters
    (a,e,mu,sigma,theta, p1,p2,r) the critical value of
    c that exhausts the budget is:
def cf(t,b,a,k,mu,sigma,theta,p1,p2,r):
    if b+r>=t*z(t,0,a,k,mu,sigma,theta,p2):
        return 0
    else:
        return brentq(benefitbudget,0,2,
    args=(t,b,a,k,mu,sigma,theta,p1,p2,r))

#Tax revenue in the economy is:
def taxrevenue(t,b,a,k,mu,sigma,theta,p1,p2,r):
    return t*z(t,cf(t,b,a,k,mu,sigma,theta,p1,p2,r),
    a,k,mu,sigma,theta,p2)

#Let x=t. The optimisation problem that simulates the
    revenue maximising tax rate is given by:
def lafferresults(b,a,k,mu,sigma,theta,p1,p2,r,s1):
    def objlaffer(x,b,a,k,mu,sigma,theta,p1,p2,r):
        return
            -taxrevenue(x,b,a,k,mu,sigma,theta,p1,p2,r)
    res=minimize(objlaffer,s1,
    args=(b,a,k,mu,sigma,theta,p1,p2,r),
    method='SLSQP',options={'ftol':1e-10,'disp':False})
    return res.x[0],
        taxrevenue(res.x[0],b,a,k,mu,sigma,theta,p1,p2,r)

#To check that the output from 'lafferresults'
    corresponds to a global optima, we define the below
    function which plots tax revenue over successively
    higher values of t:
def taxrevenueplots(b,a,k,mu,sigma,theta,p1,p2,r):
    print "(p1=",p1,"p2=",p2,")","[",
    taxrevenue(0.1,b,a,k,mu,sigma,theta,p1,p2,r),",",
    t=0.1
    while t<0.89:
        t=t+0.01
```

```
        print
            taxrevenue(t,b,a,k,mu,sigma,theta,p1,p2,r),",",
    print
        taxrevenue(0.9,b,a,k,mu,sigma,theta,p1,p2,r),"]"


#3)OPTIMISATION PROBLEM
#======================================================

#Let x=(x[0],x[1],x[2])=(t,b,c).

#Benefit expenditure is defined by:
def expenditure(x,a,k,mu,sigma,theta,p1,p2):
    return x[1]+(theta*(1-p1)+(1-theta)*p2*
    appprop(x[0],x[2],a,k,mu,sigma))*x[2]

#Net tax revenue is given by:
def budget(x,a,k,mu,sigma,theta,p2,r):
    return x[0]*z(x[0],x[2],a,k,mu,sigma,theta,p2)-r

#The optimisation problem is:
def results(a,k,mu,sigma,theta,p1,p2,r,s1,s2):
    consfe=({'type':'eq',
'fun':lambda
    x:np.array([expenditure(x,a,k,mu,sigma,theta,p1,p2)
-budget(x,a,k,mu,sigma,theta,p2,r)])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])},
{'type':'ineq',
'fun':lambda x: np.array([x[2]])},
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([0.95-x[0]])})
    def obj(x,a,k,mu,sigma,theta,p1,p2):
        return -(theta*((1-p1)*u(x[1]+x[2])+p1*u(x[1]))+
        (1-theta)*(wa(x[0],x[1],a,k,mu,sigma)+
        p2*wadiff(x[0],x[1],x[2],a,k,mu,sigma)))
    res=minimize(obj,[s1,s1*z(s1,0,a,k,mu,sigma,
    theta,p2),s2],
    args=(a,k,mu,sigma,theta,p1,p2),
    bounds=((0,0.95),(0,None),(0,None)),
    constraints=consfe,
    method='SLSQP',options={'ftol':1e-10,'disp':False})
    return (res.x[0],res.x[1],res.x[2],
    appprop(res.x[0],res.x[2],a,k,mu,sigma),
    -obj(res.x,a,k,mu,sigma,theta,p1,p2))

#The dictionary of constraints 'consfe' specifies that:
    (i) benefit expenditure should equate with net tax
    revenue; (ii) the optimal benefits should be
    non-negative; and (iii) the tax rate must lie
    between zero and 0.95 (this upper bound avoids the
    program trying to search for an optimum near t=1).

#The function 'results' outputs the optimal tax rate
    (res.x[0]); the optimal universal benefit
    (res.x[1]); the optimal categorical benefit
```

```
    (res.x[2]); the number of able applicants at the
    optimum (appprop); and the value function (-obj).

#Note that s1 and s2 are starting in the search for the
    optimal tax rate and benefits.

#The below function loops 'results' over different
    values of p2'. To choose which result to output:
    set i=0 for res.x[0]; set i=1 for res.x[1]; set i=2
    for res.x[2]; set i=3 for appprop; set i=4 for -obj.

def resultplots(a,k,mu,sigma,theta,p1,r,s1,s2,i):
    print "[p1=",p1,"]",results(a,k,mu,sigma,theta,p1,0,
    r,s1,s2)[i],",",
    p2=0
    while p2<0.5975:
        p2=p2+0.0025
        print results(a,k,mu,sigma,theta,p1,
        p2,r,s1,s2)[i],",",
    print
        results(a,k,mu,sigma,theta,p1,0.6,r,s1,s2)[i],"]"

#OPTIMISATION PROBLEM FOR GIVEN VALUE OF t
#====================================================

#To check that the optimal tax rate returned from
    'results' corresponds to a global optima it useful
    to solve the optimisation problem taking t as
    given, and then determining the level of t for
    which welfare is maximised.

#Let x=(x[0],x[1])=(b,c).

#Benefit expenditure is now written as:
def expendituret(x,t,a,k,mu,sigma,theta,p1,p2):
    return (x[0]+(theta*(1-p1)+(1-theta)*
    p2*appprop(t,x[1],a,k,mu,sigma))*x[1])

#Whilst net tax revenue is:
def budgett(x,t,a,k,mu,sigma,theta,p2,r):
    return t*z(t,x[1],a,k,mu,sigma,theta,p2)-r

#The optimisation problem is:
def resultst(t,a,k,mu,sigma,theta,p1,p2,r,s1):
    consfe=({'type':'eq',
'fun':lambda x:np.array([expendituret(x,t,a,k,mu,
sigma,theta,p1,p2)
-budgett(x,t,a,k,mu,sigma,theta,p2,r)])},
{'type':'ineq',
'fun':lambda x: np.array([x[0]])},
{'type':'ineq',
'fun':lambda x: np.array([x[1]])})
    def objt(x,t,a,k,mu,sigma,theta,p1,p2):
        return -(theta*((1-p1)*u(x[0]+x[1])+p1*u(x[0]))
        +(1-theta)*(wa(t,x[0],a,k,mu,sigma)
        +p2*wadiff(t,x[0],x[1],a,k,mu,sigma)))
    res=minimize(objt,[t*z(t,s1,a,k,mu,sigma,theta,p2)-r,
    s1],
```

```
    args=(t,a,k,mu,sigma,theta,p1,p2),
    bounds=((0,None),(0,None)),
    constraints=consfe,
    method='SLSQP',
    options={'ftol':1e-10,'disp':False})
    return (res.x[0], res.x[1],
    appprop(t,res.x[1],a,k,mu,sigma),
    budgett(res.x,t,a,k,mu,sigma,theta,p2,r),
    -objt(res.x,t,a,k,mu,sigma,theta,p1,p2))

#The below function loops the function 'resultst' over
    successively higher values of the tax rate.
def resultstplots(a,k,mu,sigma,theta,p1,p2,r,s1,i):
    print
        "[p1=",p1,"p2=",p2,"]",resultst(0.1,a,k,mu,sigma,
    theta,p1,p2,r,s1)[i],",",
    t=0.1
    while t<0.6:
        t=t+0.0025
        print resultst(t,a,k,mu,sigma,theta,p1,
        p2,r,s1)[i],",",
```

2

Appendix C Optimal Tax Under No Enforcement

In the main text we assume that the *ex-post* no-work condition is fully enforced. It is also of interest, however, to derive the optimal tax rate for the polar case where the *ex-post* condition is not enforced. In this case all able individuals choose to apply for the categorical benefit. Further, under the preferences specified in (4.1) labour supply is independent of unearned income such that an individual's labour supply is independent of whether or not they receive the categorical benefit. Consequently, aggregate earnings in the economy are a function of solely the tax rate and given by:

$$Z(1-t;\theta) = (1-\theta)\int_0^\infty y(1-t)dF(n) \tag{C.1}$$

The optimisation problem becomes *much* simpler and is given by:

$$\begin{aligned} \max_{t,B,C} W = &\ \theta[(1-p_I)u(B+C) + p_I u(B)] \\ &+ (1-\theta)\int_0^\infty \langle p_{II} v[n(1-t), B+C] + (1-p_{II})v[n(1-t), B]\rangle\, dF(n) \\ \text{s.t.} \quad & B + [\theta(1-p_I) + (1-\theta)p_{II}]C = t\cdot Z(1-t;\theta) - R\ , \\ & t \in (0,1)\ ,\ B \geq 0\ ,\ C \geq 0\ . \end{aligned} \tag{C.2}$$

The resulting FOCs for B and C are:

$$\begin{aligned} (B):\ & \theta[(1-p_I)u'(\hat{B}+\hat{C}) + p_I u'(\hat{B})] \\ &+ (1-\theta)\int_0^\infty \left\langle p_{II} v_M[n(1-\hat{t}), \hat{B}+\hat{C}] + (1-p_{II})v_M[n(1-\hat{t}), \hat{B}]\right\rangle dF(n) \\ &\leq \hat{\lambda}\ ;\quad \hat{B} \geq 0 \end{aligned} \tag{C.3}$$

$$(C): \frac{\theta(1-p_I)u'(\hat{B}+\hat{C}) + (1-\theta)p_{II}\int_0^\infty v_M[n(1-\hat{t}), \hat{B}+\hat{C}]dF(n)}{\theta(1-p_I) + (1-\theta)p_{II}} \leq \hat{\lambda}\ ;\ \ \hat{C} \geq 0 \tag{C.4}$$

As usual, we test two hypotheses: (i) ($\hat{B} > 0$, $\hat{C} = 0$); and (ii) ($\hat{B} = 0$, $\hat{C} > 0$)

(i) ($\hat{B} > 0$, $\hat{C} = 0$)

Setting $\hat{C} = 0$ the two FOCs become:

$$(B): \ \theta u'(B) + (1-\theta)\int_0^\infty v_M[n(1-t), B]dF(n) = \lambda \tag{C.5}$$

$$(C): \ \frac{\theta(1-p_I)u'(B) + (1-\theta)p_{II}\int_0^\infty v_M[n(1-t), B]dF(n)}{\theta(1-p_I) + (1-\theta)p_{II}} \leq \lambda \tag{C.6}$$

Substituting (C.6) into (C.5) gives:

$$u'(B)[1-p_I-p_{II}] \leq \bar{s}[1-p_I-p_{II}] \tag{C.7}$$

which can clearly only hold when $p_I + p_{II} \geq 1$. So $\hat{C} > 0 \ \forall \ p_I + p_{II} < 1$; whilst $\hat{C} = 0 \ \forall \ p_I + p_{II} = 1$.

Given that $\hat{C} > 0$ whenever the test administering C has positive discriminatory power we an combine (C.3) and (C.4) to obtain:

$$\begin{aligned} &\frac{\theta p_I u'(\hat{B}) + (1-\theta)(1-p_{II})\int_0^\infty v_M[n(1-t), \hat{B}]dF(n)}{\theta p_I + (1-\theta)(1-p_{II})} \\ \leq \hat{\lambda} = &\frac{\theta(1-p_I)u'(\hat{B}+\hat{C}) + (1-\theta)p_{II}\int_0^\infty v_M[n(1-\hat{t}), \hat{B}+\hat{C}]dF(n)}{\theta(1-p_I) + (1-\theta)p_{II}}\ ;\ \hat{B} \geq 0 \end{aligned} \tag{C.8}$$

(i) ($\hat{B} = 0$, $\hat{C} > 0$)

From (C.8) we can see that $B > 0$ only if the average smvi of non-categorical recipients exceeds the average smvi of categorical recipients when $B = 0$. If $p_I = 0$ then this certainly holds because $\lim_{x \to 0} u' = +\infty$.

From (C.8) it must hold that for $p_I > 0$ and $p_{II} > 0$:

$$u'(\hat{B}+\hat{C}) > \hat{\lambda} > \int_0^\infty v_M[n(1-\hat{t}),\hat{B}]dF(n) \tag{C.9}$$

This will assist us in interpreting the optimal tax expression which follows.

The FOC for the optimal tax rate is:

$$\begin{aligned}(t):\ &\int_0^\infty \left\langle p_{II}\cdot -nv_\omega[n(1-\hat{t}),\hat{B}+\hat{C}] + (1-p_{II})\cdot -nv_\omega[n(1-\hat{t}),\hat{B}+\hat{C}]\right\rangle dF(n)\\ &-\hat{\lambda}\int_0^\infty \left\langle y-\hat{t}\frac{\partial y}{\partial(1-t)}\right\rangle dF(n)=0\end{aligned}$$

and thus:

$$\frac{\hat{t}}{1-\hat{t}} = \frac{\int_0^\infty y\left\langle \hat{\lambda}-\left\{p_{II}v_M[n(1-\hat{t}),\hat{B}+\hat{C}]+(1-p_{II})v_M[n(1-\hat{t}),\hat{B}]\right\}\right\rangle f(n)dn}{\hat{\lambda}\bar{Z}\mathcal{E}_Z} \tag{C.10}$$

Notice that by (C.9) the numerator is unambiguously positive.

Part II

Individual Decisions and Risk Taking

Chapter 5

Enforcing *Ex-Post* Conditionality: Categorical Benefit Size and Risk

5.1 Introduction

Type II (false award) classification errors in the administration of welfare benefits give rise to a range of enforcement issues because they provide incentives for abuse of the welfare system.[1] This abuse can take a number of forms, some more detectable than others. First, individuals may choose to apply for a benefit and, if incorrectly awarded it, comply with any subsequent *ex-post* conditions. This behaviour is very difficult to detect (Yaniv, 1986). In this regard, a growing empirical literature analyses the work capability of recipients on the U.S. Social Security Disability Insurance programme, where the two most prevalent recipient categories are those with the difficult to detect and monitor conditions of musculoskeletal disease (back pain) or mental illness (Autor and Duggan, 2003; Von Wachter et al., 2011) [2]. Second, ineligible benefit recipients may

[1]Two principal studies provide estimates of the propensity of the U.S. Social Security Administration to make Type II errors in awarding disability insurance (see Benitez-Silva et al., 2004; Nagi, 1969). Both generate quantitatively similar estimates of around 20% (22% and 19%, respectively).

[2]Modern welfare programmes typically do not focus exclusively on strict medical criteria, but also on the subjective assessment of ability to function in the workplace. For example, the US Social Security Administration places significant emphasis on an applicant's reported discomfort (Autor and Duggan, 2006)

choose to violate *ex-post* conditions, taking into account the risk of being detected and sanctioned in some way. This may take the form of working when receiving a disability benefit or unemployment benefit. Whilst it may seem natural to assume that all working recipients work 'cash-in-hand' in the informal economy, Fuller et al. (2015) provide evidence that a substantial degree of fraud occurs through official/registered employment. This may arise due to inadequacies in the integration of I.T. systems across tax authorities, benefit authorities and local government.[3]

This chapter analyses the decision of individuals who are able to work - but differ continuously in their wage - to apply for a categorical benefit, C, that is *ex-ante* conditional on an applicant being unable to work; and *ex-post* conditional on a recipient not working. Benefit recipients may be also be required to spend a fraction of the working day at the benefit office. It is assumed that there are no checks or penalties in place for an able individual who is incorrectly awarded C but does not work when receiving it. There are two reasons for making this assumption. First, as discussed above, such behaviour is likely to be highly difficult to detect as an able recipient exactly mimics an unable recipient. Second, in reality some individuals may be unsure as to their own eligibility and, consequently, their being awarded the benefit by administrative error does not constitute fraud. However, the act of working when receiving C is detectable and provides the 'smoking gun' necessary for the benefit authority to identify, *ex-post*, a benefit recipient as able to work and, importantly, actively violating the rules. This constitutes detectable fraud in this chapter.[4] An individual who is detected working is made to repay the benefit in its entirety, in addition to paying a fine *proportional* to the benefit size. This falls in line with the type of sanction benefit authorities may 'offer' in reality.[5]

[3]Many governments now cite the need to improve the degree of information sharing across relevant authorities. In the U.K., for example, this takes the form of the 'Single Fraud Investigation Service'.

[4]Yaniv (1986) refers to this as 'outright fraud'; whilst Fuller et al. (2015) refer to it simply as 'concealed earnings'.

[5]In reality, sanctions can take a number of non-mutually exclusive forms, including exclusion from the welfare programme for a given period; a fine of a fixed absolute amount; a fine proportional to the monetary amount incorrectly received; the repayment of benefits; or even prosecution. Note that each type of sanction imposes a financial loss on the individual: either directly through a fine or indirectly through the loss of future benefit payments. For example, in the U.K. the Department of Work and Pensions may, in addition to recovering overpayments (i) disqualify a recipient for periods including 13 weeks, 26 weeks or 3 years, depending on the number of previous offences committed; (ii) offer a fine proportional to 50 percent of the overpayment (up to a maximum of £2000); or (iii) prosecute. The administrative fine is offered as an alternative to prosecution. Fines and repayments may be either deducted from benefit payments (if still receiving the benefit) or, if disqualified, paid directly

In the main framework labour supply is modelled in the *extensive* margin, such that individuals either work a fixed amount or they do not work at all. Such an assumption seems reasonable given that many studies emphasise the importance of the extensive margin relative to the intensive margin for labour supply responses to tax/benefit programmes, in particular for those with lower incomes (Eissa and Liebman, 1996; Saez, 2001; Jacquet, 2006, 2014). As will be detailed below, this assumption greatly simplifies the analysis of individual decisions to engage in the risky activity of working when receiving the categorical benefit.

Within this framework, we ask the following related questions.

(i) For cases where the standard enforcement parameters (detection probability, penalty rate) do not alone deter applications from individuals who would choose to work when receiving the categorical benefit, how does the benefit level affect deterrence? In particular, are there enforcement parameter-benefit level combinations such that full deterrence can be achieved?

(ii) How does the decision of an able individual to apply for the categorical benefit and, if awarded it, their subsequent work decision, differ with the wage rate?

(iii) How do the answers to the above two questions differ depending on whether preferences exhibit constant absolute risk aversion (CARA) or decreasing absolute risk aversion (DARA)?

To analyse these questions, we capture the risk attached to working when receiving the categorical benefit by the *risk premium* associated with the variance in benefit income, as approximated by standard methods (Arrow, 1970; Pratt, 1964). The assumption that a working individual's labour supply is *fixed* in the intensive margin allows us to use these methods.

In answering these questions, this chapter relates to two strands of literature. First, drawing on the economics of crime (Becker, 1968), a small number of papers model the decision of individuals to claim unemployment insurance when actually employed (Yaniv, 1986; Fuller et al., 2015). The closest precursor to this chapter is Yaniv (1986), who employs a static model in which individuals with preferences over income that exhibit decreasing absolute risk aversion must choose the number of days of their time endowment to fraudulently claim. Recipients must spend a fraction of the working day

by the individual (see Department for Work and Pensions, 2010, 2011, 2015).

at the benefit office, in addition to waiting a fixed period before benefits are received. The author considers two alternative fine structures from the tax evasion literature (see Allingham and Sandmo, 1972; Yitzhaki, 1974). When the fine is proportional to the number of claiming days - and thus independent of the benefit size - a *ceteris paribus* increase in the benefit level increases the claiming duration via positive substitution and income effects. Contrastingly, when the fine is proportional to total benefit income, an increase in the benefit level increases the expected fine, thus generating an overall ambiguous effect on incentives. Notably, Yaniv does not allow for the possibility that the benefit itself may induce voluntary unemployment. In reality, the size of a benefit is likely to not only influence which ineligible individuals choose to apply, but also whether or not they choose to fully comply with *ex-post* conditions when receiving it.

A related second strand of literature focuses on the design of welfare benefits when the benefit authority has no formal technology to determine eligibility, but instead chooses consumption bundles/transfer levels that satisfy incentive compatibility constraints and induce self-revelation (Besley and Coate, 1992; Blackorby and Donaldson, 1988; Cuff, 2000; Diamond and Mirrlees, 1978; Kreiner and Tranaes, 2005; Nichols and Zeckhauser, 1982). In particular, the works of Besley and Coate (1992), Cuff (2000), and Kreiner and Tranaes (2005) use unproductive workfare - which is analogous to the *ex-post* condition of spending time at the benefit office - as a tool to deter non-needy individuals from applying for a given benefit. An important distinction that separates the current chapter from this literature is that, given an exogenous eligibility test and a continuum of potential ineligible applicants, there will always be some individuals who apply for a benefit and, if incorrectly awarded it, choose the riskless activity of not working. Even if individuals were required to spend the full working day at the benefit office, those with a sufficiently low wage may still choose to apply. Focus in our setting is therefore placed on investigating conditions under which individuals can be deterred from breaking *ex-post* conditionality.

The analysis in this chapter proceeds via backwards induction. Conditional on receiving C, we first determine which individuals would choose voluntary unemployment and which individuals would choose to work; taking into account the risk of being detected and fined. Then given this behaviour, we determine which individuals would choose to apply for C. In the main analysis the only condition placed on benefit recipients is that they do not work. An individual who would choose to work conditional on receiving C

will only apply for C if the expected benefit income exceeds the risk premium associated with the variance in benefit income. In a subsequent section we then adopt a simpler framework to analyse the case where benefit recipients are also required to spend a fraction of the 'working day' at the benefit office. The time requirement is taken to be fully enforced. In this framework individuals have preferences only over consumption (see Yaniv, 1986). An individual who would choose to work when receiving C will only choose to apply for it if the expected benefit income *net* of earnings foregone through time spent at the benefit office exceeds the risk premium. With this by way of background, the key results are as follows:

1. **Constant Absolute Risk Aversion**: Under the class of utility functions satisfying $-u_{xx}/u_x = \eta$, where $0 < \eta < 1$ is a constant and x is consumption, the risk premium associated with the variance in benefit income is independent of an individual's wage and *convex-increasing* in C. So whilst an increase in C (i) linearly increases a working benefit recipient's expected income; it also (ii) exposes them to increasingly greater risk through raising the risk premium. For any enforcement parameters (detection probability, penalty rate) which do not fully deter applications from those who would choose to work when receiving C, there is a *critical* level of C set *above* which full deterrence can in fact be attained. Further, the lower the level of enforcement provided by the standard enforcement parameters, the higher this critical level is. If C is set below this critical level, all able individuals above a threshold productivity level will apply for C and continue to work if awarded it.

2. **Decreasing Absolute Risk Aversion**: Under the class of utility functions satisfying $-u_{xx}/u_x = \eta/x$ [6], the risk premium associated with the variance in benefit income is a decreasing function of an individual's wage and *convex-increasing* in C. A necessary though not sufficient condition to deter individuals who would work when receiving C from applying is that the enforcement parameters be set sufficiently high - where the required level is independent of C and productivity. Conditional on this being achieved, full deterrence can only be attained if C exceeds its critical value at *each* productivity level.

3. **Extension with fully enforced time requirement**: When the only *ex-post* condition imposed on recipients is that they do not work, an unsatisfactory im-

[6] i.e. Constant Relative Risk Aversion preferences.

plication of both the CARA and DARA analyses is that higher earners will apply for C if the enforcement of the no-work condition is too lenient. In reality, higher earners are unlikely to apply for benefits even if they face a positive probability of being awarded them. One of many explanatory factors for this is that receiving benefits can be a time consuming activity. Within a framework where individuals have CARA preferences, we impose the requirement that recipients of C must spend a fraction of the working day at the benefit office. This (i) preserves the result that a benefit set sufficiently high can fully enforce the no-work requirement; but now also (ii) generates a critical wage above which no able individual will apply for C because the opportunity cost of foregone earnings is too high.

The remainder of this chapter is structured as follows. In Section 5.2 we set up the model. Section 5.3 then presents the main analysis under both CARA and DARA preferences. Section 5.4 then extends the analysis to a framework where there is a fully enforced condition that benefit recipients spend a fraction of the day at the benefit office. Concluding remarks are provided in Section 5.5.

5.2 The Model

5.2.1 Individuals

Individual preferences over consumption, $x \geq 0$, and leisure, $l \in [0,1]$, are represented by the utility function $u(x,l)$. The standard assumptions apply: u is continuous, differentiable, increasing in both arguments ($u_x > 0$, $u_l > 0$) and strictly concave ($u_{xx} < 0$, $u_{ll} < 0$, $u_{xx}u_{ll} - u_{xl}^2 > 0$). Any additional properties will be later determined by assumptions placed on risk aversion.

Let there be a subpopulation of individuals who are able to work, but differ continuously in their net wage, $\omega \geq \omega_0$; where $\omega_0 > 0$ is the lowest wage in the economy. Individual labour supply is modelled in the *extensive* margin[7], such that an individual either works the fixed amount $\mathcal{H} \in (0,1)$ of their time endowment, or enjoys full leisure. The assumption is made that, in absence of any form of unearned income, all able

[7]Once more, extensive margin labour responses to tax/benefit programmes are shown to be important relative to intensive margin responses, particularly for those commanding lower wages (Eissa and Liebman, 1996; Saez, 2001; Jacquet, 2006, 2014).

individuals will work and have consumption $x = \omega\mathcal{H}$.

However, these individuals may choose to apply for a categorical benefit, $C \geq 0$, that is targeted at an unable subpopulation of individuals who cannot work, but which is administered with Type II (false award) classification errors. The categorical benefit is *ex-ante* conditional on an applicant being unable to work; and *ex-post* conditional on a recipient not working.[8]

For simplicity, it is assumed that applications for C are costless in terms of money, stigma and time. Under this frequently employed assumption (see Jacquet, 2006, 2014) the utility of a rejected applicant coincides with their utility from having not applied in the first place. Application decisions are therefore independent of the positive Type II error propensity, instead depending solely on whether the (expected) utility from receiving C exceeds that from not. This renders the analysis more tractable.

5.2.2 Enforcement Issues

Type II errors in the awards process violate *ex-ante* conditionality, and in turn generate enforcement issues with respect to *ex-post* conditionality because individuals who are able to work are receiving the categorical benefit. Whether or not they will in fact choose to work will depend on the enforcement mechanisms in place.

It is assumed that there are no checks or penalties in place for an able recipient of C who does not work and who thus complies with *ex-post* conditionality. There are two reasons for making this assumption. First, in reality it is not immediately clear that such behaviour is 'fraudulent' because an applicant may be uncertain of their own eligibility. Second, and of more importance for this chapter, such behaviour is likely to be highly difficult to detect because the truly ineligible recipient exactly mimics the behaviour of a truly eligible recipient (Yaniv, 1986).

However, the act of working whilst receiving C is detectable and provides the 'smoking gun' necessary for the benefit authority to identify, *ex-post*, a benefit recipient as able to work and, importantly, actively violating the rules. This constitutes detectable fraud in this chapter. Any recipient of C who does work risks being detected with probability ρ ; $0 \leq \rho \leq 1$. If detected the ineligible recipient is made to repay C in its entirety

[8] In a subsequent section we will explore a framework where recipients are also required to spend a fraction of the working day at the benefit office.

and, in addition, is imposed with a fine of size ϕC, where $0 \leq \phi \leq 1$. Putting this all together, the expected categorical benefit when working is given by:

$$\tilde{C}(C, \rho, \phi) = C[1 - \rho(1 + \phi)] \geq 0 \tag{5.1}$$

We note immediately the partial derivatives: $\tilde{C}_C \geq 0$, $\tilde{C}_\rho \leq 0$ and $\tilde{C}_\phi \leq 0$. For any enforcement pairs (ρ, ϕ) satisfying $\rho(1+\phi) > 1$ it is straightforward to see that no able individual will apply for C with the intention of working because $\tilde{C} < 0$. Accordingly, we henceforth restrict attention to *lenient enforcement* pairs satisfying $\rho(1 + \phi) \leq 1$ (and thus $\tilde{C} \geq 0$) and wish to identify the conditions under which individuals will apply for C and, if awarded it, violate *ex-post* conditionality by working.

5.3 Analysis

Given the assumptions laid out, the purpose of the main analysis is to determine how the enforcement parameters and benefit size determine which able individuals will choose to apply for the categorical benefit and, importantly, whether or not they will choose to work when receiving it. By backwards induction we first establish how an able individual would behave (work/do not work) conditional on receiving C. To determine application decisions, we then compare their utility when receiving C with that when not receiving C.

To proceed, however, we need a convenient way to write the expected utility of a working recipient of C. We denote the risk associated with working - as captured by deviations around $\tilde{C}$ - by the random variable:

$$\alpha = \begin{cases} \alpha_0 & : \text{Prob}(\rho) \\ \alpha_1 & : \text{Prob}(1 - \rho) \end{cases} \tag{5.2}$$

where:

$$\alpha_0 = -(1 - \rho)(1 + \phi)C \ \ , \ \alpha_1 = \rho(1 + \phi)C \tag{5.3}$$

and thus $E(\alpha) = 0$.

The detected state therefore corresponds to $\tilde{C} + \alpha_0 = -\phi C$, whilst the undetected state corresponds to $\tilde{C} + \alpha_1 = C$. To save on notation, let $A \equiv \omega\mathcal{H} + \tilde{C}$ denote the expected income of an individual who works when receiving C. Because labour supply is exogenously fixed in the intensive margin, we can employ the standard methods of Arrow (1970) and Pratt (1964) to show that the risk premium, $\chi(A, \sigma)$, associated with the variance in benefit income, σ, satisfies:

$$u(A - \chi, 1 - \mathcal{H}) \equiv \rho u(A + \alpha_0, 1 - \mathcal{H}) + (1 - \rho)u(A + \alpha_1, 1 - \mathcal{H}) \tag{5.4}$$

where:[9]

$$\chi(A, \sigma) = \frac{1}{2} r(A) \cdot \sigma(\rho, \phi, C);$$

$$\sigma(\rho, \phi, C) = E(\alpha^2) = \rho(1 - \rho)(1 + \phi)^2 C^2, \tag{5.5}$$

$$r(A) = -\frac{u_{xx}(A, 1 - \mathcal{H})}{u_x(A, 1 - \mathcal{H})}$$

So, in general, the risk premium is the coefficient of absolute risk aversion[10] , r, multiplied by half the variance in categorical income[11] , σ (see Pratt, 1964, p.125). Once more, because labour supply is constant in the intensive margin, the curvature of the utility function over income is given by the standard formula for r (Chetty, 2006). Note that σ is *convex-increasing* in both C and ϕ, but may be increasing or decreasing in ρ. Formally:

[9] Our use of the mean-zero random variable α follows directly from Pratt (1964, p.124, see equations (1) and (2)).

[10] Given that labour supply is constant, we could define $a(x) \equiv u(x, 1 - \mathcal{H})$, so that $r = -a''/a'$.

[11] Note that the variance in α is equivalent to the variance in benefit income. Formally:

$$E[\alpha - E(\alpha)]^2 = \rho{\alpha_0}^2 + (1 - \rho){\alpha_1}^2 = \rho[(\tilde{C} + \alpha_0) - \tilde{C}]^2 + (1 - \rho)[(\tilde{C} + \alpha_1) - \tilde{C}]^2$$

$$\sigma_C > 0,\ \sigma_{CC} > 0,\ \sigma_\phi > 0,\ \sigma_{\phi\phi} > 0\ ,$$

$$\sigma_\rho \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \Leftrightarrow \rho \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \frac{1}{2} \quad . \tag{5.6}$$

We can also see from (5.5) that individual productivity - the sole source of heterogeneity across able individuals - only enters χ via r. The decision of an able recipient of C to work will therefore depend on how r changes with ω, via changes in A. To proceed, we analyse both the cases where risk aversion is (i) independent of ω and thus constant across individuals; and (ii) a decreasing function of ω.

5.3.1 Constant Absolute Risk Aversion (CARA)

In this section we eliminate the dependence of r on A (and thus ω) through assuming $r = \eta$, where $0 < \eta < 1$ is a constant. By standard methods, the class of utility functions satisfying this assumption is given by the solution to the second-order linear differential equation $u_{xx} + \eta u_x = 0$. Formally, let preferences take the CARA form[12]:

$$u(x, l) = 1 - \psi(l)e^{-\eta x} \tag{5.7}$$

where $r = -u_{xx}/u_x = \eta$, $\psi(l) > 0$; $\psi' < 0$; $\psi'' > 0$ and $\psi\psi'' > (\psi')^2$. The last assumption guarantees that u is strictly concave (see Appendix).[13] An implication of these preferences is that $u_l u_{xx} - u_x u_{xl} = 0$, such that the marginal utility of consumption is independent of leisure along an indifference curve. Further, no individual with $\omega \leq \bar{\omega} \equiv -(1/\eta)\text{Log}[\psi(1)/\psi(1 - \mathcal{H})]/\mathcal{H} > 0$ will work. For simplicity we assume $\omega_0 = \bar{\omega}$, such that all able individuals who do not receive C choose to work.[14]

[12]This functional form is also employed by Berloffa and Simmons (2003). For an application of CARA preferences in an analysis of benefit fraud, see Fuller et al. (2015).

[13]Suppose, for example, that $\psi(l) = 1/l$. In this case $\psi'(l) = -1/l^2$ whilst $\psi''(l) = 2/l^3$. We therefore have $\psi\psi'' = 2/l^4 > (\psi')^2 = 1/l^4$ and the assumption is readily satisfied.

[14]The main results remain unchanged if we allow $\omega < \bar{\omega}$: in this case voluntarily unemployed

The risk premium in (5.5) now becomes:

$$\chi(\sigma) = \frac{1}{2}\eta\sigma \tag{5.8}$$

So χ is now a *constant* multiplied by half the variance in categorical income. It is independent of the individual wage rate and thus takes the same value for all individuals. Under this assumption then, the enforcement parameters and benefit level will affect the risk premium solely through their effect on the variance in benefit income. Indeed, from (5.5) we can see that χ is *convex-increasing* in both ϕ and C, but may be increasing or decreasing in ρ.

Work Decision Conditional on Receiving C

An able recipient of C will only choose to work if $u(A-\chi, 1-\mathcal{H}) > u(C,1)$. Given that A is unambiguously increasing in ω, let the critical net wage $\underline{\omega}(\rho,\phi,C)$ satisfy:

$$u\left[\underline{\omega}(\rho,\phi,C)\cdot\mathcal{H} + (\tilde{C}-\chi), 1-\mathcal{H}\right] \equiv u(C,1) \tag{5.9}$$

We assume that at the point of indifference an individual will choose *not* to work. An able recipient of C will therefore not work if $\omega \leq \underline{\omega}$; but will work if $\omega > \underline{\omega}$. Differentiating (5.9) with respect to C, ρ and ϕ, respectively, we obtain:

$$\begin{aligned}
\underline{\omega}_C &= \frac{1}{\mathcal{H}}\left\{\left[\frac{u_x(C,1)}{u_x(\underline{\omega}\mathcal{H}+\tilde{C}-\chi, 1-\mathcal{H})} - \tilde{C}_C\right] + \chi'\sigma_C\right\} > 0 \\
\underline{\omega}_\phi &= \frac{1}{\mathcal{H}}(\chi'\sigma_\phi - \tilde{C}_\phi) > 0 \\
\underline{\omega}_\rho &= \frac{1}{\mathcal{H}}(\chi'\sigma_\rho - \tilde{C}_\rho)\begin{Bmatrix}\geq \\ <\end{Bmatrix} 0
\end{aligned} \tag{5.10}$$

individuals with $\omega \in [\omega_0, \bar{\omega}]$ will all apply for C and, if awarded it, remain voluntarily unemployed (if they do not work when receiving no benefit income, they will certainly not work when receiving benefit income).

where $\underline{\omega}_\rho > 0 \;\forall\; \rho \leq 1/2$, whilst $\forall\, \rho > 1/2$:

$$\underline{\omega}_\rho \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 0 \Leftrightarrow C \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \bar{C}(\rho, \phi) \equiv \frac{2}{\eta(2\rho - 1)(1 + \phi)} \tag{5.11}$$

From (5.10) it can be readily verified (see Appendix) that an increase in C unambiguously increases $\underline{\omega}$. There are two effects at work. First[15], an increase in C increases both the certain consumption of a non-working recipient and the expected consumption of a working recipient. However, because $\tilde{C}_C < 1$ the increase in consumption for a non-worker exceeds that for a worker, thus driving $\underline{\omega}$ to rise. Second, an increase in C increases χ and thus the risk that a working individual faces. To compensate for this, $\underline{\omega}$ must further rise.

Turning to the enforcement parameters, an increase in ϕ unambiguously increases $\underline{\omega}$ because $\chi'\sigma_\phi > 0$ and $\tilde{C}_\phi < 0$ (i.e. it increases the risk premium and lowers the expected benefit), whilst the affect of an increase in ρ depends on whether $\rho \leq 1/2$ or $\rho > 1/2$. In the former case, $\underline{\omega}$ unambiguously increases for the same reasons as for ϕ. In the latter case, however, we have $\chi'\sigma_\rho < 0$, such that $\underline{\omega}$ only increases if the reduction in χ is offset by the reduction in $\tilde{C}$, thereby causing $(\chi - \tilde{C})$ to rise. Given that $\chi'\sigma_{CC} > 0$, this will only hold if C is sufficiently small.

Decision to Apply for C

With knowledge of individual behaviour conditional on receiving C, we now determine who will apply for C and under what conditions. Because applications are taken to be costless, an individual's utility from having an application for C rejected is identical to that from having not applied. Accordingly, application decisions are independent of the propensity of the benefit authority to make Type II errors (which is assumed to be positive). An individual commanding wage ω will therefore make an application if the utility from receiving C exceeds that from working, given that all able individuals work when not receiving C.[16]

[15]Note that under the CARA preferences in (5.7) the implication of $u_l u_{xx} - u_x u_{xl} = 0$ is that $u_x(C,1)/u_x(\underline{\omega}\mathcal{H}+\tilde{C}-\chi, 1-\mathcal{H}) = 1$. We could thus alternatively write $\underline{\omega}_C = (1/\mathcal{H})\left\{[1 - \tilde{C}_C] + \chi'\sigma_C\right\}$.

[16]Following the assumption that $\omega_0 = \bar{\omega} \equiv (\frac{1}{\eta})\text{Log}\left[\frac{\psi(1)}{\psi(1-\mathcal{H})}\right] /\mathcal{H}$.

As established, there are two important subgroups to consider: (i) those with $\omega < \underline{\omega}$ who would not work conditional on receiving C; and (ii) those with $\omega > \underline{\omega}$ who would work conditional on receiving C.

(i) $\omega < \underline{\omega}$

First off, an individual with $\omega < \underline{\omega}$ will apply for C only if $u(C, 1) \geq u(\omega\mathcal{H}, 1 - \mathcal{H})$, and thus if $\omega \leq \bar{\bar{\omega}}(C)$, where $\bar{\bar{\omega}}$ is implicitly defined by:

$$u[\bar{\bar{\omega}}(C) \cdot \mathcal{H}, 1 - \mathcal{H}] \equiv u(C, 1) \tag{5.12}$$

It follows immediately from (5.12) that $\bar{\bar{\omega}} > C/\mathcal{H} > C$ and:

$$\bar{\bar{\omega}}_C = \frac{1}{\mathcal{H}} \frac{u_x(C, 1)}{u_x(\bar{\bar{\omega}}\mathcal{H}, 1 - \mathcal{H})} > 1 \tag{5.13}$$

Unsurprisingly then, an increase in C increases the critical wage at or below which an able individual chooses to apply for C and not work if awarded it.

(ii) $\underline{\omega} < \omega$

Turning to those with $\underline{\omega} < \omega$, an application for C will be made only if $u[\omega\mathcal{H} + (\tilde{C} - \chi), 1 - \mathcal{H}] > u(\omega\mathcal{H}, 1 - \mathcal{H})$, which holds if $\chi < \tilde{C}$. Given that χ is independent of ω, it follows that *all* individuals with $\underline{\omega} < \omega$ will apply for C if $\chi < \tilde{C}$.

Combining the definitions of $\underline{\omega}$ and $\bar{\bar{\omega}}$ in (5.9) and (5.12), respectively, we have:

$$u[\underline{\omega}\mathcal{H} + (\tilde{C} - \chi), 1 - \mathcal{H}] \equiv u[\bar{\bar{\omega}}\mathcal{H}, 1 - \mathcal{H}] \equiv u(C, 1) \tag{5.14}$$

Recall that $\underline{\omega}$ is a function of the enforcement parameters and the benefit level, whilst $\bar{\bar{\omega}}$ is a function of only the benefit level. Accordingly, the combination of enforcement parameters and the benefit size will determine the position of $\underline{\omega}$ relative to $\bar{\bar{\omega}}$ on the net wage continuum. It follows immediately from (5.14) that the relationship between $\underline{\omega}$ and $\bar{\bar{\omega}}$ is:

$$\underline{\omega} - \bar{\bar{\omega}} \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0 \Leftrightarrow (\chi - \tilde{C}) \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0 \tag{5.15}$$

If $\chi > \tilde{C}$ then $\underline{\omega} > \bar{\bar{\omega}}$ and so (i) all those with $\omega \in [\omega_0, \bar{\bar{\omega}}]$ apply for C and do not work if awarded it; whilst (ii) no individual with $\omega \in (\bar{\bar{\omega}}, \infty)$ applies. Notice that in this case there are individuals who belong in the wage interval $\omega \in (\bar{\bar{\omega}}, \underline{\omega}]$ who would not work if awarded the benefit, but choose not to apply for it. If $\chi = \tilde{C}$ then $\underline{\omega} = \bar{\bar{\omega}}$ and applications are otherwise as described above. However, if $\chi < \tilde{C}$ then $\underline{\omega} < \bar{\bar{\omega}}$ such that (i) all those with $\omega \in [\omega_0, \underline{\omega}]$ apply for C and do not work if awarded it; whilst (ii) all those with $\omega \in (\underline{\omega}, \infty)$ *apply* for C and *do work* if awarded it. Figure 5.1 graphically depicts the relationship between these two critical wages.

The parameters under which *ex-post* conditionality is fully enforced

Following (5.15), a natural question to investigate is the conditions under which $\chi \geq \tilde{C}$ (respectively $\chi < \tilde{C}$). Intuitively, this will depend on the relationship between the enforcement parameters and the benefit level in determining the relative sizes of χ and $\tilde{C}$. Indeed, equating χ and $\tilde{C}$, it is straightforward to establish the following result.

Result 1: *For any lenient enforcement pair (ρ, ϕ);*

$$(\chi - \tilde{C}) \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0 \Leftrightarrow C \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} \overset{*}{C}(\rho, \phi) \tag{5.16}$$

where:

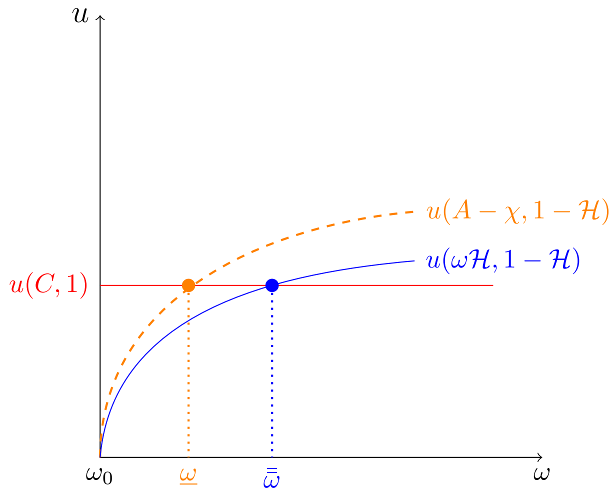
$$\overset{*}{C}(\rho, \phi) \equiv \frac{2[1 - \rho(1 + \phi)]}{\eta\rho(1 - \rho)(1 + \phi)^2} \tag{5.17}$$

An immediate corollary from (5.16) is:

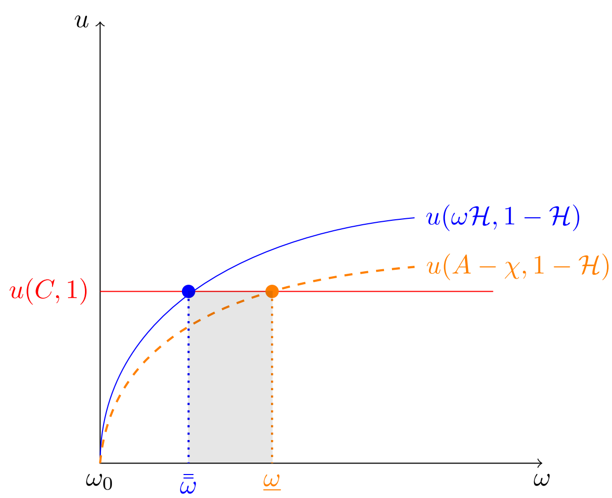
Corollary 1: *For enforcement parameters which either (i) fully deter , or (ii) provide zero deterrence, against applications from those with $\underline{\omega} < \omega$, the size of C has no impact*

Figure 5.1: The relationship between $\underline{\omega}$ and $\bar{\bar{\omega}}$.

(a) Utility and the critical wages ($\tilde{C} > \chi$)



(b) Utility and the critical wages ($\tilde{C} < \chi$)



Notes. This figure provides graphical intuition for the critical net wages $\underline{\omega}$ and $\bar{\bar{\omega}}$. Subplot (a) depicts the case where $\tilde{C} > \chi$, whilst subplot (b) depicts the case where $\tilde{C} < \chi$. Both subplots plot individual utility over the net wage continuum. The horizontal line in each subplot gives the utility level of an individual who receives the categorical benefit and does not work. If $\tilde{C} > \chi$ then the (expected) utility of a working categorical recipient must, at all wage levels, exceed the utility of a worker who does not receive the categorical benefit. It must therefore hold that $\underline{\omega} < \bar{\bar{\omega}}$. Contrastingly, if $\tilde{C} < \chi$ then the (expected) utility of a working categorical recipient must be less than that of a worker who does not receive the categorical benefit. In this case it must therefore hold that $\underline{\omega} > \bar{\bar{\omega}}$. The filled region between $\bar{\bar{\omega}}$ and $\underline{\omega}$ illustrates that some individuals who would choose not to work conditional on receiving C do not in fact apply for C.

on application decisions. Formally:

$$\lim_{\rho(1+\phi)\to 1} \overset{*}{C} = 0 \quad ,$$
$$\lim_{\rho(1+\phi)\to 0} \overset{*}{C} = +\infty.$$

Much more interesting however, are the intermediate cases, $0 < \rho(1+\phi) < 1$, where the enforcement parameters in isolation provide neither full or zero deterrence against applying for C and subsequently working. In these cases, the size of C matters in deterring such applications. In particular, (5.16) and (5.17) show that C must be set sufficiently *high* to achieve deterrence. The intuition for this initially surprising result rests on the fact that:

$$\frac{d^2}{dC^2}(\chi - \tilde{C}) = \frac{d^2}{dC^2}\chi = \chi' \sigma_{CC} > 0 \tag{5.18}$$

So whilst $\tilde{C}$ is linearly increasing in C, χ is *convex-increasing* in C and this drives the main result. Intuitively, increases in C expose working benefit recipients to greater and greater risk such that, for C set above the critical level $\overset{*}{C}$, χ more than offsets $\tilde{C}$. Figure 5.2 illustrates why this result arises. Exactly how high C would need to be set to achieve deterrence depends on the leniency of the enforcement parameters. In particular:

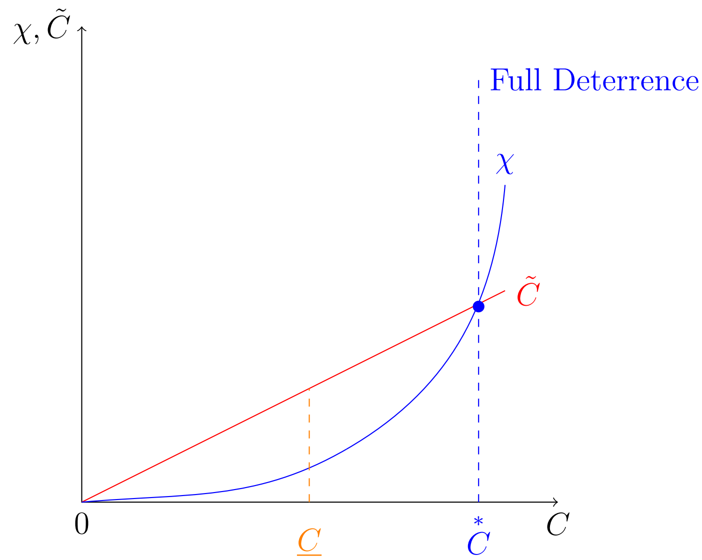
$$\overset{*}{C}_\rho < 0,\ \overset{*}{C}_\phi < 0,\ \overset{*}{C}_{\rho\phi} > 0 \quad \forall\, 0 < \rho(1+\phi) < 1 \tag{5.19}$$

So $\overset{*}{C}$ is increasing in the leniency of enforcement. The less effective the enforcement parameters are at deterring able benefit recipients from working, the higher the benefit level would need to be set. Note that $\overset{*}{C}_\rho < 0$ implies that $d(\chi - \tilde{C})/d\rho > 0$ whenever $C < \overset{*}{C}$. Indeed, from (5.11) and (5.17) it is straightforward to show that:

$$\bar{C}(\rho,\phi) \leq \overset{*}{C}(\rho,\phi) \tag{5.20}$$

The detection probability is therefore effective at raising $\underline{\omega}$ - and thereby reducing the

Figure 5.2: The Critical Categorical Benefit, $\overset{*}{C}(\rho, \phi)$



Notes. The point $\underline{C} \equiv \overset{*}{C}/2$ simply satisfies $d(\chi - \tilde{C})/dC = 0$.

number of individuals who would choose to apply for C and subsequently work - for all cases where $\chi \leq \tilde{C}$.

5.3.2 Decreasing Absolute Risk Aversion (DARA)

Under the assumption of CARA, changes in the enforcement parameters and the benefit level only affect the risk premium via changes in the variance in benefit income. Let us now relax this assumption and instead suppose that risk aversion is *decreasing* over income and, in turn, the individual wage. We first re-write the general expression for the risk premium in (5.5) as:

$$\chi(A, \sigma) = \frac{1}{2} r(A) \sigma(\rho, \phi, C) \ ; \ r'(A) < 0 \tag{5.21}$$

Given that χ is now a function of ω, where ω is continuously distributed, there must now be a continuum of risk premia. Note that $d\chi/d\omega = \chi_A \cdot \partial A/\partial \omega = (1/2) r'(A) \sigma \partial A/\partial \omega < 0$. Differentiating χ with respect to C, there are now two conflicting effects:

$$\frac{d\chi(A,\sigma)}{dC} = \frac{r(A)\sigma}{2C}\left\{\frac{C}{\sigma}\sigma_C + \frac{C}{r(A)}\left(r'(A)\cdot\tilde{C}_C\right)\right\}\begin{Bmatrix}\geq\\<\end{Bmatrix}0 \tag{5.22}$$

The first term in braces is the elasticity of σ with respect C. This term is positive because $\sigma_C > 0$ and it captures the standard effect from the CARA analysis. The second term, meanwhile, is the elasticity of r with respect to C. This term must be negative because $\tilde{C}_C > 0$, thus capturing the fact that an increase in expected income lowers the coefficient of risk aversion. So whether or not an increase in C will increase $\chi(A,\sigma)$ will depend on the aggregate of these two opposing effects.

We can also write the affect of the enforcement parameters on χ in terms of elasticities:

$$\begin{aligned}\frac{d\chi(A,\sigma)}{d\phi} &= \frac{r(A)\sigma}{2\phi}\left\{\frac{\phi}{\sigma}\sigma_\phi + \frac{\phi}{r(A)}(r'(A)\cdot\tilde{C}_\phi)\right\} > 0\ ,\\ \frac{d\chi(A,\sigma)}{d\rho} &= \frac{r(A)\sigma}{2\rho}\left\{\frac{\rho}{\sigma}\sigma_\rho + \frac{\rho}{r(A)}(r'(A)\cdot\tilde{C}_\rho)\right\}\begin{Bmatrix}\geq\\<\end{Bmatrix}0\ .\end{aligned} \tag{5.23}$$

Given that $\sigma_\phi > 0$ and $r'(A)\tilde{C}_\phi > 0$, an increase in ϕ unambiguously increases χ at each productivity level. In words, an increase in the penalty rate serves to (i) increase the variance in benefit income, which in isolation increases the risk premium; and (ii) increase the coefficient of absolute risk aversion through lowering expected benefit income, which in isolation also serves to increase the risk premium. Once more, the effect of an increase in ρ on χ depends on whether $\rho \leq 1/2$ or $\rho > 1/2$. If $\rho < (=)1/2$ then $\sigma_\rho > (=)0$ and, because $r'(A)\tilde{C}_\rho > 0$, χ unambiguously increases for the same two reasons as discussed for ϕ. However, if $\rho > 1/2$ then $\sigma_\rho < 0$ such that the variance in benefit income and coefficient of absolute risk aversion move in opposite directions, thereby generating an overall ambiguous effect. So unlike in the CARA case, it may not always hold that χ is decreasing in ρ for all $\rho > 1/2$, where this arises because high values of ρ decrease expected assets which, in turn, increase risk aversion.

To trace through in more detail some of the implications of DARA on individual decisions to apply for C, a more explicit expression for the risk premium than that in

(5.22) is required. We now assume that:

$$r = \frac{\eta}{A};\ 0 < \eta \leq 1 \tag{5.24}$$

By standard differential equation methods once more, we thus assume that preferences take the form:

$$u(x, l) = \frac{x^{1-\eta}\psi(l)}{1-\eta}\ ;\ \ \eta \neq 1\ ,\ \psi' > 0\ \ ,\ \ \psi'' < 0 \tag{5.25}$$

Substituting (5.24) into (5.21) gives:

$$\chi(A, \sigma) = \frac{1}{2}\left(\frac{\eta}{A}\right)\sigma \tag{5.26}$$

Differentiating (5.26) respect to C, it can be readily established that the risk premium is once more *convex-increasing* in C. Formally:

$$\begin{aligned} \frac{d\chi}{dC} &= \frac{1}{2}\eta\left\{\frac{A\sigma_C - \sigma\tilde{C}_C}{A^2}\right\} = \frac{1}{2}\eta\left\{\frac{(A-\tilde{C})\sigma_C + \sigma\tilde{C}_C}{A^2}\right\} > 0\ , \\ \frac{d^2\chi}{dC^2} &= \frac{1}{2}\eta\frac{(A-\tilde{C})^2}{A^3}\sigma_{CC} > 0\ . \end{aligned} \tag{5.27}$$

Turning to the affect of the enforcement parameters on the risk premium, we have $d\chi/d\phi > 0$ as established in (5.23), whilst:

$$\frac{d\chi}{d\rho} > 0\ \forall\rho \leq \frac{1}{2},\ \ \frac{d\chi}{d\rho}\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \Leftrightarrow C \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \frac{(2\rho-1)\omega\mathcal{H}}{(1-\rho)^2 + \rho^2\phi}\ \forall\rho > \frac{1}{2} \tag{5.28}$$

So for all $\rho > 1/2$, C must be set sufficiently high at each wage level in order for an increase in ρ to raise χ. When C is large an increase in ρ generates a reduction in A - and thus an increase in r - that is sufficiently large so as to offset the reduction in σ.

Work Decision Conditional on Receiving C

A recipient of C will only choose to work if $u(A-\chi, 1-\mathcal{H}) > u(C,1)$.[17] Differentiating the left side with respect to ω gives $u_x \cdot \mathcal{H}(1-\chi_A) > 0$, which implies that we can once more define a critical net wage $\underline{\omega}^D(\rho, \phi, C)$ above which an individual will work when receiving C. Formally, $\underline{\omega}^D$ satisfies:

$$u\left[A(\underline{\omega}^D) - \chi(A(\underline{\omega}^D), \sigma), 1-\mathcal{H}\right] \equiv u(C,1) \tag{5.29}$$

An individual with $\omega \in [\omega_0, \underline{\omega}^D]$ will thus not work when receiving C; whilst someone with $\omega \in (\underline{\omega}^D, \infty)$ *will* work when receiving C.

Differentiating (5.29) with respect to C,ρ and ϕ, respectively, gives:

$$\begin{aligned}
\underline{\omega}^D_C &= \frac{1}{\mathcal{H}(1-\chi_A)}\left\{\left[\frac{u_x(C,1)}{u_x(\underline{\omega}^D\mathcal{H}+\tilde{C}-\chi, 1-\mathcal{H})} - \tilde{C}_C\right] + [\chi_A\tilde{C}_C + \chi_\sigma\sigma_C]\right\} > 0 \\
\underline{\omega}^D_\phi &= \frac{1}{\mathcal{H}}\left\{\frac{\chi_\sigma\sigma_\phi}{1-\chi_A} - \tilde{C}_\phi\right\} > 0 \\
\underline{\omega}^D_\rho &= \frac{1}{\mathcal{H}}\left\{\frac{\chi_\sigma\sigma_\rho}{1-\chi_A} - \tilde{C}_\rho\right\}\left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} 0
\end{aligned} \tag{5.30}$$

An increase in C unambiguously increases $\underline{\omega}^D$. Analogous to (5.10), there are two effects at work. First, an increase in C increases both the certain consumption of a non-working categorical recipient and the expected consumption of a working categorical

[17] Note that under the preferences specified in (5.25) we have $u_l u_{xx} - u_x u_{xl} < 0$, such that leisure is a normal good. An immediate implication of this is that there will be a critical wage $\bar{\omega}(C)$ satisfying $u(C,1) \equiv u(\bar{\omega}\mathcal{H} + C, 1-\mathcal{H})$ at or below which no individual will choose to work. Explicitly, this critical wage is given by:

$$\bar{\omega}(C) = \frac{C}{\mathcal{H}}\left\{\frac{1}{(1-\mathcal{H})^{\frac{1}{1-\eta}}} - 1\right\} \quad ; \quad \bar{\omega}'(C) > 0\,.$$

recipient. However, because $\tilde{C}_C < 1$ the increase in consumption for the non-worker exceeds the increase in expected consumption for the worker, thus causing $\underline{\omega}^D$ to rise.[18] Second, we know from (5.27) that the total effect of an increase in C on χ is positive, which further acts to increase $\underline{\omega}^D$. Next, an increase in ϕ unambiguously increases $\underline{\omega}^D$ because a working recipient's expected benefit income falls; whilst the risk premium increases. Finally, if (i) $\rho \leq 1/2$ an increase in ρ will unambiguously increase $\underline{\omega}^D$; but if (ii) $\rho > 1/2$ an increase in ρ will only increase $\underline{\omega}^D$ if $\chi_\sigma \sigma_\rho > \tilde{C}_\rho(1-\chi_A)$.

Decision to Apply for C

We now turn to address application decisions. Analogous to the the CARA analysis, all those with $\omega \leq \underline{\omega}^D$ ($\omega \leq \underline{\omega}$ in the CARA analysis) will apply for C if $\omega \leq \bar{\bar{\omega}}$, where $\bar{\bar{\omega}}$ is as defined in (5.12). Further, an individual with $\underline{\omega}^D < \omega$ will once more only apply for C if $\chi(A,\sigma) \leq \tilde{C}$. A key difference between the CARA and DARA cases is that this condition is now a function of the net wage rate. It is straightforward to show that:

$$(\chi - \tilde{C}) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \Leftrightarrow \Gamma(\rho,\phi) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \frac{\omega \mathcal{H}}{C} \;\; ; \; \Gamma \in [-1, +\infty) \tag{5.31}$$

where the function Γ is explicitly given by (see Appendix for the derivation);

$$\begin{aligned} \Gamma(\rho,\phi) &\equiv \frac{\frac{1}{4}\eta\sigma_{CC} - \tilde{C}_C^2}{\tilde{C}_C} \\ &= \frac{-(1+\phi)^2(\frac{\eta}{2}+1)\rho^2 + (1+\phi)[\frac{\eta}{2}(1+\phi)+2]\rho - 1}{[1-\rho(1+\phi)]} \;\; \forall \; \rho(1+\phi) \leq 1 \end{aligned} \tag{5.32}$$

[18] Under the preferences in (5.25) the marginal utility of consumption is increasing in leisure, such that $u_x(C,1)/u_x(\underline{\omega}\mathcal{H} + \tilde{C} - \chi, 1-\mathcal{H}) > 1$.

and satisfies:

$$\lim_{\rho(1+\phi)\to 0} \Gamma(\rho,\phi) = -1 \ ,$$

$$\lim_{\rho(1+\phi)\to 1} \Gamma(\rho,\phi) = +\infty \ , \tag{5.33}$$

$$\Gamma_\rho > 0 \ , \ \Gamma_\phi > 0 \ \forall \ \rho(1+\phi) \leq 1 \ .$$

Notice that Γ is a function of *only* the enforcement parameters (and the constant η). Furthermore, it is increasing in both of these parameters for all lenient enforcement cases. An immediate implication of (5.31) is therefore that, for any enforcement parameters that render $\Gamma < 0$, there will be no level of the categorical benefit that can deter an individual with $\underline{\omega}^D < \omega$ from applying for C and subsequently working if awarded it. That is, if enforcement is sufficiently lenient so as to render $\Gamma < 0$, there will be no way to achieve deterrence because the right side of (5.31) is always positive.

Given that Γ is a quadratic equation in ρ (equivalently, a quadratic equation in ϕ), let the function $\breve{\rho}(\phi)$ satisfy $\Gamma[\breve{\rho}(\phi),\phi] \equiv 0$. From the properties of Γ in (5.33) it must hold that $\breve{\rho}(\phi)$ is the first, or lower, root of Γ for a given ϕ. The second root occurs for enforcement parameters satisfying $\rho(1+\phi) > 1$ and is thus not considered. Figure 5.3 graphically depicts the function Γ and provides some numerical examples of $\breve{\rho}$ for various values of ϕ and η.

Drawing this all together, the following result must hold:

Result 2:

(i)

$$\rho \leq \breve{\rho}(\phi) \Rightarrow \chi < \tilde{C} \quad \forall \, \omega \tag{5.34}$$

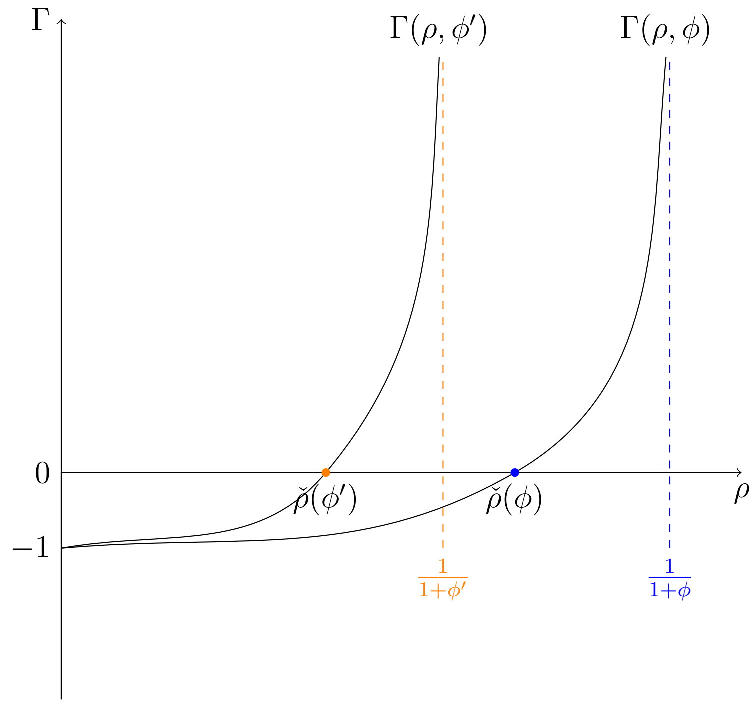
(ii)

$$\rho > \breve{\rho}(\phi) \Rightarrow (\chi - \tilde{C}) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \ \text{ if } \ C \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \overset{*}{C}{}^D(\omega,\rho,\phi) \equiv \frac{\omega\mathcal{H}}{\Gamma} \tag{5.35}$$

where $\lim_{\Gamma\to 0} \overset{*}{C}{}^D = +\infty$ *and* $\lim_{\Gamma\to +\infty} \overset{*}{C}{}^D = 0$

Figure 5.3: Properties of the Function $\Gamma(\rho, \phi)$

(a) Graphical Depiction of Γ



Notes: $\phi' > \phi$ and the points $(1+\phi)^{-1}$ and $(1+\phi')^{-1}$ are the lenient enforcement upper bounds

(b) Critical Detection Probability, $\check{\rho}$

ϕ	$\check{\rho}$ (η=0.25)	$\check{\rho}$(η=0.5)	$\check{\rho}$(η=0.75)
0	0.889	0.800	0.727
0.1	0.758	0.675	0.610
0.2	0.667	0.587	0.528
0.3	0.596	0.519	0.464
0.4	0.538	0.465	0.413
0.5	0.490	0.420	0.371
0.6	0.450	0.382	0.336
0.7	0.414	0.350	0.306
0.8	0.384	0.322	0.280
0.9	0.357	0.298	0.258
1	0.333	0.276	0.239

It follows directly from this result that:

Corollary 2: *A necessary - though not sufficient - condition to deter an individual with* $\underline{\omega} < \omega$ *from applying for* C *is that* $\rho > \check{\rho}(\phi)$*, and thus that the level of enforcement be set sufficiently high.*

The key message from this analysis is Result 2(i) and the proceeding Corollary 2. The intuition for this can be seen from the denominator in the definition of χ in (5.26): Higher levels of the enforcement parameters limit the impact of the size of C on A and thus limit the extent to which an increase in C reduces the coefficient of absolute risk aversion. Contrastingly, in the CARA case A is absent from the coefficient of absolute risk aversion and consequently this necessary condition does not arise.

5.3.3 Discussion

An unrealistic implication of the results obtained in both Section 5.3.1 (CARA analysis) and Section 5.3.2 (DARA analysis) is that, for enforcement-benefit level combinations that expose working recipients to insufficient risk, *all* higher earners will apply for the categorical benefit and, if awarded it, work. Yet, in reality higher earners are unlikely to want to apply for a categorical benefit even if they face some positive probability of being awarded it. This is likely to arise because neither (i) applying for the benefit, or (ii) receiving the benefit, are actually costless as so far assumed. Whilst applications are a one-time cost and abstracting from them renders the analysis more tractable, actually receiving C is likely to involve non-negligable time costs because recipients are typically required to spend a fraction of their day at the benefit office or engaging in certain activities. We turn to discuss this in the following section.

5.4 Time Opportunity Costs

The purpose of this section is to take seriously the real-word facet of welfare programmes that there are time costs associated with receiving categorical benefits. These time costs can take a multitude of forms depending on the nature of the benefit, such as regular meetings with a benefit officer; medical reassessments; work-capability tests;

work-preparation activities or providing evidence of job search.[19] Under the realistic assumption that such requirements are imposed during the working day, they also present *opportunity costs* to fraudulent recipients who work since their net earnings are lowered through taking time off work to fulfil these conditions (Yaniv, 1986)[20].

To analyse the implications of time opportunity costs we adopt a simpler framework where preferences are defined over consumption only[21] and given by $u(x)$; where $u' > 0$ and $u'' < 0$ (see also Yaniv, 1986). We further assume that $-u''/u' = \eta$ such that, analogous to Section 5.3.1, preferences exhibit constant absolute risk aversion.

In this setting, an individual who does not receive the categorical benefit has consumption $x = \omega$. Meanwhile, an individual who receives the categorical benefit *and* works has expected consumption $x = \omega(1-k) + \tilde{C}$; where $k \in (0,1)$ denotes the fraction of the day that a benefit recipient must spend attending meetings/reassessments with a benefit officer. This time requirement is taken to be fully enforced such that any recipient who does not conform with it automatically loses the benefit.

The analysis proceeds as in the previous sections: we first establish which individuals would choose to work/not work conditional on receiving C; and then, given this behaviour, determine which individuals will choose to apply for C. Given our assumption of CARA preferences, the risk premium in (5.8) applies.

[19]Whilst some benefits - such as incapacity payments - do not require frequent reassessment, many others do. For example, in the U.K. individuals who have some form of (non-terminal) illness or disability and receive Employment and Support Allowance (ESA) are required to attend frequent meetings with a 'benefits adviser' if they are placed in a 'work-related activity group'. An aim of such programmes is to eventually transfer recipients - via work capability assessments - into job search programmes and ultimately employment. Indeed, individuals receiving Jobseekers Allowance must 'sign on' every two weeks and provide evidence of job search. The transition between these benefit programmes is characterised by a shift from 'work preparation conditionality' (i.e. ESA) to 'full conditionality' (i.e. Jobseekers allowance) (see Department for Work and Pensions, 2010).

[20]Benefit offices such as JobCentre Plus in the U.K. will typically be open 9am-5pm (i.e. the typical working day.)

[21]In the context of the framework employed in Section 5.3.1, where preferences were defined over both over consumption and leisure, there would be a certain inconsistency in maintaining the assumption of fixed working hours but then allowing individuals to take the fraction k of their working day off from work (to attend the benefit office). In this case some individuals would have incentives to take time off and simply enjoy more leisure, thus getting closer to their optimal labour supply. Whilst the framework adopted in the current section is more restrictive, it avoids these conceptual issues.

5.4.1 Work decision when receiving C.

When receiving C an individual may choose to either comply with the no-work condition and have utility $u(C)$, or instead work and have (expected) utility $u[\omega(1-k)+\tilde{C}-\chi]$. Let $\underline{\omega}^k(\rho,\phi,C)$ denote the critical wage at which an individual is indifferent between these two choices. Formally:

$$\underline{\omega}^k = \frac{(C-\tilde{C})+\chi}{1-k} \tag{5.36}$$

We once more assume that at the point of indifference an individual will choose *not* to work. All those with $\omega \in [\omega_0, \underline{\omega}^k]$ will therefore not work when receiving C; whilst all those with $\omega \in (\underline{\omega}^k, \infty)$ will work when receiving C. The properties of $\underline{\omega}^k$ with respect to ρ, ϕ and C parallel that for $\underline{\omega}$ in (5.9):

$$\underline{\omega}^k_C > 0 \;\; ; \;\; \underline{\omega}^k_{CC} > 0 \quad ; \quad \underline{\omega}^k_\phi > 0 \;\; \underline{\omega}^k_{\phi\phi} > 0 \quad ; \quad \underline{\omega}^k_\rho \begin{Bmatrix} \geq \\ < \end{Bmatrix} 0 \;\; ; \;\; \underline{\omega}^k_{\rho\rho} < 0$$

It is also straightforward to see that $\underline{\omega}^k$ is increasing in k.

5.4.2 Application decisions

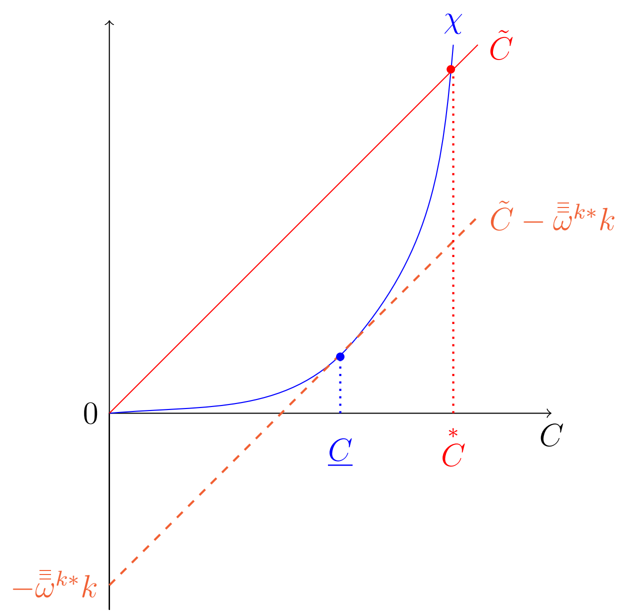
Turning to application decisions, there are thus two groups to consider: (i) those with $\omega \leq \underline{\omega}^k$ who do not work when receiving C; and (ii) those with $\omega > \underline{\omega}^k$ who do work when receiving C. The decision to apply for the former group is straightforward; they will apply if $\omega \leq \bar{\omega}^k \equiv C$. Contrastingly, an individual with $\omega > \underline{\omega}^k$ will only apply for C if $\tilde{C} - \omega k > \chi$, and thus if $\omega \leq \bar{\bar{\bar{\omega}}}^k$, where $\bar{\bar{\bar{\omega}}}^k$ satisfies:

$$\bar{\bar{\bar{\omega}}}^k(\rho,\phi,C,k) \equiv \frac{\tilde{C}-\chi}{k} \tag{5.37}$$

The intuition is as follows.[22] An individual who would choose to work when receiving C

[22]The assumption of CARA preferences greatly simplifies the analysis of the condition $\tilde{C}-\omega k > \chi$. Indeed, because χ is independent of ω we can solve for a unique $\bar{\bar{\bar{\omega}}}$ above which an individual who would work when receiving C will not apply. If χ were to take the DARA form in (5.26) there may be multiple solutions to $\tilde{C}-\omega k > \chi$.

Figure 5.4: The Upper Bound $\bar{\bar{\bar{\omega}}}^{k*}$.



Notes. This figure graphically illustrates the upper bound $\bar{\bar{\bar{\omega}}}^{k*}$. This is simply $\bar{\bar{\bar{\omega}}}^{k} = (\tilde{C} - \chi)/k$ evaluated at the value of C which maximises $(\tilde{C} - \chi)$, i.e. $\underline{C}$.

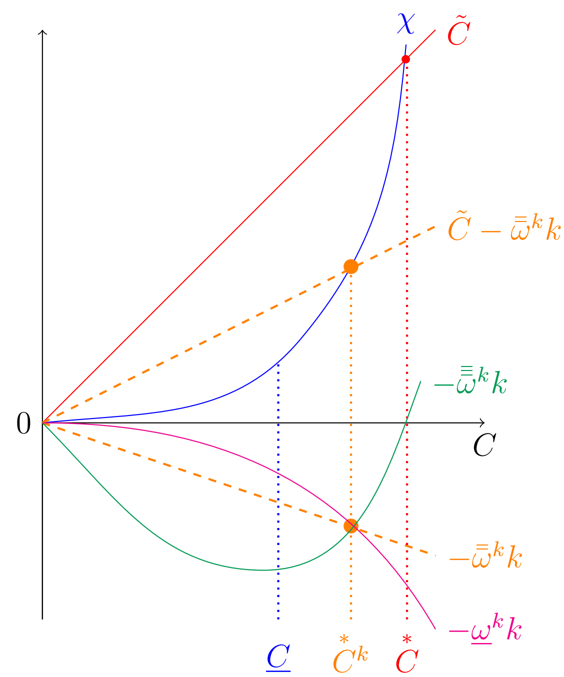
will only apply if the expected benefit $(\tilde{C})$ *net* of foregone earnings (ωk) exceeds the risk premium associated with the variance in benefit income. However, for an individual with $\omega > \bar{\bar{\bar{\omega}}}^{k}$ the opportunity cost of foregone earnings is too high, thus rendering an application suboptimal. Straight away then, we can see that the imposition of a time requirement generates more realistic application decisions because higher wage individuals will not apply.

An upper bound on $\bar{\bar{\bar{\omega}}}^{k}$. From (5.37), we can take the analysis a step further by recognising that $\bar{\bar{\bar{\omega}}}^{k}$ must, *ceteris paribus*, take its maximum value at the value of C which maximises $\tilde{C} - \chi$. For any given enforcement parameter pair (ρ, ϕ), the expected benefit net of the risk premium is maximised at $\underline{C}(\rho, \phi)$, where:

$$\underline{C}(\rho, \phi) \equiv \frac{\overset{*}{C}(\rho, \phi)}{2} = \frac{1 - \rho(1+\phi)}{\eta\rho(1-\rho)(1+\phi)^2} \tag{5.38}$$

Given that $\underline{C}$ is as stationary point it will hold that $\chi' \sigma_C(\rho, \phi, \underline{C}) = \tilde{C}_C(\rho, \phi, \underline{C})$. From

Figure 5.5: The critical categorical benefit $\overset{*}{C}^k$



Notes. This figure illustrates that the critical benefit size $\overset{*}{C}^k$ occurs at the point where $\underline{\omega}^k = \bar{\omega}^k = \bar{\bar{\omega}}^k$. The negative portion of the figure captures this in terms of the earnings that these marginal individuals would forego were they to work when receiving the categorical benefit. The value of C at which these foregone earnings intersect will intuitively depend on the size of k. Indeed, one can readily establish that $\partial \underline{\omega}_C^k k / \partial k > 0$, $\partial \bar{\omega}_C k / \partial k > 0$, but $\partial \bar{\bar{\omega}}_C^k k / \partial C = 0$. An increase in k thus lowers the value of the categorical benefit at which the intersection occurs.

this it follows that we can define a critical net wage which is independent of C, above which no able individual will choose to apply for C and work.

Result 3a. $\forall\ C < \overset{*}{C}(\rho, \phi)$, *$\bar{\bar{\omega}}^k$ is bounded above by a critical productivity $\bar{\bar{\omega}}^{k*}$ that depends only on the standard enforcement parameters (and k). Formally:*

$$\bar{\bar{\omega}}^{k*}(\rho, \phi, k) \equiv \bar{\bar{\omega}}^k\left(\rho, \phi, \underline{C}(\rho, \phi), k\right) \tag{5.39}$$

where $\bar{\bar{\omega}}^k(\rho, \phi, C, k) < \bar{\bar{\omega}}^{k*}(\rho, \phi, k) \quad \forall\ C \neq \underline{C}$

This upper bound is graphically illustrated in Figure 5.4.

Whether or not an individual who would work when receiving C will apply for it will clearly depend on where $\underline{\omega}^k$ lies relative to $\bar{\bar{\omega}}^k$ on the net wage continuum. Intuitively, if $\bar{\bar{\bar{\omega}}}^k < \underline{\omega}^k$ then no such individual will apply. With this in mind, we state the following result.

Result 3b.

$$\underline{\omega}^k \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \bar{\bar{\omega}}^k \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \bar{\bar{\bar{\omega}}}^k \text{ if } C \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \overset{*}{C}^k \equiv \frac{2[1-\rho(1+\phi)-k]}{\eta\rho(1-\rho)(1+\phi)^2} \tag{5.40}$$

An immediate implication of (5.40) is that $\overset{*}{C}^k < \overset{*}{C}\ \forall\ k > 0$, such that the imposition of time costs lowers the critical categorical benefit above which no individual who would work when receiving C will choose to apply. Notice further that if $k \geq 1 - \rho(1+\phi)$ then enforcement is independent of the categorical benefit and no individual who would work when receiving C will ever apply. Figure 5.5 provides the graphical intuition for this result. In addition, Figure 5.6 depicts how application choices differ over the net wage continuum for the three cases depicted in (5.40).

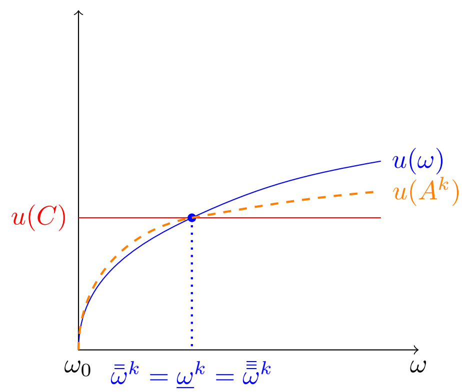
5.5 Concluding Remarks

This chapter has analysed the decisions of workers - who differ continuously in their wage - to apply for a categorical benefit that is targeted at individuals who are unable to work, but administered with Type II (false award) classification errors. The categorical benefit is *ex-ante* conditional on an applicant being unable to work; and *ex-post* conditional on a recipient not working. Recipients may also be required to spend a fraction of the working day at the benefit office. Any recipient who works risks being detected. Upon detection a fraudulent recipient is required to repay the benefit in its entirety, in addition to paying a fine that is *proportional* to the benefit size.

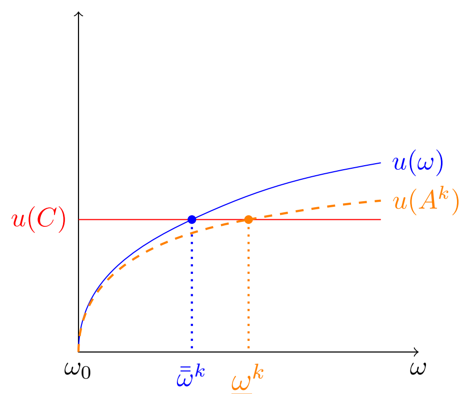
Under both CARA and DARA (CRRA) preferences, the risk premium associated with the variance in benefit income is *convex-increasing* in the benefit size. This yields the interesting (and initially surprising) result that individuals who would work conditional on receiving the categorical benefit can be deterred from applying for it through setting the benefit level sufficiently *high*. In the case of CARA preferences the risk premium is independent of the individual wage rate and so one risk premium characterises the

Figure 5.6: The relationship between $\underline{\omega}^k$, $\bar{\omega}^k$ and $\bar{\bar{\bar{\omega}}}^k$

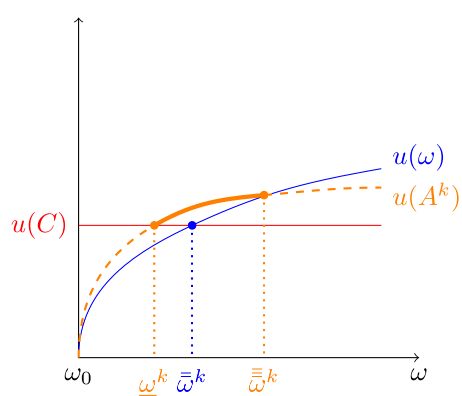
(a) $C = \overset{*}{C}^k$



(b) $C > \overset{*}{C}^k$



(c) $C < \overset{*}{C}^k$



Notes. This figure graphically depicts the three cases stated in (5.40). Note that $A^k = \omega(1-k)+\tilde{C}-\chi$. Subplots (a) and (b) illustrate that if $\underline{\omega}^k \geq \bar{\bar{\bar{\omega}}}^k$ (i) all those with $\omega \in [\omega_0, \bar{\omega}^k]$ will apply for C and, if awarded it, not work; whilst (ii) all those with $\omega \in (\bar{\omega}^k, \infty)$ will *not* apply for C. Contrastingly, subplot (c) illustrates that if $\underline{\omega}^k < \bar{\bar{\bar{\omega}}}^k$ (i) all those with $\omega \in [\omega_0, \underline{\omega}^k]$ will apply for C and, if awarded it, not work; (ii) all those with $\omega \in (\underline{\omega}^k, \bar{\bar{\bar{\omega}}}^k]$ will apply for C and, if awarded, continue to work; whilst (iii) those with $\omega \in (\bar{\bar{\bar{\omega}}}^k, \infty)$ do *not* apply for C.

attitudes of all individuals to risk. For all cases where the standard enforcement parameters (detection probability, penalty rate) alone do not provide full deterrence, there is a critical categorical benefit level above which no able individual will apply for the categorical benefit and work when receiving it. This critical benefit level is increasing the leniency of the standard enforcement parameters. Contrastingly, for DARA preferences which exhibit constant relative risk aversion, the risk premium is decreasing in the individual wage. A necessary but not sufficient condition to deter individuals from applying for the categorical benefit and subsequently working is that the standard enforcement parameters be set sufficiently high. Conditional on this being achieved, a categorical benefit set sufficiently high can be used to provide deterrence.

When the only *ex-post* condition imposed on recipients is that they do not work, an unsatisfactory implication of the above results is that higher earners will all apply for the benefit if enforcement is too lenient. In reality, however, higher earning individuals are unlikely to apply for categorical benefits even if they face a positive probability of being awarded them. One explanatory factor for why is that receiving benefits is in many cases a time consuming activity. A simple extension imposing a fully enforced time requirement on recipients (i) preserves the result that a benefit set sufficiently high can achieve deterrence; but now also (ii) generates a critical wage above which no able individual will apply because the opportunity cost of foregone earnings is simply too high.

Imposing a fine proportional to the benefit incorrectly obtained parallels how real-world systems operate. It is, however, precisely because the fine is an increasing function of the benefit size that we obtain the result that a benefit set sufficiently high can fully deter individuals from violating *ex-post* conditionality.

Appendix A Derivations and Proofs

Derivation of the risk premium $\chi(A, \sigma)$

By the standard method established in Pratt (1964), we take first- and second- order Taylor approximations around $(A, 1-\mathcal{H})$ on the left- and right- sides of equation (5.4) in the main text, respectively, to obtain:

$$u[A-\chi, 1-\mathcal{H}] \approx u[A, 1-\mathcal{H}] - \chi \cdot u_x[A, 1-\mathcal{H}] \ , \tag{A.1}$$

$$\begin{aligned} E(u) &\approx \rho \left\{ u(A, 1-\mathcal{H}) + \alpha_0 u_x(A, 1-\mathcal{H}) + \frac{1}{2} u_{xx} \alpha_0^2 (A, 1-\mathcal{H}) \right\} \\ &+ (1-\rho) \left\{ u(A, 1-\mathcal{H}) + \alpha_1 u_x(A, 1-\mathcal{H}) + \frac{1}{2} \alpha_1^2 u_{xx}(A, 1-\mathcal{H}) \right\} \\ &= u[A, 1-\mathcal{H}] + u_x[A, 1-\mathcal{H}] \cdot \underbrace{E(\alpha)}_{=0} + \frac{1}{2} u_{xx}[A(\omega), 1-\mathcal{H}] \cdot E(\alpha^2) \ . \end{aligned} \tag{A.2}$$

where, defining $\sigma(\rho, \phi, C) \equiv E[(\alpha - E(\alpha))^2] = E(\alpha^2)$:

$$\begin{aligned} \sigma &= \rho(\alpha_0)^2 + (1-\rho)(\alpha_1)^2 = [\rho(1-\rho)^2 + \rho^2(1-\rho)](1+\phi)^2 C^2 \\ &= \rho(1-\rho)(1+\phi)^2 C^2 \end{aligned} \tag{A.3}$$

Through substituting (A.3) into (A.2) and subsequently combining (A.1) and (A.2), it is straightforward to arrive at the risk premium as defined in (5.5).

CARA Preferences

Preferences satisfying $u_{xx} + \eta u_x = 0$

The solution to $u_{xx} + \eta u_x = 0$ simply follows standard second-order linear differential equation methods (see Simon and Blume, 1994, pp.647-648). Given that labour is exogenously fixed in the intensive margin, let $\tilde{u}(x) \equiv u(x, l)$. The differential equation defining absolute risk aversion is therefore $\tilde{u}''(x) + \eta\tilde{u}'(x) = 0$. Substituting in $\tilde{u} = e^{zx}$ (and thus $\tilde{u}' = ze^{zx}$ and $\tilde{u}'' = z^2e^{zx}$) yields $e^{zx}[z^2 + \eta z] = 0$ and thus the characteristic equation $z(z + \eta) = 0$, which has roots $z = 0$ and $z = -\eta$. The solution is therefore:

$$\tilde{u}(x) = ae^0 - be^{\eta x} = a - be^{-\eta x}$$

It is straightforward to show that a utility function in consumption and leisure (fixed) which satisfies these properties - and the initial assumptions placed on preferences - is:

$$u(x, l) = 1 - \psi(l)e^{-\eta x} \;\; ; \; \psi(l) > 0 \; , \; \psi' < 0 \; , \; \psi'' > 0 \; , \; \psi\psi'' > (\psi')^2$$

We can readily establish that:

$$u_x = \eta\psi(l)e^{-\eta x} > 0 \; ; \;\; u_{xx} = -\eta^2\psi(l)e^{-\eta x} < 0 \; ; \;\; -u_{xx}/u_x = \eta > 0$$

$$u_l = -\psi'(l)e^{-\eta x} > 0 \; ; \;\; u_{ll} = -\psi''(l)e^{-\eta x} < 0 \; ; \;\; u_{xl} = \eta\psi'(l)e^{-\eta x}$$

Strict concavity requires $u_{xx}u_{ll} - u_{xl}^2 > 0$. We have:

$$\begin{aligned} u_{xx}u_{ll} - u_{xl}^2 &= \eta^2\psi(l)\psi''(l)e^{-2\eta x} - \eta^2[\psi'(l)]^2e^{-2\eta x} \\ &= \eta^2e^{-2\eta x}\left\{\psi(l)\psi''(l) - [\psi'(l)]^2\right\} > 0 \end{aligned}$$

by the assumptions placed on u.[23]

[23]Normality of leisure requires $u_l u_{xx} - u_x u_{xl} < 0$. With these preferences, note however that $u_l u_{xx} - u_x u_{xl} = 0$. We can readily show that any reservation wage in the economy is independent of unearned income. Suppose all individuals in society receive an unconditional universal benefit (B): then the reservation wage ($\bar{\omega}$) satisfies:

$$\begin{aligned} 1 - \psi(1)e^{-\eta B} = 1 - \psi(1 - \mathcal{H})e^{-(\bar{\omega}\mathcal{H} + B)} &\Rightarrow \psi(1)e^{-\eta B} = \psi(1 - \mathcal{H})e^{-\eta(\bar{\omega}H)}e^{-\eta B} \\ &\Rightarrow \frac{\psi(1)}{\psi(1 - \mathcal{H}} = e^{-\eta\bar{\omega}\mathcal{H}} \\ &\Rightarrow \bar{\omega} = -\left(\frac{1}{\eta}\right)\text{Log}\left[\frac{\psi(1)}{\psi(1 - \mathcal{H})}\right]/\mathcal{H} > 0 \end{aligned}$$

The derivatives of the critical net wage $\underline{\omega}$

Totally differentiating the definition of $\underline{\omega}$ in (5.9) in the main text with respect to C, ρ and ϕ gives:

$$\begin{aligned} u_x(\underline{\omega}\mathcal{H} + \tilde{C} - \chi, 1 - \mathcal{H}) \cdot (\underline{\omega}_C\mathcal{H} + \tilde{C}_C - \chi'\sigma_C) &= u_x(C, 1) \\ u_x(\underline{\omega}\mathcal{H} + \tilde{C} - \chi, 1 - \mathcal{H}) \cdot (\underline{\omega}_\phi\mathcal{H} + \tilde{C}_\phi - \chi'\sigma_\phi) &= 0 \\ u_x(\underline{\omega}\mathcal{H} + \tilde{C} - \chi, 1 - \mathcal{H}) \cdot (\underline{\omega}_\rho\mathcal{H} + \tilde{C}_\rho - \chi'\sigma_\rho) &= 0 \end{aligned} \tag{A.4}$$

Solving for $\underline{\omega}_C$, $\underline{\omega}_\phi$ and $\underline{\omega}_\rho$ respectively, gives the derivatives in (5.10).

It is straightforward to establish from (5.10) that $\omega_C > 0$. From the properties of the CARA preferences in (5.7) we can write $\underline{\omega}_C$ as:

$$\underline{\omega}_C = \left(\frac{1}{\mathcal{H}}\right)\left[(1 - \tilde{C}_C) + \chi'\sigma_C\right]$$

Noting that $\tilde{C}_C = 1 - \rho(1 + \phi)$ is maximised at $\rho = 0$; whilst $\chi'\sigma_C$ is minimised at $\rho = 0$, it must be the case that $\underline{\omega}_C$ is minimised at $\rho = 0$, where $\underline{\omega}_C|_{\rho=0} = 0$. Since we only consider $\rho > 0$ it follows that $\underline{\omega}_C > 0$.

With respect to the enforcement parameters, it is unambiguously the case that $\underline{\omega}_\phi > 0$ and $\underline{\omega}_\rho \geq 0 \;\forall\; \rho \leq 1/2$. However, because $\chi'\sigma_\rho < 0 \;\forall\; \rho > 1/2$, we can see that the sign of $\underline{\omega}_\rho$ when $\rho > 1/2$ will depend on:

$$\begin{aligned} \underline{\omega}_\rho \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \text{ if } \frac{d}{d\rho}(\chi - \tilde{C}) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 &\Leftrightarrow \frac{1}{2}\eta\underbrace{(1 - 2\rho)}_{<0}(1 + \phi)^2C^2 + (1 + \phi)C \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \\ &\Leftrightarrow C \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \bar{C} \equiv \frac{2}{\eta(2\rho - 1)(1 + \phi)} \;\; \forall\; \rho > \frac{1}{2} \end{aligned} \tag{A.5}$$

Derivation of the Critical Benefit Level $\overset{*}{C}$ and its Properties

The relationship between χ and $\tilde{C}$ is simply given by:

$$\chi \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \tilde{C} \Leftrightarrow \frac{1}{2}\eta\rho(1-\rho)(1+\phi)^2 C^2 \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} [1-\rho(1+\phi)]C \Leftrightarrow C \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \overset{*}{C} \equiv \frac{2[1-\rho(1+\phi)]}{\eta\rho(1-\rho)(1+\phi)^2} \tag{A.6}$$

Differentiating C^* with respect to ρ and ϕ, respectively, gives:

$$\begin{aligned} \overset{*}{C}_\rho &= -\left\{\frac{2}{\eta(1+\phi)^2}\right\}\left\{\frac{\rho(1-\rho)(1+\phi)+[1-\rho(1+\phi)](1-\rho)}{\rho^2(1-\rho)^2}\right\} \\ &= -\left\{\frac{2}{\eta(1+\phi)^2}\right\}\cdot\left\{\frac{\rho^2(1+\phi)+(1-2\rho)}{\rho^2(1-\rho)^2}\right\} = -\left\{\frac{2}{\eta(1+\phi)^2}\right\}\cdot\left\{\frac{\rho^2\phi+(1-\rho)^2}{\rho^2(1-\rho)^2}\right\} < 0 \end{aligned}$$

$$\begin{aligned} \overset{*}{C}_\phi &= \left\{\frac{2}{\eta\rho(1-\rho)}\right\}\cdot\left\{\frac{-\rho(1+\phi)^2-2[1-\rho(1+\phi)](1+\phi)}{(1+\phi)^4}\right\} \\ &= \left\{\frac{2}{\eta\rho(1-\rho)}\right\}\cdot\left\{\frac{\rho(1+\phi)^2-2(1+\phi)}{(1+\phi)^4}\right\} = \left\{\frac{2}{\eta\rho(1-\rho)}\right\}\cdot\left\{\frac{\rho(1+\phi)-2}{(1+\phi)^3}\right\} < 0 \end{aligned}$$

$$\overset{*}{C}_{\rho\phi} = \left\{\frac{2}{\eta(\rho(1-\rho))}\right\}\cdot\left\{\rho(1+\phi)^3+3(2-\rho(1+\phi))(1+\phi)^2(1+\phi)^6\right\} > 0 \tag{A.7}$$

Proof that $\overset{*}{C} \leq \bar{C}$

Suppose, contrary to (5.20), that $\overset{*}{C} > \bar{C}$. This implies:

$$\begin{aligned} \left\{\frac{2[1-\rho(1+\phi)]}{\eta\rho(1-\rho)(1+\phi)^2}\right\} > \left\{\frac{2}{\eta(2\rho-1)(1+\phi)}\right\} &\Leftrightarrow [1-\rho(1+\phi)](2\rho-1) > \rho(1-\rho)(1+\phi) \\ &\Leftrightarrow (2\rho-1) > \rho^2(1+\phi) \\ &\Leftrightarrow -(1-\rho)^2 > \rho^2\phi \end{aligned} \tag{A.8}$$

This is a contradiction and thus $\overset{*}{C} < \bar{C}$. Q.E.D.

Derivation of Explicit Risk Premium under CARA preferences

The purpose of this section is to illustrate why the Taylor approximation of the risk-premium greatly increases the tractability of the analysis. Under the stated CARA preferences the condition defining the risk premium, χ, is:

$$1-\psi(1-\mathcal{H})e^{-\eta(A-\chi)} = \rho[1-\psi(1-\mathcal{H})e^{-\eta(A+\alpha_0)}] + (1-\rho)[1-\psi(1-\mathcal{H})e^{-\eta(A+\alpha_1)}]$$

which reduces to the condition:

$$\begin{aligned} e^{-\eta(A-\chi)} &= \rho e^{-\eta(A+\alpha_0)} + (1-\rho)e^{-\eta(A+\alpha_1)} \\ \Rightarrow \quad e^{-\eta A}e^{\eta x} &= \rho e^{-\eta A}e^{-\eta\alpha_0} + (1-\rho)e^{-\eta A}e^{-\eta\alpha_1} \\ \Rightarrow e^{\eta\chi} &= \rho e^{-\eta\alpha_0} + (1-\rho)e^{-\eta\alpha_1} \end{aligned}$$

Taking Logs (to the base e) thus gives:

$$\eta\chi = \text{Log}\left\{\underbrace{\rho e^{-\eta\alpha_0} + (1-\rho)e^{-\eta\alpha_1}}_{>1}\right\}$$

and finally:

$$\chi(\rho,\phi,C) = \left(\frac{1}{\eta}\right)\text{Log}\left\{\rho e^{-\eta\alpha_0} + (1-\rho)e^{-\eta\alpha_1}\right\} \tag{A.9}$$

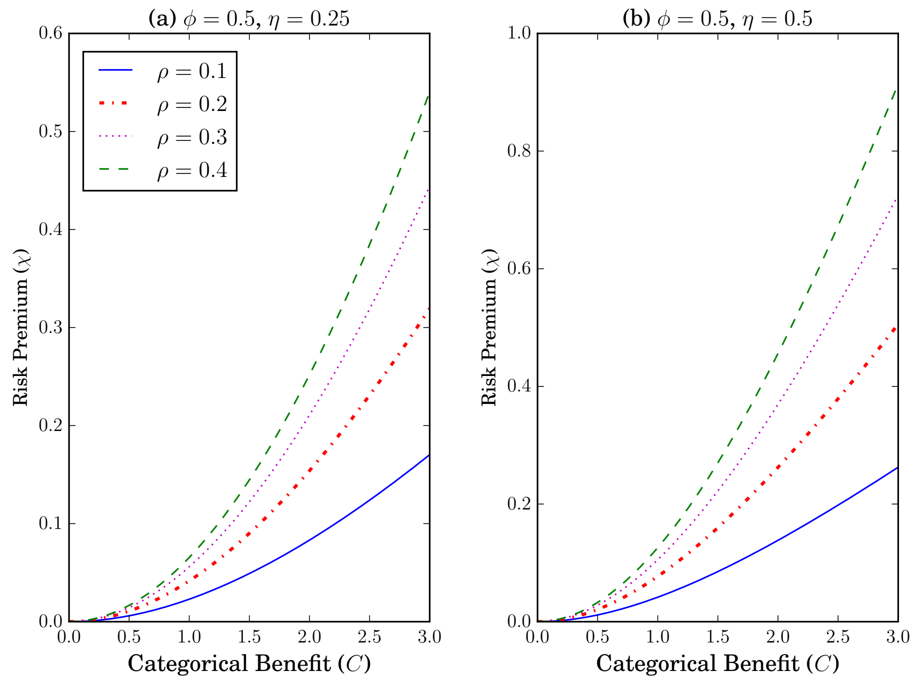
Figure 5.7 illustrates that this is *convex-increasing* in C.

DARA preferences

Preferences satisfying $u_{xx}x + \eta u_x = 0$

Drawing from Simon and Blume (1994), once more let $\tilde{u}(x) = u(x,l)$ where l is a constant. Letting $q(x) = \tilde{u}'(x)$ the second-order differential equation $\tilde{u}''(x)x + \eta u'(x) = 0$ becomes $(dq/dx)x + \eta q = 0$ and thus $dq/q = -\eta(dx/x)$. Integrating both sides yields $\ln q = -\eta(\ln x + c)$ where c is an integration constant. Taking the exponential of both

Figure 5.7: Explicit CARA Risk Premium



Notes: The values of η chosen here are 0.25 (panel a) and 0.5 (panel b). Higher values are permitted in the literature (see Berloffa and Simmons, 2003).

sides yields $q = e^{-\eta(\ln x + c)} = e^{\ln x^{-\eta} e^{-\eta c}} = x^{-\eta} b$; where $b = e^{-\eta c}$ is a constant. Given that $q = \tilde{u}'(x)$ and thus $\tilde{u}'(x) = x^{-\eta} b$ we integrate both sides to obtain:

$$\tilde{u} = \begin{cases} a + b\frac{x^{1-\eta}}{1-\eta} & : 0 < \eta < 1 \\ a + b\ln x & : \eta = 1 \end{cases}$$

A utility function in consumption and leisure (fixed) which satisfies these properties - and the initial assumptions placed on preferences - is:

$$u(x,l) = \frac{x^{1-\eta}\psi(l)}{1-\eta} \; ; \quad \eta \neq 1 \; , \quad \psi(l) > 0 \quad , \quad \psi'(l) > 0 \quad , \psi''(l) < 0$$

We can readily establish that:

$$u_x = x^{-\eta}\psi(l) > 0 \quad , \quad u_{xx} = -\eta x^{-\eta-1}\psi(l)/x < 0 \quad , \quad -\frac{u_{xx}}{u_x} = \frac{\eta}{x} > 0$$

$$u_l = \frac{x^{1-\eta}\psi'(l)}{1-\eta} > 0 \quad , \quad u_{ll} = \frac{x^{1-\eta}\psi''(l)}{1-\eta} < 0 \quad ; \quad u_{xl} = x^{-\eta}\psi'(l) > 0$$

Notice that strict concavity requires $-\frac{\eta}{(1-\eta)}\psi(l)\psi''(l) > [\psi'(l)]^2$. Formally[24]:

$$\begin{aligned} u_{xx}u_{ll} - u_{xl}^2 &= -\frac{\eta}{1-\eta}x^{-2\eta}\psi(l)\psi''(l) - x^{-2\eta}[\psi'(l)]^2 \\ &= -x^{-2\eta}\left\{\frac{\eta}{1-\eta}\psi(l)\psi''(l) + [\psi'(l)]^2\right\} > 0 \end{aligned}$$

Finally, note that normality of leisure is satisfied because:

$$u_x u_{xl} - u_l u_{xx} = x^{-2\eta}\psi(l)\psi'(l)\left\{1 + \eta x^{-1}/(1-\eta)\right\} > 0$$

Proof that χ is convex increasing in C

To proceed, we note that: $\tilde{C}_C = [1-\rho(1+\phi)] = \tilde{C}/C$; whilst $\sigma_C = \rho(1-\rho)(1+\phi)^2 C = 2\sigma/C$. Taken together we obtain $\tilde{C}\sigma_C = 2\sigma\tilde{C}_C$. It then follows that:

$$\begin{aligned} \frac{d\chi}{dC} = \frac{1}{2}\eta\left\{\frac{A\sigma_C - \sigma\tilde{C}_C}{A^2}\right\} &= \frac{1}{2}\eta\left\{\frac{(A-\tilde{C})\sigma_C + \tilde{C}\sigma_C - \sigma\tilde{C}_C}{A^2}\right\} \\ &= \frac{1}{2}\eta\left\{\frac{(A-\tilde{C})\sigma_C + 2\sigma\tilde{C}_C - \sigma\tilde{C}_C}{A^2}\right\} \\ &= \frac{1}{2}\eta\left\{\frac{(A-\tilde{C})\sigma_C + \sigma\tilde{C}_C}{A^2}\right\} \end{aligned}$$

[24]Suppose $\psi(l) = l^{\alpha}(1-\eta)$ $\alpha \in (0,1)$ such that $\psi'(l) = \alpha(1-\eta)l^{\alpha(1-\eta)-1}$ and $\psi''(l) = \alpha(1-\eta)[\alpha(1-\eta)-1]l^{\alpha(1-\eta)-2}$. Then it is straightforward to check that the condition for strict concavity becomes $\eta[1-\alpha(1-\eta)] > \alpha(1-\eta)^2$.

We can now use the fact that $\sigma_{CC} = \rho(1-\rho)(1+\phi)^2 = \sigma_C/C$ to show that:

$$\begin{aligned}
\frac{d^2\chi}{dC^2} &= \frac{1}{2}\eta\left\{\frac{A[(A-\tilde{C})\sigma_{CC} + \sigma_C\tilde{C}_C] - 2\tilde{C}_C[(A-\tilde{C})\sigma_C + \sigma\tilde{C}_C]}{A^3}\right\} \\
&= \frac{1}{2}\eta\left\{\frac{(A/C)[(A-\tilde{C})\sigma_C + 2\sigma\tilde{C}_C] - 2\tilde{C}_C[(A-\tilde{C})\sigma_C + \sigma\tilde{C}_C]}{A^3}\right\} \\
&= \frac{1}{2}\eta\left\{\frac{A[(A-\tilde{C})\sigma_C + 2\sigma\tilde{C}_C] - 2\tilde{C}[(A-\tilde{C})\sigma_C + \sigma\tilde{C}_C]}{CA^3}\right\} \\
&= \frac{1}{2}\eta\left\{\frac{(A-\tilde{C})\sigma_C(A-2\tilde{C}) + 2\sigma\tilde{C}_C(A-\tilde{C})}{CA^3}\right\}
\end{aligned}$$

But of course $2\sigma\tilde{C}_C = 2\tilde{C}\sigma/C = \tilde{C}\sigma_C$ and thus:

$$= \frac{1}{2}\eta\left\{\frac{(A-\tilde{C})[\sigma_C(A-\tilde{C}) + \tilde{C}\sigma_C - \tilde{C}\sigma_C]}{CA^3}\right\}$$

and thus finally

$$= \frac{1}{2}\eta\frac{(A-\tilde{C})^2}{A^3}\sigma_{CC} > 0 \tag{A.10}$$

Derivation and sign of $d\chi/d\rho$

$$
\begin{aligned}
\frac{d\chi}{d\rho} = \frac{1}{2}\eta \cdot \left\{\frac{A\sigma_\rho - \sigma\tilde{C}_\rho}{A^2}\right\} \Leftrightarrow \frac{d\chi}{d\rho}\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \Leftrightarrow \sigma_\rho A \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \sigma\tilde{C}_\rho & \\
\Leftrightarrow (1-2\rho)(\omega\mathcal{H} + \tilde{C}) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} - \rho(1-\rho)(1+\phi)C & \\
\Leftrightarrow (1-2\rho)\omega\mathcal{H} \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} - C \begin{Bmatrix} (1-2\rho)[1-\rho(1+\phi)] \\ +\rho(1-\rho)(1+\phi) \end{Bmatrix} & \\
\Leftrightarrow (2\rho-1)\omega\mathcal{H} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} C\left\{1 - 2\rho + \rho^2(1+\phi)\right\} & \\
\Leftrightarrow (2\rho-1)\omega\mathcal{H} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} C[(1-\rho)^2 + \rho^2\phi] &
\end{aligned}
\tag{A.11}
$$

If $\rho \leq 1/2$ it must always be the case that $d\chi/d\rho > 0$. Otherwise, $d\chi/d\rho > 0$ requires $C > (2\rho-1)\omega\mathcal{H}/[(1-\rho)^2 + \rho^2\phi]$ as stated above and in (5.28).

The derivatives of the critical net wage $\underline{\omega}^D$

Differentiating the definition of $\underline{\omega}^D$ in (5.29) with respect to C, ρ and ϕ yields:

$$
\begin{aligned}
u_x[A(\underline{\omega}^D) - \chi(A(\underline{\omega}^D),\sigma), 1-\mathcal{H}] \cdot \left\{(\underline{\omega}_C^D\mathcal{H} + \tilde{C}_C)(1-\chi_A) - \chi_\sigma\sigma_C\right\} &= u_x(C,1) \\
u_x[A(\underline{\omega}^D) - \chi(A(\underline{\omega}^D),\sigma), 1-\mathcal{H}] \cdot \left\{(\underline{\omega}_\phi^D\mathcal{H} + \tilde{C}_\phi)(1-\chi_A) - \chi_\sigma\sigma_\phi\right\} &= 0 \\
u_x[A(\underline{\omega}^D) - \chi(A(\underline{\omega}^D),\sigma), 1-\mathcal{H}] \cdot \left\{(\underline{\omega}_\rho^D\mathcal{H} + \tilde{C}_\rho)(1-\chi_A) - \chi_\sigma\sigma_\rho\right\} &= 0
\end{aligned}
\tag{A.12}
$$

Solving for $\underline{\omega}_C^D$, $\underline{\omega}_\rho^D$ and $\underline{\omega}_\phi^D$ gives (5.30). We now demonstrate that $\underline{\omega}_C^D > 0$. Given that $u_x(C,1)/u_x(A-\chi, 1-\mathcal{H}) > 1$, it suffices to show that $1 + \chi_\sigma\sigma_C > \tilde{C}_C(1-\chi_A)$ to establish $\underline{\omega}_C^D > 0$. Substituting in $\chi_A = -\frac{1}{2}\frac{\eta}{A^2}\sigma$, we can write $1 + \chi_\sigma \sigma_C - \tilde{C}(1 - \chi_A)$ as:

$$1 + \frac{\eta}{A}\frac{\sigma}{C} - \frac{\tilde{C}}{C}\left\{1 + \frac{1}{2}\frac{\eta}{A^2}\sigma\right\} = \left(1 - \frac{\tilde{C}}{C}\right) + \frac{\eta}{A}\frac{\sigma}{C}\left\{1 - \frac{1}{2}\frac{\tilde{C}}{A}\right\} > 0 \tag{A.13}$$

because $C/\tilde{C} < 1$ and $\tilde{C}/A < 1 \; \forall \; \rho > 0$.

Derivation of the Function $\Gamma(\rho, \phi)$

To derive the function $\Gamma(\rho, \phi)$ we write:

$$\begin{aligned}
\chi \geq \tilde{C} &\Leftrightarrow \frac{1}{2}\left\{\frac{\eta\rho(1-\rho)(1+\phi)^2 C}{\omega\mathcal{H} + C[1-\rho(1+\phi)]}\right\} \geq [1-\rho(1+\phi)] \\
&\Leftrightarrow \eta\rho(1-\rho)(1+\phi)^2 C \geq 2\omega\mathcal{H}[1-\rho(1+\phi)] + 2C[1-\rho(1+\phi)]^2 \\
&\Leftrightarrow \left\{\eta\rho(1-\rho)(1+\phi)^2 - 2[1-\rho(1+\phi)]^2\right\} C \geq 2\omega\mathcal{H}[1-\rho(1+\phi)] \\
&\Leftrightarrow \left\{\frac{-(1+\phi)^2(\eta+2)\rho^2 + [\eta(1+\phi)^2 + 4(1+\phi)]\rho - 2}{2[1-\rho(1+\phi)]}\right\} C \geq \omega\mathcal{H} \\
&\Leftrightarrow \left\{\frac{-(1+\phi)^2(\frac{\eta}{2}+1)\rho^2 + (1+\phi)[\frac{\eta}{2}(1+\phi) + 2]\rho - 1}{[1-\rho(1+\phi)]}\right\} C \geq \omega\mathcal{H} \\
&\Leftrightarrow \Gamma(\rho, \phi) \cdot C \geq \omega\mathcal{H}
\end{aligned} \tag{A.14}$$

where:

$$\Gamma(\rho, \phi) \equiv \frac{-(1+\phi)^2(\frac{\eta}{2}+1)\rho^2 + (1+\phi)[\frac{\eta}{2}(1+\phi) + 2]\rho - 1}{[1-\rho(1+\phi)]} \tag{A.15}$$

We can immediately establish that:

$$\begin{aligned}
\lim_{\rho(1+\phi)\to 0} \Gamma(\rho, \phi) &= \frac{-1}{1} = -1 \\
\lim_{\rho(1+\phi)\to 1} \Gamma(\rho, \phi) &= \frac{\left\{-\frac{\eta}{2} - 1 + \frac{\eta}{2}(1+\phi) + 1\right\}}{0} = \frac{\eta\phi/2}{0} = +\infty
\end{aligned} \tag{A.16}$$

The derivatives of Γ with respect to ρ and ϕ are given by:

$$\begin{aligned}
\frac{\partial\Gamma}{\partial\rho} &= \left\{\begin{matrix}[1-\rho(1+\phi)]\langle -2\rho(1+\phi)^2(\frac{\eta}{2}+1)+(1+\phi)\left[\frac{\eta}{2}(1+\phi)+2\right]\rangle \\ +(1+\phi)\langle -(1+\phi)^2(\frac{\eta}{2}+1)\rho^2+(1+\phi)\left[\frac{\eta}{2}(1+\phi)+2\right]\rho-1\rangle\end{matrix}\right\}/[1-\rho(1+\phi)]^2 \\
&= \left\{\begin{matrix}-2\rho(1+\phi)^2(\frac{\eta}{2}+1)+\frac{\eta}{2}(1+\phi)^2+2(1+\phi) \\ +2\rho^2(1+\phi)^3(\frac{\eta}{2}+1)+\frac{\eta}{2}\rho(1+\phi)^3+2\rho(1+\phi)^2 \\ -\rho^2(1+\phi)^3(\frac{\eta}{2}+1)+\frac{\eta}{2}\rho(1+\phi)^3+2\rho(1+\phi)^2-(1+\phi)\end{matrix}\right\}/[1-\rho(1+\phi)]^2 \\
&= \left\{\begin{matrix}\rho(1+\phi)^2(2-\eta)+\frac{\eta}{2}(1+\phi)^2+(1+\phi) \\ +\rho^2(1+\phi)^3(\frac{\eta}{2}+1)+\eta\rho(1+\phi)^3\end{matrix}\right\}/[1-\rho(1+\phi)]^2>0
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial\Gamma}{\partial\phi} &= \left\{\begin{matrix}[1-\rho(1+\phi)]\langle -2\rho^2(1+\phi)(\frac{\eta}{2}+1)+[\eta(1+\phi)+2]\rho\rangle \\ -\rho^3(1+\phi)^2(\frac{\eta}{2}+1)+\rho^2\left[\frac{\eta}{2}(1+\phi)^2+2(1+\phi)\right]-\rho\end{matrix}\right\}/[1-\rho(1+\phi)]^2 \\
&= \left\{\begin{matrix}-2\rho^2(1+\phi)(\frac{\eta}{2}+1)+\rho\left[\eta(1+\phi)+2\right] \\ +2\rho^3(1+\phi)^2(\frac{\eta}{2}+1)-\rho^2\left[\eta(1+\phi)^2+2(1+\phi)\right] \\ -\rho^3(1+\phi)^2(\frac{\eta}{2}+1)+\rho^2[\frac{\eta}{2}(1+\phi)^2+2(1+\phi)]-\rho\end{matrix}\right\}/[1-\rho(1+\phi)]^2 \\
&= \left\{\begin{matrix}\rho^3(1+\phi)^2(\frac{\eta}{2}+1)-\rho^2(1+\phi)[\eta(1+\phi)+\eta+2] \\ +\rho[\eta(1+\phi)+1]\end{matrix}\right\}/[1-\rho(1+\phi)]^2 \\
&= \frac{1}{2}\rho\left\{\eta(\rho-1)(1+\phi)\langle\rho(1+\phi)-2\rangle+2\langle\rho(1+\phi)-1\rangle^2\right\}/[1-\rho(1+\phi)]^2>0
\end{aligned}$$

Chapter 6

Concluding Remarks

6.1 A summary of the thesis

This thesis has focused on an economy where a fraction of the population is unable to work, whilst the remaining fraction is composed of individuals who are able to work but differ continuously in their productivity (as in Mirrlees, 1971). The unable subpopulation provides the basis for targeted, or categorical, transfers within this framework. Indeed, under the strong assumption that categorical status (unable, able) is perfectly observable, it will be optimal under a utilitarian welfare criterion to target resources solely at the unable so long as the social marginal value of income (smvi) of these individuals exceeds the average smvi of the able.[1] In reality, however, categorical status is difficult to identify and this gives rise to classification errors. Further, targeted benefits are typically conditioned not just on initial eligibility, but also in an *ex-post* dimension whereby recipients must comply with certain behavioural requirements or restrictions. These may also be imperfectly enforced. Taking seriously these real-world facets of welfare provision allows one to analyse both design and enforcement issues within this framework.

Part I of this thesis focused on the design issues. The starting point (chapter 2) was to abstract from tax revenue considerations and analyse the optimal division of a fixed benefit budget between (i) a tax-free categorical benefit that is *ex-ante* conditional on

[1]Whether or not this will be achieved will depend on the size of the budget in place for benefit expenditure.

an applicant being unable to work and *ex-post* conditional on a recipient not working; and (ii) a tax-free universal benefit that is received unconditionally by all. Importantly, the categorical benefit is administered with Type I (false rejection) and Type II (false award) classification errors. The latter error type gives rise to enforcement issues because individuals who are truly able to work are receiving the categorical benefit. Whether or not these ineligible recipients will choose to work will depend crucially on the enforcement of the *ex-post* condition. In this regard, the two binary cases of No Enforcement and Full Enforcement were considered. Under the former there are no effective mechanisms in place to deter individuals from working, such that all able individuals apply for the benefit. However, under the latter there are fully effective mechanisms in place. The application decisions of able (ineligible) individuals are thus endogenous to the benefit size and only those of lower productivity will choose to apply because they are better off receiving the categorical benefit and not working. The two enforcement structures yield different results for optimal welfare provision and, relatedly, the welfare effects of classification errors. Whilst Type I errors are always welfare reducing, the effect of Type II errors on social welfare may differ across the enforcement regimes. An increase in the propensity to make Type II errors (i) unambiguously reduces social welfare under the No Enforcement regime; but (ii) may increase social welfare under the Full Enforcement regime. The intuition is that under Full Enforcement 'leakage' of the categorical benefit is restricted only to able individuals of lower productivity and, consequently, may play a redistributive role within the able subpopulation. Numerical examples where this arises were provided.

The following two chapters (3 and 4) then proceeded to relax the assumption of a fixed benefit budget and analyse the case where income tax revenue is used to finance benefit expenditure (and any exogenous revenue requirement for spending outside of welfare). The natural starting point was the Perfect Discrimination case where categorical transfers are perfectly targeted at unable individuals. A key result in the literature is that the optimal linear income tax expression with perfectly administered categorical transfers can be written as in the uni-dimensional model where individuals differ solely in productivity (Viard, 2001a,b). This result depends, however, on the assumption that categorical transfers eliminate inequality in the average net smvi across categorical groups (in our case the unable and able). Yet, if categorical transfers are financed by tax revenue there may be cases where it is suboptimal to impose a tax rate that generates enough revenue to eliminate this inequality. This is particularly likely to hold if there is a large dependent population in need of transfers and/or the government has significant spending commitments outside of welfare. With this in mind, the contribution of Chapter 3 was to (i) provide expressions for optimal linear and piecewise linear tax rates that allow for the persistence of between group inequality at the optimum; and, importantly, (ii) provide numerical examples where between-group inequality does indeed persist at the optimum. Analytical examples where between-group inequality persists at the flat tax optimum were also provided for the special case where (i) preferences have a constant labour elasticity (see Atkinson, 1990; Saez, 2001); and (ii) taxation is purely redistributive (i.e. no revenue requirement).

Chapter 4 then reintroduced classification errors into the analysis and provided an expression for the optimal linear tax rate for the case where the *ex-post* 'no-work' condition is fully enforced. The full enforcement assumption ensures that the decision of an able individual to apply for the categorical benefit is endogenous to the tax rate. Consequently, an increase in the tax rate generates both direct and indirect behavioural effects. The direct effect is simply that found in all conventional analyses: i.e. individuals adjust their labour supply in the intensive margin in response to a change in the tax rate. The indirect effect, meanwhile, captures the fact that a *ceteris paribus* increase in the tax rate induces additional able individuals to apply for the categorical benefit. This has implications for *both* the tax revenue side of the budget constraint (because additional individuals are awarded the categorical benefit and stop working) *and* the benefit expenditure side of the budget constraint (because additional individuals are awarded the categorical benefit). In the optimal tax expression that results an important term in both equity (numerator) and efficiency (denominator) considerations is the elasticity of the distribution function with respect to individual productivity, evaluated at the critical productivity at or below which able individuals choose to apply for the categorical benefit.

Part II of this thesis focused on individual decisions under risk and contained Chapter 5. This chapter modelled the decision of able individuals to apply for the categorical benefit, conditioning on whether or not they would choose to comply with the *ex-post* 'no-work' requirement if awarded it. Drawing on the economics of crime (see Becker, 1968), working recipients risk being detected with some probability and, if detected, they are sanctioned. The sanction considered here mirrors that available to actual benefit authorities: the fraudulent recipient is required[2] to repay the benefit in its entirety in addition to paying a fine proportional to the benefit size[3]. The proportionality of the fine to the benefit size plays a crucial role in the analysis of risk. Indeed, following Arrow (1970) and Pratt (1964) this chapter captures the risk that a working recipient faces via the risk premium associated with the variance in benefit income, the latter of which is convex-increasing in the benefit size. Under CARA preferences the risk premium is therefore also convex-increasing in the benefit size. The implication is that there will be a critical benefit level above which no able individual who would choose to work conditional on receiving the benefit will choose to apply. Intuitively, this critical level is increasing in the leniency of the standard enforcement parameters (detection probability, penalty rate). For any categorical benefit set below this critical level all able individuals in the economy will choose to apply, with those of higher productivity choosing to work. This result is readily made more realistic through requiring recipients to spend a fraction of the 'working' day at the benefit office, thereby imposing an opportunity cost on ineligible claimants. In particular, this places an upper bound on the productivity type at or below which able individuals choose to apply for the categorical benefit.

6.2 Going Forwards

There are a number of interesting directions in which the framework presented in thesis could be furthered.

Heterogeneity in disability. Throughout this thesis the unable subpopulation has taken on a passive role. They are unable to provide any labour and, consequently, do not respond to incentives of the tax system. A richer - but more complex - model may therefore consider an economy where individuals differ in the extent of their disability, as could be captured through differing quantity constraints on labour supply. The benefit structure and *ex-post* conditions placed on benefit receipt would necessarily change to capture the differing degrees of disability. In such a setting Type I errors may take on a more significant role. In particular, individuals who are incorrectly denied a

[2]In many cases the financial sanction will be 'offered' as an alternative to prosecution.

[3]In the U.K., for example, fraudulent recipients may be offered an administrative penalty of 50% of the overpaid benefit (see Department for Work and Pensions, 2015).

categorical benefit may be forced to work more than they would in the optimum with no such errors. This would also introduce new equity and efficiency considerations into the optimal tax expression.

Societal concern for the unable beyond economic status. Under a standard utilitarian social welfare function an able individual who is voluntarily unemployed will have the same smvi as an unable individual when both receive the same in benefit income. It would, however, be of interest to analyse a case where society expressly cares about the distinction between voluntary unemployment and disability/quantity constraints. Indeed, it would seem likely that many people would place greater value on a transfer received by an individual who is unable to work, rather than someone who can work as much as they wish but choose to be unemployed. Some initial work in this area has been undertaken by Saez and Stantcheva (2013).

Bibliography

Akerlof, G. A. (1978). The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *American Economic Review*, 68:8–19.

Allingham, M. and Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 1:323–338.

Apps, P., Long, N. V., and Rees, R. (2014). Optimal Piecewise Linear Income Taxation. *Journal of Public Economic Theory*, 16(4):523–545.

Arrow, K. J. (1970). *Essays in the theory of risk-bearing*. Markham, Chicago.

Atkinson, A. and Sutherland, H. (1989). Analysis of a partial basic income scheme. In Atkinson, A., editor, *Poverty and Social Security*. Harvester Press, Hemel Hempstead.

Atkinson, A. B. (1990). Public economics and the economic public. *European Economic Review*, 34:225–248.

Atkinson, A. B. (1995). *Public Economics in Action: The Basic Income/ Flat Tax Proposal*. Oxford University Press, Oxford.

Atkinson, A. B. (2015). *Inequality: what Can be done?* Harvard University Press, London, England.

Atkinson, A. B. and Stiglitz, J. E. (1980). *Lectures on Public Economics*. McGraw-Hill Book Company, New York.

Autor, D. H. and Duggan, M. G. (2003). The rise in the disability rolls and the decline in unemployment. *Quarterly Journal of Economics*, 118(February):157–206.

Autor, D. H. and Duggan, M. G. (2006). The growth in the Social Security Disability rolls: a fiscal crisis unfolding. *The journal of economic perspectives : a journal of the American Economic Association*, 20(3):71–96.

Beath, J., Lewis, G., and Ulph, D. (1988). Policy targeting in a new welfare framework with poverty. In Hare, P. G., editor, *Surveys in Public Sector Economics*, pages 161–185. Basil Blackwell, Oxford, UK.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Benitez-Silva, H., Buchinsky, M., and Rust, J. (2004). How large are the classification errors in the social security disability award process? *NBER Working Paper Series*, (10219).

Berloffa, G. and Simmons, P. (2003). Unemployment Risk, Labour Force Participation and Savings. *Review of Economic Studies*, 70(3):521–539.

Besley, T. and Coate, S. (1992). Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs. *American Economic Review*, 82(1):249–261.

Blackorby, C. and Donaldson, D. (1988). Cash Versus Kind, Self-selection, and Efficient Transfers. *The American Economic Review*, 78:691–700.

Boadway, R., Marceau, N., and Sato, M. (1999). Agency and the design of welfare systems. *Journal of Public Economics*, 73(1):1–30.

Boone, J. and Bovenberg, L. (2006). Optimal welfare and in-work benefits with search unemployment and observable abilities. *Journal of Economic Theory*, 126(1):165–193.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants. *The American economic review*, 79(3):482–503.

Bound, J. (1991). The health and earnings of rejected disability insurance applicants: reply. *The American Economic Review*, 81(5):1427–1434.

Burgess, P. (1992). Compliance with unemployment-insurance job-search regulations. *Journal of Law and Economics*, 35(2):371–396.

Callan, T., O'Donoghue, C., Sutherland, H., and Wilson, M. (1999). *Comparative analysis of Basic Income proposals: UK and Ireland*. Number February. The Microsimulation Unit, Department of Applied Economics, University of Cambridge, Cambridge.

Chen, S. and van der Klaauw, W. (2008). The work disincentive effects of the disability insurance program in the 1990s. *Journal of Econometrics*, 142(2):757–784.

Chetty, R. (2006). A New Method of Estimating Risk Aversion. *The American Economic Review*, 96(5):1821–1834.

Crawford, R., Emmerson, C., and Ifs, S. K. (2014). Public Finances: Risks on Tax, Bigger Risks on Spending? In Emmerson, C., Johnson, P., and Miller, H., editors, *IFS Green Budget 2014*, chapter 2, pages 23–50. The Institute of Fiscal Studes.

Cuff, K. (2000). Optimality of workfare with heterogeneous preferences. *Canadian Journal of Economics/Revue canadienne d'Economique*, 33(1):149–174.

Currie, J. (2004). The take up of social benefits. *NBER Working Paper Series*, (10488).

Department for Work and Pensions (2010). Sanction Policy : In respect of fraudulent Social Security Benefit Claims. Technical Report April, The Department of Work and Pensions.

Department for Work and Pensions (2011). Fraud and Error Penalties and Sanctions: Equality Impact Assessment. Technical Report October, The Department of Work and Pensions.

Department for Work and Pensions (2015). Penalties Policy: In Respect of Social Security Fraud and Error. Technical Report January.

Diamond, P. and Mirrlees, J. (1978). A Model of Social Insurance with Variable Retirement. *Journal of Public Economics*, 10:295–336.

Diamond, P. and Sheshinski, E. (1995). Economic aspects of optimal disability benefits. *Journal of Public Economics*, 57:1–23.

Diamond, P. A. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *American Economic Review*, 88:83–95.

Dilnot, A., Kay, J., and Morris, C. (1984). *The Reform of Social Security.* Clarendon Press., Oxford.

Edgeworth, F. (1897). The pure theory of taxation. *The Economic Journal*, 7(25):46–70,226–38, 550–71.

Eissa, N. and Liebman, J. (1996). Labor Supply Response to the Earned Income Tax Credit. *The Quarterly Journal of Economics*, 111(2).

Fuller, D. L., Ravikumar, B., and Zhang, Y. (2015). Unemployment Insurance Fraud and Optimal Monitoring. *American Economic Journal: Macroeconomics*, 7(2):249–290.

Goodin, R. (1985). Erring on the Side of Kindness in Social Welfare Policy. *Policy Sciences*.

Goodin, R. (1992). Towards a minimally presumptuous social welfare policy. *Arguing for Basic Income, Verso, London*.

Hellwig, M. (1986). The Optimal Linear Income Tax Revisited. *Journal of Public Economics*, 31.

Helpman, E. and Sadka, E. (1978). Optimal taxation of full income. *International Economic Review*, 19(1):247–251.

Immonen, R., Kanbur, R., Keen, M., and Tuomala, M. (1998). Tagging and Taxing: The Optimal Use of Categorical and Income Information in DesigningTax/Transfer Schemes. *Economica*, 65(258):179–192.

Jacquet, L. (2006). Optimal disability assistance when fraud and stigma matter. *Queen's Economics Department Working Paper*, (1098).

Jacquet, L. (2014). Tagging and redistributive taxation with imperfect disability monitoring. *Social Choice and Welfare*, 42(2):403–435.

Kaplow, L. (2008). *The Theory of Taxation and Public Economics*. Princeton University Press, New Jersey.

Kleven, H. J. and Kopczuk, W. (2011). Transfer Program Complexity and the Take-Up of Social Benefits. *American Economic Journal: Economic Policy*, 3(February):54–90.

Kreiner, C. T. and Tranaes, T. (2005). Optimal Workfare with Voluntary and Involuntary Unemployment*. *Scandinavian Journal of Economics*, 107(3):459–474.

Lewis, G. and Ulph, D. (1988). Poverty, Inequality and Welfare. *The Economic Journal*, 98(390):117–131.

Mankiw, N. G. and Weinzierl, M. (2010). The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution. *American Economic Journal: Economic Policy*, 2(1):155–176.

Mcinnes, R. (2012). ESA and incapacity benefit statistics. Technical report, House of Commons Library.

Mirrlees, J. (1971). An Exploration in the Theory of Optimal Income Taxation. *The review of economic studies*, 38(2):175–208.

Mkandawire, T. (2005). Targeting and universalism in poverty reduction. *Social Policy and Development , United Nations Research Institute for Social Development*, (Program Paper Number 23).

Moffitt, R. (1983). An economic model of welfare stigma. *The American Economic Review*, 73(5):1023–1035.

Nagi, S. (1969). *Disability and Rehabilitation: Legal, Clinical, and Self-Concepts and Measurement*. Ohio State University Press.

Nichols, A. and Zeckhauser, R. (1982). Targeting Transfers through Restrictions on Recipients. *American Economic Review*, 72(2):372–377.

ODI (2014). Disability prevalence estimates 2011/ 12. Technical Report April 1950, Department for Work and Pensions.

ONS (2012). Statistical Bulletin - 2011 Census: Population Estimates for the United Kingdom, 27 March 2011. Technical Report March 2011, Office for National Statistics (ONS).

Parsons, D. (1980). The decline in male labor force participation. *The Journal of Political Economy*, 88(1):117–134.

Parsons, D. (1991). The health and earnings of rejected disability insurance applicants: comment. *The American Economic Review*, 81(5):1419–1426.

Parsons, D. (1996). Imperfect 'tagging'in social insurance programs. *Journal of Public Economics*, 62:183–207.

Paulus, A. and Peichl, A. (2009). Effects of flat tax reforms in Western Europe. *Journal of Policy Modeling*, 31:620–636.

Peichl, A. (2014). Flat-rate tax systems and their effect on labor markets. *IZA World of Labor*, (October):1–10.

Piketty, T. and Saez, E. (2012). Optimal labor income taxation. *National Bureau of Economic Research Working Paper 18521*.

Pratt, J. (1964). Risk Aversion in the Small and in the Large. *Econometrica: Journal of the Econometric Society*, 32(1-2):122–136.

Rhodes, C. and Mcinnes, R. (2014). The welfare cap. Technical report, House of Commons Library.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies*, 68:205–229.

Saez, E. and Stantcheva, S. (2013). Generalized Social Marginal Welfare Weights for Optimal Tax Theory. *NBER Working Paper Series*, (18835).

Salanié, B. (2002). Optimal demogrants with imperfect tagging. *Economics Letters*, 75:319–324.

Seade, J. (1982). On the Sign of the Optimum Marginal Income Tax. *The Review of Economic Studies*, 49(4):637.

Sheshinski, E. (1972). The Optimal Linear Income Tax. *The Review of Economic Studies*.

Simon, C. and Blume, L. (1994). *Mathematics for economists*. W.W. Norton & Company, Inc., first edit edition.

Skocpol, T. (1991). Targeting within universalism: Politically viable policies to combat poverty in the United States. In Jencks, C. and Peterson, P. E., editors, *The urban underclass*. The Brookings Institution, Washington, D.C.

Slack, S. and Ulph, D. (2014). Optimal Universal and Categorical Benefits with Classification Errors and Imperfect Enforcement. *University of St Andrews, School of Economics & Finance Discussion Papers*, (1411).

Slack, S. E. (2015). Revisiting the optimal linear income tax with categorical transfers. *Economics Letters*, 134:73–77.

Slemrod, J., Yitzhaki, S., Mayshar, J., and Lundholm, M. (1994). The optimal two-bracket linear income tax. *Journal of Public Economics*, 53:269–290.

SSA (2012). Annual Statistical Report on the Social Security Disability Insurance Program , 2011. Technical Report 13, Social Security Aministration.

Stern, N. (1976). On the specification of models of optimum taxation. *Journal of Public Economics*, 6:123–162.

Stern, N. (1982). Optimum taxation with errors in administration. *Journal of Public Economics*, 17:181–211.

Van Parijs, P. (2004). Basic Income: A Simple and Powerful Idea for the Twenty-First Century. *Politics & Society*, 32(1):7–39.

Viard, A. (2001a). Optimal Categorical Transfer Payments: The Welfare Economics of Limited LumpSum Redistribution. *Journal of Public Economic Theory*, 3(December 2000):483–500.

Viard, A. (2001b). Some Results on the Comparative Statics of Optimal Categorical Transfer Payments. *Public Finance Review*, 29(2):148–180.

Von Wachter, T., Song, J., and Manchester, J. (2011). Trends in employment and earnings of allowed and rejected applicants to the social security disability insurance program. *American Economic Review*, 101(December):3308–3329.

Wolf, D. and Greenberg, D. (1986). The Dynamics of Welfare Fraud : An Econometric Duration Model in Discrete Time. *Journal of Human Resources*, 21(4):437–455.

Yaniv, G. (1986). Fraudulent collection of unemployment benefits: A theoretical analysis with reference to income tax evasion. *Journal of Public Economics*, 30.

Yaniv, G. (1997). Welfare fraud and welfare stigma. *Journal of Economic Psychology*, 18(4):435–451.

Yitzhaki, S. (1974). A note on Income Tax Evasion: A Theoretical Analysis. *Journal of public economics*, 3:201–202.