ESSAYS IN PUBLIC ECONOMICS

by

GIUSEPPE MIGALI

DOCTOR OF PHILOSOPHY IN ECONOMICS

Department of Economics

University of Warwick

– June 2008 –

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am really grateful to my supervisors Ben Lockwood and Ian Walker for clear comments, helpful suggestions and all the support during my years of PhD. I also want to thank my PhD colleagues for useful discussions, the staff members of the Department of Economics that attended to my presentations providing helpful comments, and all those that gave me fruitful advices in conferences and workshops where I presented my works. Finally, all eventual errors and omissions are my own.

ABSTRACT

This thesis is structured in two parts: the first part investigates topics in economics of education and the second part topics in commodity tax competition. The common thread in the first part is the English 2004 HE reform,we present both theoretical and empirical models that compare the pre-reform and post reform financing schemes, i.e. income contingent loan and mortgage loan. In the first chapter, in a world where graduates are risk averse and receive uncertain earnings, we evaluate which method gives better insurance and returns when the investment in education is risky. In the second chapter we consider the possibility that graduates do not receive an income sufficient to repay completely their loan, that is they are in default. We still develop a theoretical model, structured as a game between government and students, to see whether an ICL or a ML give students higher incentives to put more effort in their studies and avoid the risk of default. In the third chapter we present an empirical analysis, in line with the recent studies on the return to schooling. We first estimate the returns for different levels of education and then we compute the internal rate of return of the investment in higher education versus high school, under a ML and an ICL. In the second part of the thesis we provide an empirical model to develop the key idea that completion of the Single Market in the EU can be interpreted as a kind of "natural experiment" that allows us to separate the effects of tax competition from other forms of strategic interaction. We first find the conditions of cross-border shopping and the government reaction functions in three different market regimes: duty-free market before 1993, mixed market up to 1999, single market after 1999. Second, we provide an empirical test of the theoretical predictions using a panel data set of 12 EU countries over the period 1987-2004. We find that for all excise duties that we consider (beer, ethyl alcohol, still wine, sparkling wine, cigarettes, tobacco, petrol), strategic interaction between countries significantly increased after 1993, consistent with the theoretical prediction.

INTRODUCTION

0.1 Introduction

This thesis is structured in two parts, the first part develops a topic in economics of education, the second part is an empirical analysis in horizontal commodity tax competition.

First Part

Some unanswered questions in the economics of education concern the causal effects of schooling on earnings, the optimal level of education, the effect of the liquidity constraints on the individual schooling choice, the effect of parental income and/or parental education on children educational attainment, the importance of the public intervention, and the best methods to finance higher education in order to increase schooling participation.

Much of the traditional literature focuses on these questions from an empirical point of view. For example in many studies the positive correlation between family income (or other family characteristics) and schooling attainment has been widely interpreted as evidence supporting the idea that borrowing constraints hinder educational choices. However, this is a questionable conclusion since family income is also strongly correlated with secondary school achievement and other characteristics associated with college entry. Recent contributions have attempted to better understand the determinants of schooling choices. Using very different methods these studies have found little evidence that favour the idea that borrowing constraints hinder college-going, or any other schooling choice. For example, one test has been based on the idea that when borrowing constraints are present, the opportunity costs of schooling (foregone earnings) and direct costs (such as tuition) affect schooling choices differently for credit-constrained and unconstrained individuals.

These issues are directly related to the problem of financing higher education. It is possible to distinguish two main categories: systems without fees for students (completely financed by the government) and systems with fees (that can be either up-front or delayed, subsidized or not). If students have to contribute to the cost of their education, the problem of financial assistance may arise. There are various forms of support for students, but one that has stimulated most interest among economists is student loans. The most common schemes are mortgage loans and income contingent loans. Chapman (2007) offers a comprehensive summary on the large literature on the background to the new HE reforms based on ICL in the world. Our purpose in this work is to contribute to this literature by studying publicly funded higher education based on two alternative loan schemes: income contingent loan and mortgage loan. The common thread of the first part of the thesis is the English 2004 HE reform which is our reference point both in the theoretical and the empirical analysis.

Up to 1998 there were no tuition fees for English students, their living expenses were covered by a mixture of a tax-funded grant, a loan with mortgage-type repayments (introduced in 1990), and parental contributions. In 1998 they introduced an upfront fee, there was no loan to cover the fee and living expenses were met by a mixture of parental contributions and income-contingent loan (ICL), which replaced the old mortgage loan. The recent Higher Education reform increased the tuition fee and introduced an income-contingent scheme to cover them. Graduates started to repay their loan according to the level of their earnings after graduation.

In the first chapter we compare a ML and an ICL in a world where graduates are risk averse and receive uncertain earnings. We want to evaluate which method gives better insurance and returns when the investment in education is risky. In the second chapter we consider the possibility that graduates do not receive an income sufficient to completely repay their loan, that is they are in default. We develop a theoretical model, structured as a game between government and students, to see whether an ICL or a ML give students higher incentives to put more effort into their studies and avoid the risk of default. In the third chapter we present empirical estimates of the return to qualifications that we use to try to put some empirical flesh on the theoretical bones. We first estimate the returns for different qualifications and then we compute the internal rate of return of the investment in higher education versus high school.

Chapter 1

A question that is under-researched in the literature is whether investing in education is risky. There is widespread evidence of a large variance in the returns to education and some evidence to suggest that much of this is due to risk uncertainty rather than individual heterogeneity (Hartog and Serrano, 2007 and Chen, 2001). Moreover, an individual making schooling decisions is likely to be only imperfectly aware of her abilities, the probability to success, the job that may be obtained after completing an education, and the position within the post-school earnings distribution that may be attained. In our work we contribute to the existing literature in the following way. We consider both the riskiness of the investment in higher education and the uncertainty in its returns. However we face these problems from the point of view of the optimal higher education financing systems for the students. We want to see whether the recent reforms in the funding of higher education have affected the returns to schooling of risk averse individuals receiving stochastic earnings after graduation. Uncertainty is represented by the variance of the post schooling wage, in line with the literature, and we compare the individuals outcomes under two possible funding schemes based on student loans: mortgage loan and income contingent loan. Our work provides theoretical evidence on the effect of an ICL, and we are the first to provide an analysis that compares income contingent loans and mortgage loans for graduates, taking into consideration one of the main drawbacks of an ICL: that is the hidden subsidy, which arises when the expected repayment periods between the two loan schemes are different.

In our simulation results we find that people from low educated parental background, males over females, and people working in the private sector receive higher utility from an ICL. In contrast a ML is the preferred system when the age earnings profiles are not steep and income is not too high, which is typical of public sector work. Finally we show that an ICL provides a better insurance against uncertainty for low wage earners.

Chapter 2

In this chapter we try answer a common problem in HE financing systems that adopt student loans, that is the possibility that some graduates may not completely repay their debt so then default. We address this problem from a theoretical point of view, presenting a model which is structured as a simple simultaneous game between government and students. The students choose between high and low effort, and the government decides the default premium that balances its budget. A high level of effort imposes a psychological cost on the students. We consider risk neutral individuals and we derive the analytical conditions for the equilibria under a mortgage loan and an income contingent loan, assuming a participation constraint. We finally compare the different solutions using numerical simulations. In general, we find that under an ICL the participation constraint is less binding, that is students prefer putting in more effort to stay in higher education than under a mortgage loan. Moreover, under an ICL the students are, overall, more willing to incur a higher cost of effort so as to avoid the risk of default.

Chapter 3

In this chapter we have two aims, we first estimate the causal effect of education on earnings. Second, using those estimated returns we generate the working life wage profiles for those that have a high school (HS) qualification and those with a degree equivalent qualification. The objective is to compute the internal rate of return (IRR) of the investment in HE versus HS, under different HE financing schemes, i.e. free education, mortgage loan and income contingent loan. The distinctive features of our empirical estimation is that we control for selectivity and our model does not impose the restriction that the age earnings profiles are the same for each level of education. So it does not impose separability in the earnings function between the schooling effect and the age effect. We correct for unobserved ability using the raising of the school leaving age 1973 England reform and the individual's month of birth as exclusion restrictions. These are exogenous variables that affect directly education and not earnings, and we find estimates consistent with the existing literature. We then use these estimates to evaluate the decision to go to university as a financial investment, and we compute the net present value of the additional earnings for HE graduates compared to HS graduate. Following other work in the field we find the internal rate of return. We consider the cost of the investment both as the tuition and the opportunity costs of going to school, i.e. the foregone earnings. Our further contribution is to compare the IRR for an ICL and a ML, according to the features of the English HE systems pre and post 2004 reform. We find in general that the IRR for HE graduates is higher under an ICL relative to a ML: under an ICL they exploit a larger taxpayer subsidy due to the zero real interest on the loan.

Second Part

In Chapter 4 we analyze a separate topic in Public Economics. We provide an empirical analysis of horizontal competition in excise taxes through the estimation of equations informed by the theory of strategic interactions among governments.

The introduction of the Single Market, in 1993 in the European Union, resulted in a switch from destination to origin-based taxation of cross-border transactions by individuals. Standard theory predicts that this change should intensify excise tax competition and thus strategic interaction in excise taxes overall. In this work, we first extend the basic model of Nielsen (2001), finding the conditions of cross-border shopping and the government reaction functions in three different market regimes: duty-free market before 1993, mixed market up to 1999, single market after 1999. Second, we test whether the creation of the EU single market has generated a structural break giving impulse to higher strategic interaction among countries.

The key idea of the chapter is that completion of the single market can be interpreted as a kind of "natural experiment" that allows us to separate the effects of tax competition from other forms of strategic interaction. This work represents the first attempt, to our knowledge, to directly test this observable prediction. We employ a balanced panel data set of 12 EU countries over the period 1987-2004 and excise taxes on seven commodities (beer, ethyl alcohol, still wine, sparkling wine , cigarettes, tobacco, petrol). While the control variables are taken from standard sources, the excise tax data were taken from the European Commission's Excise Duty Tables and Inventory of Taxes. Using this data set, we estimate an empirical model where the excise tax in a given country depends linearly on the weighted average of other countries' taxes, and a set of control variables. The equation is estimated separately on subsamples before and after 1993.

We find for all excise duties that we consider, a structural break in the tax setting and a significant increase in the strategic interaction between countries after 1993, consistent with the theoretical prediction.

CHAPTER 1

1. FUNDING HIGHER EDUCATION UNDER WAGE UNCERTAINTY: INCOME CONTINGENT LOAN VERSUS MORTGAGE LOAN

1.1 Introduction

Investment in education is risky: an individual making schooling decisions is likely to be only imperfectly aware of her abilities, the probability of success, and earnings that may be obtained after completing an education. In our work we consider the uncertainty in the returns to higher education that leads to be a risky investment.

However, part of the motivation for this work comes from a change in the English higher education funding system. We now briefly describe the most important Higher Education reforms in England. Before 1990 there were no tuition fees for English students, their living expenses were covered by a tax-funded grant, which was means-tested against parental income, and expectations of parental contributions. In 1990 a loan with mortgage-type repayments was introduced, to be repaid in a fixed number of monthly instalments[1] - normally 60. Individuals started to repay their debt after they graduated, however, repayments could have been deferred on a yearly basis if the repayer's income was below 85% of the national average. In 1998 they introduced an upfront fee (£1,125 in 2003-4), irrespective of subject or university. There was no loan to cover the fee and living expenses were met by a mixture of parental contributions and income-contingent loan (ICL), which replaced the previous mortgage loan. Graduates started to pay back their educational loan only when their earnings were above $\pounds 15,000$ per year and at a 9 per cent fixed repayment rate. There was a zero real interest rate and repayments were made through the tax system as a payroll deduction. The recent Higher Education reform (approved in 2004 and effective from 2006/2007) increased the tuition fee and universities were able to set their fees up to a maximum £3000 p.a. The major innovation was the extension of the income-contingent scheme, with the same repayment rules, to also cover

[1] Source: Student Loans Company Limited UK.

the tuition fee. A similar higher education system has been effective in Australia since 1989, the only difference has been a progressive repayment schedule as opposed to the linear one in England.

Looking at the existing literature on risk and uncertainty in education: Weiss (1974) studies the risk adjusted average rate of return to schooling, which is the subjective discount rate at which the individual would be indifferent between acquiring a certain level of education and having no education at all. Weiss finds that this risk adjusted return sharply decreases as risk aversion increases. Olson, White and Shefrin (1979) follow the traditional literature focusing on the returns to education but incorporate graduate earnings uncertainty, and takes into account the way higher education is funded. They assume that consumption equals earnings in each period after schooling and educated individuals get a random stream of earnings whose mean varies according to the level of education achieved. Olson, White and Shefrin allow borrowing to finance education, and in particular they consider a mortgage loan that is paid back only after the completion of schooling. They find that the estimated real returns of college are large, and the estimated risk premium for attending college is small but positive.

In the recent literature fully integrating risk into the analysis of investment in human capital where borrowing is possible has remained technically problematic. Pistaferri and Padulla (2001) extend the Olson, White and Shefrin's model to consider two types of risk: employment risk and wage uncertainty, within an imperfect credit market framework. Hogan and Walker (2007) use real options theory to model optimal schooling choices, when the returns are stochastic. They find the opposite result, that individuals stay in school longer as risk increases. The crucial difference lies in the ease with which lifelong learning is possible - being able to return to education matters. Jacobs (2007) also uses real options to analyze the possible consequences in human capital investments, showing that when students can influence the timing of their HE investment, option values will affect investment behaviour. Individuals have, in fact, the option to wait and gather better information regarding the return of their investment. If they invest immediately in HE they give up this option, loosing its value. Therefore, higher returns to HE are the proper compensation for the lost option value of waiting once individuals invest irreversibly in

HE. Hartog and Serrano (2007) develop a human capital model to analyze the effects of stochastic post schooling earnings on the optimal schooling length. They show the key role of risk parameters and risk attitudes. Using Spanish data they estimate the sensitivity of schooling decisions to variance in post-schooling earnings. They use the predicted earnings, from the estimation of the earnings equation of individuals with given levels of education in each region, to compute the expected returns and the residual variances which are interpreted as a measure of the risk associated with an education. They typically find a statistically significant negative effect of risk on investment in HE education, in the case of some estimates they present, the risk premium greatly reduces the return to schooling. Hartog and Vijverberg (2007), starting from the estimation of a standard Mincerian wage equation, group the residuals by their education-occupation classification. They take the within cell residual dispersion as a useful measure of the uncertainty in the returns to schooling. In practice, adding the variance of the residuals, obtained through this grouping, to the wage equation, they find higher wages for higher residual variance and lower wages where the distribution is more right-skewed. They interpret these results as evidence of the attitude to risk: workers exhibit risk aversion (dislike risk) and have skewness affection. Belzil and Leonardi (2007) study whether risk aversion can explain differences in schooling attainments. They use a sample of Italian individuals and they implement a discrete duration model technique to analyze how grade transition from one level to the next varies with measured risk aversion. They find that schooling decisions are mainly affected by parental education differences rather than differences in risk attitudes.

In our work we contribute to the existing literature in a different way. We consider the riskiness of investment in higher education induced by uncertainty in its returns, and investigate the effect of alternative higher education financing systems for students. We wish to see whether the recent reforms in the funding of higher education can affect the returns to schooling of risk averse individuals receiving stochastic earnings after graduation. Uncertainty is represented by the variance of the post schooling earnings, in line with the literature, but we compare the individuals outcomes under two possible funding schemes based on student loans: mortgage loan and income contingent loan. The concept of ICL's as a means to fund human capital investments started with Friedman (1955) and since then several works such as Nerlove (1975), Barr (1993), Chapman (2005) have analyzed the usefulness of ICLs as a source of funding HE. Chapman (2007) offers a comprehensive summary of the large literature concerned with HE funding reforms.

Our work adds theoretical evidence on the effects of ICLs, to provide an analysis that compares income contingent loans and mortgage loans for graduates, taking into consideration one of the main drawbacks of an ICL, that is the hidden subsidy. The latter arises because of positive discounting when the repayment periods between the two schemes are different.

Starting from the same theoretical framework as Hartog and Serrano (2007), we present a model in which graduates receive uncertain future earnings (affected by a single lifetime shock), and we measure the level of uncertainty by considering the variance of earnings. We complement Hartog and Serrano who estimate the riskiness of the educational investment, by comparing two HE financing schemes: an income contingent loan and a mortgage loan. We assume that students receive a loan from the government to finance the cost of HE, and they repay their debt after graduation. We consider only individuals in full-time higher education, with no earnings during schooling and facing a given cost of HE. The government is risk neutral and has no preference for one system of loan scheme over the other. For the moment we assume no default, for simplicity, and we assume a zero interest rate for both student and government. We do not consider any external effects of education on society as a whole. We analyze which loan scheme yields greatest welfare in terms of individual lifetime expected utility, for risk neutral and risk averse graduates. Our intuition is that if graduates expect some high variance in their earnings, an income contingent loan that allows them to repay the debt only when they have the financial resources to do it provides higher welfare, because it provides insurance against uncertainty. We also assume non graduates earnings are certain and there is no unemployment.

Our interest in the combination of earnings uncertainty and student loan design is motivated by two things: the observed uncertainty in the real earnings of graduates; and the reform of the higher education financing system in England, with a mortgage loan (ML) until 1998 and its subsequent replacement by an income contingent loan scheme (ICL).

We verify our theoretical intuition by calibrating the model using real data on graduate earnings and their volatility, combined with the features of the English financing system.

In this way we simulate different scenarios where we compare the two loan schemes, observing the changes in individual welfare. We use the British Cohort Study 1970, which although restricting the sample only on individuals borne in 1970, provides useful information on family background. Assuming no selectivity, and following the assumptions of the model, we consider only the earnings post graduation and its standard deviation given the individual characteristics (e.g. sex, marital status), the family background (e.g. parental earnings, parental occupation), and the degree choice (e.g. subjects). For illustrative purposes we use parameters taken from the past and current loan schemes of the English higher education.

There are two important differences between an ICL and a ML: the first is that a ML is not providing insurance against uncertainty and the second is the presence, under a ICL, of a potential hidden subsidy, which arises because the average repayment period could be different to a ML. The first difference does not matter if the individuals are risk neutral, while the second only matters because of discounting. We isolate the second effect by assuming risk neutral individuals. We notice that the hidden subsidy typically makes an ICL more attractive. If we rule out this possibility, the expected costs under an ICL are always higher than the costs under a ML. The effect of the English reform is to increase this expected differential in cost. In terms of utility, a ML provides higher expected utility for risk neutral graduates, and in particular for those on low earnings, e.g. female. In the case of risk aversion the model offers interesting policy implications. We consider the case whether there is no hidden subsidy, by fixing the expected repayment period to be the same under an ICL as it would be under a ML, and we notice that the effect of high uncertainty combined with risk aversion makes an ICL the preferred system.

In our simulations we typically find that people from low educated parental background, males over females, people working in the private sector receive higher utility from an ICL, because of higher variance in their earnings. A ML is the preferred system when earnings exhibit low variance, and when earnings are not too high: something which is typical of the public sector careers.

In the second part of our work, this model is extended to incorporate stochastic changes of earnings over time. In particular, we assume that the growth rate of the earnings follows a Brownian motion, as in Hogan and Walker (2007). However, they solve the problem of education choice with uncertain returns using an options approach. Instead we compute the expected utilities under the two loan schemes using a numerical method. The new framework allows us to generate earnings paths for the entire individual working life, where uncertainty affects the earnings each year. We again find, as in the previous part, that higher uncertainty increases the utility of an ICL. However, the results in this case show the importance of the initial earnings: in fact an ICL is highly preferred by individuals with low initial earnings.

In the next Section we describe the theoretical model, in Section 1.3 we analyze the case of risk neutrality, in Section 1.5 we obtain the algebraic form of the expected utilities under risk aversion, while in Section 1.6 we describe the dataset and show the results of the simulations. In Section 1.7 we set up the model with the new assumptions on the earnings and show the results of the simulations. Section 1.8 concludes.

1.2 Theoretical Model

This section presents the main assumptions of the theoretical model that hold both under a mortgage loan and under an income contingent loan. We distinguish between the schooling period and post graduation period.

1.2.1 Schooling period

Individuals go to university for s years full time and education has the same cost for everybody without distinction between courses and subjects. Earnings during the schooling are assumed to be zero. Following Olson, White and Shefrin (1979) consumption is always equal to earnings. Therefore, during university, consumption is also set to zero. There is no private market for loans, and no informal market for loans e.g from parents. There is no insurance market because it is not profitable for the private market to insure the investment in HE, since moral hazard and adverse selection cannot be avoided due to the lack of any collateral.

The only source of financing allowed is a public loan, again equal for all the students, of fixed size and that covers all the costs of attending university. The real interest rate on the loan is zero[2]. For simplicity, consumption at school is zero and equals earnings. The loan finances fees - although we could allow it to finance consumption during schooling. Thus the loan size is the same irrespective of scheme and equal to fees. The government finances a constant cost of higher education, through issuing the same amount of debt regardless of repayment scheme (therefore the subsidy is the same for all the students). The debt is repaid only by the graduates' repayments, there is no opting out choice: when the students enter a scheme he or she cannot leave it before the full repayment of the loan. This implies that there is no default[3], and in the long run all the cost of education is recovered equally by both schemes. The government is risk neutral and does not have any preference over the funding systems. Since the government could bear different costs of providing the loan according to the scheme, we assume a zero real interest rate on the borrowing[4]. Under this assumption the costs for the government are the same under an ICL and a ML, and social welfare depends only on students' utility.

Moreover, when the expected repayment periods under the two systems differ, hidden subsidies arise. The longer is the repayment period, the lower is the present value of the repayments, given a positive subjective discount rate. Therefore the system with the longer repayment period provides a larger hidden subsidy to graduates and it is more attractive than the scheme with shorter repayments. We mainly perform the analysis under the condition of no hidden subsidy, but in one case we also allow for hidden subsidies to establish what difference this makes.

1.2.2 Post graduation period

Upon graduation, the individuals start working immediately, and, as in Hartog and Serrano (2007), they obtain an uncertain wage because it is subject to a random shock. For simplicity, the shock has a single lifetime realization, after which earnings remain constant. We can imagine an initial random draw that fixes earnings at a certain level which remains unchanged for all the working life. Let $\tilde{y} > 0$ be the shock with $\mathrm{E}(\tilde{y}) = 1$ and $Var(\tilde{y}) = \sigma^2$.

Individuals cannot insure the wage risk and seek to maximize the expected lifetime

[2] This is not just a simplifying assumption, but a real feature of the income contingent loans as implemented by the national governments in England and Australia. There is only an adjustment to the inflation. In our model, for fair comparison, we assume zero real interest also for a mortgage loan.

[3] The case of the students' default is analyzed in Chapter 2.

[4] The case with positive real interest on the borrowing is also analyzed in this Chapter, but only for risk neutral individuals

utilities. Consumption is equal to earnings and is strictly positive; utility is defined over the individuals' earnings stream.

In this model we focus only on the post graduation period, but it is important in all the following analysis to distinguish between the repayment period and the post repayment period. Graduates must start repaying their educational loan straight after graduation and for T years, after that they receive their entire earnings for the rest of their life, assumed infinite for simplicity. Considering a general repayment scheme, we define R as the general per-period payment. The expected utility is:

$$V = \mathrm{E}\left\{ \int_{s}^{T+s} e^{-\rho t} u\left(\tilde{y} - R\right) dt + \int_{T+s}^{\infty} e^{-\rho t} u\left(\tilde{y}\right) dt \right\} \tag{1.1}$$

where $R < \tilde{y}$, and ρ is the subjective discount rate that measures how much the present is taken in consideration with respect to the future. A loan scheme is described fully by (T, R).

1.2.3 Mortgage Loan and Income Contingent Loan

We stress that the cost of the loan is equal to the total cost of education C, and it is the same under both systems. The way it is repaid produces different individual utilities because of the random earnings. If we assume no uncertainty and identical repayment rates the two systems are equal in terms of utility.

Mortgage Loan

The individuals take out a loan equal to C and repay through T equal, fixed and periodical installments φ, at a zero real interest rate. The repayment period is just

$$T = C/\varphi. \tag{1.2}$$

Income Contingent Loan

The individuals borrow an amount equal to C, and start to pay back their loan after graduation according to the level of their earnings. Under this scheme, if earnings are below a minimum threshold no payment is due. If earnings increase, a greater portion of the debt is repaid and all the loan is paid off in less time. Therefore, compared to a mortgage loan the ICL repayment period, $\tilde{T}$, is *random.* In our model, for simplicity, we assume no initial threshold and the total cost of schooling is given by a fixed percentage (γ) of the random graduate earnings. Note, γ is chosen ex ante by government, the repayment period $\tilde{T}$ is determined when the income shock is realized.

$$C = \gamma \int_{s}^{\tilde{T}+s} \tilde{y}\, dt \tag{1.3}$$

therefore the repayment period is:

$$\tilde{T} = \frac{C}{\gamma \tilde{y}} \quad . \tag{1.4}$$

We can now define the following assumption concerning the expected repayments under the two schemes.

Assumption 1. *With a ML,*

$$R < \tilde{y} \quad \Longleftrightarrow \quad \varphi < \tilde{y}$$

and the expected repayment is

$$\mathrm{E}[P_{ML}] = T \times \varphi = C$$

Assumption 2. *Under an ICL the annual repayment is:*
$\gamma \times \tilde{y}$ *until the loan is repaid, and the expected repayment is*

$$\mathrm{E}[P_{ICL}] = T\gamma\tilde{y} = C$$

Proposition 1. *Expected repayment periods under the two systems*

$$\begin{array}{lll} \text{(a)} & if\ \ \gamma = \varphi & \mathrm{E}[T_{ICL}] > T_{ML} \\ \text{(b)} & if\ \ \gamma = \varphi \times \mathrm{E}[\frac{1}{\tilde{y}}] & \mathrm{E}[T_{ICL}] = T_{ML} \end{array}$$

where $\mathrm{E}[\frac{1}{\tilde{y}}] > 1$

Proof. *See Appendix 1.9*

The first proposition highlights one of the main differences between the two schemes: a feature of an ICL is to spread the same cost over a longer repayment period compared to a ML, assuming the same repayment rates. However, this implies the existence of a hidden subsidy because the longer repayment period under ICL compared to ML because of positive discounting and the real interest is zero[5]. Therefore, in point (b) of Proposition 1, we find the condition that rules out the hidden subsidy and allows a comparison of the two schemes on the same basis. In particular, we notice that to have same repayment periods we require a higher repayment rate under an ICL.

1.3 Risk Neutrality and Expected Costs

When the individuals are risk neutral $u(\tilde{y}) = \tilde{y}$, and we need only consider the costs to compare the two repayment schemes. We work out the present value of the costs, substituting for each scheme the respective repayment period, T and $\tilde{T}$, and discount to $t = 0$.

Proposition 2. *Assuming risk neutral individuals,*

$$V_{ICL} > V_{ML}$$

when $\gamma = \varphi$, *instead*

$$V_{ICL} < V_{ML}$$

when $\varphi = \frac{\gamma}{\mathrm{E}[1/\tilde{y}]}$

Proof. See Appendix 1.9.1

In the Appendix 1.9.1, we prove analytically Proposition 2 when $\gamma = \varphi$. Then, we argue that for the case $\varphi = \frac{\gamma}{\mathrm{E}[1/\tilde{y}]}$ we require a numerical solution[6] which we provide using real data from our BCS70 dataset.

The result highlights, in terms of expected utilities, the differences between the two systems raised in Proposition 1. The presence of the hidden subsidies completely changes

[5] The subsidy still exists providing $r < \delta$.

[6] This requirement is clear looking at equations 1.25 and 1.26 that need to be compared in order to evaluate which system produces higher expected utility. Since we are considering risk neutral individuals we compare only the expected value of the costs, as argued in the Appendix 1.9.1

the preferences and makes an ICL more preferred. Both if we assume the same repayment periods, instead, we would need to increase the ICL repayment rate and then a ML gives higher welfare.

To give a broad intuition of the result in Proposition 2, we assume a general repayment method R, the present value of the education cost is:

$$PVC = \int_0^T R e^{-\rho t} dt = \frac{R}{\rho}[1 - e^{-\rho T}]$$

Taking the derivatives of PVC with respect to T, we can easily observe that this function is concave[7]. Consider now a loan with a certain repayment period of $T_1 = 10$ years, and an alternative loan with two equal probability repayment periods of 5 and 15 years, T_2; therefore we have $T_1 = ET_2$, which rules out the hidden subsidy, at least in the risk neutral case. The concavity property implies that the expected present value of the cost of a certain repayment period is lower than the present value of the cost of an uncertain repayment period of the same expected length (even though we are risk neutral):

$$\mathrm{E}[PVC(T_1)] < PVC[\mathrm{E}(T_2)] \quad \Longrightarrow \mathrm{E}[PVC(10)] < PVC(10).$$

We now analyze the second result of Proposition 2, through a numerical calibration. We first estimate $\mathrm{E}[1/\tilde{y}]$ by its sample analogue, then fix γ and get φ accordingly. The method is explained in detail in the Appendix 1.9.1, and we report the results of the simulations in the next section.

1.3.1 *Equal repayment periods*

We compute the expected costs under a ML and under an ICL based in data for all the graduates, and then distinguishing between males and females. After the British Higher Education reform the annual cost of education was set up to a max of £3000 pounds, while before the reform the cost was £1150 a year[8]. Assuming a 3-year degree, we consider two levels of total cost: £3450 and £9000. We consider three possible levels of the

[7] $\frac{\vartheta^2 PVC}{\vartheta T^2} = -R\rho e^{-\rho T} < 0$ for all T.
[8] In practice all institutions have charged the maximum

subjective discount $\rho = [0.08 \quad 0.15 \quad 0.3]$ and three levels of the ICL's repayment rate $\gamma = [0.02 \quad 0.09 \quad 0.2]$.

Looking at the top of Figure 1.1, we report on the vertical axes the difference between the expected costs $EC_{ICL} - EC_{ML}$. We refer to $\tilde{y}$ as the earnings of all graduates, and ym and yw are the earnings of male and female graduates respectively. Setting $\gamma = 9\%$, we observe that the expected costs under an ICL are always higher than the expected costs under a ML, and the effect of the reform is to increase this gap. This means that when there are no hidden subsidies the benefits of an ICL for risk neutral graduates decrease strongly. Moreover, although the repayment periods are equal between the two systems, they differ across graduates. For all graduates, the repayment period increases from 5.6 to 14.7 years, for male from 5.1 to 13.3 years and for females from 5.7 to 14.9 years. The males are those with the highest earnings and therefore the shortest repayment period. These differences across sex are reflected in the expected costs, and we observe that for males the gap between ICL and ML is lower than for females, although they both prefer ML.

At the bottom of Figure 1.1, we keep the cost fixed at £9000, with γ as above, and therefore the repayment periods are unchanged. However we let the subjective discount rate increase and we notice an increase of the gap in the expected costs when the individuals discount their future more. A ML is still preferred.

1.4 *Cost for the government*

The analysis presented in the previous Section and the simulations are based on the assumption that the government imposes a zero real interest on the repayment of the education loan. This means that repayments made to the government by the students are equal under both schemes. Defining P the total repayment to the government and r its interest rate; C is the cost of the loan to the students and ρ their subjective discount rate, we obtain the net benefit as $NB = P - C$. What we did earlier was to assume $r = 0$, therefore $P_{ML} = P_{ICL}$ and then we just compared the costs to the students to see which system was producing higher social net benefit:

$$NS = NB_{ICL} - NB_{ML} = C_{ML} - C_{ICL}.$$

Obviously, under risk aversion we compare the expected utilities, while under risk neutrality we can just compare the effective cost of repayments, as already explained in Section 1.3.

Our purpose in this Section is to consider a positive interest rate for the government and to observe how the net social benefit changes. For simplicity we assume risk neutrality but also the absence of hidden subsidies. This implies equal expected repayments periods, $ET_{ICL} = T_{ML}$.

1.4.1 Expected Costs and Expected Payments

As we have already seen, under risk neutrality the expected utility equation (1.1) becomes

$$V = \mathrm{E}(\int_s^{\infty} e^{-\rho t}\tilde{y}\, dt) - \mathrm{E}(\int_s^{T+s} e^{-\rho t} R\, dt) \tag{1.5}$$

So we can just compare the expected repayments, now differing between students and government, due to $r > 0$.
Knowing that $\mathrm{E}[\tilde{y}] = 1$ and $\mathrm{E}[T_{ICL}] = T_{ML}$, we require $\varphi = \frac{\gamma}{\mathrm{E}[1/\tilde{y}]}$. Under a ML the expected present value of the cost to the students and the expected present value of the payment to the government are respectively[9]

$$\mathrm{E}[C_{ML}] = \frac{\varphi}{\rho}[1 - e^{-\rho\frac{C\times\mathrm{E}[1/\tilde{y}]}{\gamma}}]e^{-\rho s} \tag{1.6}$$

$$\mathrm{E}[P_{ML}] = \frac{\varphi}{r}[1 - e^{-r\frac{C\times\mathrm{E}[1/\tilde{y}]}{\gamma}}]e^{-rs} \tag{1.7}$$

Similarly, under an ICL:

$$\mathrm{E}[C_{ICL}] = \mathrm{E}\,[\frac{\gamma\tilde{y}}{\rho}(1 - e^{-\rho\frac{C}{\gamma\tilde{y}}})e^{-\rho s}] \tag{1.8}$$

[9] To simplify the notation, by E[C] and E[P] we mean the expected present values E[PVC] and E[PVP].

$$\mathrm{E}[P_{ICL}] = \mathrm{E}\left[\frac{\gamma\tilde{y}}{r}(1 - e^{-r\frac{C}{\gamma\tilde{y}}})e^{-rs}\right] \tag{1.9}$$

We observe that the only difference between expected costs and expected repayments is given by r and ρ. We can now compute the net benefit under the two schemes:

$$NB_{ML} = \mathrm{E}[P_{ML}] - \mathrm{E}[C_{ML}]$$

$$NB_{ICL} = \mathrm{E}[P_{ICL}] - \mathrm{E}[C_{ICL}]$$

and the net social benefit

$$NS = NB_{ICL} - NB_{ML}.$$

In order to evaluate which system produces higher welfare we perform some calibrations. We consider the effects of the British HE reform and we assume an ICL repayment rate of 9%. Looking at Table 1.1, when the government's interest rate is lower than the students' time preference rate, our simulation shows an ICL is the preferred system and the effect of the reform is simply to increase the net social benefit. In our data females receive lower mean earnings and lower variance and therefore benefit from an ICL more than males. The results change when $r > \rho$, in this case a ML gives higher net social welfare.

1.5 Comparing Mortgage and Income Contingent Loans under Risk Aversion

In this Section we consider individuals who are risk averse and work out their expected utility (represented by equation (1.1)), under a mortgage loan and an income contingent loan system. We omit the majority of calculations that are shown in more detail in Appendices 1.10 and 1.10.1. We consider the assumptions stated in Section 1.2 and we develop the analysis using a constant relative risk aversion (CRRA) utility function[10]. The CRRA functional form for utility together with the CARA are each quite simple, involve just one parameter, and make analysis in many economic settings quite tractable. The risk aversion measures associated with these forms are also very simple. As a consequence, these two functional forms for utility and the risk preferences they represent have been

[10] We developed the analysis also using a constant absolute risk aversion functional form, CARA. We omit here the results for simplicity, but they can be found in a early working paper version of this work at the following link *http://www2.warwick.ac.uk/fac/soc/economics/research/papers/twerp_740.pdf*

frequently used in the literature examining expected utility based decisions. There are other more flexible functional forms for utility which allow a wider range of risk preferences to be represented than those available under CARA or CRRA. For example, the hyperbolic absolute risk averse (HARA) family. However these alternative functional forms involve more than one parameters and are less simple to manipulate and use. We do not want to add further complexity to our analysis and so we prefer using a CRRA functional form; for an extensive discussion of the other functional forms for utility functions see Meyer (2007).

1.5.1 Expected Utility with a Mortgage Loan

Under a mortgage loan, the expected utility is obtained by substituting $R = \varphi$ into equation (1.1):

$$V_{ML} = \int_{s}^{T+s} e^{-\rho t}\mathrm{E}\left[u\left(\tilde{y} - \varphi\right)\right] dt + \int_{T+s}^{\infty} e^{-\rho t}\mathrm{E}\left[u\left(\tilde{y}\right)\right] dt. \tag{1.10}$$

To get a closed-form solution for V_{ML}, we use a second order Taylor expansion around the mean $\mathrm{E}[\tilde{y} - \varphi] = 1 - \varphi$ [11] for utility during the repayment period, and around $\mathrm{E}[\tilde{y}] = 1$ for the utility after the repayment period:

$$\mathrm{E}[u(\tilde{y} - \varphi)] \simeq u\left(1 - \varphi\right) + \frac{1}{2}u^{''}\left(1 - \varphi\right)\sigma^2. \tag{1.11}$$

$$\mathrm{E}[u(\tilde{y})] \simeq u(1) + \frac{1}{2}u^{''}(1)\sigma^2. \tag{1.12}$$

We develop our analysis using a CRRA utility function.

$$u(\tilde{y}) = \frac{\tilde{y}^b}{b}.$$

where $b = 1 - a$ and a is the risk aversion parameter.

[11] See Pistaferri and Padula (2001) and Hartog and Serrano (2003).

After simplifying [12], we get:

$$\begin{aligned} V_{ML_{CRRA}} = & \frac{e^{-\rho s}}{\rho}\{\left(1 - e^{-\frac{\rho C}{\varphi}}\right)\left[\frac{(1-\varphi)^b}{b} + \frac{1}{2}(b-1)(1-\varphi)^{b-2}\sigma^2\right] \\ & + e^{-\frac{\rho C}{\varphi}}\left[\frac{1}{b} + \frac{1}{2}(b-1)\sigma^2\right]\}. \end{aligned} \tag{1.13}$$

1.5.2 Expected Utility with an Income Contingent Loan

Under an income contingent loan we do not know how long people take to repay their education debt. Therefore in the general equation of the expected utility the random earnings appear twice: first in the integral's bounds as random repayment period, second as an argument of the utility function.

$$V_{ICL} = \mathrm{E}\left\{\int_s^{\frac{C}{\gamma\tilde{y}}+s} e^{-\rho t} u\left[\tilde{y}(1-\gamma)\right] dt + \int_{\frac{C}{\gamma\tilde{y}}+s}^{\infty} e^{-\rho t} u(\tilde{y})\, dt\right\} \tag{1.14}$$

Solving the integral we get the following equation:

$$V_{ICL} = \frac{e^{-\rho s}}{\rho}\mathrm{E}\left\{\left[1 - e^{-\frac{\rho C}{\gamma\tilde{y}}}\right] u\left[\tilde{y}(1-\gamma)\right] + \left[e^{-\frac{\rho C}{\gamma\tilde{y}}}\right] u(\tilde{y})\right\}. \tag{1.15}$$

To simplify the calculations we define the expression included in the expected value operator as $g(\tilde{y})$. This trick allows us to apply a second order Taylor expansion of $\mathrm{E}[g(\tilde{y})]$ around the mean $\mathrm{E}[\tilde{y}] = 1$. Then, the equation (1.15) becomes:

$$V_{ICL} = \frac{e^{-\rho s}}{\rho}\left[g(1) + g''(1)\frac{\sigma^2}{2}\right]. \tag{1.16}$$

The remaining procedure (explained in Appendix 1.10.1) consists of calculating the value of $g(1)$ and $g''(1)$ in general, and with a CRRA utility function in particular. Finally, we substitute the expressions found in equation (1.16), and we obtain the following result. After simplifying, expected utility [13] is:

$$\begin{aligned} V_{ICL_{CRRA}} = & \frac{e^{-\left(s+\frac{C}{\gamma}\right)\rho}}{2b\gamma^2\rho}\{e^{\frac{\rho C}{\gamma}}(1-\gamma)^b\gamma^2[2+(b-1)b\sigma^2] - [(1-\gamma)^b - 1] \\ & \cdot [2\gamma^2 + ((b-1)b\gamma^2 + 2(b-1)C\gamma\rho + C^2\rho^2)\sigma^2]\}. \end{aligned} \tag{1.17}$$

[12] See Appendix 1.10 for the proof.

[13] The expected utility with an income contingent loan is equal to the expected utility with a mortgage loan if $\varphi = \gamma$ and the variance of the earnings is zero.

We can compare equation 1.17 and equation 1.13 through empirical simulations.

1.6 Empirical Background and Simulations

In this Section we first illustrate the BCS70 dataset used as a basis for calibrating the theoretical model. We describe graduate earnings and their standard deviation, in four possible environments, in order to get an idea of the wage uncertainty. An important assumption is the absence of selection bias, although we know that it could matter even for variance comparisons (Chen, 2004). However here we are more interested in observing how the theoretical model works under different potentially real situations[14]. In the second part of this Section we show the results of our calibrations and discuss the implications.

1.6.1 Data

Our statistics are based on the 1970 British Cohort Study (BCS70), that takes as its subjects around 17,000 British births in the week 5-11 April 1970. Subsequently, full sample surveys took place at ages 5, 10, 16, 26 and 30. BCS70 highlights all aspects of the health, educational and social development of its subjects as they passed through childhood and adolescence. In later sweeps, the information collected covers their transitions to adult life, including leaving full-time education, entering the labour market, setting up independent homes, forming partnerships and becoming parents. (Bynner, Butler et al., 2002). The initial sample follow-up in 1999-00 consists of 11261 respondents aged 30. The smaller sample size in the 1999-00 survey relative to the original survey in 1970 depends on sample attrition due to nonresponse and it cannot be avoided[15]. The lowest response rate in the BCS70 study was registered in the postal survey conducted in 1996 at age 26, the loss of observations was mainly due to a postal strike. However in the previous surveys, above all those based on interviews to the parents of the cohort's members, the rate of non response was quite high[16].

[14] The presence of selection bias is potentially an issue of what we are aware, however there is little evidence in the literature concerning the selection into subjects and into job sectors. We will try to address these questions in Chapter 3.

[15] In general, attrition due to non-response is low in the non-adult sweeps (1-3) and increases at the adult sweeps (4-5). For example, the response rates of the sweep 0 observed sample is over 86% at sweeps 1, 2 and 3 falling to around 73% at sweeps 4 and 5 (1970 British Cohort Study Technical Report (2004).

[16] It should be noted that the reason for non-participation at a later sweep may be because the cohort member has died or permanently emigrated. It is, for example, also possible for data to be missing for one

In general, the age 30 survey (1999-2000) was the first systematic attempt with widespread coverage to collect qualification and earnings data. It had a high response rate and it involved face to face interviews. For the purposes of our work, we merge the sweeps 1999-2000, 1980 and 1986. The last two sweeps are used because they provide information on family background: that is, family earnings and parental education. We have to stress the point that BCS is the only dataset that has family background. But it BCS70 seems useful for looking at earnings variance in graduation.

In our sample we include observations if: respondents have a NVQ 4 equivalent qualification in 2000[17]; they are in the labour market and earn a positive wage after graduation[18]. In particular, we consider those that got a degree from 1987 to 2000 and start working not earlier than the same year of graduation. This implies that the longest working period is 13 years, but we only consider the earnings in 1999-2000 and for full time or part time employees[19]. According to these criteria in the final sample there are 1177 respondents.

1.6.2 *Descriptive Statistics*

We observe the average annual gross wage and its standard deviation according to the individual characteristics, family background, degree subjects and job sector.
The average earnings in the sample are around £24000 with a standard deviation of £18300. Male average earnings are around 40% higher than female earnings, but also more than twice their variance (Table 1.2). Married graduates are around 63% of the sample and their earnings are slightly higher than single graduates, but somewhat more uncertain (Table 1.3). Those who have one or more children (22% of the total) have a lower earnings but those earnings are more uncertain (Table 1.4). It is useful to consider a breakdown of data from family background because it is one of the determinants of participation in HE and because family earnings determine how much the individuals are

part of the schedule especially as, during the years of childhood, data were obtained from different sources (parents, teachers and medical personnel)(1970 British Cohort Study Technical Report (2004)).

[17] The variable has been generated according to the UK national qualifications framework, NVQ equivalent level 4 includes academic qualifications (Degree and HE Diploma), vocational qualifications (BTEC Higher Certificate/ Diploma, HNC/HND) and occupational qualifications (NVQ level 4, Professional degree level qualifications, Nursing/paramedic, Other teacher training qualification, City & Guilds Part 4, RSA Higher Diploma).

[18] We exclude those working before and during education because this is a specific assumption in the theoretical model.

[19] This restriction allow us to clean from many inconsistencies in the earnings, and it is based on work undertaken by Lorraine Dearden and Alissa Goodman, Institute for Fiscal Studies.

allowed to borrow in the English loan schemes. We consider then the family earnings of the cohort members in 1980, when they are 10 years old (Table 1.5). We define "low" family earnings below £99 per week in 1980 prices; "medium" family earnings between £100 and £200 per week; high family earnings above £200 per week. Unfortunately, we discount the graduates from low earnings families because they just are 2% of the sample. The data for medium and high earnings family look more reasonable and with a relatively low uncertainty compared to the graduate average. Observing the graduate earnings given the mother qualifications in 1980 (Table 1.6), those with a graduate mother get the highest earnings, but the most uncertain is when the mother hold an 'O level' (secondary school qualification). Controlling for father occupation in 1986 (cohort members are 16 years old), see Table 1.7, those with a father in a professional occupation have a higher earnings than the graduate average and this is quite stable. Rather, when the father is unskilled the individuals get the highest earnings but on average these are also the most uncertain. We have to stress in this case as well that the number of observations is small (4% of the sample).

Table 1.8 shows three degree subjects, the earnings are above the average in all the cases, and quite similar to each other. However, those that took a degree in sciences (around 25% of the sample) have the lowest standard deviation. Finally, looking at the job sectors (Table 1.9): 62% of the graduates work in the private sector and earn around 30% more than those in the public sector. However, in the latter the age earnings profile is flat and this is reflected by a very low level of uncertainty.

1.6.3 Simulations under Risk Aversion

We perform the simulations according to the different levels of earnings and uncertainty described in Section 1.6.2. The method is the same followed in Section 1.3.1, however in this case we calibrate the equations (1.13) and (1.17), both when the repayment periods are equal and when they differ. We assume that the college premium is around 30%, consistent with most of the OLS estimations of the college premium, and the loan scheme does not affect the wage distribution.

We divide the simulations in 4 broad categories, using the average wage and standard deviation provided by the BCS70 sample:

1. individual characteristics: sex, marital status, children;
2. family background: family earnings 1980, mother's qualifications 1980, father's social class 1986;
3. degree subjects: science, social science, art and humanity.
4. public sector versus private sector.

In all the computations we set the following parameters:

- risk aversion: $a = \{0.25, \quad 0.5, \quad 0.75, \quad 1.5\}$, following the literature (Olson, White and Shefrin, Weiss);
- subjective discount factor: $\rho = \{0.08, \quad 0.15, \quad 0.3\}$, following the literature;
- cost of education: $C = \{£3450, \quad £9000\}$, before and after the English Higher Education reform;
- ICL repayment rate: $\gamma = \{0.02, \quad 0.09, \quad 0.2\}$, where 0.09 is the current rate fixed by the English reform;

In the following analysis when we change one parameter we keep the others constant at these levels: ICL repayment rates 9% (for English relevance), subjective discount rate 8%, cost £9000 (English current cost for 3-year degree), risk aversion 0.5. We do not consider every conceivable combination of parameters. However our qualitative results on the effect of any parameter is not sensitive to the assumed values of the other parameters.

Individual Characteristics

In the top left graph of Figure 1.2, we compare the expected utilities for increasing costs of education. We notice two initial important results: when the cost is low the difference between the two systems is small, although females prefer a ML and males an ICL. For increasing costs, uncertainty matters more: compared to females, standard deviation of males earnings is almost double (see Table 1.2). We observe that the gap in the strength of preference gets bigger for higher costs. Looking at the graph on the top right of Figure 1.2, we assume increasing subjective discount rate. For female the preferences are unchanged, instead for males the utility from an ICL reduces. We notice that the repayment periods, although equal for the two systems, differ according to the earnings. Therefore females with lower earnings have longer repayment periods.

The central left graph of Figure 1.2 shows that when the earnings are relatively high, an ICL is preferred and increasingly so for higher costs. The observed gap is due to the effect of the uncertainty, since the married have larger standard deviation of earnings they receive higher utility from an ICL. The graph on the central right of Figure 1.2 highlights the effect of risk aversion. For both group of graduates, higher risk aversion strengthens the preference for an ICL, and the gap remains unchanged.

The bottom left graph of Figure 1.2 shows how the effect of the uncertainty implies a preference for an ICL over a ML for those with children and those without. However, the higher earnings of those without children (see Table 1.4) gives higher utilities within the ICL system. In the graph on the bottom right of Figure 1.2 we set the highest cost and we observe that when risk aversion is very high, for those with children the utility from an ICL reduces sharply. A possible explanation is the big difference in the repayment periods, for those with no children it is around 12 years, for graduates with children around 18 years. Moreover, φ is endogenous and depends on both γ and $\mathrm{E}[1/\tilde{y}] = \overline{h}$, when $\overline{h}$ is high, as in the case of graduates with children, the ML utility decreases. Smaller installments and high risk aversion makes an ML more advantageous.

Family Background

The top left graph of Figure 1.3 shows the variation of the expected utilities for increasing costs of education, assuming that the graduates come from families with different earnings in 1980. The effect of low family earnings is not relevant because there are few observations. However it can be used to see what happens when the standard deviation is very high (see Table 1.5). An ICL is highly preferred for higher cost, and also looking at the graph to the right, when the risk aversion is very large the ICL expected utility drops sharply. This is due, as mentioned above, to the high $\overline{h}$ and the long repayment period (19 years compared to around 13 under the higher level of family earnings). When graduates come from families with high earnings, they obtain a low variance wage and this is reflected in a preference for a ML. Graduates from medium family earnings obtain lower earnings on average but more uncertain: therefore an ICL provides higher utility. The same results are confirmed for increasing risk aversion.

Controlling for mother's education (central left graph Figure 1.3 and Table 1.6), those with a graduate mother get the highest earnings and least uncertain, therefore they prefer a ML. Comparing those whose mothers have no qualification with those whose mothers have secondary school qualification, earnings and standard deviation are above the average for both, but the first prefer an ICL with more intensity. The same is true for increasing risk aversion, although when it becomes too high utility from ICL reduces.

We now consider the father's social class in 1986 (bottom left graph Figure 1.3 and Table 1.7). For unskilled fathers the observations are again too few, but since the standard deviation is very high there is a strong preference for an ICL. Those with a father in a skilled occupation get earnings above the average and very stable instead, and they are almost indifferent between the two systems. If the father is professional an ICL is preferred. For increasing risk aversion we observe a slightly increasing preference for an ICL.

Degree Subjects and Sectors

The graduates that took the three subject groups observed (see Table 1.8) have earnings higher than the all graduate average. The lowest earnings variance is with a Science degree and in fact the science graduates are almost indifferent between the two systems (Figure 1.4). Those with a degree in Art and Humanities have the most uncertain earnings, and they strongly prefer an ICL, with higher utility for increasing cost and risk aversion. For higher subjective discount rates the preference for an ICL reduces. The opposite happens for higher risk aversion.

We now consider the effects on expected utilities of graduates working in the private and public sector (see descriptive statistics in Table 1.9). In Figure 1.5 the graphs on the left side assume no hidden subsidies, while the graphs on the right side do not. In the latter case, according to Proposition 1, we set $\gamma = \varphi$ and we have $E(T_{ICL}) > T_{ML}$[20]. In the top graphs of Figure 1.5 we compare the two loan schemes for increasing cost. When there are no hidden subsidies, as expected, in the private sector graduates prefer an ICL, since they get a higher but more uncertain earnings. In the public sector, income is lower

[20] We set a fixed ML installment to £900 pounds, that means a repayment period with the lowest cost of around 4 years, and with the highest of 10 years.

but also less uncertain and a ML is more preferred instead. In presence of hidden subsidies an ICL is much more preferred in the private sector. However, the interesting result is that also in the public sector they also prefer an ICL, with increasing intensity as costs rise. Therefore, the effect of the hidden subsidy is very strong: enough to change the preferences when the level of uncertainty is low. The same behavior is confirmed in the graphs in the middle of Figure 1.5 where we increase risk aversion. We finally compare the expected utilities under the same system in the two different sectors (graph in the bottom left corner). The difference between ICL in the private sector and ICL in the public sector is negative, and the same happens under a ML, with a slight change for increasing costs. This means that the effect of low uncertainty prevails, and under the same scheme a more stable career gives higher utility.

1.6.4 Conclusion

The analysis suggest some clear effects: when there is very high uncertainty an ICL is preferred in all cases because it gives better insurance. Excluding extreme situations, the preference for an ICL depends strongly on the combination of level of earnings and its standard deviation, compared to the all graduates average. Therefore, if earnings are high, and the standard deviation below average, a ML gives higher utility (e.g. mother with a degree). If both income and standard deviation are slightly above the average (e.g. one or more children, or mother with O level), an ICL is preferred but with less intensity. When the costs of education are high and uncertainty is high as well, ICL gives higher utility. Moreover, the effect of risk aversion in general is to increase the preference for an ICL. However, it is particular interesting if combined with high cost and long repayment periods, because for very high risk aversion the utility of an ICL drops drastically. If we consider age earnings profile with low standard deviation and above average earnings (for example people with a degree in the public sector), what matters is the relative level of uncertainty: if low a ML is preferred. Finally, looking at the job sectors we notice that when we compare a stable career with another which features high earnings uncertainty, in the first case a ML gives always higher utility. However, if we don't rule out the hidden subsidy we observe a sharp increase in the utility under an ICL, sufficient to invert the initial preference for a ML, even for low variance earnings.

1.7 Increasing earnings

In this Section we extend our model to incorporate stochastic changes of earnings over time. We make the model more realistic and verify which conditions still hold relative to the case of static earnings. We assume that graduate earnings are no longer affected by a single life time shock, but there is a shock each year throughout the individual working life. To model this assumptions we consider the earnings growth rate following a geometric Brownian motion $W(t)$[21]. This means that $y(t)$ satisfies

$$dy(t)/y(t) = \lambda dt + \sigma dW(t). \tag{1.18}$$

This expression can be interpreted heuristically as expressing the relative, or percentage, increment dy/y in y during an instant of time dt. λ is the deterministic growth rate and σ its standard deviation. Solving the stochastic differential equation (1.18) we obtain the stochastic earnings:

$$y(t) = y(0)\exp[(\lambda - \frac{1}{2}\sigma^2)t + \sigma W(t)] \tag{1.19}$$

The solution is explained in the Appendix 1.12. Equation (1.19) represents the new earnings we use to compute the expected utilities under the two loan schemes. Since it is not straightforward to obtain an algebraic solution for the expected utilities under an ICL we adopt a numerical method. We consider a discrete form of equation (1.19) because it is more relevant to our problem. The method is reported in detail in the Appendix (1.12); briefly, we generate many earnings paths of the same length (equal to a working life period of 40 years), and we use them to compute the expected utilities. Each earnings path produces one level of utility, therefore we average over the number of paths created. Ultimately we get the average expected utilities under an ICL (A_{ICL}) and ML (A_{ML}). What is important to stress under this new approach is that two new parameters enter the model. Looking at the equation (1.19), they are the initial earnings (y_0) and its deterministic growth rate (λ). Moreover, σ is the volatility of the Brownian motion and represents the maximum variation of the earnings in the interval t (for us one year). As we did in the first part of this Chapter we use real data to calibrate the model and simulate different scenarios.

[21] $W(t)$ is Normally distributed with $\mathrm{E}(W(t)) = 0$ and $Var(W(t)) = t$.

1.7.1 Setting the new model

Under this new stochastic framework the only variable that we fix using actual data is the initial earnings. In our BCS70 dataset, we generate a new variable for initial earnings, using graduate earnings within four years of graduation. As shown in Table 1.10, there are only 405 observations for the initial wage with an average of almost £24000. We use £8600, which is the mean of the bottom decile, and £23987 which are the median earnings. We allow different earnings growth for different individuals.

Uncertainty now affects the growth of earnings over years and a single individual. We choose three levels of σ (0.02 0.05 0.15) in order to have different intensities of the effect of the stochastic shock on earnings. For example, a $\sigma = 0.05$ means that the maximum annual variation of the earnings can be 5% around the trend growth, which occurs if $\sigma = 0$.

In our dataset we do not have information on the growth rate of the earnings because we observe a cross section in 1999-00. If we had merged this with the sweep of 1996 we would have lost a lot of observations, since that survey was conducted through mail questionnaires and many people did not respond. Therefore, we would have ended up with just a few observations. We therefore simply set three values for the deterministic growth rate ($\lambda = 0.5 \quad 1 \quad 1.5$). Assuming no uncertainty, for example a $\lambda = 0.5$ corresponds to a increase of 1.2% per year of initial earnings over 35 years. If $\lambda = 1$ the increase of the initial earnings is 2.4% p.a.

Extension from the model in the previous Section would be simple if earnings never fell below a threshold - effectively moving the bankruptcy constraint from 0 to £15000. The £15000 threshold is really just an institutional detail of the English system - there for equity reasons. We simulate with different thresholds: first zero threshold in order to have a similar case as that analyzed in the previous section, then a threshold of £10000 and finally a threshold of £15000 as in the English reform. This means that if in any year the earnings are below the threshold no payment is due. We perform the simulations setting first an equal repayment period[22] between the 2 systems. Then we fix the ML installment to £1000 for a cost of £9000; and to £1500 and £2400 for a cost of £12000. In this way we get three fixed repayment periods of 10, 8, 5 years respectively. The costs are chosen

[22] Since the earnings are stochastic, we first get the repayment period under an ICL and then set this value also for a ML.

according the English Reform, that is £3000 per year for a 3-year degree. We finally set the other parameters as in the simulation in Section1.6.

1.7.2 Simulations

The results of the simulations are reported in Tables 1.11-1.14, where each time we change one parameter keeping the others equal to those in Table 1.11. We show the results for both low and high earnings. Looking at Table 1.11, we notice a first important effect: the preference for one system over the other depends strongly on the level of the initial earnings. For low starting earnings an ICL is the favored system, but for high starting earnings the utility under an ICL declines sharply and a ML can be preferred. Another important result is that the level of uncertainty matters less than in the case with static earnings, although the direction of the effect is the same. We observe, in fact, that for higher earnings volatility the utility of an ICL increases but at a very slow rate. When the earnings are high the effect of uncertainty is more evident, in fact for $\rho = 8\%$ the initial preference for a ML is replaced by a preference for an ICL.

For a higher subjective discount rate, we notice a reduction of the gap between the two systems if the earnings are low. For high earnings the preference for an ICL is increasing but the two systems give almost equal utility. In Table 1.12, we first increase and then decrease the fixed ML installment with respect to the baseline case. When the ML repayment period is very long, an ICL is still preferred for low earnings, but not for high earnings. Conversely, short ML repayment periods always make an ICL preferable.

The effect of increasing risk aversion is to strengthen the preferences for one system. Low income earners prefer an ICL, and becoming more risk averse increase their utility under this system. Instead for high income earners the effect of risk aversion is correlated to the level of uncertainty. If σ is low, they prefer a ML even at high risk aversion. If the level of σ is high the two systems are more or less equal, although for high risk aversion a ML becomes the favourite. In Table 1.14 we consider an increase of the deterministic growth rate, from 1% to 4% per year. For low income earners the utility from an ICL reduces, and for high income earners the preference for a ML increases. In Table 1.15 we show the effects of different ICL thresholds for increasing levels of uncertainty. We first notice that an ICL is less preferred, as expected, when there is no repayment threshold because low income earners must also repay their loan. For a higher threshold the ICL becomes more preferred, and for very high uncertainty those earning a high income also prefer an ICL. So at higher thresholds the extent to which an ICL provides insurance is greater.

In Table 1.16 we finally consider the effects of the English HE reform: if the level of the initial earnings is low, an ICL is always preferred and increasingly when the costs rise. When the initial earnings are low the effect is reversed and a ML is the preferred system.

1.7.3 HE Participation

So far we have assumed that individuals will always participate in HE irrespective of the loan scheme but HE participation may itself depend on the parameters of the loan schemes available. So we now consider the level of participation in HE under the two systems. By participation we mean individuals that went to HE.

We assume again a HE education wage premium of around 30%, consistent with most of the OLS estimations of the college premium. We model this assuming constant earnings for those not going to college, which is 30% lower than the initial earnings of those going to college. We compute the total present values of the net earnings under a ML and an ICL and we compare them with the total present value of the non schooling earnings. We consider the same parameters as in the previous simulations.

Observing Table 1.17, when the initial earnings are low and we have low σ the participation is almost the same under the 2 systems. Instead when σ increases participation under an ICL is higher relative to ML. Keeping σ constant, e.g. 5%, using the pre-reform cost of education, the participation in HE with an ICL is 7% higher than under a ML. Using post-reform cost of education, which is £3000 p.a., we see a bigger effect of the loan scheme on participation, which is 26% higher under a ICL compared to a ML. When we consider a high level of earnings the difference between the two systems is very small, with only a slightly larger participation rate under a ML.

1.7.4 Conclusion

The simulations show that the preference for one system over the other are driven by the level of the initial earnings and the size of the ML installment. Low wage earners have a strong preference for an ICL, but when we change the parameters that affect lifetime earnings (such as the growth rate) to make it higher, then the utility of an ICL decreases. For increasing uncertainty the utility of an ICL rises, but by very small amounts. Higher risk aversion increases the preferences for the desired system. The size of the ML installments can change the preference: higher installments imply that the utility of a ML reduces sharply. Finally, participation in HE is higher under an ICL than under a ML, for low initial income earners and rising uncertainty.

1.8 Conclusion

In this Chapter we presented a theoretical model to compare two loan schemes for higher education, when graduate earnings are uncertain. The findings of the model have been calibrated using real data on graduate earnings, obtained from the 1970 British Cohort Survey. In the first part of the Chapter we assume that the graduate earnings are affected by a single lifetime shock, and we compute the individual expected utilities under an ICL and a ML for risk neutral and risk averse people. Our first result, supported by the empirical simulation, is that for risk neutral individuals the preference depends strongly on the presence of hidden subsidies. Assuming different repayment periods the expected costs under a ML are higher than the costs under an ICL. But in the case of same repayment periods, the preference order is inverted. For risk averse individuals we evaluated the effects of our model under different possible scenarios. We used information on graduate earnings and relative uncertainty controlling for individual characteristics, family background, degree courses and job sector.

Considering the English Higher Education reform, the main result is that for high earnings uncertainty an ICL is the preferred system. Therefore, people from low educated parents background, males over females, people working in the private sector prefer an ICL. Instead those with high earnings and low uncertainty, such as those with highly educated mother, prefer a ML. Another important result is that for increasing risk aversion an ICL is more preferred, this confirms the conviction that this system provides better insurance against an uncertain future. Moreover, a flat age earning profile in the public sector gives modest earnings but also low uncertainty, therefore a ML produces higher utility. However, if we don't rule out the hidden subsidy an ICL becomes preferred even when the degree of uncertainty is very low. In the second part of the Chapter, we changed the assumptions on earnings, allowing stochastic growth across the working life. The results of the new framework have in common with the previous one a preference for an ICL with increasing uncertainty. However the factors that affect the choice of one system over the other are different. The size of the starting earnings is the main discriminant: when the earnings are low an ICL is the preferred system. For high initial earnings a ML gives higher utility instead. Finally, increasing the size of the ML installment makes an ICL the preferred system.

1.9 *Appendix: Proof of Proposition 1*

We know that $T_{ICL} = \frac{C}{\gamma\tilde{y}}$ and $T_{ML} = \frac{c}{\varphi}$, and given the assumption $\mathrm{E}[\tilde{y}] = 1$ we compute the expected value of the repayment period under an ICL.

$$\mathrm{E}(T_{ICL}) = \mathrm{E}(\frac{C}{\gamma\tilde{y}}) = \frac{C}{\gamma} \times \mathrm{E}(\frac{1}{\tilde{y}})$$

by the Jensen's inequality we know that

$$\frac{C}{\gamma} \times \mathrm{E}(\frac{1}{\tilde{y}}) > \frac{C}{\gamma \mathrm{E}(\tilde{y})}$$

that implies

$$\mathrm{E}(\frac{1}{\tilde{y}}) > 1.$$

Given this result it is straightforward to prove the point (a):

if $\gamma = \varphi$ then

$$\frac{C}{\gamma} \times \mathrm{E}(\frac{1}{\tilde{y}}) > \frac{C}{\varphi} \rightarrow \mathrm{E}(T_{ICL}) > T_{ML}.$$

Point (b)

We assume $\mathrm{E}(T_{ICL}) = T_{ML}$ that means

$$\frac{C}{\gamma} \times \mathrm{E}(\frac{1}{\tilde{y}}) = \frac{C}{\varphi}$$

we get γ:

$$\gamma = \varphi \times \mathrm{E}(\frac{1}{\tilde{y}}) \quad \Longrightarrow \gamma > \varphi$$

since $E(\frac{1}{\tilde{y}}) > 1$.

1.9.1 Proof of Proposition 2

Under risk neutrality equation (1.1) becomes

$$V = E(\int_s^{\infty} e^{-\rho t}\tilde{y}\, dt) - E(\int_s^{T+s} e^{-\rho t} R\, dt) \tag{1.20}$$

So we can compare only the expected costs. Under a ML the present value of the cost of size C is:

$$\begin{aligned} PVC_{ML} &= \int_s^{T+s} \varphi e^{-\rho t}\, dt \\ &= e^{-\rho s}\frac{\varphi}{\rho}[1 - e^{-\rho\frac{C}{\varphi}}]. \end{aligned} \tag{1.21}$$

Under an ICL the present value of the cost of size C is:

$$\begin{aligned} PVC_{ICL} &= \int_s^{\tilde{T}+s} \tilde{y}\gamma e^{-\rho t}\, dt \\ &= e^{-\rho s}\frac{\gamma\tilde{y}}{\rho}\left[1 - e^{-\rho\frac{C}{\gamma\tilde{y}}}\right] \quad . \end{aligned} \tag{1.22}$$

Knowing that $E(\tilde{y}) = 1$, we take the expected value of both the equations above.

$$E(PVC_{ML}) = \frac{\varphi}{\rho}[1 - e^{-\rho\frac{C}{\varphi E(\tilde{y})}}]e^{-\rho s} \tag{1.23}$$

$$E(PVC_{ICL}) = E\,[\frac{\gamma\tilde{y}}{\rho}(1 - e^{-\rho\frac{C}{\gamma\tilde{y}}})e^{-\rho s}] \tag{1.24}$$

Case $\gamma = \varphi$

Under this condition we have $E[T_{ICL}] > T_{ML}$. We can easily observe that the expected values of the costs can be written:

$$E(PVC_{ML}) = f[E(\tilde{y})]$$

$$E(PVC_{ICL}) = Ef(\tilde{y})$$

Since $f(\tilde{y}) = \frac{\gamma \tilde{y}}{\rho}\left(1 - e^{-\frac{\rho C}{\gamma \tilde{y}}}\right)e^{-\rho s}$ is a concave function [23] by the Jensen's inequality we obtain that the expected costs under an ICL are lower than the expected costs under a ML: $\mathrm{E}(PVC_{ICL}) < \mathrm{E}(PVC_{ML})$. According to equation (1.20) the expected utility under an ICL is higher than the expected utility under a ML.

Case $\mathrm{E}[T_{ICL}] = T_{ML}$

To verify which costs are higher we have to substitute $\varphi = \frac{\gamma}{\mathrm{E}(1/\tilde{y})}$ in the equation (1.23) and then compare the two equations of the present value of the costs under the two systems:

$$\mathrm{E}(PVC_{ML}) = \frac{\varphi}{\rho}\,[1 - e^{-\rho \frac{C}{\frac{\gamma}{\mathrm{E}(1/\tilde{y})}\mathrm{E}(\tilde{y})}}]e^{-\rho s} \tag{1.25}$$

$$\mathrm{E}(PVC_{ICL}) = \mathrm{E}\,[\frac{\gamma\,\tilde{y}}{\rho}\,(1 - e^{-\rho \frac{C}{\gamma \tilde{y}}})e^{-\rho s}] \tag{1.26}$$

As we can see the comparison is not straightforward, therefore we adopt a numerical solution using the real data on graduate earnings provided in our dataset from BCS70. In order to be consistent with the assumption $\mathrm{E}(\tilde{y}) = 1$, we first standardize the annual gross earnings of the graduates. We call w_i the earnings in the sample and divide each of the them by the sample mean, and we call this new variable z.

$$z_i = \frac{w_i}{\frac{1}{n}\sum_{i=1}^{n} w_i} \quad for \;\; i = 1...n$$

then

$$\overline{z} = \frac{1}{n}\sum_{i=1}^{n} z_i = 1$$

and we use the sample analogue $\overline{z}$ to estimate $\mathrm{E}(\tilde{y}) = 1$. Instead, to estimate $\mathrm{E}(\frac{1}{\tilde{y}})$ we generate its sample analogue

$$\overline{h} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{z_i} \;.$$

In our sample $\overline{h} > \frac{1}{\overline{z}}$, then the Jensen's inequality holds.

[23] $f''(\tilde{y}) = -\frac{\rho C^2 e^{-\left(s + \frac{C}{\gamma \tilde{y}}\right)\rho}}{\gamma \tilde{y}^3}$. It is reasonable to assume that γ, ρ and C are all greater or equal than zero. Therefore, the second derivative of $f(\tilde{y})$ is always negative when the shock on earnings is positive: $f''(\tilde{y}) < 0, \quad \forall \tilde{y} > 0$.

1.10 Appendix: Expected Utility with a Mortgage Loan

The Taylor approximation in equation (1.11) is the following

$$\begin{aligned} \mathrm{E}[u(\tilde{y}-\varphi)] &= \mathrm{E}\left\{u(1-\varphi)+u'(1-\varphi)(\tilde{y}-1)+\frac{1}{2}u''(1-\varphi)(\tilde{y}-1)^2\right\} \\ &= u(1-\varphi)+u'(1-\varphi)\,\mathrm{E}(\tilde{y}-1)+\frac{1}{2}u''(1-\varphi)\,\mathrm{E}(\tilde{y}-1)^2 \\ &= u(1-\varphi)+\frac{1}{2}u''(1-\varphi)\,\sigma^2. \end{aligned} \tag{1.27}$$

Plugging the equations (1.11) and (1.12) in the equation (1.10), substituting $T = C/\varphi$ and solving the integral, we obtain:

$$\begin{aligned} V_{ML} &= \frac{e^{-\rho s}}{\rho}\left(1-e^{-\frac{\rho C}{\varphi}}\right)\left[u(1-\varphi)+\frac{1}{2}u''(1-\varphi)\,\sigma_s^2\right] \\ &+\frac{e^{-\rho s}}{\rho}e^{-\frac{\rho C}{\varphi}}\left[u(1)+\frac{1}{2}u''(1)\,\sigma_s^2\right]. \end{aligned} \tag{1.28}$$

Finally, substituting a CRRA utility function in equation(1.28) and simplifying we get equation (1.13).

1.10.1 Expected Utility with an Income Contingent Loan

In Section (1.5.2) we defined a new function $g(\tilde{y})$ as:

$$g(\tilde{y}) = \left[1-e^{-\frac{\rho C}{\gamma\tilde{y}}}\right]u\left[\tilde{y}(1-\gamma)\right]+\left[e^{-\frac{\rho C}{\gamma\tilde{y}}}\right]u(\tilde{y}) \tag{1.29}$$

We rewrite the equation (1.15)

$$V_{ICL} = \frac{e^{-\rho s}}{\rho}\mathrm{E}[g(\tilde{y})] \tag{1.30}$$

and we apply a second order Taylor expansion to $\mathrm{E}[g(\tilde{y})]$, around the mean $\mathrm{E}[\tilde{y}] = 1$, then:

$$\begin{aligned} \mathrm{E}[g(\tilde{y})] &= \mathrm{E}\left\{ g(1) + g'(1)(\tilde{y}-1) + g''(1)\frac{(\tilde{y}-1)^2}{2} \right\} \\ &= g(1) + g'(1)\mathrm{E}(\tilde{y}-1) + \frac{g''(1)}{2}\mathrm{E}\,(\tilde{y}-1)^2 \\ &= g(1) + g''(1)\frac{\sigma_s^2}{2}. \end{aligned} \tag{1.31}$$

The equation (1.30) becomes

$$V_{ICL} = \frac{e^{-\rho s}}{\rho}\left[g(1) + g''(1)\frac{\sigma_s^2}{2} \right] \tag{1.32}$$

From now on we follow this procedure:

1. we work out the value of g(1), in general and with a CRRA utility function;
2. we work out the first derivative and the second derivative of $g(\tilde{y})$, both in general and with a CRRA utility function;
3. we calculate $g'(1)$ and $g''(1)$ using a CRRA utility function;
4. we substitute the equations of $g(1)$ and $g''(1)$, using a CRRA utility function, in the equation (1.32) and we obtain equations (1.39) and (1.17).

- **Value of g(1)**

In general,

$$g(1) = \left[1 - e^{-\frac{\rho C}{\gamma}}\right] u\left[(1-\gamma)\right] + \left[e^{-\frac{\rho C}{\gamma}}\right] u\,(1) \tag{1.33}$$

Using a CRRA utility function we have

$$g(1)_{CRRA} = \frac{1}{b}[-e^{-\frac{\rho C}{\gamma}}((1-\gamma)^b - 1) + (1-\gamma)^b]. \tag{1.34}$$

- **Value of g'($\tilde{y}$)**

In general,

$$g'(\tilde{y}) = u'[\tilde{y}(1-\gamma)](1-\gamma)\left[1-e^{-\frac{\rho C}{\gamma \tilde{y}}}\right] + u[\tilde{y}(1-\gamma)]\left[\frac{-\rho C e^{-\frac{\rho C}{\gamma \tilde{y}}}}{\gamma \tilde{y}^2}\right] + u'(\tilde{y})[e^{-\frac{\rho C}{\gamma \tilde{y}}}] + u(\tilde{y})\left[\frac{\rho C e^{-\frac{\rho C}{\gamma \tilde{y}}}}{\gamma \tilde{y}^2}\right] \tag{1.35}$$

using a CRRA utility function:

$$g'(\tilde{y})_{CRRA} = (\tilde{y}(1-\gamma))^{b-1}(1-\gamma)\left[1-e^{-\frac{\rho C}{\gamma \tilde{y}}}\right] + (\tilde{y}(1-\gamma))^{b}\left[\frac{-\rho C e^{-\frac{\rho C}{\gamma \tilde{y}}}}{b\gamma \tilde{y}^2}\right] + \tilde{y}^{b-1}[e^{-\frac{\rho C}{\gamma \tilde{y}}}] + \left[\frac{\tilde{y}^{b-2}\rho C e^{-\frac{\rho C}{\gamma \tilde{y}}}}{b\gamma}\right]. \tag{1.36}$$

- **Value of g"($\tilde{y}$)**

$$g''(\tilde{y}) = \frac{e^{-\frac{\rho C}{\gamma \tilde{y}}}\rho C(2\gamma\tilde{y}-\rho C)}{\tilde{y}^4\gamma^2}u[\tilde{y}(1-\gamma)] + \frac{e^{-\frac{\rho C}{\gamma \tilde{y}}}\rho C(-2\gamma\tilde{y}+\rho C)}{\tilde{y}^4\gamma^2}u(\tilde{y}) - \frac{2e^{-\frac{\rho C}{\gamma \tilde{y}}}\rho C(1-\gamma)}{\tilde{y}^2\gamma}u'[\tilde{y}(1-\gamma)] + \frac{2e^{-\frac{\rho C}{\gamma \tilde{y}}}\rho C}{\tilde{y}^2\gamma}u'(\tilde{y}) + \left[1-e^{-\frac{\rho C}{\gamma \tilde{y}}}\right](1-\gamma)^2u''[\tilde{y}(1-\gamma)] + \left[e^{-\frac{\rho C}{\gamma \tilde{y}}}\right]u''(\tilde{y}). \tag{1.37}$$

Now we work out $g''(\tilde{y})$ using a a CRRA and evaluating in $\tilde{y} = 1$

$$g''(1)_{CRRA} = \frac{1}{b\gamma^2}\{e^{-\frac{\rho C}{\gamma}}[(b-1)b\gamma^2[1+(e^{\frac{\rho C}{\gamma \tilde{y}}}-1)(1-\gamma)^b] + 2\rho C(b-1)\gamma(1-(1-\gamma)^b) + C^2\rho^2(1-(1-\gamma)^b)]\}. \tag{1.38}$$

- **Results**

Substituting $g(1)$ and $g''(1)$ in equation (1.32) we get the general expected utility under an income contingent loan:

$$
\begin{aligned}
V_{ICL} = & \left[1-e^{-\frac{\rho C}{\gamma}}\right] u\left[(1-\gamma)\right] + \left[e^{-\frac{\rho C}{\gamma}}\right] u\left(1\right) \\
& + [\frac{e^{-\frac{\rho C}{\gamma}} \rho C(2\gamma-\rho c)}{\gamma^2} u[1-\gamma] + \frac{e^{-\frac{\rho C}{\gamma}} \rho C(-2\gamma+\rho c)}{\gamma^2} u(1) \\
& - \frac{2e^{-\frac{\rho C}{\gamma}} \rho C(1-\gamma)}{\gamma} u'[1-\gamma] + \frac{2e^{-\frac{\rho C}{\gamma}} \rho C}{\gamma} u'(1) \\
& + \left[1-e^{-\frac{\rho C}{\gamma}}\right] (1-\gamma)^2 u''[1-\gamma] + \left[e^{-\frac{\rho C}{\gamma}}\right] u''(1)] \frac{\sigma_s^2}{2}.
\end{aligned}
\tag{1.39}
$$

Substituting in equation (1.32) the equations for $g(1)$ and $g''(1)$ with a CRRA utility function, we obtain equation (1.17).

1.11 Appendix: Equation Brownian Motion

We assume that the growth rate of the earnings is affected by a white noise process, formally defined as the derivative of the standard Brownian motion, or standard Wiener process, $W(t)$

$$\epsilon = dW(t)/dt.$$

The derivative does not exist in the usual sense, since the Brownian motion is nowhere differentiable. However this process is used with the convention that its meaning is given by integral representation. If $\sigma(\tilde{y},t)$ is the intensity of the noise at point y at time t, then it is common agreement that $\int_0^T \sigma(y(t),t)\epsilon(t)dt = \int_0^T \sigma(y(t),t)dW(t)$. In our case, adding a white noise $\sigma\epsilon(t)$ to the constant growth rate of the earnings λ, we obtain the stochastic differential equation, SDE

$$dy(t)/dt = (\lambda + \sigma\epsilon)y(t). \tag{1.40}$$

where $\sigma \geq 0$, $\lambda > 0$, are some constants, and $y(0)$ is the deterministic component of the earnings. This means that $y(t)$ satisfies

$$dy(t)/y(t) = \lambda dt + \sigma dW(t). \tag{1.41}$$

To solve the SDE (1.41) we introduce the Itô process $R(t)$ given by

$dR(t) = \lambda dt + \sigma dW(t)$. We rewrite the SDE as

$$dy(t) = y(t)dR(t) \tag{1.42}$$

this means that $y(t)$ is the stochastic exponential, $\varepsilon(R)$, of $R(t)$. The solution of equation(1.42) is

$$\begin{aligned} y(t) &= y(0)\varepsilon(R)(t) \\ &= y(0)\exp[R(t) - R(0) - \frac{1}{2}[R,R](t)]. \end{aligned} \tag{1.43}$$

$R(t)$ is easily found to be $R(t) = \lambda t + \sigma W(t)$, $R(0) = 0$, and its quadratic variation $[R,R](t)$ is the quadratic variation of an Itô process and equal to

$$[R,R](t) = \int_0^t \sigma^2 ds = \sigma^2 t.$$

Substituting these expressions in equation(1.43) we get

$$\begin{aligned} y(t) &= y(0)\exp[\lambda t + \sigma W(t) - \frac{1}{2}\sigma^2 t] \\ &= y(0)\exp[(\lambda - \frac{1}{2}\sigma^2)t + \sigma W(t)]. \end{aligned} \tag{1.44}$$

This process is a geometric brownian motion.

1.12 Appendix: Numerical Method - Brownian Motion

1. We generate a path of annual earnings for an individual working life. Since the problem requires a discrete solution, we apply the Euler-Maruyama method that takes the form

$$y_j = y_{j-1} + y_{j-1}\lambda\Delta t + y_{j-1}\sigma(W(\tau_j) - W(\tau_{j-1})). \tag{1.45}$$

To generate the increments $W(\tau_j) - W(\tau_{j-1})$ we compute discretized Brownian motion paths, where $W(t)$ is specified at discrete t values. As explained in Higham (2001) we first discretize the interval $[0, I]$. We set $dt = I/N$ for some positive integer N, and let W_j denote $W(t_j)$ with $t_j = jdt$. According to the properties of the standard Brownian motion $W(0) = 0$ and

$$W_j = W_{j-1} + dW_j \tag{1.46}$$

where dW_j is an independent random variable of the form $\sqrt{dt}N(0,1)$. The discretized brownian motion path is a 1-by-N array, where each element is given by the cumulative sum in equation (1.46). To generate equation (1.45), we define $\Delta t = I/L$ for some positive integer L, and $\tau_j = \Delta t$. As in Higham (2001) we choose the step-size Δt for the numerical method to be an integer multiple $R \geq 1$ of the Brownian motion increment dt: $\Delta t = Rdt$. Finally, we get the increment in equation (1.45) as cumulative sum:

$$W(\tau_j) - W_{(\tau_{j-1})} = W(jRdt) - W((j-1)Rdt) = \sum_{h=jR-R+1}^{jR} dW_h. \tag{1.47}$$

The Brownian motion of equation (1.46) is produced setting $I = 1$ and $N = 160$ in order to have a small value of dt. Using a random number generator we produce 160 "pseudorandom" numbers from the N(0,1) distribution. The increments of equation (1.47) are computed setting $R = 4$, in order to have 40 annual earnings.

2. Income contingent loan. We work out the yearly repayments as fixed percentage of the stochastic earnings generated. If the earnings are higher than £15000 the payments are positive, otherwise they are zero. We then built a vector whose elements are the cumulative sum of the repayments, in order to see the amount of loan repaid. To obtain the repayment period, we observe the years in which the cumulative sum of the payments is equal[24] to the cost of education. We work out the individual utility as discounted sum of the net earnings during and after the repayment period, up to the end of the working life. We use a CRRA utility function.

3. Mortgage loan. We set the fixed repayment period as the ratio between the cost of education and the annual installment. The individual utility is given by the discounted sum of the net earnings during and after the repayment period. We use a CRRA utility function. However, it can happen that the annual earnings are

[24] Since it is almost impossible to get a value equal to the cost, when the repayment is slightly greater than it we assume the debt has been paid off.

lower than the installment, in a usual mortgage loan the individual repays in the subsequent years at a higher interest rate. Here to highlight the loss of utility in the case of no repayment in one year, we compute the level of the utility for that year as a negative percentage[25] of the annual earnings.

4. From steps (2) and (3) we obtain a single value for the utility for an individual earnings path generated in point (1). We generalize our method generating a high number of earnings paths (1000) and for each path we compute a level of utility. We then work out the average utility under both financing scheme and the difference of the average in order to compare the two systems.

5. We let the various parameters change and we repeat steps (1) to (4), observing the trend of the difference of the average utility under the two funding schemes.

[25] We set this percentage equal to the average-low interest rate for a typical mortgage loan e.g. around 5%.

Tab. 1.1: Net Social Benefit

$r = 5\%$, $\rho = 8\%$	$NB_{ICL} - NB_{ML}$		
Costs	All	Man	Woman
Pre-reform	2.843	1.198	3.506
Post-reform	15.252	10.707	18.165

Tab. 1.2: Annual Gross Earnings Graduates

	Mean	**Std. Dev.**	**N**	**Percent**
male	27897.81	22576.99	623	52.93
female	19665.85	10406.5	554	47.07
Total	24023.12	18368.97	1177	100.00

Tab. 1.3: Marital Status

	Mean	**Std. Dev.**	**N**	**Percent**
single/divorced	23345	15786.7	437	37.13
married/cohabiting	24423.58	19735.23	740	62.87
Total	24023.12	18368.97	1177	100.00

Tab. 1.4: children 0-16 years old

	Mean	Std. Dev.	N	Percent
no kid/ non answ	24973.56	17884.25	913	77.57
1 or more children	20736.22	19638.16	264	22.43
Total	24023.12	18368.97	1177	100.00

Tab. 1.5: Family earnings 1980

	Mean	Std. Dev.	N	Percent
not stated	25383.69	20007.07	144	12.23
low	32384.32	56743.81	25	2.12
medium	23052.64	16881.78	759	64.49
high	25355.04	13181.55	249	21.16
Total	24023.12	18368.97	1177	100.00

Tab. 1.6: Mother qualifications 1980

	Mean	Std. Dev.	N	Percent
no quals	22306.22	18956.85	354	30.08
O-level	25773.47	22907.52	265	22.51
degree	27149.5	15615.49	72	6.12
other quals	23856.14	15164.33	486	41.29
Total	24023.12	18368.97	1177	100.00

Tab. 1.7: Father social class 1986

	Mean	Std. Dev.	N	Percent
missed/not stated	22796.19	15091.37	275	23.36
profes/interm	25969.63	19536.18	491	41.72
skilled occupation	21821.74	12973.58	364	30.93
other occupation	27916.26	42101.93	47	3.99
Total	24023.12	18368.97	1177	100.00

Tab. 1.8: Degree subjects

	Mean	Std. Dev.	N	Percent
does not apply	21249.42	14876.29	554	47.07
Sciences	26782.23	16828.23	292	24.81
Social Sciences	25857.94	21385.41	146	12.40
Art and humanities	26526.32	25277.39	185	15.72
Total	24023.12	18368.97	1177	100.00

Tab. 1.9: Job sector

	Mean	**Std. Dev.**	**N**	**percent**
other / not answ	18456.66	6864.195	83	7.05
private firm	26434.32	21703.36	736	62.53
public sector	20356.57	9910.831	358	30.42
Total	24023.12	18368.97	1177	100.00

Tab. 1.10: Annual gross initial earnings

	Mean	**Std. Dev.**	**N**
low (5th perc)	8640.983	2652.527	40
high (95th perc)	53002.737	20505.394	38
total	23987.433	13147.324	405

Tab. 1.11: Increasing Income $AU_{ICL} - AU_{ML}$ - ρ and σ changing

	$\varphi = £1000$	$C = £9000$	$T_{ML} = 9$
	$\gamma = 9\%$	$\lambda = 1\%$	$ra = 0.5$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$\rho = 8\%$	67.2721	67.2994	67.4311
$\rho = 15\%$	51.5189	51.5387	51.6331
$\rho = 30\%$	32.7445	32.7559	32.8088
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$\rho = 8\%$	-0.3704	-0.2421	0.0227
$\rho = 15\%$	0.7800	0.8651	1.1011
$\rho = 30\%$	1.4404	1.4808	1.6262

Tab. 1.12: Increasing Income $AU_{ICL} - AU_{ML}$ - φ and σ changing

		$C = £9000$	$\rho = 8\%$
	$\gamma = 9\%$	$\lambda = 1\%$	$ra = 0.5$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$\varphi = £500$ $T_{ML} = 18$	48.8080	48.8332	48.9630
$\varphi = £1000$ $T_{ML} = 9$	67.2721	67.2994	67.4311
$\varphi = £3000$ $T_{ML} = 3$	90.6606	90.6905	90.8231
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$\varphi = £500$ $T_{ML} = 18$	-10.9630	-10.8355	-10.5687
$\varphi = £1000$ $T_{ML} = 9$	-0.3704	-0.2421	0.0227
$\varphi = £3000$ $T_{ML} = 3$	10.9286	11.0558	11.3084

Tab. 1.13: Increasing Income $AU_{ICL} - AU_{ML}$ - ra and σ changing

	$\varphi = £1000$	$C = £9000$	$T_{ML} = 9$
	$\gamma = 9\%$	$\lambda = 1\%$	$\rho = 8\%$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$ra = 0.25$	6.9826	6.9870	7.0090
$ra = 0.5$	67.2721	67.2994	67.4311
$ra = 1.2$	38271.00	38265.00	38242.00
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$ra = 0.25$	-0.0224	-0.0121	0.0106
$ra = 0.5$	-0.3704	-0.2421	0.0227
$ra = 1.2$	-743.0753	-594.7867	-340.0252

Tab. 1.14: Increasing Income $AU_{ICL} - AU_{ML}$ - λ and σ changing

	$\varphi = £1000$	$C = £9000$	$T_{ML} = 9$
	$\gamma = 9\%$	$\rho = 8\%$	$ra = 0.5$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$\lambda = 1\%$	67.2721	67.2994	67.4311
$\lambda = 2.4\%$	65.3690	65.2274	63.9582
$\lambda = 4\%$	54.9972	55.0309	55.2131
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$\lambda = 1\%$	-0.3704	-0.2421	0.0227
$\lambda = 2.4\%$	-2.3783	-2.2142	-1.7014
$\lambda = 4\%$	-2.8460	-2.8486	-3.2024

Tab. 1.15: Increasing Income $AU_{ICL} - AU_{ML}$ - ICL threshold and σ changing

	$\varphi = £1000$	$C = £9000$	$T_{ML} = 9$
	$\gamma = 9\%$	$\rho = 8\%$	$ra = 0.5$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
threshold=0	4.1188	3.8858	3.2633
threshold=£10000	67.2721	67.2902	65.6918
threshold=£15000	67.2721	67.2994	67.4311
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
threshold=0	-18.2505	-16.8020	-14.5048
threshold=£10000	-5.6606	-5.6371	-6.1650
threshold=£15000	-0.3704	-0.2421	0.0227

Tab. 1.16: Increasing Income $AU_{ICL} - AU_{ML}$ - Effects English Reform

	$\varphi = £1000$	$T_{ML} = 9$	
	$\gamma = 9\%$	$\lambda = 1\%$	$\rho = 8\%$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$cost - pre$	2244	2268	2278
$cost - post$	6246	6101	6022
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$cost - pre$	-539.46	-562.13	-615.31
$cost - post$	-411.55	-311.82	-304.39

Tab. 1.17: Participation in HE (% of individuals)- Difference ICL-ML

	$\varphi = £1000$	$T_{ML} = 9$	
	$\gamma = 9\%$	$\lambda = 1\%$	$\rho = 8\%$
Low Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$cost - pre$	0	7.6	8.5
$cost - post$	0.3	26.0	23.0
High Income			
	$\sigma = 2\%$	$\sigma = 5\%$	$\sigma = 15\%$
$cost - pre$	0	-0.8	-0.5
$cost - post$	0	-0.1	-0.5

Fig. 1.1: $EC_{ML} - EC_{ICL}$ - Expected costs and Risk Neutrality

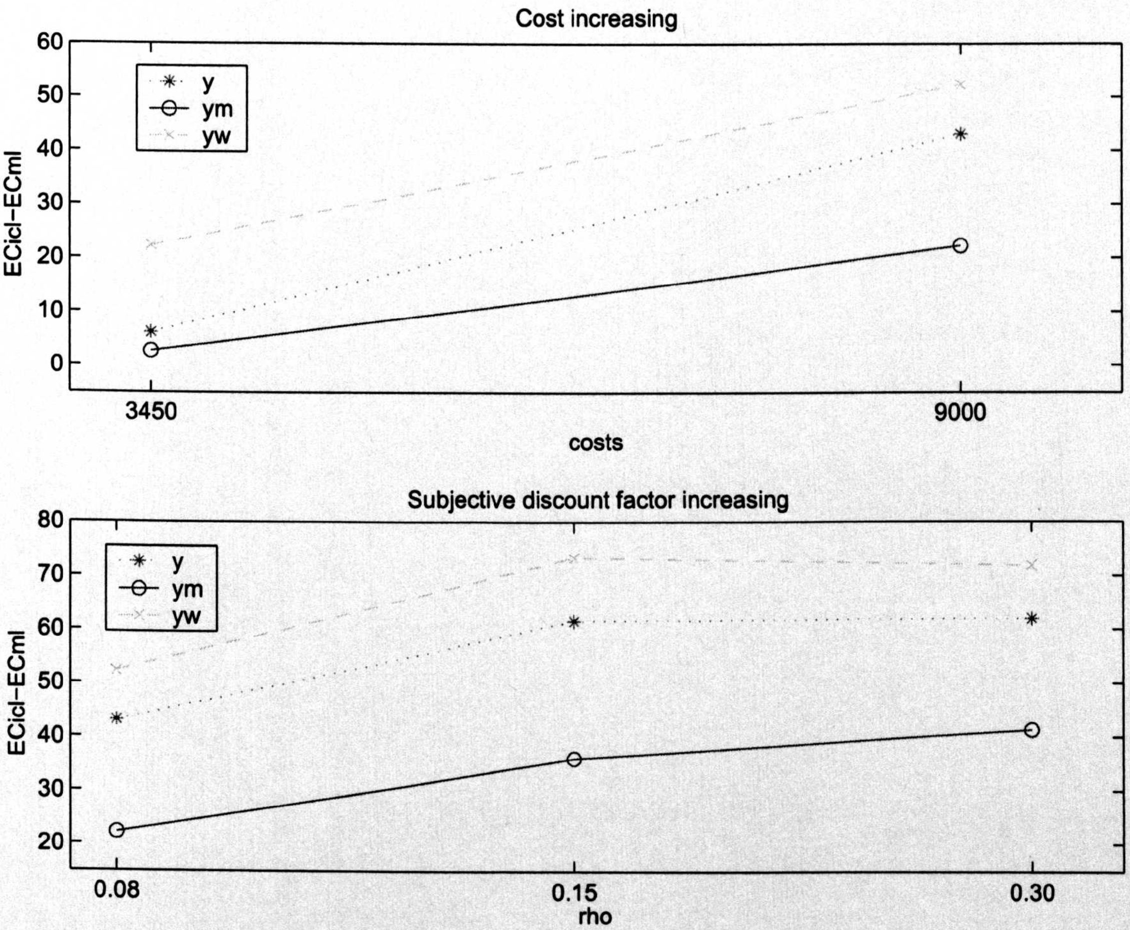


Fig. 1.2: $EU_{ICL} - EU_{ML}$ - Individual Characteristics

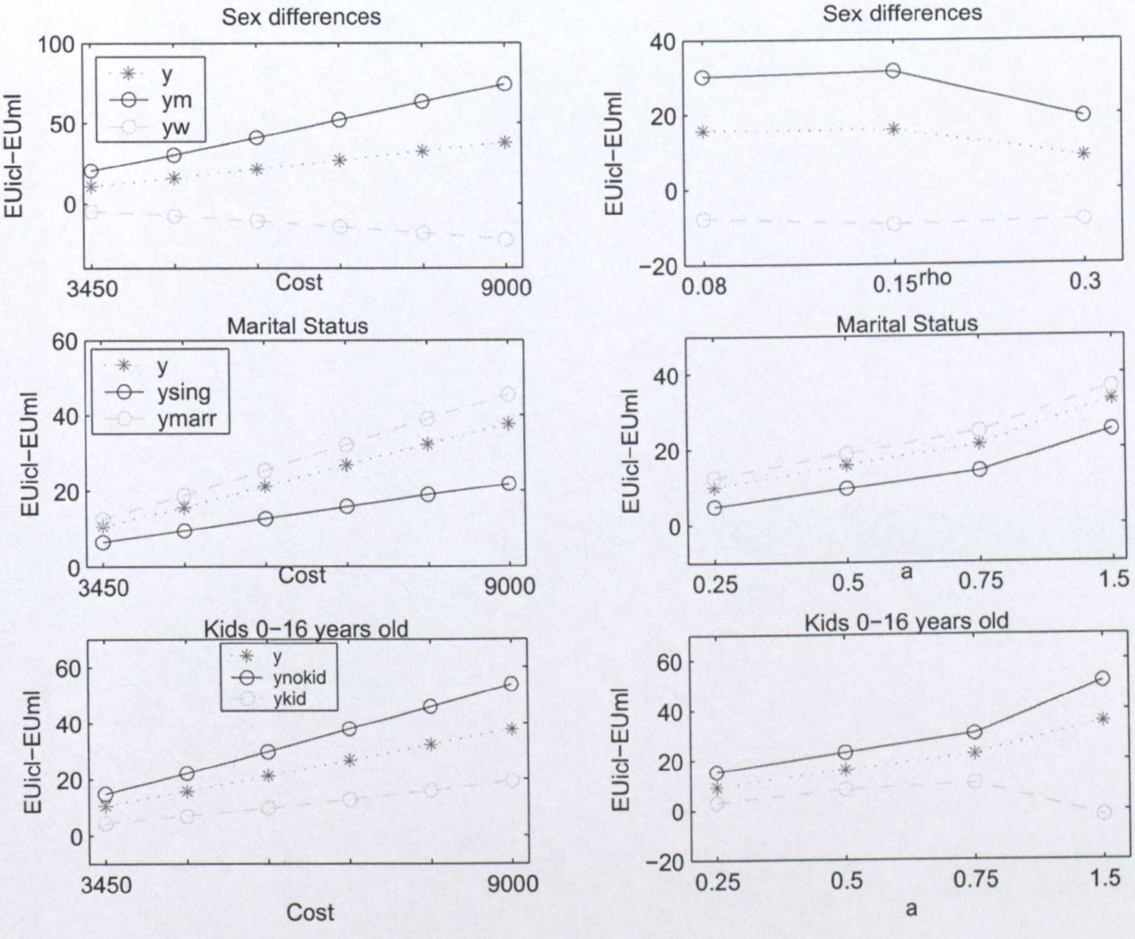


Fig. 1.3: $EU_{ICL} - EU_{ML}$ - Family Background

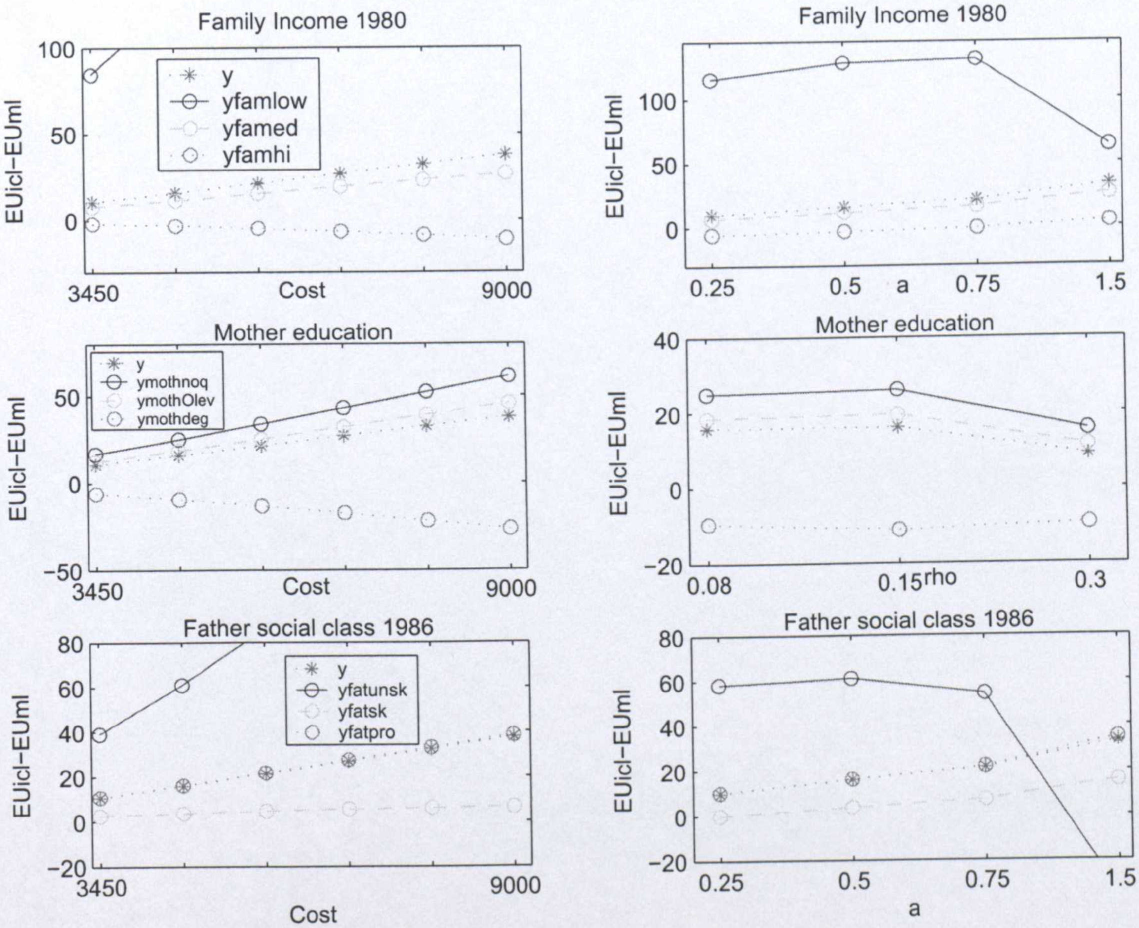


Fig. 1.4: $EU_{ICL} - EU_{ML}$ - Degree Subjects

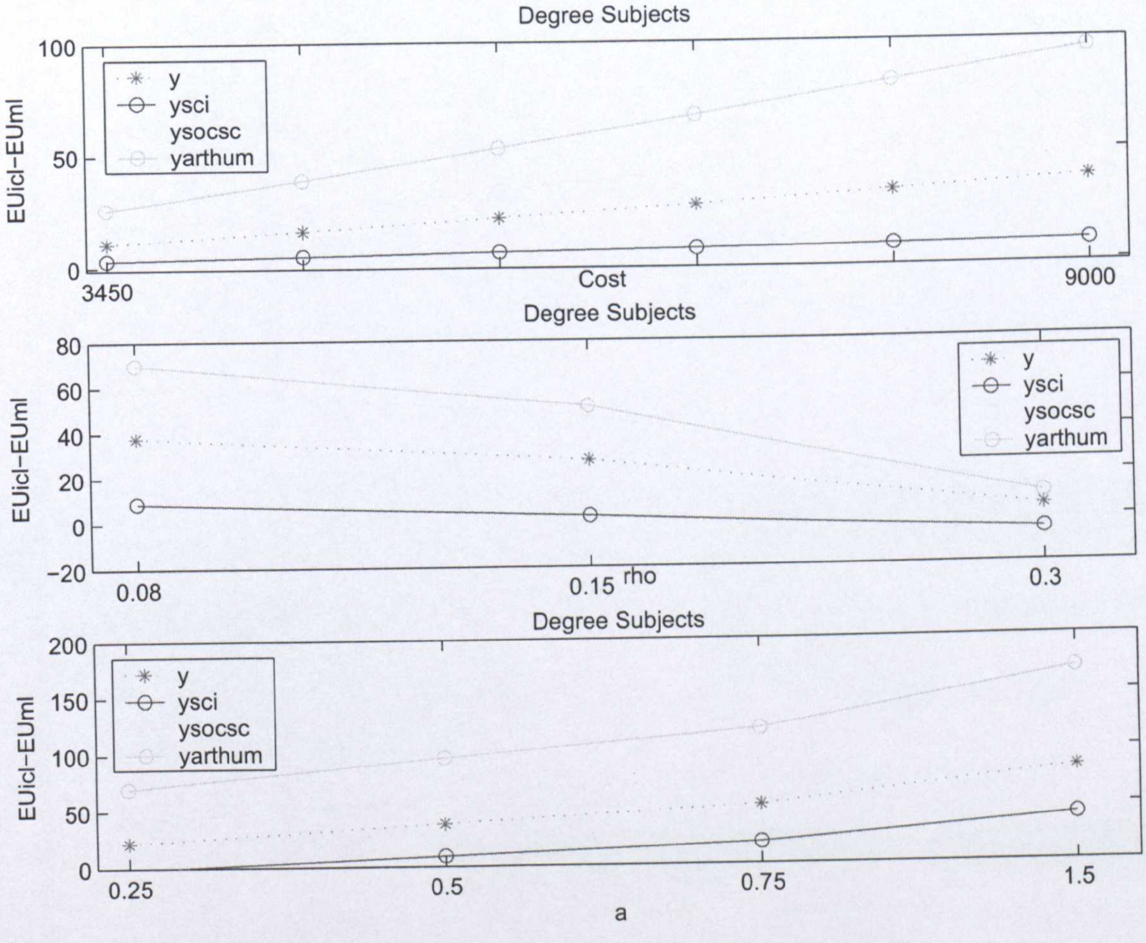
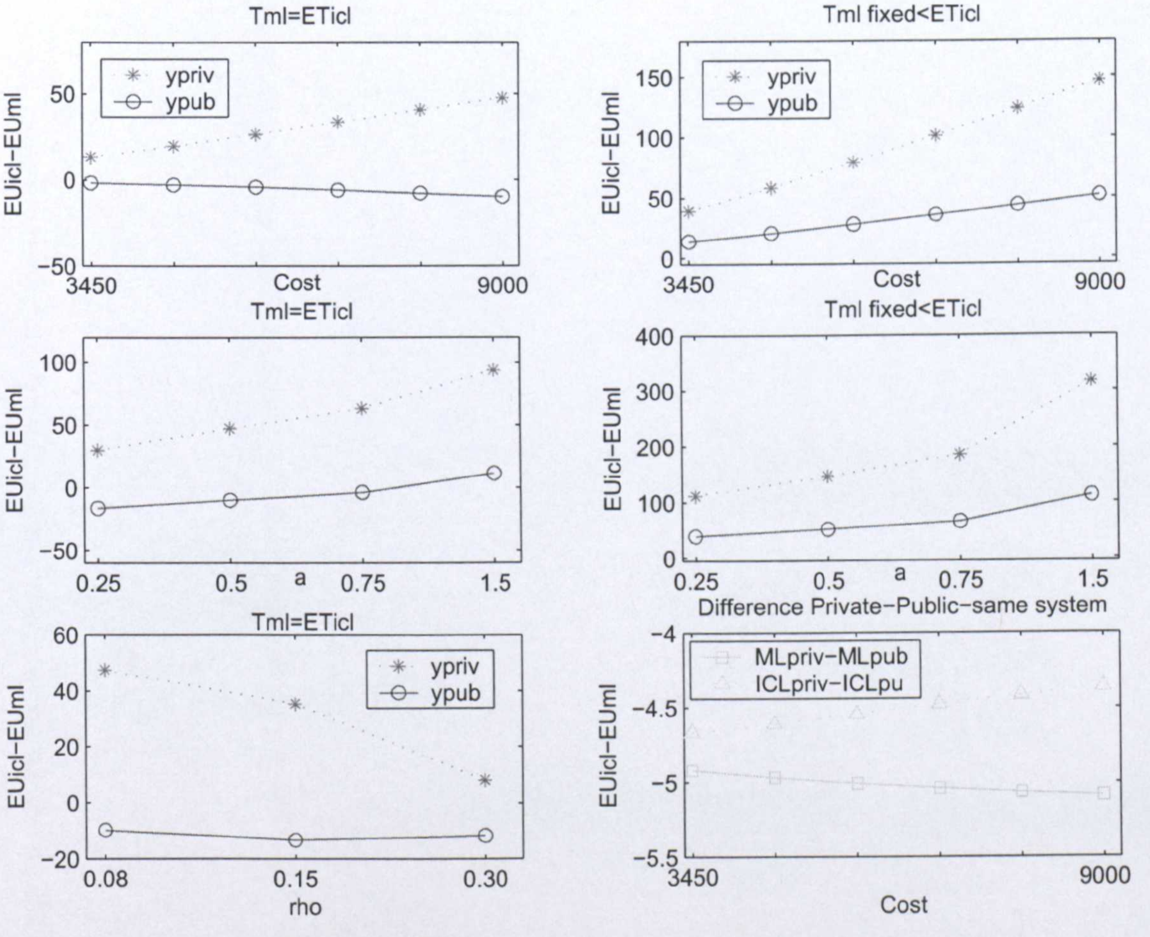


Fig. 1.5: $EU_{ICL} - EU_{ML}$ - Public versus Private sector



CHAPTER 2

2. EFFORT, DEFAULT AND ALTERNATIVE HIGHER EDUCATION FUNDING SCHEMES

2.1 Introduction

In this chapter we try to address a common problem in HE financing systems that adopt student loans: some of the graduates may not repay the complete debt and so default occurs[1]. As noted by Jacobs (2000) there are two ways to protect against graduates' default: risk pooling among students and risk shifting to society. In the first case the risk of default is shared among graduates. For example, the interest rate on the loan could be increased by a premium to cover the cost of default of people not able to repay. However, in this case the successful students bear the costs of non repayment of people on default. With a risk pooling arrangement the usual problems of asymmetric information may arise when providing insurance. In fact, if the government cannot observe the riskiness type of the students, the good risk types may not participate in the program to avoid the additional premium caused by the bad risks (see Nerlove, 1975). The proportion of risky students increases and so do the repayment rates. The adverse selection problem may ultimately reduce investment in higher education.

If the government instead cannot observe the actions of the students and graduates, moral hazard occurs. Jacobs (2000) considers some types of moral hazard under an income contingent system with a default premium. For example, graduates may reduce labour supply or work effort in order to avoid earning enough to repay their educational debt. Moreover, students may reduce their learning efforts while at college since they know that the costs of default are borne by their fellow students. The moral hazard raises the default premium and makes an ICL system less attractive than it would otherwise be. In addition if there are unobserved ability differences between students then a higher default premium will cause an adverse selection problem whereby the high ability students do not

[1] By default we mean inability to pay, not unwillingness to pay.

participate. In contrast financing systems where the risk of default is borne by society as whole, there will be no redistribution from the hard working to the lazy students, except for the fact that the payments of graduates that pay off their debt contribute to the government budget from which the cost of default are paid. Examples of risk shifting arrangements are the English and the Australian higher education financing schemes.

In the model presented in Chapter 1 we assumed that students could not default; here we illustrate a model that is an extension of our basic work since it considers the possibility that graduates cannot completely repay their debt. Under a ML default happens if the loan is not repaid by the fixed repayment period (5 years under the old English system), and the government starts a recovery procedure. Under an ICL, default usually implies that the debt is written off. This means if the graduates do not make an income exceeding the threshold (£15000 in England) for any reason including working overseas, unemployment or sickness, he or she does not make any repayments that year. After a fixed number of years (25 in England or when the graduate is 50 years old or 60 if he or she was over 40 at time of last borrowing) any remaining debt is written off. In a recent project edited by Schwarzenberger (2008) on funding higher education in Europe, we find a simulation of the potential rates of non repayment of the ICL loan in England. Since this new student loan scheme has been recently introduced, there is no empirical information on repayments yet. Looking at Figure 2.1 we notice that for a loan of £9000, that covers only the fees for a three-year undergraduate course, the expected rate of non repayment is around 4%. For higher amounts of the loan we observe that the non repayment rate increases dramatically reaching a value above 28% for a loan of £25000.

In our work we present a theoretical model to take into account of default, and it is the first to our knowledge that explicitly compare a ML and an ICL (see e.g. Nerlove, 1975, Garcia Penalosa and Walde, 1999, Jacobs, 2000). For a ML we maintain the English structure for the loan, for an ICL we need some simplifying assumption for computational purposes. At this stage of our work we do not consider the initial threshold and we assume the debt's write-off coinciding with the end of the working life. Therefore under an ICL when the individual gets a low income, the repayments over the whole of their working life are not sufficient to completely repay the loan. We still have a non repaid part of the loan which is effectively written off, and the problem of default, in our view does not loose generality.

We structure the model as a simultaneous game between students and government, where the students decide the level of effort to put into their studies, where high effort results in high probability to get a good qualification and high earnings, and the government chooses the default premium that keeps its budget balanced. By comparing the individuals payoff under all the possible cases and taking into account the participation constraint, we derived, analytically, the cutoffs for the costs of effort required to implement the Nash equilibria. The equilibria are high effort, low effort or multiple equilibria. We play the same game for a ML and an ICL. In general, we find that under an ICL the HE participation rate is higher, that is students prefer putting more effort to stay in higher education than under a mortgage loan. Moreover if the objective of the government is to increase the students' effort and reduce the default, an ICL fulfills better this objective. Students, under an ICL, are willing to bear a higher cost of effort in order to get a high income. This because, according to our model, the loss they will have in case of default is higher under an ICL than under a ML.

2.2 Assumptions

Individuals going into higher education can choose two levels of effort: high $\overline{e}$ and low $\underline{e}$. If they choose $\overline{e}$ they get a high wage $\overline{y}$ with certainty. If they choose $\underline{e}$ they obtain a low wage $\underline{y}$ with probability p, and a high wage $\overline{y}$ with probability $1-p$.

$$\begin{cases} \overline{e} \Rightarrow \quad \overline{y} \text{ with } P(\overline{y} \mid \overline{e}) = 1 \ ; \\ \\ \underline{e} \Rightarrow \quad \begin{cases} \underline{y} \quad \text{with } P(\underline{y} \mid \underline{e}) = p \ ; \\ \\ \overline{y} \quad \text{with } P(\overline{y} \mid \underline{e}) = 1 - p \ . \end{cases} \end{cases}$$

Students making high effort during the schooling period bear a psychological cost Δ, for those exerting a low level of effort this cost is zero:

$$\begin{cases} \bar{e} \Rightarrow \quad \Delta > 0; \\ \\ \underline{e} \Rightarrow \quad \Delta = 0\ . \end{cases}$$

The government finances investment in higher education by issuing debt that is paid back only with the graduates' repayments. The students take out a loan of fixed size that cover all the costs of the university. The government decides the repayment scheme before the student decision to attend higher education or not. The government can choose between the two financing systems: Mortgage Loan and Income Contingent Loan. Each scheme is fully funded, student participation in the chosen scheme is obligatory and repayments cannot be avoided.

We assume that the graduates have a maximum period, coinciding with the end of their working life, of T_{max} years to pay off the debt, after which any repayment is stopped. T_{max} is a fixed parameter, and retirement is exogenous in this model[2]. All individuals study for s years. If after graduating they get a high wage $\overline{y}$, they fully repay their loan within a period T less than T_{max}. Instead if they receive a low wage $\underline{y}$, they are unable to pay off the debt within T_{max}, so they are in default. Assuming a general per period repayment R we have two cases:

$$\begin{cases} \overline{y} \quad \Rightarrow \quad \int_s^{T+s} (\overline{y} - R)\, e^{-\rho\tau} d\tau > 0\ ; \\ \\ \underline{y} \quad \Rightarrow \quad \int_s^{T_{max}} \left(\underline{y} - R\right) e^{-\rho\tau} d\tau < 0 \rightarrow \text{default}\ . \end{cases}$$

We assume that the government wants to balance the budget, it is risk neutral and has a zero real interest on the loan. To avoid a deficit caused by the students' default it has the option to add a premium to the loan: $\delta \in [0, \infty)$. The model is structured as a game with the following stages.

[2] If we try to incorporate any labour supply decision e.g. endogenous retirement we have the problem that an ICL because it is an income tax would induce an income substitution effect of opposite sign thus any comparative static proposition would become ambiguous.

1. The government chooses the repayment scheme: ML or ICL.

2. Given the government's decision of the loan scheme, each student decides whether to participate in HE or not. If they do not, they get a payoff normalized to zero. If they do, play moves to stage 3.

3. The students, who have chosen to participate, choose a level of effort ($\bar{e}$ or $\underline{e}$), and simultaneously, the government chooses a default premium $0 \leq \delta \leq \infty$.

Solve backwards:

- game at stage 3 given ML;

- game at stage 3 given ICL.

We remark that the students are N identical rational agents, and when choosing their action they are unaware of the choice of the government. This means that the latter rationally anticipates the level of effort of the students. The government imposes a positive premium if it expects a low effort choice by the students, otherwise if it expects a high effort choice the public budget is balanced without any premium. In the first case the students have to pay back $C(1+\delta)$, in the second case just C. The third stage is similar to the well known game "The Battle of the Sexes" with multiple equilibria. In our game we find two Nash equilibria: $(\bar{e}, \delta = 0)$ and $(\underline{e}, \delta > 0)$.

Equilibrium

Given a continuum of players $[0,1]$

$$e^* \in \{\underline{e}, \bar{e}\} \quad \text{and} \quad \delta^* \in [0, \infty\}$$

s.t

$$(i) \qquad e^* \ \text{ is optimal given } \ \delta^*$$

$$(ii) \qquad \delta^* \ \text{ achieves budget balance given } \ e^*.$$

2.3 Payoffs

We derive the students' payoffs under the two levels of effort and for the two different repayment schemes. Since the individuals are identical, the choice of the effort involves all of them at the same time. Once they have chosen the effort they cannot change their decision.

2.3.1 Payoff with a Mortgage Loan

We start with a mortgage loan, and we define T^{ML} the fixed repayment period, C is the fixed cost of the education (i.e. the loan), φ is the fixed installment. Finally, Δ is a positive cost of effort paid only by the students choosing high effort, we can think of it as a psychic cost required to get a good degree.

Payoff under High effort

The students choosing high effort receive a high income $\bar{y}$ for all their working life and they repay the cost of their education, C, in T^{ML} years. The repayment period is $T^{ML} = \frac{C}{\varphi}$, where C and φ are fixed. By assumption $T^{ML} < T^{\max}$, therefore under the choice of high effort there is no default. After repaying the loan they receive the full income until they retire, that is from T^{ML} to $T^{\max}$.

Defining PH^{ML} as the payoff of high effort with a mortgage loan, we have two cases according to the government choices: $(PH^{ML}|\delta = 0)$ and $(PH^{ML}|\delta > 0)$. Note that the two payoffs are not equal, when the government adds a positive default premium the cost of the loan is higher and consequently the repayment period longer.

We derive the payoff in general[3]:

$$\begin{aligned} PH^{ML} &= \int_0^{T^{ML}} (\bar{y} - \varphi - \Delta)\, e^{-\rho\tau} d\tau + \int_{T^{ML}}^{T^{\max}} \bar{y} e^{-\rho\tau} d\tau \\ &= (\bar{y} - \varphi - \Delta)\frac{1}{\rho}[1 - e^{-\rho T^{ML}}] + \frac{\bar{y} e^{-\rho T^{ML}}}{\rho}[1 - e^{-\rho(T^{\max} - T^{ML})}]. \end{aligned} \tag{2.1}$$

[3] For simplicity we put Δ inside the integral and we normalize to zero the initial year of schooling, s=0.

If the government sets a zero default premium, the cost of loan is C and the repayment period is $T^{ML} = \frac{C}{\varphi}$. If the government chooses a positive default premium, the cost of loan is $C(1+\delta)$ and the repayment period is $T^{ML} = \frac{C(1+\delta)}{\varphi}$.

Payoff under low effort

If the students choose $\underline{e}$, they will receive a high income $\bar{y}$ with probability $(1-p)$ and a low income $\underline{y}$ with probability p. Both types of wages are fixed and constant for all the working life. If the students obtain $\bar{y}$ they are able to fully repay their loan in $T^{ML} < T^{\max}$ years, and there is no default. If they receive $\underline{y}$, this wage is not sufficient to fully repay the loan during the repayment period T^{ML}. This implies a default. We assume that the low income is lower than the ML installment φ:

$$if \quad p > 0 \iff \underline{y} < \varphi.$$

The repayment of the loan is compulsory and graduates cannot avoid the payments, therefore we assume they will repay each year an amount smaller than φ and exactly equal to their income $\underline{y}$, until the end of the repayment period. This means they repay $\underline{y} \times T^{ML}$ years. After T^{ML} years the cost of the loan is not fully repaid and the individuals are in default:

$$\underline{y} \times T^{ML} < C(1+\delta)$$

replacing $T^{ML} = \frac{C(1+\delta)}{\varphi}$

$$\underline{y} \times \frac{C(1+\delta)}{\varphi} < C(1+\delta) \iff \underline{y} < \varphi$$

If this condition does not hold because the earned income is higher than the ML installment then there is no default.

Example

Consider the fees fixed by the English government for a 3-year degree, $C = £3000$ per year, and $\delta = 2\%$. If the installment of a ML is set to $\varphi = £1836$ per year, the repayment period is $T^{ML} = 5$ years. The total cost to repay is £9180 (it includes the default premium).

The English HE financing system before the 2004 Reform was based on a mortgage loan with repayment in 60 monthly installment.

Our default condition implies that the earned income must be lower than £1836 per year; assuming for example a low income to be equal to $\underline{y} = £1500$ per year, we have

$$\underline{y} \times T^{ML} < C(1+\delta) \Longleftrightarrow 1500 \times 5 = 7500 < 9000(1+0.02) = 9180 \Longrightarrow default$$

Its is clear that default (which here simply means an inability to pay the installment) occurs when income is less than the installment. Thus, given a 5-year repayment period, someone with an income of £15000 would not default unless the loan size had been at least more than £75000 - not an empirically relevant magnitude.

When individuals are in default, according to the English rules, the government starts a debt recovery procedure. We model this default procedure imposing a fixed penalty D on the income ($\underline{y}$) received after the end of the repayment period T^{ML}.

Let PL^{ML} be the payoff under low effort with a ML. According to the government's move we can have: $(PL^{ML}|\delta = 0)$ and $(PL^{ML}|\delta > 0)$. We derive the payoff in general.

$$\begin{aligned} PL^{ML} = (1-p)[\int_0^{T^{ML}} (\bar{y} - \varphi)\, e^{-\rho\tau} d\tau + \int_{T^{ML}}^{T^{\max}} \bar{y} e^{-\rho\tau} d\tau] \\ + p\,[\int_0^{T^{ML}} (\underline{y} - \varphi)\, e^{-\rho\tau} d\tau + \int_{T^{ML}}^{T^{\max}} (\underline{y} - D) e^{-\rho\tau} d\tau]. \end{aligned} \tag{2.2}$$

Since $\underline{y} < \psi$ the expression $\int_0^{T^{ML}} (\underline{y} - \varphi)\, e^{-\rho\tau} d\tau$ cancels out and we get

$$\begin{aligned} PL^{ML} = (1-p)[(\bar{y} - \varphi)\frac{1}{\rho}(1 - e^{-\rho T^{ML}}) + \frac{\bar{y} e^{-\rho T^{ML}}}{\rho}(1 - e^{-\rho(T^{\max} - T^{ML})}) \\ + p(\underline{y} - D)\frac{e^{-\rho T^{ML}}}{\rho}(1 - e^{-\rho(T^{\max} - T^{ML})}). \end{aligned} \tag{2.3}$$

If the government chooses $\delta = 0$ the repayment period is $T^{ML} = \frac{C}{\varphi}$, while if the government chooses a positive default premium the repayment period is $T^{ML} = \frac{C(1+\delta)}{\varphi}$.

2.3.2 Payoff with an ICL

Under an income contingent loan the graduates repay each year a fixed percentage γ of their income. We define T^{ICL} to be the repayment period, C is the fixed cost of the education (i.e. the loan), Δ is a positive cost of effort paid only by the students choosing high effort, and δ is the default premium chosen by the government.

Payoff under high effort

The students choosing high effort receive a high income $\bar{y}$ for all their working life, and they pay each year a fixed percentage γ of $\bar{y}$ until the loan is paid off. In order to fully repay the loan the following condition must be satisfied

$$C = \gamma \int_0^{T^{ICL}} \bar{y} d\tau = \gamma \bar{y} T^{ICL}$$

Therefore the repayment period is $T^{ICL} = \frac{C}{\gamma \bar{y}}$ years, where C and γ are fixed. By assumption $T^{ICL} < T^{\max}$, and under the choice of high effort there is no default. After the repayment of the loan the graduates will get their full income until they retire, that is from T^{ICL} to $T^{\max}$.

Let PH^{ICL} be the payoff of high effort with an income contingent loan, we have two cases according to the government choices: $(PH^{ICL}|\delta = 0)$ and $(PH^{ICL}|\delta > 0)$. Note that the two payoffs are not equal, when the government adds a positive default premium the cost of the loan is higher and consequently the repayment period longer.

Deriving the payoff in general:

$$\begin{aligned} PH^{ICL} &= \int_0^{T^{ICL}} [\bar{y}(1-\gamma) - \Delta] e^{-\rho\tau} d\tau + \int_{T^{ICL}}^{T^{\max}} \bar{y} e^{-\rho\tau} d\tau \\ &= [\bar{y}(1-\gamma) - \Delta]\frac{1}{\rho}[1 - e^{-\rho T^{ICL}}] + \bar{y}\frac{1}{\rho}[e^{-\rho T^{ICL}} - e^{-\rho T^{\max}}] \\ &= [\bar{y}(1-\gamma) - \Delta]\frac{1}{\rho}[1 - e^{-\rho T^{ICL}}] + \bar{y}\frac{e^{-\rho T^{ICL}}}{\rho}[1 - e^{-\rho(T^{\max} - T^{ICL})}]. \end{aligned} \tag{2.4}$$

If the government chooses a zero default premium, the cost of the loan is C and the repayment period is $T^{ICL} = \frac{C}{\gamma\bar{y}}$. If the government chooses a positive default premium, the cost of loan is $C(1+\delta)$ and the repayment period is $T^{ICL} = \frac{C(1+\delta)}{\gamma\bar{y}}$.

Payoff under low effort

If the students choose $\underline{e}$, they will receive a high income $\bar{y}$ with probability $(1-p)$ and a low income $\underline{y}$ with probability p. Both types of wages are fixed and constant for all the working life. If the students obtain $\bar{y}$ they are able to fully repay their loan in $T^{ICL} < T^{\max}$ years, and there is no default. If they receive $\underline{y}$, this wage is not sufficient to fully repay the loan during the working life, that is within $T^{\max}$ years. In fact, we assume that the repayment period required for a full repayment of the debt is longer than the working life[4]:

$$T_{fr}^{ICL} = \frac{C(1+\delta)}{\gamma\underline{y}} > T^{\max}$$

This is the default condition under an income contingent loan, the debt is written off at the end of the working life. The main difference with a mortgage loan is that graduates must repay a fixed percentage γ of their low income $\underline{y}$ until retirement, and the sum of all these payments is not enough to repay the cost of their education. Relatively to the current structure of the ICL, where after a certain number years (25) the loan is written off, we have the simplifying assumption that the write-off coincides with the retirement. However, we want to stress the point that the loan remains still unpaid and there is a loss for the government.

Let PL^{ICL} be the payoff under low effort with a ML, according to the government's move we can have: $(PL^{ICL}|\delta = 0)$ and $(PL^{ICL}|\delta > 0)$. We derive the payoff in general.

[4] fr stands for full repayment.

$$
\begin{aligned}
PL^{ICL} &= (1-p)\{\int_0^{T^{ICL}} \bar{y}(1-\gamma)e^{-\rho\tau}d\tau + \int_{T^{ICL}}^{T^{\max}} \bar{y}e^{-\rho\tau}d\tau\} \\
&+ p\{\int_0^{T^{\max}} (\underline{y} - \gamma\underline{y})e^{-\rho\tau}d\tau\} \\
&= (1-p)\{\bar{y}(1-\gamma)\frac{1}{\rho}[1-e^{-\rho T^{ICL}}] + \bar{y}\,\frac{e^{-\rho T^{ICL}}}{\rho}[1-e^{-\rho(T^{\max}-T^{ICL})}]\} \\
&+ p\,\underline{y}(1-\gamma)\frac{1}{\rho}[1-e^{-\rho T^{\max}}].
\end{aligned}
\tag{2.5}
$$

If the government chooses $\delta = 0$, the repayment period is $T^{ICL} = \frac{C}{\gamma\bar{y}}$. If the government chooses $\delta > 0$, the repayment period is $T^{ICL} = \frac{C(1+\delta)}{\gamma\bar{y}}$.

Notation

In equations (2.1) and (2.3) we define:

$$
\begin{aligned}
\alpha &= 1 - e^{-\rho T^{ML}} \\
\beta &= 1 - e^{-\rho(T^{\max} - T^{ML})}.
\end{aligned}
$$

if $T^{ML} = \frac{C}{\varphi}$. When $T^{ML} = \frac{C(1+\delta)}{\varphi}$ we redefine the two expressions above α_1 and β_1.

In equations (2.4) and (2.5) we define:

$$
\begin{aligned}
\epsilon &= 1 - e^{-\rho T^{ICL}} \\
\theta &= 1 - e^{-\rho(T^{\max} - T^{ICL})} \\
\iota &= 1 - e^{-\rho T^{\max}}.
\end{aligned}
$$

if $T^{ICL} = \frac{C}{\gamma\bar{y}}$. When $T^{ICL} = \frac{C(1+\delta)}{\gamma\bar{y}}$ we redefine the first two expressions above ϵ_1 and θ_1.

2.4 Equilibria

To implement the two Nash equilibria, which are the solutions of the game described in Section 2.2, we have to compare the payoffs analytically derived in the Section 2.3 given the different strategies of students and government. In particular, to obtain a high effort equilibrium $(\bar{e}, \delta = 0)$ we require that $(PH|\delta = 0) > (PL|\delta = 0) > 0$. To obtain a low effort equilibria $(\underline{e}, \delta > 0)$, we require $(PL|\delta > 0) > (PH|\delta > 0) > 0$. We also include a participation constraint, that is under each equilibria the lowest payoff has to be greater than zero. We assume that students do not attend higher education if they do not get a positive return, and therefore there is no game in the second stage.

In order to derive the equilibrium conditions we solve for the cost of effort Δ. Within both financing systems we derive two levels of cost of effort, i.e. for high effort and low effort equilibria.

2.4.1 High effort equilibrium under ML

To implement a high effort equilibrium under a mortgage loan, we require

$$(PH^{ML}|\delta = 0) > (PL^{ML}|\delta = 0) \geq 0 \tag{2.6}$$

Equation (2.6) implies two conditions:

1. participation constraint $(PL^{ML}|\delta = 0) \geq 0$: in equilibrium the lowest payoff cannot be negative otherwise the students do not go to college;

2. payoff condition $(PH^{ML}|\delta = 0) > (PL^{ML}|\delta = 0)$: the payoff to the students choosing high effort has to be higher than the payoff to those choosing low effort, given the strategy of the government of choosing a zero default premium.

Proposition 1. *Under a ML there is high effort equilibrium if*

$$\Delta < \Delta(\bar{e})^{ML}$$

$$0 < \alpha \leq 1$$

$$\beta \geq 0.$$

where

$$\Delta(\bar{e})^{ML} = p[(\bar{y} - \varphi) + (\bar{y} - (\underline{y} - D))\frac{(1-\alpha)}{\alpha}\beta]$$

Proof

Knowing that the repayment period is $T^{ML} = \frac{C}{\varphi}$, we consider equation (2.3) and set the participation constraint:

$$(1-p)[(\bar{y} - \varphi)\alpha + \bar{y}(1-\alpha)\beta] + p(\underline{y} - D)(1-\alpha)\beta \geq 0.$$

Given that $\bar{y} > \varphi$, $\underline{y} \geq D$, $1 \leq p \leq 0$ by assumption, the second and third conditions are sufficient to show straightforwardly that the participation constraint holds.
To prove the payoff condition we set equation (2.1) greater than equation (2.3):

$$\begin{aligned}&(\bar{y} - \varphi - \Delta)\frac{\alpha}{\rho} + \frac{\bar{y}}{\rho}(1-\alpha) > \\ &(1-p)[(\bar{y} - \varphi)\frac{\alpha}{\rho} + \bar{y}\frac{1-\alpha}{\rho}\beta] + p(\underline{y} - D)\frac{1-\alpha}{\rho}\beta.\end{aligned}$$

we solve for the cost of effort Δ and after some calculation we get

$$\Delta\alpha < \alpha p(\bar{y} - \varphi) + p\bar{y}(1-\alpha)\beta - p(\underline{y} - D)(1-\alpha)\beta$$

then we can easily obtain on the right hand side $\Delta(\bar{e})^{ML}$.

2.4.2 Low effort equilibrium under ML

To implement a low effort equilibrium under a mortgage loan, we require

$$(PL^{ML}|\delta > 0) > (PH^{ML}|\delta > 0) \geq 0 \tag{2.7}$$

Equation (2.7) implies two conditions:

1. participation constraint $(PH^{ML}|\delta > 0) \geq 0$: in equilibrium the lowest payoff cannot be negative otherwise the students do not go to college;

2. payoff condition $(PL^{ML}|\delta > 0) > (PH^{ML}|\delta > 0)$: the payoff to the students choosing low effort has to be higher than the payoff to those choosing high effort, given the strategy of the government of choosing a positive default premium.

Proposition 2. *Under a ML there is low effort equilibrium if*

$$\Delta(\underline{e})^{ML} \leq \Delta \leq \overline{\Delta}^{ML}$$

where

$$\overline{\Delta}^{ML} = (\bar{y} - \varphi) + \bar{y}\frac{(1-\alpha_1)}{\alpha_1}\beta_1$$
$$\Delta(\underline{e})^{ML} = p[(\bar{y} - \varphi) + (\bar{y} - (\underline{y} - D))\frac{(1-\alpha_1)}{\alpha_1}\beta_1].$$

Proof

Knowing that the repayment period is $T^{ML} = \frac{C(1+\delta)}{\psi}$ we consider equation (2.1) and set the participation constraint:

$$(\bar{y} - \varphi - \Delta)\alpha_1 + \bar{y}(1-\alpha_1)\beta_1 \geq 0$$

Solving for Δ it is straightforward to obtain on the right hand side $\overline{\Delta}^{ML}$. We also require the latter to be greater than $\Delta(\underline{e})^{ML}$, which implies the following condition

$$(\bar{y} - \varphi)(p-1) + \frac{1-\alpha_1}{\alpha_1}\beta_1[p(\bar{y} - (\underline{y} - D)) - \bar{y}] < 0$$

which is certainly true given our assumptions.

The derivation of the payoff condition is similar to the case with high effort equilibrium. We change the order of the inequality by setting equation (2.3) greater than equation (2.1) and using a different repayment period.

Proposition 3. *Under a ML if* $\Delta(\underline{e})^{ML} < \Delta(\bar{e})^{ML} < \overline{\Delta}^{ML}$

- *for* $\Delta < \Delta(\underline{e})^{ML}$ *there is high effort equilibrium;*
- *for* $\Delta(\underline{e})^{ML} < \Delta < \Delta(\bar{e})^{ML}$ *there are multiple equilibria;*
- *for* $\Delta(\bar{e})^{ML} < \Delta \leq \overline{\Delta}^{ML}$ *there is low effort equilibrium.*

Proof

Assuming the validity of Proposition 1 and Proposition 2, we know already that $\Delta(\underline{e})^{ML} < \overline{\Delta}^{ML}$. The first condition $\Delta(\underline{e})^{ML} < \Delta(\bar{e})^{ML}$ holds if

$$\frac{1-\alpha}{\alpha}\beta > \frac{1-\alpha_1}{\alpha_1}\beta_1.$$

which is true under our assumptions. The condition $\Delta(\bar{e})^{ML} < \overline{\Delta}^{ML}$ holds if

$$(\bar{y}-\varphi)(p-1) + p[(\bar{y}-(\underline{y}-D))\frac{1-\alpha}{\alpha}\beta] - \bar{y}\frac{1-\alpha_1}{\alpha_1}\beta_1 < 0.$$

Since the default rate p is lower than 1 (it is usually around $20 - 25\%$) this condition is very likely to hold. However we test it with some numerical simulation.

2.4.3 High effort equilibrium under ICL

To implement a high effort equilibrium under an income contingent loan, we require

$$(PH^{ICL}|\delta = 0) > (PL^{ICL}|\delta = 0) \geq 0 \tag{2.8}$$

Equation (2.8) implies two conditions:

1. participation constraint $(PL^{ICL}|\delta = 0) \geq 0$: in equilibrium the lowest payoff cannot be negative otherwise the students do not go to college;

2. payoff condition $(PH^{ICL}|\delta = 0) > (PL^{ICL}|\delta = 0)$: the payoff to the students choosing high effort has to be higher than the payoff to those choosing low effort, given the strategy of the government of choosing a zero default premium.

Proposition 4. *Under a ICL there is high effort equilibrium if*

$$\Delta < \Delta(\bar{e})^{ICL}$$

$$0 < \epsilon \leq 1$$

$$\theta \geq 0 \quad and \quad \iota \geq 0.$$

where

$$\Delta(\bar{e})^{ICL} = p[(1-\gamma)(\bar{y} - \frac{\underline{y}\,\iota}{\epsilon}) + \bar{y}\frac{(1-\epsilon)}{\epsilon}\theta]$$

Proof

Knowing that the repayment period is $T^{ICL} = \frac{C}{\gamma \bar{y}}$, we consider equation (2.5) and set the participation constraint:

$$(1-p)[\bar{y}((1-\gamma)\epsilon + (1-\epsilon)\theta)] + p\underline{y}(1-\gamma)\,\iota \geq 0.$$

Given that $\bar{y} > \underline{y} > 0$, $0 \leq p \leq 1$ and $0 \leq \gamma \leq 1$ by assumption, the condition holds. To prove the payoff condition we set equation (2.4) greater than equation (2.5):

$$(\bar{y}(1-\gamma) - \Delta)\frac{\epsilon}{\rho} + \bar{y}\frac{(1-\epsilon)}{\rho}\theta >$$
$$(1-p)[\bar{Y}(1-\gamma)\frac{\epsilon}{\rho} + \bar{y}\frac{(1-\epsilon)\theta}{\rho} + p\underline{y}(1-\gamma)\frac{\iota}{\rho}.$$

we solve for the cost of effort Δ and after some calculation we get

$$\Delta\epsilon < \epsilon p\bar{y}(1-\gamma) + p\bar{y}(1-\epsilon)\theta - p\underline{y}(1-\gamma)\iota$$

then we can easily obtain on the right hand side $\Delta(\bar{e})^{ICL}$.

2.4.4 Low effort equilibrium under ICL

To implement a low effort equilibrium under an income contingent loan, we require

$$(PL^{ICL}|\delta > 0) > (PH^{ICL}|\delta > 0) > 0 \tag{2.9}$$

Equation (2.9) implies two conditions:

1. participation constraint $(PH^{ICL}|\delta > 0) \geq 0$: in equilibrium the lowest payoff cannot be negative otherwise the students do not go to college;
2. payoff condition $(PL^{ICL}|\delta > 0) > (PH^{ICL}|\delta > 0)$: the payoff of the students choosing low effort has to be higher than the payoff of those choosing high effort, given the strategy of the government of choosing a positive default premium.

Proposition 5. *Under a ICL there is low effort equilibrium if*

$$\Delta(\underline{e})^{ICL} \leq \Delta \leq \overline{\Delta}^{ICL}$$

where

$$\overline{\Delta}^{ICL} = \bar{y}(1-\gamma) + \bar{y}\frac{(1-\epsilon_1)}{\epsilon_1}\theta_1$$
$$\Delta(\underline{e})^{ICL} = p[(1-\gamma)(\bar{y} - \frac{\underline{y}\,\iota}{\epsilon_1}) + \bar{y}\frac{(1-\epsilon_1)}{\epsilon_1}\theta_1].$$

Proof

Knowing that the repayment period is $T^{ICL} = \frac{C(1+\delta)}{\gamma\bar{y}}$ we consider equation (2.4) and set the participation constraint:

$$(\bar{y}(1-\gamma) - \Delta)\epsilon_1 + \bar{y}(1-\epsilon_1)\theta_1 \geq 0$$

Solving for Δ it is straightforward to obtain on the right hand side $\overline{\Delta}^{ICL}$. We also require the latter to be greater than $\Delta(\underline{e})^{ICL}$, which implies the following condition

$$\frac{\bar{y}(1-\epsilon_1)\theta_1}{\epsilon}(p-1) + \bar{y}(1-\gamma)(p-1) - p\underline{y}\,\frac{\iota}{\epsilon_1} < 0$$

which is certainly true given our assumptions.

The derivation of the payoff condition is similar to the case with high effort equilibrium. We change the order of the inequality setting equation (2.5) greater than equation (2.4) and using a different repayment period.

Proposition 6. *Under an ICL if* $\Delta(\underline{e})^{ICL} < \Delta(\bar{e})^{ICL} < \overline{\Delta}^{ICL}$

- *for* $\Delta < \Delta(\underline{e})^{ICL}$ *there is high effort equilibrium;*
- *for* $\Delta(\underline{e})^{ICL} < \Delta < \Delta(\bar{e})^{ICL}$ *there are multiple equilibria;*
- *for* $\Delta(\bar{e})^{ICL} < \Delta \leq \overline{\Delta}^{ICL}$ *there is low effort equilibrium.*

Proof

Assuming the validity of Proposition 4 and Proposition 5, we know already that $\Delta(\underline{e})^{ICL} <$

$\overline{\Delta}^{ICL}$. The first condition $\Delta(\underline{e})^{ICL} < \Delta(\bar{e})^{ICL}$ holds if

$$\bar{y}(1-\epsilon)\frac{\theta}{\epsilon} - \frac{a}{\epsilon} > \bar{y}(1-\epsilon_1)\frac{\theta_1}{\epsilon_1} - \frac{a}{\epsilon_1}.$$

where $a = (1-\gamma)\underline{y}\ \iota$.

The condition $\Delta(\bar{e})^{ICL} < \overline{\Delta}^{ICL}$ holds if

$$\frac{\bar{y}(1-\epsilon_1)\theta_1}{\epsilon_1} > p\,[\frac{\bar{y}(1-\epsilon)\theta}{\epsilon} - (1-\gamma)\underline{y}\ \frac{\iota}{\epsilon}]$$

Since the default rate p is lower than 1 (it is usually around $20-25\%$) this condition is very likely to hold. However we test all these conditions with some numerical simulations.

2.5 *Simulations*

We assign numerical values to each parameter in the equations of the costs of effort in order to compare them under the two funding systems. Using the same BCS70 data set of Chapter 1, we consider the income of graduates aged 30 in 2000. We set the low level of the income ($\underline{y}$) equal to the first percentile of the income distribution, and the highest income ($\bar{y}$) equal to the 95th percentile. We distinguish between men and women:

Tab. 2.1: Earnings (in £)

	All	Males	Females
$\underline{y}$	2892	6518	1760
$\bar{y}$	48000	54000	36000

In our baseline calibration we consider a probability of default equal to 25%, a default premium of 9% and a subjective discount rate of 8%. Following Dearden et al (2008), the highest amount available on loan students in a 3-year degree (including leaving expenses) is £12450 before the English 2004 reform[5], and £18000 post-reform[6]. The ML installment is set to £3000, an amount greater than the low income therefore those graduates will default. The ICL repayment rate is 9%. We then repeat the simulations increasing only one parameter and keeping all the others constant. We use a probability of default equal

[5] The pre-reform fess are £3450.
[6] The post-reform fess are £9000.

to 70%, a default premium of 15% and a ML installment of £3000. We consider only risk neutral individuals.

2.5.1 Results

In Table 2.2 we report the simulated cutoff points of the cost of effort, obtained by calibrating the corresponding equations. We conduct our analysis for the whole sample, and separately for males and females. In our baseline model, we first notice that the participation constraint always binds, and so HE participation is greater under an ICL than a ML. This means that under a ML the individual prefers the no schooling option when Δ is low, in contrast to an ICL. An important result is that under an ICL the students prefer to spend more effort when they participate in HE.

Now to compare the two loan schemes, we first observe that the cutoffs' points under a ML are always smaller than those under an ICL. This implies that the students are willing to spend less effort to implement all the types of equilibria under a ML.

So starting from the cutoffs $\Delta(\underline{e})^{ML}$, for $\Delta < \Delta(\underline{e})^{ML}$ a high effort equilibrium is implemented under both financing schemes. For $\Delta(\underline{e})^{ML} < \Delta < \Delta(\bar{e})^{ML}$, under a ML we observe multiple equilibria. But under an ICL, until we do not reach $\Delta(\underline{e})^{ICL}$ we can still implement a high effort equilibrium. Then for $\Delta(\underline{e})^{ICL} < \Delta < \Delta(\bar{e})^{ICL}$ there are multiple equilibria. Under a ML after $\Delta(\bar{e})^{ML}$, which is lower than $\Delta(\bar{e})^{ICL}$, we have only low effort equilibrium. Finally, when the costs of effort are very high, $\Delta > \Delta(\bar{e})^{ICL}$ the low effort equilibrium is also the unique solution under an ICL.

Before analyzing the other cases a general conclusion is that under an ICL compared to a ML the individuals prefer to pay higher costs of effort to get a high wage with certainty, which means a high effort equilibrium can be implemented for high Δ. Under a ML the level of the cost of effort that the students are willing to spend to select $\bar{e}$ is small, and therefore they prefer to choose $\underline{e}$ and so gamble on obtaining a low income with probability p. For this reason we observe a low effort equilibrium implemented for low Δ.

The effect of the English 2004 reform has a negative impact on the costs of effort: increasing the cost of education reduces both the participation constraint and the cutoff points under an ICL and a ML. Thus a low effort equilibrium is reached at a lower cost of effort. Looking at the gender differences, for females the cutoff points are smaller than for males, which implies that females are willing to spend less effort to choose $\bar{e}$. They prefer the low effort choice and the risk of default than spending more effort to get a high wage with certainty. Consequently a low effort equilibrium is more likely to occur for females at a given value of Δ relative to males.

We performed other simulations but for simplicity we do not show the values but we nearly comment on the results. Keeping all the parameters constant, for higher probability of default (increasing p) the cutoff points get bigger approaching the participation constraint. The individuals, given a higher probability of getting a low wage, prefer putting more effort into their studies in order to get a high wage. Naturally, if the probability of default is one when low effort is exerted then there will be only a high effort equilibrium.

Similar results hold for increasing default premiums, as the net wage gets smaller the individuals prefer choosing $\bar{e}$ at higher costs of effort. It is less preferred to choose no effort; because they face a probability p of default, and because, if they have a good draw and get a high wage, their wage is consistently reduced by the default premium.

Finally we increase the installments under a ML and the effect is an increase of the cutoff points. If the students have to repay their loan more quickly, then they prefer to get a high wage with certainty instead of risking default. The disadvantage of a ML over an ICL greatly reduces, and under both systems a high effort equilibrium can be implemented at higher costs of effort.

2.6 Default Premium

In this Section we derive analytically the default premium that the government can choose in case of low effort equilibrium. The amount of the premium is decided in order to keep the public budget constraint balanced. We compute the equilibrium premium under the two funding systems.

2.6.1 Mortgage Loan

If the graduates receive a high income $\bar{y}$ with probability (1-p), the revenue for the government is the payment of the fixed installment φ for T^{ML} periods. Instead, if the graduates get a low income $\underline{y}$ with probability p, they must devote it completely for T^{ML} periods to paying back the loan. However, in the latter case the revenue for the government is not sufficient since $\underline{y} < \varphi$. The budget constraint of the government is

$$p\,\underline{y}\,T^{ML} + (1-p)\,\varphi T^{ML} = C \tag{2.10}$$

substituting the installment $\varphi = \frac{C(1+\delta)}{T^{ML}}$

$$p\,\underline{y}\,T^{ML} + (1-p)C(1+\delta) = C \tag{2.11}$$

Solving for δ, the equilibrium default premium under a ML is

$$\delta^{ML} = \frac{p\,(C - \underline{y}\,T^{ML})}{(1-p)C} \tag{2.12}$$

2.6.2 Income Contingent Loan

If the graduate receives a high income $\bar{y}$ with probability (1-p), the government receives a percentage of their earnings for T^{ICL} years. If the graduates obtain a low income $\underline{y}$ with probability p, he/she pays a percentage of those low earnings for T^{max} years; but the revenue is not sufficient to recover the loan. The budget constraint of the government is

$$p\,\gamma\,\underline{y}\,T^{\max} + (1-p)\;\gamma\,\bar{y}\,T^{ICL} = C \tag{2.13}$$

substituting $T^{ICL} = \frac{C(1+\delta)}{\gamma\bar{y}}$ and solving for δ

$$(1-p)\,C\,(1+\delta) = C - p\,\gamma\,\underline{y}\,T^{\max}$$

$$\delta = \frac{C - p\,\gamma\,\underline{y}\,T^{\max}}{(1-p)\,C} - 1$$

then the equilibrium default premium under an ICL is

$$\delta^{ICL} = \frac{p\,(C - \gamma\,\underline{y}\,T^{\max})}{(1-p)\,C} \quad . \tag{2.14}$$

Comparing the equilibrium premiums under the two systems, that is equation (2.12) with equation (2.14), we can easily observe that if $T > \gamma T^{max}$ then $\delta^{ML} < \delta^{ICL}$. This happens for $\gamma < 1$, which holds by assumption because the individuals cannot repay more than their wage.

2.7 Conclusion

In this Chapter we extended the schooling model illustrated in Chapter 1, allowing for the possibility of default. We considered the possibility that the individuals are not able to repay completely their loan with the government because their earnings after graduation are too low.

The new model is structured as a two-stage game between students and government, where the government first decides the financing scheme and then plays a simultaneous game. The strategies for the government are to set a positive default premium or no premium at all, and the students can choose between positive effort and no effort in their studies. We also assume a participation constraint, that implies a no schooling choice when the costs of effort become too high.

By comparing the individuals payoff under all the possible cases, we derived analytically the cutoffs of the costs of effort required to implement the Nash equilibria. The equilibria are high effort, low effort or multiple equilibria. We play the same game for a ML and an ICL.

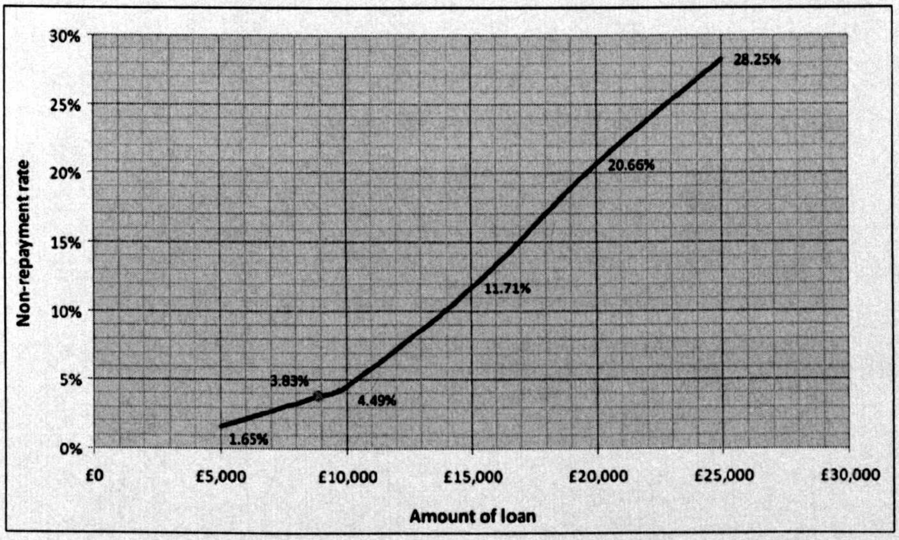
Our first finding is that under an ICL, compared to a ML, the students are willing to spend more effort to participate in higher education. Second, under an ICL a high effort equilibrium can be implemented for higher costs of effort relative to a ML. This means that under an ICL the individuals prefer putting higher effort in order to receive with certainty a high wage. Conversely, under a ML for relatively lower costs of effort the students prefer the no effort choice and hence the risk of getting a low income and then being in default.

The results we found have some policy implications. If the government wants to avoid default and having students putting the highest effort possible into their studies, the choice of an ICL is more likely to achieve this objective than a ML. For the government, in fact, there is no difference among the two financing schemes when the required level of effort in studying is low. Therefore in courses that do not require particularly high psychic costs the government will be indifferent between a ML and an ICL. When the costs of effort are too high, for example for students with special education needs, an ICL is the preferred system since the students are more likely to want to pay the higher costs of effort and avoid the possibility of default. Moreover, assuming that the earnings received by the individuals represent their returns to education, the earnings received by males are perceived as a good return in order to decide to bear very high costs of effort. For females the returns to education in term of earnings are lower, therefore they prefer to bear lower costs of effort and face the possibility of default for lower delta than men.

Tab. 2.2: Simulated cost of effort cutoff

	All			Male			Female		
Costs	$\Delta(\underline{e})$	$\Delta(\bar{e})$	$\overline{\Delta}$	$\Delta(\underline{e})$	$\Delta(\bar{e})$	$\overline{\Delta}$	$\Delta(\underline{e})$	$\Delta(\bar{e})$	$\overline{\Delta}$
	ICL								
Pre-reform	4.78	5.17	20.28	5.63	6.09	25.38	2.81	3.03	11.80
Post-reform	3.46	3.73	14.68	4.05	4.37	18.28	2.06	2.21	8.67
	ML								
Pre-reform	3.68	3.96	14.86	4.11	4.43	16.76	2.74	2.95	11.07
Post-reform	2.73	2.92	11.00	3.05	3.27	12.41	2.03	2.17	8.17

Fig. 2.1: Non-repayment rate according to total loan amount at graduation - (Source: Schwarzenberger (2008))



CHAPTER 3

3. EVALUATING INVESTMENT IN HIGHER EDUCATION UNDER ALTERNATIVE FINANCING METHODS

3.1 Introduction

The existing literature in the economics of education has many examples that focus on policy intervention in HE such as tuition subsidies (e.g. Kean and Wolpin (1997) and Lee (2001)) and many examples that estimate returns to schooling (see Belzil and Hansen (2004) for a comprehensive survey). Our aim in this work is twofold: we combine empirical estimation of the returns to HE with an analysis of the most recent HE financing reforms in England. The objective is to compute the internal rate of return (IRR) of the investment in HE versus high school (HS), under different HE financing schemes, i.e. free education, and public loans (mortgage loan and income contingent loan). In this way we add empirical evidence both to the literature on income contingent loans (e.g. Barr (1993), Chapman (2005)) and to the literature on the new HE reforms in the UK (see Barr and Crawford (2005) for a review).

There are, in particular, two works very close in spirit to ours. Dearden et al. (2008) generate lifetime wage profiles using the UK LFS dataset to study the distributional effects of the English HE Reform. We notice two main differences relative to our work. They consider only graduates versus all non graduates without further distinguishing the education level of the latter group. Here we compare graduates with non graduates that had the education level that would have allowed them to attend university. Secondly, they combine the wage distributions of graduates and non graduates, and this distribution is assumed to grow according to a Markov process. They do not consider the problem of potential selection bias, although they take into account whether individuals are unemployed. In our work the age earnings profile of individuals at each education level is generated from econometric estimates of the real returns of education. We correct for endogeneity and we do not impose separability in the earnings function between the schooling effect and the age effect. Moreover, Dearden et al. compute the net present value of the debt repayments using a fixed discount rate. Their interest is in the years that individuals take to repay their debts, the government subsidy provided by an ICL and the difference between the ICL and ML systems. Their analysis distinguishes between different parental background and between low and high earners. Their findings suggest a strong increase in welfare under an ICL that provides insurance for individuals with low returns and poor family background. As we said we are interested in computing the IRR of the investment in HE, by contrasting free education with two public loan schemes, ML and ICL.

For our estimations we use an ordered probit selection model to correct for endogeneity. In the literature we find several examples of estimation using treatment effect models, Willis and Rosen (1979) consider college education as an endogenous dummy variable; Kenny et al. (1979) extend this to the case in which schooling is a Tobit specification; Harmon and Walker (1995) use an extension of the Heckman two-step approach where the latent variable for schooling is treated as an ordered probit. Our work is an extension of Heckman et al. (2008): they also compute the IRR of the investment in education for different levels of schooling, starting from a non parametric estimation of the determinants of log earnings. In contrast to Heckman we adopt a parametric selection model. Parametric models have some advantages over non parametric: they converge faster, they do not require the estimation of smoothing parameters, they are easy to interpret, and non parametric estimations cannot be used to extrapolate out of sample. In our selection model the educational variables can be ordered (from the lower secondary education up to the postgraduate qualifications) and we estimate the probability that individuals have a particular level of education exploiting the fact that they are mutually exclusive.

Using the UK Labour Force Survey, a very large and flexible dataset, we estimate the mean of the distribution of the causal effect of education in two-steps. The distinctive features of this work are: we control for selectivity and our model does not impose the restriction that the age profiles (and the region and race effects) are the same at each level of education. We correct for unobserved ability using, as exclusion restrictions, the raising of the school leaving age reform (rosla) imposed on the English education system in 1973 and the month of birth. Crawford et al. (2007), analyzing the English education system, show that children born later in the school year perform significantly worse in exams than those born earlier in the school year. Both rosla and month of birth are in principle uncorrelated with the unobservable determinants of earnings.

In a previous analysis we considered the expected utilities of the individuals receiving an uncertain wage after graduation, and we compared their utilities under the two loan schemes. We calibrated our theoretical model using the mean and the standard deviation of the earnings of graduates, distinguishing among individual characteristics, family background and degree subjects. We used the BCS70 data set, that provides useful information on the family characteristics of the individuals but it is limited only to the cohort of people born in 1970. We also generated an income flow along the individuals working life, assuming a Brownian motion growth rate and using as initial income the 10th percentile and the median of the graduate wage distribution in BCS70.

Here we now treat education as a financial investment and we evaluate the returns of this investment for HE graduates versus HS graduates. We compute the returns to education as the internal rate of return from a flow of future earnings (not as a single parameter from a schooling coefficient). That is we find the interest rate that makes the net present value of the earnings of the HE graduates equal to the net present value of the earnings of HS graduates, minus the costs of schooling. Heckman et al. show that the internal rates of return to schooling for a general earnings function differ from those obtained with the traditional Mincerian returns which are based upon strong assumptions; and they also take into account taxes and tuition but not foregone earnings. We use a similar framework to generate the IRR, but we take into account both tuition and foregone earnings. When the individuals start working they keep their entire earnings if education is free; otherwise their earnings are reduced by an amount that depends on the structure of the loan scheme decided by the government.

Our results suggest values of the IRR in line with the recent literature. In particular, focusing on the most recent educational reform in the English HE education financing system, we find the highest IRR is associated with free education. The IRR under an ICL is always higher than under a ML. Depending on the structure of the loan under an ICL, the IRR can get very close to the IRR under free education.

This Chapter is structured as follows: Section 3.2 describes the data, in Section 3.3 we explain the estimation method of the returns to education and we illustrate the results, Section 3.4 computes the IRR and shows the results for the different HE loan systems, Section 3.5 concludes.

3.2 Data description

The LFS is a quarterly sample survey of households living at private addresses. Its purpose is to provide information on the UK labour market that can then be used to evaluate labour market and educational policies. The survey seeks information on respondents' personal circumstances and their labour market status during a specific reference period, normally a period of one week or four weeks (depending on the topic) immediately prior to the interview. The UK LFS has existed since the mid 1970's but it is only since 1993 that data on gross earnings has been collected.

Our selected sample from the raw LFS data consists of: working age men[1] interviewed in Wave 5 of the Quarterly LFS in calendar years 1993-2005 inclusive. The following groups were dropped from the analysis: residents of Northern Ireland, people who came to the UK after age 5 (and so are likely to have had some education outside of the UK), and those still in full-time education or never had education.

The main variables of interest in this analysis are earnings, education and individual characteristics. To obtain our final sample we impose the following restrictions:

1. The hourly gross earnings information in the data is used to define the annual earnings. Our constructed wage variable is transformed into a real wage rate by dividing by the Retail Price Index (All items) such that September 2000 is set as the base. Thus all earnings are measured in September 2000 prices.
   In addition, the top and bottom 1% of the annual wage distribution by category of highest academic qualifications were trimmed to avoid outliers arising from measurement error in the wage rate influencing the results unduly.
   Finally, it is generally more useful to consider the proportional differences in earnings across different groups of individuals, rather than the absolute difference. That is, it makes more sense to investigate the percentage, rather than absolute, difference in

[1] The analysis of women is the subject of further work.

earnings associated with a higher qualification. To this end it is usual to consider the log of the wage rate rather than the wage rate itself. This is because the difference in the log of earnings is (approximately) the percentage difference in earnings. And this allows us to interpret our estimates as the percentage effects on earnings.

2. Our analysis concentrates on education qualifications, rather than the age at which individuals leave education. In England, compulsory education is organized from the age of 5 to 16. Age 11 is the end of primary school. Pupils undertake examinations for the General Certificate of Secondary education[2] (GCSEs) at age 16. After the age of 16 they can enter the labour market or continue to post compulsory education. Students can choose between academic and vocational qualifications. The academic track consists of GCSEs at 16, followed by A-levels at 18 and university degree usually at age 21, followed by possible postgraduate degree. The vocational track is less easy to characterize, typically students would leave formal education at the age of 16 and engage in some occupational training perhaps on a part-time basis while at work. Many vocational qualifications are taken by rather small group of individuals. Thus we adopt the grouping of qualifications into National Vocational Qualifications (NVQ) equivalent[3] defined in Table 3.1. There are five NVQ levels, from NVQ1 (GCSE equivalent qualifications) to NVQ5 (master and PHD equivalent qualifications). Table 3.1 gives examples of the most important non vocational qualifications and their associated NVQ levels. We follow established practice in how NVQs are defined with the exception that we pool together the NVQ4 and NVQ5 qualifications due to the small number of observations at NVQ5 level[4]. So NVQ1 is the omitted category.

3. The economic activities are restricted to employees and we further restrict the total number of hours worked in the reference week to the range [0...94].

The total sample size is 169,811 males, aged from 16 to 70 years old, born between 1935 and 1990. Only those earning a positive wage are included in the sample and the distribution of the earnings given the NVQ levels is shown in Table 3.2. We observe that

[2] This in the education system after 1973.
[3] SeeMakepeace et al. (2003).
[4] We refer to NVQ4 for both NVQ4 and NVQ5 in the future.

the majority of the individuals, 38% of the sample, have a NVQ3 and their wage is more than 16% higher than those with a NVQ2 qualification. The individuals with NVQ4 are around 34% of the sample and they get the highest earnings. The college premium is the earnings for those with NVQ4 (or 5) relative to NVQ3, and Figure 3.1 shows the college premium and other differentials by age from the raw data. It is evident that the earnings profile for NVQ4 is higher and steeper than NVQ3.

3.3 Estimation Method

In the empirical literature it is common to estimate the returns to schooling, Belzil and Hansen (2004) show that since 1970 more than 200 published articles and working papers have been devoted to this topic. It is well known that OLS methods produce unbiased estimates only if realized schooling and unobservable attributes that affect earnings are uncorrelated. The presence of ability bias implies that estimates of the effects of education on earnings reflect the unobservable difference across individuals as well as the effect of education per se, so OLS are biased upwards.

We can think that more highly educated individuals might earn more because they have unobservable attributes, like ability, that are valued in the labour market and which are correlated with education. As stressed by Walker and Zhu (2007), we need to distinguish the average effect due to an additional year of schooling, from the differences in earnings that occur, on average, across individuals that have different levels of education.

The empirical problem of estimating the true causal effect of education is to exploit that part of the variation in education levels across individuals that is not due to self-selection. In the literature many studies use the IV approach, and Card (2000) surveys a number of studies which use a variety of instruments. We can mention among others: parental background (Willis-Rosen, 1979); quarter of birth (Angrist and Krueger, 1994); nearby college (Card, 1995b); raising of the school leaving age (Hogan and Walker, 1995); WWII (Ichino and Winter-Ebmer, 1998).

IV methods find, in general, returns to schooling between 20% and 40% above the OLS estimates, which are typically slightly larger than the estimates from selection models. One of the explanations for IV estimates systematically exceeding the estimates of other methods, is the combination of small ability bias (in IV), with downward bias in the

OLS estimates due to measurement error in reported schooling. For example, imagine there is no noise in the education data, then the estimated correlation between earnings and education would be correct. At the other extreme, imagine there was only noise in the education data, then education is reported randomly and cannot be correlated with earnings (or anything else) and the coefficient on education will be zero. So more noise biases the estimates downward (Griliches (1979); Angrist and Krueger (1991), Walker and Zhu (2007)). However it seems unlikely that measurement errors in qualifications will be very large.

An alternative explanation is that IVs provide, in terms of the Angrist, Imbens and Rubin (1996) causal model, a Local Average Treatment Effect (LATE) estimator, which estimates the effect for those whose treatment status is affected by the instrument. LATE is the average effect of the treatment for those who change treatment status because of a change in the value of the instrument.

An intervention does not affect all individuals in the same way. Typically there is heterogeneity in the response to the treatment across individuals. Consequently, there are different potential questions that evaluation methods attempt to answer, the most common is the average effect on individuals of a certain type. Heckman and Vytlacil (1999) and Blundell and Dias (2002) distinguish between: the population average treatment effect (ATE), which would be the outcome if individuals were assigned at random to treatment, the average effect on individuals that were assigned to treatment (TT), the effect of treatment on agents that are indifferent to participation, which is the marginal version of the local average treatment effect (LATE) discussed above, or the effect of treatment on the untreated (TU) which is typically an interesting measure for decisions about extending some treatment to a group that was formerly excluded from treatment. Except for the case of homogeneous treatment effect[5] all these measures are conceptually different and each has a different set of conditions for identification.

Moffit (1999) gives a simple graphical interpretation of these estimators. Figure 3.2 shows a hypothetical density of α_i (with mean $\bar{\alpha}$) in the population, where α_i is the heterogenous treatment response for individual i, and it is also known as random coefficient.

[5] Under this assumption all these measures are identical.

Moffit assumes that selectivity takes place strictly on α_i and that selection is positive. This means that individuals with higher α_i are more likely to participate in the treatment. More precisely those with α_i above the cutoff point α^* participate and the other individuals do not. In Figure 3.2, the parameter $\bar{\alpha}$ is the ATE. Whereas α^{TT} is the TT, the average gain of those who participate $E(\alpha|\alpha_i > \alpha^*)$, and α^{TU} is the mean gain of non participants in case of treatment $E(\alpha|\alpha_i < \alpha^*)$. To represent the LATE we need to consider a discrete changes in α^* to $\alpha^{*'}$, which changes α^{TT} to $\alpha^{TT'}$. The LATE is given by the the difference $\alpha^{TT} - \alpha^{TT'}$ divided by the area under the curve between α^* and $\alpha^{*'}$, which is the change in the probability of participating.

In our analysis we use an ordered probit selection model, which estimates the mean of the distribution of the causal effect of education in two-steps. With our model we do not impose the restriction that the selectivity works in the same way at all NVQ levels. That is we do not impose the restriction that the age profiles (and the region and race effects) are the same at each level of NVQ. So we allow for non separability in the earnings function between the schooling effect and the age effect.

A selection model is more restrictive than IV since we assume that the distribution of unobservables in both earnings and schooling are jointly normal distributed. Our educational variables can be ordered - so that NVQ4 is higher than 3, and higher than 2 and so on. Therefore, we can estimate the probability that individuals have an NVQ at any particular level and exploit the fact that they are mutually exclusive - so that the probabilities of having each level sum to one across levels. The main difference with IV methods, is that by making an assumption about the distribution of the unobservable determinants of earnings we estimate the effect of qualifications on earnings across the whole distribution of unobservables, in particular at the mean. Thus, the selection model method yields estimates that are comparable to the much simpler least squares regression method.

3.3.1 Exclusion Restrictions

Our selection model is the following:

$$\begin{aligned} Y_{ij} &= \alpha_j + \beta_j D_{ij} + \gamma_j X_{ij} + \epsilon_{ij} \\ for\ i &= 1 \ldots N \quad and \quad j = NVQ1 \ldots NVQ4. \end{aligned} \tag{3.1}$$

$$\begin{aligned} D^*_{ij} &= a_j + b_j Z_{ij} + v_{ij} \\ with \quad D_{ij} &= 1 \quad if \quad D^*_{ij} = j \\ for \quad j &= 1 \ldots 4. \end{aligned} \tag{3.2}$$

where N is the sample size, $E[v_{ij}] = E[\epsilon_{ij}] = 0$, but $cov(v_{ij}, \epsilon_{ij}) \neq 0$, then $E[D_{ij}\epsilon_{ij}] \neq 0$. This is a more general model because we are not assuming separability between age and schooling.

The selection model estimates an ordered probit as the first step , equation 3.2, where the vector **Z** includes at least some variables that are not contained in **X** . As discussed by Heckman (1990) and Card (1993) identification in selection models (as for IV) has to be able to include variables that affect education and do not also affect earnings directly, the so-called exclusion restrictions.

Our identification strategy is based on the exclusion in the wage equation 3.1 of two variables that we can reasonably consider to be exogenous. The first variable is the raising of the school leaving age reform (rosla) imposed on the English education system in 1973. Those born before August 1958 faced a minimum school leaving age of 15 years, those born after that date were required to stay in school until at least 16 years of age. As shown by Harmon and Walker (1995) and Walker and Zhu (2007) rosla can be used both as IV and as an exclusion restriction in a selection model (see also Oreopoulos (2008)).

The second exclusion restriction is month of birth. In the literature it is well documented that there is an impact of date of birth on cognitive test scores, with the youngest children in each academic cohort year performing poorer, on average, than the older members of their cohort. Puhani and Weber (2005) use a sample of German children and investigate the impact of age at school entry on test scores at the end of primary school (age 10). They find that children who start school aged 7 rather than aged 6 have test scores that are 0.42 standard deviations higher at the end of primary school. Bedard and Dhuey (2006) use internationally comparable data for OECD countries to estimate the impact of relative age on test scores at ages 9 and 13. They find that children being one month older get higher test scores at the age of 9 than at age 13. Ashwood and Heyndels (2007) consider the effects of month of birth in soccer education programs. They find systematic differences in players' performance depending on the months in which they are born. These differences then produce productivity and wage differences in adulthood. A work that is very close to the purposes of our analysis is Crawford et al. (2007), which consider the English education system. They show that children born later in the school year perform significantly worse in exams than those born earlier in the school year, even up to GCSE level (NVQ2 level). A child born in September will, on average, perform significantly better in academic tests than a child born in the following August, simply because they start school (and sit the tests) up to a year younger. This means that access to further and higher education, and hence future success in the labour market, is likely to be significantly affected by the month in which they are born.

3.3.2 First step results

We exclude month of birth and rosla from our wage equation. Both variables are in principle uncorrelated with the unobservable determinant of the earnings. The two exclusion restrictions satisfy the condition of the random assignment to treatment, in terms of the Rubin et al. (1996) causal model. This means that all the individuals have the same probability to be born before or after the introduction of the rosla. And in the same way the month of birth of an individual is completely random[6]. In particular, both rosla and month of birth have only an indirect effect on earnings, that it is through the level of education of the individual.

We also include, in both the wage and participation equations, an ethnicity variable to consider the effect of non-white people, and region of residence dummies to take into account of region specific effects.

[6] We assume that parents do not plan the date of birth of their children.

In Table 3.3, we report the results of the first step of our estimation, i.e the ordered probit. At the bottom of this Table we show the tests for the joint significance of rosla and months of birth together and months of birth only. In both cases we reject the null hypotheses of non significant coefficients. Looking at the months of birth, we notice that in September there is a change of sign, the coefficient turns on significantly positive while it was significantly negative in the previous months, especially in August. Our results are therefore consistent with the literature.

Although the total effect of rosla is negative in Table 3.3, we show in Table 3.4 the marginal effects of rosla evaluated by each level of education. The coefficient of rosla is significantly positive from NVQ1 to NVQ3, and it has its maximum at NVQ2. The intention of the government and the consequent effect of the policy was to increase the participation at the lower secondary levels of education, and it's there where we find the highest effect of the rosla. Note that if we had used rosla as an IV then the estimated effect of the NVQ would be for those at the bottom of the education distribution - e.g. for those who wanted to leave at 15 and the policy forces to remain in school until 16. In contrast, in our selection model, we estimate the effect on the mean.

Second Step Results

In the second step, we use the estimated probit coefficients from equation 3.2, to generate four new variables known as the Inverse Mills Ratios (IMR), which capture the likelihood that an individual has a particular level of qualification. We include the IMRs in the wage equation 3.1 to correct for the fact that individuals with a particular level of education will have a particular unobserved component to earnings.

We keep the ethnic variable and the regional dummies used in the first step and we add a series of controls. We include dummies for each LFS survey's year, to control for macro economic shocks; and we assume that log earnings profiles are quadratic in experience for each education classification. There is no a good measure of experience, if we consider a conventional measure, i.e age minus years of education, we end up with an endogenous variable (since education is endogenous). Therefore we use age and age square as proxy for experience, and for computational purposes we consider the deviation from age 30, to keep the mean values of age and age squared closer to zero. We also include a group of dummies to control for individual characteristics, such as marital status, children and a variable for health problems lasting at least 12 months and affecting the work the individual might have done. Our final specification allows the intercept and the coefficients on the controls to vary by schooling qualification. We estimate log earnings for each NVQ level using regressions of the form

$$\begin{aligned} Y_{ij} =& \alpha_j + \gamma_j X_{ij} + \lambda_j \widehat{IMR}_j + \epsilon_{ij} \\ & for\ i = 1 \ldots N \quad and \quad j = NVQ1 \ldots NVQ4. \end{aligned} \tag{3.3}$$

which allows heterogeneity in the effect of qualifications on earnings[7]. As we mentioned before the identification strategy requires an exclusion restriction in the vector $\mathbf{X}$. If the selection variables coincide with the independent variables, $\mathbf{Z} = \mathbf{X}$, we could still use the two-step procedure and get the estimates of the IMR, but identification would come only from the distributional assumptions. Only because of these assumptions the IMR would be a non-linear transformation of the regressors $\mathbf{X}$ in the outcome equations, and this would be sufficient to identify the all the parameters in principle. However in practice this is seldom sufficient.

In Table 3.5 we show the estimates of the wage equations for each NVQ level. A standard t-test on the coefficient, λ, of the IMR is a valid test of the null hypothesis of no selection bias. For the lowest levels of education λs are negative or non significant, whereas for NVQ3 and NVQ4 λs are positive and significant. This is consistent with our expectations, i.e. more highly educated individuals might earn more because they have unobservable attributes, like ability and perseverance, that are valued in the labour market and which are positive correlated with education. So since the IMRs estimate the extent of selectivity we would, if individuals always selected the qualifications that were "best" for them, expect them to be positive.

We test whether the wage distribution at NVQ4 is parallel to the wage distribution at NVQ3. At the bottom of Table 3.5 we show the the F test on restrictions[8], and we reject

[7] we use the Stata routine "hoeckman" written by Chiburis and Lokshin (2007)

[8] The restricted model is given by the regression of the log earnings on age and age square for NQ4 graduates only. To obtain the unrestricted model is slightly more complicated. We first generate the wage residuals at NVQ4 level, given by the difference between the log earnings minus the fitted values of age and age square obtained by the estimation of the wage equation 3.3 at NVQ3 (see Table 3.5). We then regress these residuals on the same controls used in the wage equation 3.3. We construct the F statistic using the squared sum of the residuals in the restricted and the unrestricted models.

the null hypotheses of parallel distributions. This also clear looking at the raw data in Figure 3.1.

The estimates of the other covariates are in line with the recent literature (e.g. Walker and Zhu (2007)). The effect of age (centered at age thirty[9]) is below 3% with an NVQ1 and it slightly increases with higher qualifications, in particular from NVQ3 to NVQ4 it raises from 3.3% to more than 4%. The coefficient of age square is practically unchanged for all levels of education, it is negative and around 0.1% per annum. Looking at the health problems effects our results are quite consistent with the expectations, we find in fact a negative impact on wage ranging between 18% and 23%, the highest value being with an NVQ2. Moreover white people earn more than non white, this effect is quite strong for lower levels of education but it drastically drops by ten percentage points (from around 17% to 7%) from NVQ3 to NVQ4. The ethnic effect therefore is much lower when individuals are highly educated. Finally, we find a positive effect of being married (note we have a sample of males only), the coefficients of marital status vary between 12% and 16%. In Table 3.6 we show the estimates from applying OLS and we notice that the slope coefficients on the exogenous characteristics are virtually the same, apart from slightly smaller standard errors. The substantial differences are the higher constant term at NVQ4, and the higher effect of age at NVQ4 [10].

In our ordered probit selection model the returns to schooling do not appear among the estimated coefficients of the wage equation, but they are obtained through the estimated increments to predicted earnings that are associated with the successive levels of NVQ. In Table 3.7 we compare the results of our ordered probit selection model OPS (that assumes that educational qualifications are correlated with the unknown factors that determine earnings) with the OLS (that do not allow for such a correlation). These figures are the average predicted log earnings across each sample. With OPS the returns of an NVQ2 relative to NVQ1 is a 3% wage increase, while with OLS is around 12%. Whereas the effect of NVQ3 relative to NVQ2 is 8%, same as OLS. The so called "college premium"

[9] The estimates should be insensitive to this, the constant term adjusts to deal with it. But it is a convenient normalization and improves the computational precision of the estimates; it is also convenient for future work when we will be merging with the BCS70 dataset.

[10] We estimated the same model including cohort controls (i.e. a cubic effect of year of birth). We obtain effectively the same results. As we would expect, the only substantial difference is that the age effects on earnings are now larger reflecting the fact the the old age earnings profiles are contaminated by cohort effects. But this does not matter for our subsequent analysis, what matters is that we still have a substantially steeper age earnings profile for NVQ4 compared to NVQ3.

that is the higher earnings of HE graduates relative to high school graduates is under OPS around 9%, whereas with OLS is almost 28%. These results are consistent with our expectations, we find OLS estimates higher than the selection models because we have corrected for ability bias. Similar results are in Walker and Zhu (2007), while Harmon and Walker (1995), using a selection model with year of schooling as endogenous variable and rosla as exclusion restrictions, find in general results higher than OLS and more in line with IV. It seems that allowing for non separability and nonlinearity in the effects of education on earnings is very important.

3.4 Evaluation of the investment in HE

In the second part of this chapter we evaluate the internal rate of return (IRR) of the investment in higher education for different HE financing schemes: free education, mortgage loan and income contingent loan.

We restrict our attention to NVQ3 and NVQ4 qualifications only, since our target is to generate lifetime earning profiles to be used to compute the IRR of NVQ4. The lifetime earnings distributions are constructed using the parameters estimated in the wage equations. We make our calculations for the default individual, so then we need to consider only the intercept and the coefficients of age and age square, each one selected according to the level of education. We generate two different age earnings profiles for those having an NVQ3 and those having an NVQ4 [11],

$$\begin{aligned} W_{tj} &= \alpha_j + t \cdot \beta_{1j} + t^2 \cdot \beta_{2j} \\ for\ t &= -11 \ldots 28 \quad and \quad j = NVQ3 \end{aligned} \tag{3.4}$$

$$\begin{aligned} W_{th} &= \alpha_h + t \cdot \beta_{1h} + t^2 \cdot \beta_{2h} \\ for\ t &= -8 \ldots 31 \quad and \quad h = NVQ4 \end{aligned} \tag{3.5}$$

Since we have centered the age at 30, we start the profile back to the year following the awarding of the qualifications: age 19 for NVQ3 and age 22 for NVQ4. In both equations, β_1 is the coefficient of $(age-30)$ and β_2 is the coefficient of $(age-30)^2$ of the corresponding NVQ level.

[11] If we include the IMR it would not make any difference.

We notice two main differences relative to Dearden et al (2008). They generate the wage profiles only for graduates versus all non graduates without further distinction in the education levels, we instead allow for a comparison between university graduates and high school graduates. They construct the lifetime earnings distributions using copula functions[12] and assuming that earnings grow according to a Markov process (2% per year), and they do not consider the problem of selection bias. In our case, we first correct the returns to education for sample selection bias. Then, given the estimated parameters, we generate the earnings profile, where the earnings grow according to the values of the experience coefficients.

The next step is to use the wage profiles of HE graduates, W_{NVQ4}, and of high school (HS) graduates, W_{NVQ3}, to compute the internal rate of return. Dearden et al. compute the net present value of the debt repayments using a given discount rate, we use instead the same framework as Heckman et al. (2008) to solve for the IRR. The latter estimate first log earnings using a non parametric method, and then they compute the IRR by solving for r in an equation like the following

$$\begin{aligned} &\sum_{t=1}^{40}(1+r)^{-(t+5)}\left[W_{NVQ4}-\tau(t)\right]-\sum_{s=1}^{5}(1+r)^{-s}\left[C_{NVQ4}\right] \\ &=\sum_{t=1}^{40}(1+r)^{-(t+2)}\left[W_{NVQ3}\right]-\sum_{s=1}^{2}(1+r)^{-s}\left[C_{NVQ3}\right] \\ &\textit{where} \\ &\tau(t)=\tau \quad if\, ML \\ &\tau(t)=\tau(W_{NVQ4})) \quad if\, ICL \end{aligned} \tag{3.6}$$

The IRR is the rate of interest that equates the net present value of income of HE graduates and HS graduates[13]. On the left hand side, equation 3.6 discounts HE graduates' earnings net of loan repayments and costs of education back to the age of 16, because at this age the individuals decide whether to undertake high school or start working. The parameter τ is a per period repayment of the loan, and its amount depends on the HE financing scheme chosen by the government (ML or ICL). On the right hand side

[12] The name being a reference to the fact that they couple together k univariate distributions to form a k-variate distribution.

[13] In financial terms, the IRR is the annualized effective compounded return rate which can be earned on the invested capital, i.e., the yield on the investment.

equation 3.6 discounts HS graduates' earnings and costs education back to the age of 17.

We assume the same length of the working life (40 years) both for HE graduates and HS graduates. The assumption is unimportant, because the earnings at the end of the life cycle are heavily discounted and so small differences in working life have non relevant impact on the present value of lifetime earnings.

The costs of education are discounted for the number of years of education. Heckman et al. consider the US education system and they assume no costs for the first 12 years of education and then they discount the tuition for the remaining years according to the education level reached. We consider the English system and assume no costs up to NVQ2 level of education which corresponds to students aged 16 that have studied 11 years. Afterwards, they study two years, from age 17 to age 18, to get an NVQ3 and three years, from age 19 to age 21, to get an NVQ4.

Now particular attention is required for the definition of costs. Heckman et al. divide the total revenue from tuition and fees across all higher education institutions in the US, by the total enrollment to obtain estimates of the average private tuition cost of college. In our case we assume that the initial cost of the educational investment is given by both tuition and foregone earnings during the years of schooling. We further exploit the results of our estimation to compute those foregone earnings.

For simplicity we assume no tuition for HS graduates, and C_{NVQ3} in equation 3.6 is equal to foregone earnings only, that is the average wage of those having an NVQ2 qualification. In practice, we use the earnings for the first two years of the age earnings profile[14] of the individuals with an NVQ2.

For HE graduates we consider both tuition and foregone earnings. In terms of equation 3.6, C_{NVQ4} is equal to the foregone earnings during HS (C_{NVQ3}) plus the foregone earnings during HE. This last term is given by the average predicted earnings of those having an NVQ3 qualification. In practice, we take the earnings for the first three years of the predicted age earnings profile of those with an NVQ3.

Since we are analyzing the English system, we contrast the old free HE education system with the subsequent loan scheme (first a mortgage loan, then an income contingent loan). Therefore in both cases students are subsidized during university, because they are

[14] The age earnings profile for NVQ2 is computed using an equation similar to 3.4 and 3.5, but taking the estimates in Table 3.5 for NVQ2.

only expected to repay the tuition fee (if required) with their earnings after graduation[15].

3.4.1 HE Financing schemes

Our setting is an extension of Heckman et al. (2008) and Dearden et al. (2008), where we study the alternative HE financing systems adopted with the most recent educational reforms in England. Before 1990 there were no tuition fees for English students, their living expenses were covered by a tax-funded grant, which was means-tested against parental income, and expectations of parental contributions. In 1990 a loan with mortgage-type repayments was introduced, to be repaid in a fixed number of monthly installments[16] - normally 60. Individuals started to repay their debt after they graduated, however, repayments could have been deferred on a yearly basis if the repayer's income was below 85% of the national average. In 1998 they introduced an upfront fee (£1,125 in 2003-4), irrespective of subject or university. There was no loan to cover the fee and living expenses were met by a mixture of parental contributions and income-contingent loan (ICL), which replaced the previous mortgage loan. Graduates started to pay back their educational loan only when their earnings were above $\pounds 15,000$ per year and at a 9 per cent fixed repayment rate. There was a zero real interest rate and repayments were made through the tax system as a payroll deduction. The recent Higher Education reform (approved in 2004 and effective from 2006/2007) increased the tuition fee and universities were able to set their fees up to a maximum £3000 p.a. The major innovation was the extension of the income-contingent scheme, with the same repayment rules, to also cover the tuition fee.

In our simulations, we assume a zero real interest rate on the loan as in the English system, and we compare a mortgage loan (ML) and an income contingent loan scheme (ICL) for different amounts of loan. The basic loan covers only the fees, which under the English Reform are currently fixed to a maximum of £3000 per year by almost all institutions. This implies that students taking a three-year undergraduate course will graduate with a debt of £9000. Then, we consider a loan of £18000 (£9000 of this is for fees and think of the remainder being for maintenance) representing the maximum amount available to those with low parental income (assumed below £35000 in Dearden et al.) in

[15] The criteria to compute the repayments are based on gross annual earnings (source Student Loans Company Limited UK). We for the moment ignore the non linear tax system. Since we are comparing two loan schemes the effect of non linear tax system will cancel out from IRR differences.

[16] Source: Student Loans Company Limited UK.

the existing scheme. This compares with a maximum loan of £12000 (all for maintenance) under the old ML system. All amounts in the computations are expressed in 2000 prices.

We set the mortgage loan parameters according to the English scheme. The threshold above which graduates start to repay their loans is fixed at 85% of the weighted average of the predicted earnings across each NVQ level. Given the earnings profile generated by our estimates, under a ML the repayments start after five years. The annual installment depends on the loan: in the basic setting (£9000) we consider a five-year repayment period which means an annual installment of £1800. We also use installments of £2400 and £3600, which imply correspondingly different repayment periods.

Our hypothetical income contingent loan is structured according to the English HE Reform 2004. Graduates start to repay their loan if they have earnings higher than £15000 per year, and the amount due is equal to 9% of the earnings in excess of the threshold. We also try different structures with no initial threshold and with a higher threshold of £20000. If repayments take longer than 25 years, the debt is written off.

In figure 3.3 we show the two discounted net earnings profiles of HE graduates taking the highest loan (£18000), under an ICL with initial threshold of £15000 and a ML with annual installment of £3600 for five years. The fall in net earnings is evident, during the five-year repayment period of the mortgage loan, while under an ICL we observe a smooth decrease along the whole working profile[17].

3.4.2 IRR results

In Table 3.8 we report the computed internal rates of return of the investment in higher education versus high school. In Table 3.9 we show the taxpayer subsidy and the repayment periods. The individuals that take longer to repay can exploit a bigger hidden subsidy due to discounting and the zero real interest on the loan. This is also called the taxpayer subsidy, expressed as proportion of the face value of the loan. Note that to calculate the value of taxpayer subsidies, we assume a real discount rate of 2.2% per year as in Dearden et al. (this follows the government's present convention for discounting[18]). In our analysis, we simulate different situations according to the size of the loan and the

[17] All this is being done at mean predictions of NVQ4 and NVQ3. In future work we intend to make these calculations for the whole wage distribution so we can show the range of the IRR

[18] See Department for Education and Skills (2007).

structure of the financing systems. However in each case the foregone earnings are always assumed unchanged.

Looking at Table 3.8 we notice that the IRR, under all the different specifications, ranges between a minimum of 6.2% and a maximum of 7.1%. When education is free, as expected, we get the highest IRR, around 7.12%, but this is almost matched by an ICL with a threshold of £20000. In the latter case the IRR is unchanged at 6.95% for any size of the loan. This is because the individuals start the repayments ten years after graduation and they do not completely repay the loan, which is written off after 25 years. Looking at Table 3.9, the repayment periods are longer than 25 years and the graduates exploit the highest hidden subsidies. For example, with a loan of £18000, almost 71% of it is subsidized by the taxpayer.

It is clear that the effect of the initial threshold is of course to increase the IRRs. In fact, without it, graduates start repaying an amount equal to 9% of their total annual earnings soon after graduation. This means that they pay off their £9000 debt in just seven years, with an IRR of around 6.66% and a hidden subsidy a little more than 8%, which is the lowest in our simulations.

Under the current English system with fees of £9000 and an initial threshold of £15000, the individuals start to repay seven years after graduation, but only 9% of the part of earnings exceeding the threshold. They take 23 years to pay off the loan yielding an IRR of 6.86%. These figures are only slightly smaller at higher levels of loan, because the hidden subsidy increases.

The effects of this dramatic differences in the IRR and taxpayer subsidies arise because the earnings at the end of the life cycle are heavily discounted, they have a small impact on the total value of the lifetime earnings and have little effect on the IRR. When the graduates repay without delay, the impact on the total present value is much larger, they pay off their loan earlier with a reduced effect of the discounting. When we consider increasing costs of education the IRR decreases slowly, because once again they take longer to repay and the effect of higher costs is smoothed by the discounting.

Looking at a ML, we first notice the lower IRR compared to an ICL under the current structure of the English system (threshold £15000 and costs £9000), for any level of the installment. Moreover, an ICL with a positive threshold always gives greater returns than a ML. Naturally a ML is always preferred when the ICL threshold is zero . The taxpayer subsidies under a ML are much lower compared to those under an ICL. With our estimated wage profile and the English structure of the scheme the graduates start to repay the loan five years after graduation. However, the length of the repayment period is fixed and shorter, thus the graduates cannot exploit the favorable effects of the heavy discounting. We notice that for the same repayment period (e.g. 5 years) and different combinations of loan and installment, the subsidy is always the same, around 6.3%, whereas the IRRs are decreasing from 6.72% to 6.33%.

Comparing an ICL under the current system (maximum loan £18000 and threshold £15000) and a ML under the pre-reform settings (maximum loan £12000, repayment 5 years, installment £2400), we notice a gain under an ICL of around 0.2% (6.76% versus 6.59%). The gain is even higher if we consider a loan of £18000 under a ML, for any level of the installment.

In conclusion, in terms of taxpayer subsidy our results are a little higher than those of Dearden et al., but we generate our age earnings profiles using a different method and the resulting distribution are not directly comparable. Compared to Heckman et al. (2008) our results are much lower for the same levels of education (i.e. high school vs higher education). They find IRRs ranging from a minimum of 14% to a max of 29%, according to the specification adopted. Instead we get IRRs ranging between 6% and 7%. One explanation is the differences in the data used (US vs UK), the estimates, and in the educational system analyzed. A more important reason is the different approach in computing the IRR, Heckman et al. consider the tuition as the only initial cost of the investment in education although they then discount the earnings net of taxes. In our case the cost of the investment in education is much higher since it is given by both foregone earnings and tuition, and this keeps the IRRs low.

3.5 Conclusion

In this chapter we first estimated the causal effect of education on earnings using a selection model. We considered four different education levels from NVQ1 to NVQ4. To correct for sample selection bias, we used, as exclusion restrictions, the raising of the leaving school age 1973 reform and the month of birth, which are variables directly affecting education and not earnings. We found estimates of the return to schooling controlling for selectivity which were lower than the existing empirical literature. In the second part of the chapter we used our estimations to generate working life wage profiles for those that have a high school qualification and those with a degree equivalent qualification. Our intent was to compute the internal rate of return of the investment in HE versus high school, under different HE financing schemes. In particular, we analyzed the cases of free education and public loans, such as income contingent loan and mortgage loan. Focusing on the structure of the English HE education financing system before and after the 2004 reform, we found that the IRR under an ICL is always higher than under a ML and the graduates exploit higher government subsidies due to the zero real interest on the loan.

Tab. 3.1: NVQ Equivalent Qualifications

NVQ level	Examples of academic qualifications	Examples of vocational qualifications
5	PhD Masters degree	PGCE (Teaching) Non-masters postgraduate quals
4	Undergraduate degree	Other teaching quals, HE below degree, RSA Higher, HNC/HND, Nursing
3	2+ A-levels 3+ SCE Highers AS-level or equivalent (4+)	GNVQ - advanced, RSA advanced dip/cert, OND, ONC, BTEC, City and Guilds Advanced Craft , Scot. Cert 6th year Studies, Higher national qualification or equivalent (Scotland) , ,(other) NVQ Level 3
2	1 A-level, 1-3 AS levels, 1-2 Highers 5+ GCSEs at A-C	GNVQ – intermediate, RSA diploma, City and Guilds - Craft , BTEC, SCOTVEC etc. first or general diploma , Scottish CSYS , Intermediate 2 Higher qualification (Scotland), (other) NVQ Level 2
1	1-4 GCSEs/SCE passes (but below Level 2)	GNVQ, GSVQ foundation level, CSE below grade1, GCSE below grade C, BTEC, SCOTVEC first or general certificate, SCOTVEC modules, RSA other, City & Guilds other , YT/YTP certificate , Key Skills, Basic Skills
Apprenticeships	n.a.	Modern and Traditional
Other	n.a.	Unspecified in the data

Tab. 3.2: NVQ Equivalent and log annual earnings - LFS 1993-2005

NVQ level	**mean**	**sd**	**per cent**
NVQ1	9.55	0.52	6.40
NVQ2	9.66	0.57	21.43
NVQ3	9.82	0.51	38.20
NVQ4	10.16	0.54	33.97
Total	9.88	0.57	100

Tab. 3.3: First Step - Selection model: ordered probit

Variable	Coefficient	(Std. Err.)
	Equation: nvqtot	
rosla	-0.339***	(0.005)
Feb	-0.016	(0.013)
Mar	-0.016	(0.013)
Apr	0.003	(0.013)
May	-0.007	(0.013)
Jun	-0.023*	(0.013)
Jul	-0.025*	(0.013)
Aug	-0.043***	(0.013)
Sep	0.029**	(0.013)
Oct	0.037***	(0.013)
Nov	0.017	(0.013)
Dec	-0.001	(0.013)
cut1	-1.700***	(0.029)
cut2	-0.734***	(0.029)
cut3	0.298***	(0.029)

N	169811
Log-likelihood	-207973.855
$\chi^2_{(31)}$	5808.617

LR[a] $\chi^2_{(12)}$	3950.61
Prob > χ^2	0.000
LR[b] $\chi^2_{(11)}$	68.73
Prob > χ^2	0.000

Additional controls: non white, regions of residence.

Significance levels : * : 10% ** : 5% *** : 1%

[a] Testing joint significance of rosla and months of birth.
[b] Testing joint significance of months of birth.

Tab. 3.4: Marginal Effects oprobit - rosla

Variable	dy/dx	(Std. Err.)
	Equation 1 : NVQ level = NVQ1	
rosla	0.039***	(0.00064)
	Equation 2 : NVQ level = NVQ2	
rosla	0.073***	(0.00041)
	Equation 3 : NVQ level = NVQ3	
rosla	0.012***	(0.005)
	Equation 4 : NVQ level = NVQ4	
rosla	-0.124***	(0.002)

Significance levels : * : 10% ** : 5% *** : 1%

Tab. 3.5: Second Step - Selection model

Dep var: log earnings	NVQ1 (Std. Err.)	NVQ2 (Std. Err.)	NVQ3 (Std. Err.)	NVQ4 (Std. Err.)
$age - 30$	0.0290***	0.0335***	0.0333***	0.0422***
	(0.001)	(0.001)	(0.001)	(0.001)
$(age - 30)^2$	-0.00112***	-0.00138***	-0.00128***	-0.00147***
	(0.000)	(0.000)	(0.000)	(0.000)
health problems	-0.186***	-0.227***	-0.184***	-0.205***
	(0.026)	(0.016)	(0.012)	(0.016)
non white	-0.163***	-0.129***	-0.168***	-0.068***
	(0.043)	(0.027)	(0.026)	(0.021)
married	0.162***	0.165***	0.128***	0.153***
	(0.013)	(0.008)	(0.006)	(0.006)
intercept	9.349***	9.401***	9.545***	9.596***
	(0.138)	(0.041)	(0.017)	(0.054)
λ	-0.0001	-0.084**	0.193***	0.151***
	(0.072)	(0.038)	(0.026)	(0.038)
N	169811			
$F(35, 45414)^a$	130.156			
Prob $> F$	0.000			

Additional controls: LFS survey years 1994-05, regions of residence.
Significance levels : * : 10% ** : 5% *** : 1%

[a] Testing parallel distribution NQ4 and NVQ3.

Tab. 3.6: OLS model

Dep var: log earnings	**NVQ1** (Std. Err.)	**NVQ2** (Std. Err.)	**NVQ3** (Std. Err.)	**NVQ4** (Std. Err.)
$age - 30$	0.0287***	0.0342***	0.0311***	0.0407***
	(0.001)	(0.001)	(0.001)	(0.001)
$(age - 30)^2$	-0.00112***	-0.00136***	-0.00128***	-0.00147***
	(0.000)	(0.000)	(0.000)	(0.000)
health problems	-0.186***	-0.227***	-0.184***	-0.205***
	(0.023)	(0.015)	(0.010)	(0.014)
non white	-0.163***	-0.128***	-0.171***	-0.070***
	(0.035	(0.021)	(0.020)	(0.017)
married	0.163***	0.163***	0.133***	0.154***
	(0.011)	(0.007)	(0.005)	(0.005)
intercept	9.349***	9.473***	9.55***	9.776***
	(0.035)	(0.021)	(0.015)	(0.020)
N	169811			

Additional controls: LFS survey years 1994-05, regions of residence.
Significance levels : $*$: 10% $**$: 5% $***$: 1%

Tab. 3.7: Predicted Logearnings

OPS	**nvq1**	**nvq2**	**nvq3**	**nvq4**
mean	9.617	9.650	9.828	9.920
sd	0.257	0.287	0.353	0.386
OLS	**nvq1**	**nvq2**	**nvq3**	**nvq4**
mean	9.612	9.733	9.815	10.087
sd	0.257	0.290	0.270	0.331

Tab. 3.8: IRR - HE versus High school graduates

		ICL		
Cost	£0	£9000	£12000	£18000
$d = £0$	7.12	6.66	6.49	6.23
$d = £15000$	7.12	6.89	6.83	6.76
$d = £20000$	7.12	6.95	6.95	6.95
		ML		
Cost	£0	£9000	£12000	£18000
$\varphi = £1800$	7.12	6.72	6.60	6.44
$\varphi = £2400$	7.12	6.68	6.59	6.36
$\varphi = £3600$	7.12	6.61	6.58	6.33

d: ICL threshold, φ: ML installment.

Tab. 3.9: Taxpayer Subsidy (%)- ICL and ML

		ICL	
Cost	£9000	£12000	£18000
$d = £0$	8.2	13.58	17.74
	(7)	(9)	(12)
$d = £15000$	29.44	31.16	54.11
	(23)	(25)	(30)
$d = £20000$	41.27	55.95	70.64
	(26)	(29)	(35)
		ML	
Cost	£9000	£12000	£18000
$\varphi = £1800$	6.28	8.25	11.11
	(5)	(7)	(10)
$\varphi = £2400$	5.27	6.28	8.77
	(4)	(5)	(7.5)
$\varphi = £3600$	3.72	4.24	6.28
	(2.5)	(3)	(5)

d: ICL threshold, φ: ML installment.

Repayment period in brackets.

Fig. 3.1: Average Earnings by NVQ and Age

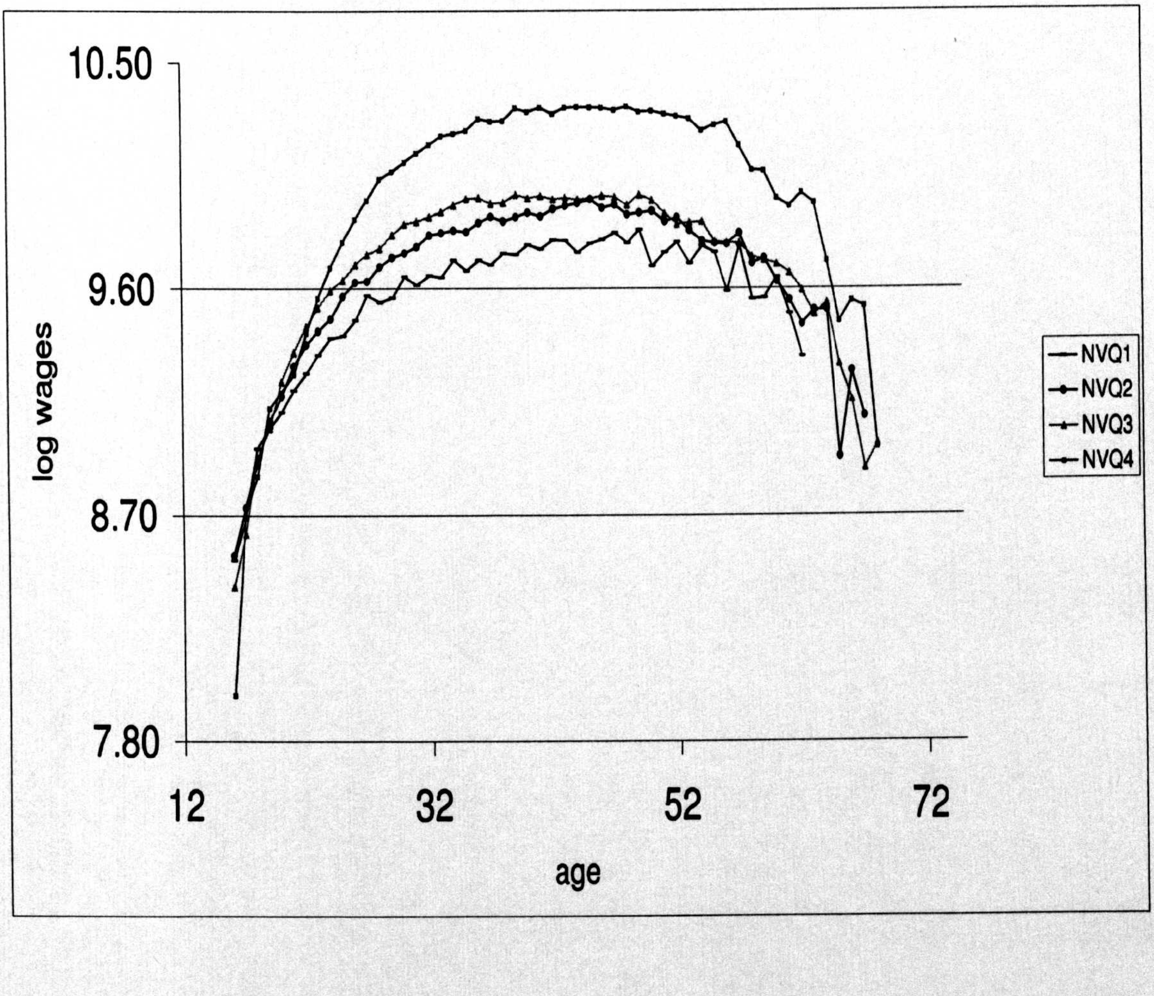


Fig. 3.2: Density of Treatment gains in population

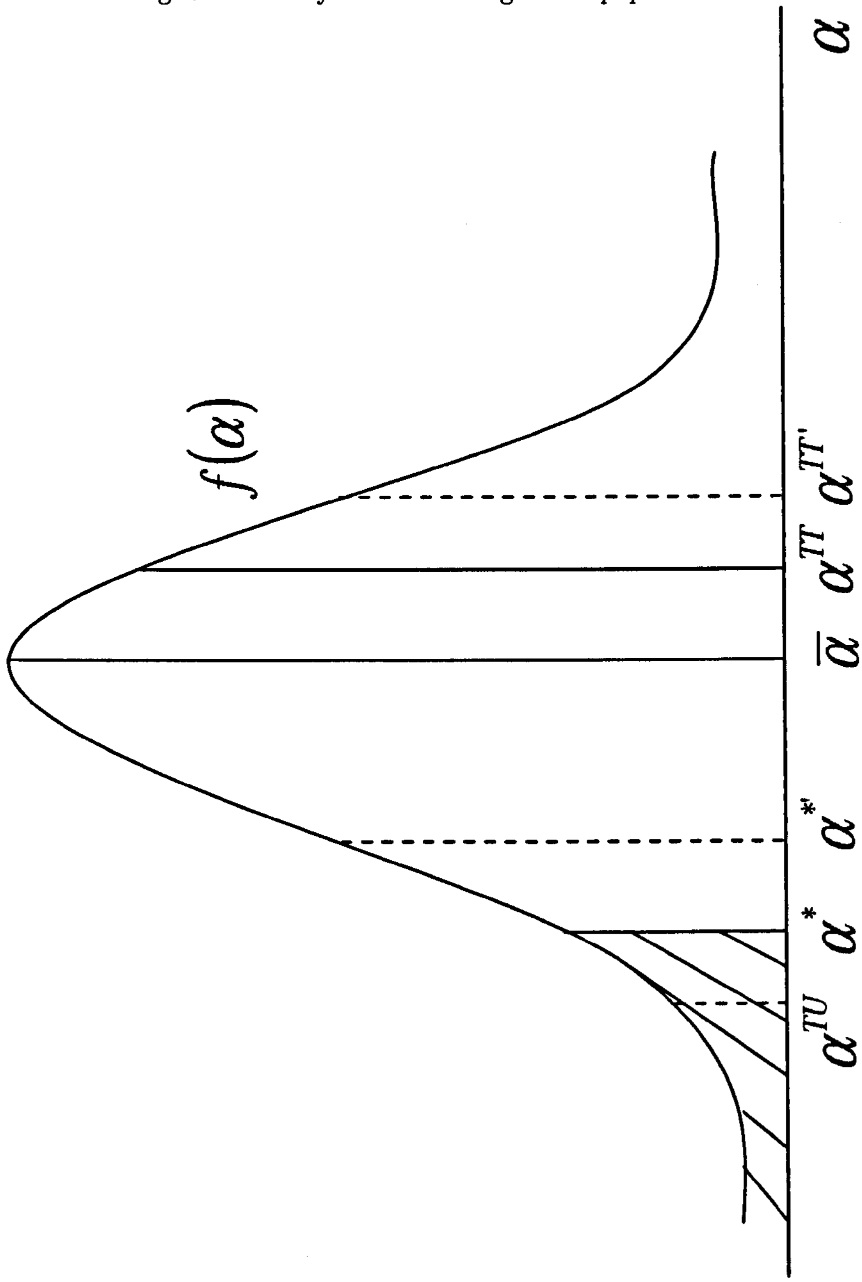
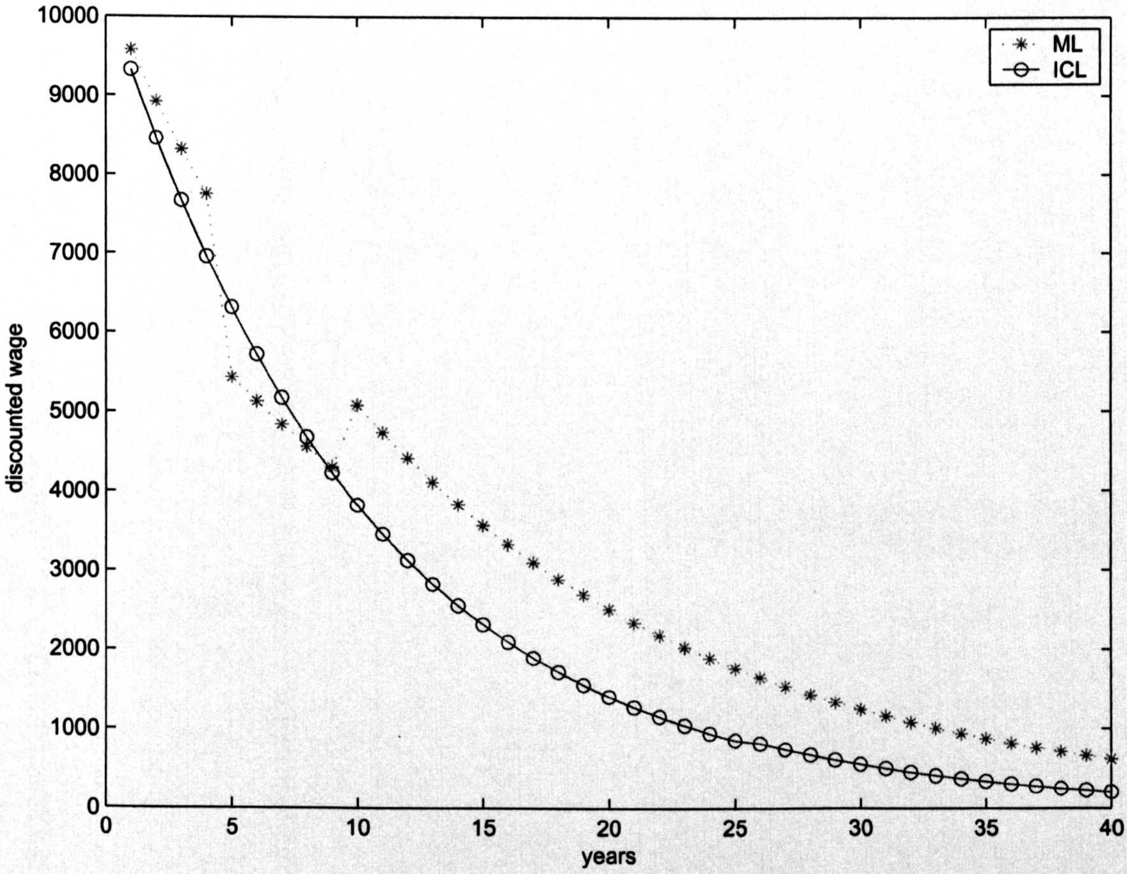


Fig. 3.3: PV Earnings - HE graduates



CHAPTER 4

4. DID THE SINGLE MARKET CAUSE COMPETITION IN EXCISE TAXES? EVIDENCE FROM EU COUNTRIES

4.1 Introduction

This Chapter provides an empirical analysis of horizontal competition in excise taxes trough the estimation of equations informed by the theory of strategic interactions among governments. In terms of Brueckner (2003) tax competition models belong to the category of resource-flow models, which relates the tax rates of a country to a weighted average of other countries's tax rates, in order to generate a reaction function (see Kanbur and Keen 1993, Lockwood 1993, Ohsawa 1999, Nielsen 2001). The goal of the growing empirical studies is to estimate such functions. In this chapter we use a panel data set of 12 EU countries over the period 1987-2004 and taxes on 7 commodities. We set up a spatial empirical model and estimate the impact of spatial tax competition on commodity taxation. Our purpose is to test whether the creation of the EU single market has generated a structural break giving impulse to higher strategic interaction among countries.

The Single European Act, which came into force in July 1987, initiated a vast legislative programme involving the adoption of hundreds of directives and regulations; which gradually established the Single Market amongst EU member states over a period up to the end of 1992. Two of most important provisions of the Single Market were (i) to allow individuals to import relatively large quantities of goods purchased abroad, which had previously been subject to the importing country's rate of tax; (ii) the abolition of physical border controls, being replaced by random spot checks.

Specifically, before 1 January 1993, all imports to EU countries (either from EU countries or from outside EU) were subject to a duty-free regime. That is, a small quantity of the product could be bought (usually, but not necessarily at duty-free shops in airports, on boats, etc.) without any tax payable, and then brought into the country. Any excess imports were taxed at the destination country's rate (in practice, both duties and VAT are imposed). It is important to note that duty-free allowances are quite small[1].

On 1 January 1993, all imports to EU countries from other EU countries were subject to no restrictions, except (i) that tax must have been paid in the country of purchase of the good; and (ii) that goods are not for resale. Condition (ii) is enforced by (generous) upper limits, plus random customs checks at borders. According to the UK Custom and Excise, for example, "if you bring back large quantities of alcohol or tobacco, a Customs Officer is more likely to ask about the purposes for which you hold the goods. This will most likely be the case if you appear at the airport with more than: 3200 cigarettes, 400 cigarillos, 200 cigars, 3kg of smoking tobacco, 110 litres of beer, 10 litres of spirits, 90 litres of wine, 20 litres of fortified wine i.e.: port or sherry". The above allowance is more than enough for the annual consumption of the average two-adult household, and depending on the item, 15-40 times the duty-free allowance. Moreover, imports in excess of these levels do not automatically trigger fines or prosecution: the levels are indicative only, and the onus is on Customs officials to prove smuggling.

Obviously, these changes in the rules create incentive both for legitimate tax-induced cross-border shopping and for smuggling. There is evidence that both these activities are occurring on a large scale at some borders. For example Belgians, French, and Germans have traditionally stocked up on cigarettes when passing through Luxembourg because of the low price of cigarettes there. In Luxembourg (population: 418 000) 5,165 million cigarettes were sold in 1997. It is estimated that 85% of these sales were due to cross border shopping (*Ministry of Finance, Brussels, personal communication*).

Furthermore, the rates of excise duty on alcoholic drinks and tobacco products in the UK are significantly higher than those in most other EU Member States, especially France. Five years after the start of the Single Market, the annual cost of revenue lost from legal cross border shopping was estimated to be £375 million for 1998: £55m from beer, £180m from wine, £50m from spirits, and £85m from tobacco products[2]. In 2003/04 in the UK, 10.5 billion cigarettes were successfully smuggled and the further 6.5 billion cross-border shopped implying a loss of 3.1 billion pounds of revenue for Government (see

[1] For example, in the UK, the allowances were: 200 cigarettes, or 100 cigarillos, or 50 cigars, or 250 grams of tobacco, 2 litres of still table wine, 1 litre of spirits or strong liqueurs over 22% of volume, or 2 litres of fortified wine, sparkling wine or other liqueurs.

[2] House of Commons Debate 26 November 1999 cc 254-5W.

Fig. 4.1: UK: increased use of cheap cigarettes and hand-rolling tobacco

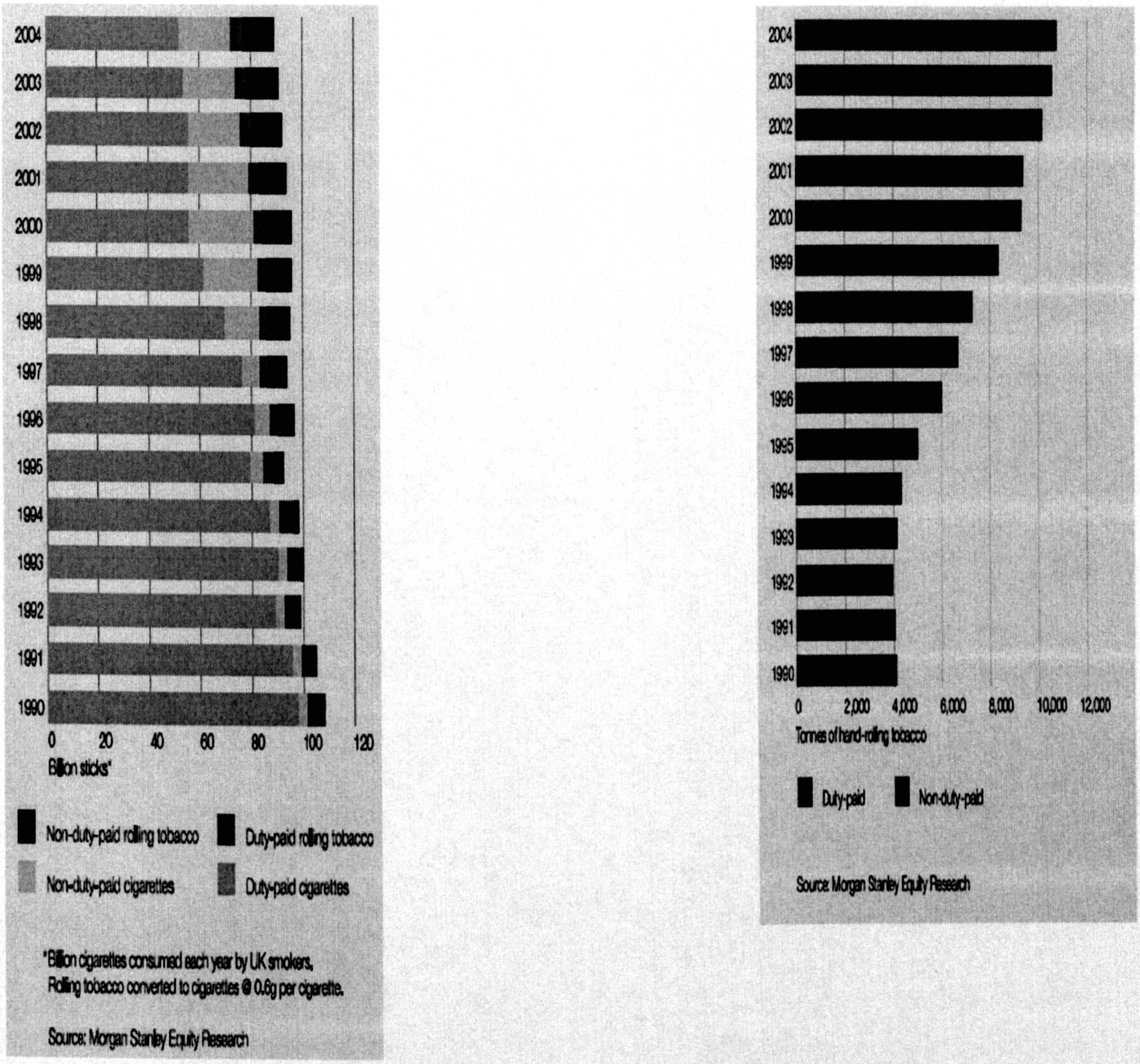


Figure 4.1). Moreover, 57% of hand-rolling tobacco was smuggled, while a further 14% was cross-border shopped (CMO Annual Report 2004).

What is less clear is whether these changes, and the subsequent excise revenue losses in high-tax countries, have caused tax competition between EU member states to occur or intensify. Certainly, the theory (Kanbur and Keen 1993, Lockwood 1993, Ohsawa 1999, Nielsen 2001)) suggest that this should happen. Technically, the Single Market resulted in a switch from destination to origin-based taxation of cross-border transactions by individuals. These models predict that tax competition only occurs with origin-based taxation. So, the models predict, unambiguously and generally, that we should observe competition in excise taxes between EU countries only after 1993.

Of course, strategic interaction can occur for other reasons e.g. yardstick competition, or common intellectual trends. So, the observable implication of the theory is that strategic interaction between EU countries in the setting of excise taxes should intensify after 1993. How much it intensifies depends on the scale of cross-border shopping and potential revenue losses from high taxes: as the above discussion indicates, these might be quite large. The key idea of the chapter is that completion of the single market can be interpreted as a kind of "natural experiment" that allows us to separate the effects of tax competition from other forms of strategic interaction. This work represents the first attempt, to our knowledge, to directly test this observable prediction.

In the first part of this Chapter, we propose a theoretical model that is an extension of the basic model of Nielsen (2001). We find the conditions for cross border shopping and the reaction functions of the government under three different market regimes: duty-free market before 1993, mixed market up to 1999, single market after 1999. The theoretical prediction is no strategic interaction before 1993, and increasing strategic interaction after 1993. In the second part of the Chapter we employ a balanced panel data set of 12 EU countries over the period 1987-2004 and excise taxes on seven commodities. While the control variables are taken from standard sources, the excise tax data were taken from the European Commission's Excise Duty Tables and Inventory of Taxes. Using this data set, we estimate an empirical model where the excise tax in a given country depends linearly on the weighted average of other countries' taxes, and a set of control variables. The equation is estimated separately on subsamples before and after 1993.

First, we find robust evidence that the coefficient estimates differ on the two subsamples. Specifically, looking across all seven taxes, and a variety of different weighting matrices suggested by the theory, the null hypothesis of equality of coefficients can be rejected in every single case. Second, again looking across all seven taxes, and a variety of different weighting matrices (i) coefficient measuring strategic interaction is *always* significantly positive post-1993, whereas before 1993, sometimes it is not; and (ii) whenever both are significant, the post-1993 coefficient is higher. From these results we can conclude that the introduction of the Single Market has determined a structural break in the commodity tax setting after 1993, and it has also intensified the tax competition among neighbouring countries.

Related literature is as follows. First, there is a small empirical literature on spatial interactions in excise taxes in the US (e.g. Nelson (2002), Rork (2003), Lockwood at al. (2005)). But in the US, there has been no "natural experiment" similar to the completion of the single market in the EU in recent times. Within the US, importation of commodities e.g. cigarettes subject to excises by individuals for private consumption is essentially unrestricted[3], meaning that the origin regime is firmly in place for these kinds of transactions.

There are also a couple of cross-country empirical studies of strategic interaction in commodity taxes (Egger et al. (2005), Evers et. al. (2004)). Egger et al. (2005) test some of the predictions of Ohsawa's theoretical model of commodity tax competition on commodity tax data for a panel of 22 OECD countries. But, unlike our study, they use an aggregate indicator of commodity taxation taken from national accounts data, which, relative to our chapter, obviously has the disadvantage that it does not measure very precisely the setting of individual tax instruments by governments. The paper by Evers et. al. (2004), in contrast, studies strategic interaction in the setting of diesel excises in EU countries, plus Norway and Switzerland, and so is closest to our chapter. But, almost by definition, the treatment of imports of fuel in the tank of a vehicle must be on an origin basis[4], and so completion of the single market has no predicted effect on the setting of this excise, except possibly through the introduction of a minimum EU excise; the latter effect is the focus of Evers et. al. (2004).

In Section 4.2 we illustrate our theoretical extension. In Section 4.3 and Section 4.3.2, we explain our econometric method and estimation procedure. Section 4.4 describes the data, Sections 4.5-4.6 show the results and Section 4.7 provides a robustness check. Section 4.8 concludes.

[3] The borderline of legality in the case of cigarettes is provided by the Contraband Cigarette Act of 1978, which prohibits single shipments, sale or purchase of more than 60,000 cigarettes not bearing the tax stamp of the state in which they are found.

[4] That is, even with border controls, customs officials have no way of knowing where the fuel in the tank of a vehicle has been bought.

4.2 Theoretical Model

The key facts exposed above show how intra-EU trade basically changed from destination to origin in 1993. But a small complication is that the duty-free was only abolished in 1999. We model these three regimes using a modified version of the Nielsen (2001) model. We consider two countries $i = 1, 2$ located on a line with a common border. Each country composed of identical households uniformly distributed on the line with unit density[5]. Country i is thus of both length and size n_i, see Figure 4.2. There is one taxed and one numeraire good. Each household has a unit demand for the taxed good: specifically, it is willing to pay $v > 0$ for up to 1 unit of the good, and nothing for additional units. Commodity taxes are specific taxes, levied at the rates t_i, the production cost is p_i and the total price is $q_i = p_i + t_i$. In this model, we have perfect symmetry so it does not really matter the size of the country and the direction of the cross-border shopping (CBS). Also a household living d distance from the border incurs travel costs dc of traveling to the border.

Given these assumptions we define three market regimes: duty-free (DF), single market (SM), mixed market (M). We procede, in each of the three cases, finding the cross-border shopping conditions and the reaction functions for a government that maximizes the social welfare, W. It would be desirable to have different population densities in two countries, but this gives rise to non differentiable reaction functions (see Kanbur and Keen). This would be difficult to estimate. In this literature people estimate linear reaction functions and we want a model that is consistent with that.

4.2.1 Duty-free Regime

Let the home country be 1, and assume w.l.o.g. that $t_1 > t_2$, so that there will be outward cross-border shopping (CBS). Each household in country *1* can buy at home at price $q_1 = p_1 + t_1$ or buy up to $0 \leq \alpha \leq 1$ units of the good at a duty-free shop at the border, at only the production cost p_2. Since a destination principle is applied any units in excess of α is taxed at the destination country's tax rate, $q_2 = p_2 + t_1$.

A household in the home country at distance $\widehat{d}_{DF}$ from the border is indifferent about

[5] Kanbur and Keen (1993) assume different population densities.

cross-border shopping, i.e. its surplus from buying at home is equal to the surplus obtained at the border. This $\widehat{d}_{DF}$ must satisfy

$$v - q_1 = v - \alpha p_2 - (1-\alpha)q_2 - c\widehat{d}_{DF} \tag{4.1}$$

Solving for $\widehat{d}_{DF}$ we get:

$$\widehat{d}_{DF} = \frac{(p_1 - p_2) + \alpha t_1}{c} \tag{4.2}$$

Then, all households at distance $d < \widehat{d}_{DF}$ will cross-border shop, and the others will buy at home.

Now consider the government. We assume that the government maximizes the social welfare, defined as an average utility, thus extending the Nielsen model. It is given by the net utility of those buying at home, plus the net utility of the cross-border shoppers and plus the total revenue weighted according the value of public good provision, θ. The value of theta has to be greater than one otherwise optimal provision should be zero in equilibrium and hence no taxes. The marginal cost of public funds is greater than one in the model because there is distortionary tax. If $\theta > 1$ the reaction function slope is greater than zero.

Generally, the revenue for the government of country 1 is the number of households minus the cross-border shoppers, i.e. $n_1 - \widehat{d}$ times the tax rate

$$R_1 = (n_1 - \widehat{d}_{DF})t_1.$$

The maximization problem is the following

$$\max_{t_1} W_1 = (n_1 - \widehat{d}_{DF})(v - q_1) + \int_0^{\widehat{d}_{DF}} (v - \alpha p_2 - (1-\alpha)q_2 - cx)dx + \theta R_1. \tag{4.3}$$

The first-order condition implicitly defining the tax reaction function is:

$$\begin{aligned} \frac{\partial W_1}{\partial t_1} =& -(n_1 - \widehat{d}_{DF}) - (1-\alpha)\frac{\partial q_2}{\partial t_1}\widehat{d}_{DF} \\ &+ \theta(n_1 - \widehat{d}_{DF}) - \theta\frac{\partial \widehat{d}_{DF}}{\partial t_1}t_1 = 0. \end{aligned} \tag{4.4}$$

Proposition 1. *The reaction function for the duty free regime is:*

$$t_1 = \frac{c(\theta-1)n_1 - (\theta-\alpha)(p_1-p_2)}{\alpha(2\theta-\alpha)} \tag{4.5}$$

Proof. *See Appendix 4.9*

4.2.2 Single Market Regime

Each household in country *1* can buy at home at price $q_1 = p_1 + t_1$ or at the border, at price $q_2 = p_2 + t_2$. A source principle is applied and there is no duty free.
A household in the home country at distance $\widehat{d}_{SM}$ from the border is indifferent about cross-border shopping i.e. its surplus from buying at home is equal to the surplus obtained at the border.

$$v - q_1 = v - q_2 - c\widehat{d}_{SM} \tag{4.6}$$

Solving for $\widehat{d}_{SM}$ we get:

$$\widehat{d}_{SM} = \frac{(p_1-p_2)+(t_1-t_2)}{c} \tag{4.7}$$

Then, all households at distance $d < \widehat{d}_{SM}$ will cross-border shop, and the others will buy at home. We notice that equation (4.7) differs from the duty free condition (4.2) not only for the absence of α but also for the dependence on country 2 tax rate, t_2.
The government, whose revenue is $R_1 = (n_1 - \widehat{d}_{SM})t_1$, maximizes a social welfare function, modified according the new market conditions:

$$\max_{t_1} W_1 = (n_1 - \widehat{d}_{SM})(v - q_1) + \int_0^{\widehat{d}_{SM}} (v - q_2 - cx)dx + \theta R_1. \tag{4.8}$$

The first-order condition implicitly defining the tax reaction function is:

$$\frac{\partial W_1}{\partial t_1} = -(n_1 - \widehat{d}_{SM}) + \theta(n_1 - \widehat{d}_{SM}) - \theta\frac{\partial \widehat{d}_{SM}}{\partial t_1}t_1 = 0. \tag{4.9}$$

Proposition 2. *The reaction function for the single market regime is*

$$t_1 = \frac{c(\theta-1)n_1 - (\theta-1)(p_1-p_2)}{2\theta-1} + \frac{\theta-1}{2\theta-1}t_2 \tag{4.10}$$

Proof. *See Appendix 4.9.1*

Note that now, countries do react to each others' taxes: the slope of the reaction function is $\frac{\theta-1}{2\theta-1}$. This slope has to be between one half and zero (which is satisfied) for $\theta > 1$.

4.2.3 Mixed Market Regime

Each household in country *1* can buy at home at price $q_1 = p_1 + t_1$ or buy up to $0 \leq \alpha \leq 1$ units of the good at a duty-free shop at the border, at only the production cost p_2. Since in this regime a source principle is applied, any units in excess of α that are bought in the foreign country 2 are taxed at that country's tax rate, $q_2 = p_2 + t_2$.
A household in the home country at distance $\widehat{d}_M$ from the border is indifferent about cross-border shopping i.e. its surplus from buying at home is equal to the surplus obtained at the border.

$$v - q_1 = v - \alpha p_2 - (1-\alpha)q_2 - c\widehat{d}_M \tag{4.11}$$

solving for $\widehat{d}_M$ we get

$$\widehat{d}_M = \frac{(p_1 - p_2) + t_1 - (1-\alpha)t_2}{c} \tag{4.12}$$

Then, all households at distance $d < \widehat{d}_M$ will cross-border shop, and the others will buy at home.
We notice that the presence of α makes this regime similar to the duty free, however the application of an origin principle maintains the dependence of $\widehat{d}_M$ on country 2 tax rate, as in the single market regime.
The government, whose revenue is $R_1 = (n_1 - \widehat{d}_M)t_1$, maximizes a social welfare function

$$\max_{t_1} W_1 = (n_1 - \widehat{d}_M)(v - q_1) + \int_0^{\widehat{d}_M} (v - \alpha p_2 - (1-\alpha)q_2 - cx)dx + \theta R_1. \tag{4.13}$$

The first-order condition implicitly defining the tax reaction function is:

$$\frac{\partial W_1}{\partial t_1} = -(n_1 - \widehat{d}_M) + \theta(n_1 - \widehat{d}_M) - \frac{\theta t_1}{c} = 0. \tag{4.14}$$

Proposition 3. *The reaction function for the mixed market regime is*

$$t_1 = \frac{c(\theta-1)n_1 - (\theta-1)(p_1-p_2)}{2\theta-1} + \frac{(\theta-1)(1-\alpha)}{2\theta-1}t_2 \tag{4.15}$$

The Proof is similar to the Single Market regime. Note that the slope of the reaction function is now positive but smaller than in the SM regime.

4.3 Empirical Implementation

In the theoretical Section we found that origin based commodity taxation generates positively sloped reaction functions i.e. that in country i, the excise tax, t_i, is an increasing - in fact, piecewise linear - function of the tax rate in the other countries, t_j, $j \neq i$. Moreover, under realistic assumptions[6], the response of t_i to t_j will be non-zero only if i and j are contiguous i.e. share a common border. Finally, this response will depend on the length of the border between i and j and also on the population sizes in the two countries (Ohsawa, 1999).

Our empirical specification is therefore the following:

$$t_{is} = f_i + \sum_{j\neq i} \beta_{ij} t_{js} + \delta' \mathbf{Z}_{is} + \epsilon_{is} \tag{4.16}$$

where $i = 1,..n$ denotes a country, and $s = 1,...S$ a time-period, f_i a country fixed effect, $\mathbf{Z}_{is}$ a $k \times 1$ vector of relevant characteristics of country i at time s, and δ a $k \times 1$ vector of coefficients. However, this cannot be estimated as it stands, as there are too many parameters β_{ij} to be estimated. The usual approach is to modify (4.16) as:

$$t_{is} = f_i + \phi \sum_{j\neq i} \omega_{ij} t_{js} + \delta' \mathbf{Z}_{is} + \epsilon_{is} \tag{4.17}$$

where the ω_{ij} are exogenously chosen weights, that aggregate the tax rates in other countries into a single variable, which has coefficient ϕ . The ω_{ij} are usually normalized so that $\sum_{j\neq i} \omega_{ij} = 1$. This is a widely used procedure and there is considerable discussion of the appropriate weights in the literature e.g. Brueckner(2003)[7].

[6] That is that prices are such that consumers do not wish to drive through a third country to buy in a low-tax country.

[7] See Brueckner (2003) for a survey of empirical techniques.

Our key theoretical hypothesis is that ϕ changes with the regime, being higher when the Single Market regime is in place. In fact, if only tax competition is present, and no other form of strategic interaction, we expect $\phi = 0$ before 1993. We test for this dependence by estimating (4.17) on two sub-samples, 1987-1992 and 1993-2004. Let the estimates of ϕ on the earlier and later subsample be ϕ_1, ϕ_2 respectively. So, our basic hypothesis is that $\phi_2 > \phi_1$. The theoretical model does predict intercept shift and we allow it, so intercept are also changed after 1993.

Stacking the observations by time and space we can rewrite the equation (4.17) in matrix form and specify what is known as general spatial process model:

$$\begin{aligned} t &= \phi \mathbf{W}_1 t + \mathbf{Z}\delta + \epsilon \\ \epsilon &= \lambda \mathbf{W}_2 \epsilon + \mu. \end{aligned} \tag{4.18}$$

where μ is a well-behaved normal error. The error structure allows spatial correlation in shocks across countries - common shocks to EU countries. Note all diagonal elements of $\mathbf{W}$ are zero and a representative off-diagonal element is ω_{ij}. A reliable estimation of the separate parameters ϕ and λ may be difficult (see Anselin, 1988; et al.) and for identification purposes require the definition of two weighting matrices.

There are three special cases derived imposing restrictions to model (4.18). Setting $W_2 = 0$, and $W_1 = W$, we obtain the mixed spatial autoregressive model (SAR)

$$\begin{aligned} t &= \phi \mathbf{W} t + \mathbf{Z}\delta + \epsilon \\ \epsilon &\sim N(0, \sigma^2 I). \end{aligned} \tag{4.19}$$

The parameter ϕ is a coefficient on the spatially lagged dependent variable, Wt, and the parameter δ reflects the influence of the explanatory variables on variation in the dependent variable t. The model is termed as mixed regressive - spatial autoregressive model because it combines the standard regression model with a spatially lagged dependent variable, reminiscent of the lagged dependent variable model from time-series analysis.

Letting $W_1 = 0$, $W_2 = W$ in equation (4.18) produces a regression model with spatial autocorrelation in the disturbances, called spatial error model (SEM)

$$\begin{aligned} t &= \mathbf{Z}\delta + \epsilon \\ \epsilon &= \lambda \mathbf{W}\epsilon + \mu. \end{aligned} \tag{4.20}$$

The parameter λ is a coefficient on the spatially correlated errors, analogous to the serial correlation problem in time series models, and μ is well-behaved.

A related specification is the spatial Durbin model (as referred by Anselin, 1988 and Lesage, 1999), where W is the spatial weight matrix and the parameter λ is a coefficient on the spatial lag of the dependent variable. An additional set of explanatory variables is added to the model by constructing a spatial lag of the explanatory variables using the matrix product $W\mathbf{Z}$, with associated parameter δ_2. This set of variables represent explanatory variables constructed as averages from neighboring observations. The WZ's are artificial regressors - they have no meaning - they are just an alternative way of expressing the spatial error model.

$$\begin{aligned} t &= \lambda \mathbf{W}t + \mathbf{Z}\delta_1 + \mathbf{W}\mathbf{Z}\delta_2 + \epsilon \\ \epsilon &\sim N(0, \sigma^2 I). \end{aligned} \tag{4.21}$$

The latter model is formally equivalent to a spatial error model, rewriting the error specification in (4.20) and replacing in t[8]:

$$\begin{aligned} \epsilon &= (I - \lambda W)^{-1}\mu \\ t &= \mathbf{Z}\delta_1 + (I - \lambda W)^{-1}\mu. \end{aligned}$$

that can be rewritten as

$$\begin{aligned} (I - \lambda W)t &= (I - \lambda W)\mathbf{Z}\delta_1 + \mu \\ t &= \lambda W t + \mathbf{Z}\delta_1 - \lambda \mathbf{W}\mathbf{Z}\delta_1 + \mu. \end{aligned} \tag{4.22}$$

The model in (4.21) could be viewed as imposing a restriction that $\delta_2 = -\lambda\delta_1$ (known as the common factor hypothesis) on the model in (4.22). This model has been used in

[8] We assume $\delta = \delta_1$.

spatial econometric analysis by Pace and Barry (1998).

4.3.1 Weighting Matrix

Finally, we turn to the specification of the weighting matrix. Following the theoretical literature, and several empirical studies, our baseline weighting matrix uses contiguity weights. These weights capture the idea that tax bases are mobile only between geographically neighbouring countries. Specifically, we define *contiguity weights* as:

$$\omega_{ij} = \begin{cases} \frac{1}{n_i} & \text{if } j \in N_i \\ 0 & \text{if } j \notin N_i \end{cases}$$

where N_i is the set of states that border state i, and $n_i = \#N_i$. This assigns equal weight to all countries on the border of country i, and weight zero to the other countries. The matrix is row-normalized to have sum equal one.

It is also difficult to specify the spatial weights matrix when a country is an island, or part of an island, or has no direct EU neighbors. These problems arise for three of the eight countries in our data-set such as the UK, Ireland, or Greece[9]. A strict imposition of contiguity weights for the UK, for example, would give only Ireland as the neighbor for the UK, and vice versa. This is inaccurate, because it does not account for the considerable tax-induced cross-border shopping between the UK and France. Our solution was to say that if i is an island, a positive contiguity weight was given to country j when j could be directly reached from country i by crossing only over water i.e. without passing through a third country. But obviously, this is a rather arbitrary extension of the idea of contiguity.

Moreover, earlier work by Devereux et al. (2007) shows that the refining contiguity weights by adjusting for population around border is unimportant.

We consider a second very simple weighting matrix: weights are assumed to be *uniform* i.e. $\omega_{ij} = \frac{1}{n-1}$, all i, j. While giving a useful benchmark, this is unlikely to work well, especially for commodities the tax base is mobile due to cross-border shopping and smuggling.

A modification of contiguity weights often considered in the literature are distance

[9] If i is an island, a positive contiguity weight was given to country j when j could be directly reached from country i by crossing only over water i.e. without passing though a third country. But obviously, this is a rather arbitrary extension of the idea of contiguity.

weights. Contiguity can be constructed artificially using a Delauney triangle algorithm, which subdivides an area into triangles. The Delauney triangulation of a point set is a collection of edges satisfying an "empty circle" property: for each edge a circle can be found containing the edge's endpoints but not containing any other points[10]. Once an algorithm such as Delauney triangulation has been applied to determine neighbors to each country, a weight matrix can be defined. For all countries i and j, $w_{ij} = 1$ if i and j are contiguous observations, and zero otherwise. The resulting weight matrix W is then row normalized in the usual fashion. Kelley Pace has written the code (FDELW2.m) to generates this form of weighting matrix[11]. The latitude and the longitude coordinate of the capital of each country are used to allow the FDELW2 Matlab function to convert the Delauney algorithm results into a contiguity matrix. This approach is followed to estimate W, in particular we consider two specifications, the first yields a symmetric spatial weight matrix and the second specification yields a row-stochastic spatial weight matrix.

Finally, following Case, Hines and Rosen (1992) we adopt a robustness check. Specifically, we want to test whether *any* weighting matrix will generate positive strategic interaction between countries; if not, i.e. only contiguity-type weighting matrices work, this is more evidence that the strategic interaction we observwe is no due to something inherent in the econometric tax competition. Following Case, Hines and Rosen (1992), we construct a weighting matrix based on a "nonsense" procedure; $\omega_{ij} = 1$ only if the name of country j comes after country i in the alphabet.

4.3.2 Estimation procedure

The OLS estimation is inappropriate for models that incorporate spatial dependence. As explained by Anselin (1988) in presence of spatially lagged dependent variable (SAR), unlike what holds for the time series counterpart of this model, the spatial lag term Wt is correlated with the disturbances, even when the latter are i.i.d.[12]. Consequently the OLS

[10] Delaunay triangulation: given a set of points in a plane, a Voronoi polygon about point P is the set of points closer to, or as close to, P as any other point. The Delaunay triangulation is the dual of the Voronoi tesselation, and has the following properties: no point is contained in the circumcircle of any triangle; maximises the minimum angle for all triangular elements (note: we would like to minimise the maximum).

[11] The code is included in her Spatial Statistics toolbox 2.0 for Matlab, which can be downloaded from www.spatial-statistics.com.

[12] Writing the reduced form of (4.19), $t = (I - \phi W)^{-1} Z\delta + (I - \phi W)^{-1}\epsilon$, each inverse can be expanded into an infinite series which implies that shocks at any location affect all other locations, through a spatial multiplier effect (or, global interaction).

estimator is biased and inconsistent for the parameters of the spatial model.

In the presence of spatial residual autocorrelation (SEM) the properties of OLS are more in line with the time series results. OLS remains unbiased, but it is no longer efficient and the classical estimators for standard errors will be biased, due to the non diagonal structure of the disturbance variance matrix.

Given these problems the attention has been focused on two estimation methods: Instrumental Variables and Maximum Likelihood. Using IV we can only estimate the SAR model since we cannot model IV in order to take into account of the spatial correlated errors. At the first stage, the endogenous variable $\sum_{j\neq i} \omega_{ij} t_{js}$ is instrumented by the weighted averages of the controls i.e. $\sum_{j\neq i} \omega_{ij} z^c_{js}$, for control $c = 1, ..k$. So, our maintained hypothesis is that in country the controls are exogenous to the setting of excise taxes on tobacco, alcohol products and petrol; given our list of controls, this seems reasonable.

The maximum likelihood approach permits the estimation of SAR, SDM and SEM, it involves the maximization of a log likelihood function (concentrated with respect to δ and σ) with respect to the spatial autoregressive parameter. For the SAR model (4.19) the steps are enumerated in Anselin (1988) as:

1. perform OLS for the model: $t = \mathbf{Z}\delta_0 + \epsilon_0$
2. perform OLS for the model: $Wt = \mathbf{Z}\delta_L + \epsilon_L$
3. compute the residuals: $e_0 = t - \mathbf{Z}\widehat{\delta}_0, \quad e_L = Wt - \mathbf{Z}\widehat{\delta}_L$
4. given e_0 and e_L, find ϕ that maximizes the concentrated likelihood function

$$L_C = -(n/2)ln(\pi) - (n/2)ln(1/n)(e_0 - \phi e_L)'(e_0 - \phi e_L) + ln|I - \phi W|$$

5. given $\widehat{\phi}$ that maximizes L_C, compute $\widehat{\delta} = \widehat{\delta}_0 - \phi\widehat{\delta}_L$ and $\widehat{\sigma}^2_\epsilon = (1/n)(e_0 - \phi e_L)'(e_0 - \phi e_L)$.

This same approach can be applied to the SDM (4.21) repeating the steps using the augmented matrix $\widetilde{\mathbf{Z}} = [\mathbf{Z} \quad \mathbf{WZ}]$. The SEM model estimation is a slightly more complex estimation since a one-time optimization of L_C with respect to λ is not enough to obtain ML estimates, and an iterative procedure that involves to carry out estimated GLS can be followed (Anselin, 1988).

In general, the concentrated likelihood for the spatial process is a nonlinear function of one parameter only, therefore a direct search method is applied. The values of the spatial parameter are restricted to the range -1 to 1, and a local maximum is obtained evaluating the L_C at small intervals over that range. Repeating the searches around the initial optimum will lead to more precise estimates.

4.4 The Data

We construct a balanced panel of data from 12 EU countries over 17 years, 1987 and 1989-2004 inclusive[13]. Data definitions are given in the Appendix 4.10 and summary statistics in Table 4.1. We consider only the countries which were members of the EU in 1987, excluding those that joined the EU later on. The final sample size has 204 observations. Data are based on the Excise Duty Table issued by the European Commission[14], cross-checked against the available issues of the Inventory of Taxes (only available for 1994, 1999, 2002). In the case of a discrepancy, which were not many, we took the data from the Inventory of Taxes as being authoritative, as this data is directly supplied by the member countries.

We study seven kinds of products: cigarettes, loose or rolling tobacco[15], petrol, still wine, sparkling wine, beer[16], and ethyl alcohol, the last being effectively an excise tax on spirits, such as whiskey, brandy, etc. All of these products, except for cigarettes and tobacco, are only subject to a specific or unit excise tax i.e. levied per unit of physical quantity. Where there are several rates of tax e.g. standard and reduced rates, we use only the standard rates.

In the case of cigarettes and tobacco, all countries also levy an ad valorem excise tax. Moreover, depending on the country, either the specific or ad valorem component of the tax can be the more important one and so we cannot safely ignore either. On the other hand, we do not have data separately on the retail price of cigarettes and tobacco, so we

[13] Data are not available for the year 1988.

[14] In the Appendix we discuss all the problems we had in collecting the data.

[15] There are in fact two categories of smoking tobacco which can be taxed differently by some countries (fine and other) . When the two taxes were different, we took a simple average of the two.

[16] In the case of beer, there were two kinds of physical unit used in the Excise Duty Tables: degree Plato, and degree of alcohol by volume. According to Directive 92/84/EEC it has been accepted that a tax of 0.748 Euros Plato is equal to a tax of 1.87 Euros alcohol by volume, and we applied this conversion factor.

are constrained by data in the Excise Duty Tables and we consider only the specific excise tax. We report for each of the seven goods the time series plot of the tax rates in Euro[17] and in national currency, in order to highlight the differences due to the fluctuations of the exchange rates. In both cases, the taxes are in nominal terms.

When we run the regressions, adjusted for inflation by dividing through by the RPI for the relevant country. We do this for two reasons. First, conceptually, rational governments will be concerned with the real, rather than nominal, value of the taxes they set. Second, perhaps for this reason, we did not find any evidence of strategic interaction when we used nominal taxes. Our interpretation is that the tax competition is not passive (accidental, through inflation) but it is sophisticated.

Another point is that the regressions are performed using the dependent variables in Euro. In estimating the determinants of the taxes, we need to control for other factors. We use a parsimonious set of controls that can be found in most of the existing empirical literature. These include: economic variables (GDP, and government consumption expenditure as a percentage of GDP) and demographic variables (population). All of these variables are taken from World Bank WDI (2004). We add two political dummy variables for the ideological orientation of governments, derived from the Schmidt Index variable included in Comparative Political Data Set 1960-2004 (Armigeon et al., 2006). We decomposed the Schmidt Index defining a dummy for right-wing cabinets, a dummy for stand-off between left and right cabinets, and a dummy for left-wing cabinets. The second dummy is used as reference category in the estimation. The descriptive statistics for the controls are also given in Table 4.2.

Tab. 4.1: Summary Statistics - Excise Duties in Euro

Variable	Mean	Std. Dev.	Min.	Max.
beer	2.386	2.501	0.209	9.24
ethyl alcohol	1455.779	869.194	42.33	3925
still wine	55.69	87.663	0	281.48
sparkling wine	121.931	158.26	0	562.96
cigarettes	32.721	37.15	0.61	154.34
tobacco	22.588	37.8	0	165.74
petrol unleaded	438.206	123.222	139.83	789.85
N		204		

[17] Before 1999, we converted national currencies to ECU using the exchange rates provided in the Excise Duty Tables.

Tab. 4.2: Summary Statistics - Control variables

Variable	Mean	Std. Dev.	Min.	Max.
poptot	292.13	268.52	3.72	825.41
$gdp - pc$	365.40	561.59	73.95	2484.32
govcon	19.86	3.30	13.74	26.76
govright	0.44	0.50	0.00	1.00
govleft	0.26	0.44	0.00	1.00

4.5 IV - Results

The results are given by commodity in Tables 4.3-4.9. All have the same format i.e. for each of the four weighting matrices, estimates of equation (4.17) over the period 1987-92, and 1993-2004 are given separately. After the R^2 statistics a number of diagnostic statistics are reported. The Pagan and Hall's (1983) is a test of heteroskedasticity for instrumental variables (IV) estimation. This statistic is distributed as chi-squared under the null of no heteroskedasticity, and under the maintained hypothesis that the error of the regression is normally distributed. The F test in the first stage of the estimation tests the null hypothesis of whether the instruments are not correlated with the endogenous variable. A rejection means that there is such a correlation. The Anderson canonical correlations likelihood-ratio test of whether the equation is identified[18]. The statistic provides a measure of instrument relevance, and rejection of the null indicates that the model is identified. The Hansen-Sargan test is a test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as chi-squared in the number of overidentifying restrictions. A rejection casts doubt on the validity of the instruments. We finally report the results of the Chow test that test a structural change in the parameters in the two subsamples considered, before 1993 and after 1993. We have to say that time series is not long and large squared residuals in the unrestricted model make easy to find structural breaks (also after 1993).

[18] The null hypothesis of the test is that the matrix of reduced form coefficients has rank=K-1, where K=number of regressors, i.e, that the equation is underidentified. Under the null of underidentification, the statistic is distributed as chi-squared with degrees of freedom=(L-K+1), where L=number of instruments (included+excluded).

4.5.1 Beer

Comparing Figure 4.3 and Figure 4.4 is evident the effect of the tax fluctuations on the tax rates. We first notice that all the countries have positive excise duties on beer, and only PT and UK have a clear constant increasing trend. Some countries (such as ES, DE, FR and LU) react to the market unification with strong and straight increase of the tax rates in 1993. Other countries (DK, IT, NL) have the opposite behavior decreasing their tax rates.

The first two columns in table 4.3 give estimates of equation (4.17) over the period 1987-92, and 1993-2004, for our baseline case of contiguity weights. First, note that we can reject the null hypothesis that the two regressions are the same. Generally speaking, the second regression performs better than the first; the controls are jointly significant, we find heteroskedasticity therefore we use robust standard errors. The F test in the first stage shows absence of correlation between the endogenous variable and the instruments. The rejection of the Anderson test indicates that the model is identified, and finally the model passes the Hansen test of OID restrictions that means that the instruments used are valid.

Turning on the key coefficient of the variable WTAX, it is insignificant before 1993, and larger and significant after 1993, consistent with the main hypothesis of this chapter. Also, note that as the taxes are in levels, the interpretation of the coefficients is that after 1993, for example, a 1 Euro increase in the specific tax on beer in all other countries causes a given country to raise its tax by about 0.3 Euros. Looking across the different weighting matrices, we see that this picture is similar for Wdis1, Wdis2. But now, the slope of the reaction function is much larger. Finally, in the case of Wrand, we see that the spatial lag coefficient is only positive before 1993, and negative, very small in absolute value, and insignificant afterwards. The difference between the results based on distance weights on the one hand, and random weights on the other, gives us some confidence that we are picking up some real effect.

4.5.2 Ethyl alcohol

Comparing Figure 4.5 and Figure 4.6, we notice an increasing tax rate trend in any country. After 1993, the main reaction is to put up the rate, some countries (e.g. LU, FR, IT) show big jumps. DK and DE reduce their rate after 2000, IE and NL instead increase it in the same period.

Looking at the first two columns of Table 4.4 we notice first that there is a structural break after 1993 since the equations pass the Chow test. Focusing on the second equation, the controls are jointly significant and the standard errors are robust to heteroskedasticity. The F test in the first stage says that the instruments are correlated with the WTAX variable, and the Anderson test indicates that the model is identified. The Hansen test is not rejected therefore the excluded instruments are correctly excluded from the estimated equation.

The coefficient of WTAX after 1993 is significant at 10% and almost the double of the coefficient before 1993, although the latter is not significant. The same behavior is confirmed with a uniform matrix, with Wdis1 and Wdis2 instead the coefficient of WTAX is decreasing after 1993, however with Wdis2 the spatial lag coefficient before 1993 is not significant. When we use Wrand the coefficient is positive and significant before 1993 but is not significant after 1993, this could confirm our expectation to pick some real effect.

4.5.3 Still wine

Comparing Figure 4.7 and Figure 4.8, six countries over twelve have zero tax rate on wine still, LU reduced the excise to zero in 1993. Therefore, it is not easy to analyze a situation of tax competition. Moreover, among the countries with positive rates, we notice that the exchange rate plays an important role. Looking at the rates in national currency, the UK only increased slowly the rate over the 17 year-period, some of the others (BE, IE, NL, FR) increased it just once or twice after 1992 and some countries (DK, LU) decreased the rate in 1993 and keeping it constant after.

The first two columns of Table 4.5 show the equation using a contiguity weight matrix. The equation after 1993 performs much better, note first the Chow test is rejected and there is a structural break after 1993. The controls are joint significant and the standard errors correct for heteroskedasticity according to the Pagan test. We reject the F test of no correlation between instrument and endogenous variables. The equation pass both tests on IV validity, i.e. Anderson and Hansen tests. Turning on the spatial lagged coefficient: WTAX is increasing, positive and significant after 1993, before it is negative and insignificant. Looking across the other matrix specifications the same picture is confirmed. The WTAX coefficient is always positive, highly significant and increasing after 1993. A further confirm is given by the estimation with Wrand, since the WTAX coefficients are negative and insignificant.

4.5.4 Sparkling Wine

The tax rates of the excise on sparkling wine reproduce a similar pattern as for still wine, that is some country with zero rate and others increasing or reducing the rates after 1993 (see Figure 4.9 and Figure 4.10). Observing, for example, the rate for DE or IE it is clear also the effect of the exchange rate's fluctuations, that seem to give a trend when in reality the rates are kept constant for many years.

Looking at the second column of Table 4.6 note that estimation of the equation after 1993 gives better results than the equation before 1993. We reject the Chow test therefore the two regressions are not the same. The controls are joint significant and the s.e. robust. All the diagnostic tests pass, this means validity of the instruments and identification of the model.

Using a contiguity matrix, the coefficient of WTAX after 1993 is positive and significant, while before 1993 it is negative and insignificant. Across the other matrix specifications we get the same results, further confirmed using Wrand where the coefficient after 1993 is negative and not significant.

4.5.5 Cigarettes

We observe in Table 4.1 that on average among the 12 countries the specific tax rate on cigarettes is around 33 Euros, but with large variations from a minimum of 60 cents to a max of 154 Euros for 1000 pieces. Looking at Figure 4.11 and Figure 4.12, we notice a constant increase of the tax rates from 1987 to 2004. In general, the increase is in the range of $10 - 15$ Euros in 17 years, however for Ireland and the UK the excises have been augmented from 50 to 150 Euros, and the effect of the exchange rate is small since the same increase can be observed in national currency.

The first two columns of Table 4.7 give the estimates using a contiguity matrix, the second equation performs better. We notice first that there is a structural break after 1993. Moreover the controls are joint significant and the s.e. corrected for heteroskedasticity. Observing the diagnostic test, we reject the F test in the first stage at 5%, there is correlation between the instruments and WTAX. We reject also the Anderson test, the model is identified. The equation passes the Hansen test of OID restrictions that means that the instruments used are valid.

Turning on the coefficient of the spatial lagged dependent variable: it is positive, significant and increasing after 1993. A 1 Euro increase in the specific tax rate on cigarettes in all other countries causes a given country to raise its tax by about 1.3 Euros. Looking across the other matrix specifications, the same picture is confirmed. Note that WTAX is always highly significant and increasing after 1993. When using Wrand we have positive and significant coefficients, but the model after 1993 is not valid since it does not pass the Hansen test on OID restrictions.

4.5.6 Tobacco

In Figure 4.13 and Figure 4.14 we observe that the majority of the countries set a zero tax rate on tobacco. Among those with a positive tax rates, IE, NL and UK have a continuous increasing trend from 1992 to 2004, instead DE and DK change their taxes almost every 5 years.

The second column in Table 4.8 shows the estimated equation after 1993, and it performs better than the equation before 1993. First note that we reject the Chow test, that confirms a structural break after 1993. The controls are joint significant and the s.e robust. The F test in the first stage passes at 10%, and the Anderson test is rejected. The Hansen test passes, the instruments are valid.

Looking at the coefficient of WTAX we notice that is positive, significant and increasing after 1993. The same picture holds using different weighting matrices. When considering Wrand we get coefficients insignificant and negative after 1993.

4.5.7 Petrol

Comparing Figure 4.15 and Figure 4.16, only few countries show a clear increasing pattern, such as BE, LU, FR and NL. Some countries have strong ups and bumps, for example

IT in 1995, PT in 2000 or IE in 1993 and 2002. These changes are not depending on the fluctuations of the exchange rates (as it is evident for ES instead) since we observe the same trend also with the tax rates in national currency. The analysis using petrol is likely not to pick a real tax competition effect, it is difficult to imagine people smuggling petrol at the border and then the control is almost impossible to detect where the individuals bought the fuel for their car. However we present our results.

Looking at Table 4.9, the second column shows the estimated equation with a contiguity matrix after 1993. The controls are joint significant and there is a structural break after 1993. All tests pass, therefore the model is identified and the instruments are valid.

The coefficient of WTAX is increasing, positive and significant after 1993. Looking across the different matrix specifications, the same results are confirmed. When using Wrand we get a positive significant coefficient before 1993, but after 1993 it becomes insignificant and the Hansen test fails.

4.6 Result - ML estimations

In addition to the Instrumental Variable estimation we show also the results using a Maximum Likelihood approach (Tables 4.10-4.13). The model that best suits the theoretical implications is the SAR, since we observe a closer similarity with the reaction functions derived in the theoretical Section 4.2. However, the results with ML are different relative to IV, we get in fact a decreasing coefficient of WTAX after 1993. We should say that the codes in MATLAB (written by Lesage and improved by Elhorst for panel data) used to implement the ML estimation, do not allow for heteroskedasticity correction and it is not possible to use more than one endogenous variable. We think that this could be a reason of our results. In addition, using Maximum Likelihood we can estimate also the SEM and SDM, which in econometric terms are formally equivalent given the common factor hypothesis. When using a SDM the coefficient ϕ of WTAX is significantly positive and increasing after 1993. We perform a Wald test in each regression to see whether ϕ is statistically significant.

In the SDM we have also the estimates of the weighted regressors, an effect that is not directly included in the theoretical specification. However, Ohasawa (1999) has proved that the equilibrium tax rates are affected not only by the neighboring tax rates, but also by the sizes and positions of countries. Kanbur and Keen (1993) consider the effect of the countries' size in terms of population densities. The prediction of the two models is in contrast, for Ohsawa the effect of the size is positive whereas it is negative for Kanbur and Keen. Egger et al. (2005) provide an empirical estimation of this effect including in their econometric specification the weighted size of the country's neighbors, and they find confirmation of Kanbur and Keen result. In our empirical specification, using a SDM we take into account not only of the weighted size of the neighboring countries (that could be measured by the population) but also of the weighted effect of the economic and political variables.

For simplicity, we report only the estimates using a contiguity weighting matrix and for SAR and SEM models we only show the coefficients of WTAX. We consider the same controls as in the IV estimation.

4.6.1 Beer

The first two columns of Table 4.10 show the estimated coefficients of equation 4.21 for beer, before and after 1993. Note first that the Chow test is rejected and there is a structural break after 1993. The controls are joint significant, and looking at the coefficient of WTAX, it is positive, significant and increasing after 1993. The interpretation remains unchanged, a 1 Euros increase in the specific tax on cigarettes in all other countries causes a given country to raise its tax by about 0.58 Euros.

Given the situation above, the results in Table 4.13 for SAR and SEM models are not clear. With SAR the coefficient of WTAX is decreasing while with SEM is negative and insignificant. However, there is always a structural break.

4.6.2 Ethyl alcohol

Looking at columns three and four in Table 4.10, using a SDM there is structural break and increasing tax competition after 1993. The coefficient of WTAX before 1993 is not significant, while after 1993 becomes significantly higher. The fit of the model is very high, R^2 above 75%.

In terms of SAR and SEM (Table 4.13) there is a structural break after 1993. Under a

SAR model the coefficient of WTAX is positive, significant but decreasing after 1993. The spatial lagged coefficient is negative with SEM instead.

4.6.3 Still wine

In Table 4.11, columns one and two show the results for still wine. Note first that there is a structural break after 1993. The coefficient of WTAX is increasing from 0.53 Euros to 0.68 Euros after 1993. The fit of the model is very good with a R^2 above 87%. In general, we find high tax competition among countries that set positive tax rates on still wine.
In Table 4.13 we report the coefficients of WTAX with SAR and SEM models: they are positive, significant, big in magnitude (above 70%), but decreasing after 1993. There is no evidence of increasing strategic interaction after 1993, but there is a structural break.

4.6.4 Sparkling wine

Looking at columns three and four in Table 4.11, we notice a strong effect of strategic interaction, the coefficient of WTAX increases more than twice, from 0.32 Euros before 1993 to 0.66 Euros after 1993. There is a structural break and the fit of the model is good. Observing Table 4.13, under a SAR model the spatial dependent variable coefficients pre and post 1993 are significantly high and closer in magnitude (above 0.75 Euros). Under a SEM the values are still high but decreasing after 1993. Although there is no increasing strategic interaction, we observe a structural break.

4.6.5 Cigarettes

Observing the first and the second column in Table 4.12 we notice first that there is a structural break after 1993. The coefficient of WTAX is positive and significant after 1993, while it is not significant before. The fit of the regression is quite good with a R^2 above 70%. Using SAR and SEM models, we observe in Table 4.13 a structural break. The spatial autocorrelation coefficients are always positive and significant, however they are not increasing after 1993.

4.6.6 Tobacco

The results with a SDM in Table 4.12 (col.s three and four) confirm the structural break, and show an increasing strategic interaction after 1993. The coefficient of WTAX increases from 0.38 Euros to 0.6 Euros. In Table 4.13 note first a structural break under both SAR and SEM models. With a SAR model the spatial dependent variable is positive and significant, although decreasing after 1993. Same picture under a SEM, but in this case the spatial autoregressive error before 1993 is closer to the coefficient of WTAX after 1993.

4.6.7 Petrol

The analysis of the excise duties on petrol with a SDM does not produce good results and we omit them. We don't get in fact significant and positive coefficients of WTAX, although we observe a structural break after 1993. Looking at Table 4.13, with SAR and SEM models we get positive and significant coefficients of the spatial dependent variable after 1993, while they are significantly negative before 1993.

4.7 Minimum Tax Rates

So far in the analysis, we have ignored any possible effects of minimum tax rates. Evers et. al.(2005), based on the theoretical literature, argue that such rates, if they affect the Nash equilibrium at all, will generally cause rates to rise. For example, in Nieslen's (2001) model, it is easily verified in the two-country case that if the minimum tax is binding on the lower-tax country, it will not only raise the tax in that country, but also in the other country, as the latter country is moved along its upward-sloping tax reaction function. So, we should expect, other things equal, the minimum tax to increase the intercepts of the reaction functions.

It is less clear how the minimum tax will affect the amount of strategic interaction. Again, in the same simple two-country model of Nielsen, a minimum tax that binds on the low-tax country (say country B) will make the country B's tax locally unresponsive to A's tax. Evers et. al.(2005) allow for such effects by interacting the minimum tax with the weighted average of other countries' taxes. So, given that minimum taxes did not come in force until 1993, one might estimate, over the period 1993-2004, an augmented version of (4.23) i.e.

$$t_{is} = f_i + \phi \sum_{j \neq i} \omega_{ij} t_{js} + \delta' \mathbf{Z}_{is} + \theta m_t + \gamma \left(m_t \times \sum_{j \neq i} \omega_{ij} t_{js} \right) + \epsilon_{is} \qquad (4.23)$$

where m_t is the minimum tax at time t. We expect $\theta > 0$ and possibly $\gamma \neq 0$. But there are some complications.

First, for wine (still and sparkling) the minimum tax rate is zero, so (4.23) cannot be estimated for these products. Second, for cigarettes, the minimum tax rate (measured as a percentage of the retail price) has been unchanged since 1993, at 57%. So, θ, γ cannot be identified from regression (4.23) just over the period 1993-2004.

For beer the minimum tax rate has been unchanged since 1993, and it is equal to 0.7448 Euros per hl/degree Plato or 1.87 Euros per hl/degree of alcohol of finished product. In Table 4.14 we report the the estimation of equation 4.23 for beer, using a contiguity weight matrix and IV method. We notice that the coefficient ϕ of WTAX is negative as expected but insignificant. The coefficient θ of the minimum tax rate is negative and insignificant as well. However, the coefficient γ of the interactive variable, obtained by multiplying the minimum tax time the spatial lagged dependent variable, is positive and significant.

4.8 Conclusion

In this Chapter we analyzed the presence of tax competition after the introduction of the Single Market in EU in 1993. First, extended the basic model of Nielsen (2001), finding the conditions of cross-border shopping and the government reaction functions in three different market regimes: duty-free market before 1993, mixed market up to 1999, single market after 1999. The theoretical prediction was no strategic interaction before the Single Market introduction and increasing tax competition from the mixed market regime to the single market regime.

Second, we specified our empirical model that tests for the presence of strategic interaction among neighboring countries in a framework of spatial econometrics. We employed a panel data set of 12 EU countries over a period of 17 years from 1987 to 2004. We estimated our equation separately on subsamples before and after 1993 and we found robust evidence that the coefficient estimates differ. This result has been confirmed across all seven taxes and a variety of different weighting matrices suggested by the theory. This means that the introduction of the Single Market has determined an important structural break in the commodity tax setting. Moreover, we found that the coefficient measuring strategic interaction is *always* significantly positive and, in most of the cases, rising after 1993. This suggests an increasing tax competition among neighbouring countries, consistently with the theoretical prediction.

4.9 Appendix: Duty-free case

In order to compute the derivative $\frac{\partial W_1}{\partial t_1}$ we split equation (4.3) in three parts:

$$\begin{aligned}&\frac{\partial}{\partial t_1}[(n_1-\widehat{d}_{DF})(v-q_1)] = \\ &= -(n_1-\widehat{d}_{DF})-(v-q_1)\frac{\partial \widehat{d}_{DF}}{\partial t_1}.\end{aligned} \tag{4.24}$$

$$\begin{aligned}&\frac{\partial}{\partial t_1}\int_0^{\widehat{d}_{DF}}(v-\alpha p_2-(1-\alpha)q_2-cx)dx = \\ &= (v-\alpha p_2)\frac{\partial \widehat{d}_{DF}}{\partial t_1}-(1-\alpha)\frac{\partial q_2}{\partial t_1}\widehat{d}_{DF}-(1-\alpha)q_2\frac{\partial \widehat{d}_{DF}}{\partial t_1}-c\widehat{d}_{DF}\frac{\partial \widehat{d}_{DF}}{\partial t_1} \\ &= -(1-\alpha)\frac{\partial q_2}{\partial t_1}\widehat{d}_{DF}+[v-\alpha p_2-(1-\alpha)q_2-c\widehat{d}_{DF}]\frac{\partial \widehat{d}_{DF}}{\partial t_1}.\end{aligned} \tag{4.25}$$

To compute the derivative of an integral when the limit of integration is a function of the variable of integration ($\widehat{d}_{DF}(t_1)$), we used the fundamental calculus theorem and the chain rule.

The derivative in the third part is:

$$\theta\frac{\partial R_1}{\partial t_1} = \theta(n_1-\widehat{d}_{DF})-\theta\frac{\partial \widehat{d}_{DF}}{\partial t_1}t_1 \tag{4.26}$$

Now summing (4.24), (4.25), (4.26) and setting equal to zero we get:

$$\begin{aligned}\frac{\partial W_1}{\partial t_1} &= \\ &- (n_1 - \widehat{d}_{DF}) - (v - q_1)\frac{\partial \widehat{d}_{DF}}{\partial t_1} \\ &- (1-\alpha)\frac{\partial q_2}{\partial t_1}\widehat{d}_{DF} + [v - \alpha p_2 - (1-\alpha)q_2 - c\widehat{d}_{DF}]\frac{\partial \widehat{d}_{DF}}{\partial t_1} \\ &+ \theta(n_1 - \widehat{d}_{DF}) - \theta\frac{\partial \widehat{d}_{DF}}{\partial t_1}t_1 = 0.\end{aligned} \tag{4.27}$$

Noting that as $v - q_1 = v - \alpha p_2 - (1-\alpha)q_2 - c\widehat{d}_{DF}$ by definition, equation (4.27) reduces to equation (4.4). Substituting in the latter, equation (4.2) and

$$\frac{\partial q_2}{\partial t_1} = \frac{\partial}{\partial t_1}[p_1 + t_1] = 1$$

$$\frac{\partial \widehat{d}_{DF}}{\partial t_1} = \frac{\alpha}{c}$$

we obtain

$$\begin{aligned}\frac{\partial W_1}{\partial t_1} &= (\theta - 1)(n_1 - \widehat{d}_{DF}) - (1-\alpha)\widehat{d}_{DF} - \theta\frac{\alpha}{c}t_1 = 0 \\ &= c(\theta - 1) - c[(\theta - 1) + (1-\alpha)]\widehat{d}_{DF} - \theta\alpha t_1 = 0.\end{aligned} \tag{4.28}$$

then

$$c(\theta - 1)n_1 - (\theta - \alpha)(p_1 - p_2) = \alpha[(\theta - \alpha) + \theta]t_1.$$

Solving for t_1 we get equation (4.5).

4.9.1 Single Market case

The maximization procedure is similar to that of the duty free regime, we solve separately the derivatives in equation (4.13) and we add them all. The first and the third part of the derivative are similar to equation (4.24) and equation (4.26), we need just to replace $\widehat{d}_{DF}$ with $\widehat{d}_{SM}$. The derivative of the integral is a bit different because now P_{2g} is not function of t_2, and the cross-border utility has changed.

$$\begin{aligned}&\frac{\partial}{\partial t_1}\int_0^{\widehat{d}_{SM}}(v-q_2-cx)dx = \\ &= -\frac{\partial q_2}{\partial t_1}\widehat{d}_{SM} + [v-p_2-c\widehat{d}_{SM}]\frac{\partial \widehat{d}_{SM}}{\partial t_1}.\end{aligned} \tag{4.29}$$

Summing the three derivatives and using $v - q_1 = v - q_2 - c\widehat{d}_{DF}$, we obtain equation (4.9). Substituting in the latter $\frac{\partial q_2}{\partial t_1} = 0$, $\frac{\partial \widehat{d}_{SM}}{\partial t_1} = \frac{1}{c}$ and equation (4.7), we get

$$\begin{aligned}\frac{\partial W_1}{\partial t_1} = -(n_1 - \widehat{d}_{SM}) + \theta(n_1 - \widehat{d}_{SM}) - \frac{\theta}{c}t_1 = 0 \\ c(\theta-1)n_1 - c(\theta-1)\widehat{d}_{SM} - \theta t_1 = 0 \\ c(\theta-1)n_1 - (\theta-1)(p_1-p_2) - (\theta-1)(t_1-t_2) - \theta t_1 = 0 \\ c(\theta-1)n_1 - (\theta-1)(p_1-p_2) + (\theta-1)t_2 = (2\theta-1)t_1.\end{aligned} \tag{4.30}$$

solving for t_1 we obtain the equation (4.10).

4.10 Appendix: Notes on the Excise Duty Tables

The Excise Duty Tables that we use are issued by the European Commission - Directorate general Taxation and Customs Union Tax Policy. We are interested in the position of each Member State on yearly basis, therefore we consider (when available) the tables issued on Jan 1st, which represents the situation at end of the previous year. Up to 2004 the EU published the tables quarterly, they now only publish the tables twice a year (January and June). The reason they publish the tables at intervals of more than once a year is that Member States have their annual Budgets at different times during the year. Therefore, by publishing them periodically they can ensure that the information is reasonably up to date.

In detail, and following the advice of Peter Couzins (European Commission Taxation and Customs Union - Unit D2), we used the following tables:

- For the year 1987: the tables issued on 1.4.1987.
- For the year 1988: no table was produced.

- For the year 1989: the tables issued on 1.1.1990.
- For the year 1990: the tables issued on 1.1.1991.
- For the year 1991: the tables issued on 1.1.1992.
- For the year 1992: the tables issued on 1.1.1993.
- For the year 1993: the tables issued on 1.1.1994.
- For the year 1994: the tables issued on 29.03.1995 (no other tables produced to describe the situation in 1994).
- For the year 1995: the tables issued on 1.1.1996.
- For the year 1996: the tables issued on 1.1.1997
- For the year 1997: the tables issued on 23.06.1998 (no other tables produced to describe the situation in 1998).
- For the year 1998: the tables issued on 12.1998.
- For the year 1999: the tables issued on 11.1999.
- For the year 2000: the tables issued on 11.2000.
- For the year 2001: the tables issued on 11.2001.
- For the year 2002: the tables issued on 08.2002 (a later edition for that year was not published).
- For the year 2003: the tables issued on 12.2003.
- For the year 2004: the tables issued on 01.01.2005.

One concern of ours is that the use of 1998 tables to get tax rates for both 1997 and 1998 gives exactly the same taxes in both years for DK,FR,IT,UK. In the remaining countries, changes from 1997 to 1998 are very minor.

4.10.1 Modifications to Excise Duty Table Data

The panel is formed by the 12 countries that belong to the EU in 1987. The period analyzed is from 1987 to 2004. The data are only available since the countries joined the EU. From 1987 to 1999 the currency values are reported in ECU, from 1999 the figures are in Euro. The ECU was replaced by the Euro on 1 January 1999 with a conversion rate of 1:1.

Beer: excise in ECU/Euro per hl/degree Plato.

The tables reports the values of the excise both in degree Plato per hl of beer and degree alcohol per hl of beer. The choice of the unit of measure depends on the country. However, according to Directive 92/84/EEC it has been accepted that 2.5 degrees Plato is broadly equivalent to one degree alcohol by volume (or 0.748 Euros Plato is equal to 1.87 Euros alcohol by volume). For the reasons mentioned above, it is not pretended that the two rates relate with absolute exactitude. On the contrary, the most that can be said is that no great difference will normally arise from applying either charge to a given beer. We report the values only per hl per degree Plato, using the accepted conversion explained above.

For the year before the EU Directive, the general procedure is:

in 1987 we consider the figures per hl of 12.5 degree Plato, and we divide by 12.5 to get hl/degPlato. In 1989-1991 the values are per hl of 12 degree Plato, and we divide by 12 to get hl/degPlato.

With some exceptions:

EL 1989-1991: the value are per hl in 12.5 degree Plato, so we divide by 12.5.

IE 1989-1991: the values are per hl in 12.5 degree Plato, so we divide by 12.5.

IE 1995-6 the values are both per degree Plato and per degree alcohol, we consider only degree Plato.

IE 1997-2004: per degree alcohol only, we convert in degree Plato.

For the countries that present several values of excise tax depending on beer strength we take the tax on beer of a standard strength of 5% alcohol which is equal to 12.5° Plato, and we divide the values reported (that is in hl of beer) by 12.5 to get hl/degree Plato.

Specifically:

NL 1992-2004: there are several taxes, we consider 11-15 degree Plato and we divide by

12.5.

PT 1992-2004: there are several taxes, we consider 11-13 degree Plato and we divide by 12.5.

UK 1992: the value in degree Plato is strange, therefore we consider the value in degree alcohol of 1.6.1993 and convert it to degree Plato. From 1993 to 2004, all the values are per degree alcohol and we convert them in per degree Plato.

DK 1993-2003: there are several ranges, we consider 11-14 degree Plato and we divide by 12.5. DK 2004: only 1 figure reported and per degree alcohol, we convert in per degree Plato.

FR 1993-2004: values per degree alcohol converted in per degree Plato. FR 1996-2004: only taxes on beer of strength > 2.8% of alcohol are reported.

IT : the rates don't change from 1992 to 2001 but in Euro we notice many variations due to large variations of the exchange rate ECU-LIT in that period.

The rates for 1997 are in the Table 1.6.1998, and it is the same Table used for 1998 that is contained in Table 20.10.1998.

Ethyl alcohol: excise in ECU/Euro per hl of ethyl alcohol.

As from 1 January 1993, the minimum rate of excise duty on alcohol and alcohol contained in beverages is fixed at ECU 550 per hectoliter of pure alcohol. However, Member States which apply to alcohol and alcoholic beverages a rate of duty not exceeding ECU 1000 per hectoliter of pure alcohol may not reduce their national rate. In addition Member States which apply to the said products a rate of duty exceeding ECU 1000 per hectoliter of pure alcohol may not reduce their national rate below ECU 1000.

EL 1987 : the value is low in 1987 with respect to the values in the subsequent years, but in the table there is a special note for EL that says: " several very high ad valorem taxes are levied besides the excise duty especially on imported products". The higher ad valorem taxes are not reported in the Tables.

EL 1992-1993: 2 rates reported, we consider the higher which applies to non-ouzo alcoholic beverages.

EL 1994: there are 2 rates reported, the first for ouzo (same as in the previous years) and second for non-ouzo, but it seems low. We use the same rate for non ouzo as in the previous years.

EL 1993-1995: for the departments of Dodecanese the rate is reduced by 50% against the rate applicable in the rest of Greece: we use the figure for the rest of Greece.

EL 1996-2004: the lower rates for the departments of Dodecanese are reported in the reduced rates table.

EL 1995-2004: only 1 rate reported and for non ouzo.

IT 1987-1991: 2 rates, the lower applies to alcohol produced form distillation of wine, potatoes, fruit ,etc; le higher rate applies only to synthetic alcohol or alcohol derived from sugar. We consider an average between the 2 rates.

IT 1992-1996: may maintain its existing system of taxation of alcohol until 30 June 1996 provided that the rate is not lower than the rates of Directive 92/84/EEC. We consider an average between the 2 rates.

For DK, we have a problem because the overall tax is a percentage of the average wholesale price excluding all taxes, plus a specific excise, and we only know the specific excise.

In more detail:

1987: a rate of 60% on the average wholesale price excluding all taxes, plus 22.880 DKR/hl.

DK 1989-1991 and 1994-5: the excises reported in the Tables exclude 37.5% of the wholesale price (excluding VAT).

DK 1992-1993: the excises include 37.5% of the wholesale price (excluding VAT), therefore we consider only the fixed excise (14300 DKR), unchanged with respect the previous 3 years, and we convert it in ECU (exchange rate at 1.10.1992 and at 1.10.1993).

DK 1996: we input the fixed excise of 14300 DKR, converted in ECU at 1.1.1997. This because the rate showed in the table is lower than the fixed excise considered in the previous years, that is 14300 DKR. Moreover, a special note in the Table 29.03.1995 says: "DK may maintain its existing system of taxation of alcohol until 30 June 1996 provided that the rate is not lower than the minimum rate as set out in the Directive 92/84/EEC". We assume therefore that in 1996 the taxation system is still based on a fixed excise of 14300 DKR.

DK 1997-2002: from July 1st 1997 the fixed excise is increased to 27500 DKR, the value in ECU is its conversion each year.

DK 2003-2004: the fixed excise is reduced to 15000 DKR, the value in ECU is its conversion each year.

FR 1992: there are 2 rates in that year, we consider the more recent.

IE 2003: : big increase in the excise with respect to the previous year, no explanation given.

The rates for 1997 are in the Table 1.6.1998, and it is the same Table used for 1998 that is contained in Table 20.10.1998.

Still wine : excise in ECU/Euro on still wine.

In 1987,1989,1990, 1991 the excises for wine not > 12% of alcohol from 1992 % of alcohol are not specified, but the minimum excise is 0 ECU per hl.

DK 1987-1991 the excise includes wine >= 8.5% and < 15% of alcohol.

DK 1996 several excises, we consider 6% to 15% of alcohol.

DK 1999-2004 two excises are reported and we consider 6% to 15% of alcohol.

LU 1992-2004: the excise is 0% for wine < 13% and >= 13% of alcohol.

IE 1995-6 two excises are reported and we consider 5.5% to 15% of alcohol.

Sparkling wine : excise in ECU/Euro on sparkling wine.

In 1987 the excise for wine is > 6% of alcohol.

ES and PT 1987 assumed 0% : for ES and PT the value for wine > 6% of alcohol is not reported, but for all the other types of wine with different alcoholic degrees the excise is zero.

In 1989,1990, 1991 the excises for wine are not > 12%.

DK 1987-1991 the excise includes wine >= 8.5% and < 15% of alcohol.

EL 1989-1991: 12% of the producer's selling price, value not reported. FR 1989-1991: the value in ECU in the Tables is the sum of two excises. We consider only the excise on wine not > 12%, that is 54FF, and we convert for the years 1989-1991 this value in ECU (using the rate on Jan 1st of each year).

SE 1994-1996: the value considered is 8.5% to 15% of alcohol.

DK 1996 several excises, we consider 6% to 15% of alcohol.

UK 1996 several excises, we consider 8.5% to 15% of alcohol.

DK 1999-2004 two excises are reported and we consider 6% to 15% of alcohol.

For both type of wine, the rates for 1997 are in the Table 1.6.1998, and it is the same Table used for the rates in 1998 (contained in Table 12.1998). This is a problem that affects all the tax rates for wine, beer and ethyl alcohol that consequently are equal in 1997 and 1998.

Cigarettes: specific excise (per 1000 cigarettes) in ECU/Euro.

Tobacco: average between specific excise for "fine cut smoking tobacco" and "other smoking tobacco", in ECU/Euro.

DE, NL, are the only countries with both ad valorem and specific excise. For the other countries, if they have the ad valorem excise the specific is zero, and vice versa.

DE 1987: the value is not available.

DK 1987 : the only value for "fine cut tobacco" is in local currency DKR and we convert it using the exchange rate of ECU on 1.4.1987: 7.93076 DKR per 1 ECU.

IE 1997: the only value for "fine cut tobacco" is in local currency IRL and we convert it using the exchange rate of ECU on 1.4.1987: 0.776488 IRL per 1 ECU.

UK 1987: the values are only in local currency UKL and we convert them using the exchange rate of ECU on 1.4.1987: 0.715957 UKL per 1 ECU.

NL 1987: the value is not available.

SE 1994, 1995, 1996: there is one value for fine cut tobacco, and for the same year two different values for other smoking tobacco. We consider an average among the value for fine cut tobacco and the two values for other smoking tobacco. From 1997 there is the same specific excise for "fine cut" and "smoking tobacco".

In the years 1989, 1990, 1991 although the tables are only for smoking tobacco, for the countries that have a positive specific excise it is possible to distinguish between "fine cut" and "other smoking tobacco"; except for NL, whose value is for "other smoking tobacco".

Petrol : excise in ECU/Euro on unleaded petrol.

For 1987 the values are available only for: DK, DE, EL, NL, and UK.

BE 2003, 2004: we consider the average between the excises < 1998 and >= 1998.

EL 1998-2004: average between excise <= 1996 and > 1996. ES 1996-2004: average between excise <= 1997 and > 1997. FR 2000-2004: average between excise < 1995 and

>= 1997. IE 1998-2004: average between ordinary unleaded and high-octane. NL 2003: average between normal and low-sulphur. UK 1997-2003: average between ordinary unleaded and ULSP (ultra low sulphur). UK 2004: average unleaded, ULSP and sulphur free.

Fig. 4.2: Country size

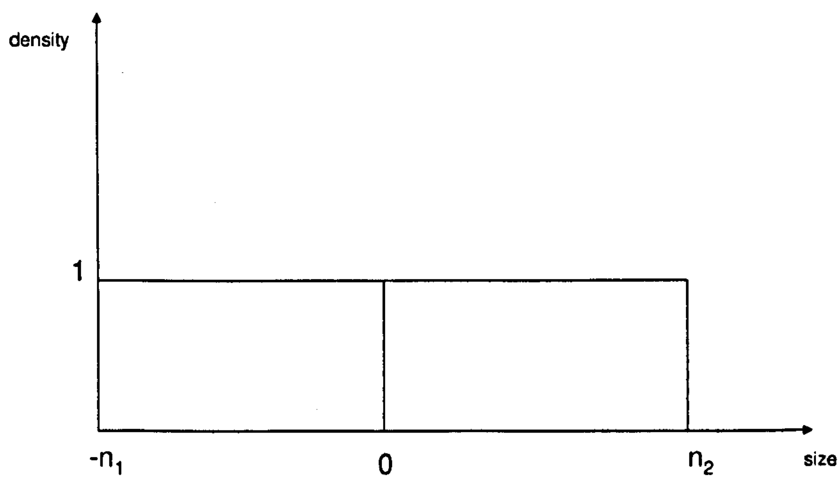


Fig. 4.3: Beer - Specific Excise in Euro

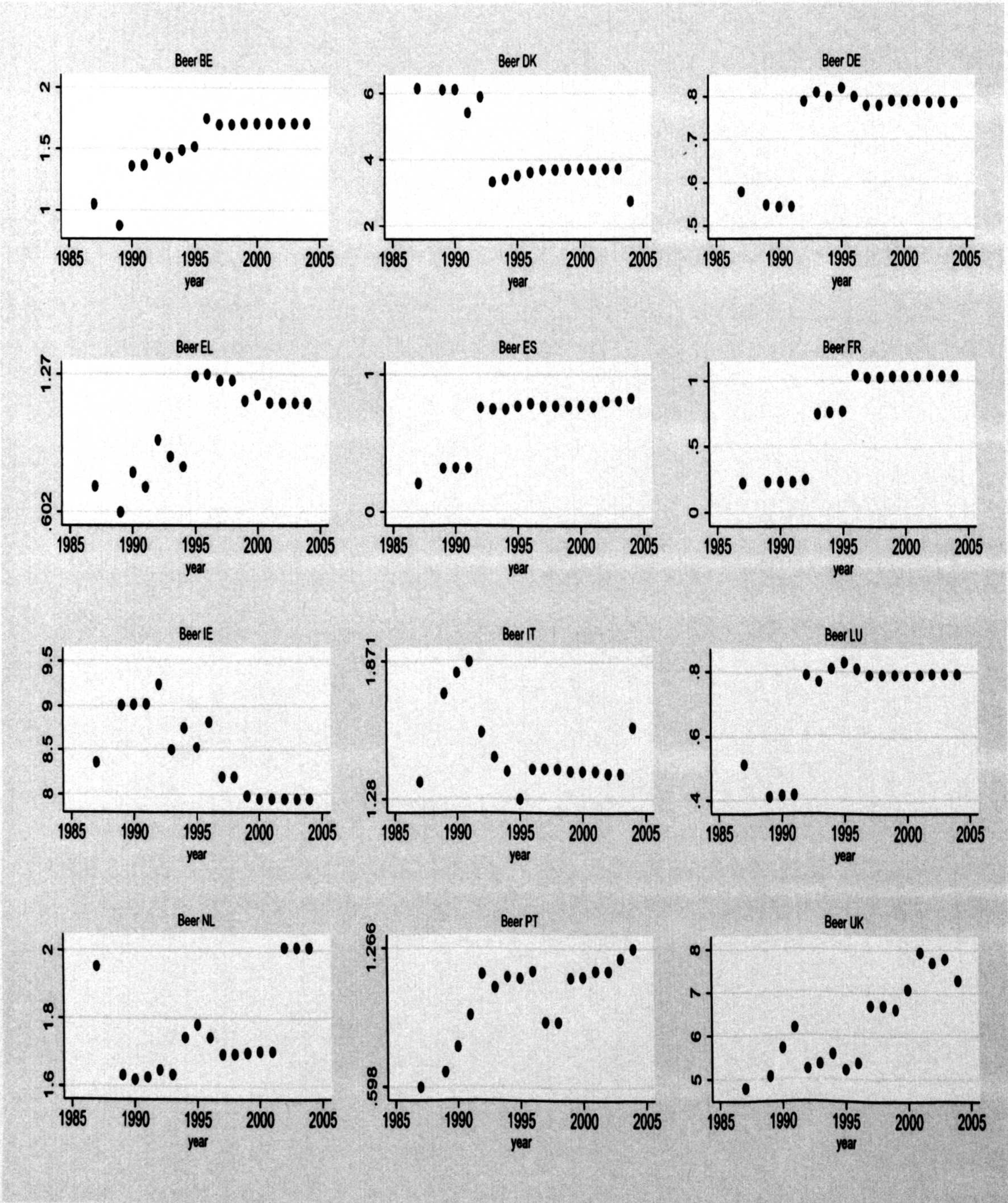


Fig. 4.4: Beer - Specific Excise in National Currency

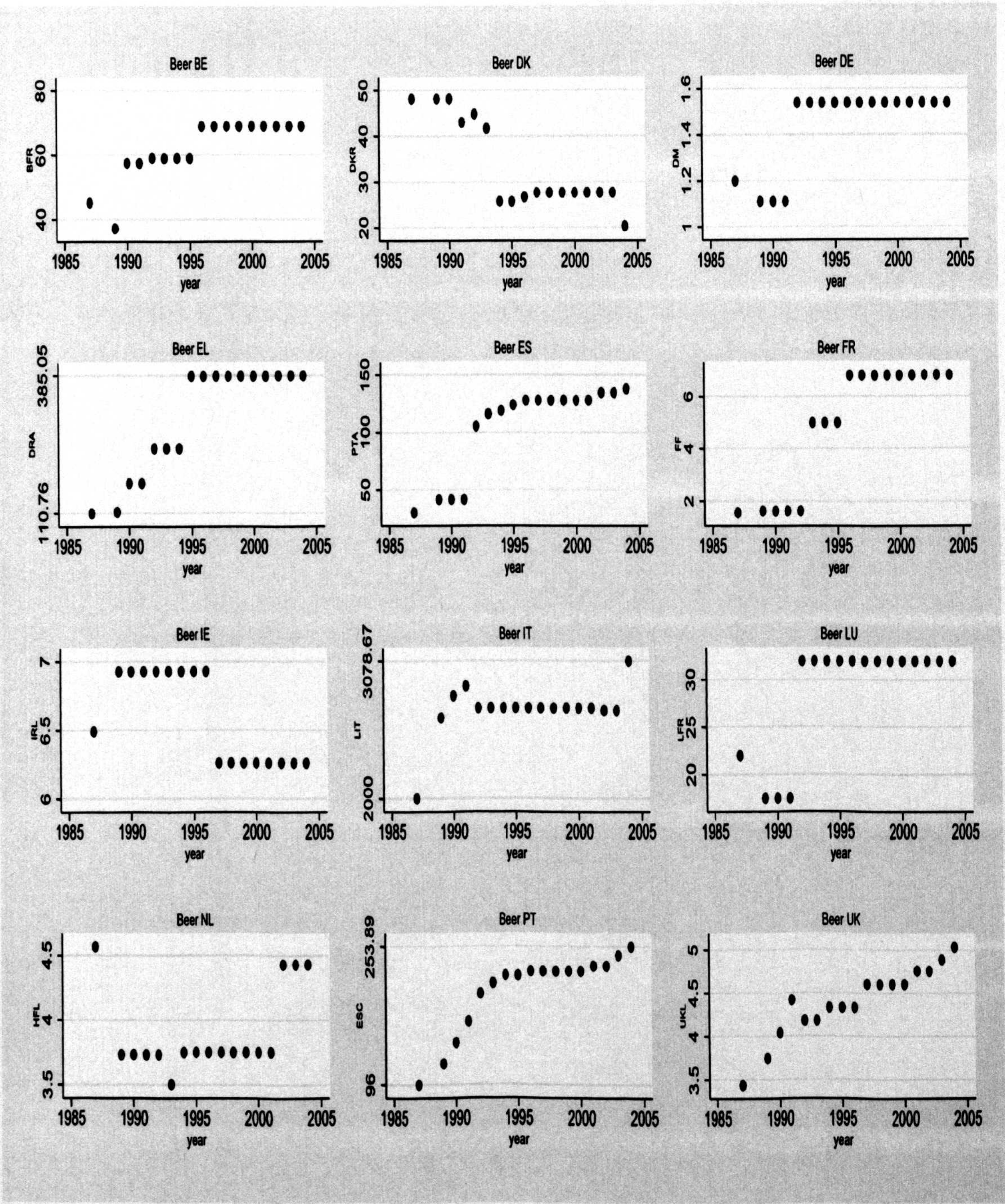


Fig. 4.5: Ethyl Alcohol - Specific Excise in Euro

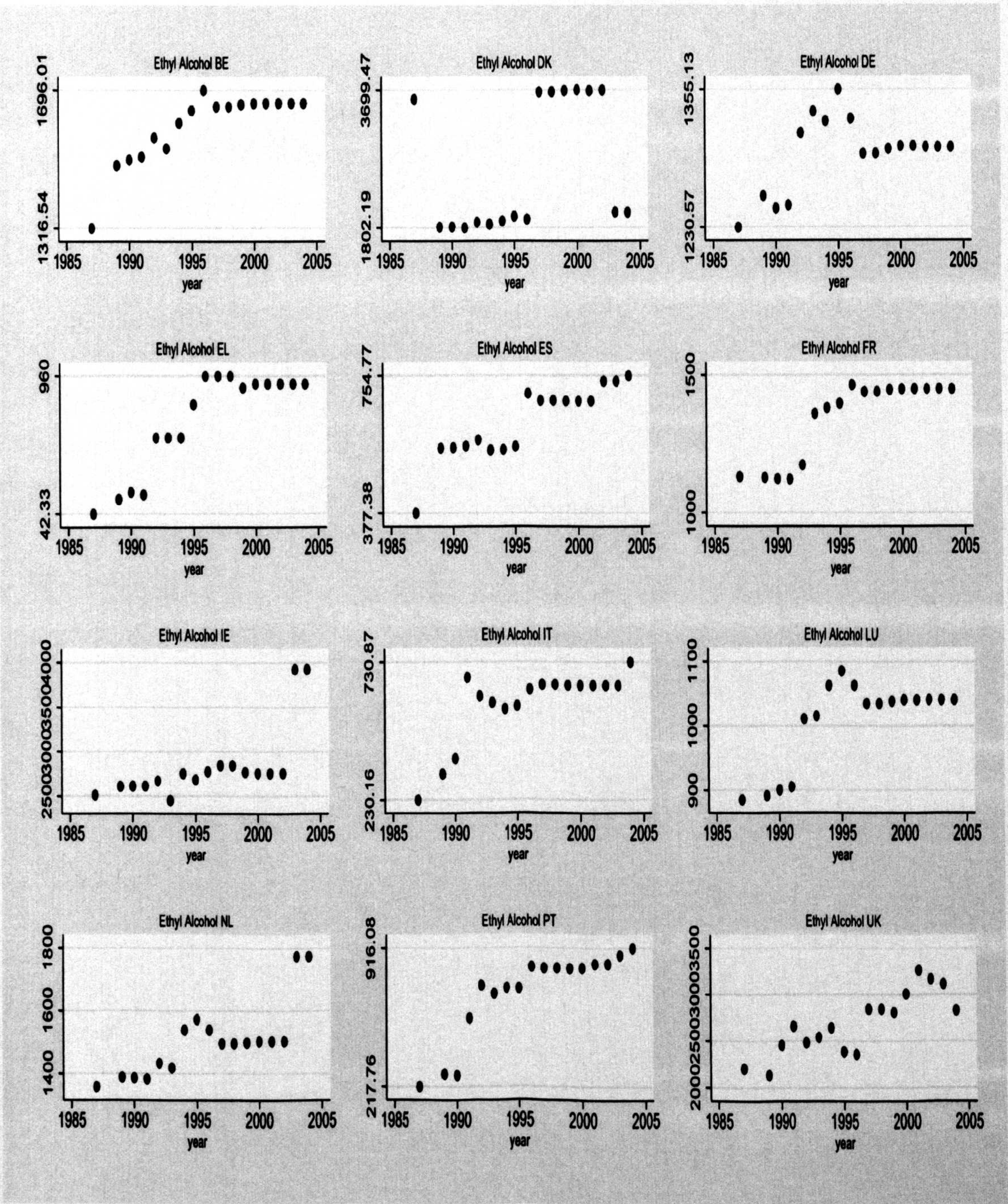


Fig. 4.6: Ethyl alcohol - Specific Excise in National Currency

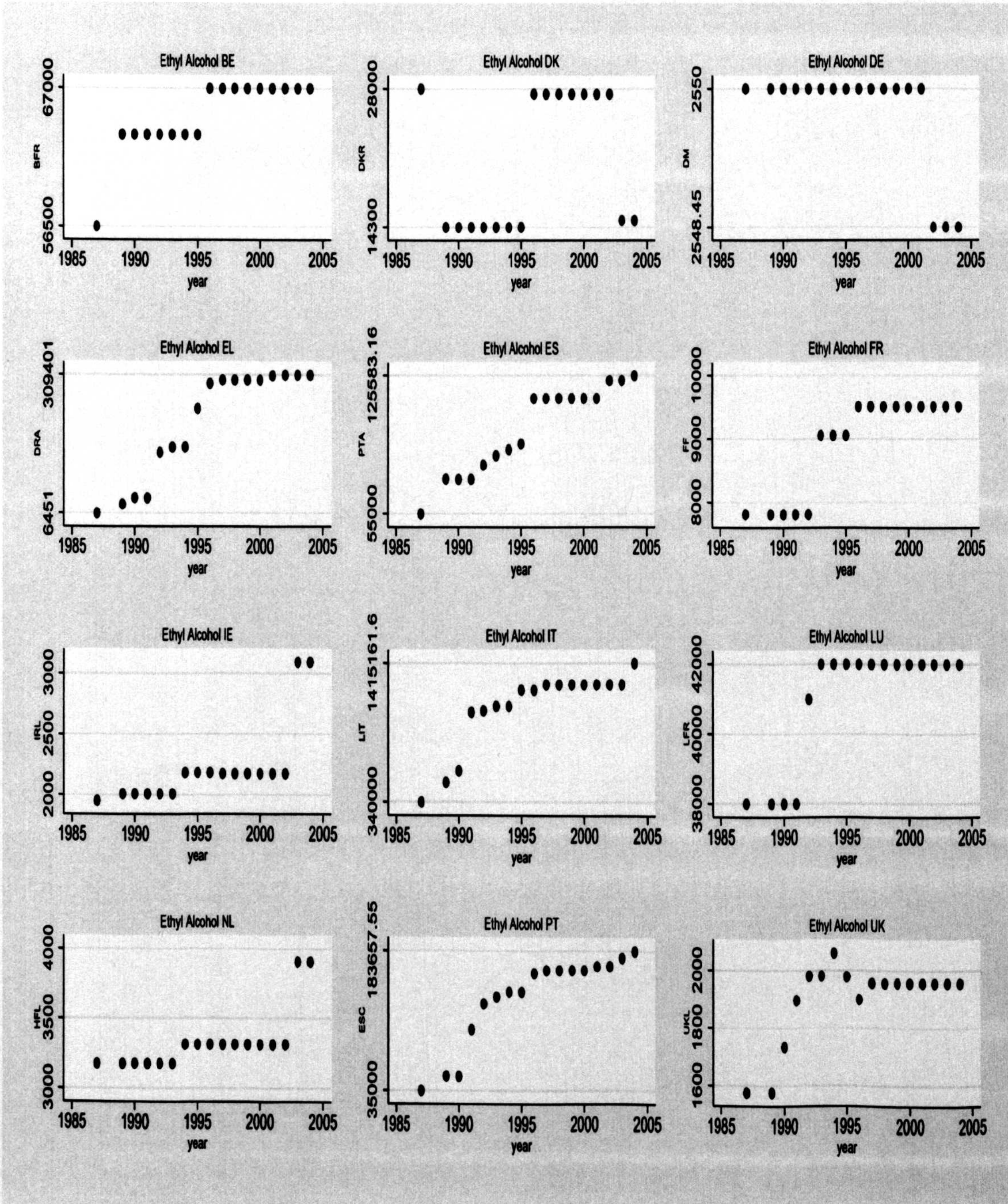


Fig. 4.7: Still wine - Specific Excise in Euro

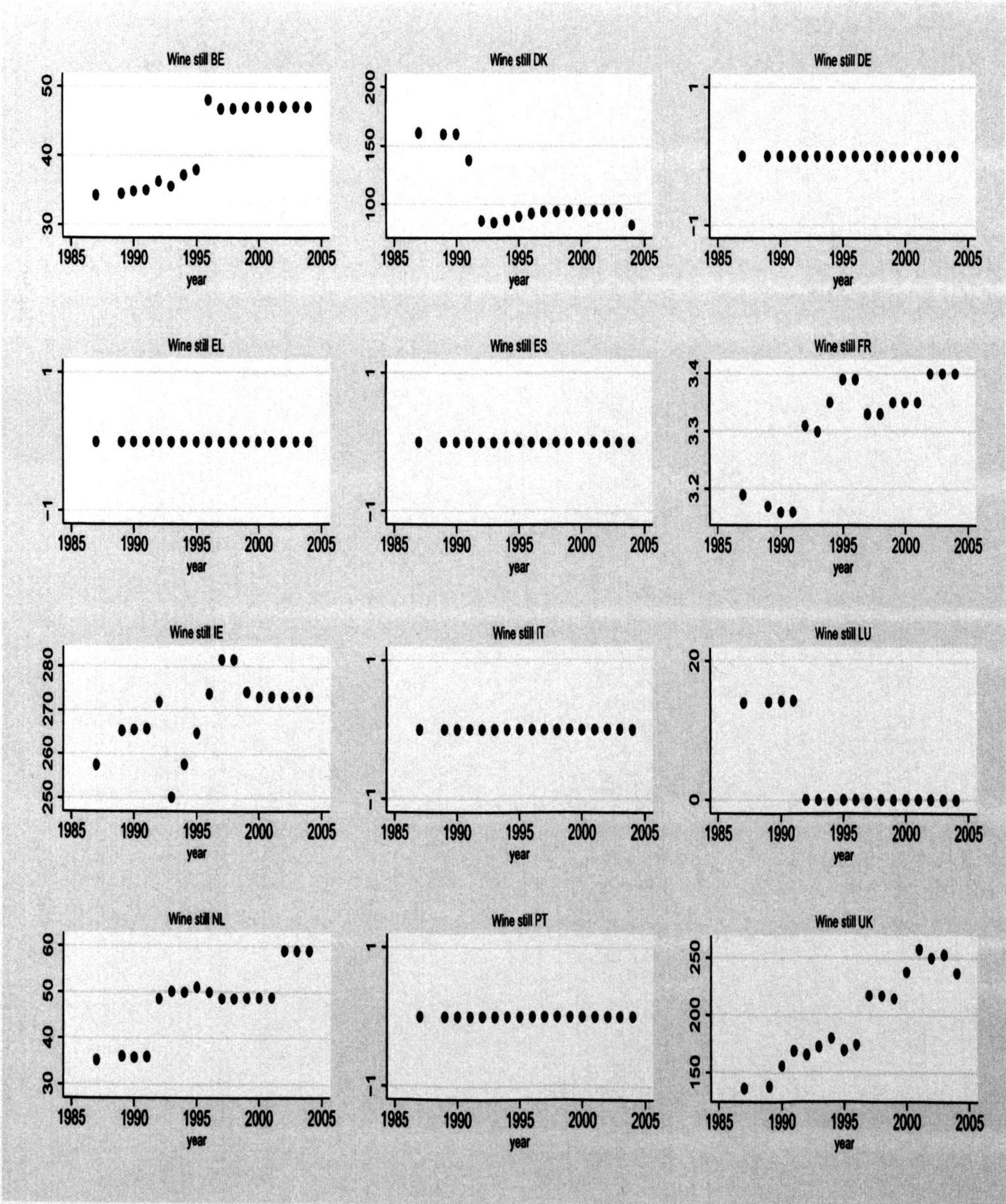


Fig. 4.8: Still wine - Specific Excise in National Currency

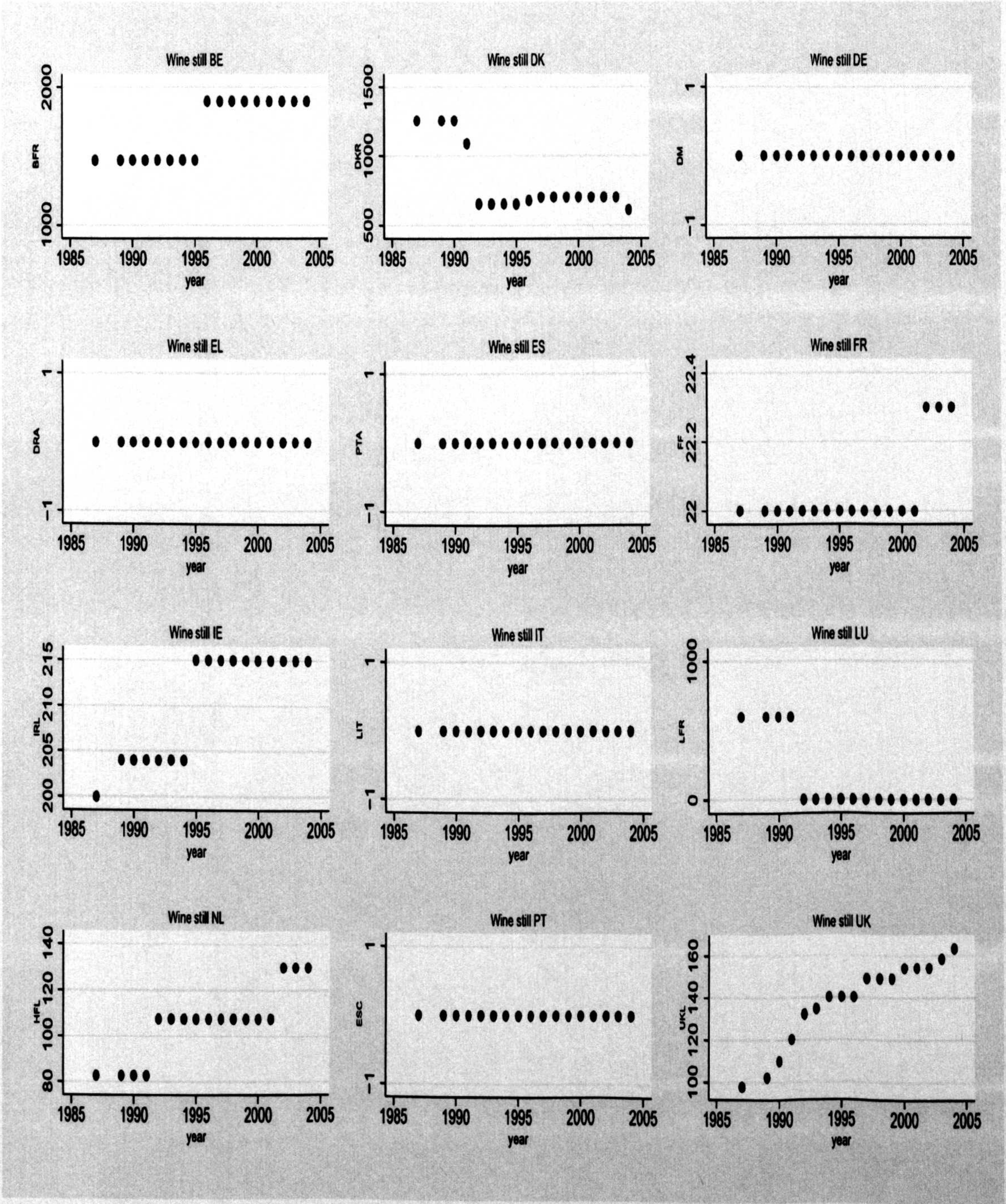


Fig. 4.9: Sparkling wine - Specific Excise in Euro

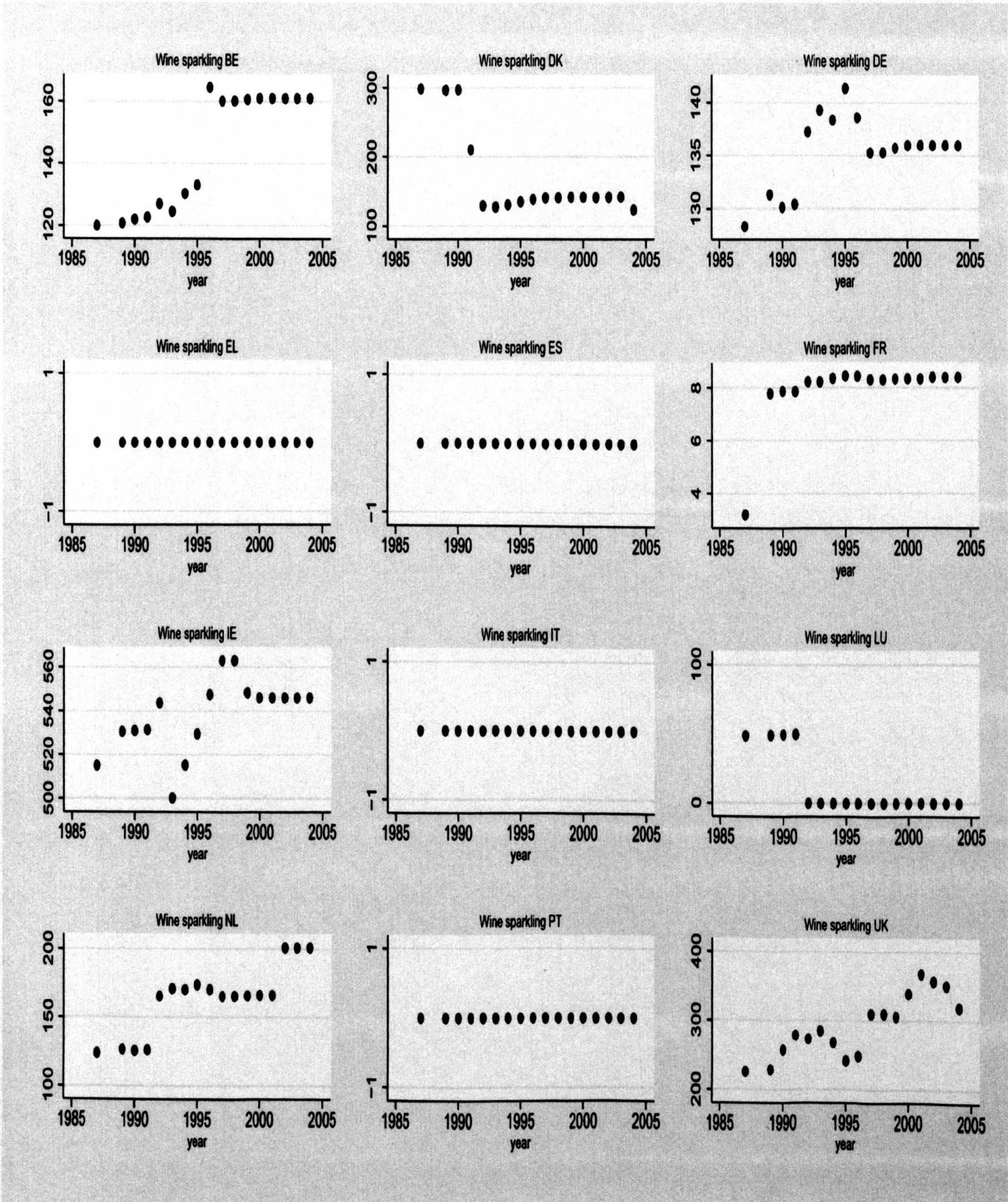


Fig. 4.10: Sparkling wine - Specific Excise in National Currency

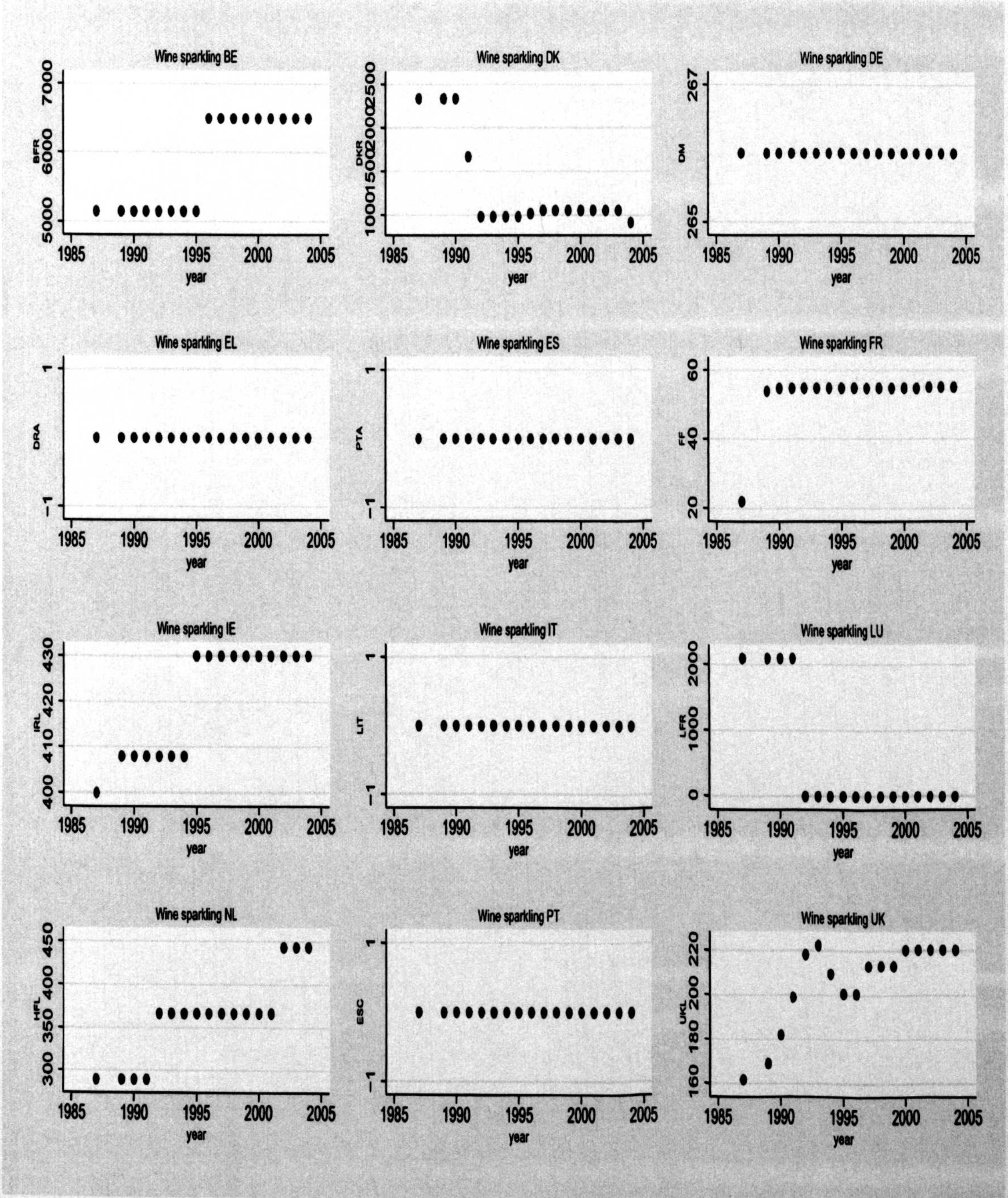


Fig. 4.11: Cigarettes - Specific Excise in Euro

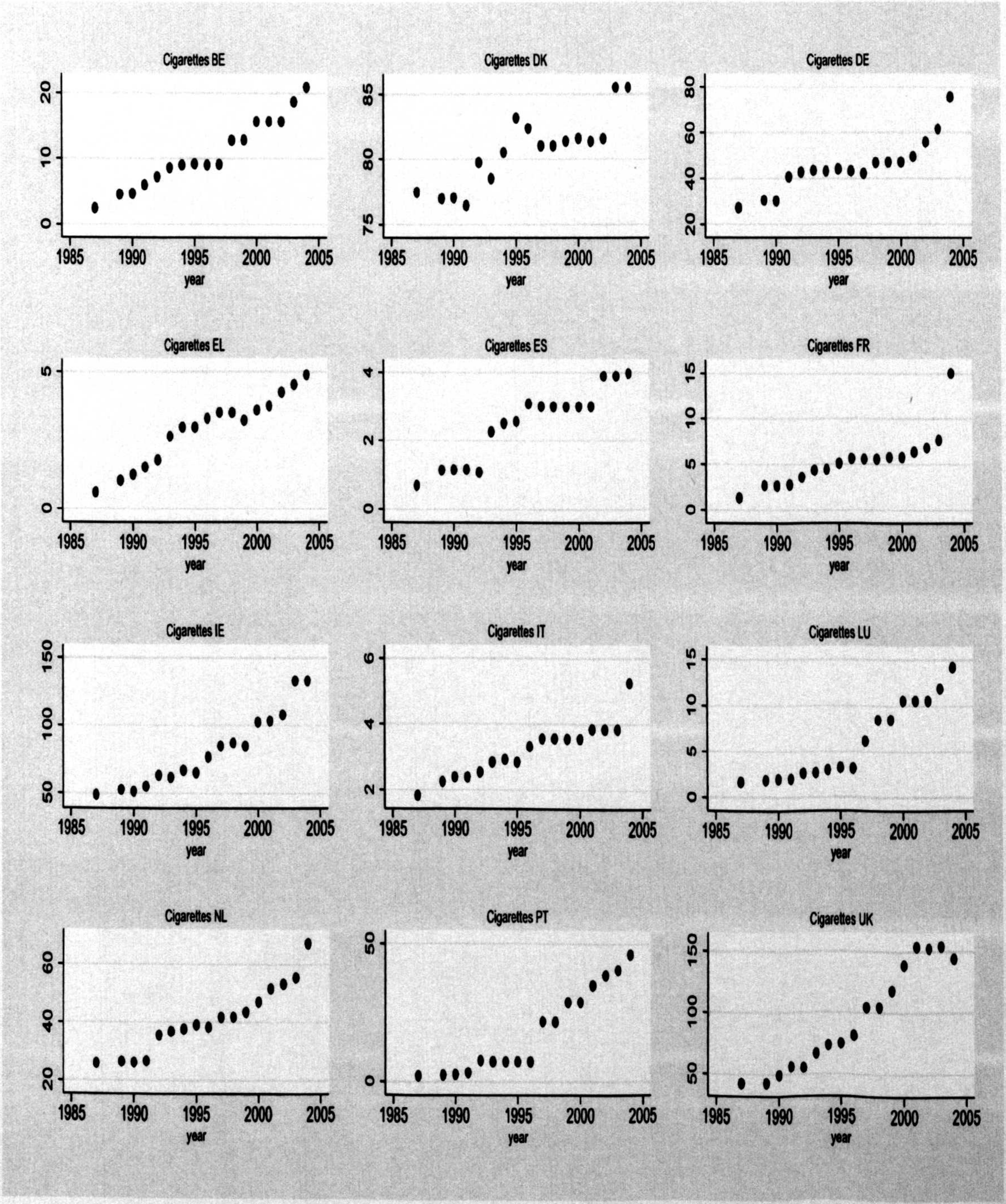


Fig. 4.12: Cigarettes - Specific Excise in National Currency

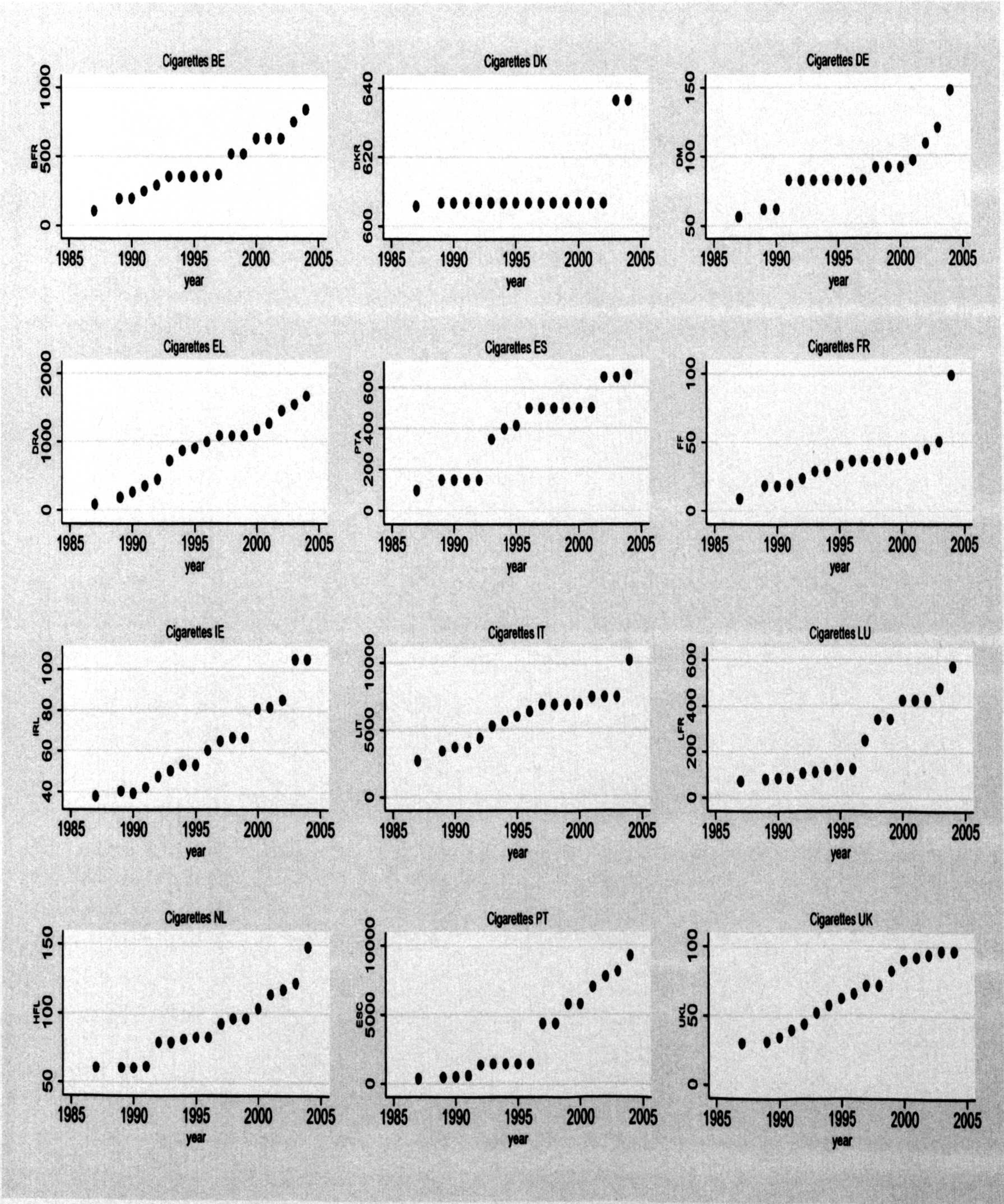


Fig. 4.13: Tobacco - Specific Excise in Euro

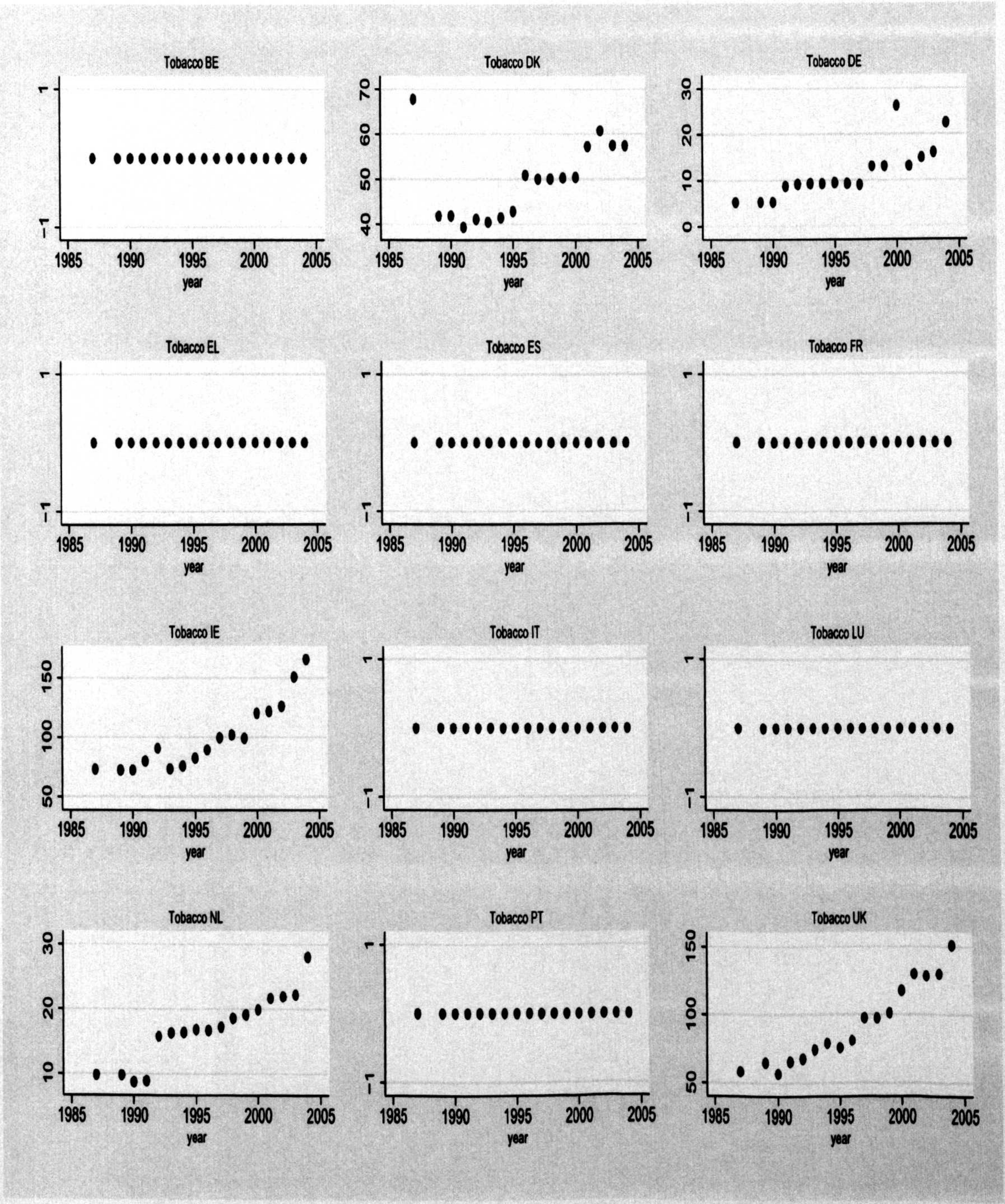


Fig. 4.14: Tobacco - Specific Excise in National Currency

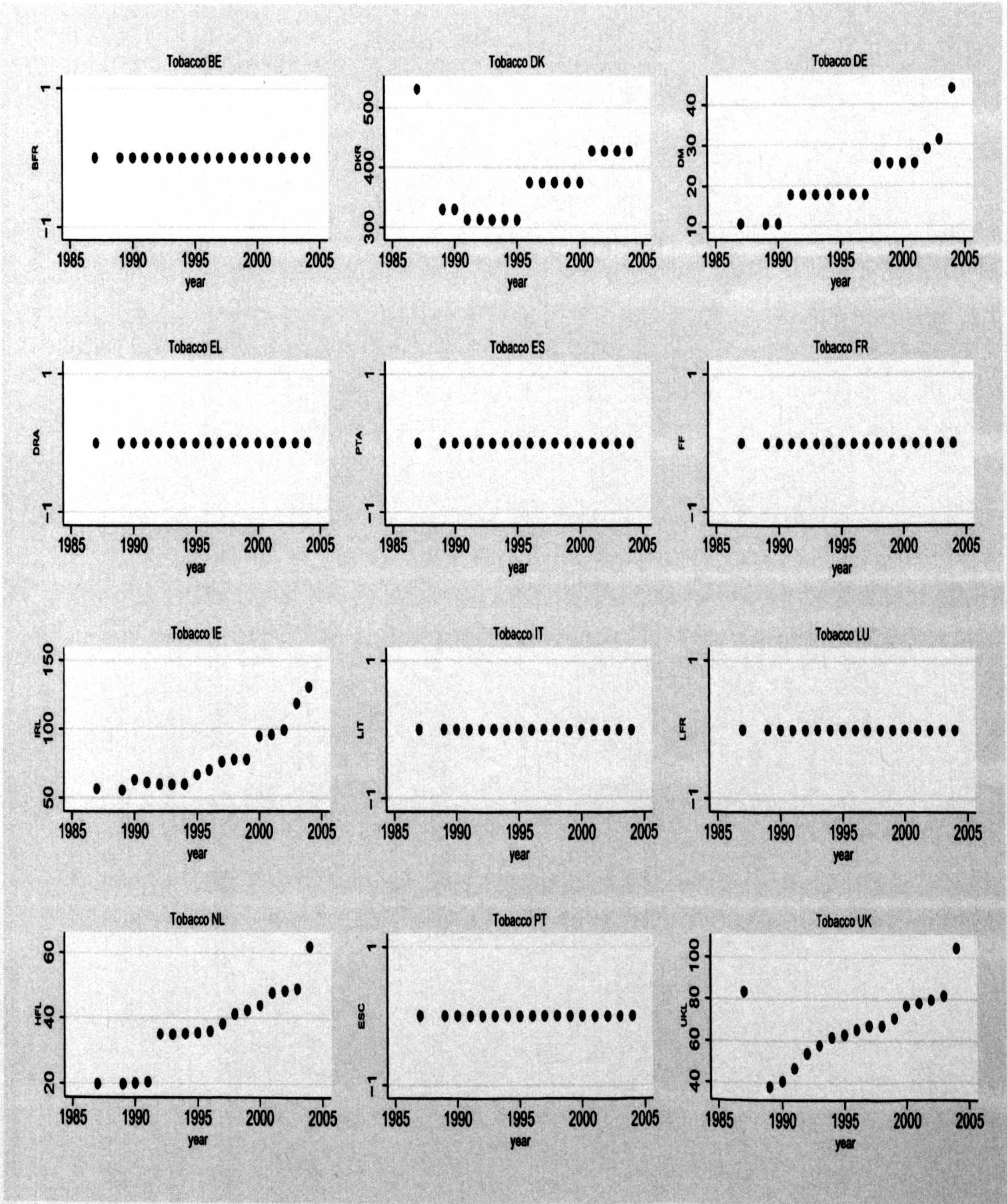


Fig. 4.15: Petrol - Specific Excise in Euro

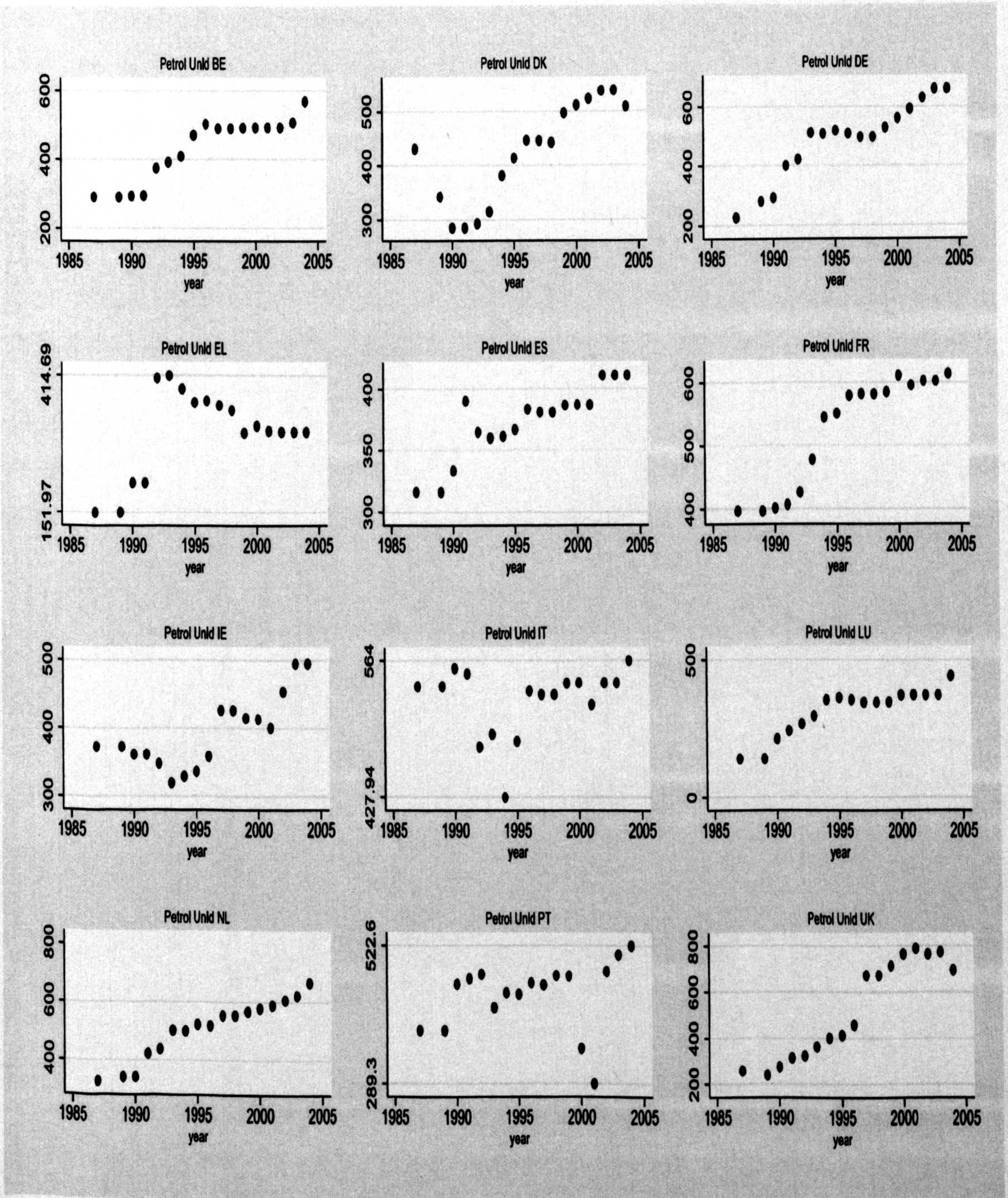
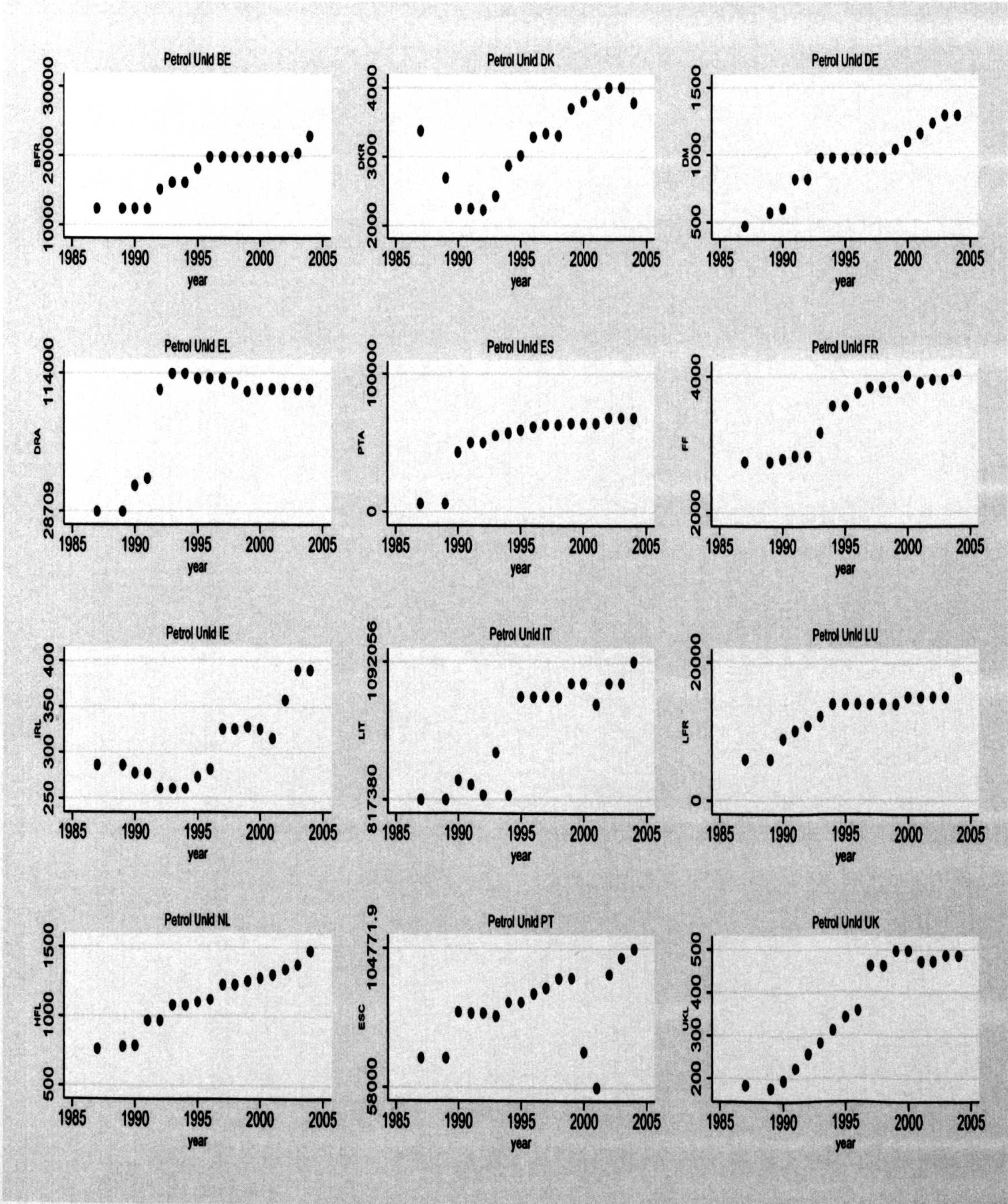


Fig. 4.16: Petrol - Specific Excise in National Currency



Tab. 4.3: Pooled model with spatially lagged dependent variable and time period fixed effects

Dependent Variable: Beer specific tax rates (2000 prices) Spatial Autoregressive Model - IV

Coefficient Estimates

Explanatory Variables

	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	0.265	0.309*	0.995	1.301***	1.610***	1.685***	1.415***	1.465***	0.743**	-0.062
	(0.384)	(0.175)	(0.889)	(0.307)	(0.554)	(0.467)	(0.524)	(0.391)	(0.302)	(0.119)
total population	0.015	0.020**	-0.001	0.009	-0.014	-0.009	-0.005	-0.003	0.002	0.028**
	(0.010)	(0.009)	(0.024)	(0.007)	(0.016)	(0.008)	(0.015)	(0.007)	(0.010)	(0.012)
gdppc[a]	0.003	0.000	-0.004	-0.002*	-0.006	-0.001*	-0.006	-0.001	-0.007	0.001
	(0.004)	(0.001)	(0.009)	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.005)	(0.001)
govcons[b]	0.080*	0.036	0.032	0.056	-0.042	0.142**	-0.047	0.110**	-0.038	-0.008
	(0.042)	(0.037)	(0.077)	(0.037)	(0.077)	(0.055)	(0.087)	(0.047)	(0.079)	(0.038)
govright	0.039	0.046	0.023	0.027	0.040	0.085	0.053	0.058	0.015	-0.024
	(0.051)	(0.095)	(0.058)	(0.085)	(0.137)	(0.107)	(0.110)	(0.101)	(0.070)	(0.079)
govleft	0.072	0.016	0.040	-0.038	0.081	0.105	0.082	0.045	0.116	-0.035
	(0.117)	(0.144)	(0.136)	(0.144)	(0.174)	(0.162)	(0.155)	(0.156)	(0.128)	(0.138)
Intercept	-4.228	1.861	9.683	4.330***	0.465	4.168	15.465	-1.004	-1.678	-1.315
	(6.421)	(2.025)	(17.943)	(0.792)	(2.052)	(5.771)	(9.293)	(1.963)	(1.821)	(1.007)
N	60	144	60	144	60	144	60	144	60	144
R^2	0.987	0.985	0.985	0.973	0.983	0.964	0.984	0.967	0.981	0.987
Pagan-H	17.735	66.717	25.646	74.369	28.265	2.123	27.124	73.006	27.947	67.039
	(0.604)	(0.000)	(0.140)	(0.000)	(0.132)	(0.000)	(0.166)	(0.000)	0.1417	(0.000)
F	1.35	3.94	9.18	49.15	11.30	11.74	13.41	13.88	11.60	11.52
	(0.270)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anderson	5.508	27.154	9.065	112.914	45.378	69.692	45.571	69.848	49.582	83.204
	(0.239)	(0.000)	(0.028)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hansen	4.779	7.768	0.640	0.374	5.574	7.193	2.683	9.320	6.618	4.794
	(0.188)	(0.051)	(0.726)	(0.829)	(0.233)	(0.126)	(0.612)	(0.053)	(0.157)	(0.3091)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% robust std. err. in brackets, p-values for tests.

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.4: Pooled model with spatially lagged dependent variable and time period fixed effects

Dependent Variable: Ethyl Alcohol specific tax rates (2000 prices) Spatial Autoregressive Model - IV

Coefficient Estimates

Explanatory Variables	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	0.357	0.678*	1.330***	1.824**	1.496***	1.149***	1.260	1.171***	468.042***	25.795
	(0.313)	(0.366)	(0.148)	(0.844)	(0.470)	(0.270)	(0.752)	(0.281)	(170.743)	(72.316)
total population	20.349**	1.421	-2.519	12.191**	-2.550	-9.311**	1.678	-8.593*	12.153	8.560*
	(9.269)	(6.952)	(3.794)	(5.819)	(5.814)	(4.035)	(10.735)	(4.705)	(8.208)	(4.538)
gdppc[a]	-3.655	2.304***	-2.597	-3.207*	-2.981	1.024	-2.742	0.799	-4.116	2.190*
	(5.341)	(0.568)	(1.755)	(1.800)	(4.754)	(1.222)	(4.828)	(1.201)	(5.627)	(1.304)
govcons[b]	48.082	10.101	-0.788	103.332**	-8.804	21.846	-1.872	12.921	-10.665	29.982
	(30.938)	(37.158)	(17.968)	(43.173)	(31.958)	(21.983)	(44.541)	(24.591)	(36.827)	(26.137)
govright	22.697	-20.973	90.682***	70.979	86.778	-78.569	83.953	-77.961	21.921	-83.333
	(49.042)	(82.089)	(34.011)	(86.651)	(86.248)	(75.338)	(80.581)	(74.522)	(55.786)	(69.742)
govleft	152.522	-182.880*	142.977*	-123.399	200.393*	-180.119*	196.353*	-178.296*	126.596	-227.056**
	(105.022)	(99.795)	(75.454)	(97.334)	(116.895)	(101.844)	(116.29)	(99.766)	(108.519)	(113.950)
Intercept	6296.901	-3627.501**	2223.847	-944.815	1239.071	7147.908**	5479.077	-1326.798**	-1818.842**	121.114
	(10365.494)	(1443.639)	(2977.00)	(1676.693)	(852.079)	(3251.876)	(9759.017)	(668.68)	(787.610)	(508.953)
N	60	144	108	96	60	144	60	144	60	144
R^2	0.922	0.897	0.941	0.938	0.933	0.905	0.939	0.908	0.917	0.904
Pagan-H	36.086	78.192	50.017	39.610	18.580	47.865	15.531	43.211	28.247	66.821
	(0.010)	(0.000)	(0.000)	(0.005)	(0.549)	(0.000)	(0.688)	(0.001)	(0.1333)	(0.000)
F	13.10	14.70	263.49	68.74	9.10	35.00	3.01	33.44	11.60	11.52
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.041)	(0.000)	(0.000)	(0.000)
Anderson	31.196	43.810	272.296	105.668	36.991	95.697	14.073	79.809	49.582	83.204
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)
Hansen	5.220	0.422	1.852	1.472	8.491	1.621	6.741	1.517	10.683	6.921
	(0.073)	(0.809)	(0.603)	(0.688)	(0.036)	(0.654)	(0.034)	(0.468)	(0.0304)	(0.140)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% robust std. err. in brackets, p-value for the tests

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.5: Pooled model with spatially lagged dependent variable and time period fixed effects

Dependent Variable: Still wine specific tax rates - (2000 prices) Spatial Autoregressive Model - IV

Coefficient Estimates

Explanatory Variables	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	-0.550	0.428*	2.792	1.154***	1.088	1.705***	1.784*	1.935***	-0.099	-0.176
	(1.131)	(0.156)	(1.668)	(0.352)	(1.088)	(0.537)	(1.040)	(0.665)	(0.224)	(0.137)
total population	0.672	0.437*	0.107	0.082	0.847*	-0.532*	0.882*	-0.611	0.772	0.795*
	(0.597)	(0.255)	(0.527)	(0.219)	(0.462)	(0.314)	(0.466)	(0.423)	(0.513)	(0.408)
gdppc[a]	-0.044	0.005	-0.422	-0.023	-0.290	-0.028	-0.413	-0.042*	-0.102	0.045**
	(0.248)	(0.015)	(0.259)	(0.018)	(0.267)	(0.021)	(0.285)	(0.025)	(0.155)	(0.018)
govcons[b]	0.502	-0.454	0.906	-1.441	-0.922	1.122	-2.250	1.197	1.075	-2.505
	(2.501)	(1.166)	(2.225)	(1.404)	(2.398)	(1.677)	(2.788)	(2.024)	(2.521)	(1.581)
govright	5.653	-0.961	3.327	-3.729	6.247	-1.778	6.737	-1.438	6.198	-4.204
	(5.945)	(2.992)	(4.538)	(3.054)	(5.783)	(3.432)	(6.458)	(3.481)	(5.239)	(2.793)
govleft	6.822	-1.926	6.127	-4.937	10.162	-0.848	11.842	-0.733	7.398	-4.414
	(7.127)	(4.537)	(5.990)	(4.878)	(8.042)	(5.227)	(8.509)	(5.354)	(6.297)	(4.591)
Intercept	149.506	-365.351*	789.022*	112.577**	-386.615	380.532**	831.052	436.829	-81.251	-1.168
	(484.233)	(208.341)	(449.569)	(47.361)	(279.784)	(162.335)	(530.112)	(319.065)	(68.407)	(36.593)
N	60	144	60	144	60	144	60	144	60	144
R^2	0.966	0.981	0.975	0.978	0.978	0.974	0.974	0.972	0.978	0.979
Pagan-H	7.603	73.245	26.781	73.394	25.025	73.711	19.901	70.815	32.130	65.549
	(0.990)	(0.000)	(0.141)	(0.000)	(0.200)	(0.000)	(0.400)	(0.000)	(0.0568)	(0.000)
F	1.12	3.52	15.33	82.31	1.47	10.30	1.23	6.29	7.79	11.84
	(0.3511)	(0.017)	(0.000)	(0.000)	(0.229)	(0.000)	(0.312)	(0.000)	(0.000)	(0.000)
Anderson	3.875	22.721	59.179	179.607	11.199	53.329	7.801	27.158	51.91	86.72
	(0.275)	(0.000)	(0.000)	(0.000)	(0.024)	(0.000)	(0.050)	(0.050)	(0.000)	(0.000)
Hansen	0.046	0.280	5.257	0.223	7.089	1.640	7.801	0.179	7.261	7.938
	(0.9880)	(0.869)	(0.153)	(0.973)	(0.069)	(0.650)	(0.124)	(0.914)	(0.122)	(0.093)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% std. err. in brackets, p-values for the tests.

[a] GDP per capita (constant LCU)
[b] General government final consumption expenditure (% of GDP)

Tab. 4.6: Pooled model with spatially lagged dependent variable and time period fixed effects

Dependent Variable: Sparkling wine specific tax rates - (2000 prices) Spatial Autoregressive Model - IV

Explanatory Variables

Coefficient Estimates

	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	-0.763	1.208***	3.148	1.164***	1.248	1.565***	-0.349	1.452***	0.053	-0.019
	(2.513)	(0.339)	(1.875)	(0.281)	(0.823)	(0.437)	(1.625)	(0.337)	(1.491)	(0.169)
total population	1.831	-0.012	0.624	-0.062	2.714**	-0.637	1.965	-0.428	1.235	0.916
	(1.392)	(0.234)	(1.109)	(0.294)	(1.007)	(0.574)	(1.368)	(0.382)	(1.404)	(0.533)
gdppc[a]	-0.086	-0.061	-1.005*	-0.032	-0.859	-0.061**	-0.307	-0.063**	-0.631	0.058**
	(1.043)	(0.049)	(0.561)	(0.0232)	(0.571)	(0.026)	(0.737)	(0.029)	(0.391)	(0.027)
govcons[b]	-0.479	4.325	1.419	-2.295	-4.090	0.992	0.850	-0.098	-0.736	-4.892
	(5.748)	(3.143)	(4.351)	(2.307)	(5.439)	(2.276)	(5.412)	(2.397)	(4.747)	(2.496)
govright	12.373	5.309	6.451	-5.939	15.586	-3.312	14.205	-3.505	14.747	-6.597
	(10.725)	(6.791)	(9.307)	(4.748)	(13.534)	(4.973)	(10.637)	(4.839)	(9.449)	(4.615)
govleft	10.386	2.353	7.283	-8.726	19.741	-3.056	11.780	-3.893	13.684	-6.244
	(14.515)	(7.103)	(11.412)	(6.805)	(17.083)	(6.849)	(15.305)	(6.684)	(16.206)	(11.379)
Intercept	362.944	1.377	1814.085*	145.230**	-1358.000**	400.457	730.387	413.905	-177.544	9.884
	(1774.486)	(89.921)	(966.203)	(62.704)	(579.341)	(289.677)	(1252.846)	(307.862)	(122.735)	(50.763)
N	60	144	60	144	60	144	60	144	60	144
R^2	0.955	0.972	0.968	0.985	0.968	0.985	0.973	0.985	0.974	0.984
Pagan-H	6.233	39.733	24.507	56.467	17.926	52.699	26.354	49.833	24.536	58.582
	(0.997)	(0.003)	(0.220)	(0.000)	(0.592)	(0.000)	(0.091)	(0.000)	(0.267)	(0.000)
F	0.32	4.75	12.56	78.54	1.41	7.53	2.42	24.86	9.99	13.06
	(0.814)	(0.003)	(0.000)	(0.000)	(0.249)	(0.000)	(0.101)	(0.000)	(0.577)	(0.000)
Anderson	1.642	14.402	45.608	178.283	12.14	43.942	9.039	58.185	50.35	66.77
	(0.649)	(0.002)	(0.000)	(0.000)	(0.016)	(0.000)	(0.010)	(0.010)	(0.000)	(0.000)
Hansen	0.108	0.284	6.865	0.599	7.788	1.727	2.994	1.405	7.459	10.115
	(0.947)	0.867)	0.076)	0.896)	(0.050)	(0.630)	(0.083)	(0.235)	(0.113)	(0.038)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% std. err. in brackets, p-values for the tests.

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.7: Pooled model with spatially lagged dependent variable and country fixed effects

Dependent Variable: Cigarettes specific tax rates (2000 prices) - Spatial Autoregressive Model - IV

Coefficient Estimates

Explanatory Variables	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	0.500***	1.316***	0.797***	1.281*	0.902***	1.946***	0.679**	1.718***	0.423**	0.366***
	(0.181)	(0.287)	(0.396)	(0.239)	(0.312)	(0.353)	(0.308)	(0.288)	(0.184)	(0.102)
total population	0.559***	-0.273	0.379	-0.367 **	0.365**	-1.353***	0.502***	-0.852***	0.498***	0.364
	(0.126)	(0.327)	(0.167)	(0.264)	(0.150)	(0.383)	(0.138)	(0.321)	(0.144)	(0.354)
gdppc[a]	-0.036	-0.079**	-0.030	-0.038**	-0.032	-0.074***	-0.032	-0.082***	-0.016	0.017
	(0.033)	(0.039)	(0.037)	(0.018)	(0.028)	(0.025)	(0.036)	(0.026)	(0.039)	(0.016)
govcons[b]	0.723*	3.382**	0.160	-1.568	0.250	-0.667	0.279	-1.832	0.018	-1.369
	(0.428)	(1.597)	(0.680)	(1.531)	(0.525)	(1.300)	(0.609)	(1.284)	(0.745)	(1.505)
govright	-0.071	3.909	-0.777	-4.432	-0.444	-0.511	-0.430	-2.305	-0.691	-3.870
	(0.810)	(3.478)	(0.915)	(3.078)	(0.954)	(2.648)	(0.875)	(2.661)	(0.844)	(2.856)
govleft	1.352	5.250	0.514	-2.799	0.776	3.121	0.871	0.599	0.971	-3.384
	(1.238)	(4.356)	(1.322)	(4.428)	(1.341)	(4.349)	(1.285)	(4.234)	(1.341)	(4.163)
Intercept	72.936	119.609	86.016	185.304***	13.676	1145.768***	78.598	78.165***	30.238	-24.635
	(61.855)	(81.460)	(80.972)	(70.041)	(10.916)	(313.818)	(72.961)	(21.221)	(0.0873)	(44.816)
N	60	144	60	144	60	144	60	144	60	144
R^2	0.988	0.902	0.984	0.913	0.987	0.904	0.986	0.911	0.983	0.926
Pagan-H	21.060	55.911	28.90	68.621	30.077	74.39	23.243	73.408	29.021	67.369
	(0.393)	(0.000)	(0.116)	(0.000)	(0.068)	(0.000)	(0.331)	(0.000)	(0.113)	(0.000)
F	36.307	2.69	144.45	351.79	59.16	13.59	20.06	18.15	48.64	7.70
	(0.000)	(0.0341)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anderson	8.11	20.59	179.76	354.13	85.068	68.590	68.584	96.354	92.641	99.296
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hansen	0.51	5.07	1.84	1.32	1.26	3.46	4.32	5.14	6.437	16.485
	(0.916)	(0.166)	(0.764)	(0.856)	(0.738)	(0.325)	(0.363)	(0.272)	(0.168)	(0.002)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% std. err. in brackets, p-values fo the tests.

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.8: Pooled model with spatially lagged dependent variable and time period fixed effects

Dependent Variable: Tobacco specific tax rates (2000 prices) Spatial Autoregressive Model - IV

Explanatory Variables

Coefficient Estimates

	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	0.478	0.807***	1.082	0.812***	0.251	2.748***	1.909**	2.554***	0.040	-0.148
	(0.464)	(0.196)	(0.727)	(0.297)	(0.518)	(0.862)	(0.896)	(0.794)	(0.253)	(0.129)
total population	0.379**	0.184	0.248	-0.016	0.358**	-1.732***	-0.007	-1.405**	0.357*	0.6980
	(0.155)	(0.234)	(0.200)	(0.315)	(0.165)	(0.637)	(0.199)	(0.555)	(0.162)	(0.360)
gdppc[a]	-0.079	0.018	-0.087	0.032**	-0.057	-0.076	-0.072	-0.088*	0.800	0.083***
	(0.081)	(0.014)	(0.092)	(0.016)	(0.093)	(0.046)	(0.087)	(0.046)	(1.056)	(0.021)
govcons[b]	0.766	0.024	0.036	-2.499	0.469	-3.908**	-1.596	-4.940***	-0.058	-1.647
	(0.562)	(0.870)	(0.989)	(1.647)	(0.885)	(1.621)	(1.312)	(1.656)	(0.807)	(1.456)
govright	-0.045	-0.374	0.604	-4.147	-0.225	-1.543	1.929	-2.290	-0.388	-3.392
	(1.640)	(0.747)	(1.230)	(3.049)	(0.740)	(3.367)	(2.175)	(3.157)	(0.672)	(3.118)
govleft	2.334	1.014	2.679	-2.852	2.282	1.477	3.348	-0.066	2.258	-1.961
	(1.724)	(2.613)	(2.110)	(3.985)	(1.822)	(4.524)	(2.513)	(4.218)	(1.822)	(3.703)
Intercept	150.985	-12.629	-190.583	120.183***	-167.127*	1140.598***	167.701	184.408***	-47.570	-550.809
	(156.579)	(41.399)	(157.931)	(32.095)	(93.438)	(372.150)	(172.860)	(33.411)	(24.520)	(310.387)
N	60	144	60	144	60	144	60	144	60	144
R^2	0.977	0.967	0.964	0.934	0.968	0.905	0.967	0.910	0.967	0.940
Pagan-H	25.735	53.904	29.912	65.535	29.396	40.210	20.573	44.823	31.256	47.757
	(0.174)	(0.000)	(0.071)	(0.000)	(0.060)	(0.003)	(0.360)	(0.000)	(0.070)	(0.000)
F	3.56	2.07	89.92	254.86	3.90	4.36	8.48	9.66	10.48	23.48
	(0.014)	(0.088)	(0.000)	(0.000)	(0.015)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)
Anderson	21.036	21.804	133.426	309.468	20.368	18.907	25.016	33.314	57.666	151.75
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hansen	1.605	3.590	2.912	2.991	5.262	2.899	5.760	1.970	9.931	14.713
	(0.658)	(0.309)	(0.405)	(0.559)	(0.072)	(0.234)	(0.056)	(0.373)	(0.041)	(0.005)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% robust std. err. in brackets, p-values for the tests.

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.9: Pooled model with spatially lagged dependent variable and time period fixed effects

Dependent Variable: Petrol specific tax rates (2000 prices) Spatial Autoregressive Model - IV

Coefficient Estimates

Explanatory Variables	Wcont		Wunif		Wdis1		Wdis2		Wrand	
	87-92	93-04	87-96	97-04	87-92	93-04	87-92	93-04	87-92	93-04
WTAX	0.228	0.838***	0.857***	0.898**	0.876***	0.987***	0.958***	1.006***	0.731***	0.101
	(0.320)	(0.188)	(0.097)	(0.341)	(0.310)	(0.163)	(0.325)	(0.163)	(0.182)	(0.119)
total population	7.026***	1.711	3.760***	-0.175	2.702	0.065	2.428	0.306	4.223***	5.430**
	(1.532)	(1.272)	(0.880)	(0.749)	(1.792)	(0.957)	(1.797)	(0.942)	(1.387)	(2.306)
gdppc[a]	0.457	-0.071	-0.056	0.252**	-0.067	0.089	-0.201	0.046	-0.643	0.444***
	(0.882)	(0.180)	(0.222)	(0.099)	(0.809)	(0.092)	(0.839)	(0.095)	(0.772)	(0.072)
govcons[b]	14.773	1.249	-0.056	-6.488	8.459	-4.803	5.869	-5.950	-17.160***	-1.877
	(11.300)	(6.715)	(0.222)	(5.446)	(8.734)	(5.425)	(8.625)	(5.246)	(6.951)	(5.379)
govright	12.320	12.225	14.816	-4.230	12.161	-1.282	15.272	-2.148	17.259	-13.594
	(13.116)	(14.021)	(13.089)	(7.239)	(13.007)	(10.369)	(13.528)	(10.119)	(14.418)	(9.904)
govleft	32.373	-5.617	22.176	-32.402**	26.515	-14.057	28.390	-15.506	29.568*	-17.397
	(21.974)	(20.368)	(15.287)	(13.062)	(16.879)	(18.382)	(16.983)	(18.177)	(17.086)	(20.210)
Intercept	-1451.193	44.472	-168.797	-341.890*	-195.870	113.205	147.608	-40.039	-344.350	99.151
	(1887.293)	(367.035)	(499.900)	(186.498)	(199.707)	(753.912)	(1745.423)	(101.696)	(229.260)	(197.538)
N	60	144	108	96	60	144	60	144	60	144
R^2	0.835	0.753	0.878	0.927	0.866	0.804	0.867	0.809	0.878	0.986
Pagan-H	27.673	65.666	29.227	36.780	26.364	76.618	25.704	74.320	23.378	70.540
	(0.117)	(0.000)	(0.083)	(0.0124)	(0.192)	(0.000)	(0.218)	(0.000)	(0.324)	(0.000)
F	24.26	6.04	230.29	172.26	53.01	37.45	44.19	33.88	118.18	33.92
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anderson	76.678	41.623	277.299	195.515	129.665	177.307	116.503	173.625	168.385	125.48
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hansen	6.396	2.657	0.644	2.915	1.485	4.578	1.355	3.412	16.90	30.46
	(0.093)	(0.447)	(0.886)	(0.405)	(0.829)	(0.333)	(0.851)	(0.491)	(0.002)	(0.003)
Chow	Rej Ho		Rej Ho		Rej Ho		Rej Ho		Rej Ho	

Significance levels : * : 10% ** : 5% *** : 1% std. err. in brackets, p-values for the tests.

[a] GDP per capita (constant LCU)
[b] General government final consumption expenditure (% of GDP)

Tab. 4.10: Spatial Durbin Model with country fixed effects

Dependent Variable: specific tax rate

Explanatory Variables	Coefficient Estimates Beer		Ethyl alcohol	
	87-92	93-04	87-92	93-04
WTAX	0.481***	0.584***	0.160	0.280***
	(0.00)	(0.00)	(0.19)	(0.00)
total population	0.000	-0.000	-0.877***	0.247
	(0.37)	(0.48)	(0.00)	(0.44)
gdppc[a]	0.009	0.033***	7.80***	9.444***
	(0.20)	(0.00)	(0.00)	(0.00)
govcons[b]	0.001	-0.026***	7.152***	-3.902***
	(0.59)	(0.00)	(0.00)	(0.00)
govright	0.053	0.404**	37.89	73.03
	(0.79)	(0.05)	(0.65)	(0.44)
govleft	0.001	0.446*	132.882	-3.394
	(0.99)	(0.07)	(0.15)	(0.97)
weighted total population	0.014***	-0.007***	-2.739*	0.226
	(0.00)	(0.00)	(0.07)	(0.82)
weighted gdppc	0.006***	-0.004***	0.028	-0.745
	(0.00)	(0.00)	(0.96)	(0.06)
weighted govcons	0.039***	-0.063***	15.21***	-15.72***
	(0.00)	(0.00)	(0.00)	(0.00)
weighted govright	-0.056	-0.400	-186.5	504.4***
	(0.89)	(0.28)	(0.18)	(0.00)
weighted govleft	0.183	-0.586	67.38	-592.1***
	(0.59)	(0.12)	(0.63)	(0.00)
R-squared	0.971	0.889	0.952	0.897
Rbar-squared	0.956	0.862	0.926	0.874
Nobs,Nvar,TNvar	60, 16, 22	144, 16, 29	60, 12, 18	144, 16, 29
log-likelihood	-54.7804	-186.751	-399.390	-1055.15
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	

Significance levels : * : 10% ** : 5% *** : 1% p-values in brackets, contiguity matrix used.

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.11: Spatial Durbin Model with time period fixed effects

Dependent Variable: specific tax rate

Explanatory Variables	Coefficient Estimates			
	Still wine		Sparkling wine	
	87-92	93-04	87-92	93-04
WTAX	0.527***	0.68***	0.326***	0.665***
	(0.00)	(0.00)	(0.00)	(0.00)
total population	-0.005	-0.038*	-0.133**	-0.096**
	(0.79)	(0.10)	(0.02)	(0.03)
gdppc[a]	0.336*	1.143***	-0.641	0.478**
	(0.08)	(0.00)	(0.13)	(0.05)
govcons[b]	0.009	-0.852***	0.752***	-1.689***
	(0.92)	(0.00)	(0.00)	(0.00)
govright	6.055	13.80**	11.82	21.23
	(0.28)	(0.04)	(0.33)	(0.11)
govleft	10.78	17.53**	22.02	14.48
	(0.13)	(0.03)	(0.15)	(0.36)
weighted total population	0.305***	-0.259***	0.770***	0.030
	(0.00)	(0.00)	(0.00)	(0.82)
weighted gdppc	0.103***	-0.163***	0.858***	0.244***
	(0.00)	(0.00)	(0.00)	(0.00)
weighted govcons	1.508***	-1.975***	2.092***	-5.544***
	(0.00)	(0.00)	(0.00)	(0.00)
weighted govright	3.200	-21.30***	15.11	-9.912
	(0.733)	(0.08)	(0.45)	(0.67)
weighted govleft	-0.267	-30.97***	10.61	-5.214
	(0.97)	(0.01)	(0.60)	(0.82)
R-squared	0.977	0.914	0.952	0.897
Rbar-squared	0.964	0.893	0.926	0.874
Nobs,Nvar,TNvar	60, 16, 22	144, 16, 29	60, 12, 18	144, 16, 29
log-likelihood	-256.940	-695.452	196.081	-786.955
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	

Significance levels : * : 10% ** : 5% *** : 1% p-values in brackets, contiguity matrix used.

[a] GDP per capita (constant LCU)

[b] General government final consumption expenditure (% of GDP)

Tab. 4.12: Spatial Durbin Model with time period fixed effects

Dependent Variable: specific tax rate (2000 prices)

Explanatory Variables	Coefficient Estimates			
	Cigarettes		Tobacco	
	87-92	93-04	87-92	93-04
WTAX	0.171	0.318***	0.384***	0.603***
	(0.12)	(0.00)	(0.00)	(0.00)
total population	0.015	-0.014	0.036***	0.017
	(0.17)	(0.37))	(0.00)	(0.16)
gdppc[a]	0.161	0.656***	0.139	0.636***
	(0.137)	(0.00)	(0.13)	(0.00)
govcons[b]	0.20***	-0.326***	-0.021	-0.520***
	(0.00)	(0.00)	(0.67)	(0.00)
govright	-1.196	11.16**	1.192	8.163**
	(0.70)	(0.02)	(0.65)	(0.02)
govleft	2.920	12.85**	4.516	12.30***
	(0.46)	(0.02)	(0.18)	(0.00)
weighted total population	0.148***	-0.218***	0.138***	-0.155***
	(0.00)	(0.00)	(0.00)	(0.00)
weighted gdppc	0.118***	-0.040	0.0131	-0.104***
	(0.00)	(0.04)	(0.56)	(0.00)
weighted govcons	0.798***	-1.261***	1.060***	-0.732***
	(0.00)	(0.00)	(0.00)	(0.00)
weighted govright	-4.279	-20.48***	0.151	-19.50***
	(0.40)	(0.01)	(0.97)	(0.00)
weighted govleft	-3.148	-28.72***	2.643	-28.80***
	(0.54)	(0.01)	(0.56)	(0.00)
R-squared	0.927	0.787	0.952	0.881
Rbar-squared	0.887	0.736	0.926	0.852
Nobs,Nvar,TNvar	60, 16, 22	144, 16, 29	60, 12, 18	144, 16, 29
log-likelihood	-201.616	-626.992	196.081	-599.014
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	

Significance levels : * : 10% ** : 5% *** : 1% p-values in brackets, contiguity matrix used.

[a] GDP per capita (constant LCU)
[b] General government final consumption expenditure (% of GDP)

Tab. 4.13: Spatial models with country fixed effects

	Coefficient Estimates			
	SAR		SEM	
	pre-1992	post-1992	pre-1992	post-1992
Dependent Variable: **Beer**				
WTAX	0.658***	0.556***	0.568***	-0.030***
Wald Test	(0.00)	(0.00)	(0.00)	(0.69)
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	
Dependent Variable: **Ethyl alcohol**				
WTAX	0.812***	0.553***	0.814***	-0.104***
Wald Test	(0.00)	(0.00)	(0.00)	(0.00)
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	
Dependent Variable: **Still wine**				
WTAX	0.812***	0.711***	0.756***	0.674***
Wald Test	(0.00)	(0.00)	(0.00)	(0.00)
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	
Dependent Variable: **Sparkling wine**				
WTAX	0.781***	0.767***	0.797***	0.656***
Wald Test	(0.00)	(0.00)	(0.00)	(0.00)
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	
Dependent Variable: **Cigarettes**				
WTAX	0.527***	0.682***	0.772***	0.443***
Wald Test	(0.00)	(0.00)	(0.00)	(0.00)
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	
Dependent Variable: **Tobacco**				
WTAX	0.659***	0.527***	0.721***	0.679***
Wald Test	(0.00)	(0.00)	(0.00)	(0.00)
Chow Test at 5%	rej Ho → structural break		rej Ho → structural break	
Dependent Variable: **Petrol**				
WTAX	-0.557***	0.161***	-0.839***	0.127***
Wald Test	(0.00)	(0.00)	(0.00)	(0.00)
Chow Test at 5%	rej Ho → structural break		do not rej Ho →no structural break	

Significance levels : * : 10% ** : 5% *** : 1% p-values in brackets, contiguity matrix used.

Tab. 4.14: Estimation with Minimum Tax Rate

Spatial Autoregressive Model - IV

Dependent variable: beer (2000 prices)

Variable	**Coefficient** (Std. Err.)
WTAX	-0.327
	(0.262)
mintax × WTAX	0.479**
	(0.196)
mintax	-0.037
	(0.218)
total population	0.023***
	(0.007)
gdppc	-0.002**
	(0.001)
govcons	0.040
	(0.032)
govright	-0.048
	(0.085)
govleft	-0.059
	(0.126)
N	204
R^2	0.96
$F_{(19,184)}$	314.036

Significance levels : * : 10% ** : 5% *** : 1%

CONCLUSION

5. CONCLUSION

The main objective of the first part of this thesis was to outline both theoretical and empirical models that could address issues concerning the financing of higher education. On the common thread of the English 2004 HE reform we analyzed two public funded student loans: income contingent loan and mortgage loan.

In Chapter 1 we compared an ICL and a ML, when graduate incomes are uncertain. The findings of the model have been calibrated using real data on graduate wages, obtained by the 1970 British Cohort Survey. We assumed that the graduate income is affected by a single lifetime shock, and we computed the individual expected utilities under an ICL and a ML for risk neutral and risk averse individuals. Our first result, supported by empirical simulations, is that for risk neutral individuals the preference for one system relative to the other depends strongly on the presence of hidden subsidies. For risk averse individuals we evaluated the effects of our model under different possible scenarios. We used information on graduate wages and relative uncertainty controlling for individual characteristics, family background, degree courses and job sector. Considering the English Higher Education reform for policy implications, the main result is that for high wage uncertainty an ICL is the preferred system. Therefore, people from low educated parents background, males over females, people working in the private sector prefer an ICL. Instead those with high income and low uncertainty, such as those with highly educated mother, prefer a ML. Another important result is that for increasing risk aversion an ICL is more preferred, this confirms the conviction that this system is a better guarantee for uncertain future. Moreover, a stable career in the public sector gives an income not too high but also the most stable in terms of uncertainty, therefore a ML produces higher utility. However, if we don't rule out the hidden subsidies an ICL becomes preferred also when the degree of uncertainty is very low.

In the second part of Chapter 1, we modified the assumptions on the income, allowing a stochastic growth across the working life. The results of the new framework have in common with the previous the preference for an ICL for increasing levels of uncertainty. However the factors that affect the choice of one system over the other are different. The size of the starting wage is the main discriminant: when the income is low an ICL is the preferred system. For high initial income a ML gives higher utility instead.

In Chapter 2 we extended our theoretical model introducing the possibility of default. This happens when graduates do not earn enough to fully repay their loan. The new model has been structured as a simultaneous game between students and government, where the students decide the level of effort to put in their studies and the government chooses the default premium that keeps its budget balanced. By comparing the individuals payoff under all the possible cases and taking into account of a participation constraint, we derived analytically the cutoffs of the costs of effort required to implement the Nash equilibria. The equilibria are high effort, low effort or multiple equilibria. We play the same game for a ML and an ICL. We find that under an ICL relative to a ML, the students are willing to spend more effort to participate in higher education. Moreover, under an ICL the individuals prefer putting higher effort in order to receive with certainty a high wage and avoid the default. Conversely, under a ML the levels of effort are lower and the students gamble on obtaining a low income and being in default.

In Chapter 3 we followed the common thread of the comparison of a ML and an ICL on the basis of the English 2004 HE reform, but assuming a financial perspective. We evaluated the investment in HE versus high school in terms of internal rates of return. We computed the net present values of the earning profiles for HE graduates and HS graduates. However these profiles are generated given the empirical estimation of the wages for different levels of education. We used an ordered probit selection model to correct for endogeneity, and we allowed for non separability in the earnings function between the schooling effect and the age effect. We then computed the net present value of the investment in HE and HS taking into account both the tuition and the opportunity cost of education; we considered the foregone earnings as sunk cost of the investment in schooling. We finally derived the IRR for free HE education, and for financing schemes based on ICL and ML. In particular, we found that the current English ICL gives higher IRR than the previous ML. Moreover it is evident the big benefit for graduates in terms of high government subsidies due to the zero real interest on the loan.

In the second part of the thesis we undertook an empirical analysis on commodity tax competition. In Chapter 4 we test for the presence of tax competition after the introduction of the Single Market in EU in 1993. We first extended the basic model of Nielsen (2001), finding the conditions of cross-border shopping and the government reaction functions in three different market regimes: duty-free market before 1993, mixed market up to 1999, single market after 1999. The theoretical prediction is no strategic interaction before the start of the Single Market and increasing tax competition after. We employed a balanced panel data set of 12 EU countries over the period 1987-2004 and excise taxes on seven commodities. We specified our empirical model in a framework of spatial econometrics, where the excise tax in a given country depends linearly on the weighted average of other countries' taxes, and a set of control variables. The equation has been estimated separately on subsamples before and after 1993. We always found a structural break after 1993 with all the seven excise duties, which means that the introduction of the Single Market has modified the tax setting among the EU countries. Moreover, looking across all seven taxes and a variety of different weighting matrices suggested by the theory, we found that the coefficient measuring strategic interaction is *always* significantly positive and, in most of the cases, rising after 1993. This suggests an increasing tax competition among neighbouring countries, consistently with the theoretical prediction.

BIBLIOGRAPHY

[1] Anselin L., (1988) Spatial Econometrics: Methods and Models, *Dordrecht: Kluwer Academic Publishers.*

[6] Angrist, J. and Imbens G. (1995), Two Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity, *Journal of the American Statistical Association*, Vol. 90, pp. 431442.

[3] Angrist, J. and Inbens, G. and Rubin, D. (1996) Identification of Causal effects Using Instrumental Variables, *Journal of Econometrics*, Vol. 71, No. 1-2, pp. 145-160.

[6] Angrist, J. and Krueger A. (1991), Does Compulsory Schooling Attendance Affect Schooling and Earnings?, *Quarterly Journal of Economics*, Vol. 106, pp. 979-1014.

[6] Angrist, J. and Krueger, A. (1999) Empirical strategies in labor economics, *The Handbook of labor economics*, Vol.III, ch. 23.

[6] Angrist, J. and Krueger, A. (2001), Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments, *Journal of Economic Perspectives*, Vol. 15, N. 4, pp. 6985.

[7] Arbia G., (2006) Spatial Econometrics, Statistical Foundations and Applications to Regional Convergence,*Springer Berlin Heidelberg.*

[8] Arellano, M. and Meghir, C. (1992) Female Labour Supply and On-the-Job Search: An Empirical Model Estimated using Complementary Data Sets, *The Review of Economic Studies*, Vol. 59, No. 3., pp. 537-559.

[9] Armingeon, K. and Leimgruber, P. and Beyeler, M. and Menegale, S. (2006) Comparative Political Data Set 1960-2004, *Institute of Political Science*, University of Berne 2006.

[10] Ashenfelter, O. and Krueger A. (1994), Estimates of the Economic Return to Schooling from a New Sample of Twins, *American Economic Review*, 84(5).

[11] Ashworth J. and Heyndels B. (2007) Selection Bias and Peer Effects in Team Sports, *Journal of Sport Economics*, Vol. 8, No. 4, pp. 355-377.

[12] Baltagi, B. H. and D. Levin (1986) Estimating dynamic demand for cigarettes using panel data, *Review of Economics and Statistics* 68, 148-155.

[13] Baltagi, B. H. (2005) Econometric Analysis of Panel Data, third edition *John Wiley. & Sons.*

[14] Barr, N. (1993), The Economics of the Welfare State, *Second Edition*, Weidenfeld and Nicholson, London.

[16] Barr, N. (1993), Alternative Funding Resources for Higher Education, *The Economic Journal*, No. 103, pp.718-728.

[16] Barr, N. and Crawford, I. (2005). Financing Higher Education: Answers from the UK. London and New York: Routledge.

[17] Baumgartner, H. J. and Steiner, V. (2003), Student Aid, Repayment Obligations and Enrolment into Higher Education in Germany - Evidence from a Natural Experiment, *Working Paper*, DIW Berlin and Free University of Berlin.

[18] Becker, Gary (1975), Human Capital, *The University of Chicago Press.*

[19] Bedard, K. and Dhuey, E. (2006), The persistence of early childhood maturity: international evidence of long-run age effects, *Quarterly Journal of Economics*, Vol. 121, pp. 143772.

[22] Belzil, C. and Leonardi, M. (2007) Can risk aversion explain schooling attainments? Evidence from Italy, *Labour Economics*, Vol. 14, Issue 6.

[22] Belzil, C. (2002), Unobserved Ability and the Return to Schooling, *Econometrica*, 70 (5), 2075-2091.

[22] Belzil, C. and Hansen, J. (2004), "Earnings Dispersion, Risk Aversion and Education", *Research in Labor Economics*, Vol. 23, 335-358, 2004.

[23] Besley, T., and H. S. Rosen (1998), Vertical externalities in tax setting:evidence from gasoline and cigarettes, *Journal of Public Economics* 70, 383-398.

[24] Blanden J. and Gregg P. (2004) Family Income and Educational Attainment: a Review of Approaches and Evidence from Britain, *Oxford Review of Economic Policy*, Vol. 20, N. 2.

[25] Blundell R. and Costa Dias M. (2002), Alternative Approaches to Evaluation in Empirical Microeconomics, *The Institute of Fiscal Studies* working paper CWP10/02.

[26] Brueckner, J. K. (2003) Strategic Interaction Among Governments: an Overview of empirical studies, *International Regional Science Review* 26, (2) 175-88.

[27] Brueckner, J.K., and L. A. Saavedra (2000) Do Local Governments Engage in Strategic Property-Tax Competition?, *National Tax Journal* vol. LIV, No. 2.

[28] Byner, John and Butler, Nevil et al. (2002), The design and conduct of the 1999-2000 surveys of the 1970 British Cohort Study, Centre for Longitudinal Studies, Institute of Education, University of London.

[29] Calderwood, L. Hawkes, D. Nathan, G. Plewis, I. (2004), 1970 British Cohort Study Technical Report, Centre for Longitudinal Studies, Institute of Education, University of London.

[36] Cameron, Stephen and Taber, Christopher (1999), Borrowing Constraints and the Returns to Schooling, *Working Paper* Department of Economics Columbia University, Department of Economics and Institute for Policy Research Northwestern University.

[31] Cameron, Stephen and Heckman, James (1998), Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males, *Journal of Political Economy*, vol. 106, n.2.

[32] Card, D. (1999), The Casual Effect of Education on Earnings, in: Orley Ashenfelter and David Card, eds., *Handbook of Labour Economics*, Vol. 3, North Holland, Amsterdam.

[33] Card, D. (2001), Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica*, vol.69 n.5, pp. 1127-1160

[34] Card, D. (1994), Earnings , Schooling, and Ability Revisited, *Working Paper n°331* Industrial Relations Section Princeton University.

[35] Card, D. (1995b), Using Geographic Variation in College Proximity to Estimate the Returns to Schooling, in: Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp. University of Toronto Press, pp. 201222.

[36] Cameron, S. and Heckman, J. (1998) Life Cycle Schooling and Educational Selectivity: Models and Choice, *Journal of Political Economy*, Vol. 106, N. 2.

[37] Carneiro, Pedro and Heckman, James (2002), The evidence on credit constraints in post-secondary schooling, *Working paper n.9055*, National Bureau of Economic Research.

[38] Carneiro, Pedro and Heckman, James and Vytlacil Edward (2001), Estimating the return to education when it varies among individuals, *Research supported by NFS*, University of Chicago, Stanford University.

[39] Case, A. C. Hines, James R. Jr. and Rosen, Harvey S. (1992) Budget spillovers and fiscal policy interdependence. Evidence from the States. *Journal of Public Economics* 52, 285-307.

[40] Chapman, B. (2007) Government Managing Risk - Income Contingent Loans for Social and Economic Progress, ed. Routledge

[41] Chapman, B. (1997), Some Financing Issues for Australian Higher Education Teaching, *Discussion Paper No. 384*, Centre for Economic Policy Research Australian National University, Canberra.

[42] Chapman, B. (1997) Conceptual Issues and the Australian Experience with Income Contingent Charges for Higher Education', *The Economic Journal* , Vol. 107 (442), 1, pp. 738-751.

[43] Chen, S. H. (2001), Is Investing in College Education Risky?, *Working Paper*, University of Rochester and State University of New York at Albany.

[44] Cheung, Bryan (2003), Higher Education Financing Policy: Mechanism and Effects, *Working Paper*, University of South Australia.

[45] Chiburis R. and Lokshin M. (2007) Maximum likelihood and two-step estimation of an ordered-probit selection model, *Stata Journal*, Vol. 7, issue 2, pp. 167-182

[46] Clifton-Fearnside, Alex and Whitmarsh, Alyson (2004), Labour Force Survey Reweighting and Seasonal Adjustment Review, *Office for National Statistics*, London.

[47] Coates, R. M. (1986) Estimating dynamic demand for cigarettes using panel data, *Review of Economics and Statistics* Vol. 68, pp. 148-155.

[48] Conlon, G. (2005) The Determinants of Undertaking Academic and Vocational Qualifications in the United Kingdom, *Education Economics*, Vol. 13, N.3, pp. 299-313.

[49] Crawford C., Dearden L. and Meghir C. (2007) When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England, *The Institute for Fiscal Studies.*

[50] Dearden, L. Fitzsimons, E. Goodman, A. and Kaplan, G. (2008) Higher Education Funding Reforms in England: The Distributional Effects and the Shifting Balance of Costs, *Economic Journal*, Vol. 118, Issue 526, pp. F100-F125.

[51] Department of Health, Chief Medical Officer (2005) On the state of the public health: Annual report of the Chief Medical Officer 2004, Product number 269349, Copyright holder Crown.

[52] Devereux, M. P. and Lockwood, B. and Redoano, M (2007) Horizontal and Vertical Indirect Tax Competition: Theory and Some Evidence from the USA, *Journal of Public Economics*, vol. 91(3-4), pp. 451-479.

[53] Diacu, Florin (2000), An Introduction to Differential Equations: Order and Caos, *W. H. Freeman and Company*, New York.

[54] Dixit, Avinash K. and Pindyck, Robert S. (1994), Investment under Uncertainty, *Princeton University Press*, Princeton, New Jersey.

[55] Eaton, Jonathan and Rosen, Harvey (1980), Taxation, Human Capital, and Uncertainty, *The American Economic Review*, September, pp. 705-714.

[56] Egger, P. and Pfaffermayr, M. and Winner, H. (2005) Commodity taxation in a 'linear' world: a spatial panel approach, *Regional Science and Urban Economics* 35, 527-541.

[57] Elhorst, J. P. (2001) Dynamic models in space and time, *Geographical Analysis* 33, 119-140.

[58] Elhorst, J. P. and Allers, M. (2005) Tax mimicking and yardstick competition among local governments in the Netherlands, *International Tax and Public Finance*, 12 493-513.

[59] European Commission, Directorate General XXI Customs and Indirect Taxation, Excise Duty tables.

[60] Friedman, M. (1975) The role of government in education, in (R.A. Solo ed.), *Economics and the Public Interest*, pp. 123-44, New Brunswick, NJ: Rutgers University Press.

[61] Garcia-Penalosa, Cecilia and Walde, Kalus (2000), Efficiency and Equity Effects of Subsisdies to Higher Education, *Oxford Economic Paper*, No. 2, pp. 702-722.

[62] Goodman, Alissa and Kaplan Greg (2003), 'Study Now, Pay Later' or 'HE for Free'? An Assessment of Alternative Proposals for Higher Education Finance, *Commentary 94* The Institute For Fiscal Studies.

[63] Griliches, Z. (1977), Estimating the Returns to Schooling: Some Econometric Problems, *Econometrica*, Vol. 45, pp. 121.

[64] Harmon, C. and Walker, I. (1995), Estimates of the Economic Returns to Schooling for the United Kingdom, *American Economic Review*, Vol. 85, N. 5, pp. 1278-1286.

[65] Hartog, J. and Vijverberg, W. P. M. (2007) On Compensation for Risk Aversion and skewness affection in wages, *Labour Economics*, Vol. 14, Issue 6.

[66] Hartog, J. and Serrano, L. D. (2007), "Earnings Risk and Demand for Higher Education: A Cross-Section Test for Spain, *Journal of Applied Economics*, Vol. X, pp. 1-28

[67] Haufler, A. (1996), Tax competition with different preferences for public goods: conflict or harmony of interest? *International Tax and Public Finance* 3, 5-28.

[68] Hause, John C. (1974), "The Risk Element in Occupational and Educational Choices: Comment", *The Journal of Political Economy*, Vol. 82, No. 4 (Jul.- Aug.), pp. 803-807.

[69] "Higher Education Bill" (2004), Published by Authority of The House of Commons London: The Stationery Office.

[70] Heckman, Lochner L. and Todd P. (2008) Earnings Functions and Rates of Return, *Journal of Human Capital*, Vol. 2, no. 1

[71] Heckman, J. (1978), Dummy Endogenous Variable in a Simultaneous Equations System, *Econometrica*, Vol. 46, N. 4, pp. 93160.

[72] Heckman, J. (1990), Varieties of Selection Bias, *AEA Papers and Proceedings*, Vol. 80(2), pp. 313318.

[73] Heckman, James and Vytlacil, Edward (1998), Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is correlated with Schooling, *The Journal of Human Resources*, Vol. 33, No. 4. , pp. 974-987.

[74] Heckman, James and Vytlacil, Edward (1999) Models of treatment effects when responses are heterogenous, *Proceeding National Academic Science* vol.96, pp.4730-4734, USA.

[75] HEFCE (2002), Funding Higher Education in England: How the HEFCE Allocates its Funds, *Report*, Higher Education Funding Council, UK.

[76] Heyndels, B., Vuchelen, J.(1998) "Tax mimicking among Belgian municipalities", *National Tax Journal*, LI (1), 89-100.

[77] Hartley, H.O. Hocking, R.R. (1971), The analysis of incomplete data, *Biometrics* 27, 783-823.

[78] Higham, D. J. (2001), An Algorithmic Introduction to the Numerical Simulation of Stochastic Differential Equations, *SIAM Review*, Vol.43, No.3, pp.525-546.

[79] Hogan, V. and Walker, Ian (2007), Education Choice under Uncertainty and Public Policy, *Labour Economics Paper*, Vol.14, Issue 6, pp. 894-912.

[80] Ichino, A. and WinterEbmer, R. (1999), Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments, *European Economic Review*, Vol. 43, pp. 889901.

[81] Ichino, A. (2006) The Problem of Causality in the Analysis of Educational Choices and Labor Market Outcomes, *University of Bologna and CEPR)*, Slides for Lectures.

[82] Kanbur, R., and M. Keen (1993), Jeux sans Frontiers: Tax competition and Tax Coordination when countries differ in size, *American Economic Review* 83, 887-892.

[83] Keane, M. and Wolpin, K. (1997) The career decisions of young men, *Journal of Political Economy*, Vol. 105, no. 3, pp. 473522.

[84] Kenny, L.W. and Lee, L.-F. and Maddala, G.S. and Trost, R.P. (1979) Returns to college education: An investigation of self-selection bias based on the project talent data, *International Economic Review* Vol. 20, pp. 775-789.

[85] Klebaner, Fima C. (1999), Introduction to Stochastic Calculus with Applications, Imperial College Press.

[86] Jacobs, Bas (2002), An Investigation of Education Finance Reform, *CPB Discussion Paper*, No. 9.

[87] Jacobs, Bas (2003), Optimal Financing of Education with capital Insurance Market Imperfections, *Working Paper* University of Amsterdam and CPB Netherlands.

[88] Ladd, H. F. (1992) Mimicking of local tax burdens amomg neighboring countries, *Public Finance Quarterly*, 20 No.4, 450-67.

[89] Lalonde, R. (1986), Evaluating the Econometric Evaluations of Training Programs with Experimental Data, *American Economic Review*, 76, 604-620

[90] Lee, D. (2001) An estimable dynamic general equilibrium model of work, schooling and occupational choice. Ph.D. thesis, University of Pennsylvania

[91] LeSage, J. P. (1999), The Theory and Practice of Spatial Econometrics, *Department of Economics University of Toledo.*

[92] Lockwood, B. (1993) Tax competition in a customs union under destination and origin principle, *Journal of Public Economics* 53, 141-162.

[93] Makepeace G., Woods L., Joshi H., Dolton P. and Galinda-Rueda F. (2003) From School to the Labour Market, (eds) Changing Britain, Changing Lives Three Generations at the Turn of the Century, Institute of Education.

[94] Maddala, G.S. (1971) The Use of Variance Components Models in Polling Cross Section and Time Series Data, *Econometrica* 39, pp. 341-358.

[95] Maddala, G.S. (1993) Limited Dependent and Qualitative Variables in Econometrics, *Cambridge University Press.*

[96] Mas-Colell, Andreu and Whinston, Michael D. and Green, Jerry R. (1995), Microeconomic Theory, *Oxford University Press*, New York.

[97] Meyer, J. (2007) Representing Risk Preferences in Expected Utility Based Decision Models, Economics and Management of Risk in Agriculture and Natural Resources Annual Meeting, Gulf Shores, AL March 15-17, 2007.

[98] Moffit, R. A. (1999) Models of treatment effects when responses are heterogenous, Commentary *Proceeding National Academic Science* vol.96, pp.6575-6576, USA.

[99] Nelson, M. A. (2002) Using excise taxes to finance state government: do neighboring state taxation policy and cross-border markers matter?, *Journal of Regional Science* 42, 731-752.

[100] Nerlove, Marc (1975), Some Problems in the Use of Income Contingent Loans for the Finance of Higher Education, *The Journal of Political Economy*, vol.83, No. 1.

[101] Nielsen, S. B.(2001) A Simple Model of Commodity Taxation and Cross-Border Shopping, *Scandinavian Journal of Economics* 103, 599-623.

[102] OECD (2002) Education at Glance, Organisation for Economic Cooperation and Development, Paris.

[103] Olson, Lawrence and White, Halbert and Shefrin, H. M. (1979), Optimal Investment in Schooling when Incomes are Risky, *The Journal of Political Economy*, Vol. 87, No. 3(Jun.), pp. 522-539.

[104] Ohsawa, Y. (1999) Cross-border shopping and commodity tax competition among governments, *Regional Science and Urban Economics* 29, 33-51.

[105] Oreopoulos, P. (2008) Should We Raise the Minimum School Leaving Age to Help Disadvantaged Youth? Evidence from Recent Changes to Compulsory Schooling in the United States, *Book chapter in "An Economic Framework for Understanding and Assisting Disadvantaged Youth", Jonathan Gruber (ed)*, NBER, forthcoming.

[106] Pace, R.K. and Barry, R. (1998) Spatial Statistics Toolbox 1.0, *Real Estate Research Institute*, Louisiana State University, Baton Rouge, LA.

[107] Padula, M. and Pistaferri L. (2001), Education, Employment and Wage Risk, *Working Paper No. 67*, Centre for Studies in Economics and Finance (CSEF), University of Salerno, Italy.

[108] Puhani, P. and Weber, A. (2005), Does the early bird catch the worm? Instrumental variable estimates of educational effects of age of school entry in Germany, *IZA Discussion Paper no. 1827.*

[109] Rork, Jonathan C. (2003) Coveting Thy Neighbors' Taxation, *National Tax Journal*, 56(4), pp. 775-787.

[110] Schwarzenberger A. (2008) Public-Private funding of higher education: a social balance, *European Commission under the Socrates programme.*

[111] Walker, I. and Zhu, Y. (2007) The Labour Market Effects of Qualifications with Special Reference to Scotland, *Report* N.1 for the Government of Scotland.

[112] Weiss, Y. (1974) The Risk Element in Occupational and Educational Choices, *The Journal of Political Economy*, Vol. 80, No. 6 (Nov.- Dec.), pp. 1203-1213.

[113] Willis, R. and Rosen, S. (1979), Education and Self Selection, *Journal of Political Economy*, University of Chicago Press, vol. 87(5).

[114] Wilson, J.D. (1999) Theories of Tax Competition, *National Tax Journal* 52, 269-304.

[115] J.M.Wooldridge, (2002), *Econometric Analysis of Cross-Section and Panel Data*, MIT Press.

[116] Yor, Marc (2001), Exponential Functionals of Brownian Motion and related Processes, *Springer Finance*, Germany.