INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
NOTE TO USERS

This reproduction is the best copy available.

UMI
RICE UNIVERSITY

Essays in Microeconomics:
Wage Subsidy in an Optimal Redistribution Program
and
Bundling Hardware and Software

by

Alexei Zarovnyi

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Peter Mieszkowski, Chairman
Allyn R. and Gladys M. Cline
Professor of Economics and Finance

George R. Zodrow
Professor of Economics

Suchan Chae
Associate Professor of Economics

Steven C. Curran
Associate professor of Management
And Psychology

Houston, Texas
October, 2000
ABSTRACT

Essays in Microeconomics:
Wage Subsidy in an Optimal Redistribution Program
and
Bundling Hardware and Software

By

Alexei Zarovnyi

My dissertation consists of two unrelated essays. In the first essay, "Wage Subsidy in an Optimal Redistribution Program", I analyze the efficiency of income transfers and wage subsidies as instruments of income redistribution in an optimal taxation framework. I extend the Mirrlees model (1971) of income inequality by specifying a model in which individuals' productivity and wages depend on investment in skill acquisition in addition to ability. The principal result of the research is that a wage subsidy has an important role to play in an optimal system of income maintenance. In the second essay, "Bundling Hardware and Software", a class of simple models is analyzed to explain prevailing bundling practices in computer markets. A profit-maximizing monopolist may over-provide hardware-software bundles, practice "pure bundling" when preferences are symmetric with respect to software, and under-bundle and under-produce software when preferences are asymmetric.
ACKNOWLEDGMENTS

Many people helped me to complete this research successfully. My deepest thanks and appreciation to all of them! First and foremost I would like to thank my dissertation adviser Dr. Peter Mieszkowski. Technical skills, even very advanced, and ability to work hard are not sufficient to become a good researcher, the most important factors are economic intuition - ability to see economics of the problem beyond pages of the math equations and computer programs - something that cannot be learnt simply by taking courses and could only come as a result of years of working with an experienced researcher if he is willing to teach. Dr. Mieszkowski became such a person in my life. I had privilege to work with him for the last 5 years both as his student and his research assistant. From the selection of the topic for the to the final stages of writing my “wage subsidy” paper his help is difficult to overestimate. These are his efforts, time and devotion that made this dissertation possible.

Another person I want to thank is Dr. Suchan Chae, “Bundling Hardware and Software” is our joint work. I’d like to thank him for his patience, persistence and all the research skills I learned while working with him on this paper.

Also I am very grateful to Dr. George R. Zodrow and Dr. Steven C. Currah for reading through several versions of wage subsidy manuscript and for valuable comments and suggestions.

Finally I want to thank my wife Lena for her support and understanding.
# TABLE OF CONTENTS

CHAPTER I. WAGE SUBSIDY IN AN OPTIMAL REDISTRIBUTION PROGRAM .......................................................... 1

1.1. Introduction ............................................................................................................................................... 1

1.2. An extended Mirrlees model of income inequality .......................................................... 6
   1.2.1. Wage subsidy in the original Mirrlees model ......................................................... 6
   1.2.2. The basic model ............................................................................................................... 12
   1.2.3. Completed simulations and results obtained ....................................................... 17

1.3. Extensions of and other experiments with the basic model ........................................... 36
   1.3.1. Different objective functions .................................................................................. 36
   1.3.2. The efficiency of subsidies for education ............................................................... 42
   1.3.3. The financing of a wage subsidy with a two bracket income tax ......................................................... 48
   1.3.4. Analysis of an Earned Income Tax Credit (EITC) ........................................ 51

1.4. A two-market model of a wage subsidy ........................................................................ 54
   1.4.1. A two market model .................................................................................................... 57
   1.4.2. Two-market model with no education ................................................................. 58
   1.4.3. Two-market model with linear informal market production ........................................ 62
   1.4.4. Two-market model with formal market production and variable leisure valuation ......................................................... 68

1.5. Concluding remarks ......................................................................................................................... 76

CHAPTER II. Bundling hardware and Software ................................................................. 78

2.1. Introduction ............................................................................................................................................... 78

2.2. Basic model ............................................................................................................................................ 80
   2.2.1. Profit maximization ..................................................................................................... 83
   2.2.2. Welfare maximization ................................................................................................. 91

2.3. Software bundling with symmetric preferences ............................................................. 96

2.4. Software bundling with asymmetric preferences ......................................................... 98

2.5. Conclusion ............................................................................................................................................. 105

Appendix ......................................................................................................................................................... 106

References ...................................................................................................................................................... 113
LIST OF TABLES

Table 1.1. Wage subsidy in the original Mirrlees model ................................................. 9

Table 1.2a. Comparison of linear income tax and linear income tax with wage subsidy structures. Non-deductible private education expenses .......................... 26

Table 1.3a. Comparison of linear income tax and linear income tax combined with wage subsidy structures. Tax deductible private education expenses ................................................................................................................. 28

Table 1.2b. Linear income tax and linear income tax with wage subsidy structures, no demogrant. Non-deductible private education expenses ................. 30

Table 1.3b. Linear income tax and linear income tax with wage subsidy structures, no Demogrant. Tax deductible private education expenses ................. 30

Table 1.4. Comparison of universal demogrant and universal demogrant with wage subsidy income maintenance structures. Minimum income requirement ........................................................................................................ 39

Table 1.5. Comparison of universal demogrant and universal demogrant with wage subsidy income maintenance structures. Poverty index minimization ........................................................................................................ 40

Table 1.6. Comparison of phased-out demogrant and phased-out demogrant with wage subsidy income maintenance structures. Minimum income requirement ........................................................................................................ 41

Table 1.7. Comparison of phased-out demogrant and phased-out demogrant with wage subsidy income maintenance structures. Poverty index minimization ........................................................................................................ 41

Table 1.8. Comparison of linear income tax and linear income tax with private cost of education subsidy structures .......................................................................................................................... 46

Table 1.9. Comparison of one and two-bracket linear income tax structures .......... 50

Table 1.10. Comparison of universal demogrant and EITC in linear income tax with the wage subsidy redistribution scheme .............................................. 53

Table 1.11. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures .................................................. 59

Table 1.12. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures .................................................. 65
Table 1.13. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures. Distribution A ................. 73

Table 1.14. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures. Distribution B .................. 73

Table 2.1. Reservation prices .......................................................................................................................... 82

Table 2.2. Infeasible system choices ............................................................................................................... 84

Table 2.3. Feasible and undominated system choices .................................................................................... 85

Table 2.4. System choices for welfare maximization ...................................................................................... 93

Table 2.5. Reservation prices with asymmetric preferences ........................................................................ 99

Table 2.6. Reservation prices with asymmetric preferences ...................................................................... 100

Table A.1. Infeasible system choices ............................................................................................................. 107

Table A.2. Dominated system choices .......................................................................................................... 107

Table A.3. Feasible and undominated system choices .................................................................................. 108

Table A.4. Feasible and undominated system choices .................................................................................. 108

Table A.5. Feasible and undominated system choices for profit maximizer .............................................. 110

Table A.6. Feasible and undominated system choices for welfare maximizer ........................................... 111
LIST OF FIGURES

Fig. 1.1. Abilities .............................................................................................................. 8

Fig. 1.2. Utility comparison for no-tax, linear income tax alone and linear income tax with wage subsidy structures ................................................................. 10

Fig. 1.3. No-tax, linear income tax alone and linear income tax with wage subsidy labor supply decisions ............................................................................................ 10

Fig. 1.4. Before and after subsidy wages ........................................................................ 11

Fig. 1.5. Utility comparison for no-tax, linear income tax alone and linear income tax with wage subsidy structures ................................................................. 34

Fig. 1.6. No-tax, linear income tax alone and linear income tax with wage subsidy labor supply decisions ............................................................................................ 34

Fig. 1.7. No-tax, linear income tax alone and linear income tax with wage subsidy education choices ...................................................................................................... 35

Fig. 1.8. No tax and linear income tax alone observable wages. Linear income tax with wage subsidy observable and real (after subsidy) wages ......................... 35

Fig. 1.9. No-tax, before and after price of education subsidy utility .................................. 47

Fig. 1.10. No-tax, before and after price of education subsidy labor supply decisions .... 47

Fig. 1.11. No-tax, before and after price of education subsidy education choices ............ 48

Fig. 1.12. No-tax wage distribution .................................................................................. 59

Fig. 1.13a. No-tax, linear income tax and linear income tax with wage observable (formal) market labor supply .............................................................................. 60

Fig. 1.13b. No-tax, linear income tax and linear income tax with wage unobservable (informal) market labor supply ................................................................. 61

Fig 1.14. No-tax, linear income tax and linear income tax with wage subsidy utility .......... 61

Fig. 1.15. No-tax wages .................................................................................................. 63
CHAPTER I.
WAGE SUBSIDY IN AN OPTIMAL REDISTRIBUTION PROGRAM.

1.1. Introduction.

At least three different arguments have been developed in favor of wage subsidies or related work-contingent income maintenance systems. The first, put forward by Ballard (1988, 1996) and Triest (1994), is that the marginal efficiency cost (MEC) of a wage subsidy is quite low relative to general income transfers when a system of income redistribution is added to general government expenditures on defense and/or education. The common sense behind this result is that wage subsidies produce an efficiency gain as they encourage workers who are net recipients of subsidies to increase their work effort, thus offsetting the tax distortions associated with general revenue requirements. This efficiency gain offsets the incremental efficiency loss of the tax increase necessary to finance the wage subsidy.

Beaudry and Blackorby (1998) have developed a different case for wage subsidies as a replacement for or supplement to income transfers. They postulate a two sector economy consisting of a formal sector in which earnings are observed by the government and an informal (underground) economy in which earnings are not reported. Income transfers will be limited and ineffective in this situation as taxpayers will increase their activity in the informal economy. Moreover, individuals who are productive in both sectors will report low official earnings and collect income transfers while earning a substantial amount of income in the informal sector. Beaudry and Blackorby point out that a wage subsidy will partly offset these shortcomings by serving as a self-selection
mechanism. Persons whose productivity is low in both sectors will be induced by wage subsidies to shift their activity to the formal sector and to work longer hours there. In addition, high productivity persons who remain in the informal economy will no longer receive public transfers.

A third general argument for wage subsidies is suggested by the recent criticism of the welfarist approach to income maintenance by Kanbur, Keen and Tuomola (1994) and by Besley and Coate (1995). These authors argue that the welfarist approach, by valuing both net income and leisure, has little relevance to the contemporary policy debate as politicians and the public give little, if any, weight to the leisure of the poor. Instead, taxpayers are concerned with the maintenance of the net income of the poor above some minimum standard and/or with the minimization of the number of persons who are living in poverty. Besley and Coate show that the adoption of a non-welfarist objective function implies a general work requirement for the poor. It seems self-evident that the disregard of leisure in welfare supports the adoption of wage subsidies in place of income transfers. As they are a compromise between unrestricted transfers and a general work requirement, wage subsidies also give poor persons more flexibility than does a uniform work requirement.

This paper evaluates these three arguments by extending the Mirrlees model of income determination in which abilities are the sole determinant of wages. I do so by allowing for skill acquisition and for variations in work exertion while on the job. My formulation preserves Mirrlees' key informational assumption -- i.e., that a person's ability is non-observable by the government while his/her wage rate (productivity) is appraisable by both employers and public agencies. The introduction of endogenous skill
acquisition into the determination of wages allows for two incentive margins: the amount of time spent on the job and the amount of education and/or the amount of effort expended while on it. The generalization allows for a richer and more realistic analysis of the incentive effects of wage subsidies as a redistributive instrument.

Kesselman (1969, 1976, 1977), was the first to analyze the effects of wage subsidies on labor supply. His later work for a graded earnings tax (GET) was a direct extension of the Mirrlees model. As in this system one factor — ability — determines a person's productivity. If the government observes a person's wage it also observes his/her ability.

Not surprisingly, Kesselman found that the wage-related tax/subsidy scheme is much more redistributive and produces a significantly higher level of welfare than possible under an optimal income tax. Similarly, Blomquist (1981) concluded, "a tax which is a linear affine function of each individual's wage rate will give a social welfare optimum very close to the first-best optimum." I have confirmed these results for the basic Mirrlees' model by substituting a linear wage subsidy financed by a linear income tax for the graded earnings tax. Such a system equalizes net wages and utilities across the population for an additive social welfare function. Every person whose productivity is below the average will have their wage rate raised to the mean while the net wage rate of more productive individuals will be reduced to the average. This strong equalitarian result is questionable on grounds of political feasibility and ethical "just dessert." Moreover, it is based on the highly unrealistic assumption, as is the analysis at Kesselman and Blomquist, that persons will produce at their natural ability level, independent of
what they actually receive in net compensation. The inclusion of work exertion and skill acquisition in a more general model yields a more sensible, useful formulation.

A second stimulus for endogenizing worker productivity is the result of the well-known literature which uses estimates of labor supply to quantify the efficiency-equity trade-off involved in different systems of redistribution. Browning and Johnson (1984) use the consumer surplus measure developed by Harberger (1971) to estimate the welfare effects of an incremental increase in taxes on high-income groups which is used to finance a small universal demogrant. The efficiency cost per dollar of net transfers to the poor of such a system is estimated to be $2.50, i.e., the welfare cost to high-income tax payers of a dollar transfer to the poor is $3.50. One explanation of this large dead-weight loss is that the redistributive program is introduced into an economy already badly distorted by general non-redistributive taxes. In addition, a demogrant is not a "targeted" means-tested form of redistribution.

Ballard (1988, 1996) obtains a smaller estimate for the efficiency cost of a incremental demogrant system. He uses a computational general equilibrium model with explicitly specified utility and production functions and calculates equivalent variation measures of the welfare loss for different income groups. Significantly, Ballard also finds the efficiency cost of a small, exogenously determined wage-subsidy program to be zero or even negative. The net-of-tax wage increase for low income groups reduces the tax distortion from existing general revenue taxes; these efficiency gains offset the increased distortions imposed on high income groups.

Triest (1994), using a closely related methodology, confirms Ballard's general conclusions in his analysis of a small tax increase in an economy distorted by taxes. He
finds that the efficiency cost of a marginal increase in a work-related system of redistribution is small relative to an incremental increase of a general income transfer.

Although these results are for small incremental changes in taxes and transfers, they have two general implications. First, the findings suggest that, since work-related forms of redistribution are more efficient, they should be used to implement a given amount of redistribution or at least supplement the existing system of redistribution. Second, as the dead-weight loss of a wage subsidy may be close to zero, the total amount of redistribution a society will choose to make with this instrument will be significantly larger than when it uses general income transfers to maximize social welfare.

The model developed here, which allows for skill acquisition, qualifies the conclusion that a work-contingent system of redistribution is necessarily more efficient than income transfers. I account for two incentive margins, (1) the time spent on the job and (2) the investment in skill acquisition. ¹ The advantage of a wage subsidy is that it encourages workers to spend more time at work. The disadvantage of this instrument is that individuals whose wages are subsidized will face smaller incentive to acquire skill and to work intensively while on the job.

These two offsetting considerations produce a redistributive "package" within an optimal taxation framework. In an optimal system a wage subsidy will be combined with a universal income transfer (demogrant) or with a variant of the earned income tax credit (EITC). The relative importance of the two instruments will depend on the relative magnitudes of the work leisure trade off and the effect of wage subsidies on skill

¹ An alternative way to introduce the second incentive margin is to allow for work exertion rather than skill acquisition.
acquisition. Other things equal, the lower (higher) the elasticity of substitution between the wage rate and leisure the larger will be the role of income transfers (wage subsidies).

Another important conclusion, consistent with the incremental analysis of Ballard and Triest, is that high public revenue requirements for non-redistributive purposes limit the optimal amount of redistribution and increase the relative role of wage subsidies in an income maintenance system. The second result, however, needs to be carefully interpreted as a wage subsidy is most effective as a redistributive instrument when it is complemented by a negative demogrant (a head tax). Thus, a linear income tax with an intercept term and wage subsidies complement each other even when initial tax rates are high and the wage subsidy is the sole redistributive instrument.

1.2. An extended Mirrlees model of income inequality.

1.2.1. Wage subsidy in the original Mirrlees model.

The early contributions on wage subsidies are limited in a number of ways. Although they are a direct extension of Mirrlees model they violate its most characteristic feature - the incompleteness of information. If we assume for the Mirrlees formulation that a tax is a function of wages then we are assuming that wages are observable. In the simplest model used by both Kesselman and Blomquist, this means that abilities are observable and a screening problem simply does not exist. Not surprisingly, the solution they obtained for a wage subsidy is close to the first best. To demonstrate the results of this formulation, I use a model similar to the one used by Kesselman and Blomquist except that the graded earnings tax is replaced by the linear wage subsidy financed by a linear income tax with a demogrant.
Individuals maximize a CES utility function:

$$\max \ U_i = (\alpha \cdot C_i^{(q-1)/q} + (1-\alpha) \cdot L_i^{(q-1)/q})^{q/(q-1)}, \alpha, q \in [0,1] \quad (1.2.1.1)$$

$$C_i, L_i$$

Where $L_i$ is $i$'s leisure $L_i = (1-l_i)$ and $C_i$ is consumption of $i$'s individual.

Individuals are subject to a linear income tax with a tax rate $t$ and a guaranteed income (demogrant) $a$:

$$C_i = Y_i \cdot (1-t) + a \quad (1.2.1.2)$$

Individuals receive a linear wage subsidy if their observable wage $w_{obsi}$ is less than some “target” wage $w_{argi}$, the subsidy is equal to some fraction of the difference between the target and observable wages, so the “real” (after subsidy) wage is

$$w_{ri} = \begin{cases} w_{obsi}, & w_{obsi} > w_{argi} \\ w_{obsi} + \delta \cdot (w_{argi} - w_{obsi}), & w_{obsi} \leq w_{argi} \end{cases} \quad (1.2.1.3)$$

Individuals’ observable wages are equal to their abilities

$$w_{obsi} = w_i^* \quad (1.2.1.4)$$

Finally, income

$$Y_i = w_{ri} \cdot l_i \quad (1.2.1.5)$$

The government maximizes the social welfare function:

$$Wf = \sum_{V} \frac{1}{V} [U_i(C_i, L_i)]^r \quad (1.2.1.6)$$

In addition there may be a revenue requirement: $\sum (T_i + WS_i) = R$, where $T_i$ is an individual’s tax returns (including the demogrant), and $WS_i$ is expenditure on the subsidized wage.
A numerical simulation is conducted. The population consisted of 1000 individuals with lognormally distributed abilities $w^* \sim \log N(-1.0,0.39)$\(^2\) This distribution is shown in Figure 1.1.

![Figure 1.1. Abilities.](image)

The following results are obtained:

- an optimal wage subsidy is characterized by a high target wage higher than the largest wage in the economy combined with the subsidy rate of $\delta = 100\%$.
- there is a complete equalization in the model. All the individuals have identical wages, labor supplies and utilities.
- social welfare increases significantly relatively to a linear income tax as shown in Table 1.1\(^3\).

---

\(^2\) Such a distribution of abilities is widely used in the literature. See for example Stern (1976) or Slemrod et. al. (1994).

\(^3\) 18 simulations were conducted for three values of the elasticity $q$ of the utility function, $q = \{0.2, 0.4, 0.99\}$, two values of parameter $\nu$ of social welfare function (6) $\nu = \{1,-2\}$ and three values of the revenue requirement $R = \{0, 0.05, 0.1\}$. Coefficient $\alpha$ of the utility function is chosen so that in the absence of taxation the individual with mean skills supplies labor $\ell = 2/3$. 
- The welfare gains resulting from the addition of a wage subsidy increase with an increase in society's preference for egalitarianism and with an increase in general revenue requirements. The addition of a wage subsidy allows the demogrant to become a non-distortionary lump-sum tax as the wage subsidy redistributes income towards low earning individuals.

Table 1.1. Wage subsidy in the original Mirrlees model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Linear Income Tax</th>
<th>Linear Income Tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\nu$</td>
<td>$R$</td>
<td>$q$</td>
<td>Demo grant</td>
</tr>
<tr>
<td>1.0</td>
<td>0</td>
<td>0.2</td>
<td>0.0897</td>
</tr>
<tr>
<td>1.0</td>
<td>0</td>
<td>0.4</td>
<td>0.0561</td>
</tr>
<tr>
<td>1.0</td>
<td>0</td>
<td>0.99</td>
<td>0.0310</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.2</td>
<td>0.0575</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.4</td>
<td>0.0178</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.99</td>
<td>-0.0120</td>
</tr>
<tr>
<td>1.0</td>
<td>0.10</td>
<td>0.2</td>
<td>0.0289</td>
</tr>
<tr>
<td>1.0</td>
<td>0.10</td>
<td>0.4</td>
<td>-0.0175</td>
</tr>
<tr>
<td>1.0</td>
<td>0.10</td>
<td>0.99</td>
<td>-0.0544</td>
</tr>
<tr>
<td>-2.0</td>
<td>0</td>
<td>0.2</td>
<td>0.1590</td>
</tr>
<tr>
<td>-2.0</td>
<td>0</td>
<td>0.4</td>
<td>0.1230</td>
</tr>
<tr>
<td>-2.0</td>
<td>0</td>
<td>0.99</td>
<td>0.0796</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.2</td>
<td>0.1340</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.4</td>
<td>0.0967</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.99</td>
<td>0.0482</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.10</td>
<td>0.2</td>
<td>0.1097</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.10</td>
<td>0.4</td>
<td>0.0731</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.10</td>
<td>0.99</td>
<td>0.0212</td>
</tr>
</tbody>
</table>

*If there is a revenue requirement then "no-tax" state does not exist and corresponding welfare comparisons cannot be performed.*
Figures 1.2-1.4 depict utility levels and labor supply for situations with no taxes imposed, a linear income tax alone and for a linear wage subsidy and before and after subsidy wages.5

Fig. 1.2. Utility comparison for no-tax, linear income tax alone, and linear wage subsidy structures.

Fig. 1.3. No-tax, linear income tax alone and linear income tax with wage subsidy labor supply decisions.

---

5 for the case of r=1, q = 0.2, α=0.44, R=0., similar results were obtained for q = 0.4 and q=0.99 and revenue requirement R=\{0.05, 0.1\}. 
The main result of this experiment is that individual utility is equalized and social welfare increases by a large amount as the result of the introduction of a wage subsidy. The basic problem with this formulation is that there is no incentive for individuals to change their jobs. Suppose there are two types of individuals with differential abilities and observed wages of $5/hour and $40/hour. Let the government impose an extreme policy of a target wage of $50/hour and a 100% wage subsidy. All individuals now receive the same subsidized wage of $50/hour independently of their observable wage. Common sense suggests that higher paying jobs require more “effort” and/or more “education/training”. The higher effort or training results in disutility and/or financial costs. Since there will be no reward for choosing higher-paying jobs which require more effort or involve a “less pleasant environment” all individuals will tend to work at the

---

6 One can argue that no government will ever impose 100% wage subsidy with such a high target wage rate. The example however, is not intended to be realistic but rather to point out to the obvious problem. This problem of non-existing incentive structure will manifest itself at any target wages and subsidy rates.

7 We put aside “creative” jobs such as high managerial positions and so on. What we are considering is a $8/h job in a grocery store vs. $25/h job at the car assembly line.
lowest wage rate observed in the economy, a suboptimal outcome. This will not occur in a Mirrlees-Kesselman-Blomquist formulation in which individuals always work at their maximum ability level as long as the observable wage is positively correlated with the after-subsidy wage\(^8\). On the basis of these results we conclude that an analysis of an optimal wage subsidy should have two features:

- an individual's wage rate and individual's ability should not be related in such way that knowing the former will uniquely identify the latter. In reality, the observed wage rate does not perfectly reveal ability. If it did there would be no screening problem and a first best outcome could be achieved.
- the model should have the feature that higher-paying jobs require more effort and/or more training.

1.2.2. The basic model.

The economy consists of 1000 individuals. Each person is characterized by a natural ability \(w^*\).

Individuals maximize a CES utility function:

\[
\max \ U_i = (\alpha \cdot C_i^{(q-1)/q} + (1 - \alpha) \cdot L_i^{(q-1)/q})^{q/(q-1)}, i = 1..1000, \alpha, q \in [0,1] \ (1.2.2.1)
\]

\[E_i, l_i\]

Where \(L_i\) is \(i\)'s leisure, \(l_i\) is \(i\)'s labor time, \(C_i\) is \(i\)'s consumption and \(E_i\) is \(i\)'s level of education.

\(^8\) This is true as long as the that standard incentive compatibility constraint of the Mirrlees model is satisfied, i.e., higher income individuals are not worse off then those with lower incomes.
The CES utility function is chosen as it allows for a variety of behavioral responses. The variable elasticity of substitution can generate a backward bending labor supply curve while the Cobb-Douglas utility function cannot.

Unlike the Mirrlees model where an individual’s observable wage, \( w_{obs} \), depends only on an individual’s ability, \( w_{obs} \), is a function of individuals’ natural ability and their chosen level of education:

\[
  w_{obs} = w^0_i + (k_e \cdot (w^*_i)^{(y-1)/r} + (1 - k_e)(E_i)^{(y-1)/r})^{\beta \gamma (y-1)} \\
  i = 1..1000, \ \gamma, k_e \in [0,1], \ \beta \geq 1.
\]

Where \( E_i \) denotes the level of education chosen by an individual. Thus, the observable wage of individuals, or observed productivity, is “produced” with natural ability \( w^*_i \) and acquired education \( E_i \) as inputs.

The variable \( w^0_i \) denotes the wage an individual can earn with a zero level of education. In reality, because of compulsory education and child labor laws elementary education is determined by government regulation. Education becomes elective at the high school level. It is assumed that everybody receives the same amount\(^9\) of elementary education. Thus \( w^0_i \) depends on an individual’s ability and the amount of compulsory education.

The CES function is chosen to characterize the production of skill for the following reasons:

\( ^9 \) i.e. number of years.
(i) It is unlikely that education and ability are close substitutes. Even very able individuals are not considered for some jobs if they do not have a college degree and employees have to acquire certain vocational skills or specialized university level training to be eligible for many jobs regardless of their natural ability. Moreover, low-ability individuals will not profit from the large investments in education if the required skill level is high and their natural ability is low. Unlike a Cobb-Douglas function or a high degree polynomial approximation\(^{10}\) a CES function with its variable ability-education complementarity allows these considerations to be modeled.

(ii) Individuals with higher ability levels have an enhanced capacity to absorb (benefit from) education\(^{11}\). Therefore, the effect of education is larger for those with greater ability. The CES production function (1.2.2.2) has a positive second derivative of the observable wage with respect to education and natural ability \(\partial^2(w_{obs})/\partial(E_i)\partial(w^*)\) for any \(E\) and \(w^*\) so this property is always satisfied.

(iii) It seems reasonable to assume that a proportional increase of a person's natural ability and the time they devote to education will increase their observable wages by more than a proportional amount\(^{12}\). Thus, unlike most of its other applications, the CES function used in the production of skill is postulated to exhibit increasing returns to scale. This feature is specified by assuming \(\beta \geq 1\) (\(\beta = 1\) corresponds to the case of constant returns to scale).

---

\(^{10}\) Fair (1971).

\(^{11}\) Cronbach (1957, 1975).

\(^{12}\) i.e. if we double the ability and education simultaneously the acquired ability (or observable wage) will increase by more than two.
Education has a unit cost $p_e$. The government can choose between letting educational expenses be tax deductible or not.

The time spent on education cannot be used for either work or leisure\(^{13}\). The time spent on leisure can be expressed as

$$L_i = (1 - l_i - E_i)$$ \hspace{1cm} (1.2.2.3)

where $l_i$ is an individual's labor time.

Phelps (1997) argues that it is not optimal to make work-contingent income transfers to very low-income, low-ability individuals. Such a system increases labor supply and decreases the consumption of leisure without significantly increasing output. To test this proposition, I determine the optimal value of the minimum observed wage that individuals have to earn before they become eligible for the wage subsidy. Thus, an individual with an observable wage less than some target wage $w_{\text{arg}}$ and more than some minimum wage $w_{\text{min}}$ (could be zero) receives a linear wage subsidy, which is equal to a fraction $\delta$ ($\delta \in [0,1]$) of the difference between the observable wage and the target wage and is paid in addition to the observable wage. Call $\delta$ - the rate of subsidy. As the observable wage approaches the target level the wage subsidy decreases to zero.

$$w_{Ri} = \begin{cases} w_{\text{obs}}, & w_{\text{obs}} > w_{\text{arg}}, w_{\text{obs}} \leq w_{\text{min}} \\ w_{\text{obs}} + \delta \cdot (w_{\text{arg}} - w_{\text{obs}}), & w_{\text{min}} < w_{\text{obs}} \leq w_{\text{arg}} \end{cases}$$ \hspace{1cm} (1.2.2.4)

A wage subsidy, which can be supplemented by a cash transfer demographant, is financed by a linear income tax.\(^{14}\) A tax on wage rates, the alternative to the income tax, has several shortcomings. Such a tax would have an especially adverse effect on job

\(^{13}\) Learning-by-doing is not addressed in this paper.

\(^{14}\) As it turned out, addition of the two-bracket tax does not change the results significantly. A simple linear income tax is much easier to compute so the model can be analyzed for bigger set of parameters.
choice and investment in education. In addition, a tax on wage rates is informationally inefficient. In any problem with incomplete information, all the available informational signals available, such as the observed wage rate, the supply of labor or the income level should be utilized. Any policy which disregards a part of the available information by utilizing a subset of observable variables, such as income tax with a demogrant alone or the wage tax with wage subsidy alone, will be suboptimal.

Individuals' gross income is denoted as $Y_i = w_{ni} \cdot l_i$, the tax rate as $t$, demogrant as $a$, consumption $C_i$ is

$$C_i = Y_i \cdot (1-t) + a - p_e \cdot E_i$$ \hspace{1cm} (1.2.2.5)

if educational expenditures are not tax deductible and is equal to

$$C_i = (Y_i - p_e \cdot E_i) \cdot (1-t) + a$$ \hspace{1cm} (1.2.2.5')

if the cost of education is tax-deductible.

For an infinitely high price of education and with $w_i^0 = w^*$ individuals' choice of education is zero and the wage is determined solely by ability as in the original Mirrlees model.

Individuals choose their education level $E_i$ and their labor supply $l_i$.

The government chooses an optimal tax policy which consists of

- an income tax rate $T = \{t\}$
- a demogrant $a$ which could be zero or negative (i.e., a lump-sum tax)
- a wage subsidy scheme $WS = \{w_{min}, w_{arg}, \delta\}$

In addition, the government decides whether to allow education to be tax deductible.
The government's objective is to maximize the social welfare function

$$\max \ Wf = \sum_{i}^{1000} \frac{1}{\nu} [U_i(C_i, L_i)]^\nu$$  \hspace{1cm} (1.2.2.6)

Subject to the budget constraint

$$\sum_{i}^{1000} \ (T_i + WS_i) = R$$  \hspace{1cm} (1.2.2.7)

Where \( T_i \) is an individual's tax liability (including the demogrant), \( WS_i \) is expenditure on the subsidized wage, and \( R \) is the revenue requirement. The government can observe the individuals' wage rate, the individuals' income and can infer individuals' labor supply.

1.2.3. Completed simulations and results obtained.

I adopt the numerical simulation approach employed by Stern (1976) in his analysis of the optimal linear tax and by Slemrod et al. (1994) in their analysis of two-bracket income tax. A lognormal distribution \((\mu = -1., \ and \ \sigma = 0.39)\) of abilities consisting of 1000 individuals is generated. Each simulation is conducted for three values of the elasticity \( q \) of the utility function (1.2.2.1), \( q = \{0.2, 0.4, 0.99\} \), two values of parameter \( \nu \) of social welfare function (1.2.2.6) \( \nu = \{1, -2\} \)\(^{15}\), and three values of the revenue requirement \( R = \{0, 0.05, 0.1\} \). In addition, I follow Stern by choosing the coefficient \( \alpha \) of the utility function so that, in the absence of taxation, when the education level is zero and wage depends only on natural ability, the individual with

---

\(^{15}\)These values are often used in optimal taxation literature, see for example Slemrod et al. (1994): \( \nu = 1 \) is simple additive welfare function. \( \nu = -2 \) puts more weight on low income individuals.
mean skills supplies labor \( l = 2/3 \). It follows that \( \alpha = 0.44 \) for \( q = 0.2, \alpha = 0.59 \) for \( q = 0.4, \alpha = 0.67 \) for \( q = 0.99 \).

The calibration of the earning function (1.2.2.2.) is a more complicated problem. In an optimal taxation framework, Fair (1971) is the only researcher to make earnings a function of an individual's ability and of education. He used statistical data on the earnings of high school graduates, college graduates and advanced degree holders to approximate an earning function as a high degree polynomial. This approach can be used only as an approximation as there is evidence of a high correlation between the level of education and an individual's ability.

In all of the simulations the price of education, \( p_e \), is set equal to half the yearly income of an individual with a mean skill level. The elasticity of substitution between education and ability \( \gamma \), coefficient \( k_e \) and the returns to scale parameter \( \beta \) in (1.2.2.2) are set so that the implied levels of education and the income and wage distributions approximate those which are actually observed.

The model is static and does not deal with time explicitly. However, to approximate reality and to calibrate the model it is assumed that individuals have 50 years at their disposal after achieving legal working age. This endowment is represented by one in equation (1.2.2.3).

The main results of the simulations are presented in Tables 1.2 a, b for the case of non-deductible education expenses and in Tables 1.3 a, b for the case of tax-deductible education expenses.

Tables 1.2a and 1.3a report the results for an optimal linear income tax alone and for an optimal linear income tax combined with a wage subsidy. The results allow a no-
tax situation to be compared with each of form of redistribution. In these calculations the linear income tax is introduced first and the wage subsidy is then added. In contrast, Tables 1.2b and 1.3b compare the no tax situation financed with a proportional income tax without a demogramt and with a wage subsidy financed with a proportional income tax without a demogramt. The two base cases are (1) the no tax situation where there is a zero revenue requirement and (2) the situation where a given revenue requirement is financed with a proportional income tax without a demogramt.

Tables 1.2-1.3a,b present the percentage increases in welfare relative to a linear income tax. The welfare improvement of adding a wage subsidy to a linear income tax is also calculated. In Tables 1.2-1.3b the base case level of welfare is compared with a wage subsidy financed by a proportional income tax without a demogramt. The welfare level for this case is compared with a wage subsidy financed with a linear income tax which allows for a demogramt.

The first three columns of Tables 1.2-1.3 report the distributional parameter \( \nu \), the revenue requirement \( R \), and the elasticity of substitution between consumption and leisure \( q \). The next column records the level of welfare without a redistribution system. As a tax is required when a general revenue requirement is imposed, this column (as well as corresponding welfare comparison) is designated as N/A. The next three columns of the table present the optimal demogramt, the tax rate and the value of the maximized welfare function for an optimal linear tax. The next five columns present the results for a wage subsidy combined with linear income tax. I report the optimal demogramt, the tax rate, the target wage of the wage subsidy \( w_{\text{target}} \), and the wage subsidy rate \( \delta \). The final two columns present two welfare comparisons: the increase in welfare associated with
the introduction of a linear income tax and the increase in welfare resulting from adding a wage subsidy to an optimal linear income tax. The results for the simulation presented in the first row of the Table 1.2a are used as a reference point. For this case the elasticity of substitution between consumption and leisure is \( q = 0.2 \) \((\alpha = 0.44)\), \( v = 1 \) and the revenue requirement is \( R = 0 \). The price of education, \( p_e \), is set equal to 0.08 so that a year of education costs approximately 50% of one year of income for the mean ability individual.\(^{16}\) For an elasticity of substitution between education and ability \( \gamma = 0.4 \), \( k_e = 0.85 \), and \( \beta = 1.5 \) the least able individual chooses to acquire 3 years of education, beyond primary school, and the most able individual acquires 14 years of education, an advanced degree. The person with the mean ability obtains 8 years of education. Wages vary by a factor of 14 (e.g. \$5-\$70/h) and incomes by the factor of 6 (e.g. \$12-\$72K a year). The wage an individual earns with a zero level of education is calculated by substituting the individual's ability level and 4 years of elementary education into the CES “wage-production” function.

The results for the case of a zero revenue requirement are clear cut. The linear tax alone and the wage subsidy alone both improve welfare relative to the base case. Also the addition of the second instrument improves welfare somewhat, indicating that both redistributive mechanisms should be used. When a wage subsidy is added to a linear income tax the size of the demigrant can be reduced and work effort is encouraged. However, when persons receive a wage subsidy, they invest less in education and/or exert less effort while on their jobs. Consequently, the demigrant is partially retained and the

\(^{16}\) This value is on the high side as it includes the price of high school, public education at the secondary and college level is heavily subsidized. Ideally, I should test the sensitivity of the rest to varying values of this parameter.
size of the wage subsidy is limited. The wage subsidy partially substitutes for the
demogrant as the optimal demogrant is always smaller when combined with a wage
subsidy. Similarly, an optimal wage subsidy is always smaller when combined with a
positive demogrant.

For one case an optimal linear tax is calculated to be \( t = 27.6\% \), and the
demogrant, \( a = 0.045 \). Welfare is increased by approximately 0.34\%. In contrast, the
Mirrlees case for the same utility function without investment in education yields an
optimal tax rate of \( t = 35.6\% \) and a welfare increase of 0.5\%.\(^{17}\) The welfare increase
associated with the wage subsidy is an additional 0.04\%. The welfare increase due to
linear income tax with a demogrant and a wage subsidy is 12\% higher than the welfare
increase resulting from an income tax with the demogrant. With a wage subsidy and a
demogrant the tax rate is lower and the demogrant is less than the in no-subsidy case,
\( t = 25.7\% \), \( a = 0.034 \).

An optimal wage subsidy is approximately a fixed per-hour-worked grant. The
target wage is \( w_{\text{arg}} = 1.1 \) (twice the highest observable wage in the economy) and the
subsidy rate is only \( \delta = 0.02 \) (or 2\%). The optimal wage subsidy is not restricted to the
low-income, low-ability population, but is approximately constant in amount across the
population. The explanation for this unexpected result is that the important effect of the
wage subsidy is to encourage work effort. Redistributive taxes discourage work effort
and the subsidy serves to partially offset this disincentive effect and encourages persons
to utilize their accumulated human capital more fully.

\(^{17}\) All the results of numerical simulation of Mirrlees model in Stern (1976) and Slemrod (1994) are
verified and confirmed by the current model with the price of education set equal to infinity and no-
education wage set equal to individual’s ability.
The optimal minimum wage for wage subsidy eligibility is estimated to be zero. It is suboptimal to exclude very low wage individuals from the wage subsidy in a mixed demogrant–subsidy system. If low-income individuals are excluded from the wage subsidy their income will fall and the demogrant must be increased to preserve the same level of equality.

Figures 1.5-1.8 show the effects of the optimal wage subsidy relative to both the no-tax and the linear-tax cases. The graph for utility levels, Figure 1.5, indicates that the wage subsidy has greater redistributing power than the income tax alone. As shown in Figure 1.6 labor supply for a wage subsidy system with a demogrant is higher for all ability levels than labor supply for an income tax alone. Figure 1.7 for educational choices demonstrates that a wage subsidy combined with a demogrant discourages education more than the linear income tax alone. It is noteworthy that the small welfare improving effect of the wage subsidy is accompanied by a relatively large decline in education. Moreover, observable wages, shown in Figure 1.8, change in the same direction as education. Labor productivity decreases when a linear income tax is imposed and decreases further if a wage subsidy is added. The graph for the after-subsidy wage illustrates the “flat” shape of the optimal wage subsidy.\(^{18}\)

\(^{18}\) An explanatory note on the optimal wage subsidy is needed. In a number of cases shown in Tables 1.2-1.3 an optimal wage subsidy is represented by multiple solutions with only slight variation in maximized social welfare. For example in just illustrated case of \(v=1.0, R=0, q=0.2\) the optimal wage subsidy could be \(\text{target wage rate} = \{(0.7,0.03); (1.0,0.02); (2.2,0.01)\}\). Recall the expression for real wage \(w_R = w_{obs} + \delta \cdot (w_{arg} - w_{obs}) = w_{obs} (1-\delta) + \delta w_{arg}\). Note now that when optimal \(\delta\) is small (a few per cent in most of our cases) and optimal \(w_{arg}\) is large, \(w_{obs} (1-\delta)\) component of real wage changes little while what is important \(\delta w_{arg}\) product. Thus, an increase in the target wage will produce a policy similar to the existing one if it is accompanied by a proportional decrease in subsidy rate. This is what in generally happens if the set of solutions with similar/identical maximized welfare is obtained. It should be noted however that in all cases the wage subsidy is relatively flat, i.e. includes both “lower” and the “middle” classes of our distribution, though sometimes exclude the upper tail of the abilities/wages. Also, in order to keep computations manageable the target wage is maximized over the range from zero to
If the policy choice is restricted to one instrument the income transfer (linear income tax) is relatively efficient when the labor supply elasticity is low and the wage-subsidy is the preferred instrument for moderate and higher values of labor supply response. Closely related to this finding is the result that the magnitude of welfare improvement resulting from the addition of a wage subsidy to a demogrant increases with an increase in the elasticity of substitution between consumption and leisure. The more elastic the supply of labor the larger is the role of the wage subsidy in an optimal system of income redistribution.

The interpretation of the results is more complicated when there are positive revenue requirements. First, it should be noted that when the demogrant is positive the use of income transfers and wage subsidies both enhance welfare. Also, as the labor supply elasticity increases, the wage subsidy becomes relatively more effective. But my results also bring out the fact that at higher revenue requirements the optimal demogrant may be negative, i.e. the optimal linear income tax consists of a constant lump sum tax and a flat-rate tax. When this occurs there can be no redistribution to low income persons through income transfers. For these cases it appears that a wage subsidy will be the only redistributive instrument. However, the results also bring out the critical importance of tax system that is used to finance a revenue requirement and income transfers.

A wage subsidy has a much larger potential to redistribute income when it is financed by a linear income tax with a lump sum tax component. When a proportional income tax is used to finance general revenue the redistributive potential of a wage

 approximately five maximal observed no-tax wages in the economy. In some cases going out of this domain can give slight increase in welfare. Since the computational burden of extending maximization range for the target wage is significant and welfare increases are small on the margin of simulation precision, such solutions are not considered.
subsidy will be limited, even negligible. One example of this outcome is the case of an additive welfare function, a high revenue requirement and a high labor supply response. This is a situation for which the optimal wage subsidy is zero when financed by a proportional income tax. A switch to a linear income tax with a negative demogrant allows the introduction of a wage subsidy for this case and increases welfare significantly. Although the wage subsidy is the only redistributive instrument, most of the welfare improvement relative to the base case is the result of the adoption of a more efficient form of tax revenue. Wage subsidies and demogrants (income transfers) complement each other as policy instruments for two quite different reasons. When the demogrant is positive (income transfer) the appropriate mix of transfers and wage subsidies strikes a balance between the two incentive margins, the incentive to work and the incentive to acquire skill.

When the optimal demogrant alone is approximately zero the introduction of a redistributive wage subsidy allows the partial adoption of a more efficient, regressive, lump-sum tax component. Consider the interesting case of a society with strongly egalitarian preferences, a higher revenue requirement and a high labor supply response. In this situation the linear income tax and a wage subsidy financed by a proportional tax both have a very limited potential to improve welfare when introduced individually. When combined, however, they produce a substantial increase in welfare relative to the base case as a portion of the required revenues can be financed in part with a lump sum tax.

For the vast majority of the simulations (the exception is the case of nondeductible education expenses with a low elasticity of labor supply \( q = 0.2 \)) an
increase in the revenue requirement leads to increase in the welfare effect of the wage subsidy. Thus, when the government has to collect more revenue for general expenditures it should rely more heavily on the wage subsidy. When the work-leisure choice is already distorted by the tax imposed for revenue requirements, the wage subsidy becomes more effective relative to a demogrant.

As expected, when private educational expenses are tax deductible, the welfare effects of the wage subsidy and linear income tax increase. Also, as expected, allowing the cost of education to be deducted from taxable income eliminates the distortionary effect of taxation on human capital expenditures and increases wages and redistribution. Welfare is significantly increased relative to the case where education expenditures are non-deductible.
### Table 1.2a. Comparison of linear income tax and linear income tax with wage subsidy structures. Non-deductible private education expenses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>v</td>
<td>R</td>
<td>q</td>
<td>Welfare</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.2</td>
<td>172.292</td>
<td>0.0449</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.4</td>
<td>169.992</td>
<td>0.0294</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.99</td>
<td>171.140</td>
<td>0.0161</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
<td>0.0090</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
<td>-0.0091</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
<td>-0.0273</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
<td>-0.0240</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
<td>-0.0293</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
<td>-0.0268</td>
</tr>
</tbody>
</table>
Table 1.2a. Continued.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Welfare</td>
<td>Demo Grant</td>
<td>Tax rate</td>
<td>Welfare</td>
</tr>
<tr>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.0</td>
<td>0.00</td>
<td>0.2</td>
<td>-20816.443</td>
<td>0.0885</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.00</td>
<td>0.4</td>
<td>-21143.274</td>
<td>0.0700</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.00</td>
<td>0.99</td>
<td>-20071.207</td>
<td>0.0406</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
<td>0.0588</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
<td>0.0430</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
<td>0.0122</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
<td>0.0305</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
<td>0.0194</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
<td>-0.0056</td>
</tr>
</tbody>
</table>
Table 1.3a. Comparison of linear income tax and linear income tax combined with wage subsidy structures. Tax deductible private education expenses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Welfare</td>
<td>Demo Grant</td>
<td>Tax rate</td>
<td>Welfare</td>
</tr>
<tr>
<td>( v )</td>
<td>1.0</td>
<td>0.</td>
<td>0.2</td>
<td>172.292</td>
</tr>
<tr>
<td>( R )</td>
<td>1.0</td>
<td>0.</td>
<td>0.4</td>
<td>169.992</td>
</tr>
<tr>
<td>( q )</td>
<td>1.0</td>
<td>0.</td>
<td>0.99</td>
<td>171.140</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
</tr>
<tr>
<td>Parameters</td>
<td>No Tax</td>
<td>Linear Income Tax</td>
<td>Linear Income tax with Wage Subsidy</td>
<td>Comparison of the tax structures</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>-------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td>Welfare</td>
<td>Demo Grant</td>
<td>Tax rate</td>
<td>Welfare</td>
</tr>
<tr>
<td>$v$</td>
<td>$R$</td>
<td>$Q$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.2</td>
<td>-20816.443</td>
<td>0.0944</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.4</td>
<td>-21143.274</td>
<td>0.0709</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.99</td>
<td>-20071.207</td>
<td>0.0385</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
<td>0.0681</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
<td>0.0452</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
<td>0.0108</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
<td>0.0440</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
<td>0.0246</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
<td>-0.0066</td>
</tr>
</tbody>
</table>
Table 1.2b. Linear income tax and linear income tax with wage subsidy structures, no demogrant. Non-deductible private education expenses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Base Case</th>
<th>Linear Income Tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Tax</td>
<td>Linear Income Tax</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Welfare</td>
<td>Tax rate</td>
<td>Welfare</td>
</tr>
<tr>
<td>$v$</td>
<td>$R$</td>
<td>$Q$</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.2</td>
<td>172.292</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.4</td>
<td>169.992</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.99</td>
<td>171.140</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Table 1.2b. Continued.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Base Case</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Tax</td>
<td>Welfare</td>
<td>Tax rate</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Table 1.3b. Linear income tax and linear income tax with wage subsidy structures, no demogrant. Tax deductible private education expenses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Base Case</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Tax</td>
<td>Linear Income Tax</td>
<td></td>
</tr>
<tr>
<td>$\nu$</td>
<td>$R$</td>
<td>$q$</td>
<td>Welfare</td>
</tr>
<tr>
<td>1.0</td>
<td>0.</td>
<td>0.2</td>
<td>172.292</td>
</tr>
<tr>
<td>1.0</td>
<td>0.</td>
<td>0.4</td>
<td>169.992</td>
</tr>
<tr>
<td>1.0</td>
<td>0.</td>
<td>0.99</td>
<td>171.140</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
</tr>
</tbody>
</table>
### Table 1.3b. Continued.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Base Case</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Tax</td>
<td>Linear Income Tax</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Welfare</td>
<td>Tax rate</td>
<td>Welfare</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$\gamma$</td>
<td>$\delta$</td>
<td></td>
</tr>
<tr>
<td>-2.0</td>
<td>0</td>
<td>0.2</td>
<td>-20816.443</td>
</tr>
<tr>
<td>-2.0</td>
<td>0</td>
<td>0.4</td>
<td>-21143.274</td>
</tr>
<tr>
<td>-2.0</td>
<td>0</td>
<td>0.99</td>
<td>-20071.207</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.05</td>
<td>0.99</td>
<td>N/A</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.2</td>
<td>N/A</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.4</td>
<td>N/A</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.1</td>
<td>0.99</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Fig. 1.5. Utility comparison for no-tax, linear income tax alone and linear income tax with wage subsidy structures.

Fig. 1.6. No-tax, linear income tax alone and linear income tax with wage subsidy labor supply decisions.
Fig. 1.7. No-tax, linear income tax alone and linear income tax with wage subsidy education choices.

Fig. 1.8. No tax and linear income tax alone observable wages. Linear income tax with wage subsidy observable and real (after subsidy) wages.
1.3. Extensions of and other experiments with the basic model.

1.3.1. Different objective functions.

This paper and much of the work on income maintenance based on an optimal income tax framework adopts the welfarist approach which takes into account the value of net income and leisure in individual utility functions. This approach also allows for decreasing marginal utility of net income for the individual and calculates the optimal amount of redistribution by assigning welfare weights to the utilities of persons with different levels of skill.

The welfarist objective function have been challenged by Kanbur, Keen, and Tuomola (1994) and by Besley and Coate (1995). They argue that contemporary policy debate give little if any weight to the value of leisure of the poor. Taxpayers and politicians are concerned with maintaining the net income of the poor above some minimum standard and/or attempting to minimize that number of persons who are living in poverty. These writers argue that while utility maintenance maybe more consonant with welfare economics, they doubt the value of this objective as a positive model that could help shape the formulation of current policy. I follow these researchers and analyze the relative merits of a wage subsidy and income transfers for a non-welfarist objective function. Besley and Coate demonstrate the value of adopting a general work requirement for some members of the poverty population. But they do not consider hourly wage subsidies nor do they make skill endogenous.

The income maintenance objective sets a target level of income that must be achieved by all of the poor and chooses between policies that minimizes the budgeting
cost of meeting this objectives. A closely related approach, a variant of the one adopted by Kanbur et.al. (1994), is the minimization of a poverty index of the form

\[ P_\alpha = \sum n_i \frac{(x_i - x^*)^\alpha}{x^*}, \quad \alpha > 1 \]  

(1.3.1.1)

where \( x_i \) is the income of group \( i \) whose income is below the poverty level \( x^* \) and \( n_i \) is the number in group \( i \). The parameter \( \alpha \) measures alternative degrees of aversion to inequality among the poor. The objective of policy is to minimize this poverty index subject to a budget constraint; the government must raise a fixed amount of revenue to finance other non-redistributive expenditures. For both non-welfarist objective functions income transfers and wage subsidies will be concentrated on the poverty population. For a linear income tax the demografant will decrease with income and its effects can readily be compared with a wage subsidy that will also decrease with the wage rate.

Common sense suggests that when the leisure of the poor does not count in the objective function greater reliance will be placed on a wage subsidy that encourages work effort. However, this remains to be demonstrated in an approach which allows for endogenous skill acquisition and assumes a strong complementarity between ability and education, a proxy for investment in skill. If a wage subsidy is targeted toward the more able members of the poverty population it will encourage them to work harder but discourage from acquiring additional skills.

The model developed in Section 1.2 is modified as follows.

Individuals chose their education level and labor supply to maximize their utilities.\(^{19}\)

\(^{19}\) The variables are defined and the notation is specified in Section 1.2.
\[
\text{max } U_i = (\alpha \cdot C_i^{(q-1)/q} + (1 - \alpha) \cdot L_i^{(q-1)/q})^{q/(q-1)}, i = 1..1000
\]  
\[E_i, l_i\]

subject to

\[Y_i = w_{\text{Ri}} \cdot l_i\]

The observed wage is a function of individuals' natural abilities and their education levels:

\[w_{\text{obsi}} = w_i^0 + (k_x \cdot (w_i^*)^{(\gamma-1)/\gamma} + (1 - k_x)(E_i)^{\gamma} \cdot N^{\gamma-1})\]

The government imposes a wage subsidy in addition to a universal demogrant.

\[w_{\text{Ri}} = \begin{cases} w_{\text{obsi}}, w_{\text{obsi}} > w_{\text{arg}} \\ w_{\text{obsi}} + \delta \cdot (w_{\text{arg}} - w_{\text{obsi}}), w_{\text{obsi}} \leq w_{\text{arg}} \end{cases}\]

Consumption is:

\[C_i = Y_i + a - p_e \cdot E_i\]

A tax is not imposed on the individuals who receive the transfers. Their consumption is the sum of their wage incomes plus transfers received minus their educational expenses.\(^{20}\) To justify this framework we can think of a society consisting of a group of the working poor and a group of more affluent taxpayers committed to income maintenance. The latter group, represented by the government, seeks to minimize the cost of meeting the objective of achieving a minimum income level for the poor.

The government objective is:

\[\min \sum_{i=1}^{1000} (a + WS_i)\]

subject to the constraint: \[Y_i \geq Y_{\text{min}}, i = 1..1000.\]

\(^{20}\) An apparent problem with this formulation is that it leaves out who will pay for the tax.
Where $a$ is a universal demogrant, $WS_i$ is amount of the wage subsidy (if any) received by $i$'s individual and $Y_{maa}$ is the minimum income which must be maintained. The government can observe wage rates, individuals' income and to infer an individual's labor supply.

Two types of experiments were carried out. The first experiment is a comparison of a universal demogrant with an optimal wage subsidy for the two welfare criteria under consideration.

The minimum incomes are set so that in the state with no transfers approximately 10\% of the population earns less than the poverty level of income\(^{21}\). Alternatively, it is assumed that the budget constraint is set at 5\% of GDP. The results are presented in Tables 1.4 and 1.5.

**Table 1.4.** Comparison of universal demogrant and universal demogrant with wage subsidy income maintenance structures. Minimum income requirement.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Universal Demogrant</th>
<th>Universal demogrant with Wage Subsidy</th>
<th>Univ. Demogr. to Univ. Demogr. with Wage Subsidy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q$</td>
<td>Demogrant</td>
<td>Total Transfer</td>
<td>Demogrant</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0682</td>
<td>68.272</td>
<td>0.0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0694</td>
<td>69.398</td>
<td>0.0</td>
</tr>
<tr>
<td>0.99</td>
<td>0.0930</td>
<td>93.000</td>
<td>0.0</td>
</tr>
</tbody>
</table>

\(^{21}\) Minimum incomes are 0.1115, 0.1076 and 0.093 for $q = \{0.2, 0.4, 0.99\}$ correspondingly.
Table 1.5. Comparison of universal demigrant and universal demigrant with wage subsidy income maintenance structures. Poverty index minimization.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No-Transf. Povrt. Index</th>
<th>Universal Demigrant</th>
<th>Universal demigrant with Wage Subsidy</th>
<th>No-Transfer To Univ. Demogr. Poverty Index Decrease %</th>
<th>Univ. Demogr. to Univ. Demogr. &amp; Wage subsidy Poverty Index Decrease %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.2</td>
<td>14.708</td>
<td>0.0008</td>
<td>14.127</td>
<td>0.000047</td>
</tr>
<tr>
<td>1.</td>
<td>0.4</td>
<td>14.917</td>
<td>0.0008</td>
<td>14.340</td>
<td>0.000059</td>
</tr>
<tr>
<td>1.</td>
<td>0.99</td>
<td>16.770</td>
<td>0.0008</td>
<td>16.225</td>
<td>0.000022</td>
</tr>
<tr>
<td>2.</td>
<td>0.2</td>
<td>0.400</td>
<td>0.0008</td>
<td>0.380</td>
<td>0.000047</td>
</tr>
<tr>
<td>2.</td>
<td>0.4</td>
<td>0.394</td>
<td>0.0008</td>
<td>0.376</td>
<td>0.000059</td>
</tr>
<tr>
<td>2.</td>
<td>0.99</td>
<td>0.420</td>
<td>0.0008</td>
<td>0.403</td>
<td>0.000022</td>
</tr>
</tbody>
</table>

Not surprisingly, the universal demigrant is very ineffective relative to a more targeted wage subsidy. The revenue requirements to meet the income objective are between 15 to 35% of what they are under a universal demigrant. For the poverty index objective function, and a fixed budget constraint, the percentage decrease in the poverty index is 4-5 times larger for a wage subsidy relative to a demigrant.

An obvious limitation of the universal demigrant is that it is paid to every member of the society. The revenue is wasted being paid to richer individuals whose incomes are above minimum or poverty level simply because the distribution scheme is absolutely inflexible. To investigate this issue the model is modified by replacing the universal demigrant by a more flexible income based transfer – a phased-out demigrant.

Equation 1.3.1.6 is replaced with:

$$C_i = \begin{cases} 
Y_i + a \left(1 - \frac{Y_i}{Y_{brk}}\right) - p_e \cdot E, & \text{if } Y_i \leq Y_{brk} \\
Y_i - p_e \cdot E, & \text{if } Y_i > Y_{brk}
\end{cases}$$

(1.3.1.6')
The demogrant declines with income and it is zero after some optimally determined income level $y_{brk}$.

The results for the minimum income objective function are summarized in Table 1.6 for the minimum income and in Table 1.7 for the poverty index minimization objectives.

**Table 1.6.** Comparison of phased-out demogrant and phased-out demogrant with wage subsidy income maintenance structures. Minimum income requirement.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Phased-out Demogrant</th>
<th>Phased-out Demogrant with Wage Subsidy</th>
<th>Phase-out Demogrant to Phased-out Demogrant &amp; Wage Subsidy total transfer decrease, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q$</td>
<td>Intercept</td>
<td>Income bracket $^{22}$</td>
<td>Total Transf.</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1048</td>
<td>0.134</td>
<td>9.845</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1048</td>
<td>0.141</td>
<td>20.030</td>
</tr>
<tr>
<td>0.99</td>
<td>0.0930</td>
<td>0.236</td>
<td>59.719</td>
</tr>
</tbody>
</table>

**Table 1.7.** Comparison of phased-out demogrant and phased-out demogrant with wage subsidy income maintenance structures. Poverty index minimization.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No-Transfer Poverty Index</th>
<th>Phased-out Demogrant</th>
<th>Phased-out Demogrant with Wage Subsidy</th>
<th>No-Transfer To Phased-out Demogrant Poverty Index Decrease %</th>
<th>Phased-out Demogrant to Phased-out Demogrant with Wage Subsidy Poverty Index Decrease %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$q$</td>
<td>Intercept</td>
<td>Poverty index</td>
<td>Intercept</td>
<td>Income bracket</td>
</tr>
<tr>
<td>1. 0.2</td>
<td>14.708</td>
<td>0.0304</td>
<td>0.121</td>
<td>10.825</td>
<td>0.0304</td>
</tr>
<tr>
<td>1. 0.4</td>
<td>14.917</td>
<td>0.0208</td>
<td>0.125</td>
<td>12.106</td>
<td>0.0050</td>
</tr>
<tr>
<td>1. 0.99</td>
<td>16.770</td>
<td>0.0063</td>
<td>0.150</td>
<td>15.941</td>
<td>0.0002</td>
</tr>
<tr>
<td>2. 0.2</td>
<td>0.400</td>
<td>0.0427</td>
<td>0.112</td>
<td>0.2070</td>
<td>0.0404</td>
</tr>
<tr>
<td>2. 0.4</td>
<td>0.394</td>
<td>0.0476</td>
<td>0.100</td>
<td>0.2918</td>
<td>0.0171</td>
</tr>
<tr>
<td>2. 0.99</td>
<td>0.420</td>
<td>0.0071</td>
<td>0.144</td>
<td>0.3819</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

$^{22}$ Income varies from approximately 0.05 to 0.35 in absolute terms.
Here the results are much less clear cut; they depend critically on the elasticity-of-substitution between income and leisure (ES). The addition of a wage subsidy to a phased-out demografant when the ES is equal to 0.2 decreases revenue requirements by only 5 percent under the income maintenance objective. The optimal wage subsidy for the minimization of poverty objective is zero for an ES of 0.2. In this case, sole reliance should be placed on the phased-out demografant.

Once the value of ES is raised to 0.4 and 1.0, the wage subsidy comes into its own. The wage subsidy plays the major role under an income maintenance objective, the revenue requirement falls by 45 percent when this instrument is added. When the poverty index is minimized the addition of a wage subsidy makes a significant difference and it is the primary redistributive instrument for the higher ES.

These results are interesting as they are qualitatively very similar to those for the welfarist approach in Section 1.2. Moreover, while they show the benefit of adding a wage subsidy they reiterate the importance of the magnitude of the elasticity of labor supply in the determining the relative role for the wage subsidy.

1.3.2. The efficiency of subsidies for education.

In extending the basic model we analyze the effects of introducing a subsidy for educational expenditures, a decrease of tuition to the student. To isolate the contribution of an educational subsidy the analysis is carried out for a model without a wage subsidy. As before individuals chose their level of education and their labor supply to maximize utilities.

\[
\max_{E_i, t_i} U_i = (\alpha \cdot C_i^{(q-1)/q} + (1-\alpha) \cdot L_i^{(q-1)/q})^{\frac{q}{q-1}}, i = 1..1000
\]  

(1.3.2.1.)
subject to: \( L_i = (1 - l_i - E_i) \) \tag{1.3.2.2.}

Since there is no wage subsidy the real wage is equal to the observable wage:

\[
w_{it} = w_{\text{obs}} = w_i^0 + (k_e \cdot (w_i^*)^{(r-1)/r} + (1 - k_e)(E_i)^{(r-1)/r})^{\beta r / (r-1)} \tag{1.3.2.3.}
\]

Consumption for non-deductible education is:

\[
C_i = Y_i \cdot (1 - t) + a - p_e \cdot E_i \tag{1.3.2.4.}
\]

Consumption for tax-deductible education is:

\[
C_i = (Y_i - p_e \cdot E_i) \cdot (1 - t) + a \tag{1.3.2.4'}
\]

where \( p_i \leq p_e \) is private cost of education.

The government observes wages \( w_{\text{obs}} \), incomes \( Y_i \) and labor supply decisions \( l_i \) for every individual \( i = 1..1000 \), and chooses the optimal tax rate \( t \), the demogrant \( a \), and subsidized price of education \( p_e \) to maximize social welfare function for the two cases where the costs of education are tax deductible and when they are not.

\[
\max Wf = \sum_{i=1}^{1000} \frac{1}{\nu} [U_i(C_i, L_i)]^\nu \tag{1.3.2.5.}
\]

subject to \( \sum_{i=1}^{1000} (T_i + ES_i) = R \) \tag{1.3.2.6.}

Where \( T_i \) is total tax linear income tax including demogrant, \( R \) is revenue requirement (if any), and \( ES_i \) is amount of education subsidy for \( i \)'s individual equal to \((p_e - p_i)\) multiplied by the amount of education chosen by the individual.

Numerical simulation were conducted for three values of elasticity \( q \) of utility function \( (5.1.1) \), \( q = \{0.2, 0.4, 0.99\} \), two values of parameter \( \nu \) of social welfare function \( (6) \nu = \{1, -2\} \), and for the case of tax deductible and non-deductible educational
expenses. The revenue requirement is taken to be zero so that taxes are imposed solely for redistributive purposes.

The results of the simulations are presented in Table 1.8. The first four columns of Table 1.8 show the "egalitarian" parameter $\nu$ of the social welfare function, the revenue requirement $R(=0)$, the elasticity of substitution between consumption and leisure $\eta$, and the welfare attained in the absence of redistribution. The next three columns present the optimal demogrant, the tax rate and value of social welfare function. The following four columns present the result for the optimal education subsidy combined with a linear income tax and consist of the optimal demogrant, the linear tax rate, and education subsidy rate or the fraction of the original education cost paid by the government. The last two columns give welfare comparisons for a linear income tax policy relative to no redistribution and a linear income tax relative to this policy combined with an education subsidy.

The results of Table 1.8 indicate that:

- An educational subsidy combined with an optimal linear income tax is a slight improvement on an optimal linear income tax if the costs of education are not tax deductible.

- The rate of educational subsidy and the welfare effects are higher for a more egalitarian social welfare function ($\nu = -2$).

- The rate of subsidy and their welfare effects are significantly higher for smaller elasticities of substitution between consumption and leisure.

- If the private cost of education is tax deductible, the educational subsidy cannot improve social welfare.
The effects of the educational subsidy on education, the labor supply decision and utility are shown in Fig.1.9-1.11. We see that educational subsidy encourages education and discourages labor supply, a marked contrast with the effects of the wage subsidy.

The key result of this section is that if the private cost of education is tax deductible the educational subsidy cannot improve social welfare. Allowing private educational expenditures to be tax deductible eliminates under-investment in human capital and reduces the optimal educational subsidy to zero. This confirms that tuition subsidies are not a useful policy instrument in a one-period framework where individuals have no family history and tuition aid cannot be varied with family income. When educational expenditures are non-deductible a public subsidy for educational tuition is an improvement on a pure demogrant, as a redistributive tax will discourage work effort and investments in education. The resulting increase in welfare is quite small as the subsidy encourages education and increases wages; offsetting these positive effects are the higher taxes required to finance the public expenditures on education. So, while the educational subsidy encourages education it also discourages labor supply. Furthermore, as the educational subsidy induces more able individuals to spend more on education the degree of inequality will increase.

---

\(^{23}\) For the case of \( \nu = 1 \) and elasticity \( q = 0.2 \).
Table 1.8. Comparison of linear income tax and linear income tax with private cost of education subsidy structures.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with education subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non deductible cost of education</td>
<td>1.0</td>
<td>0.0</td>
<td>0.2</td>
<td>172.292</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
<td>0.4</td>
<td>169.992</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
<td>0.99</td>
<td>171.140</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>-20816.443</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.4</td>
<td>-21143.274</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.99</td>
<td>0.99</td>
<td>-20071.207</td>
</tr>
<tr>
<td>Tax deductible cost of education</td>
<td>1.0</td>
<td>0.0</td>
<td>0.2</td>
<td>172.292</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
<td>0.4</td>
<td>169.992</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
<td>0.99</td>
<td>171.140</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>-20816.443</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.4</td>
<td>-21143.274</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.99</td>
<td>0.99</td>
<td>-20071.207</td>
</tr>
</tbody>
</table>

* Increase in welfare is beyond precision of the simulation.
Fig. 1.9. No-tax, before and after price of education subsidy utility.

Fig. 1.10. No-tax, before and after price of education subsidy labor supply decisions.
1.3.3. The financing of a wage subsidy with a two bracket income tax.

One extension of the basic model is an analysis of a two-bracket income tax combined with a linear wage subsidy. This system is chosen as a compromise between simplicity and generality as it allows us to characterize a system as progressive or regressive. In order to simplify calculations the target wage and the subsidy rate are fixed at the values found optimal in Section 1.2.

The model is identical to the basic model of Section 1.2 except the equations for consumption (1.2.2.5) and (1.2.2.5') are replaced with equations (1.3.3.1) and (1.3.3.1') below:

Consumption for non-deductible education is now:

\[ C_t = \begin{cases} Y_t \cdot (1-t_1) + a - p_s \cdot E_t, & Y_t \leq Y_c \\ Y_t \cdot (1-t_1) + (Y_t - Y_c) \cdot (1-t_2) a - p_s \cdot E_t, & Y_t > Y_c \end{cases} \quad (1.3.3.1) \]

Consumption for tax-deductible education is:
\[ C_i = \begin{cases} (Y_i - p_e \cdot E_i) \cdot (1-t_1) + a, & Y_i \leq Y_c \\ Y_i \cdot (1-t_1) + (Y_i - Y_c - p_e \cdot E_i) \cdot (1-t_2)a, & Y_i > Y_c \end{cases} \]  

(1.3.3.1*)

Where \( Y_c \) is income bracket and \( t_1, t_2 \) are tax rates.

The results of the simulation are presented in Table 1.9.

The two conclusions first that an optimal two-bracket tax is regressive, and second that the welfare gains associated with adding a second tax bracket are small.²⁵

However, there is no guarantee that the result will hold if the two-bracket tax and the wage subsidy are jointly optimized. Computational complexity rules out this experiment at this time, though the problem is high on my future research agenda.

²⁵ Slemrod's et.al. (1994) made similar conclusions for the original Mirrlees model.
### Table 1.9. Comparison of one and two-bracket linear income tax structures.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Wage Subsidy</th>
<th>Linear Income Tax with Wage Subsidy</th>
<th>Two-bracket Linear Income tax with Wage Subsidy</th>
<th>Welfare Increase due to Two-bracket tax structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target wage</td>
<td>Subsidy rate</td>
<td>Demogr.</td>
<td>Tax rate</td>
</tr>
<tr>
<td>V 1.0</td>
<td>0.0</td>
<td>0.2</td>
<td>1.1</td>
<td>0.02</td>
</tr>
<tr>
<td>R 0.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.8</td>
<td>0.04</td>
</tr>
<tr>
<td>Q 0.99</td>
<td>1.3</td>
<td>0.03</td>
<td>0.084</td>
<td>0.156</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.45</td>
<td>0.1</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.5</td>
<td>0.16</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.99</td>
<td>0.5</td>
<td>0.18</td>
</tr>
</tbody>
</table>
1.3.4. **Analysis of an Earned Income Tax Credit (EITC).**

The EITC provides assistance through the U.S. tax system to low earning families with children. Currently, a family with two dependent children is paid a credit at a rate of 40 percent on wage earnings up to a total of $9400 a year. The maximum credit remains constant up to earnings of $12,400 and is phased out at a rate of 20 percent per dollar of income reaching zero at earnings of $30,000.

In discussions of welfare reform, (Blank [1997], Solow [1998]) the EITC is praised as an important anti-poverty devise especially as a time limitation has been imposed on eligibility for the cash component of some welfare programs.

I have analyzed a simple variant of an EITC, combining it with an hourly wage. The EITC is substituted for a uniform demographic. It is modeled without demographic detail and without an income range where the credit is constant. The EITC increases at a constant rate to a level of earnings, \( Y_1 \), and then declines between \( Y_1 - Y_2 \). If \( Y_1 \) is small the program is equivalent to a phased out demographic. If in addition \( Y_2 \) is large the EITC is equivalent to a universal demographic. Equation (1.3.4.1) below replaces equation (1.2.2.5) in the basic model for the consumption in the case of non-deductible education expenses.\(^\text{26}\) The rest of the model remains unchanged.

\[
C_i = \begin{cases} 
Y_i (1-t) + a \frac{Y}{Y_1} - p_e \cdot E, & \text{if } Y_i \leq Y_1 \\
Y_i (1-t) + a \left( \frac{Y_2 - Y_1}{Y_2 - Y_1} \right) - p_e \cdot E, & \text{if } Y_1 < Y_i \leq Y_2 \\
Y_i(1-t) - p_e \cdot E, & \text{if } Y_i > Y_2
\end{cases}
\](1.3.4.1)

\(^{26}\)The equation for tax deductible education expenses is similar.
As the joint optimization of an EITC and the wage subsidy is intractable, I fix the target wage and the subsidy rate at the optimal values found in the simulations of Section 1.2. The results are summarized in Table 1.10.\textsuperscript{27} For the assumed values of the skill distribution and educational levels cases the optimal value of $Y_1$ is found to be small; there is no individual below $Y_1$. The increase in welfare implies that the optimal income maintenance system is a phased out demogrant combined with a wage subsidy.

This result is not surprising as we expect the optimal income transfer portion of the system to be non-linear. To analyze the magnitude of the welfare improvement for a more precise form of the vanishing demogrant a more general version of the EITC needs to be considered.

\textsuperscript{27} For the non deductible educational expenses.
Table 1.10. Comparison universal demogrant and EITC in linear income tax with the wage subsidy redistribution scheme.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Wage subsidy</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Universal Demogrant to EITC welfare increase, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V</td>
<td>R</td>
<td>Q</td>
<td>Target wage</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.2</td>
<td>1.1</td>
<td>0.02</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.8</td>
<td>0.04</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.99</td>
<td>1.3</td>
<td>0.03</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.45</td>
<td>0.10</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.5</td>
<td>0.16</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>0.99</td>
<td>0.5</td>
<td>0.18</td>
</tr>
</tbody>
</table>
1.4. A two-market model of a wage subsidy.

Beaudry and Blackorby (1998) model a wage subsidy in a way which will not lead to wage equalization, as in the original Mirrlees model. They modify this model by abstracting from leisure and by specifying the existence of a taxable formal sector and a non-taxable informal sector. Income earned in the non-taxable sector is assumed to be a perfect substitute for income earned in the formal sector. The authors also make the important assumption that the productivity in the informal sector varies across individuals and is not necessarily equal to their productivity in the formal sector.

The authors develop a general argument that it will be efficient to subsidize and encourage the work of the relatively unskilled and to finance the subsidy with taxes on the more able. The effect of such a system is quite different from a universal demogrant which decreases the work effort of the unskilled.

Beaudry and Blackorby note that the wage subsidy will deter persons of moderate ability from passing themselves off as unskilled by earning most of their income in the informal sector, reporting a low taxable income and collecting an income transfer. Molefsky (1982) reports an IRS study which estimates that 25 percent of the underground economy comes from welfare recipients who do not report earned income - a fact also noted by Blank (1997,p.79). In addition, there is a growing literature on the underground (shadow) economy recently surveyed by Schneider and Enste (2000). These authors estimate that about 10 percent of GDP in the US is not recorded in officially recorded statistics. This percentage is as high as 30 percent for some other OECD countries, notably Italy and Spain.
I combine the Beaudry-Blackorby formulation with the basic model of the previous section by allowing for three activities: work in the formal and informal sectors and leisure. I first analyze this model on the assumption that the capacity to enjoy leisure is identical across individuals and that the skill distributions are exogenous, i.e., there is no education. After-tax wages cannot be equalized in this situation as high ability persons will leave the formal sector. To avoid the unrealistic outcome where no individual works simultaneously in both sectors I assume that earnings for each individual are subject to decreasing returns in the informal sector.

I also analyze a model with endogenous skill acquisition and constant returns for individual earnings in both sectors. Wage rates in the formal market depend on the ability distribution and education, while wage rates in the informal sector are exogenous and are correlated with the individuals' native abilities.

In another version of the two-market model I specify leisure and a single wage market and assume that individuals have differential capacities to enjoy leisure. The bivariate distribution of abilities is assumed to be highly correlated across markets. Moreover, it is postulated that the ability of low-ability persons to "produce" leisure is higher, relative to their goods-producing ability, than it is for high ability persons. The rest of the model is the same as for the base case of Section 1.2, in which marketable skill is partially acquired through education.

Several considerations motivate these extensions. Variations in leisure valuation as well as of non-taxable activities are observed in the real world. The results of Section 1.2, where individuals have the same valuation of leisure, indicate that the use of a wage subsidy is somewhat limited. It adversely affects investment in education which
counteracts the positive effect of the wage subsidy on labor supply. It is expected that the wage subsidy will be a more efficient instrument of redistribution when non-taxable activity is an employment option, as in this setting a wage subsidy can be used as an instrument of self-selection.

Suppose there is a group of individuals, \( L \), with low observable incomes and low (unobservable) utilities because they have low taxable market and non-taxable market abilities. Suppose there is another group, \( H_{nm} \), with low taxable market potential but with higher non-taxable market abilities. This second group of individuals will have higher utility levels than the first group. Ideally we would like to redistribute some income from the \( H_{nm} \) to \( L \) group. But there is no way to do this with an income based redistribution system. As the \( H_{nm} \) group devotes most of its time to the non-taxable market its observable incomes are low. In order to carry out an income transfer in this setting a labor contingent income transfer or wage subsidy can be used as an instrument for self-selection.

The results for these models are similar to those obtained in the previous section. The addition of a wage subsidy increases welfare significantly but the demogrant remains an important instrument of redistribution. The wage subsidy becomes increasingly more important if abilities are weakly correlated across the markets. In this case observed incomes do not convey much information on individual utilities and income based redistribution is inefficient.
1.4.1. A two market model.

Individuals can earn their incomes in two markets. In the formal market earnings are observed by the government and are taxed. In this market an individual's income is a product of his/her real wage and the amount of labor supplied

\[ Y_{t1} = w_{t1} \cdot l_{t1} \]  
(1.4.1.1)

The informal market is a market of unreported earnings, home, illegal or other non-taxable activity. The government does not observe individuals' wages or incomes earned in this market. A key assumption about the second market is that it exhibits decreasing return to scale in labor supply\(^{28}\), so

\[ Y_{t2} = w_{t2} \cdot l_{t2}^\eta, \quad \eta < 1. \]  
(1.4.1.2)

As a consequence, only a few individuals will work entirely in the second market. A large portion of the population will supplement its first-market taxable earnings with some amount of unreported second market income. The incomes earned in the two markets are assumed to be perfectly substitutable so that consumption is:

\[ C_i = Y_{t1} \cdot (1-t) + a - p_e \cdot E_i + Y_{t2} \]  
(1.4.1.3)

when private education expenses are not tax deductible, and:

\[ C_i = (Y_{t1} - p_e \cdot E_i) \cdot (1-t) + a + Y_{t2} \]  
(1.4.1.3')

when education expenses are tax deductible.

The observed wage in the formal market is a function of individuals' natural abilities and their education levels. Wages in the informal market are independent of the

\(^{28}\) This somewhat ad hoc assumption is made to ensure that most individuals work in both markets. It captures the idea that it becomes progressively more difficult for any individual to earn money in non-market activity.
level of education and are equal to individuals’ ability to earn income in the informal sector.

Leisure is the time an individual has after working in both markets and after acquiring education:

\[ L_i = (1 - I_{i1} - I_{2i} - E_i) \]  \hspace{1cm} (1.4.1.4)

As the general model is very difficult to solve computationally I develop three submodels or special cases each of which contains a key feature of the general model. The first case is one where there is no education and wages are exogenous. In the second case wages are endogenous and production is linear in both markets (i.e. the model with education) and in the third case only the formal market exists but the valuation of leisure varies across individuals.

1.4.2. Two-market model with no education.

In this section I examine a Mirrlees type model with a wage subsidy and two markets. Individuals’ observable wages in the first market are set equal to their natural ability \( w_{obsi} = w_i^* \). The education level is set to zero \( E_i = 0 \).\textsuperscript{29}

The following lognormal bivariate distribution of abilities was generated to solve the model: \( \mu_1 = -1., \mu_2 = -1.2, \sigma_1 = 0.39, \sigma_2 = 0.2 \) and correlation \( \rho = 0.999 \). Abilities or wages in the two markets are assumed to be highly correlated to each other. Wages in the informal market are smaller and exhibit less variance than the wages in the formal market. The no-tax wage distribution is shown on the Figure 1.12.

\textsuperscript{29} Technically, in the general model the price of education is set equal to infinity \( p_e = \infty \), so that individuals will choose no education.
Note that wages for the low ability individuals in the informal market can be higher than their wages in the formal one. A set of two simulations was completed. The revenue requirement was set to zero. The elasticity of the CES utility function has the parameter values \( q = (0.2, 0.99) \) and \( \eta = 0.5 \). The welfare function is additive (\( \nu = 1 \)). The results are presented in Table 1.11.

![Fig. 1.12. No-tax wage distribution.](image)

Table 1.11. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \nu )</td>
<td>( R )</td>
<td>( q )</td>
<td>Welfare</td>
<td>Demo grant</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.2</td>
<td>331.055</td>
<td>0.0085</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.99</td>
<td>326.605</td>
<td>0.0089</td>
</tr>
</tbody>
</table>
Figures 1.13a,b present the labor supplies observed in the formal and informal markets for the cases of no-tax, a linear income tax unassisted and a linear income tax with a wage subsidy. The wage subsidy discourages work in the informal market and encourages work in the formal, taxable market. The linear income tax with only a demigrant has an opposite effect. It encourages work in the informal market and discourages work in the formal, taxable market. The complete equalization of wages cannot occur in this model as workers will leave the formal sector. The wage subsidy significantly increases welfare relatively to the income tax alone.\textsuperscript{30}

This model shares some of the limitations of the Mirrlees model: workers have no incentive to change jobs and ability in the formal sector is observable. Nevertheless, it illustrates the essential points of the two-market model. The linear income by itself has a limited redistributive capacity and the addition of the wage subsidy to a demigrant increases redistribution quite significantly.

\textsuperscript{30} 0.35\% vs. 0.052\% for the case of $q = 0.2 \ 0.21\%$ vs. 0.08\% for the case of $q = 0.99$
Fig. 1.13b. No-tax, linear income tax and linear income tax with wage unobservable (informal) market labor supply.

Fig. 1.14. No-tax, linear income tax and linear income tax with wage subsidy utility.
1.4.3. **Two-market model with linear informal market production.**

The model examined in this section is a constant returns version of the one introduced at the beginning of Section 1.4.1. The model is described by equations 1.4.1.1–1.4.1.4 with $\eta = 1$. Production in the informal market is now linear in labor so that $Y_{2i} = w_m \cdot l_{2i}$.

Lognormal bivariate distribution of abilities was generated to solve the model with $\mu_1 = -1$, $\mu_2 = -1.2$, $\sigma_1 = 0.39$, $\sigma_2 = 0.2$ and correlation $\rho = 0.999$. The no-tax wage distribution is depicted in Fig 1.15.

Though this ability distribution is the same as in the previous Section 1.4.2 the no-tax-wage distribution is different for this model. An individual never works in the formal and informal markets simultaneously. Individuals who work exclusively in the informal market do not invest in skills as by assumption education does not affect informal market wages. Consequently, individuals' formal market wages depicted in Fig 1.15 “jump” at the point where individuals' abilities in the first market are high enough relative to their abilities in the second market so that individuals who switch to the formal market begin to invest in education.

---

31 In this sense the second market is of the Mirrlees type. This is an obvious simplification as education will also affect informal market wages. However, we deviate from the traditional Mirrlees model as both the income tax and wage subsidy are used. The informal market is not a subject to taxes so the original Mirrlees behavioral assumptions do not lead to the same results as they would if they were employed to characterize the formal taxable sector analysis and are a quite reasonable approximation of reality.
A set of eight simulation was conducted for two different values of the elasticity of the CES utility function $q = \{0.2, 0.099\}$, two revenue requirements $R = \{0, 0.1\}$, and two welfare functions $(\nu = \{1, -2\})$. The private cost of education is assumed to be non-deductible for all the simulations. The results are summarized in Table 1.12.

Figures 1.16-1.18 show markets labor supplies, education choices and utilities for this simulation. The graphs are for an elasticity $q = 0.2$ and an additive social welfare function $(\nu = 1.)$

Figures 1.16a and 1.16b depict labor supply effect of the income tax and the wage subsidy in both markets. As expected, an income tax discourages labor supply in the formal taxable market while the wage subsidy encourages it. This effect is more complex than one observed in the base model. A wage subsidy affects the labor supply of the people working in the formal market and also induces more people to quit the informal sector. The wage subsidy discourages education by reducing its level for individuals who
already work in the formal market. However low-income individuals who switch to the formal sector will now invest in education.

These considerations explain the unusual shape of the graphs: there are jumps in both education and labor supply as individuals leave the informal and enter the formal market. A higher wage subsidy will induce more individuals with low ability to work in the formal sector. The point at which the formal market labor supply becomes positive shifts left along the first market ability axe. Also the curve shifts as a result of an increase in the labor supply of those working in the formal sector. A similar explanation can be given for changes in the labor supply in the informal sector and in the amount of education.

There is a significant increase in welfare following the addition of a wage subsidy to the demogrant. The effect is higher for an additive welfare function. The introduction of a revenue requirement enhances the relative effectiveness of the wage subsidy. The explanation of this result is straightforward. When a demogrant is used for redistribution individuals shift out of the taxable sector and revenue is harder to collect. In contrast a wage subsidy creates more “taxpayers”.

We conclude that the wage subsidy increases welfare, encourages labor supply in the formal taxable market and encourages low ability people to acquire additional education. However, the wage subsidy discourages investment in education and reduces the productivity of high ability individuals. A wage subsidy is especially effective when a revenue requirement is added to the redistributive objective.
Table 1.12: Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\nu$</th>
<th>$R$</th>
<th>$q$</th>
<th>Welfare</th>
<th>Demo Grant</th>
<th>Tax rate</th>
<th>Welfare</th>
<th>Demo Grant</th>
<th>Tax rate</th>
<th>Target wage</th>
<th>Subs. rate</th>
<th>Welfare</th>
<th>No-tax to Linear Income Tax Welfare increase %</th>
<th>Lin. Income Tax to Linear Income Tax &amp; Wage subsidy Welfare increase %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.2</td>
<td></td>
<td>172.862</td>
<td>0.0064</td>
<td>0.042</td>
<td>172.989</td>
<td>0.0173</td>
<td>0.185</td>
<td>0.9</td>
<td>0.04</td>
<td>173.467</td>
<td>0.073</td>
<td>0.28</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.99</td>
<td></td>
<td>172.034</td>
<td>0.0058</td>
<td>0.042</td>
<td>172.266</td>
<td>0.0120</td>
<td>0.170</td>
<td>1.7</td>
<td>0.02</td>
<td>172.685</td>
<td>0.14</td>
<td>0.24</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>0.2</td>
<td></td>
<td>-19987.085</td>
<td>0.0170</td>
<td>0.131</td>
<td>-18856.337</td>
<td>0.0270</td>
<td>0.342</td>
<td>1.6</td>
<td>0.05</td>
<td>-17754.905</td>
<td>5.66</td>
<td>5.84</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>0.99</td>
<td></td>
<td>-19286.210</td>
<td>0.0128</td>
<td>0.114</td>
<td>-18430.409</td>
<td>0.0187</td>
<td>0.258</td>
<td>0.7</td>
<td>0.10</td>
<td>-17704.173</td>
<td>4.44</td>
<td>3.94</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
<td>N/A</td>
<td>-0.0772</td>
<td>0.137</td>
<td>88.830</td>
<td>-0.0740</td>
<td>0.226</td>
<td>0.9</td>
<td>0.05</td>
<td>90.488</td>
<td>N/A</td>
<td>1.8</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.99</td>
<td></td>
<td>N/A</td>
<td>-0.0772</td>
<td>0.150</td>
<td>87.008</td>
<td>-0.0765</td>
<td>0.231</td>
<td>3.9</td>
<td>0.01</td>
<td>89.201</td>
<td>N/A</td>
<td>2.5</td>
</tr>
<tr>
<td>-2</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
<td>N/A</td>
<td>-0.0721</td>
<td>0.211</td>
<td>-149971.647</td>
<td>-0.0743</td>
<td>0.342</td>
<td>0.5</td>
<td>0.29</td>
<td>-79701.212</td>
<td>N/A</td>
<td>46.8</td>
</tr>
<tr>
<td>-2</td>
<td>0.1</td>
<td>0.99</td>
<td></td>
<td>N/A</td>
<td>-0.0750</td>
<td>0.195</td>
<td>-199025.447</td>
<td>-0.0788</td>
<td>0.331</td>
<td>0.5</td>
<td>0.29</td>
<td>-81305.098</td>
<td>N/A</td>
<td>59.1</td>
</tr>
</tbody>
</table>
Fig. 1.16a. No-tax, linear income tax alone and linear income tax with wage subsidy observable (formal) market labor supply.

Fig. 1.16b. No-tax, linear income tax alone and linear income tax with wage subsidy unobservable (informal) market labor supply.
Fig. 1.17. No-tax, linear income tax alone and linear income tax with wage subsidy education choices.

Fig. 1.18. No-tax, linear income tax alone and linear income tax with wage subsidy utility.
1.4.4. **Two-market model with formal market production and variable leisure valuation.**

In the model examined previously in Sections 1.2 and 1.3 the valuation of leisure was constant across the population. We now modify the analysis by allowing the value of leisure to vary across the population.

With one important except all the equations in this section as well as the social welfare function are the same as for the basic model of Section 1.2. The difference is in the functional form of the second term of the utility function (or leisure). In the one-market models examined so thus far leisure was defined as: \( L = (1 - l - E) \). In this section the basic model is modified so that \( L = \theta_i \cdot (1 - l - E) \).

This second term can now be interpreted as:

- Leisure. In this case every individual now has his own valuation of free time used for leisure \( \theta_i \).

- Income from non-taxable activity such as home employment, or any non-declared earning or non-taxable market activities. \(^{32}\) In this case \( \theta_i \) is individuals' productivity or wage in the market the government cannot tax.

An important feature of the model examined in this chapter is that what is "produced" in the second market is not perfect substitute for formal market earnings.

The model was solved for two distributions of abilities:

---

\(^{32}\) There are some evidence that the non-taxable activities or tax evasion are quite significant. One study conducted in a low income Chicago neighborhood showed that people with declared earnings of less than $10,000 a year spend full 254% of what they make according to official data while those with earnings between $10,000 and $20,000 a year spend 140% of what they make ( "Up From Poverty", The Wall Street Journal, Dec 28 1998, pp. A14.)
Distribution A is identical to the distribution used in sections 1.4.2, 1.4.3: $\mu_1 = -1.1, \mu_2 = -1.2, \sigma_1 = 0.39, \sigma_2 = 0.2$ and correlation $\rho = 0.999$. The no-tax wages for this distribution are shown on Figure 1.19.

Distribution B is identical to Distribution A except there is virtually no correlation between abilities and leisure valuations: $\mu_1 = -1.1, \mu_2 = -1.2, \sigma_1 = 0.39, \sigma_2 = 0.2$ and correlation $\rho = 0.001$. The no-tax wages for this distribution are shown in Figure 1.20.

Distribution A is used in all the models of this section. It reflects the assumptions that earnings on the informal market are in general lower than in the formal market\textsuperscript{33}, that informal market earnings exhibit less variation\textsuperscript{34}, and that a person who can earn more in the formal market will earn more if he does a similar job and underreports earnings, i.e. wages in the two markets are strongly correlated.

Distribution B is introduced for two reasons. First, to test the implication of the assumption that leisure valuation is not correlated with income. Also, Distribution B allows as to test the effectiveness of a wage subsidy as a self-selection device. Assume the government designs a redistribution scheme based solely on income transfers. As it is shown in Figure 1.21 for Distribution A income and utility are strongly correlated. The use of income as a signal allows equalization to be achieved. With Distribution B the situation is quite different. Figure 1.22 shows a particular income level of 0.08 can be exhibited by a number of individuals (each at the intersection of the “blue line” of income and 0.08 level line). All of these persons have different utilities. It is very difficult in this case to carry redistribution on the basis of observable income alone.

\textsuperscript{33} e.g. in reality one can easily find $7$/hour job with no “paper trail”. $50$/h job is another matter.
Fig. 1.19. No-tax wages. Distribution A.

Fig. 1.20. No-tax wages. Distribution B.

34 One can work "for cash" at probably 5-25$ per hour, formal market presents much more variation.
A set of four simulations was conducted for each distribution for $q = \{0.2, 0.99\}$, a revenue requirement $R = 0.$, and two welfare functions ($\nu = \{1., -2\}$). The results of the simulations are presented in Table 1.13 for the Distribution A and Table 1.14 for Distribution B. In addition Figures 1.23-1.25 present the results.
choices and utility. The graphs are for Distribution A, elasticity of the CES utility function is $q=0.2$ and additive social welfare function ($\nu = 1.$)

We conclude that all the results of the base model hold: the wage subsidy improves welfare, encourages labor supply and discourages education.

When abilities are not correlated and if consumption and leisure are not substitutable the introduction of a wage subsidy to an income transfer will increase welfare more significantly than in one-market case of Section 1.2. For Distribution B and elasticity $q = 0.2$ the wage subsidy increases welfare 241% for $\nu = 1$ and 19.4% $\nu = -2$. In the one-market case the corresponding welfare increases are 12% and 2%.\textsuperscript{35} This is explained by the presence a number of individuals who have identical incomes but different utilities. Exclusive reliance on income based redistribution is inefficient in this situation.

\textsuperscript{35} 27% and 2.41% for Distribution A
Table 1.13. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures. Distribution A.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Welfare</td>
<td>Demo grant</td>
<td>Tax rate</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$R$</td>
<td>$q$</td>
<td>Welfare</td>
<td>Demo grant</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.2</td>
<td>77.698</td>
<td>0.0135</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.99</td>
<td>93.060</td>
<td>0.0091</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>0.2</td>
<td>-99520.116</td>
<td>0.0415</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>0.99</td>
<td>-75364.138</td>
<td>0.0431</td>
</tr>
</tbody>
</table>

Table 1.14. Two-market economy. Comparison of linear income tax and linear income tax with wage subsidy structures. Distribution B.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Tax</th>
<th>Linear Income Tax</th>
<th>Linear Income tax with Wage Subsidy</th>
<th>Comparison of the tax structures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Welfare</td>
<td>Demo grant</td>
<td>Tax rate</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$R$</td>
<td>$q$</td>
<td>Welfare</td>
<td>Demo grant</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.2</td>
<td>76.041</td>
<td>0.0120</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.99</td>
<td>91.946</td>
<td>0.0155</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>0.2</td>
<td>-95055.568</td>
<td>0.0333</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>0.99</td>
<td>-70269.049</td>
<td>0.0400</td>
</tr>
</tbody>
</table>
Fig. 1.23. No-tax, linear income tax and linear income tax with wage subsidy observable (formal) market labor supply.

Fig. 1.24. No-tax, linear income tax and linear income tax with wage subsidy education choices.
Fig. 125. No-tax, linear income tax and linear income tax with wage subsidy utility.
1.5. **Concluding remarks.**

In this paper I have extended the Mirrlees model of optimal taxation by allowing for endogenous skill acquisition or job choice. My analysis concludes that in an optimal income maintenance system a wage subsidy should be combined with an income transfer. These two instruments are complementary as there are two incentive margins: the wage subsidy affects job choice while the demogrant affects the labor-leisure decision. The relative efficiency of the wage subsidy improves with increases in the elasticity of labor supply. The wage subsidy is paid to most workers as it encourages individuals to work more and better utilize their accumulated human capital.

When a general revenue requirement is introduced the two instruments have quite limited redistributive potential individually as the marginal efficiency cost of an incremental redistributive tax will be higher in presence of an existing tax distortion. However, a wage subsidy financed by a negative demogrant (a head tax) and a small increase in a flat rate tax is able to redistribute income and to increase welfare quite significantly.

A wage subsidy is also quite effective if utilized when a non-observable non-taxable labor market co-exists with a formal sector. The wage subsidy complements a demogrant by acting as an effective screening mechanism for different persons with different ability who have low observed income. The screening effect is clearly indicated by the result that both the optimal demogrant and the optimal tax rate are increased when the subsidy is added to the demogrant. This is in marked contrast to the results for the one-sector economy discussed in Section 1.2 where the wage subsidy and the demogrant are “substitutes”. 
Experiments for non-welfarist objective function such as income maintenance and the minimization of a poverty index also provide support for the use of wage subsidies in combination with a phased-out demogrant. Nevertheless, it is important to reiterate one of the basic points of this essay. A wage subsidy has an efficiency cost: it decreases skill acquisition and distorts job choice.
CHAPTER II.

BUNDLING HARDWARE AND SOFTWARE.36

2.1. **Introduction.**

One of the most significant developments in the late 20th century is the growth of knowledge-based industries. Many products used to be mere combinations of steel, rubber, glass, plastic, and paper glued together by human sweat and toil. Today more and more products embody human knowledge and art as essential components. Most computer products, for instance, are bundles of electronics, called hardware, and information processing algorithms, called software.

The software components incur the fixed costs of development while the hardware components incur marginal costs for incremental consumers. Based primarily on this observation, one may extend the usage of the terms "software" and "hardware" to products outside of the computer industry. One may say for instance that a music CD is a bundle of music, the software, and a compact disc, the hardware. The same can be said for a video cassette containing a movie. These are rather obvious examples. But there are other more subtle example. A PCS wireless phone based on the CDMA technology is a bundle of the CDMA technology, the software, and a portable device, the hardware.

A question of economic interest is what determines a supplier's bundling choice, that is, the selection of the exact types of hardware and software (out of many possible types) produced, combined, and sold in a single "system" such as a computer system or music CD. Even though consumers sometimes bundle the hardware and software compo-

36. This essay is written in collaboration with Dr. Suchan Chae.
ments themselves, the predominant practice is that they purchase ready-made bundles from suppliers.

The literature on bundling identifies three broad motives for bundling: monopolistic price discrimination, strategic considerations, and natural bundling due to technology.

The first motive, bundling due to a producer's desire to extract consumer surplus, has been analyzed by Adams and Yellen (1976) and later by Spence (1980), Paroush and Peles (1981), Schmalensee (1982, 1984), Lebwel (1985), McAfee, McMillan and Whinston (1989). The strategic incentives for bundling have been examined in Carbajo, De Meza and Seidman (1990), Whinston (1990), Matutes and Regibeau (1990). Natural bundling due to economies of scope in distribution technology was examined in Chae (1992).

In this paper, we primarily deal with products that are in a sense bundled "out of necessity". Thus we focus on the question of "how to bundle" rather than "why to bundle". Nevertheless, two of the three motives for bundling identified above, a producer's desire to extract consumer surplus and the economies of scope in distribution technology, are useful in analyzing the problem of "how" to bundle hardware and software.

Note the unique features of products bundled. Software requires a high fixed cost of research and development. Once, however, the first copy is produced, the cost of reproducing additional copies can be realistically taken to be zero. Hardware is a different kind of product. The marginal cost of production is relatively high.

In this paper, we construct and analyze a class of simple theoretical models reflecting the cost characteristics of hardware and software to explain some bundling practices. We also draw welfare implications of such practices from these models.
In Section 2.2, we present a basic model with a monopolistic producer and two types of consumer with symmetric preferences and analyze the profit-maximizing and welfare-maximizing outcomes. By comparing the two outcomes, we find that a profit-maximizing monopolist may over-provide hardware-software bundles. In Section 2.3, we extend this model by allowing more than one software type to be bundled within a single system. Assuming that preferences are symmetric with respect to the two software types, we find that both the profit-maximizing and welfare-maximizing monopolists practice "pure bundling". In Section 2.4, we extend the model further to allow for asymmetric preferences. Here we find that a profit-maximizing monopolist under-bundles and under-produces software. In Section 2.5, we provide the conclusion.

2.2. Basic model.

There is a monopolist that produces both hardware and software. Alternatively, one may think of multiple hardware and software producers colluding to maximize their joint profits. The monopolist can produce two types of hardware: \( l \) and \( h \). Type \( l \) is low-quality and type \( h \) high-quality hardware.

The marginal cost of producing the low-quality hardware is \( c > 0 \). The marginal cost of the high-quality hardware is \( c - d \), where \( d > 0 \). Call \( d \) the hardware cost differential. The fixed cost for hardware is assumed to be zero.

There are two types of software: \( a \) and \( b \). Both types of software have zero marginal cost and identical fixed cost of development \( f > 0 \).

Each system consists of one type of hardware bundled with one type of software. Thus, the firm can produce four possible systems:
1) low-quality hardware with type \( a \) software installed, denoted \( la \);

2) low-quality hardware with type \( b \) software installed, denoted \( lb \);

3) high-quality hardware with type \( a \) software installed, denoted \( ha \);

4) high-quality hardware with type \( b \) software installed, denoted \( hb \).

Note that systems \( la \) and \( lb \) have identical marginal costs of production equal to \( c \), and systems \( ha \) and \( hb \) have identical marginal costs of production equal to \( c - d \). The fixed cost of any type of system is \( f \), the development cost of software.

In order to motivate the above set-up with real-life examples, consider the computer industry. In this industry, there are monopolistic producers that produce both hardware and software such as Apple Computer Corp. Based on the mid 90's state of technology, an example of low-quality hardware is MC68XX CPU for Apple Macintosh (or AMD K6/IBM Cyrix/Intel Celeron CPU for IBM PC), and an example of high-quality hardware is Power PC CPU (or Intel Pentium II CPU for IBM PC).

As far as the software types are concerned, think of type \( a \) as "office" software such as word processors, spreadsheets, and communication software and type \( b \) as "research" software such as programming languages, CAD, and statistical software. In the Appendix, we provide an example of a producer offering "office" and "home" software. Obviously, other examples can also be found. What we need is any two types of software such that different groups of consumers have different preference rankings over the types.

There are two types of consumers in the market: \( A \) and \( B \). Each type consists of identical people of measure 1. A consumer of type \( A \) prefers software of type \( a \) to software of type \( b \). A consumer of type \( B \) prefers software of type \( b \) to software of type \( a \). The high-quality hardware is preferred to the low-quality hardware by both consumers.
Assume that consumers have the following reservation prices for different systems: $x (> 0)$ for the low-quality hardware bundled with the less preferred software; $xv$, where $v > 1$, for the low-quality hardware bundled with the preferred software; $qx$, where $q > 1$, for the high-quality hardware bundled with the less preferred software; $q xv$ for the high-quality hardware bundled with the preferred software. Call parameter $q$ quality preference and $v$ preference for one's own variety. Observe that the preference structure is such that a consumer's marginal value of software is increasing in hardware quality level. This reflects the idea that, using computer products for illustration, the value of the same software is higher if it is installed on a high-speed computer.

If one denotes by $U_j(i)$ the reservation price for system $j \in \{ \emptyset, la, ha, lb, hb \}$ of consumer $i (= A, B)$, where $\emptyset$ denotes a null system (that is, purchasing nothing), the values of $U_j(i)$ are as in Table 2.1.

<table>
<thead>
<tr>
<th>$j$</th>
<th>$U_A(j)$</th>
<th>$U_B(j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$la$</td>
<td>$xv$</td>
<td>$x$</td>
</tr>
<tr>
<td>$ha$</td>
<td>$xqv$</td>
<td>$xq$</td>
</tr>
<tr>
<td>$lb$</td>
<td>$x$</td>
<td>$xv$</td>
</tr>
<tr>
<td>$hb$</td>
<td>$xq$</td>
<td>$xqv$</td>
</tr>
</tbody>
</table>

We assume $c < x$ and $c + d < xq$. The first inequality says that the marginal cost of the low-quality hardware is less than a consumer's reservation price for the low-quality hardware bundled with the less preferred software. The second inequality says that the
marginal cost of high-quality hardware is less than a consumer's reservation price for the high-quality hardware bundled with the less preferred software.

Each consumer makes her purchase decision to maximize her consumer surplus. The monopolist maximizes its profits.

2.2.1. **Profit Maximization.**

Notice that since there are only two types of consumers, the firm will supply at most two systems out of four possible systems. Observing the prices of the systems supplied, a consumer has the option of purchasing one of the systems supplied or nothing.

Denote the system intended for consumers of type \( i \) by \( J(i) \) \((i = A, B)\) and the cost of producing the line of products \( \{J(A), J(B)\} \) by \( C\{J(A), J(B)\} \). Note that \( J(A) \) and \( J(B) \) can be identical. The firm's problem is first to choose systems \( J(A), J(B) \in \{\emptyset, la, ha, lb, hb\} \) and then to choose prices \( P_{J(A)}, P_{J(B)} \) to solve the maximization problem:

\[
\text{Max } P_{J(A)} + P_{J(B)} - C\{J(A), J(B)\}
\]

\[
s.t. \ U_i(J(i)) - P_{J(i)} \geq U_i(j) - P_j \text{ for } j = \emptyset, J(A), J(B) \text{ for } i = A, B.
\]

Denote the maximum profit associated with the above maximization problem by \( \Pi(J(A), J(B)) \). Then the firm's product decision is the solution to the maximization problem

\[
\text{Max } \Pi(J(A), J(B))
\]

\[
s.t. J(A), J(B) \in \{\emptyset, la, ha, lb, hb\}
\]
The product decision depends on the cost and preference parameters. There are 25 possible choices of \((J(A), J(B))\). Not all these choices are feasible. For instance, if \((J(A), J(B)) = (la, ha)\), there are no prices \(P_{J(A)}, P_{J(B)}\) that would induce consumers \(A\) and \(B\) to purchase \(J(A)\) and \(J(B)\), respectively, for the constraints in the maximization problem are
\[
x_v - P_{la} \geq x_q v - P_{ha}, \quad x_q - P_{ha} \geq x - P_{la}, \quad x_v - P_{la} \geq 0, \quad x_q - P_{ha} \geq 0.
\]

The first two inequalities, which are incentive-compatibility constraints, are not compatible with each other if consumers prefer their own software types as assumed (i.e., \(v = 1\)). It is easy to check that the following product choices are infeasible in this sense:

<table>
<thead>
<tr>
<th>((J(A), J(B)))</th>
<th>Incentive Compatibility Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>((la, ha))</td>
<td>(P_{ha} - P_{la} \geq x v (q - 1), P_{ha} - P_{la} \leq x (q - 1))</td>
</tr>
<tr>
<td>((lb, la))</td>
<td>(P_{la} - P_{lb} \geq x (v - 1), P_{la} - P_{lb} \leq (-x)(v - 1))</td>
</tr>
<tr>
<td>((lb, ha))</td>
<td>(P_{ha} - P_{lb} \geq x (q v - 1), P_{ha} - P_{lb} \leq x (q - v))</td>
</tr>
<tr>
<td>((hb, la))</td>
<td>(P_{hb} - P_{la} \leq x (q - v), P_{hb} - P_{la} \geq x (q v - 1))</td>
</tr>
<tr>
<td>((hb, ha))</td>
<td>(P_{ha} - P_{hb} \geq x q (v - 1), P_{ha} - P_{hb} \leq (-x q)(v - 1))</td>
</tr>
<tr>
<td>((hb, lb))</td>
<td>(P_{hb} - P_{lb} \leq x (q - 1), P_{hb} - P_{lb} \geq x v (q - 1))</td>
</tr>
</tbody>
</table>

It is also easy to check that selling nothing to one or both types of consumers is not optimal because it is dominated by selling something to both types of consumers.

For example, the product choice \((J(A), J(B)) = (la, \emptyset)\) is dominated by the product choice \((J(A), J(B)) = (la, lb)\). Indeed, one has
\[ \Pi(la, lb) = 2(xv - c - f) = 2\Pi(la, \emptyset) > \Pi(la, \emptyset). \]

The remaining product choices and the corresponding profits are shown in Table 2.3. For instance, the maximization problem associated with the product choice \((J(A), J(B)) = (ha, la)\) is

\[
\begin{align*}
\text{Max} & \quad P_{ha} + P_{la} - 2c - d - f \\
\text{s.t.} & \quad \begin{cases} 
xqv - P_{ha} \geq xv - P_{la} \\
x - P_{la} \geq qx - P_{ha} \\
x - P_{la} \geq 0 \\
xqv - P_{ha} \geq 0
\end{cases}
\end{align*}
\]

The solution is

\[ P_{la} = x, \quad P_{ha} = xqv - xv + x. \]

\[ \Pi(ha, la) = x\{v(q - 1) + 2\{ - 2c - d - f. \}
\]

The same maximized profit obtains if the product choice is \((J(A), J(B)) = (lb, hb). \)

**Table 2.3. Feasible and undominated system choices**

<table>
<thead>
<tr>
<th>System Type</th>
<th>((J(A), J(B)))</th>
<th>((P_{J(A)}, P_{J(B)}))</th>
<th>Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>((ha, lb))</td>
<td>((q xv, xv))</td>
<td>((q + 1) xv - 2c - d - 2f)</td>
</tr>
<tr>
<td></td>
<td>((la, hb))</td>
<td>((xv, q xv))</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>((ha, ha))</td>
<td>((qx, qx))</td>
<td>(2qx - 2c - 2d - f)</td>
</tr>
<tr>
<td></td>
<td>((hb, hb))</td>
<td>((qx, qx))</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>((ha, hb))</td>
<td>((q xv, q xv))</td>
<td>(2q xv - 2c - 2d - 2f)</td>
</tr>
</tbody>
</table>
### Table 2.3. Feasible and undominated system choices

<table>
<thead>
<tr>
<th>System Type</th>
<th>((J(A), J(B)))</th>
<th>((P_{J(A)} - P_{J(B)}))</th>
<th>Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>((la, la))</td>
<td>((x, x))</td>
<td>(2x - 2c - f)</td>
</tr>
<tr>
<td></td>
<td>((lb, lb))</td>
<td>((x, x))</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>((la, lb))</td>
<td>((x\nu - x\nu - x, x))</td>
<td>(2x\nu - 2c - 2f)</td>
</tr>
<tr>
<td></td>
<td>((lb, hb))</td>
<td>((x, qx\nu - x\nu - x))</td>
<td></td>
</tr>
</tbody>
</table>

Comparing the maximized profits associated with the above six product choices, one can find the firm's optimal product choice.

**Theorem 2.1.** The optimal choices corresponding to different cost and preference parameters \((c, d, f, x, q, v)\) of a profit maximizing firm are as shown in Figure 2.1.

![Fig. 2.1. Profit-maximizing product choices.](image-url)
Figure 2.1 depicts five different regions in the cost space. Each region corresponds to a particular profit-maximizing choice of systems. Note that system type 0 in Table 2.3 is never optimal while system types 1-5 are optimal for Regions 1-5, respectively. The kinked boundary of the graph depicts the non-negative profit constraint. If the costs are sufficiently high, some of the regions depicted in Fig 2.1 may not be feasible. It can be shown, however, that if the marginal cost of hardware, c, is small and the preference for a consumer's own variety, v, is less than 2, then all five regions are feasible.37

To obtain Figure 2.1, the profits associated with different product choices have to be compared with each other. For instance, the product choice \((J(A), J(B)) = (ha, hb)\) is preferred to \((J(A), J(B)) = (ha, la)\) when

\[
\Pi(ha, hb) = 2xvq - 2c - 2d - 2f > \Pi(ha, la) = x(v(q - 1) + 2) - 2c - d - f ,
\]

i.e.,

\[
d < (q - 1)xv + 2x(v - 1) - f = (q - 1)(2 - v)x + (v - 1)2qx - f ,
\]

which is the equation of the frontier between regions 2a and 5b on Figure 2.1. The non-negative profit constraint is found by comparing profits of Regions 3a, 5a and 1 to zero. For instance, Region 1 is bounded from the right by the inequality

\[
\Pi(hb, hb) = 2qx - 2c - 2d - f \geq 0 ,
\]

i.e.,

\[
f \leq 2qx - 2c - 2d .
\]

We will not present here all the calculations but provide sufficient intuition for Figure 2.1.

37. The same is true for the welfare maximizing solution examined in the next subsection.
In Region 1, the firm produces either system \textit{ha} only or system \textit{hb} only so that both types of consumers purchase an identical system with the high-quality hardware, one type of consumers with their more preferred software and the other type with their less preferred software.

In Region 2, two different systems \textit{ha} and \textit{hb} are produced so that both types of consumers purchase their most preferred systems, that is, their preferred software bundled with the high-quality hardware.

In Region 3, the firm produces either system \textit{la} only or system \textit{lb} only. Both types of consumers purchase an identical system with the low-quality hardware, one type with the preferred software and the other type with less preferred software.

In Region 4, systems \textit{la} and \textit{lb} are produced so that both types of consumers purchase their preferred software bundled with the low-quality hardware.

In Region 5, either consumers of type \textit{A} purchase system \textit{ha} and consumers of type \textit{B} purchase system \textit{la}, or consumers of type \textit{A} purchase system \textit{lb} and consumers of type \textit{B} purchase system \textit{hb}. In this case, one type of consumers purchase their most preferred system configuration while the other type of consumers purchase their least preferred system configuration, that is, their less preferred software bundled with the low-quality hardware.

To understand the derivation of the partition intuitively, consider three different intervals for the fixed cost \( f \):

Interval \( L \): \( f \leq (v - 1)2x \)

Interval \( M \): \( (v - 1)2x < f \leq (v - 1)2qx \)

Interval \( H \): \( f > (v - 1)2qx \)
On interval \( L \), software (diversity) separation and hardware (quality) pooling occur. It corresponds to Regions 2c, 2d, and 4 in Figure 2.1. In this case, the development cost of software, \( f \), is small relative to the increase in consumers' reservation prices due to consuming the more preferred software. It is profitable for the firm to produce both types of software and offer each type of consumers a system of their most preferred type since the gains from increases in prices outweigh the costs.

Whether the software is installed on the high- or low-quality hardware depends on how large the hardware cost differential, \( d \), is relative to the increase in consumers' reservation prices associated with the improvement of quality. If it is not too costly to produce the high-quality hardware, as is the case in Regions 2c and 2d, both types of consumers are offered their most preferred system, if the high-quality hardware is so costly that the cost differential can not be compensated with increases in reservation prices (and consequent increases in prices), as is the case in Region 4, both types of consumers are offered the low-quality hardware.

Interval \( M \) corresponds to Regions 2a, 2b, 3b and 5b. In Region 5b, the monopolist faces binding self-selection constraints and has to choose instruments for the separation of consumers.

In the case of very high cost differential as in Region 3b, the firm produces only the low-quality hardware. In the case of low cost differential as in Regions 2a and 2b, only the high-quality hardware is produced. Recall that the reservation prices are multiplicative both in quality and diversity. Because of this preference structure, the high-quality hardware makes the increase in reservation values associated with the preferred software greater than the low-quality hardware does. Thus, in Regions 2a and 2b, two different
types of software are produced (so that both types of consumers get their most preferred systems), while in Region 3b software pooling occurs.

If the cost differential is in the middle range, as is the case in Region 5b (as well as Region 5a which belongs to Interval $H$ explained below), hardware quality is used as an instrument for separation. In this case, one type of consumers get their most preferred system while the other type of consumers get their least preferred system. The two types of consumers are separated because the increase in the value of the system when switching to the high-quality hardware is higher for the consumer type who gets his preferred software than the consumer type who gets his less preferred software.

On interval $H$, software pooling occurs. It corresponds to Regions 1, 3a, and 5a. Consumers’ variety preference is too small relative to the cost of developing software. It is not profitable for the firm to incur the costs of developing two different types of software since the small increase in reservation prices will not compensate for the costs. The type of hardware produced again depends on the cost differential relative to changes in reservation prices associated with quality increase. When the cost increase due to the high-quality hardware is very small, as is the case in Region 1, the firm produces the high-quality hardware only. If the cost increase is very large, as is the case in Region 3a, only the low-quality hardware is produced. In the intermediate case of Region 5a, the monopolist practices price discrimination using hardware quality as a self-selection device.

The result can be summarized as follows: First, when the values of the relevant parameters are extreme, bundling hardware and software does not affect production decisions that would obtain in the absence of hardware-software bundling. As far as hardware is concerned, the high-quality hardware would be produced if the cost differential is suffi-
ciently small, and the low-quality hardware would be produced if the cost differential is sufficiently large. As for software, both types of software would be produced if the cost is sufficiently low, and only one type of software would be produced if the cost is sufficiently high. Second, when the parameter values are in the intermediate range, the firm might use the hardware-software bundling as a price-discriminating device. The hardware-software bundling provides effective discrimination devices, for the firm can offer one type of consumers a bundle that they prefer most and offer another type of consumers a bundle that they prefer least, that is, \((ha, la)\) or \((lb, lh)\).

2.2.2. Welfare Maximization.

In the previous subsection, we assumed that the firm maximizes its profits. We now want to compare the profit-maximizing solution with the welfare-maximizing solution using total surplus, which is the sum of consumer surplus and profit, as the measure of welfare.

Regarding the product choice, recall that a profit maximizer has thrown out some system choices such as \(J(A), J(B)) = (la, ha)\) because there were no prices that would satisfy incentive-compatibility conditions, and has also thrown out system choices including null systems such as \(J(A), J(B)) = (la, \emptyset)\) because they were always dominated by system choices that do not include null systems. It is easy to see that a welfare maximizer would not want to make any of these system choices for similar reasons. Thus a welfare maximizer's system choices are limited to those in Table 2.3.

Regarding the pricing decision, note that for a profit maximizer's prices serve two different functions. One function is to transfer revenues from consumers to the monopo-
list, and the other is to separate or pool the consumer types. A welfare maximizer is indifferent with respect to the transfer of revenues, and thus can use prices freely as devices for consumers' self-selection. Since in our model the prices of the systems intended for the two consumer types are sufficient devices to separate or pool those two types, the first best solution can always be implemented by a welfare maximizer even if the welfare maximizer cannot distinguish between the two types a priori. The question then is how to find the first best solution.

It is well known that the first best solution can be achieved by a monopolist who has the ability to perfectly price discriminate, for such a monopolist can charge the whole reservation prices of consumers, which constitute the total potential welfare, as revenue. Thus in solving the welfare-maximization problem in this subsection, we will consider the analytically equivalent problem of profit maximization without incentive-compatibility constraints (but with participation constraints). This will make the comparison between the profit-maximizing and welfare maximizing solutions more intuitive.

A perfectly price-discriminating monopolist's problem is first to choose systems $J(A), J(B) \in \{ \emptyset, la, ha, lb, hb \}$ and then to choose prices $P_{J(A)}, P_{J(B)}$ to solve the maximization problem

$$\text{Max } P_{J(A)} + P_{J(B)} - C\{J(A), J(B)\}$$

s.t. $U_i(J(i)) - P_{J(i)} \geq 0$ for $i = A, B$.

Denote the maximum profit associated with the above maximization problem by $W(J(A), J(B))$. Then the welfare maximizing product decision is the solution to the maximization problem:
Max $W(J(A), J(B))$

s.t. $J(A), J(B) \in \{\emptyset, la, ha, lb, hb\}$

It is straightforward to calculate welfare for feasible and undominated product choices as we did for the profit maximizing case. Table 2.4 shows the product choices and corresponding welfare.

**Table 2.4.** System choices for welfare maximization.

<table>
<thead>
<tr>
<th>System Type</th>
<th>$(J(A), J(B))$</th>
<th>Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(ha, ha)$</td>
<td>$q(x(v + 1) - 2c - 2d - f)$</td>
</tr>
<tr>
<td></td>
<td>$(hb, hb)$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$(ha, hb)$</td>
<td>$2q(xv - 2c - 2d - 2f)$</td>
</tr>
<tr>
<td>3</td>
<td>$(la, la)$</td>
<td>$x(v + 1) - 2c - f$</td>
</tr>
<tr>
<td></td>
<td>$(lb, lb)$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$(la, lb)$</td>
<td>$2xv - 2c - 2f$</td>
</tr>
<tr>
<td>5</td>
<td>$(ha, la)$</td>
<td>$qv(x + 2c - 2d - f)$</td>
</tr>
<tr>
<td></td>
<td>$(lb, hb)$</td>
<td></td>
</tr>
</tbody>
</table>

The optimal decision will again depend on the cost and preference parameters.
Theorem 2.2. The optimal choices corresponding to different cost and preference parameters (c, d, f, x, q, v) of a welfare maximizing planner are as shown in Figure 2.2.

![Diagram](Image)

As mentioned earlier, welfare maximization leads to the same set of systems produced as profit maximization. The difference lies in the conditions under which particular systems are produced. Notice that although the regions of production are similar, the boundaries of these regions in the cost parameter space have shifted. To facilitate the comparison, the boundaries are shown in Figure 2.2 with solid lines for the welfare maximizing solution and with dashed lines for the profit maximizing solution. The latter was shown before in Figure 2.2.
One implication of Theorem 2.2 is that a profit-maximizing monopolist produces "too many" systems. To see this, notice that Region 1 is smaller under profit maximization. Both of the regions \( R2/1 \) and \( R5/1 \) are regions where a welfare maximizing firm produces a single type of software (either \( a \) or \( b \)) for both types of consumers bundled with the high-quality hardware. In the same regions, a profit maximizing firm will produce either two types of software bundled with high-quality hardware as in the region \( R2/1 \) or two types of hardware bundled with a single type of software as in the region \( R5/1 \). That is, a profit maximizing firm uses either hardware quality or software variety as an instrument for discrimination. As a result, for certain values of cost differential and variety preference parameters, too many hardware types are produced or too many software types are developed by a profit-maximizing monopolist. The same is true for the region \( R4/3 \). While a single type of software would have been produced in the socially optimal case, an additional software type is developed by a monopolistic profit maximizer.

Another implication of Theorem 2.2 is that a profit-maximizing monopolist may produce nothing when a welfare maximizer would produce something, for the boundary of the feasible region for the profit maximizer is to the left of that for the welfare maximizer in Figure 2.2. The reason is obvious: a welfare maximizer would produce as long as the total surplus is positive while a profit maximizer would stop producing as soon as the profit becomes negative.

The two implications of Theorem 2.2 mentioned above are consistent with two phenomena broadly observed in the literature. First, a monopolist often over-provides product variety in an effort to create more instruments of price discrimination. Second, an
unconstrained welfare maximizer may produce something despite losses, which a profit maximizer would not do.

2.3. **Software bundling with symmetric preferences.**

One rigidity of the basic model considered in the previous section was that in the case where two types of software were developed, it did not allow the two software types to be installed simultaneously on a system. In this section, we extend the model by relaxing this constraint.

Let each system consist of one type of hardware bundled with either one or two types of software. The firm can now produce six possible systems: four presented in the basic model \{la, lb, ha, hb\} and two additional systems with software bundles:

i) low-quality hardware with both types of software installed: lab

ii) high-quality hardware with both types of software installed: hab

Note that the last two systems result from the firm’s deliberate decision to combine both types of software in a single package. For the first four systems, bundling one type of software with one type of hardware type was out of technological necessity. Even though there were choices regarding how to bundle, the decision to bundle itself was exogenously determined by the complementary nature of the products.

The marginal costs of the two new bundles, lab and hab, are c and c - d, respectively while the fixed cost of each one is 2f.

Assume that the reservation price of a consumer for a software bundle is the sum of reservation prices for its components. That is, the reservation price for the system hab is \( xq(v - l) \) and that for the system lab is \( x(v - l) \).
Both the profit and welfare maximizers face essentially the same problem as in the basic model of Section 2.2. The process of solving the problem is virtually identical but more tedious, for two new systems are added so that 20 new possible supply decisions should be considered in addition to those already examined in the previous case. We will prove the following in the Appendix:

**Theorem 2.3.** The optimal choices corresponding to different cost and preference parameters \((c, d, f, x, q, v)\) of both a profit maximizer and a welfare maximizer are as shown in Figure 2.3.

![Diagram](image_url)

**Fig. 2.3.** Product choices with software bundling.

The profit and welfare maximizing choices coincide. Only systems with software bundles are produced, and both consumers always get identical bundles. The hardware
type depends on the value of the hardware cost differential. Regions are restricted with the right hand side by the nonnegative-profit constraint.

Thus, if the preferences are symmetric and bundling two types of software with a single type of hardware is allowed, it is always profitable and socially optimal to produce systems with bundles of software.

The reason is that when two types of software are developed, bundling them together within one system does not increase the cost of the system but increases consumers' reservation prices and thus social welfare. Due to the symmetry of the preferences, incentive-compatibility constraints are not binding so a monopolist's profits are also increased. Therefore, while in the model with no software bundles, regions such as 2 and 4 in Figure 2.2 with system choices \((lu, lb)\) and \((hu, hb)\) could be optimal, bundles \((lab, lah)\) and \((hab, hah)\) replace them when the software bundling is allowed.

In summary, using the terminology of the bundling literature, "pure bundling" occurs. The result is interesting, for it has been generally thought in the bundling literature that "mixed bundling" dominates pure bundling from the profit point of view.\(^{38}\)

2.4. Software bundling with asymmetric preferences.

The "pure-bundling" result of the previous section is partly due to the assumption that both consumers have symmetric preferences for software and hardware. The predictions are not sufficiently rich to explain hardware-software bundling practices often observed. In order to generate richer predictions, we will investigate a model with asymmetric preferences in this section.

\(^{38}\) See, for instance, McAfee, McMillan, and Whinston (1989)
Assume that different types of consumers can now have potentially different variety and quality preference parameters. Let $q_i$ be the quality preference of consumer $i$ and $v_i$ the preference of consumer $i$ for one’s own variety ($i = A, B$). Assume $x > 0$, $v_A, v_B > 0$. The reservation prices of the consumers for different systems are now as in Table 2.5.

<table>
<thead>
<tr>
<th>$j$</th>
<th>$U_A(t)$</th>
<th>$U_B(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$la$</td>
<td>$xv_a$</td>
<td>$x$</td>
</tr>
<tr>
<td>$ha$</td>
<td>$xq_a v_a$</td>
<td>$xq_b$</td>
</tr>
<tr>
<td>$lb$</td>
<td>$x$</td>
<td>$xv_b$</td>
</tr>
<tr>
<td>$hb$</td>
<td>$xq_a$</td>
<td>$xq_b v_b$</td>
</tr>
<tr>
<td>$lab$</td>
<td>$x(l - v_a)$</td>
<td>$x(l - v_b)$</td>
</tr>
<tr>
<td>$hab$</td>
<td>$xq_a(1 - v_a)$</td>
<td>$xq_b(1 - v_b)$</td>
</tr>
</tbody>
</table>

Recall that $U_i(j)$ is the reservation price of system $j \in \{\emptyset, la, ha, lb, hb, lab, hab\}$ for consumer $i$. Note that in the case of symmetric preferences, one has $v_a = v_b = v$ and $q_a = q_b = q$.

Assume that consumers of type $A$ value both hardware quality and software type while consumers of type $B$ value only hardware quality and are indifferent between the software types. In order to illustrate this assumption using computer products, think of basic (e.g., Win95 + MS Works) versus extended (e.g., Win95 + MS Office) software packages. Consumers of type $B$ might not need additional software of the extended pack-
age and be equally happy with both packages, while consumers of type $A$ need additional software of the extended package and thus value it more.

Formally, let $v_A = v$, $v_B = l$, $q_A = q_B = q$. Consumers' reservation prices then are as in Table 2.6:

<table>
<thead>
<tr>
<th>$j$</th>
<th>$U_A(j)$</th>
<th>$U_B(j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$la$</td>
<td>$xv$</td>
<td>$x$</td>
</tr>
<tr>
<td>$ha$</td>
<td>$xqv$</td>
<td>$xq$</td>
</tr>
<tr>
<td>$lb$</td>
<td>$x$</td>
<td>$x$</td>
</tr>
<tr>
<td>$hb$</td>
<td>$xq$</td>
<td>$xq$</td>
</tr>
<tr>
<td>$lab$</td>
<td>$x(l - v)$</td>
<td>$2x$</td>
</tr>
<tr>
<td>$hab$</td>
<td>$xq(l - v)$</td>
<td>$2xq$</td>
</tr>
</tbody>
</table>

Assume $v > 2$, that is, the degree of asymmetry in preferences regarding to software is sufficiently high. Call consumers $A$ and $B$ high- and low-preference consumers respectively.
Theorem 2.4. The optimal choices corresponding to different cost and preference parameters \((c, d, f, x, q, v)\) of a profit maximizer and a welfare maximizer are as shown in Figure 2.4.

Fig. 2.4. Product choices with asymmetric preferences.

The figure depicts regions in the "marginal cost differential" - "fixed cost" space. Each region corresponds to a distinct profit- or welfare-maximizing supply decision. The boundaries for a profit-maximizer's regions are shown with dashed lines while those for the welfare maximizer's regions are shown with solid lines.
In Region 1, a profit-maximizing firm sells the bundle $hab$ to consumers of type $A$ while it sells the system $lb$, which contains a single software type $b$, to consumers of type $B$. A welfare maximizing decision is to supply both types of consumers with bundles $hab$ and $lab$, respectively. Region 1 is the region of very low fixed cost of software so that both types of software are developed. The hardware marginal cost differential here is such that the high-preference consumers are offered the high-quality hardware while the low-preference consumers are offered the low-quality hardware.

In comparison, Region 2 is the region of very high marginal cost differential so that only the low-quality hardware is offered. In Region 3, high- and low-quality hardware have similar marginal costs so that both consumers are offered high-quality hardware.

In all of the above regions, the software bundle is under-supplied by a profit-maximizer, for consumers of type $B$, who would have been supplied with the software bundle in the welfare-maximizing solution, are supplied only with one type of software. Adding a second software type to the system sold to consumers of type $B$ increases their consumer surplus. But it leaves the cost unchanged because the fixed R&D cost for both software types has already been incurred and the marginal cost of is zero. Thus a welfare maximizer offers both types of software to both types of consumers. But a profit maximizer does not have the same incentives. Suppose that a profit maximizer offered the software bundle to consumers of type $B$ as well. They can be charged a higher price for a system with $ab$ than for a system with only $b$. However, adding type $a$ software to the second system would also make it more attractive for type $A$ consumers so that the price for her system with $ab$ has to be lowered to satisfy the incentive-compatibility constraint. If consumers of type $A$

39 Given the assumption of $v>2$ and sufficiently small values of cost parameters, all regions are feasible.
have much higher preference regarding software than consumers of type $B$, that is, if preferences are sufficiently asymmetric ($v > 2$), then supplying bundles to both types of consumers will decrease total profits of the monopolist.\footnote{Note that the model now predicts (hab, lb), a supply strategy observed in the example of the Appendix.}

Consider now the intermediate regions $R2/4$, $R1/2$, $R1/5$, $R3/5$. In these regions, a welfare maximizer offers systems with the software bundle to both types of consumers as before: (lab, lab), (hab, lab), (hab, hab). A profit maximizer, however, does not serve consumers of type $B$ at all. Due to the high cost of R&D (that is, large $f$), the monopolist develops only one type of software. The monopolist supplies the single software to the high preference consumers (that is, type $A$) only. The system supplied is $lu$ in regions $R2/4$ and $R1/4$ and $ha$ in regions $R1/5$ and $R3/5$. The choice of hardware to be bundled with software $a$ is dictated by the marginal costs of hardware. If the single software, bundled with some hardware, is to be supplied to both types of consumers, the price has to be lowered sufficiently. When the reservation prices of consumers are significantly different, the monopolist would rather extract full surplus from high-preference consumers than adjust prices to serve both types of consumers. As a result, the monopolist does not offer software bundles. In fact, it does not even develop type $b$ software. Thus the monopolist under-provides variety.

In Regions 4-6, the R&D cost is too high so only one type of software is produced by both the profit and welfare maximizers. The monopolist serves only the high-preference consumer type $A$. A welfare maximizer serves both consumer types.

The result can be summarized as follows: A profit-maximizing monopolist withholds at least one type of software from one consumer type. For some parameter values of
the model, the monopolist may supply both types of software to one type of consumer but withhold one type of software from the other consumer type, resulting in "mixed bundling". For other parameter values, the monopolist may actually avoid developing one type of software altogether and furthermore serve only one type of consumers, supplying nothing to the other type of consumer.

The result that a profit-maximizing monopolist under-bundles and under-produces software variety is interesting. A similar observation has been made in Chae [1992] in the context of bundling channels for subscription television. That is, a profit-maximizing monopolist may engage in mixed bundling of subscription television channels when it would be socially optimal to have pure bundling. Also, a profit maximizer produces less channels than a welfare maximizer. It is worthwhile to note that these statements are valid for the model of this section if one replaces "channels" by "software", even though the model of this section is quite different from that of Chae (1992).

The under-production of software by a profit-maximizing monopolist in the model of this section leads to under-provision of product variety. This contrasts with the result in Section 2.2 that a monopolist over-provides product variety. The paradox can be resolved once we realize that over-provision or under-provision is only the end result. It is perhaps more important to recognize two underlying forces. One is a profit-maximizing monopolist's desire to create more instruments for price discrimination. The other is the tendency of such a monopolist to under-produce software.
2.5. Conclusion.

In this paper, we have investigated software-hardware bundling based on a class of simple models with a monopolistic producer and two consumers. The predictions of the models are consistent with some observed bundling practices.

Interesting results are as follows: Even though bundling hardware and software itself is done out of technological necessity, alternative combinations of hardware and software provide effective devices for price discrimination for a profit-maximizing monopolist. Such a monopolist may over-provide product variety in terms of hardware-software bundles, practice "pure bundling" when preferences are symmetric with respect to software, and under-bundle and under-produce software (and thus under-provide product variety in some cases) when preferences are asymmetric.

Two other variations of the model were studied but not presented in this paper. First, the model of asymmetric preferences regarding software discussed in the previous section was solved for small degrees of preference asymmetry ($v < 2$). Second, asymmetric hardware preferences were investigated. These models, however, did not produce results that are significantly different from those obtained from the basic model of Section 2.2. In both models, the monopolist discriminates by producing too much product variety or quality.
Appendix

Real-Life Bundling Example

A typical example would be two computer systems simultaneously marketed by a major producer/reseller USA-Flex in the 1996 "Computer Shopper", which is one of the largest mail-order catalogs for computer products in the USA. The first system bundles low-speed hardware (486 CPU) with the "life-style collection" of software for home use. The second system bundles 120 Mhz Pentium CPU with the "business collection CD software package" and "Novell Perfect Office Professional Package" in addition to all the software installed on the first system. Using our notation for computer systems, USA-Flex's product choice can be described as \((hab, lb)\).

Proof of Theorem 2.3.

A firm can now produce six different systems \(j \in \{\emptyset, la, ha, lb, hb, lab, hab\}\). Thus there are 24 additional product choices which were not present in Section 2.2:

\[(J(A), J(B)) = (\emptyset, lab), (\emptyset, hab), (hab, \emptyset), (lb, \emptyset), (la, hab), (ha, lab), (ha, hab), (lb, lab), (hb, lab), (hb, hab), (lab, la), (lab, ha), (lab, lb), (lab, hb), (hab, la), (hab, ha), (hab, lb), (hab, hb), (hab, hab), (lab, lab), (hab, lab), (lab, hab).\]

Four of these choices involve selling nothing to one type of consumers. These choices are suboptimal. Note that both consumer types have the same reservation values for the bundles. Thus, if one bundle is sold to one type of consumers, selling an identical bundle to the second type of consumers will always increase profits and welfare. The
remaining 20 choices should be compared to each other and to the choices in the basic model without software bundling described in Table 2.2.

We start with the profit maximizer’s solution. As in the case of Section 2.2, some choices are infeasible and some others are dominated by other systems. These are shown in Tables A.1 and A.2, respectively.

Table A.1. Infeasible system choices.

<table>
<thead>
<tr>
<th>(J(A),J(B))</th>
<th>Incentive compatibility constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>(lb, lab)</td>
<td>$P_{lab} - P_{lb} \geq xv$, $P_{lab} - P_{lb} \leq x$</td>
</tr>
<tr>
<td>(lab, la)</td>
<td>symmetric to (lb, lab)</td>
</tr>
<tr>
<td>(lb, hab)</td>
<td>$P_{hab} - P_{lb} \geq xq(v + 1) - x$, $P_{hab} - P_{lb} \leq xq(v + 1) - xv$</td>
</tr>
<tr>
<td>(hab, la)</td>
<td>symmetric to (lb, hab)</td>
</tr>
<tr>
<td>(hb, lab)</td>
<td>$P_{lab} - P_{hb} \geq x(v + 1) - xq$, $P_{lab} - P_{hb} \leq x(v + 1) - xqv$</td>
</tr>
<tr>
<td>(lab, hu)</td>
<td>symmetric to (hb, lab)</td>
</tr>
<tr>
<td>(hb, hab)</td>
<td>$P_{hab} - P_{hb} \geq xq(v + 1) - xq$, $P_{lab} - P_{hb} \leq x(v + 1) - xqv$</td>
</tr>
<tr>
<td>(hub, hu)</td>
<td>symmetric to (hb, hab)</td>
</tr>
</tbody>
</table>

Table A.2. Dominated system choices.

<table>
<thead>
<tr>
<th>(J(A),J(B))</th>
<th>Profits</th>
<th>Dominated by</th>
</tr>
</thead>
<tbody>
<tr>
<td>(hab, lab)</td>
<td>$x(q + 1)(v + 1) - 2c - d - 2f$</td>
<td>(hab, hab) or (lab, lab)</td>
</tr>
<tr>
<td>(lab, hab)</td>
<td>$x(2v + 1) - 2c - 2f$</td>
<td>(lab, lab)</td>
</tr>
<tr>
<td>(la, lab)</td>
<td>$x(v + 1) - 2c - 2f$</td>
<td>(lab, la)</td>
</tr>
<tr>
<td>(lab, lb)</td>
<td>$xv + xv(q + 1) - 2c - d - 2f$</td>
<td>(hab, hab) or (lab, lab)</td>
</tr>
</tbody>
</table>
Table A.2. Dominated system choices.

<table>
<thead>
<tr>
<th>$(J(A), J(B))$</th>
<th>Profits</th>
<th>Dominated by</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(hab, lb)$</td>
<td>$xqv + x(v + 1) - 2c - d - 2f$</td>
<td>$(la, hab)$ or $(hab, lb)$</td>
</tr>
<tr>
<td>$(lab, hb)$</td>
<td>$2xqv + xq - 2c - 2d - 2f$</td>
<td>$(hab, hab)$</td>
</tr>
<tr>
<td>$(ha, hab)$</td>
<td>$2xqv + xq - 2c - 2d - 2f$</td>
<td>$(hab, hab)$</td>
</tr>
</tbody>
</table>

We thus have only two additional system choices shown in Table A.3 in addition to the five systems in Table 2.3.

Table A.3. Feasible and undominated system choices

<table>
<thead>
<tr>
<th>System Type</th>
<th>$(J(A), J(B))$</th>
<th>$(P_{J(A)}, P_{J(B)})$</th>
<th>Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>$(hab, hab)$</td>
<td>$(xq(v - 1), xq(v - 1))$</td>
<td>$2xqv(v + 1) - 2c - 2d - 2f$</td>
</tr>
<tr>
<td>7</td>
<td>$(lab, lab)$</td>
<td>$(x(v - 1), x(v - 1))$</td>
<td>$2x(v + 1) - 2c - 2f$</td>
</tr>
</tbody>
</table>

We can now compare profits to find the optimal system choice for every region in the "hardware cost differential" -"software fixed cost" space. For example, comparing profits for $(hab, hab)$ with those for $(ha, hb)$, one can conclude that the former always dominates the latter. Similarly, $(lab, lab)$ always dominates $(la, lb)$. It turns out that $(ha, ha)$ or $(hb, hb)$ is preferred to $(hab, hab)$ if $f > 2xqv$. In this case, however, the profits are

$$
\Pi(ha, ha) = \Pi(hb, hb) < -2xq(v - 1) - 2c - 2d < 0,
$$

$$
\Pi(hab, hab) < -2xq(v - 1) - 2c - 2d < 0.
$$
Thus, they are both dominated by \((\emptyset, \emptyset)\). Therefore, \((ha, ha)\) and \((hb, hb)\) are always suboptimal. For small hardware cost differential \(d\), the profit-maximizing monopolist's choice changes from \((hab, hab)\) to \((\emptyset, \emptyset)\) as fixed cost of software \(f\) increases. It is straightforward to compare the remaining system choices to obtain the result of Theorem 2.3.

We now examine the welfare-maximizing solution. First, due to the symmetry of preferences, a welfare maximizer will never serve a single type of the consumers only. Second, if two types of software are developed, the optimal system choices will always be software bundles.

Thus the welfare maximizing solution might include the following systems with two software types only:

\[
(J(A), J(B)) = (hab, hab), (lab, lab), (hab, lab), (lab, hab).
\]

For the profit maximizing problem, these bundles did not have binding incentive-compatibility constraints. Therefore, the welfare-maximizing solution must be equal to the profit-maximizing solution obtained above. In particular, since \((hab, lab)\) and \((lab, hab)\) are dominated in the profit-maximizing case, they are also dominated in the welfare-maximizing case. Combining this with the result of the welfare-maximization problem for the model with no bundles, the following system choices are obtained:
Table A.4. Feasible and undominated system choices.

<table>
<thead>
<tr>
<th>System Type</th>
<th>((J(A), J(B)))</th>
<th>Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((ha, ha))</td>
<td>(qx(v + 1) - 2c - 2d - f)</td>
</tr>
<tr>
<td></td>
<td>((hb, hb))</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>((la, la))</td>
<td>(x(v + 1) - 2c - f)</td>
</tr>
<tr>
<td></td>
<td>((lb, lb))</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>((ha, la))</td>
<td>(qvx + x - 2c - d - f)</td>
</tr>
<tr>
<td></td>
<td>((lb, hh))</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>((hab, hab))</td>
<td>(2xq(v + 1) - 2c - 2d - 2f)</td>
</tr>
<tr>
<td>5</td>
<td>((lab, lab))</td>
<td>(2x(v + 1) - 2c - 2f)</td>
</tr>
</tbody>
</table>

Comparing profits to each other and zero, the partition in Figure 2.3 is obtained.

Proof of Theorem 2.4.

For a profit maximizer, there are 49 possible system choices \((J(A), J(B)) = \)
\((\emptyset, \emptyset), (\emptyset, la), (\emptyset, ha), (\emptyset, lb), (\emptyset, hb), (\emptyset, \emptyset), (ha, \emptyset), (lb, \emptyset), (hb, \emptyset), (la, la), (la, ha), (la, lb), (la, hb), (la, \emptyset), (ha, ha), (ha, lb), (ha, hb), (hb, ha), (hb, lb), (hb, hb), (lb, la), (lb, ha), (lb, hb), (hb, la), (hb, ha), (hb, lb), (hb, hb), (\emptyset, lab), (\emptyset, hab), (hab, \emptyset), (lab, \emptyset), (la, lab), (la, hab), (la, ha), (la, lb), (lab, ha), (lab, hb), (lab, la), (hab, ha), (hab, lb), (hab, hb), (hab, hab), (hab, lab), (hab, hab), (lab, lab), (hab, lab), (lab, hab).\)

For a welfare maximizer, some system choices can be excluded. It is not optimal from the welfare standpoint to serve only low-preference consumers. Also, it is subopti-
mal not to bundle software when two types of software are developed. A welfare maximizer therefore chooses from 19 system choices: \((J(A), J(B)) = (\emptyset, \emptyset), (la, \emptyset), (hu, \emptyset), (lb, \emptyset), (hb, \emptyset), (la, ha), (la, la), (ha, ha), (lb, lb), (lb, hb), (hb, lb), (hb, hb), (hab, \emptyset), (lab, \emptyset), (hab, hab), (lab, lab), (hab, lab), (lab, hab)\).

It is easy to verify that the profits and welfare for feasible and undominated system choices are as shown in Tables A.5 and A.6. The product choices depicted in Figure 2.4 then follows from comparisons between profits or welfare expressions in these tables.

<table>
<thead>
<tr>
<th>System Type</th>
<th>((J(A), J(B)))</th>
<th>((P_{J(A)}, P_{J(B)}))</th>
<th>Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(lab, lb)</td>
<td>((x(v - l), x))</td>
<td>(x(v + 2) - 2c - 2f)</td>
</tr>
<tr>
<td>2</td>
<td>(hab, lb)</td>
<td>((xq(v - l), x))</td>
<td>(xq(v + 1) + x - 2c - d - 2f)</td>
</tr>
<tr>
<td>3</td>
<td>(hab, hb)</td>
<td>((xq(v - l), xq))</td>
<td>(xq(v + 1) + xq - 2c - 2d - 2f)</td>
</tr>
<tr>
<td>4</td>
<td>(la, \emptyset)</td>
<td>((xv, \emptyset))</td>
<td>(xv - c - f)</td>
</tr>
<tr>
<td>5</td>
<td>(hu, \emptyset)</td>
<td>((xqv, \emptyset))</td>
<td>(xqv - c - d - f)</td>
</tr>
</tbody>
</table>

Table A.5. Feasible and undominated system choices for profit maximizer.

<table>
<thead>
<tr>
<th>System Type</th>
<th>((J(A), J(B)))</th>
<th>Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(lab, lab)</td>
<td>(x(3 + v) - 2c - 2f)</td>
</tr>
<tr>
<td>2</td>
<td>(hab, lab)</td>
<td>(xq(1 + v) + 2x - 2c - d - 2f)</td>
</tr>
</tbody>
</table>

Table A.6. Feasible and undominated system choices for welfare maximizer.
<table>
<thead>
<tr>
<th>System Type</th>
<th>$(J(A), J(B))$</th>
<th>Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$(hab, hab)$</td>
<td>$xq(v + 3) - 2c - 2d - 2f$</td>
</tr>
<tr>
<td>4</td>
<td>$(la, la)$</td>
<td>$x(v + 1) - 2c - f$</td>
</tr>
<tr>
<td>5</td>
<td>$(hu, la)$</td>
<td>$x(qv + 1) - 2c - d - f$</td>
</tr>
<tr>
<td>6</td>
<td>$(ha, hu)$</td>
<td>$xq(v + 1) - 2c - 2d - f$</td>
</tr>
</tbody>
</table>
References.


Blank, Rebecca M. (1997), It Takes a Nation, Princeton University, Princeton N.J.


