



Faculty of Economics and Business Administration

2004-2005

Economic Growth, Income Inequality and Welfare States

Dissertation

Submitted at Ghent University
to the faculty of Economics and Business Administration
in fulfilment of the requirements for the degree of Doctor in Economics

By

Niko Gobbin

Supervisor: Prof. dr. Glenn Rayp

UNIVER
SITEITS
BIBLIO
THEEK
GENT

Supervisor:

prof. dr. Glenn Rayp

Universiteit Gent

Members of the jury:

prof. dr. Roland Paemeleire

Universiteit Gent

prof. dr. Eddy Omev

Universiteit Gent

dr. Bas van Aarle

Katholieke Universiteit Leuven, LICOS

prof. dr. David de la Croix

Université Catholique de Louvain, IRES

prof. dr. Rafael Doménech

Universidad de Valencia

prof. dr. Gerdie Everaert

Universiteit Gent

prof. dr. Freddy Heylen

Universiteit Gent

prof. dr. Dirk Van de gaer

Universiteit Gent

*"Poor man wanna be rich, rich man wanna be king.
And a king ain't satisfied, till he rules everything"*

Bruce Springsteen, *Badlands*, 1978

Voorwoord

De voorbije zes jaren waren mooie jaren. Ondanks een immens doctoraal obstakel. De charmes van langdurige stress, oneindig repetitief werk en een kantoor buiten het bereik van het fijnste straaltje zonlicht blijven gereserveerd voor de onderzoeker in hart en nieren. Mijn hart en nieren zijn geen exclusief onderzoeksterrein. Voor mij is dit doctoraal proefschrift in de eerste plaats een bewijs van volharding, volgens de volkse wijsheid een oeroude West-Vlaamse deugd.

Dit proefschrift weerspiegelt mijn evolutie van student tot onderzoeker. Een vat vol kennis werd in 1998 de universiteit uitgerold. Net zoals dat het geval is bij een goede Rodenbach, verbeterde een rijpingsproces het eindproduct vele malen. Maar geen goed bier zonder goede brouwers. Mijn dank gaat uit naar Gerdie en Freddy, die alleen maar hop van de beste kwaliteit toelieten. Ook Dirk ben ik zeer erkentelijk. Door zijn inbreng veranderde een eenvoudige Rodenbach in een Grand Cru. De laatste aroma's werden aangebracht door een select comité van fijnproevers: bedankt, Bas; merci, David; gracias Rafael.

Tot slot een woord van dank voor Glenn, de Meester-Brouwer hemzelve. Zoals het een vakman betaamt, hanteert hij een gecompliceerd vakjargon en beheerst hij de finesses van de 'creatieve organisatie'. Timing en planning zijn voor kookboekjes, een meesterwerk verdraagt ze niet.

Doctoreren is veel meer dan het schrijven van een doctoraal proefschrift. Zes jaar universiteit betekent ook (zeer) veel collega's. Enkelen onder hen wil ik speciaal vermelden. Annelies, Stephen, Katrien en Delphine die stuk voor stuk fantastische bureaugenoten waren en uitzonderlijke vrienden gebleven zijn. Liesbet en Martine, veelvuldig slachtoffers van een acute nood aan een ontspannende babbel (of kritische analyse) tussen de wiskundige afleidingen in.

Doctoreren was ook het vertegenwoordigen van de assistenten in de faculteitsraad. Een leerrijke kennismaking met de werking en structuur van de universiteit. Maar soms ook de confrontatie met een gebrek aan solidariteit, verantwoordelijkheidszin en onbaatzuchtig

engagement bij jonge mensen. Doctoreren is echt wel meer dan het schrijven van een doctoraal proefschrift.

Klagen tussen pot en pint over universitaire wantoestanden, een gezamenlijk ongenoegen uiten omtrent de filmische vermassacratie van Tolkiens epos, een vooruitblik naar het komende wielerseizoen, ... Het hoort er natuurlijk ook allemaal bij. Dank aan alle gewillige slachtoffers!

Het proefschrift draag ik aan niemand op. Niet aan mijn ouders, niet aan mijn broers, schoonzussen en nichtjes, niet aan mijn vriendin, Elke. Ze zijn me allen meer waard dan deze saaie wetenschappelijke uiteenzetting.

Ik beëindig het voorwoord op een kritische noot. Sinds enkele jaren waait een krachtige onderzoekswind door de universitaire gebouwen. Wetenschappers moeten internationaal publiceren. Sterker nog: 'internationaal publiceren en niets anders dan internationaal publiceren' wordt meer en meer de slagzin. Maatschappelijke dienstverlening wordt afgestraft door het universitaire bestuur. Hierbij ziet men het belastingsgeld dat onderzoek financiert gemakkelijheidshalve over het hoofd. Even onbegrijpelijk is de houding van beleidsverantwoordelijken die een verdere scheef trekking van de middelen in het voordeel van de 'exacte' wetenschap oogluikend toestaan. En dat alles in een periode waarin belangen Vlaams worden, een bevolking grijzer wordt, ...

Voorwaar, dit zijn vreemde tijden ...

Niko Gobbin
Gent, 1 januari 2005

Contents

Introduction, non-technical summary and conclusions

1. Orientation and motivation	0.1
2. A first glance at inequality, redistribution and growth in the OECD	0.3
3. Outline and main results	0.8
4. Policy implications	0.11
5. Directions for future research	0.15
6. References	0.16

Chapter 1 Economic growth in welfare states: a survey

1. Introduction	1.1
2. Negative effects of the welfare states on economic growth	1.2
3. Negative effects of the welfare states on economic growth	1.5
3.1. Imperfect capital markets and human capital externalities	1.5
3.2. Human capital externalities	1.6
3.3. Risk and the welfare state	1.7
3.4. Socio-political instability	1.10
3.5. Redistribution and innovation	1.11
4. Preferences for inequality and redistribution	1.11
5. Conclusions	1.14
6. References	1.15

Chapter 2 Income inequality data in growth empirics: from cross-sections to time series

1. Introduction	2.1
2. Measuring incomes ... part one: what income?	2.2
3. Measuring incomes ... part one: whose income?	2.7
4. What does one number say?	2.12
5. Half a pound of data	2.18
6. From cross-section to time series: a new secondary dataset	2.22
6.1. Description of the dataset	2.24
6.2. Trends in income inequality	2.26

6.3. Limitations of the dataset	2.29
6.4. Time series characteristics of the data	2.29
7. Conclusions	2.32
8. References	2.33

Chapter 3 *Inequality and growth: does time change anything?*

1. Introduction	3.1
2. The econometrics of inequality and growth	3.2
3. The models	3.3
3.1. Set up	3.6
3.2. Complete markets model	3.8
3.3. Imperfect markets model	3.10
4. The data	3.13
4.1. Income inequality	3.14
4.2. Enrolment rates	3.15
4.3. Economic growth	3.16
4.4. Social security	3.17
5. Cointegration analysis	3.18
5.1. The cointegrating rank	3.19
5.2. Identification of the long run relationships	3.20
5.3. Limitations of the time series approach	3.25
6. Conclusions	3.26
7. References	3.27
Appendix A: Mathematical derivations	3.31
Appendix B: Enrolment rates (data description)	3.34
Appendix C: Graphs	3.37

Chapter 4 *Inequality and growth: from micro theory to macro empirics*

1. Introduction	4.1
2. The models	4.3
2.1. General set up	4.3
2.2. The complete markets model with and without median voter	4.5
2.3. The Imperfect markets model	4.7
3. Simulation set up and calibration	4.8
4. Measures of inequality and poverty	4.13

4.1. Inequality measures	4.14
4.2. Poverty measures	4.15
5. Discussion and results	4.16
5.1. A comment on the use of the log normal distribution	4.16
5.2. The complete markets model without median voter	4.17
5.3. The complete markets model without median voter	4.17
5.4. The imperfect markets model	4.21
5.5. Some further comments	4.24
6. Conclusions	4.26
Appendix A: Mathematical derivations	4.28
Appendix B: Correlations	4.31
Appendix C: Cumulative distribution function	4.32

Introduction, Non-technical Summary and Conclusions

0.1. Orientation and motivation

In the past years an increasing number of economists have started questioning the desirability of an extensive welfare state. Employers' federations, popular media and politicians have picked up their arguments and are increasingly calling for a roll back of the social security safety net. They argue that the expansion of the welfare state has gone too far. European economies are caught in excessive government intervention. Workers are incited to shirk, as they can always fall back on the unemployment benefits. For the same reason, the unemployed have no reason to actively look for a job. Firms are discouraged to innovate as their profits are pruned away by the government. An example of how things should be done, can be found on the other side of the Atlantic. Economic agents in the United States are not impeded by high taxes and social security-induced disincentives. Hence, they are able to swiftly accumulate wealth while European economies are stagnating. Even worse, if acted upon, the call for EU-wide social security standards will export inefficiencies to new EU member states.

If this dismal description is true, why then have welfare states grown that large? Lindert (2004) looks at the history of the welfare state and notices that no country before the end of the eighteenth century had even 3 percent of its national product devoted to redistributive social programmes. In 1930 social transfers amounted to only 2.59% of the national product in Sweden, which is currently the prototype of the extensive welfare state. By the end of the 1990s that figure had grown to over 30%.

Moreover, the general statement that the United States have outpaced all welfare states is not supported by the data. The correlation between the size of the welfare state and economic growth is very weak. Hence, it appears that it is not solely the size of the welfare state that determines its impact on economic growth.

There are some objective reasons for a re-examination of the welfare state. These reasons do not question the welfare state itself, but are linked to its long run feasibility. Social and economic circumstances have changed enormously since social security programmes were first introduced. European integration and globalisation can lead to social security competition between states (Meeusen and Rayp (2000)). If eligibility for social security benefits extends to immigrants, free movement of people within the enlarged EU can turn the

extensive welfare states into the poor houses of Europe (Sinn (2000)). This fear could trigger a social race to the bottom. Recently, a number of politicians of the “old” EU member states have indeed been urging for a more social Europe.

European integration has also directly affected the fiscal freedom of the EU member states. Changes in the welfare state are part of a general pattern of fiscal adjustment since the beginning of the 1990s in the light of the Maastricht criteria and the Stability and Growth Pact (see Gobbin and van Aarle (2001)).

Population aging has raised the costs of retirement pensions and health expenditure. Moreover, demographic evolutions in Europe are such that a decreasing group of workers have to bare the increasing costs. This might undermine intergenerational solidarity and popular support for an extensive social safety net, thereby threatening the survival of the welfare state.

In this dissertation we do not question the need for a broad debate concerning the future of the welfare state. Many different issues will have to be addressed in such a debate. In the dissertation we focus on one of those issues: the trade off between income equality and economic growth. The dissertation bundles 4 contributions that are somehow related to the following question: “Would rolling back the welfare state automatically stimulate economic growth?”. The dissertation contributes to the literature on inequality, redistribution and economic growth in developed countries. It contains a state of the art of both empirical and theoretical work on the subject. Moreover, it presents a thorough examination of data and methodology. We also identify and remedy some weak spots in the literature.

We end this short introduction with 2 things we do *not* do in the dissertation.

We do not assess the optimality of the welfare state. Doing so would require to take a country’s social welfare function into account. The welfare state was not designed to promote growth. Its key goals are poverty and inequality reduction, increasing social cohesion and improving human well being. Judging the welfare state should involve an evaluation of all its objectives and costs. We depart from a very narrow perspective: we only look at the welfare state in growth terms. This does in no way reflect an ideological standpoint from our part.

Contrary to most authors we focus on a small group of developed countries. As we experienced during our research, data quality and country heterogeneity already gravely hamper the analysis for a group of developed countries. Previous studies included developing countries mainly to obtain sufficient data points. However, we seriously doubt the reliability of studies that include e.g., Germany and Benin in the same estimation.

In the next section we define some key concepts and take a first look at the data.

0.2. A first glance at inequality, redistribution and growth in the OECD

The welfare state – In his survey in the Journal of Economic Literature Barr (1992) states that “*defining the welfare state continues to baffle writers and, as with poverty, much high-grade effort has been wasted in the search*” (p. 742). The term unites the state’s activities in health care, education, food, housing, cash benefits, etc. But what about e.g., private insurance for which the state offers tax relief? Lindert (2004) looks at the welfare state from a historical perspective and describes how the functions of the welfare state have expanded from basic poor relief and providing public education to a much broader range. Atkinson (1999) notes that “*the welfare state serves to even out differences in life chances, to achieve greater equity between generations and to redress inequality by race, gender, or health status. More generally, these programs are intended to help people reallocate income over the lifecycle, to insure against events which cause income loss, and to provide a sense of security to all citizens*” (p.5-6).

Since no universally accepted definition of the welfare state exists, it is obvious that a uniform way of quantifying it is also lacking. We look at the public and mandatory private social security expenditure as a percentage of GDP. This is an imperfect proxy (e.g., student loans are not included), but it seems as good as any other candidate proxy. The measure (or a closely related one) is also commonly used in the empirical literature (e.g., Korpi (1985), Persson and Tabellini (1994)).

Table 1 gives an overview of the magnitude of the welfare states in the OECD countries. We divide the countries into 4 categories: ‘Scandinavian’, ‘Continental European’, ‘South European’ and ‘Liberal’ welfare states. The division reflects a common view on the different types of welfare states.

The Scandinavian countries have an extensive welfare state that mainly aims at active measures (e.g., child care, active labour market policy). The continental European countries also maintain a large social security safety net, but spend a larger fraction on passive measures (e.g., unemployment benefits). The evolution in the Southern European economies should be seen in the light of the European integration and conditional

convergence. With the exception of Italy¹, their welfare state is still evolving quite rapidly. The remaining countries have opted for a more limited design of the welfare state.

Table 1: The magnitude of the welfare states (Public and private mandatory social security expenditure as % of GDP, 1995)

'Scandinavian' welfare state	SWE	DEN	FIN	NOR			
	33.39	32.26	31.38	28.56			
'Continental European' welfare state	FRA	BEL	GER	AUT	NET	CH	
	29.29	28.52	28.19	27.24	26.29	25.17	
'Southern European' welfare state	ITA	GRE	ESP	POR			
	25.79	21.28	20.64	17.89			
'Liberal' welfare state	UK	IRE	NWZ	CAN	AUS	US	JAP
	22.9	19.28	19.12	18.21	18.02	16.9	14.25

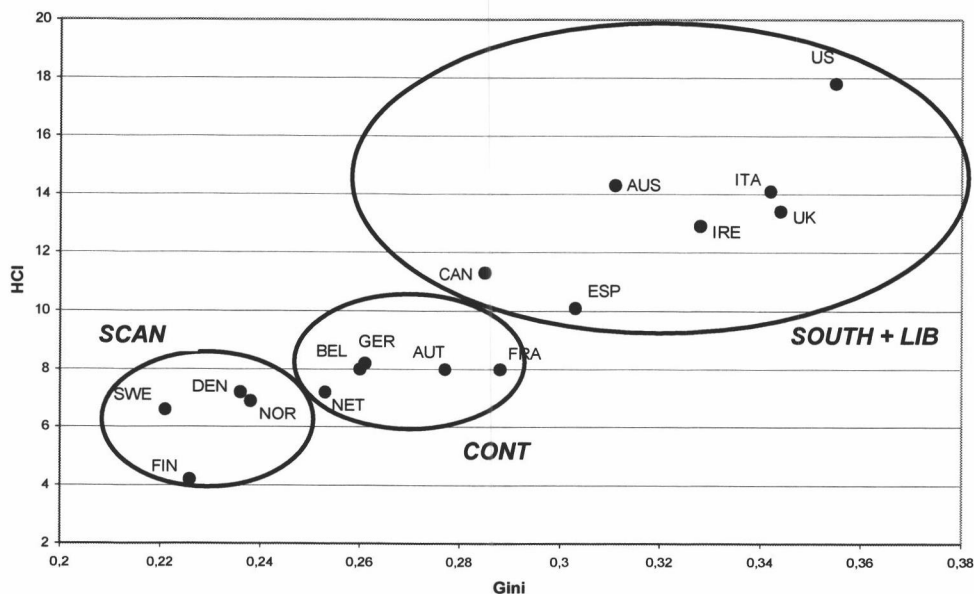
Source: OECD (2003)

Table 1 shows that there exist large differences in the size of the welfare state between OECD countries. The Swedish social security safety net is twice as large as that of the US. In the remaining part of this section we will look both at average data for the different types of welfare states and at the individual data for Sweden and the US, two 'extreme' cases.

Inequality and poverty – Figure 1 plots the gini coefficient of household disposable income against a head count index for a group of OECD countries. The gini coefficient is an inequality measure. The higher its value, the more unequal the underlying income distribution. The head count index is a poverty measure. It is the percentage of people earning less than 50% of the median income. Clearly, inequality and poverty are positively correlated in the included countries.

The Scandinavian welfare states achieve the lowest levels of inequality and poverty. Most Continental European welfare states have a similar poverty level, but their income distribution is a bit more unequal. Poverty and inequality level are highest in the Southern European countries and the liberal welfare states. Hence, the welfare states succeed in reducing inequality and poverty: the countries in the lower – left quadrant of figure 1 are also the countries with the highest share of social security expenditure in GDP (table 1).

¹ The division is not unique. E.g., it would be equally justifiable to treat Italy as a continental European welfare state.

Figure 1: Inequality and poverty in the OECD

Source: LIS (2004), year closest to 1995

The US is 'record-holder' for both inequality and poverty. But we might overestimate inequality in the US relative to the European countries. Next to inequality within European countries, there exists considerable income disparities between countries. Maybe it would be more appropriate to compare German income inequality to that of e.g., Texas.

Papatheodorou and Pavlopoulos (2003) show that more than 92% of overall EU inequality can be attributed to income disparities within the member states. The US Census Bureau (2004) reports gini ratios at the US state level. The data indicate that the state with the lowest inequality (Alaska) has a gini coefficient that equals 87% of the overall value for the US. This is still high relative to continental Europe. The state of New York has the highest inequality (108% of the US value). So overall US inequality seems to be mirrored in the individual states and our initial conclusion remains valid.

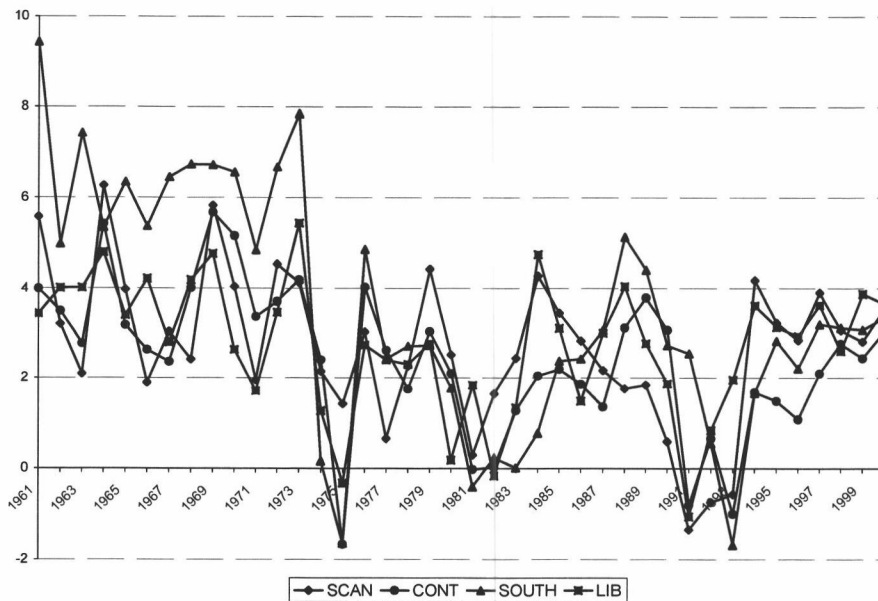
Economic growth – In the dissertation we focus on the growth rate and not on the level of GDP per capita. Clearly if one is interested in welfare, the latter indicator provides relevant information. However, we base our research on the endogenous growth literature and not on the neoclassical 'Solow-type' growth models. In the latter, economic variables have an impact on the equilibrium level of GDP. The growth rate is only temporarily affected during

the transition from the old to the new equilibrium. The former models assume that not the level, but the growth rate of GDP is permanently affected by changes in the explanatory variables.

Figure 2 gives an overview of the average annual growth rate of real GDP per capita in the different groups of welfare states over the period 1961-2000. There are large fluctuations in the growth rates. Since the mid-1970s the growth rates seem to be fluctuating around a lower mean. Using individual countries results in a similar picture.

Only limited lessons can be drawn from the figure. Country-specific growth rates are affected by many factors. E.g., the fiscal consolidations that preceded the start of the European Monetary Union have affected the business cycle in the EMU member countries (see among others Gobbin and van Aarle (2001)). It is more informative to look at average growth rates over a longer period to filter out these business cycle effects.

Figure 2: Economic growth in welfare states (1961-2000) (unweighted average)



Source: OECD (2003)

Notes: SCAN: Denmark, Finland, Norway and Sweden; CONT: Austria, Belgium, Germany, France, the Netherlands and Switzerland; SOUTH: Italy, Greece, Portugal and Spain; LIB: Australia, Canada, Ireland, Japan, New-Zealand, the UK and the US

The first line of table 2 reports the average growth rate of real GDP per capita over the period 1965-2000 for different types of welfare states. Although there are substantial differences between the growth rates, one cannot attribute these differences to the welfare state. The liberal welfare states do not systematically outperform the Scandinavian and Continental European ones.

The welfare state has evolved quite a lot since 1965. As one can see in table 2, social security expenditure as a percentage of GDP has more than doubled in most countries over a 35-year period. In 1965 the average values for the Scandinavian and the liberal welfare states were still very similar. Hence, the picture that emerges in more recent years might be more telling. If we look at the period 1980-2000, we see that the liberal countries have on average recorded the highest growth rates. But the picture remains blurred, especially if we look beyond the average. The Finnish economy grew at an average rate of 2.8% over the period 1980-2000, which is higher than the average growth rate in the US (2.3%).

Table 2: Growth and redistribution in the OECD (1965-2000)

	SCAN	CONT	SOUTH	LIB	SWE	US
Average real per capita GDP growth: 1965-2000	2.5	2.2	3.1	2.6	2.0	2.3
Average real per capita GDP growth: 1965-1980	3.0	3.0	4.4	2.8	2.2	2.4
Average real per capita GDP growth: 1980-2000	2.1	1.6	2.1	2.4	1.7	2.0
Social security expenditure as % of GDP (1965)	11.7	15.0	9.1	9.3	12.9	7.6
Social security expenditure as % of GDP (1995)	31.4	27.5	21.4	18.3	33.4	16.9

Source: OECD (2003)

The inequality-growth trade off – At first sight, the data do not indicate that a large social security system necessarily hampers growth. Over a long period of time, average cross-country growth rates are very similar despite large differences in the size of the welfare state. However, it is impossible to obtain any conclusion about the relationship between inequality, the size of the welfare state and economic growth based on a simple inspection of the data. Growth is determined by multiple variables which have to be taken into account when one is

interested in the effects of one potential determinant alone. This dissertation contains the description of the quest for a more robust conclusion.

0.3. Outline and main results

The dissertation is made up of four chapters. The first chapter presents a survey of the literature on the effects of social insurance and redistribution on growth.

The remaining chapters focus on the empirical literature linking income inequality, redistribution and economic growth. According to most authors (e.g., Perotti (1994, 1996)) the inequality-growth relationship is negative: more equality stimulates growth. However, more recent contributions (e.g., Forbes (2000)) reach the opposite conclusion. We look beyond these conclusions and try to account for the different outcomes. The second chapter looks into the data on income inequality. The third chapter addresses the estimation methodology. The fourth chapter assesses the impact of the choice of scalar inequality measure on the estimation results.

A survey of the literature on welfare states and growth – Most negative effects of the welfare state on economic growth are linked to its cost side. These costs are mainly associated with the allocative distortions of the choice of labour, savings and the composition of assets. In a large class of endogenous growth models, the high tax rate that is needed to maintain a high degree of social insurance affects the growth rate through a reduction of the net rate of return to investment. However, these model usually assume that there is no accumulation of human capital and that there exist no market imperfections. If one allows for both human and physical capital accumulation, the effect of social security on economic growth is indeterminate. Moreover, social security may serve as a second best solution in case of market imperfections. Firstly, private insurance is not always feasible (e.g., insurance against the risk of unemployment). Without insurance, risk averse agents will exchange some expected income for more income security and investment levels will be sub-optimal. Secondly, in case of credit market imperfections, poor people will not be able to obtain an optimal schooling level. On the downside, too much social insurance might lead to moral hazard problems and excessive risk taking.

Existing literature is not always informative since it assumes an unrealistic institutional and economic context. Moreover it seldom motivates why a certain level of redistribution is optimal.

The survey also indicates that it might be difficult to formulate general conclusions concerning the effects of *the* welfare state on economic growth. Even if the size of the

welfare state is comparable between countries, its composition can differ substantially. E.g., a country can spend its resources on unemployment benefits, or it can hand out a similar amount of study grants. Theory predicts that the effect on economic growth will differ. Similarly, a government can tax consumption or it can tax incomes. Again theory does not foretell a uniform impact on growth. Still, we concentrate on the size of the welfare state in the empirical parts of the dissertation for a number of reasons. Firstly, we want to evaluate existing evidence (also based on the size of the welfare state) that detects a significant relationship between inequality and growth. Secondly, there are substitution effects. E.g., if the government increases unemployment benefits, the unemployed will be able to invest more in human capital. Hence, size matters apart from composition. Thirdly, Lindert (2004) argues that large welfare states take more care in developing their tax system, since for them the cost of a wrong design is much higher. So composition and size can be correlated. In that case, the relationship is still present, although the slope becomes less steep. Fourthly, given the multitude of theoretical explanations, the only valid empirical alternative seems to be a case-study approach, which does not allow for generalisations.

The inequality-growth relationship – The nature of the relationship between income inequality and economic growth has inspired a lot of empirical research over the past decade. Initially results indicated that inequality was bad for growth. Later on it was argued that in rich countries there was no significant relationship. More recently, the conclusion is that more inequality stimulates growth. Given this evolution, one might conclude that the world has changed drastically over the past decade thereby disturbing well-established economics relations. However, a look beyond the results rather questions the stability of the results.

Income inequality data – To determine the relationship between income inequality and economic growth, one first has to look for suitable data. This is not a straightforward task. Which income flows matter? How can one represent an entire income distribution in a one-dimensional measure? Economic theory often offers a helping hand, but unfortunately, one cannot always follow its guidelines as reliable income data are scarce. Forced by data shortages, most researchers have focused on cross-section estimation techniques. However, data comparability across countries is not guaranteed if one uses the existing ready-made datasets (e.g., Deininger and Squire (1996)). Mixing inequality data based on different equivalence scales, different income concepts, etc. blurs the outcome of any econometric analysis. Based on the a priori belief that consistency within countries over time is likely to be larger than the consistency of data across countries, we gather income inequality time series for as many OECD countries as possible (chapter 2). We manage to

include 9 countries: Belgium, Canada, Finland, France, Italy, the Netherlands, Sweden, the UK and the US. A first inspection of the data indicates that inequality has evolved very differently in the included countries over the past decades. More robust econometric testing shows that income inequality series are characterised by either a stochastic or a deterministic trend.

The econometrics of inequality and growth – Non-stationarity of the data is only one of the problems complicating empirical work. Existing results are seriously tainted by parameter heterogeneity across countries, omitted variable bias, endogeneity problems, etc. Using the data set we develop in chapter 2, we propose an alternative estimation methodology in chapter 3. As a lot of the methodological problems are related to the cross-section approach, we investigate the relationship between income inequality and economic growth in individual countries in a cointegrated VAR-setting (Johansen and Juselius (1994)). Our approach steers clear of problems related to parameter and theory heterogeneity, and a priori restrictions with respect to stationarity, causality, etc.

We use this new approach to transpose the cross-section estimates of Perotti (1996) to a time series framework. We discriminate between two models of growth and inequality in 9 OECD countries. The imperfect markets model (Galor and Zeira (1993)) states that poor people cannot invest sufficiently in their education due to credit constraints. Redistribution stimulates growth by relaxing these constraints and facilitating human capital accumulation. The complete markets model (Persson and Tabellini (1994)) argues that more inequality will induce a higher preferred tax rate by the median voter. Since taxes hurt growth, inequality is negatively related to the growth rate. Although both the imperfect markets model (Belgium, Canada, the US) and the complete markets model (the Netherlands) find some support in the data, the former seems to be more relevant for our sample of OECD countries. We always find that social security is negatively related to growth, once we control for human capital investment. A conclusion in line with the implications of the imperfect markets model.

An indirect conclusion is that different models hold for different countries. This questions the appropriateness of cross-section and panel estimates.

Income inequality measures – A weak spot in the link between theory and empirics is that we cannot include an entire income distribution in a standard growth regression. We need to 'summarize' the distribution in a scalar measure of inequality. This cannot be done without a loss of information. Many inequality measures can be constructed (e.g., a gini coefficient, a quantile ratio, an Atkinson index, etc.). Since different measures focus on different aspects of the income distribution, the correlation between the measures will be limited. Moreover, they order income vectors differently. Hence, the choice of measure might matter for the

estimates. Again theory can offer some clues: e.g., if credit constraints matter, measures that are sensitive to changes in the bottom part of the income distribution should be preferred. In practice researchers have to settle with the available data, which means that, in general, gini coefficients and quantile ratios are used.

In chapter 4 we assess the robustness of empirical results to a change in the inequality measure. Since real-life growth regressions are simultaneously infected by a lot of data and econometric problems, we resort to a simulation approach. The advantage of this approach is that we are able to isolate the impact of the chosen measure. Our simulation results support the current practice of using the gini coefficient in case of data shortages, but should also encourage researchers to construct other inequality indices. While the gini coefficient is able to recover the relationship between inequality and growth, to discriminate between alternative theories other measures can be helpful.

These results strengthen the conclusions of chapter 3. Moreover, they should act as a stimulus for further macro-econometric research with respect to the inequality-growth relationship. Given a lack of data, some authors recently expressed doubts about the usefulness of this type of analysis. Our results indicate that one can empirically assess the validity of growth theories as long as econometric shortcomings are overcome.

0.4. Policy implications

In this section we present the main policy implications of the results outlined in the previous section. Since a large fraction of the dissertation was devoted to methodological issues, the number of policy implications remains limited. Our results are useful to assess the validity of existing work and offer guidelines for future research. The dissertation will contribute to more reliable empirical work, and hence, more trustworthy policy advice in the future.

An important conclusion of the dissertation is that one should steer clear of general statements concerning the effects of the welfare state. Our empirical analysis indicates that, even in a group of 'similar' countries, the relationship between income inequality, redistribution and economic growth can be quite different. The literature suggests that the welfare state can affect economic growth through many channels. The size of the effect is heavily influenced by a country's economic and institutional context. Also the way in which a government finances its social expenditure matters.

On the hand, the results can reassure policy makers favouring an extensive social safety net. Theory does not predict that social security expenditure necessarily hurts economic growth, nor does empirical work lead to that conclusion. Our analysis provides a strong warning

against mechanical use of the existing empirical 'evidence'. Given the huge list of problems disrupting empirical work, a fair amount of scepticism is warranted.

The results neither predict that a roll back of the welfare state is a wise strategy to follow. The outcome of such an operation is uncertain since e.g., credit constraints could be worsened by curtailing social security expenditure.

On the other hand, our results contain a warning for policy makers. There is no 'roadmap to equality'. Policy makers cannot mimic success stories. They should be very prudent in the design of the social security system in order to minimize the efficiency loss.

Our estimates indicate that in most OECD countries, inequality negatively affects investment in human capital. Social security expenditure are only negatively related to economic growth after human capital investment has been taken into account. Hence, even in 'rich' countries with well developed capital markets, government support for human capital investment is sensible.

But our empirical results also point out that different inequality-growth models might apply in different countries. For a policy maker it is important to unravel the mechanisms behind the inequality-growth relationship. *If* a government seeks to stimulate growth by means of its social security expenditure, it has to identify the channel through which inequality and redistribution matter for growth. E.g., since credit constraints seem to matter, additional measures should especially be directed towards the poor. We emphasized the "if" in the above statement because we do not imply that the welfare state should primarily be used to support growth. However, popular support for redistribution will depend on its cost side. Therefore, a government needs to carefully assess the efficiency loss caused by its policy measures.

In order to better understand the inequality-growth relationship, governments and international organisations (such as the OECD) should invest more resources in the development of consistent data sets. As in many fields of econometric research, the data shortage is a key problem. Although improvements have been made (e.g., the LIS project), a lot remains to be done. Given the importance of the welfare state in a lot of European countries and the urgent calls for its reform, it is unfortunate that the impact of the welfare is still not well understood due to data problems. Our results indicate that one does not need the "theoretically perfect data series" to obtain reliable results. However, one does need consistent data.

We should not end this section without addressing the most pressing challenge of the welfare state: the aging population. While we did not explicitly cover the subject in the

dissertation, the analysis offers some insights into this issue. Firstly, the actual cost of population aging might exceed the direct costs related to increasing health care expenditure and pension payments. Social security resources will be transferred from domains that are efficiency enhancing (e.g., educational credit constraints) towards domains that are not or less favourable for growth (e.g., pensions). Secondly, rolling back the welfare state will not necessarily mitigate the problem. The impact of a cutback will again depend on the types of expenditure that are reduced. Population aging decreases the government's freedom of movement since the relative weight of the 'unfavourable' categories increases. Rolling back the welfare state might especially affect those expenditure categories that compensate for the negative effects of social security on economic growth. If that is the case, a smaller welfare state will also imply lower growth.

On the other hand, the problems caused by population aging are too serious to be ignored. Policy makers need to formulate an adequate response. Since rolling back the welfare state offers no miracle solution, we believe that policy makers should step up their efforts to increase the participation rate.

0.5. Directions for future research

The debate concerning the future of the welfare state has only just begun. This dissertation can inspire the debate, but does not offer clear-cut answers. A lot remains to be done. In the dissertation we identified and cured a number of methodological problems. The results should stimulate future research.

We end this introduction by composing an (non-exhaustive) agenda for future research.

This dissertation mainly addresses the empirical analysis of inequality, redistribution and growth. Although we present a survey of the theory and identify a number of shortcomings in it, we do not bridge the gaps in the literature. An important finding of chapter 1 is that, in existing work, the link between reality and theory is too loose. By making theory more compatible with the European institutional framework and economic reality, one should be able to obtain more policy-relevant conclusions.

A first attempt (not included in the dissertation), in which we introduce efficiency wages in combination with imperfect competition on the goods and labour markets, seems promising. The combination of efficiency wages and trade unions causes a trade off between the positive and negative effects of redistribution. Efficiency wages imply unemployment in equilibrium and hence, cause a deterministic, 'social' risk. In most existing contributions, redistribution is instead motivated by the existence of a stochastic risk (e.g., agents differ in

endowment). Similar to those contributions, we assume that no private insurance covering the unemployment risk exists. The introduction of a social security system to cure this market imperfection has a spill over effect on the labour market. It strengthens the outside option of the trade unions, inducing higher wages. Social security has two counteracting effects. On the one hand, by improving the outside option it increases the expected income of workers, thereby increasing workers' investment. On the other hand, the increase in the worker's expected income is at the expense of entrepreneurs' profits and investment. An interesting implication of the model is that the choice for 'no social security' does not maximize growth. However, the model has not yet been explored to its full. Developing it further will be an important part of future research.

We started our work with an exploration of the data. The data set we assembled, aimed at full consistency. Still we cannot claim to have fully achieved this goal. To further improve the empirical part of this dissertation, it would be useful to resume the search. However, this will be a time-consuming task for which we need the help of "inequality specialists" in different countries.

The survey of the theory shows that the relationship between inequality, redistribution and growth depends on many factors. Our empirical analysis indicates that parameters and even models might differ across countries. The 'optimal' proxies for the variables in the estimations are also country-specific. There exist large differences in the schooling system, the design of the social security system, the tax design, etc. across countries. Hence, the performance of a proxy also differs across countries. On the one hand, the theoretical complexity and empirical heterogeneity strengthens our individual country approach. On the other hand, it probably means that we did not stretch it far enough. It might be interesting to pursue a thorough case-study approach for the countries in our data set. Based on the results we could fine-tune the estimation specification in chapter 3. Still, fully reconciling econometric work with the multitude of theories is probably unattainable.

Next to the case-study approach, we can already perform a number of robustness checks on the results in chapter 3. Following Kneller, Bleaney and Gemmell (1999), we could take the government's intertemporal budget restriction into account in the estimation of growth equations. Their results show that not social security expenditure, but distortionary taxation negatively affects growth.

In chapter 4 we use the US income distribution for all countries in the simulation. Optimally we would like to use real life income distributions from different countries to better match reality. Again data availability is the main obstacle.

0.6. References

- Barr N. (1992): *Economic Theory and the Welfare State: A Survey and Interpretation*, Journal of Economic Literature, 30, p.741-803
- Deininger K., Squire L. (1996): *Measuring Income Inequality: a New Data-Base*, World Bank Economic Review, 10 (3), p.565-591
- Forbes K.J. (2000): *A Reassessment of the Relationship Between Inequality and Growth*, The American Economic Review, 90 (4), p.869-887
- Galor O., Zeira J. (1993): *Income Distribution and Macroeconomics*, Review of Economic Studies, 60, p.35-52
- Gobbin N. , van Aarle B. (2001): *Fiscal adjustments and their effects during the transition to the EMU*, Public Choice, 109 (3-4), p.269-299
- Johansen S., Juselius K. (1994): *Identification of the long-run and the short-run Structure. An Application to the ISLM Model*, Journal of Econometrics, 63, p.7-36
- Kneller R., Bleaney M.F., Gemmell N. (1999): *Fiscal Policy and Growth: Evidence from OECD Countries*, Journal of Public Economics, 74, p.171-190
- Korpi (1985): *Economic Growth and the Welfare System: Leaky Bucket or Irrigation System?*, European Sociological Review, 1, p. 97-118
- Lindert P.H. (2004): *Growing Public*, Cambridge University Press, Cambridge
- Luxembourg Income Statistics (2004): *LIS Key Figures*, <URL: <http://www.lisproject.org/keyfigures>>
- Meeusen and Rayp (2000): *Socialezekerheidsconcurrentie in de EMU*, in: Ooghe H., Heylen F. and Vander Vennet R. (eds.): EMU: dé uitdaging, Gent, p.325-352
- OECD (2003): *Statistical Compendium on CD-Rom 2003/2*
- Papatheodorou C., Pavlopoulos D. (2003): *Accounting for inequality in the EU: Income Disparities Between and Within Member States and Overall Income Inequality*, CHER Working Paper no. 9, 25p.
- Perotti R. (1994): *Income Distribution and Investment*, European Economic Review, 38, p.827-835
- Perotti R. (1996): *Growth, Income Distribution and Democracy: What the Data Say*, Journal of Economic Growth, 1 (2), p.149-187
- Persson T., Tabellini G. (1994): *Is Inequality Harmful for Growth*, American Economic Review, 84 (3), p.600-621
- Sinn H.W. (2000): *EU Enlargement and the Future of the Welfare State*, CESifo Working Paper, no. 307, 14p.
- United States Census Bureau (USBC) (2004): *Gini Ratios by State*, <URL: www.census.gov/hhes/income/histinc/state/state4.html>

Chapter 1:

Economic Growth in Welfare States: A Survey

Abstract

In this paper we present a critical survey of the literature on redistribution and economic growth. Theory does not predict that a larger welfare state necessarily hurts growth. Potential costs of the welfare states might be neutralised by equal gains.

The survey also stresses that it might be risky to talk about '*the*' welfare state. The way in which redistribution affects economic activity depends on many factors, such as eligibility criteria, tax design, institutions and economic shocks.

JEL Classification: E62, I38, H53

1.1. Introduction

During the past few years calls to roll back the welfare state have resounded louder and louder. Critics argue that the too extensive welfare programs in some European countries suffocate the economy. The United States serve as the textbook example of how things should be done. As long as social spending is not reduced, the strong American economic performance is out of reach for the most part of Europe.

But how convincing are the arguments of these critics? Firstly, judging the welfare state solely in terms of economic growth neglects other important merits. The welfare state was not intended to serve as a growth-enhancing instrument. Hence, the question should be whether its economic cost is acceptable in terms of the gains in well-being. Secondly, the data do not seem to support the existence of a trade-off between redistribution and growth. Atkinson (1999, chapter 2) presents a survey of the empirical work. Although the presented evidence is not unanimous, the overall picture is that the extent of the welfare state and the rate of economic growth are at best (or at worst) weakly negatively correlated. Countries differ substantially in the size of the welfare state, but the differences in economic growth are much less pronounced.

Even without clear empirical evidence in support of the 'high economic cost'-hypothesis, it is important for a government to know how the welfare states operates. A good knowledge of its working will increase the effectiveness of policy measures. A government bears responsibility for the way it spends tax resources. Popular support for an extensive welfare state, and thus, high tax rates, will depend on the reliability of policy measures. If a government cannot give account of its actions, solidarity in the society will be undermined and support for the welfare state will decline. Hence, sustainability of the welfare state implies a proper knowledge of how it interferes with economic live.

In this paper we present a critical survey of the literature on redistribution and economic growth. Does theory predict that a larger welfare state necessarily hurts growth, or do counteracting forces exist? We look at the potential negative impact of redistribution on economic growth (section 1.2) and at channels through which redistribution might stimulate economic activity (section 1.3). We also look for explanations for the differences in inequality aversion between countries (section 1.4). Next to a review of the theoretical literature, we also look for empirical support.

1.2. Negative effects of the welfare state on economic growth: a survey of the literature

There exists a large literature documenting the cost side of the welfare state. Firstly, there are obvious administrative and operative costs: a larger government uses more resources to organise its operations. Secondly, government resources might be wasted if political behaviour is mainly motivated by self-interest (Mueller (1976)).

Thirdly, there are dead-weight costs created by policy-induced changes in relative prices. More specifically, these costs are created by the wedges that are driven between the firms' production costs and the net return of individuals on additional productive effort. These costs are mainly associated with the allocative distortions of the choice of hours of work, saving and the composition of assets. E.g., a pay-as-you-go social security system discourages individuals from saving privately for old age, without compensating for this via forced public savings. It also induces earlier retirement (Feldstein (1985)). In a Solow growth model a reduced saving rate does not affect steady state growth since only the income level depends on the aggregate capital stock. But a negative impact of reduced savings on the growth rate is predicted by many endogenous growth models (e.g., Rebello (1991)).

To finance an extensive welfare state, a government has to impose high taxes. The negative impact of high (average) tax rates on economic growth is again supported by a large class of endogenous growth models (Romer (1986, 1989), King and Rebelo (1990), Rebelo (1991)). In these models taxes affect the rate of growth through a direct channel: they reduce the net rate of return to private investment, making investment activities less attractive.

Once one accepts that economic growth is not solely due to physical capital accumulation, but can also originate from human capital accumulation (Lucas (1988)), the relationship between redistribution and growth can appear in a different light. Feldstein (1985) shows that if there is only accumulation of physical capital, then a pay-as-you-go system of social security or large benefits are never optimal. But if one allows for human capital accumulation the outcome is indeterminate (see also Kemnitz and Wigger (2000), cf. infra). Whether social benefits are still harmful to growth will depend on the importance of human capital accumulation versus physical capital accumulation.

Fourthly, an important class of potential problems with a too extensive welfare state is linked to moral hazard (Pauly (1974)) and disincentive problems. These problems arise where the insured person can influence the expected loss at a cost lower than the expected gain without the insurer's knowledge. E.g., once the unemployed are guaranteed a minimum income, they do no longer need a job to survive and might alter their behaviour increasing

the probability of remaining unemployed. For the government it is hard to monitor the efforts undertaken by the unemployed to find a job. Also the employed might be tempted to shirk, as the unemployment benefit offers a back-up. Similarly, health care might encourage an unhealthy way of life.

Based on the Swedish experiences, Lindbeck (1988, 1993) registers a number of disincentive effects created by marginal tax rates on labour income: (1) against hours of work in the regular labour market, (2) against intensity of work ('on-the-job-leisure'), (3) against human capital accumulation in case of progressive taxation. Moreover it creates 'wrong' incentives in favour of (1) the pursuit of do-it-yourself work, (2) production for barter, (3) the choice of occupation with relatively large non-pecuniary benefits and (4) tax avoidance.

Lindert (2004) challenges the common belief that the welfare state negatively affects incentives. First he looks at the revenue side. What matters for incentives is not the average tax rate, but the marginal one. Marginal tax rates on capital are not higher in most welfare states than in the US. Moreover, European welfare states have taxed labour incomes more heavily. Since capital is more internationally mobile, taxing labour hurts growth less. Blundell *et al.* (1998) show that the wage elasticity of labour supply is small. Welfare states have taxed consumption more than the US. More specifically tax rates on addictive and harmful goods (cigarettes, alcohol, gasoline) are much higher. Consumption taxes do not influence savings (if you consume now or later makes no difference, but if you save your earnings for future production you are taxed twice). Table 1 illustrates the differences in tax mixture in Europe and the US.

Table 1: Tax mixture in Europe (selected countries) and the US

	<i>BEL</i>	<i>ESP</i>	<i>FRA</i>	<i>GER</i>	<i>SWE</i>	<i>UK</i>	<i>US</i>
Top marginal tax rate on dividends (1998)	15.0	47.6	61.0	55.9	30.0	40.0	46.6
Average effective tax rate on labour income (1991-1997)	39.7	30.4	40.2	35.9	48.5	21.0	22.6
Average effective tax rate on consumption (1991-1997)	18.7	13.7	18.0	15.8	18.7	16.9	6.1
Taxes per pack of 20 cigarettes (in US\$, 1997)	2.4	0.9	2.7	2.3	3.1	4.2	0.7

Source: Lindert (2004)

European welfare states have designed a more pro-growth tax system. The higher the social budget as a share of GDP, the higher and more visible the costs of a bad choice. Thus the welfare states seem to have paid closer attention to the productivity consequences of the tax system. This conclusion also entails that the US could stimulate growth by rationalising its tax design.

Daveri and Tabellini (1997) state that labour taxes can be as distorting and harmful to growth as capital taxes. Higher tax rates on labour are shifted onto higher real gross wages if workers are organized in monopolistic unions. Firms will substitute capital for labour and the marginal product of capital falls, which in turn diminishes the incentive to accumulate (i.e. to grow). The difference between Europe and the US is reflected in the behaviour of the capital-labour ratio: between 1970-1995 this ratio more than doubled in the European Union, whereas it only rose by 25% in the US.

Atkinson (1999) argues that the costs of the welfare state have been greatly exaggerated because the benchmark in most theoretical contributions gravely oversimplifies reality. The analysis of the welfare state tends to ignore its institutional structure. E.g., it is often assumed that the only relevant condition for the receipt of benefit is being unemployed. However, typically an unemployment benefit is subject to contribution conditions, is paid only for a limited duration, and is conditional on making genuine efforts to seek employment. This has major implications for the standard theory: a person cannot refuse a job because he is satisfied with his unemployment benefit. Models that neglect the institutional context overestimate disincentive effects.

Van der Ploeg (2003, 2004) develops a related reasoning. In economies with competitive labour markets social policies harm employment and output. However, Europe does not match this description. Non-competitive labour markets with trade unions, efficiency wages and costly search and mismatch are more realistic. He argues that a more progressive tax system, high conditional unemployment benefits and facilitating corporatism can boost employment and output in case of uncompetitive labour markets.

Ljungqvist and Sargent (1995) conclude that progressive taxes and a high marginal tax rate can reduce unemployment. Since such a tax regime lowers the reward for a successful job search, workers are less likely to voluntarily quit their job. In addition, the unemployed will have a lower reservation wage (i.e. they are less likely to refuse a job offer), since the dispersion of after-tax wages decreases. Note that the lower unemployment rate comes at the cost of a less efficient labour allocation (the job search effort is not optimal). Hence, average labour productivity is lowered and the net impact on growth is unclear.

Numerical simulations by Boháček (2003) show that efficiency falls monotonically with the level of social insurance. The efficiency loss is small for low and medium levels of social insurance, but grows rapidly in the case of more generous insurance regimes.

Sinn (1995) argues that the net costs might be related to the extent of the welfare state. The marginal costs of redistribution rise and the marginal returns of redistribution drop if the welfare state gets bigger and bigger. Sinn (1995) attributes the large costs in the overdrawn welfare state to excessive risk taking due to moral hazard (cf. *infra*).

Lindbeck (1997) argues that in the long run the trade-off between redistribution and growth is quite steep. In 1970, Sweden had an income per head 15% higher than the OECD average. By 1995 it was 5% lower.

However, Danziger *et al.* (1981) present a survey of the literature and conclude that empirical results are far from unanimous.

1.3. Positive effects of the welfare state on economic growth: a survey of the literature

1.3.1. Imperfect capital markets and human capital accumulation

Galor and Zeira (1993) emphasize the link between borrowing constraints, the distribution of income and wealth and the aggregate investment in human capital. If people cannot borrow freely against future income, the initial distribution of resources will determine the aggregate stock of human capital. Investment in physical capital will also be affected. But since human capital is a very poor collateral, investment in human capital will suffer most. If one further assumes that there are decreasing returns to investment in education (Bils and Klenow (2000)), then the more equal the initial distribution is, the larger the aggregate stock of human capital will be. Moreover, if human capital and growth are positively related, equality will stimulate growth. In such a context more redistribution will stimulate growth.

Rillaers (2001) presents a model in which unemployment benefits relax the borrowing limit, allowing for higher levels of investment. This counteracts the negative impact of a guaranteed minimum income on the levels of effort.

Cameron and Taber (2004) find no evidence of educational borrowing constraints in the US but add that this does not imply that they would not exist in the absence of private and government programs currently available. Carneiro and Heckman (2002) find that only a very marginal fraction of the US population is credit constrained. They also mention that this is in part due to the successful operation of policies that were designed to eliminate such

constraints. Blöndal *et al.* (2002) argue that in the absence of government intervention, investment in human capital would be difficult to finance through unsecured personal loans in the OECD. Black *et al.* (1996) show that the availability of collateral is of major influence to obtain a loan in the UK.

Azariadis and de la Croix (2003) develop a model in which physical and human capital need to be developed jointly. In such a set up the poor might invest less in human capital, even without being constrained.

There is an extensive theoretical literature with respect to the beneficial aspects of human capital accumulation on economic growth. Pioneering contributions to the literature are Nelson and Phelps (1996) and Lucas (1988). Whereas the theoretical models almost unanimously predict a positive impact of human capital on economic growth, empirical work is more divided on the matter. There exists doubts about the fact that growth is positively related to human capital in a sample of OECD countries (Islam (1995), Barro (2001)). However, de la Fuente and Doménech (2002) show that once data quality issues are controlled for, average schooling indicators across 21 OECD countries are positively and significantly related to economic growth. Sianesi and Van Reenen (2003) present a survey of the macroeconomic returns to education. They conclude that *"taking the studies as a whole, there is compelling evidence that human capital increases productivity"* (p.195).

1.3.2. Human capital externalities

Kemnitz and Wigger (2000) show that if human capital accumulation is the driving force of economic growth, a properly designed unfunded social security system can stimulate growth. They consider a model in which there are intergenerational spillovers of human capital. Spending time for study increases both one's own human capital and that of succeeding generations. However, an individual has no incentive to take this intergenerational spillover into account, because only aggregate schooling matters. Hence, investment in human capital is too low. A social security system that internalises the intergenerational spillovers of human capital (e.g., pensions dependent on human capital accumulation) will enhance growth.

Sinn (1998) argues that parents will invest too little in the human capital of their children due to intrafamily moral hazard problems. Once the children are educated, they may refuse to compensate their parents for educational investments. A pay-as-you-go pension system will mitigate this problem, as it gives parents a stake in the income of their offspring.

Azariadis and Drazen (1990) argue that credit-rationed people will supply more labour and invest less in human capital in their youth. Due to externalities this reduced human capital investment will lower the growth rate of the economy.

One problem with the above theories is the lack of empirical support for the prevalence of human capital externalities. Both Rauch (1993) and Acemoglu and Angrist (2000) find little evidence for large human capital externalities. The modest externalities that are consistent with the results can only explain a marginal part of cross-country growth differences. However, the 'disappointing' results might stem from problems of measurement and shortcomings in the estimation methodology (see Sianesi and Van Reenen (2003)).

1.3.3. Risk and the welfare state

Insurance is an important function of the welfare state. Private insurance is unable to cover all social risks due to problems of adverse selection, moral hazard, interdependent and large probabilities, etc. (see Barr (1992)). E.g., applicants for unemployment benefits can conceal their risk (adverse selection). Once insured, they might alter their behaviour, thereby influencing the probability of entering and leaving unemployment (moral hazard). For some categories, e.g. low skilled, the risk of unemployment might be too high to insure. Although some of these problems can be (partially) dealt with, moral hazard seems inescapable. Given these information problems, income support for all unemployed will only arise if it is publicly provided. Similar problems arise in the case of pensions (unexpected inflation hits everybody at the same time) and health care (individuals with full insurance might choose an unhealthy way of life).

Sinn (1995) notes that timing is an important problem: before adulthood agents cannot settle private contracts. But the younger the person, the larger the uncertainty he faces and the greater the need for insurance. Adverse selection becomes a convincing allocative argument for social insurance if it covers a longer time span than private insurance. Similarly parents with young children do not know their innate abilities and do not know what will happen to them in the future (illness, unemployment, etc.). The welfare state does not eliminate the risks, but it helps mitigate the consequences by redistributing between the lucky and the unlucky.

Redistributive taxation can be efficiency enhancing, because it creates safety and stimulates income generating risk taking. If people are risk averse, the amount of risk taking is suboptimal and marginal returns to risk are positive. Social insurance makes people more

daring since the government takes an equal share in the gains or losses following their economic decision. Under the protection of the welfare state people can avoid costly private protection measures (e.g., offspring maximization, precautionary saving) and they can dare to change jobs, to seek employment in risky industries, to become self-employed, to specialize in a rare profession. On the other hand, it may make people careless (Sinn (1994, 1995)) leading to excessive risk taking.

Bird and Hagstrom (1999) find that redistribution reduces precautionary savings (see also Huggett and Ospina (2001)). Individuals need less savings to hedge their bets, allowing them to undertake larger economic gambles. Bird (2001) finds empirical support for the hypothesis that risk taking is positively correlated with the share of social spending in GDP.

Grossmann (2003) presents a model in which the poor invest less in human capital than the rich due to risk aversion. Human capital accumulation is uncertain and the risk is uninsurable. Krebs (2003) develops a model in which households invest in risk-free physical capital and risky human capital. Investment in human capital is risky due to idiosyncratic shocks to the stock of human capital (job termination, disability, etc.). His results indicate that the gains from labour insurance are of the same order of magnitude as the costs of distortionary taxation and justify the introduction of government-sponsored severance payments.

Brunello (2002) finds that the selected years of schooling decrease when absolute risk aversion increases. Checchi and Garcia-Peñalosa (2004) provide theoretical and empirical support for the hypothesis that greater aggregate production risk decreases the average level of education. Boone (2004) analyses the effects of unemployment insurance on workers' job-to-job mobility. Unemployment insurance can stimulate risk-averse agents to move to risky sectors (i.e. sectors with a higher probability of becoming unemployed) with a higher expected marginal rate of return.

Note that if a social safety net induces more risk taking, the pre-tax and pre-redistribution income distribution should become more unequal. This is, however, hard to deduce from the data. The positive effect of the welfare state on pre-redistribution income inequality through risk taking, might be counteracted by a negative effect through union wage bargaining. The larger the welfare state, e.g., the higher the unemployment benefit, the stronger the bargaining position of unions. Palokangas (1996) presents a model in which more union power, i.e. a better outside option, results in higher wages for the low skilled. This leads to higher unemployment for the low skilled but also to lower wages for the high skilled. Hence, the presence of a social safety net is sufficient to alter the pre-redistribution income distribution. The effect of these changes on inequality is dubious.

Kreickemeier (2004) argues that the higher unemployment benefits create more involuntary unemployment. Since the probability of unemployment is higher for the low skilled, the incentive to get educated increases with the unemployment benefit. But as more people get educated wage differences between workers are levelled off. A similar argument is developed by Cahuc and Michel (1996) for the case of minimum wages. A higher minimum wage increases the probability of unemployment and the incentives to obtain skills for the low skilled. If there exist positive externalities with respect to human capital accumulation (Backus *et al.* (1992)) the higher proportion of skilled workers will stimulate growth. Efficiency wages models (Shapiro and Stiglitz (1984), Summers (1988)) lead to the opposite result. Firms have to pay higher efficiency wages since a fixed net-of-tax distribution has to be maintained to discourage shirking.

Agell and Lommerud (1993) contribute the early success of the Scandinavian welfare states to egalitarian wage policies. Without such policies firms in new, risky sectors need to pay high wages to attract workers from old stagnating sectors. In the new sectors there can exist socially increasing returns and positive externalities from producing. By increasing the wage bill of expansive firms, the high wage premia impose a drag on the rate of structural change. Because of strong trade unions, the wages differences between sectors are limited, forcing people out of the traditional sectors into the modern one. Without redistribution structural change would be suboptimal. In this model redistribution forces people to take risks.

De Grauwe and Polan (2003) provide empirical support for the hypothesis that social spending increases the efficiency of the private sector. Given these positive externalities of social spending, larger social services can offer protection against increasing externally generated risks. There is less need for a costly reallocation of resources in the face of globalisation. Hence, the welfare state protects efficiency instead of hurting it.

The link between welfare states and externally generated risks, such as the globalisation of the world economy, the diffusion of new information technologies and the substitution of the manufacturing by the service sector, has been the subject of a number of contributions. Contrary to De Grauwe and Polan (2003) these studies assume a more negative stance towards the welfare state. Bertola and Ichino (1995) argue that increased instability due to external shocks decreases labour demand by firms. Given rigid wages and high firing costs this will especially be the case in large welfare states. Bertola and Ichino (1995) assume homogeneous workers and look only at labour demand. Ljungqvist and Sargent (1998) introduce heterogeneous workers and also consider labour supply. Workers lose skills if they become (remain) unemployed, i.e. there is a gradual depreciation of human capital. External shocks result in a sudden loss of skills, which lowers the firms' wage offers. In

combination with high benefits (which are related to previous wages and consequently, to previous skills) these lower wage offers reduce the search intensities of the unemployed and raise their reservation wages. A welfare regime can be perfectly sustainable under tranquil economic conditions, but become infeasible when a shock occurs. Rillaers (2001) shows that a skill biased technological shock that increases the productivity of the high skilled workers can counteract the beneficial effects of a social security system in terms of increased human capital investment (cf. supra). Den Haan *et al.* (2001) argue that although the search intensities of the unemployed are affected, economic turbulent times can lead to lower unemployment as high skilled workers are less likely to leave their job. Ljungqvist and Sargent (1998) assumed that the probability of resignation was exogenously determined and independent of economic turbulence. Marimon and Zilibotti (1999) impose heterogeneity with respect to both firms and workers. The productivity of a worker depends on the match between his skills and the job requirements. Unemployment benefit serves as a search subsidy to find the 'right' job. An external shock widens the productivity gap between the best and the worst job. If the unemployment benefit is too high, the unemployed might become too selective and reject socially acceptable matches. Note that unemployment in a laissez-faire economy is reduced at the expense of low labour productivity.

Empirical evidence concerning the interaction between shocks and institutions is provided by Blanchard and Wolfers (2000). They show that this interaction can account for the different evolution of the US and European economies since 1960.

1.3.4. *Socio-political instability*

Sala-i-Martin (1997) argues that social safety net mechanisms are a means to buy social peace, a way to reduce social unrest. The poor are bribed out of activities that are socially harmful, such as crimes, revolutions, riots, etc. Transfers provide an incentive to stay away from criminal activities by increasing the level of income outside jail. This explanation is closely related to the socio-political instability approach of, among others, Alesina and Perotti (1996), according to which a highly unequal, polarised distribution of resources creates strong incentives to pursue one's interests outside the normal market activities or the usual channels of political representation.

1.3.5. Redistribution and innovation

Within the framework of a Schumpeterian growth model, Zweimüller (2000) illustrates that inequality, redistribution and innovation are linked. Consumers with different levels of income do not only spend different total amounts on consumption goods, they also buy different goods. E.g., poor consumers will buy less and different (lower quality) goods than rich ones. As a consequence the market for various (new) products depends on the income distribution. If innovations are an important source of technological progress, the income distribution will determine productivity growth. E.g., if one assumes that consumers have hierarchic preferences (first basic goods, then conveniences and luxury goods), then innovators will attract buyers that are sufficiently rich. Hence, there is a trade-off between prices and market size. Either a small class of very rich people must have a very high willingness to pay for new products, or a large well-funded middle class guarantees sufficient market size. Redistribution of the very rich to the middle class will stimulate growth as the incentive to innovate increases (larger expected market size and future profits). However, redistribution among the richest has no effect, and redistribution from the richest to the poorest could even hurt innovation (lower willingness to pay).

Foellmi and Zweimüller (2003) observe that a lower degree of inequality is linked to a higher rate of productivity growth in Europe in comparison with the US. The authors suggest that the model by Zweimüller (2000) can account for this observation.

Saint-Paul (1997) presents a model in which economies with more rigid labour markets specialise in secondary innovation (i.e. improvement of existing products). The key idea is that countries with rigid labour markets tend to produce goods with a stable demand. Such 'stable' goods are mainly well established goods in a later stage of the product life cycle. The possibility that new products disappear is much higher. If hiring and firing of workers is costly, firms will prefer safe innovations. Innovation in welfare states will only be affected by this mechanism if the flexibility of the labour markets is reduced by an increase in the size of the welfare state. Increased trade union power is a potential explanation for such a relationship (cf. supra).

1.4. Preferences for inequality and redistribution

A straightforward explanation for an observed difference in the extent of the welfare state between two countries could be that their inhabitants value inequality differently. Formally this means that inequality appears in the utility function of both countries, but one country

attaches a lower weight to it than the other. In a democracy the public is able to demand more redistribution in the voting process. Alesina and Rodrik (1994) and Persson and Tabellini (1994) present models in which majority voting induces higher redistribution if inequality is higher.

Most contributions in this field have focussed on differences between the European welfare states and the US. Based on data from the 'World Values Survey', Alesina *et al.* (2004) show that inequality negatively affects utility in both Europe and the US. Although their estimates are more precise for Europe than for the US, the magnitude of the effect is not statistically different. They also note that there exists a left-right divide in Europe and a rich-poor divide in the US. This makes them conclude that inequality preferences do not differ between Europe and the US, but the view on social mobility does (see also Alesina and La Ferrara (2001) and van der Ploeg (2004)). In Europe the poor and the left are unhappy about inequality because they believe it is hard to climb up the social ladder. America is seen as 'the land of opportunities' where one controls one's own destiny. The poor can get rich, and the rich can get poor. Hence, rich Americans are unhappy because of inequality as they might end up at the other end of the table.

Empirical work can not confirm the higher income mobility in the United States (Atkinson *et al.* (1992), Burkhauser and Poupore (1997), Aaberge *et al.* (2002), DiPrete (2002)). Hence, beliefs have an important role in the above explanation. One possible explanation for the fact that observed mobility does not differ, is the larger degree of risk taking in the welfare states (cf. *infra*).

Alesina and Angeletos (2004) attribute the different attitude towards market mechanisms versus redistribution in Europe and the US to historical evolutions. In Europe opportunities for wealth and success have been restrained by class differences at least since medieval times. The resulting income distribution has been perceived as unfair because it was generated more by birth and nobility than by ability and effort. Therefore aggressive redistributive policies are favoured by the Europeans. History and perception in the US are different: successful people have made it on their own. Hence Americans prefer strong property protection and limited redistribution.

Corneo (2000) argues that US citizens support income inequality more than West-Germans because they resist redistribution that may bring less desirable people into their residential and consumption neighbourhoods. Contrary to Alesina *et al.* (2004) he concludes that Americans and Europeans fundamentally differ with respect to inequality preferences.

O'Connell (2004) shows that once one controls for income inequality, the income level has no significant effect on satisfaction. Equality of income is always positively and significantly associated with the level of satisfaction over the period 1995-1998 in 15 European States. He offers two potential interpretations of his results: equality increases social cohesion and it creates more challenging work opportunities for a greater proportion of individuals. But he also point out that the causation might be reversed (satisfied people create an equal society) or a third factor might drive both inequality and happiness. Unfortunately, he has no results for the US. His results neither reveal whether UK preferences differ from those of the other European states.

Bénabou and Tirole (2004) present a model in which 'belief in a just world' and 'redistribution' are linked. If enough individuals end up with the view that economic success is highly dependent on effort ('the American dream'), they will be able to persuade the government to set a low tax rate. On the other hand if the society carries out little redistribution, the costs of a deficient motivation to effort or savings are much higher than with a generous safety net. So preferences are endogenous in the model. This model can account for the fact that two very similar societies end up with a very different amount of social security provisions: European pessimism is confronted with American 'belief in a just world'. This model is closely related to the work of Alesina and Angeletos (2004): since Americans preferred limited regulation and low redistribution, there were also fewer distortions and more efficient market incomes. This reduced the impact of "luck". Hence, ex-ante beliefs were confirmed ex-post.

1.5. Conclusions

Is there a trade off between equality and growth? This survey has shown that theory does not predict that equality can be substituted for growth in a linear way. Potential costs of the welfare state (redistribution) might be neutralised by equal gains. The welfare state might mitigate credit constraints, induce productive risk taking, stimulate innovation, etc.

The cost side of the welfare state might also be overestimated. Firstly, not all redistribution is costly. If one takes into account market imperfections, some redistribution might even promote growth. Secondly, the design of the tax system and social security system in large welfare states seems more pro-growth.

The survey has also illustrated that it might be over-simplified to talk about '*the*' welfare state. The way in which redistribution affects economic activity depends upon the way social expenditure are financed, how eligibility for benefits is determined, etc. The impact of a social security system on economic life will differ across countries because of differences in institutions. Moreover, the impact might change over time due to world-wide economic shocks such as the globalisation of the world economy.

The complexity of the relationship between inequality, redistribution and growth is not only mirrored in a multitude of theoretical explanations, it is also reflected in a lack of robust empirical support for (or against) the different hypotheses.

1.6. References

- Aaberge R., Björklund A., Jäntti M., Palme M., Pedersen P., Smith N. (2002): *Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States*, *Review Of Income and Wealth*, 48 (4), p.443-469
- Acemoglu D., Angrist J. (2000): *How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws*, *NBER Macroeconomics Annual*, 15(1), p.9-59
- Agell J., Lommerud (1993): *Egalitarianism and Growth*, *Scandinavian Journal of Economics*, 95(4), p.559-579
- Alesina A., Rodrik D. (1994): *Distributive Politics and Economic Growth*, *Quarterly Journal of Economics*, 109, 465-490
- Alesina A., Perotti R. (1996): *Income Distribution, Political Instability and Investment*, *European Economic Review*, p.1203-1228
- Alesina A., La Ferrara E. (2001): *Preferences for Redistribution in the Land of Opportunities*, *NBER Working Paper no. 8267*, 41p.
- Alesina A., Angeletos G.M. (2004): *Fairness and Redistribution: US versus Europe*, *NBER Working Paper no. 9502*, 37p.
- Alesina A., Di Tella R., MacCulloch R. (2004): *Inequality and Happiness: are Europeans and Americans Different ?*, *Journal of Public Economics*, 88, p. 2009-2042
- Atkinson A., Bourguignon F., Morrison C. (1992), *Empirical Studies of Income Mobility*, Harwood Academic Publishers, London
- Atkinson A. (1999): *The Economic Consequence of Rolling Back the Welfare State*, *The Munich Lectures in Economics*, MIT Press, Cambridge
- Azariadis C., de la Croix D. (2003) : *Growth or Equality ? Losers and Gainers from Financial Reform*, mimeo, 49p.
- Azariadis C., Drazen A. (1990): *Threshold Externalities in Economic Development*, *Quarterly Journal of Economics*, 105 (2), p.501-526
- Backus D.K., Kehoe P.J., Kehoe T.J. (1992): *In Search of Scale Effects in Trade and Growth*, *Journal of Economic Theory*, 58, p.377-409
- Barr N. (1992): *Economic Theory and the Welfare State: A Survey and Interpretation*, *Journal of Economic Literature*, 30, p.741-803
- Barro R.J. (2001): *Education and Economic Growth*, in: Helliwell J.F. (ed.): *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-being*, OECD, p.14-41
- Bénabou R., Tirole J. (2004): *Belief in a Just World and Redistributive Politics*, Paper presented at the 'Conference on Economic Growth and Distribution', Lucca, 17 June
- Bertola G., Ichino A. (1995): *Wage Inequality and Unemployment: United States vs. Europe*, in: Bernanke B.S. and Rotemberg J.J. (eds): *NBER Macroeconomics Annual 10*, MIT Press, p.13-66

- Bils M., Klenow P.J. (2000): *Does Schooling Cause Growth*, American Economic Review, vol. 90 (5), p.1160-1183
- Bird E.J. (2001): *Does the Welfare State induce Risk-taking?*, Journal of Public Economics, 80, p.357-383
- Bird E.J., Hagstrom P.A. (1999): *The Wealth Effects of Means-tested Transfers*, Review of Income and Wealth, 45 (3), p.339-352
- Black J., de Meza D., Jeffreys D. (1996): *House Prices, the Supply of Collateral and the Enterprise Economy*, The Economic Journal, 106, p.60-75
- Blanchard O., Wolfers J. (2000): *The Role of Shocks and Institutions in the Rise of European Unemployment: the Aggregate Evidence*, The Economic Journal, 110, p.1-33
- Blöndal S., Field S., Girouard N. (2002): *Investment in Human Capital through Upper-Secondary and Tertiary Education*, OECD Economic Studies no. 34, p.41-89
- Blundell R., Duncan A., Meghir C. (1998): *Estimating Labor Supply Responses Using Tax Reforms*, Econometrica, 66 (4), p.827-861
- Boháček R. (2003): *The Efficiency-Equality Tradeoff of Social Insurance*, mimeo, 28p.
- Boone J. (2004): *Unemployment Insurance and Workers' Mobility*, Journal of Institutional and Theoretical Economics, 160, p.275-293
- Brunello G. (2002): *Absolute Risk Aversion and the Returns to Education*, Economics of Education Review, 21, p.635-640
- Burkhauser R.V., Poupore J.G. (1997): *A Cross-national Comparison of Permanent Inequality in the United States and Germany*, Review of Income and Statistics, 79 (1), p.10-17
- Cahuc P., Michel P. (1996): *Minimum Wage Unemployment and Growth*, European Economic Review, 40, p.1463-1483
- Cameron S., Taber C. (2004): *Estimation of Educational Borrowing Constraints Using Returns to Schooling*, Journal of Political Economy, 112 (1), p.132-182
- Carneiro P., Heckman J.J. (2002): *The Evidence on Credit Constraints in Post-Secondary Schooling*, The Economic Journal, 112 (8), p.705-734
- Checchi D., Garcia-Peñalosa C. (2004): *Risk and the Distribution of Human Capital*, Economics Letters, 82, p.53-61
- Corneo G. (2000): *Inequality and the State: Comparing US and German Preferences*, CESifo Working Paper no. 398, 18p.
- Danziger D., Haveman R., Plotnick R. (1981): *How Income Transfer Programs Affect Work, Savings, and the Income Distribution: A Critical Review*, Journal of Economic Literature, 19, p.975-1028
- Daveri F., Tabellini G. (1997): *Unemployment, Growth and Taxation in Industrial Countries*, CEPR Discussion Paper no. 1681, 63p.

- De Grauwe P., Polan M. (2003): *Globalisation and Social Spending*, CESifo Working Paper no. 885, 27p.
- De la Fuente A., Doménech R. (2002): *Human Capital in Growth Regressions: How Much Difference Does Data Quality Make? An Update and Further Results*, CEPR Discussion Paper no. 3587, 70p.
- Den Haan W., Haefke C., Ramey G. (2001): *Shocks and Institutions in a Job Matching Model*, NBER Working Paper no. 8463, 51p.
- DiPrete T.A. (2002): *Life Course Risks, Mobility Regimes and Mobility Consequences: A Comparison of Sweden, Germany and the United States*, American Journal of Sociology, 108 (2), p.267-309
- Feldstein M. (1985): *The Optimal Level of Social Security*, Quarterly Journal of Economics, 100 (2), p.303-320
- Foellmi R., Zweimüller J. (2003): *Inequality and Economic Growth – European versus US Experiences*, IERE Working Paper no. 158, 21p.
- Galor O., Zeira J. (1993): *Income Distribution and Macroeconomics*, Review of Economic Studies, 60, p.35-52
- Grossmann V. (2003): *Risky Human Capital Investment, Income Distribution, and Macroeconomic Dynamics*, IZA Discussion Paper no. 955, 41p.
- Huggett M., Ospina S. (2001): *Aggregate Precautionary Savings: When is the Third Derivative Irrelevant?*, Journal of Monetary Economics, 48, p.373-396
- Islam (1995): *Growth Empirics: A Panel Data Approach*, Quarterly Journal of Economics, 110, p.1127-1171
- Kemnitz A., Wigger B.U. (2000): *Growth and Social Security: the Role of Human Capital*, European Journal of Political Economy, 16, p.673-683
- King R.G., Rebelo S. (1990): *Public Policy and Economic Growth: Developing Neoclassical Implications*, Journal of Political Economy, 98, p.1008-1038
- Krebs T. (2003): *Human Capital Risk and Economic Growth*, Quarterly Journal of Economics, 118, p.709-744
- Kreickemeier U. (2004): *Fair Wages and Human Capital Accumulation in a Global Economy*, CESifo Working Paper no. 1185, 25p.
- Lindbeck A. (1988): *Consequences of the Advanced Welfare State*, World Economy, 11 (1), p.19-37
- Lindbeck A. (1993): *Work Disincentives in the Welfare State*, in: Lindbeck A.: *The Welfare State – The Selected Essays of Assar Lindbeck Vol. II*, Edward Elgar, Hants
- Lindbeck A. (1997): *The Swedish Experiment*, Journal of Economic Literature, 35 (3), p.1273-1319
- Lindert P.H. (2004): *Growing Public*, Cambridge University Press, Cambridge

- Ljungqvist L., Sargent T.J. (1995): *Welfare States and Unemployment*, Journal of Economic Theory, 6, p.143-160
- Ljungqvist L., Sargent T.J. (1998): *The European Unemployment Dilemma*, Journal of Political Economy, 106 (3), p.514-550
- Lucas R.E. (1988): *On the Mechanics of Economic Development*, Journal of Monetary Economics, 22 (1), p.3-42
- Marimon R., Zilibotti F. (1999): *Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits*, The Economic Journal, 109, p.266-291
- Mueller D.C. (1976): *Public Choice – Survey*, Journal of Economic Literature, 14 (2), p.395-433
- Nelson R.R., Phelps E.S. (1966): *Investment in Humans, Technological Diffusion and Economic Growth*, American Economic Review, 61, p.69-75
- O'Connell M. (2004): *Fairly Satisfied: Economic Equality, Wealth and Satisfaction*, Journal of Economic Psychology, 25, p.297-305
- Palokangas T. (1996): *Endogenous Growth and Collective Bargaining*, Journal of Economic Dynamics and Control, 20, p.925-944
- Pauly M.V. (1974): *Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection*, Quarterly Journal of Economics, 88 (1), p.44-62
- Persson T., Tabellini G. (1994): *Is Inequality Harmful for Growth*, American Economic Review, 84 (3), p.600-621
- Rauch J.E. (1993): *Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities*, Journal of Urban Economics, 34, p.380-400
- Rebelo S. (1991): *Long-Run Policy Analysis and Long-Run Growth*, Journal of Political Economy, 99, p.500-521
- Rillaers A. (2001): *Education and Income Inequality: the Role of a Social Protection System*, Journal of Population Economics, 14 (3), p.425-443
- Romer P.G. (1986): *Increasing Returns and Long-Run Growth*, Journal of Political Economy, 94, p.1002-1037
- Romer P.G. (1989): *Capital Accumulation in the Theory of Long-Run Growth*, in: Barro R. (ed.): *Modern Business Cycle Theory*, Harvard University Press, Cambridge
- Saint-Paul G. (1997): *Employment Protection, International Specialization, and Innovation*, mimeo, 23p.
- Sala-I-Martin X. (1997): *Transfers, Social Safety Nets and Economic Growth*, IMF Staff Papers, 44 (1), p.81-102
- Shapiro C., Stiglitz J.E. (1984): *Equilibrium Unemployment as a Worker-Discipline Device*, American Economic Review, 74, p.433-444

Sianesi B., Van Reenen J. (2003): *The Returns to Education: Macroeconomics*, Journal of Economic Surveys, 17 (2), p.157-200

Sinn H.W. (1994): *A Theory of the Welfare State*, NBER Working Paper no. 4856, 38p.

Sinn H.W. (1995): *Social Insurance, Incentives and Risk Taking*, NBER Working Paper no. 5335, 35p.

Sinn H.W. (1998): *The Pay-as-you-go Pension System as a Fertility Insurance and Enforcement Device*, CESifo Working Paper no. 154, 29p.

Summers L.H. (1988): *Efficiency Wages, Labor Relations, and Full Employment*, AEA Papers and Proceedings, p.383-388

Van der Ploeg F. (2003): *Do Social Policies Harm Employment and Growth?*, CESifo Working Paper no. 886, 28p.

Van der Ploeg F. (2004): *The Welfare State Redistribution and the Economy: Reciprocal Altruism, Consumer Rivalry and Second Best*, CESifo Working Paper no. 1234, 35p.

Zweimüller J. (2000): *Schumpeterian Entrepreneurs Meet Engel's Law: The Impact of Inequality on Innovation-Driven Growth*, Journal of Economic Growth, 5, p.185-206

Chapter 2:

Income Inequality Data in Growth Empirics: From Cross-Sections to Time Series

Abstract

As in any other field of applied macro-economic or econometric research, researchers that study income inequality have to look for suitable data. Although most researchers just draw on some ready-made dataset, finding reliable data is not that straightforward and can even be very troublesome. This paper highlights some of the pitfalls in the use of inequality data. We deal with sampling problems, the choice of equivalence scale, scalar measures of inequality, etc. We also introduce and describe a new secondary dataset on inequality data for a number of OECD countries. The main innovation of the dataset is that it creates the possibility to perform a time series analysis on inequality data. The data bring to the fore an additional problem: non-stationary behaviour of inequality series.

JEL Classification: C82, O15

We would like to thank Bas van Aarle, David de la Croix, Rafael Doménech, Gerdie Everaert, Freddy Heylen and Dirk Van de gaer for useful comments and suggestions. We would also like to thank Christian Valenduc, Markus Jäntti, Antonio Brandolini, Wim Kessels and all others that helped with the dataset.

2.1. Introduction

Different branches of macro-economic research are involved with income inequality. Some studies *describe* the level and evolution of income inequality in an individual country (e.g., Piketty (2001) for France, Goodman (2001) for Great Britain), a group of countries (e.g., Brandolini (1998), Gottschalk and Smeeding (2000), Jenkins and Van Kerm (2003)) or even the entire world (e.g., Sala-i-Martin (2001), Milanovic (2002), Deaton (2003)). Some authors go a step further and try to *explain* the observed trends and changes in income inequality (e.g., Gustafsson and Johansson (1997), Förster (2000), Berry and Serieux (2003)). Thirdly, income inequality in itself appears as an explanatory variable in a lot of empirical work. One specific domain of research focuses on the effects of income inequality on economic growth (see Aghion *et al.* (1999) or chapter 3 for a brief overview).

Although the above papers do not share a common goal, they face a common difficulty: the need to define and measure income inequality. At first sight this seems fairly straightforward and in fact a lot of authors just turn to some ready-made dataset. No further questions are asked. However, as we shall argue in this paper, a bit more thought is definitely justified.

The main purpose of this paper is descriptive in nature. We want to warn the reader about practical difficulties with the use of income inequality data in macro estimates and provide guidelines for applied research on a macro-level. To clarify some remarks we use examples from the economic literature with respect to the effects of income inequality on economic growth. Most remarks can easily be extrapolated to other areas of macro-economic and macro-econometric research.

In the first part of the paper we analyse the path from measuring incomes to calculating an income inequality measure. First, we look into the measurement of income flows (section 2.2). We define 'income' and look at two registration methods: tax records and household surveys. Secondly, we compare incomes that are measured at a different recipient level: how can the incomes received by the constituent members of a household be transferred in a single household income, and vice versa (section 2.3)? As one can not include an entire income distribution in a regression analysis, one needs to capture the distribution in a one-dimensional income inequality measure. We map some important problems related to these scalar measures of inequality (section 2.4). To better grasp the relevance of the problems in empirical work, we also present some simulation exercises. We want to stress that the set up of the simulations is rather crude and does not always match real life. The simulations are only included for illustrative purposes and no robust conclusions should be deduced from them. Based on the identified problems we argue that consistency of cross-country

inequality data is hard to ensure. On that account we gather inequality data for all OECD countries and sort out the countries for which a long and consistent annual time series exists. We present the resulting dataset in the second part of the paper and look for trends in income inequality over the last decades in the included countries. We make some inferences about the time series characteristics of inequality data and point out the non-stationary behaviour of these series (section 2.6). The dataset can be helpful to explore the possibilities of a time-series approach to the econometrics of inequality and growth. We end by summarizing the most important insights of the paper (section 2.7).

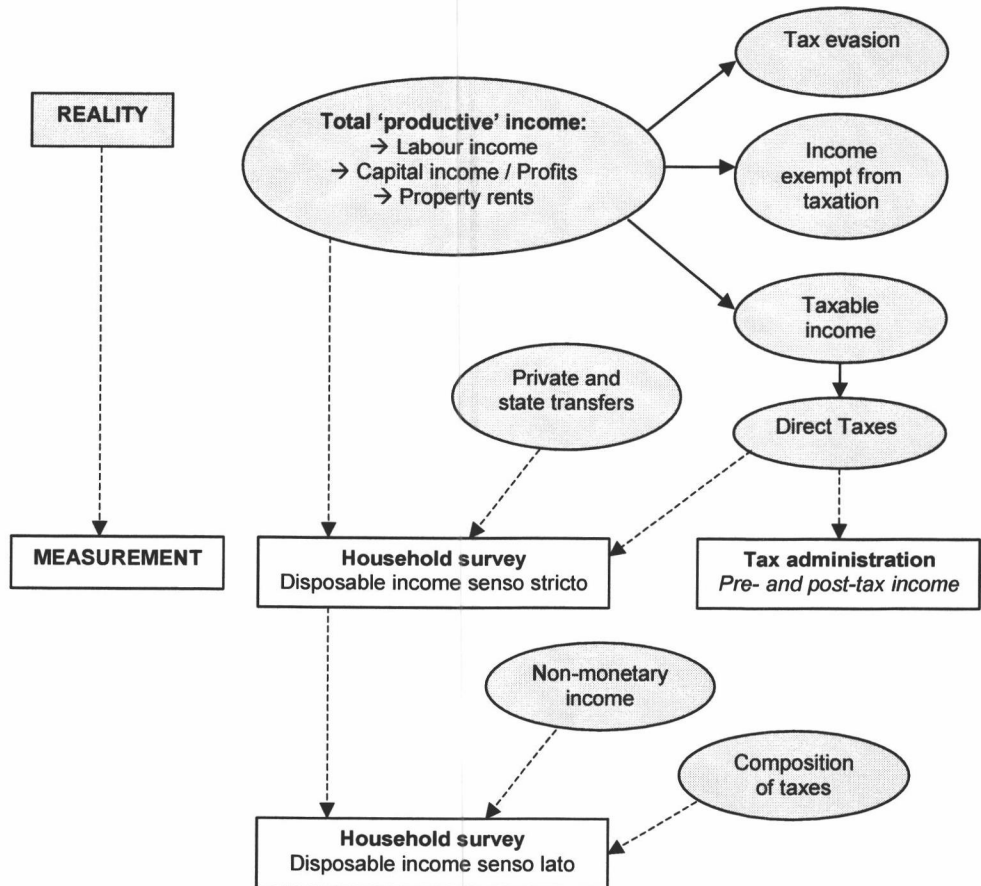
2.2. Measuring incomes ... part one: what income?

Before one can compute the amount of income inequality in a society, one needs to measure income itself. A lot of practical and conceptual difficulties can hinder the *registration* of income flows (cf. *infra*) and already the first step in the process turns out to be quite treacherous: how does one define 'income'? Which sources of income does one want to include (or exclude)?

Figure 1 gives a stylised overview of the different components that potentially matter for the measurement of income. An obvious starting point are the wages, salaries, ... and other forms of *labour income*. While this component is the main source of income on the macro-level, it will not be the main factor in the top incomes. Corporate profits, interest income and property rents must necessarily be added to complete the picture. The government will interfere in economic life by taxing the different income flows and by transferring income (e.g., unemployment benefits) to the destitute in society. Transfers might also stem from other agents in the private sector. Clearly this kind of redistribution will affect a household's disposable income. Also non-monetary income can be taken into account: e.g., if education is provided at a strongly reduced cost, the disposable income of households with children increases. In a similar way the composition of taxes might matter: the replacement of direct taxes by indirect taxes will affect households according to their respective consumption patterns.

In practice, it may not be straightforward to obtain sound data for all income flows. Two income concepts are directly measured: the tax administration registers the taxable income, household surveys aim at disposable income. For other income concepts one needs to merge different data sources. And even tax and survey data are not trouble-free.

Figure 1: Incomes and income measurement



The tax administration has (should have) information with respect to all *taxable* income. However, that is only an approximation of total income. The quality of the approximation will depend on the specifics of the tax system (e.g., tax exemptions), the amount of tax evasion, ... Capital income will typically be underestimated if tax data are used. Especially the data for the lowest and highest incomes will be less reliable. Because the importance of underreporting differs with income levels, it constitutes a sizeable threat to the quality of the inequality measures. Moreover, tax data contain only very limited information with respect to the amount of income received outside the production sphere. Income after taxes is not likely to be a good proxy for disposable income if there exists an extended welfare system.

A source of information more suited to approximate disposable income are the regular household income surveys that are organised by most countries¹. Can survey data provide a more reliable picture than tax data? Although much will depend on the way these surveys are conceived, the results will probably match the true disposable income better than those obtained by using tax data. But surveys also suffer from considerable shortages. Szekely and Hilgert (1999) show that high incomes are measured with significant error in surveys. Rich people will 'moderate' their income and poor (phone-less) people will be underrepresented. Not all sources of 'non-monetary' income will show up in the survey results: e.g., direct income support by the government will be measured, but direct consumption subsidies will not. Moreover, as the results of a survey are 'projected' on the entire population, the quality of the data will also depend on the quality and the proximity of the latest population census (next to the quality of the survey itself).

As surveys do not capture the entire population there is a risk that the sample properties do not perfectly match the properties of the entire population. Breunig (2001, 2003) shows (theoretically and empirically) that for some inequality measures (e.g., the coefficient of variation) there exists a small sample bias in case of positively skewed income distributions. However, the large sample distribution of most inequality measures is known (see Cowell (1989), Thistle (1990)). An important finding is that the asymptotic variance of the inequality estimators is proportional to the inverse of the sample size. Hence, the variance depends on the absolute size of the sample. Since inequality indices are usually derived from very large micro data sets, sampling error is probably only of minor importance for most inequality measures.

The above conclusion holds in case of random sampling error. In reality (cf. supra) the sampling error might be systematic rather than random. Samples will not be reliable as rich households underreport their income and poor households are hard to include. We perform an explorative simulation exercise to grasp the relevance of this problem.

We randomly draw 5000 incomes from a lognormal distribution $L(\mu, \sigma^2)$ with $\mu=1$ and $\sigma=0.7^2$. Following Aitchinson and Brown (1973) we can deduce that the choice of parameters implies that the income distribution has a gini coefficient of about 0.38. This resembles the gini coefficient for gross household income in the UK. The mean itself is irrelevant, as we consider two mean-invariant inequality measures: the gini coefficient and

¹ On the other hand surveys might not outperform tax data in e.g., the measurement of income before taxes. Hence, one can not make the general statement that one data source is superior the another in the measurement of 'income'.

² Consider an essentially positive variate X ($0 < X < \infty$) such that $Y = \text{Log}(X)$ is normally distributed with mean μ and variance σ^2 . We then say that X is lognormally distributed and write X is $L(\mu, \sigma^2)$ and Y is $N(\mu, \sigma^2)$.

the ratio of the income share of the tenth and first deciles³. The choice of distribution will influence the results of the simulation (e.g., the weight in the tails of the distribution will clearly matter for the value of the deciles ratio). We choose the lognormal distribution⁴ with the specified parameters because it fits reasonably well the kernel estimates of the household income distribution for a number of European countries of Papatheodorou *et al.* (2003).

Next we reduce the sample deleting all incomes below 0.86⁵ (low incomes are not included). All incomes above 8.63 are cut off at that value (high incomes are underestimated). We repeat the procedure 100 times. The gini coefficient for the unadjusted (correct) total sample has a mean of 0.379 with a standard deviation of 0.004. The gini coefficient of the adjusted data (i.e. the reduced sample) has a mean of 0.349 with a standard deviation of 0.003. The difference is quite large and significant. The deciles ratio has a respective mean of 6.032 and 6.033 with a respective standard deviation of 0.143 and 0.150. So misrepresentation seems to have a differential effect depending on the inequality measure.

Systematic sampling error could potentially affect the validity of inequality measures based on survey data.

Up to here we did not question the fact that we measure income. For many a research question the wealth distribution or asset distribution might matter more (e.g., to look into the relevance of credit constraints). But the practical difficulties linked to the measurement of assets are even larger. Deininger and Squire (1998) argue that land distribution outperforms income distribution as a proxy for asset distribution. In their sample the correlation between the gini coefficient for the initial distribution of land and that of incomes is only 39%. Deininger and Olinto (2000) also show that the correlation between income inequality and asset inequality is weak. Smith (2001) demonstrates for the US that the evolution and the level of wealth within narrow income groups is very different.

Another alternative for income inequality is human capital inequality (Castelló and Doménech (2002)). Human capital can serve as an indicator for potential future income which is again a highly relevant factor to assess the relevance of credit constraints. Gyimah-Brempong and Wilson (2004) look at health human capital.

³ Whether inequality measures should be mean-invariant is also an object of discussion. If two distributions only differ in their mean, a lot of people will consider the one with the higher mean to be less unequal as everybody is richer in absolute terms (see Sen (1973)).

⁴ We can also motivate our choice of a lognormal distribution by referring to Gibrat's law. Gibrat (1931) showed that if one excludes the top 1% of earners, the remaining 99% of incomes follow a lognormal distribution. This result is generalized into Gibrat's law or the 'law of proportionate effects'. It basically states that if a variable undergoes random independent proportionate changes, the distribution of the logarithms of this variable will eventually be approximately normal. However, Gibrat's law is not undisputed (see e.g., Kalecki (1945)). It has been shown that the top 1% of the income distribution can be represented by a Pareto distribution.

⁵ With a lognormal distribution with parameters 1 and 0.7, there is a 5% probability that an income is below 0.86 and a 5% probability that an income is above 8.63. In our simulation about 500 out of 5000 incomes are affected.

Given the choice of different income (wealth) concepts, can we determine which one is optimal? The ideal definition in empirical work will depend on the theoretical framework. To illustrate this statement we look at three theories that account for a negative relation between income inequality and economic growth: the complete market model (CMM), the imperfect market model (IMM) and the socio-political instability model (SPIM) (see Perotti (1996) for an introduction). In the CMM, growth is reduced because more income inequality induces a higher demand for redistribution. More redistribution means that average taxes have to increase, which in turn will disturb savings and investment decisions. To test the relevance of this model, one needs a measure of pre-tax and pre-redistribution inequality, as well as a measure of redistribution efforts. The IMM links more inequality to lower growth through the education channel. If people are poorer (i.e. they lack funding), they will under-invest in human capital, which will slow down growth. Now disposable income (or ideally disposable assets) will matter. Finally the SPIM links more inequality to more social unrest which negatively affects the growth performance. Again disposable income will be the more relevant income variable.

In applied work one is often restricted by the available data. Therefore, many authors do not actually test what they are claiming to test. Still their results can be informative (and in any case hard to improve upon) as long as the reader is aware of the shortcomings. The potential problems might be less relevant in a time series framework if the correlation between different income series is high enough (cf. *infra*). Unfortunately this might be 'wishful thinking': Ervik (1998) shows that the trend in the gini coefficients based on market incomes⁶, gross incomes and disposable income for 8 countries (period 1980-1995) can differ substantially. Brandolini (1998) concludes that the behaviour of inequality of market income is much more homogeneous across a sample of OECD countries than the behaviour of inequality of disposable income. Atkinson and Brandolini (1999) show that net and gross income distributions may behave differently over time within a single country. Rehme (2002) illustrates how mixing measures of gross and net income inequality can blur the estimation results if redistribution negatively affects growth. Knowles (2001) shows that many empirical results disappear once one corrects for the fact that different income concepts have been mixed together.

Table 1 compares some properties of income inequality statistics derived from tax records and of those derived from household income surveys. In view of the use of the data in

⁶ Market income is defined as "earned income of wages and salaries and self employment, cash property income and other private cash income transfers. Gross income is a somewhat broader concept and also includes social insurance cash benefits, universal cash transfers and social assistance.

growth regressions, comparability (in case of *cross-section analysis*) and frequency (in case of *time series analysis*) deserve special attention.

Table 1: Tax records versus household surveys

	Tax Records	Household Income surveys
<i>Income concept</i>	Depending on tax system Weak proxy for disposable income	Depending on questions Mainly labour income, but potentially a better proxy for disposable income
<i>Reliability</i>	Tax evasion (especially highest incomes) Low incomes are under-represented	Highest incomes are seldom correctly measured Dependent on time of year
<i>Representativeness</i>	Population with taxable income	Small sample of population Proximity of population census determines quality of extrapolations
<i>Comparability</i>	Countries: Low; very different tax systems across countries Time: Changes in tax systems over time can be considerable	Countries: Low; (very) different methodology across countries Time: Acceptable
<i>Frequency</i>	Annual. Long time series can be constructed	Mostly irregular, not annual. Only a few observations per country

2.3. Measuring incomes ... part two: whose income?

Income flows can be measured at different recipient levels: one can look at individuals, households or families. Surveys will normally aim at household income, as it is hard for individuals to discriminate between income flows of the constituent household members. Tax data are more mixed: depending on e.g., the wedded state sometimes household income and sometimes individual income is registered. Again it will be useful to have some insights in the specificities of the tax system.

Even if income flows are consistently reported at a certain level, this might not be the level preferred by the researcher. It will suffice to add the individual incomes of the different household members to compute total household income. But it is less straightforward to approximate individual disposable income on the basis of household income data or to compare households that differ in size and composition. Dividing total household income by the number of household members might not be the best approach as it will probably result in an underestimation of disposable income. Usually there will exist scale and scope effects in consumption on the household level (e.g., one television can suffice for the entire household). Thus the disposable income of both constituent parts of a couple will probably

exceed half of total household income. Moreover, needs can differ across household members (children will normally have less and less expensive needs). One needs to convert non-comparable incomes received by households, to comparable welfare imputed to individuals. *Equivalence scales* seek to answer the question “how much money does a household need to spend to be as well off as a single person living alone?”. Initially this question was answered in terms of equal consumption levels. Nowadays the focus is on equality of utility. Estimating equivalence scales is far from straightforward as one needs to specify individual and household utility functions⁷, demand functions, the extent of joint consumption within a household, ... (see Browning *et al.* (2004) for a complete methodology).

A popular method, given the complexity of the matter, is to assign different weights to the different household members to derive the ‘equivalent disposable income’. A frequently used procedure (e.g., by the Luxembourg Income Study (LIS)) is to assign a value of 1 to a single adult person, a value of 1.7 to a couple and add 0.5 to these figures for each child. Since estimated equivalence scales differ substantially, these figures can be seen as a ‘feasible compromise’. However, other (equally sensible) choices can be made. For instance, one could take into account that the needs of households with only 1 child will differ from those of households with more children. Next to household size and composition also additional influential family characteristics such as region or location can be accounted for. It is important to be aware of this if one uses ‘ready-made’ data. A different choice of equivalence scale could result in a different income distribution and introduce a bias in cross-country estimates. For a more comprehensive discussion of equivalence scales we refer to Cowell and Mercader-Prats (1999), Flückiger (1999), Atkinson *et al.* (1995) and Buhmann *et al.* (1988).

To assess the impact of different equivalence scales in inequality measurement, we resort to a simulation exercise. We consider four types of equivalence scale to adjust household income.

- (0) No correction: the sum of the incomes of both spouses equals household income, i.e. we ignore household ties and scale effects on the household level;
- (1) The 0.5-equivalence scale: total household income is divided by the square root of the number of household members;
- (2) The per capita scale: total household income is divided by the number of household members;

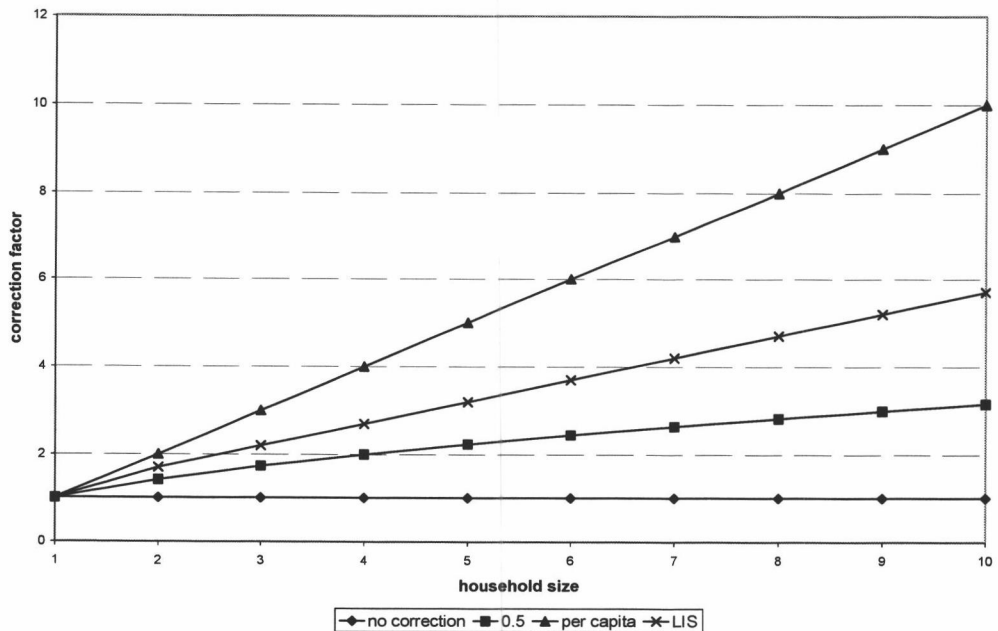
⁷ Strictly speaking not households but the individuals that compose it have utility. However, preferences of single individuals might change once they belong to a household (e.g., the utility of ‘eating at a restaurant’).

- (3) The LIS scale: total household income is corrected using the LIS scale (single adult: 1, couple: 1.7 augmented by 0.5 per child).

Again, the included equivalence scales are the ones that are commonplace in standard empirical work⁸.

Note that all corrections depend on household size, but differ in the weights that are accorded to each household member. Figure 2 illustrates how the impact on adjusted household income of different equivalence scales diverges with the household size.

Figure 2: Household size and equivalence scales



The impact of the equivalence scale will clearly depend on the probability that a person is single and on the probability that a household has a certain number of children. E.g., if everybody is single, it will not matter how one corrects for household size and composition. Hence, it is important that our simulation set up does not deviate too much from reality. We fix the a priori probability that a person is married, and the a priori probability of a household having a certain number of children on the basis of the Belgian data (NIS, 2004)⁹. We only

⁸ We do not insinuate that these equivalence scales are to be preferred. As they only take household size into account (and ignore location, composition, etc.), they can certainly be approved upon.

⁹ Singles: 37%; Couples: 63%, of which couples without children: 44%, with 1 child : 23%, with 2 children: 22%, with 3 children: 11% (the last category also includes households with more than 3 children in reality).

consider 'traditional' households. A household consist of maximum two adults. Married people can have 0, 1, 2 or 3 children, singles have no children. Children do not earn an income.

First, we randomly draw 500 incomes from a lognormal distribution with $\mu = 1$ and $\sigma = 0.7$ (cf. supra). In a second step we determine the marital status of each person. We randomly draw a number between 0 and 1 for each person. A person is single if we draw a number strictly smaller than 0.37, otherwise he is married. If somebody is married, we draw an additional income out of the same lognormal distribution and add it to his income. Next we determine the number of children in a household. Again we randomly draw a number between 0 and 1 for each household. If the number is strictly smaller than 0.44 than the household has no children. If the number is between 0.44 and 0.67 (0.67 and 0.89), the household has 1 child (2 children). If the number exceeds 0.89, the household has 3 children.

Next we compute gini coefficient and the ratio of the 10th and 1st deciles of the income distribution based on these (corrected) data. We repeat the procedure 500 times, which enables us to evaluate the differences in the inequality measures caused by the use of different equivalence scales. We have summarized the main conclusions in table 2.

The LIS equivalence scale and the 0.5-equivalence scale lead to similar results. For both the deciles ratio and the gini coefficient the differences in the mean are small and statistically insignificant. However, if one uses no correction or the per capita scale, than the outcome is significantly different. The results indicate that inequality is overestimated in both cases.

We can draw some additional insights from the simulation results. Suppose that we consider the 500 runs as 500 different countries. Then we can compute a correlation between the inequality measures and order the countries based on the degree of inequality¹⁰ (panel B and panel C). If we neglect the uncorrected measure, the correlations are fairly high for the gini coefficient but considerably lower for the deciles ratio (panel B).

In panel C we list the average differences in the ranking of the distributions if a different equivalence scale is used. If we compare e.g., the ranking of the 500 distributions based on the gini coefficient of the uncorrected data with the ranking based on the data corrected with the 0.5-equivalence scale we obtain an average difference of 64 positions (with a standard deviation of 56 positions)). The differences in ranking are substantial.

¹⁰ Recall that the income distributions of the 500 "countries" are based on the same lognormal distribution. Hence, asymptotically the income distributions are identical over "countries". The comparison only makes sense because we can make use of the sampling error: the small sample distribution differs substantially from the asymptotical one.

Table 2: The impact of different equivalence scales on inequality measurement – a simulation exercise

PANEL A: Mean and standard deviation				
	<i>No correction</i>	<i>Eq. Scale 1</i>	<i>Eq. Scale 2</i>	<i>Eq. Scale 3</i>
Gini				
Mean	0.354	0.333	0.361	0.337
Standard deviation	0.011	0.011	0.012	0.011
10th Decile/1st Decile				
Mean	6.130	4.824	5.398	4.823
Standard deviation	0.487	0.293	0.332	0.288
PANEL B: Correlations				
	<i>No correction</i>	<i>Eq. Scale 1</i>	<i>Eq. Scale 2</i>	<i>Eq. Scale 3</i>
Gini				
Eq. Scale 2	0.839			
Eq. Scale 3	0.465	0.852		
Eq. Scale 4	0.676	0.963	0.960	
10th Decile/1st Decile				
Eq. Scale 2	0.695			
Eq. Scale 3	0.213	0.535		
Eq. Scale 4	0.501	0.805	0.737	
PANEL C: Differences in rank				
	<i>No correction</i>	<i>Eq. Scale 1</i>	<i>Eq. Scale 2</i>	<i>Eq. Scale 3</i>
Gini				
	(Mean/St. Dev.)			
Eq. Scale 2	64 (56)			
Eq. Scale 3	119 (94)	62 (55)		
Eq. Scale 4	92 (77)	32 (29)	32 (28)	
10th Decile/1st Decile				
Eq. Scale 2	88 (71)			
Eq. Scale 3	144 (109)	110 (91)		
Eq. Scale 4	111 (90)	69 (61)	83 (67)	

Notes: *Equivalence scale 1: square root of total household members*
Equivalence scale 2: total household members
Equivalence scale 3: LIS scale (single adult:1, couple: 1.7, per child: +0.5)

However, one should bear in mind that all incomes were drawn from the same lognormal distribution. Therefore, income distributions are very similar and rankings will consequently be very prone to small changes. If we introduce larger differences in income distributions, the rankings become more stable¹¹.

Our conclusion mimics the one by Atkinson et al. (1995) based on LIS data. Inequality rankings at a point in time are fairly robust to the choice of equivalence scales. Buchmann et al. conclude that “equivalence scales have in general no great effect on the rank order of measured inequality across countries as long as average family size is not extremely large”.

¹¹ If we allow for a lognormal distribution that can vary in σ between 0.4 and 0.9 the mean differences in ranking drop by about 80% (for both gini coefficients and deciles ratio). So especially in case of very similar income distributions, the choice of equivalence scales is important.

The choice of equivalence scale in a cross-country or panel approach might be a less decisive factor as long as an identical scale is used for all countries. However, mixing data based on different scales is much less sensible.

2.4. What does one number say?

It is convenient to capture an entire income distribution in a single scalar in order to use it in macro estimates. On the downside one single statistic will never reveal all relevant aspects of the entire distribution. Moreover, there are different possibilities to represent a distribution. All these summary statistics can be used to rank countries, but these rankings can tell a different story (Caminada and Goudswaard (2000)).

Does one measure systematically outperform the others? Or should one make the choice dependent on the research question? We limit the discussion in this section to some selected issues and surely do not aim to provide a comprehensive overview of income inequality measurement. For further reading we refer to different chapters in the 'handbook of income distribution' (2000) and the 'handbook on income inequality measurement' (1999) next to seminal contributions by Kolm (1969), Atkinson (1970) and Sen (1973).

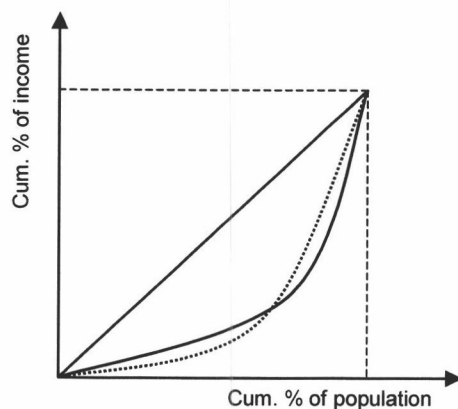
The existence of 'competing' measures of inequality stems from the dual nature of the conception of inequality. Next to an objective element in this notion (a 50-50 division of the cake is objectively more equal than giving all to one and none to the other) there is also a distinctive normative feature. In more complex problems, comparing alternative income distributions among a large number of people, the measurement of the inequality measure could be intractable without bringing in some ethical concepts (Senn (1973)). Thus inequality measures become dependent on one's ethical values. Generally one needs to make a specific choice of criteria of comparisons, properties of measures, parameters, etc. based on one's motives to make these comparisons or compute these measures. Atkinson (1970) argues that the first step in measuring inequality is specifying the social welfare function. He notes that *'while there is undoubtedly a wide range of disagreement about the form the social welfare function should take, this direct approach allows us to reject at once those [inequality measures] that attract no supporters and also serves to emphasise that any measure of inequality involves judgments about social welfare'* (Atkinson (1970), p.257).

Income distributions are frequently represented through the Lorenz curve. This curve plots the cumulative percentage of income recipients arranged in rising order of income versus the cumulative percentage of income they earn. If everybody has the same income the Lorenz

curve will be a straight (45°) line. If there exists some inequality, the line will be dented. One might be tempted to conclude that the bigger the dent is, the larger the amount of inequality will be. However, the Lorenz curve does not allow for a unique ranking of distributions based on the degree of inequality as the possibility of intersecting Lorenz curves can not be excluded¹²: no generalised statement about overall inequality can be made with a one dimensional 'Lorenz curve'-based inequality measure. In fact, no scalar measure of inequality will be able to deliver a unique ranking.

Figure 3 illustrates how different distributions can result in an identical gini coefficient. The gini coefficient measures the surface between the 45° line and the Lorenz curve. Changes in aggregate inequality can hide specific movements in the middle and the extremes of the income distribution. The gini coefficient will hardly react to the transfer of income from the poorest to the richest, because the middle part of the distribution is unaffected. The non-uniqueness of the ranking may not only mask changes in inequality as recorded by one specific measure, it may also lead to contradictory conclusions if different measures are used to compare distributions. Returning to our earlier example: while the gini coefficient will hardly change, the entropy measures will since they are much more sensitive to changes in the bottom and top part of the distribution.

Figure 3: Intersecting Lorenz curves



Cowell (1995) shows that the generalized entropy class of indices (including the Theil and Atkinson indices) has some attractive properties with respect to the measurement of inequality. They incorporate the notion of 'inequality preferences' (how strongly does one

¹² To illustrate that the possibility of intersecting Lorenz curves is not negligible: Atkinson (1970) finds that in the data on five developing countries and seven advanced countries collected by Kuznets (1963), out of 66 pairwise comparisons of the Lorenz curves, only 16 do not intersect.

dislike inequality) and are decomposable (inequality of the whole population can be decomposed in inequality between and within different subgroups). However, in empirical work, these measures are not commonly used as they are not readily available.

In table 3 we look into some influential contributions in the inequality-growth literature, and report which inequality measure they use. Next we point out some of the advantages and problems of the measures that are commonplace in empirical applications.

Table 3: Inequality measures and data sources in the inequality-growth literature

	<i>Inequality measure</i>	<i>Data source</i>
Alesina and Rodrik (1994)	1/ Gini for income distribution 2/ Gini for land distribution	Taylor and Hudson (1972); Jain (1975); Fields (1989)
Arjona <i>et al.</i> (2001)	1/ Gini for income distribution 2/ Mean log deviation of income 3/ Squared coefficient of variation 4/ Ratio of 10 th and 1 st decile	Förster (2000)
Banerjee and Duflo (2000)	Gini for income distribution	Deiningen and Squire (1996)
Barro (1999)	Gini for income distribution	Deiningen and Squire (1996)
Deiningen and Squire (1998)	1/ Gini for income distribution 2/ Gini for land distribution	Deiningen and Squire (1996)
Forbes (2000)	1/ Gini for income distribution 2/ Ratio of income share 5 th and sum of 1 st and 2 nd quintile 3/ Ratio of income share 5 th and 1 st quintile 4/ Income share 3 rd and 4 th quintile	Deiningen and Squire (1996)
Perotti (1996)	1/ Income share of 3 rd and 4 th quintile 2/ Income share of 3 rd quintile	Perotti (1994)
Persson and Tabellini	1/ Income share of 3 rd quintile 2/ Income share of 5 th quintile	Lindert and Williamson (1985); Hartog and Veenbergen (1978); Jain (1975); USDC (1975); Flora <i>et al.</i> (1987); Paukert (1973)
Weede (1997)	1/ Gini for land ownership 2/ Income share of 5 th quintile 3/ Income share of 3 rd quintile 4/ Income share of 1 st and 2 nd quintile	Persson and Tabellini (1994); Alesina and Rodrik (1994)

Note that the Deiningen and Squire dataset (DS) (cf. *infra*) has somewhat become 'the standard' in the empirical research since its conception in 1996. Also note that most studies use either the gini coefficient or some kind of deciles/quintiles ratio to measure income inequality. So they seem to disregard the theoretically preferable measures. The reason for this is fairly obvious: the availability of the gini coefficient (and to a lesser degree the ratio

measures) is far greater than that of any other inequality index. An additional motive is a bit perverse: as the use of the gini coefficient is commonplace in empirical applications using it is necessary if one wants to compare with previous studies.

Another common measure in empirical work is the 'head count index'. This measure indicates the percentage of the population that earns less than a certain amount of money a day. Hence, it is not a measure of inequality but a measure of poverty. It is especially useful to map the degree of poverty in a society. However, it is hard to measure in surveys as the poorest are usually the hardest to sample.

That the choice of inequality measure matters is illustrated by Partridge (1997). For a sample of US states he finds that a larger gini coefficient (normally interpreted as more inequality) results in more growth. However, he also notes that a larger share of income for the middle quintile (normally interpreted as less inequality) increases growth. Although these results are not necessarily irreconcilable, they provide a warning not to base conclusions on a single indicator. A similar conclusion is reported by Szekely and Hilgert (1999). In their estimates there is no significant relation between the gini coefficient and economic growth in a group of Latin American countries. However, if a bottom sensitive entropy index is used the effect of inequality on growth is negative and marginally significant. If instead a top sensitive entropy index is used, both variables are significantly positively related. So in their group of countries equally acceptable summary measures lead to significantly different conclusions about the effects of inequality on growth. One should be wary of hasty conclusions and preferably look at several inequality measures.

Again the choice of the inequality measure might be inspired by the theoretical framework. Credit constraints are at the heart of the IMM. Thus one should pay special attention at the bottom decile of the income distribution to explore that particular model. In the CMM the median voter will play an important part as the policy maker will determine the appropriate tax rate based on his relative position. An income measure that assigns more weight to the middle class (the gini coefficient or the income share of the middle quintiles) could now be preferred.

Alternatively the choice can be based on welfare considerations. If one believes that poverty is the worst of evils, one will look at a poverty line or poverty gap indicator. If one is more interested in the position of the middle class the gini coefficient will be a better alternative.

To get a first impression of the differences in estimation results caused by differences in the inequality measure used, we perform a simulation exercise. For 50 countries we randomly draw 1000 household incomes from a lognormal distribution with μ fixed at 1 and σ

randomly chosen between 0.6 and 0.8. This implies that in the gini coefficient of the most equal distribution amounts to 0.33 (close to the Swedish gini coefficient of gross household income in 2000). The gini coefficient of the most unequal distribution is about 0.43 (close to the US gini coefficient of gross household income in 2000). We also allow the mean to differ: the highest mean is 5 times the lowest mean. This will only make a difference for the value of the head count index, as all other included measures are relative to the mean income. The more the distributions resemble each other, the more important small income changes will be to discriminate between them. In particular the probability of intersecting Lorenz curves will be much higher if distributions are very similar (cf. supra) and hence a ranking of distributions based on e.g. the gini coefficient will be less robust. We only allow for limited difference in income distributions to see how sensitive the ranking of similar countries is to the use of different measures. Alternatively our 'simulated' (ranking) correlations could be considered as lower limits for a more diverse group of countries.

We repeat the exercise with samples of only 50 households to get an idea of the impact of outliers (unrepresentative samples) on the inequality measures. The differences between indices will be influenced by the size (representativeness) of the sample. Cowell and Flachaire (2002) show that the Atkinson index with an inequality aversion parameter above 1 is very sensitive to small incomes. The gini index is less sensitive to extremely high incomes. Cowell and Victoria-Feser (1996) show that data contamination, i.e. outliers in the data that are due to errors in the recording process, can have a major impact on the value of several inequality estimators. They conclude that *"the impact upon measured inequality of quite small amounts of contamination in the tails of the distribution can be disastrous"* (p.99). Cowell and Flachaire (2004) extend the analysis by also considering high leverage observations and examining the robustness of the conclusions with respect to the shape of the income distribution.

We compute 5 inequality measures from the simulated data:

- The gini coefficient;
- The Atkinson index with inequality aversion parameter 0.5 (low importance of inequality);
- The Atkinson index with inequality aversion parameter 5 (high importance of inequality);
- The ratio of the share of the tenth and first deciles in the income distribution;
- A head count index (the percentage of households earning less than 65. This number is equal to 60% of the mean of the income distribution with lowest possible mean income).

We repeat the procedure 100 times.

A glance at table 4 shows that the differences in ranking caused by using different inequality measures can be important.

Table 4: The impact of different inequality measures- a simulation exercise

PANEL A: Correlations								
Sample	Mean		St. deviation		Minimum		Maximum	
	Large	Small	Large	Small	Large	Small	Large	Small
Gini – Atk(0.5)	0.998	0.989	0.001	0.004	0.996	0.975	0.999	0.995
Gini – Atk(5)	0.859	0.625	0.039	0.083	0.740	0.325	0.941	0.798
Gini – Deciles	0.959	0.726	0.011	0.065	0.923	0.555	0.978	0.925
Gini – HCI	0.134	0.138	0.142	0.153	-0.162	-0.289	0.427	0.484
Atk(0.5) – Atk(5)	0.858	0.637	0.038	0.080	0.737	0.371	0.939	0.806
Atk(0.5) – Deciles	0.954	0.706	0.013	0.077	0.912	0.480	0.974	0.939
Atk(0.5) – HCI	0.134	0.143	0.142	0.155	-0.163	-0.263	0.435	0.540
Atk(5) – Deciles	0.844	0.651	0.041	0.075	0.726	0.482	0.939	0.827
Atk(5) – HCI	0.123	0.215	0.141	0.132	-0.229	-0.069	0.403	0.538
Deciles – HCI	0.142	0.175	0.139	0.158	-0.158	-0.169	0.444	0.548

PANEL B : Spearman Rank Correlations								
Sample	Mean		St. deviation		Minimum		Maximum	
	Large	Small	Large	Small	Large	Small	Large	Small
Gini – Atk(0.5)	0.997	0.992	0.001	0.003	0.990	0.982	0.999	0.997
Gini – Atk(5)	0.860	0.612	0.041	0.091	0.732	0.352	0.946	0.813
Gini – Deciles	0.958	0.736	0.014	0.065	0.909	0.555	0.978	0.899
Gini – HCI	0.283	0.275	0.139	0.095	-0.110	-0.289	0.572	0.451
Atk(0.5) – Atk(5)	0.863	0.639	0.041	0.087	0.736	0.371	0.947	0.834
Atk(0.5) – Deciles	0.952	0.724	0.016	0.066	0.900	0.480	0.977	0.900
Atk(0.5) – HCI	0.284	0.283	0.139	0.095	-0.128	-0.263	0.578	0.481
Atk(5) – Deciles	0.858	0.657	0.041	0.082	0.702	0.482	0.943	0.871
Atk(5) – HCI	0.304	0.379	0.133	0.079	-0.094	-0.069	0.620	0.525
Deciles – HCI	0.285	0.297	0.136	0.098	-0.111	-0.169	0.590	0.539

PANEL C : Descriptive Statistics								
Sample	Mean		Variance		Minimum		Maximum	
	Large	Small	Large	Small	Large	Small	Large	Small
Gini	0.380	0.369	0.0009	0.0023	0.311	0.231	0.454	0.557
Atk(0.5)	0.116	0.112	0.0004	0.0009	0.077	0.042	0.169	0.260
Atk(5)	0.691	0.637	0.0042	0.0088	0.516	0.329	0.934	0.923
Deciles	6.112	6.189	0.9394	3.0885	4.217	2.615	8.762	18.446
HCI	0.036	0.036	0.0363	0.0034	0	0	0.297	0.360
			Standardized var.					
Gini			0.007	0.017				
Atk(0.5)			0.027	0.072				
Atk(5)			0.009	0.022				
Deciles			0.025	0.081				
HCI			2.202	2.670				

Notes: We rescale all measures by dividing them by their respective means (so the mean of the new series equals one). The standardized variance is the variance of these rescaled measures.

The correlation between the different inequality indices and the head count index is very low. As the head count index is a *poverty* measure it does not come as a complete surprise that is only weakly correlated with *inequality* measures. Inequality measures should circumscribe the entire distribution, poverty measures only consider the bottom part. While the correlations and rank correlations between the (pure) inequality measures are on average fairly high, they sometimes differ substantially across simulation runs. Some measures are more influenced by certain particularities in the income distribution than others. In small samples the correlations are substantially lower.

Figini (2000) also finds high rank correlations between different inequality measures based on 'real life' LIS data, which indicates that our results can not be fully 'blamed' on the use of 'ad hoc' distributions.

From the point of view of empirical work, it is also interesting to look at the differences in the amount of variation in the different measures. First we rescale the measures to make them comparable. The 'standardized' variances of the gini coefficient and the Atkinson index with low inequality aversion are very similar and only about $1/4^{\text{th}}$ of the variances of the deciles ratio and the Atkinson index with high inequality aversion. Overall the measures seem very stable. Of course the variance will increase if we consider more diverse income distributions. However, our results might be informative for the group of quite similar OECD countries.

2.5. Half a pound of data!

Do there exist reliable sources for ready-made income distribution data? A look at table 2 indicates that the answer might be 'yes since 1996'. Since then most researchers have applied the 'Deininger and Squire'-dataset (DS, Deininger and Squire (1996)). The DS dataset has substantially increased the data availability and comparability, both over countries and over time. Before 1996, data had to be gathered from multiple sources and data comparability was sometimes unacceptably low. Deininger and Squire collected most of the known income distribution data, submitted them to a thorough investigation and finally divided them into different categories according to some quality standards. The end result is a subset of the data that the authors have labelled 'high quality data'. The (updated) DS dataset covers a few observations of 112 developed and developing countries (in total nearly 700 'acceptable' observations) and thus seems mainly suited for cross-section or panel estimates.

The 'high quality'-label assigned to the data is not entirely undisputed. Atkinson and Brandolini (1999) show that there remain important problems, even with the so-called 'high quality' data in the DS data. They question most 'corrections' to the raw data proposed by

Deininger and Squire as they doubt that these adjustments really solve the comparability problems. In short they are “*not convinced that at present it is possible to use secondary data-sets safely without some knowledge of the underlying sources*” and “*caution strongly against mechanical use of such data-sets*” (Atkinson and Brandolini (1999), p.35). Put differently, it is best not to mix data from different sources. Notwithstanding this clear warning, most authors have done just that.

Besides the DS dataset, two other influential secondary datasets exist. The World Institute for Development Economics Research (WIDER) of the United Nations University and UNDP maintain an extended version of the DS dataset¹³. Finally the Luxembourg Income Study (LIS) collects income survey data of its 29 members. Currently it provides data for a period covering 1970-2000 for the ‘best’ countries. This corresponds to 7 or 8 data points. But for most countries one has to settle with a shorter time period and less data points. The main achievement of the LIS is that it ‘enforces’ a strictly defined methodological framework on its members to end up with comparable data¹⁴.

Another noteworthy initiative is the University of Texas Inequality Project (UTIP, Conceição and Galbraith (1998)). They have constructed long and dense time series of the Theil inequality index for over 150 countries over the period 1963-1998 based on the UNIDO Industrial Statistics. The advantages of this dataset are a very large coverage, consistency and accuracy. However, it is doubtful that the UTIP measure is a good proxy for income inequality. It is solely based on wages¹⁵ and earnings in the industrial sector. Not only is e.g., capital income neglected, but it also assumed that earnings inequality in the industrial sector is similar to earnings inequality in other sectors of the economy. This assumption seems untenable as earnings dispersion in the services sector is normally much wider. Especially for OECD countries, with economies dominated by the services sector, the UTIP data might be an inferior proxy for total inequality.

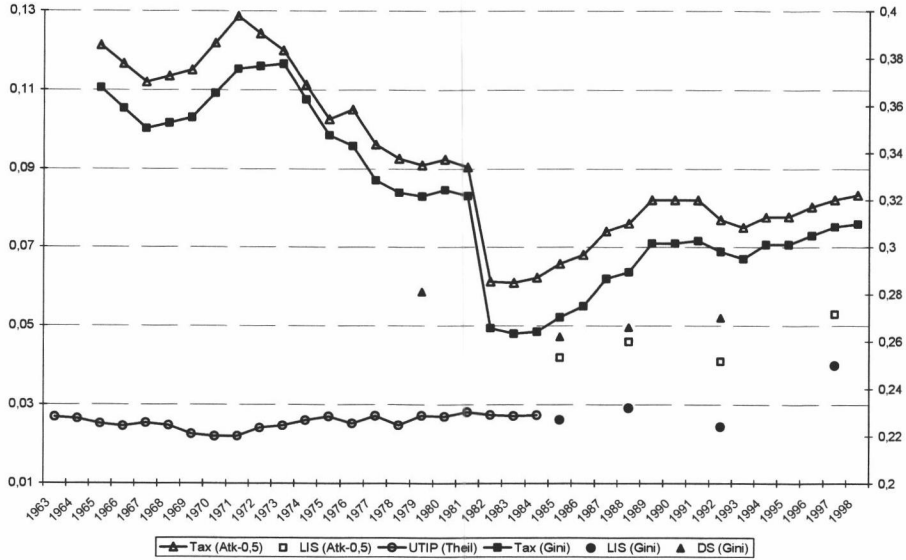
To illustrate the differences in the datasets we present some graphs for Belgium and the United States. For Belgium we look at the LIS data (2 measures), the DS data, national tax data (income after taxes, 2 measures) and the UTIP data. For the US we include the LIS data (2 measures), the US Bureau of the Census data (identical to the DS data, 3 measures) and the UTIP data.

¹³ For more details and access to the data: <http://www.wider.unu.edu/wiid/wiid.htm>

¹⁴ For more details and access to the data: <http://www.iisproject.org/keyfigures/methods.htm>

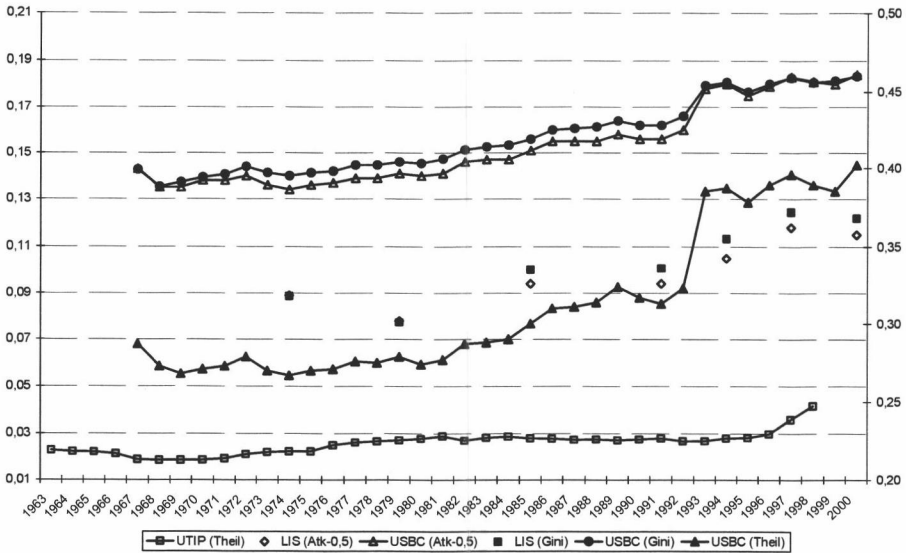
¹⁵ According to the Belgian Household Budget Survey (NIS (2003)) wages accounted for only 39,7% of the average disposable income of Belgian households in 2001. Non-wage economic activities accounted for an additional 18,1%. Social security payments and capital income accounted for respectively 29,7% and 14,7%. So while wages make up the largest component, the importance of other components can not be minimized.

Graph 1: Income inequality in Belgium – a look a different data sources



Notes: Gini coefficients are measured on the right axis (full marks), all other measures on the left axis (hollow marks).

Graph 2: Income inequality in the US – a look at different data sources



Notes: Gini coefficients and the USBC Theil index are measured on the right axis (full marks), all other measures on the left axis (hollow marks).

Graph 1 compares inequality data for Belgium. On the right axis (full marks) one can read the value for the different gini coefficients. On the left axis (hollow marks), the values for the Atkinson (with a parameter value of 0.5) and Theil index can be found. First note that the gini coefficients and Atkinson indices tell roughly the same story. Secondly, although we have very few data points, the evolution in inequality seems similar in all data series, except for the UTIP data. The correlation between the tax data and the UTIP data is -0.68 for the gini coefficient and -0.75 for the Atkinson index (but there are only 20 overlapping years)

Graph 2 plots the data for the United States. The value for the full marks are again on the right axis. Again only the UTIP data seem to change the global picture (although one would also conclude that inequality is rising). The correlation of the UTIP data and the USBC data is 0.75 for the gini coefficient, 0.71 for the Atkinson and 0.67 for the Theil index.

Although we can not deduce much from the figures, we nevertheless can conclude that the choice of database can affect the results of the analysis.

2.6. From cross-section to time series: a new secondary dataset on income inequality in OECD countries

As dense time series of inequality measures over longer time periods were unavailable, researchers have used cross-section estimates to determine the relationship between income inequality and economic growth. Recently, researchers also started exploiting the time dimension with the advent of panel data techniques. In these studies researchers combine inequality data from different primary data sources. However, this practice entails certain risks.

In the first part of this paper we have shown that inequality measures are sensitive to changes in equivalence scales and income concept. As long as the same income definition is used for all countries, the impact on the estimates of changing that definition might still be limited. However, mixing income concepts over countries will surely blur the results. Knowles (2001) provides additional support for this result. In cross-country growth regressions most authors have used the DS data on income inequality (see table 3). Although the DS dataset has the huge merit of having increased data availability, it does not solve the comparability problem. Atkinson and Brandolini state that *"in the dataset, the user is faced with a variety of different types of estimate"*. They sum up seven set of differences concerning income definitions, next to differences in data sources and processing. The LIS has strongly improved data comparability for the developed countries. But several problems remain as the underlying 'rough' data were collected for different purposes in different countries (see Gottschalk and Smeeding (2000)). This is reflected in differences in the survey set-up, which are bound to influence the survey results. Szekely and Hilgert (1999) show that previous estimation results based on survey data turn out to be quite sensitive to minor plausible corrections to the data. Tax systems differ even more across countries. Filer *et al.* (2004) conclude that *"the adjustments made to create cross-sectional comparability are complex and can seriously distort within country patterns over time"* (p. 14). Although they only investigate cross-country growth rates, their warning is universally aimed at all data that have been adjusted to increase cross-country comparability.

To control for differences in income definitions and data sources across countries, researchers typically introduce dummy variables in the regressions. But is this an effective solution? Atkinson and Brandolini (1999) do not believe that heterogeneity in inequality statistics can be eliminated by a simple additional or multiplicative adjustment. The differences between income concepts are likely to be country specific as a result of differences in government fiscal policies and tax incidence. Other studies have shown that the trend in the series of different income concepts within one country can differ substantially

(Ervik (1998), Brandolini, (1998) Atkinson and Brandolini (1999)) (cf. supra). A quick glance at the third panel of figure 3 shows that inequality of before and after tax income in Belgium has evolved differently over time. A dummy variable only takes care of the level differences. Thirdly, there is a possibility that a dummy variable captures an effect unrelated to the difference in income concept. This possibility is especially relevant in a context of growth regressions in which many effects are unaccounted for (the omitted variable bias is a well-known problem in growth regressions).

Given the comparability problems in a cross-section framework, it might be interesting to explore the possibilities of a time series approach. Although inequality measures within one country over time will also suffer from changes in the methodology, the transparency of the data series will be higher as one will probably know when and how the methodology has changed (although adjusting the data might still be difficult). Next to comparability, there are two other important advantages. Firstly, the cross-section approach bases its conclusions on a single observation. This observation might be recorded during a transitional phase or might be distorted due to some temporary influences. The importance of a single observation is lower in a time series framework and it is more likely that outliers will be spotted. Secondly, it is interesting in its own to get some insights into the time series properties of inequality series (e.g. deterministic or stochastic trends in the data).

But while cross-section estimates are not trouble-free, a time series approach seems even more troublesome (Parker (2000) and chapter 3). The datasets we described in the previous section (with the exception of the UTIP dataset) are only suited for cross-section or panel estimates, as very few data points are available (5 or 6 data points over a period of more than 30 years). Therefore one has to look for other data sources, most notably the statistical agencies, tax administrations or 'national inequality experts' of individual countries. This is a highly time consuming occupation with no a priori guarantee of success.

We have tried to accomplish the 'nearly-impossible' and have taken on the painstaking task of gathering a (limited) dataset for income inequality containing consistent annual time series. The data are not new (although some series were not publicly available). We invested a great deal of time in collecting as much data as possible on income inequality for all OECD countries. In the dataset we only include the countries for which we find at least 25 consecutive and consistent observations.

As we want an annual inequality measure over the longest possible time period, the gini coefficient turns out to be the only viable alternative. The one exception is France for which we include the income share of the 5% richest as a proxy for income inequality.

In our sample, the gini coefficients for the different countries are not all based on the same income concept, nor were they all collected in the same manner. Some gini coefficients were derived from census data, others were calculated out of income tax data. Some inequality measures concern disposable income, some after tax income, etc. However, this might not be a huge problem, as long as the user is aware of it and does not use the data in a panel regression context. This contrasts with past practices: users of the DS/LIS dataset assume that it is fully consistent although several authors have already illustrated that it is not (cf. supra). An advantage of our dataset is the consistency of the measures over time within the individual countries.

In the remaining part of this section we give a description of the dataset. Secondly, we look for some trends in income inequality in the OECD countries and finally, we check the time series characteristics of the data.

2.6.1. Description of the dataset

Nine OECD countries are included in the dataset. The choice of countries was forced upon us due to the limited data availability (we started with all OECD countries). Fortunately the sample seems to be instructive for the entire OECD. Firstly, with Canada, France, Italy, the UK and the US, 5 members of the G-7 are present as well as 4 smaller countries: Belgium, Finland, the Netherlands and Sweden. Secondly, the sample contains countries with an extensive social security system (Belgium, Finland and Sweden) and countries with a limited one (the UK and the US). Thirdly, also non-EU countries are present. However, it is certainly regrettable that we were not able to include Germany.

The series for Canada, Italy and Sweden are composite series. We have combined observations from different sources to obtain longer data series. For Canada and Italy we mixed the different series. For Sweden we used the growth rates of 2 other series to extrapolate (back- and forwards) a 'basic' series. Combining series might not be the most elegant nor the most prudent approach and potentially erodes the value added of this dataset compared to the 'conventional' ones. Remember that our main motivation to construct the dataset was the greater consistency of the data over time than across countries. Given the good correspondence of the overlapping values we believe that the series for Canada and Sweden¹⁶ remain reliable. The Italian series, however, should be looked at with a bit more prudence as we use 3 different series and the number of overlapping values is limited.

¹⁶ For Sweden the differences between the level of inequality in the overlapping years of the composite series (each time 3 years) remains fairly constant. For Canada the difference remains fully constant (0.02). Hence we added this difference to the second series.

Some missing values (Belgium, Italy, the Netherlands) were generated by linear interpolation.

Table 5 gives an overview of the data sources. We have also included some information on the countries that were not included as we did not find sufficient data.

Table 5: Data description

Country	Income concept	Source	Period
Included in dataset			
Belgium	Income before taxes	Valenduc – Tax data	1965 – 1998 4 years missing
	Income after taxes	Valenduc – Tax data	1965 – 1998 4 years missing
Canada	Disposable Household Income	Composite series: Checci and Statistics Canada – Census	1971 – 1999
Finland	Disposable Income	Erikson & Jäntti – Tax data	1960 – 1992
France	Income before taxes	Thomas Piketty (2001) – Tax data	1919 – 1998
Italy	Disposable Household Income	Composite series: Brandolini, Banca d'Italia and SHIW – Census	1967 – 1995 4 years missing
The Netherlands	Disposable Household Income	CBS – Census	1975 – 1999 4 years missing
Sweden	Net Household Income	Atkinson et al, Spant and Statistics Sweden – Census	1960 – 1999
The UK	Net Household Income	Goodman & Webb – Tax data	1961 – 1996
The US	Disposable Family Income	US Census Bureau – Census	1960 – 2000
	Disposable Household Income	US Census Bureau – Census	1967 – 2000
Not included in dataset (time series too short, less than 25 years)			
Austria	Gross Earnings	Gusenleitner <i>et al.</i> (1996) – Social security records	1972 – 1991
Denmark	Personal disposable income	Pedersen – 2 overlapping series	1975 – 1995
Germany	Net Household Income	Grabka / GSOEP – Census	1984 – 1991 (BRD) 1992 – 1999
Japan	Gross Household Income	Mizoguchi, Takayama	1962 – 1982 1 year missing
Not included in dataset (No time series found)			
Australia, Greece, Ireland, New-Zealand, Norway, Portugal, Spain, Switzerland			

For Belgium, the gini coefficients of income before and after taxes are strongly correlated (.96). For the US there exists a slightly shorter series for household incomes. The correlation between inequality of household and inequality of family income amounts to 0.97. Therefore we only look at the latter (i.e. the longest) series. Remember that we had to settle with a different inequality measure for France (income share of the richest 5% of the population).

2.6.2. Trends in income inequality

We have plotted the evolution of the inequality measures in our dataset to get an impression of the evolution of income inequality in the OECD countries over the last 3 decades. Note that it is risky to look at the absolute value of the inequality measures, as we use different definitions for different countries.

The graphs indicate that Canadian and French inequality has remained fairly stable (fluctuations within a 2.5%-points window).

In Finland inequality has steadily gone down between 1970 and 1992. In contrast, Swedish inequality has steadily increased.

Belgian inequality dropped substantially in the 1970s, but started to rise again in the early 1980s. The large decrease in 1982 can be explained by a different treatment of social benefits by the tax authorities. This change artificially inflated the taxable income of the beneficiaries, who are mostly located in the lower part of the income distribution. Hence, the drop of the gini coefficient does not reflect a real decrease of income inequality. Note that the Belgian tax system has a redistributive impact: the gini of after-tax-income is substantially lower than that of before-tax-income (up to 8%-point).

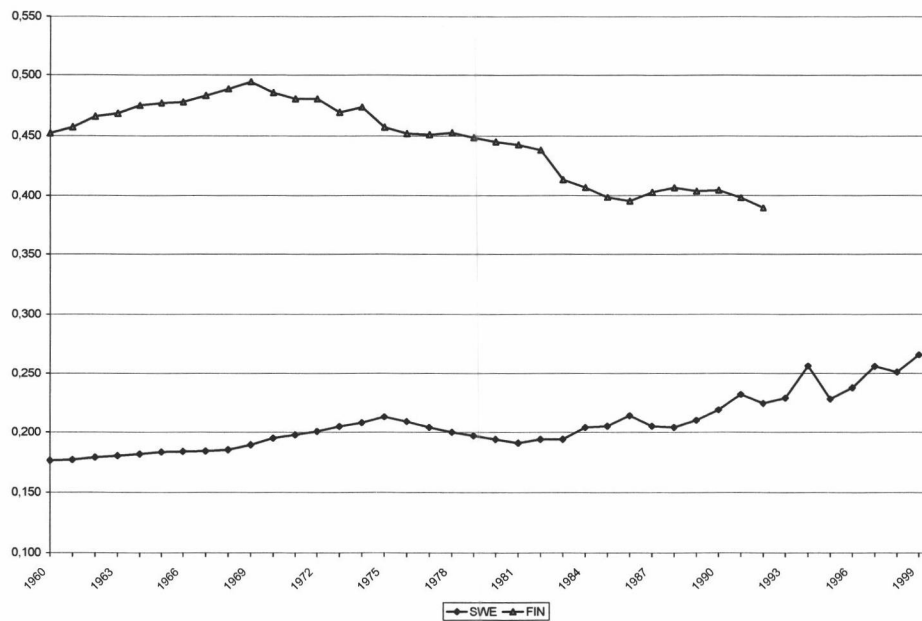
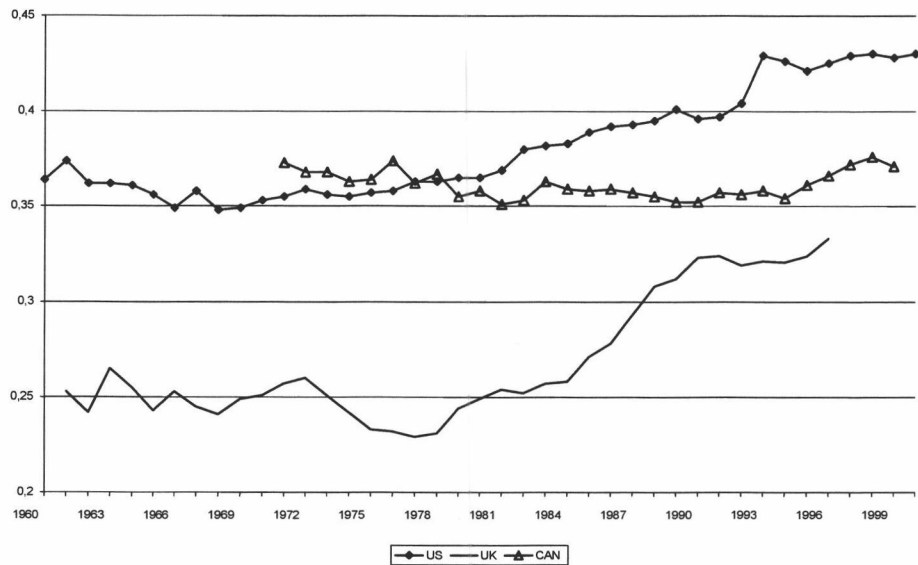
The large fluctuations in the first years of Dutch data seem to indicate that these are somewhat less reliable (although the fluctuations remain within a 4%-points window). From 1986 onwards inequality remains stable in the Netherlands.

Italian inequality dropped in the 1970s, and seems to fluctuate around a lower mean since the 1980s.

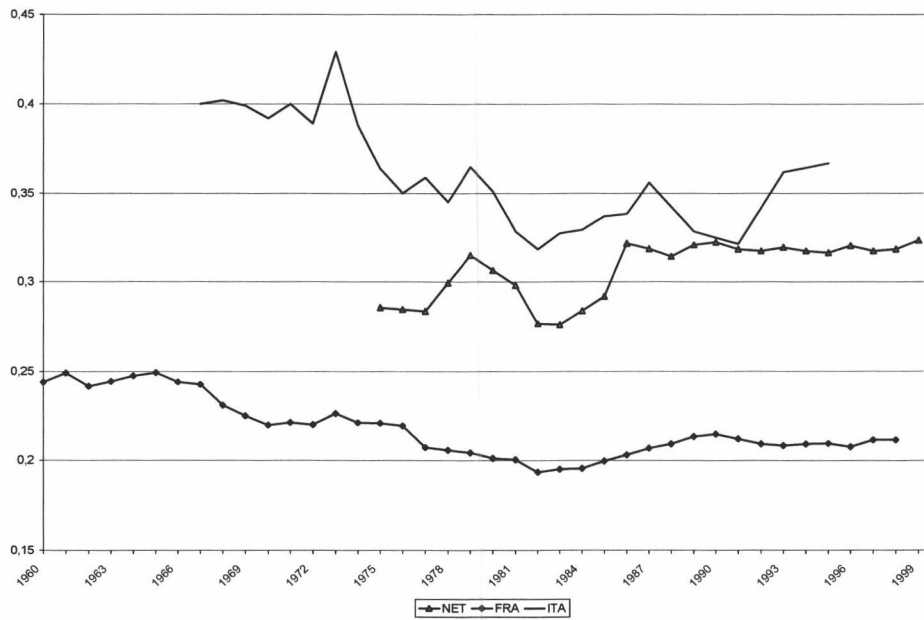
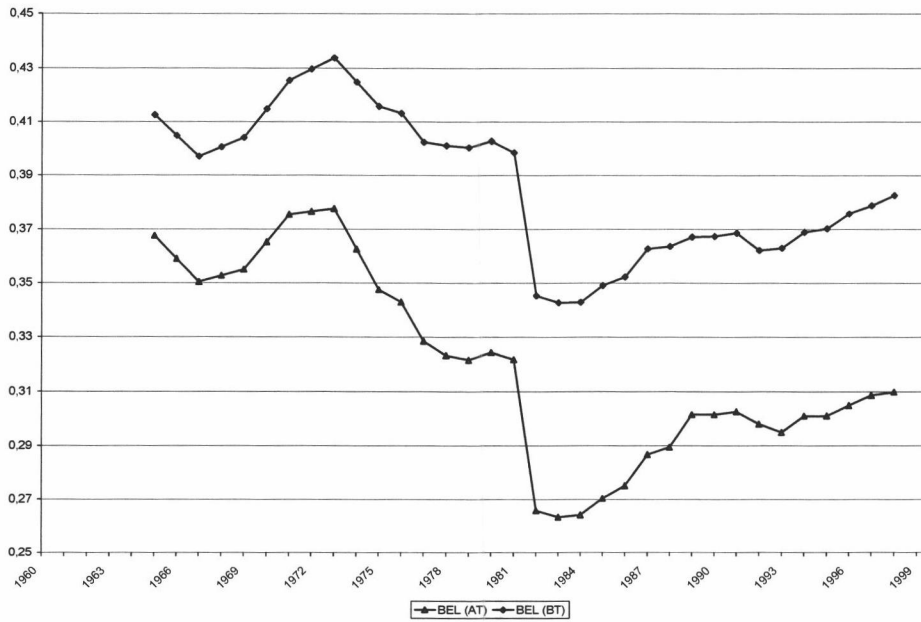
Inequality has strongly increased since the late 1970s / early 1980s in the UK and the US. In the 1960s and early 1970s inequality fluctuated around a constant mean. The trend shift roughly coincides with the start of the 'Thatcher era' in the UK (1979) and the 'Reagan reign' (1981) in the US.

In general we do not discern a global trend in income inequality in our sample of rich countries. On the contrary: the individual country experiences are quite diverse. This conclusion is similar to the one of other recent comparative studies (e.g. Brandolini (1998), Gottschalk and Smeeding (2000), Atkinson (2003)). This should not come as a surprise given the fact that the degree of income inequality in a country is determined by many factors including social and political forces as well as economic ones.

Graph 3: Income inequality in a number of OECD countries



Graph 3: Income inequality in a number of OECD countries (continued)



2.6.3. Limitations of the dataset

We have to admit that the dataset can not fully live up to our initial expectations: next to the fact that we needed to combine data from different data sources to obtain long enough time series, there is a methodological break in de Belgian data (cf. supra). Although we made sure that the consistency of the series was acceptable we can only subscribe the warning of Atkinson and Brandolini (1999) against the mechanical use of secondary datasets. Notwithstanding these critical remarks, the dataset is a useful extension to existing data sources as it enables the use of time series estimation techniques.

2.6.4. Time series characteristics of the data

In this section we check the statistic properties of the income inequality time series in our dataset. We present the results of 2 unit root tests: the Augmented Dickey Fuller test (ADF, null hypothesis: unit root) and the Kwiatkowski Phillips Schmidt Shin test (KPSS, null hypothesis: (trend) stationarity).

Note that *in theory* inequality measures cannot be non-stationary. By definition they are bounded and cannot rise or decline forever. However, over the limited time period considered in empirical work they can resemble a unit root process. In that case, one should handle them as such (Campbell and Perron (1991))¹⁷.

To test for the presence of a unit root it is important to use a regression that mimics the actual data-generating process. If not all deterministic regressors (constant, trend) are included the power of the ADF-test will drop substantially. The power will also be reduced if a regressor is inappropriately added (Enders, 1995). As we do not know the actual data-generating process we need to proceed with caution. We apply the testing procedure proposed by Enders (1995, p. 256-257). In the first step we use the least restrictive model formulation, with a constant and trend included. If we can reject the null hypothesis, there is no need to proceed (recall that misspecification will always *lower* the power of the test). Otherwise we test for the presence of a trend under the null of a unit root (using an F-test). If the trend is significant, we check for a unit root using the standard normal critical values¹⁸. Otherwise, we estimate the regression without a trend and repeat the above procedure with

¹⁷ The comment is equally valid for the analysis of unemployment rates (see e.g., Papell *et al.* (2000)). We circumvented the problem by testing for the presence of a unit root using logit-transformed inequality data, i.e. we transformed the gini coefficient into $\log[1/(1-\text{gini coefficient})]$. Since the gini coefficients is by definition between 0 and 1, the constructed variable varied between zero and plus infinity. The conclusions were unchanged.

¹⁸ As our sample is quite limited, we also looked at the Dickey-Fuller values (the procedure proposed by Enders is asymptotically valid). The conclusions were very similar.

respect to the inclusion of a constant term. The optimal lag length for the regression is determined by the Schwarz Bayesian Information Criterion (BIC).

In table 6 we only present the ADF-value for the relevant regression specification. Here we report the results for Belgian after tax inequality at length to illustrate the procedure.

Step 1: We do not reject a unit root if we allow for a constant and a trend in the specification (test statistic: -1.04, 5%-critical value: -3.50).

Step 2: Under the null of a unit root, the trend is not significant (test statistic: -0.18, 5%-critical value: -2.79). An F-test shows that we cannot reject the hypothesis that a unit root is present and the trend is insignificant (test statistic: 0.99, 5%-critical value: 6.73)).

Step 3: We re-estimate the model without a trend. Again a unit root can not be rejected (test statistic: -1.42, 5%-critical value: -2.93).

Step 4: Under the null of a unit root, the constant is not significant (test statistic: -1.56, 5% critical value: -2.54). An F-test shows that we cannot reject the hypothesis that a unit root is present and the constant is insignificant (test statistic: 1.37, 5%-critical value: 4.86).

Step 5: We use a regression specification without trend and constant. The results indicate that a unit root can not be rejected (test statistic: -0.99, 5%-critical value: -1.95). In summary table 6 we only report this final test statistic and the end conclusion of the test procedure.

As ADF-tests are known to have low power for highly persistent series (roots near unity), we check the results of the ADF-tests by means of the KPSS-test (Kwiatkowski et al., 1992). If level stationarity is not rejected at the 10%-level, we do not report the results for trend stationarity. If the conclusions of the ADF- and KPSS-test are mutually consistent, the trustworthiness of the conclusions increases.

The ADF-tests only reject a unit root in the levels for Finland. The KPSS-test always rejects level stationarity, except for Canada. For 5 countries trend stationarity is also rejected. The non-stationary evolution of inequality measures is also reported by Parker (2000).

If we look at the first differences (results not shown), the ADF-tests lead to a rejection of a unit root at the 1%-level for all countries but Italy (10%-level), the UK and the Netherlands (both at 5%-level). The KPSS-tests do not reject stationarity of the first differences, even if we look at the 10%-level, except for Finland and the US (both at 5%-level).

So while the two stationarity test do not always lead to exactly the same conclusion, level stationarity of inequality is clearly rejected. Almost all series seem to be characterised by either a stochastic or a deterministic trend.

Table 6: Unit root test for income inequality (levels)

Country	ADF	ADF (constant, trend)	KPSS (no trend)	KPSS (trend)
Belgium (after taxes)	-0.990 <i>Unit root is not rejected</i>		0.536** <i>Trend stationarity only rejected at 10%</i>	0.138*
Belgium (before taxes)	-0.566 <i>Unit root is not rejected</i>		0.499** <i>Trend stationarity not rejected</i>	0.109
Canada	-0.109 <i>Unit root is not rejected</i>		0.180 <i>Level stationarity not rejected</i>	0.168**
Finland		-2.70** <i>Reject unit root</i>	0.629** <i>Reject (trend) stationarity</i>	0.163**
France	-1.027 <i>Unit root is not rejected</i>		0.635** <i>Reject (trend) stationarity</i>	0.157**
Italy	-0.518 <i>Unit root is not rejected</i>		0.459* <i>Level stationarity only rejected at 10%</i>	0.164**
Netherlands	0.707 <i>Unit root is not rejected</i>		0.487** <i>Trend stationarity not rejected</i>	0.072
Sweden	2.237 <i>Unit root is not rejected</i>		0.775*** <i>Reject (trend) stationarity</i>	0.150**
UK	-1.777* <i>Unit root is only rejected at 10%</i>		0.594** <i>Reject (trend) stationarity</i>	0.193**
US	-0.716 <i>Unit root is not rejected</i>		0.802*** <i>Reject (trend) stationarity</i>	0.216***

Notes: Augmented Dickey Fuller test (ADF): null hypothesis is unit root
 Kwiatkowski Phillips Schmidt Sin Test (KPSS): null hypothesis is (trend) stationarity
 * / ** / *** denotes rejection of null hypothesis at 10% / 5% / 1% level

In some of the income inequality series (the US, 1993), there occurs a break in the data due to a *known* change in the data gathering methodology. For these series we adjust the series themselves, as a correction is quite straightforward. However, for other series there might be a break caused by a change in the methodology unknown to us, or caused by 'other factors'. For instance, the election of Thatcher seems to have had a clear impact on the trend in income inequality in the UK. Based on a quick glance at the figures, we can further suspect a possible level break combined with a changing trend in the Belgian inequality series. Contrary to the breaks due to a changing methodology, breaks caused by 'other factors' should not be eliminated. However, as illustrated by Perron (1989), we should take these breaks into explicit account when we test for unit roots. For instance, unit root tests could wrongly identify a level stationary process with a level shift as a unit root process. Therefore, we performed some additional tests, based on the Perron methodology (Enders (1995), p.249; results not shown). The results indicate that stationarity is not restored once level and trend shifts are taken into account.

These findings shed a new light on existing work. Early work mainly focused on cross-section data. The time series characteristics of the data were never closely examined. However, as a panel data approach gains way, the time dimension becomes more important. It has been shown that one does not need to restrict the dynamic behaviour of the data in a panel set-up in case one has a very large cross-section dimension (N) and a small and fixed time dimension (T). On the other hand, if T is large relative to N , the 'normal' asymptotics may be misleading (Wooldridge (2002), p.175). This is indeed the case in existing empirical work: e.g., Forbes (2000) looks at 45 countries over a period of 25 years. So it seems that non-stationarity has been dismissed somewhat too easily as a potential problem in the panel data approach.

2.7. Conclusions

As in any other field of applied macro-economic or econometric research, researchers that study income inequality have to look for suitable data. Although most researchers just draw on some ready-made dataset, finding reliable data is not that straightforward and can even be very troublesome. Firstly, one has to define which income flows matter. Secondly, one has to decide on the best one-dimensional representation of an entire income distribution (which inequality measure is preferable?). Economic theory can often provide guidance. But next follows the confrontation with a very grim reality: the desired data are almost never available. One has to settle with what is available and that is not a lot in the case of inequality data.

Unfortunately, the choices really matter. Mixing inequality data based on different equivalence scales, different income concepts, etc. will blur the outcome of any econometric analysis. As a consequence of the data scarcity, researchers have predominantly resorted to a cross-section estimation techniques. However data comparability across countries is sometimes low. As we a priori believed that the comparability of inequality data within one country might be better, we gathered a new dataset with an emphasis on the time dimension for income inequality for a group of OECD countries. Although we could not obtain 100% consistency, we nonetheless believe that the new dataset can be a useful tool for future empirical work. A visual inspection of the time series shows that inequality has evolved quite differently in the included countries over the past decades. More robust econometric testing shows that level stationarity does not match the behaviour of the inequality series in our dataset. While the result that all series are characterised by either a stochastic or a deterministic trend is noteworthy in itself, it is also of importance for econometric research.

2.8. References

- Aghion P., Caroli E., Garcia-Peñalosa C. (1999): *Inequality and Economic Growth: The Perspective of the New Growth Theories*, Journal of Economic Literature, 37(4), p.1615-1660
- Aitchison J., Brown J.A.C. (1973): *The Lognormal Distribution (4th Edition)*, Cambridge University Press, Cambridge
- Alesina A., Rodrik D. (1994): *Distributive Politics and Economic Growth*, Quarterly Journal of Economics, 109, p. 465-490
- Arjona R., Ladaïque M., Pearson M. (2001): *Growth, Inequality and Social Protection*, OECD Labour Market and Social Policy Occasional Paper no. 51, 80p.
- Atkinson A.B. (1970): *On the Measurement of Inequality*, Journal of Economic Theory, 2, p. 244-263
- Atkinson A.B. (2003): *Income Inequality in OECD Countries: Data and Explanations*, CESifo Working Paper no. 881, 52p.
- Atkinson A.B., Rainwater L., Smeeding T.M. (1995): *Income Distribution in OECD Countries*, OECD Social Policy Studies no. 18, 164p.
- Atkinson A.B., Brandolini A. (1999) : *Promise and Pitfalls in the Use of "Secondary" Datasets: Income Inequality in OECD countries*, mimeo, 60p.
- Atkinson A.B., Bourguignon F. (eds.) (2000) : *Handbook of Income Distribution Vol. 1*, Elsevier, Amsterdam
- Barro R.J. (1999): *Inequality, Growth and Investment*, NBER Working Paper no. 7038, 37p.
- Berry A., Serieux J. (2003): *All about the Giants: Probing the Influences on Growth and Income Inequality at the end of the 20th Century*, CESifo Working Paper no. 840, 36p.
- Brandolini A. (1998): *Pareto's Law and the Kuznets' Curve: A Bird's-Eye View of Long-Run Changes in Income Inequality*, mimeo, 46p.
- Breunig R. (2001): *An Almost Unbiased Estimator of the Coefficient of Variation*, Economics Letters, 70, p.15-19
- Breunig R. (2003): *Bias Correction for Inequality Measures: An Application to China and Kenya*, mimeo, 11p.
- Browning M., Chiappori P.A., Lewbel A. (2004): *Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power*, mimeo (January), 55p.
- Buhmann B., Rainwater L., Schmaus G., Smeeding T.M. (1988): *Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates across ten Countries using the Luxembourg Income Study (LIS) Database*, Review of Income and Wealth, 34(2), p. 115-142
- Caminada K., Goudswaard K. (2000): *International Trends in Income Inequality and Social Policy*, paper presented at the 'Year 2000 International Research Conference on Social Security', 25-27 September, Helsinki

- Campbell J.Y., Perron P. (1991): *Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots*, NBER Technical Working Paper no. 100, 67p.
- Castelló A., Doménech R. (2002): *Human Capital and Economic Growth: Some New Evidence*, The Economic Journal, 112, p.187-200
- Conceição P., Galbraith J.K. (1998): *Constructing Long and Dense Time-Series of Inequality Using the Theil Index*, UTIP Working Paper no. 1, 21p.
- Cowell F. (1989): *Sampling Variance and Decomposable Inequality Measures*, Journal of Econometrics, 42, p.27-41
- Cowell F. (2000): *Measurement of Inequality*, in: Handbook of Income Distribution Vol. 1 (eds. Atkinson A.B., Bourguignon F.), Elsevier, Amsterdam, p.87-166
- Cowell F. Mercader-Prats M. (1999): *Equivalence Scales and Inequality*, in: Handbook on Income Inequality Measurement (ed. Silber J.), Kluwer Academic Publishers, Boston / Dordrecht / London, p.405-429
- Cowell F., Flachaire E. (2002): *Sensitivity of Inequality Measures to Extreme Values*, DARP Discussion Paper no. 60, 18p.
- Cowell F., Flachaire E. (2004): *Income Distribution and Inequality Measurement: The Problem of Extreme Values*, mimeo, 24p.
- Cowell F.A., Victoria-Feser M.-P. (1996): *Robustness Properties of Inequality Measures*, Econometrica, 64 (1), p.77-101
- Deaton A. (2003): *Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)*, NBER Working Paper no. 9822, 50p.
- Deiningner K., Squire L. (1996): *Measuring Income Inequality: a New Data-Base*, World Bank Economic Review, 10 (3), p.565-591
- Deiningner K., Squire L. (1998): *New Ways of Looking at Old Issues: Inequality and Growth*, Journal of Development Economics, 57, p.259-287
- Deiningner K., Olinto P. (2000): *Asset Distribution, Inequality and Growth*, mimeo, 28p.
- Enders W. (1995): *Applied Econometric Time Series*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons Inc., New York
- Ervik R. (1998): *The Redistributive Aim of Social Policy. A Comparative Analysis of Taxes, Tax Expenditure and Direct Transfers in 8 Countries*, LIS Working Paper no. 184
- Fields G. (1989): *A Compendium of Data on Inequality and Poverty for the Developing World*, Cornell University, Unpublished Manuscript
- Figini P. (2000): *Measuring Inequality: On the Correlation Between Indices*, LIS Working Paper no. 229, 13p.
- Filer R.K., Hajkova D., Hanousek J. (2004): *The Other Side of the Moon: the Data Problem in Analysing Growth Determinants*, CEPR Discussion Paper no. 4349, 24p.

Flora P., Kraus F., Pfenning W. (1987): *State, Economy and Society in Western Europe 1815-1975: A Data Handbook*, Vol. 2, Frankfurt, Campus Verlag

Flückiger Y. (1999): *A Comment on 'Equivalence Scales and Inequality'*, in: *Handbook on Income Inequality Measurement* (ed. Silber J.), Kluwer Academic Publishers, Boston / Dordrecht / London, p.429-435

Forbes K.J. (2000): *A Reassessment of the Relationship Between Inequality and Growth*, *The American Economic Review*, 90(4), p.869-887

Förster M. (2000): *Trends and Driving Factors in Income Distribution and Poverty in the OECD Area*, *Labour Market and Social Policy Occasional Paper no. 42*, OECD

Gibrat R. (1931): *Les Inégalités Economiques*, Sirey, Paris, 296p.

Goodman A. (2001): *Inequality and Living Standards in Great Britain: Some Facts*, *Briefing Note no. 19*, The Institute of Fiscal Studies, 24p.

Gottshalk P., Smeeding T.M. (2000): *Empirical Evidence on Income Inequality in Industrialized Countries*, in: *Handbook of Income Distribution Vol. 1* (eds. Atkinson A.B., Bourguignon F.), Elsevier, Amsterdam, p.261-307

Gyimah-Brempong K., Wilson M. (2004): *Health Human Capital and Economic Growth in Sub-Saharan African and OECD Countries*, *The Quarterly Review of Economics and Finance*, 44, p.296-320

Gusenleitner M., Winter-Ebmer R., Zweimüller J. (1996): *The Distribution of Earnings in Austria, 1972-1991*, *Linz University Working Paper 9614*, 14p.

Gustafsson B., Johansson M. (1997): *In Search For a Smoking Gun: What Makes Income Inequality Vary Over Time in Different Countries*, mimeo, 35p.

Hartog J., Veenbergen J.G. (1978): *Dutch Treat: Long-Run Changes in Personal Income Distribution*, *De Economist*, 126(4), p.521-549

Jain S. (1975): *The Size Distribution of Income: a Compilation of Data*, World Bank, Washington D.C.

Jenkins S.P. and Van Kerm P. (2003): *Trends in income inequality, pro-poor income growth and income mobility*, *DIW Discussion Paper no. 377*, 18p.

Kalecki M. (1945): *On the Gibrat Distribution*, *Econometrica*, 13(2), p.161-170

Knowles S. (2001): *Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data*, *CREDIT Research Paper no. 01/03*, 32p.

Kolm S.C. (1969): *The Optimal Production of Social Justice*, in: *Public Economics* (eds. Margolis J., Guitton H.), MacMillan, New York/London

Kuznets S. (1963): *Quantitative Aspects of Economic Growth of Nations: VIII Distribution of Income by Size*, *Economic Development and Cultural Change*, 11

Lindert P.H., Williamson J.G. (1985): *Growth, Inequality, and History*, *Explorations in Economic History*, 22(4), p. 341-377

Milanovic B. (2002): *The Ricardian Vice: Why Sala-i-Martin's Calculations of World Income Inequality are Wrong*, mimeo, 34p.

NIS (2003): *Huishoudbudgetonderzoek 2001 – Uitgaven en Inkomsten*, <URL:http://www.statbel.fgov.be/downloads/hbs2001_nl.xls>, 4 December

NIS (2004): *Bevolking en huishoudens – Huishoudens en familiekeren*, 372p.

Papatheodorou C., Peristera P., Kostaki A. (2003): *Assessing Income Distribution using Kernel Estimates: A Comparative Study in Five European Countries*, paper presented at the '17th Annual Congress of the European Society for Population Economics', 13-15 June, New York City

Parker S.C. (2000): *Opening a Can of Worms: the Pitfalls of Time-Series Regression Analyses of Income Inequality*, *Applied Economics*, 32, p.221-230

Papell D.H., Murray C.J., Ghiblawi H. (2000): *The Structure of Unemployment*, *The Review of Economics and Statistics*, 82 (2), p.309-315

Partridge M.D. (1997): *Is inequality Harmful for Growth? Comment*, *The American Economic Review*, 87(5), p. 1019-1032

Paukert F. (1973): *Income Distribution at Different Levels of Development: a Survey of the Evidence*, *International Labour Review*, 108(2-3), p.97-125

Perotti R. (1994): *Income Distribution and Growth: an Empirical Investigation*, mimeo, Columbia University

Perotti R. (1996): *Growth, Income Distribution and Democracy: What the Data Say*, *Journal of Economic Growth*, 1(2), p.149-187

Perron (1989): *The Great Crash, the Oil Price Shock and the Unit Root Hypothesis*, *Econometrica*, 57 (6), p.1361-1401

Persson T., Tabellini G. (1994): *Is Inequality Harmful for Growth*, *The American Economic Review*, 84(3), p.600-621

Piketty T. (2001): *Income inequality in France 1901-98*, CEPR Discussion Paper no. 2876, 33p.

Rehme G. (2002): *(Re-)Distribution of Personal Incomes, Education and Economic Performance Across Countries*, CESifo Working Paper no. 711(5), 45p.

Silber J. (ed.) (1999): *Handbook on Income Inequality Measurement*, Kluwer Academic Publishers, Boston / Dordrecht / London

Sala-i-Martin X. (2002): *The Disturbing "Rise" of Global Income Inequality*, mimeo, 71p.

Sen A. (1973): *On Economic Inequality – The Radcliffe Lectures Delivered in the University of Warwick 1972*, Clarendon Press, Oxford

Smith J.P. (2001): *Why is Wealth Inequality Rising*, in: *The Causes and Consequences of Increasing Inequality* (ed. Welch F.), University of Chicago Press, Chicago, p.88-116

Szekely M., Hilgert M. (1999): *What's Behind the Inequality we Measure? An Investigation Using Latin American Data*, LIS Working Paper no. 234, 51p.

Taylor C.L., Hudson M.C. (1972): *World Handbook of Political and Social Indicators*, 2nd Edition, Yale University Press, New Haven

Thistle P.D. (1990): *Large Sample Properties of Two Inequality Indices*, *Econometrica*, 58 (2), p.725-728

U.S. Department of Commerce (1975): *Historical Statistics of the United States*, Washington D.C., US Government Printing Office

Wooldridge J.M. (2002): *Econometric Analysis of Cross-section and Panel Data*, The MIT Press, Cambridge / London

Chapter 3:

Inequality and Growth: Does Time Change Anything?

Abstract

The econometric analysis of economic growth has always been subject to major flaws and shortcomings. Data scarcity and reliability, parameter heterogeneity, omitted variable bias, endogeneity problems, ... have seriously tainted estimation results. In this paper we propose an alternative framework that explicitly deals with these issues. We investigate the relation between income inequality and economic growth in a number of OECD countries in a cointegrated VAR-setting. Our results suggest that different models might hold for different countries. However, all things considered, the imperfect markets model better describes reality than the complete markets model.

JEL Classification: E62, N10, O11, O15

We would like to thank Bas van Aarle, David de la Croix, Rafael Doménech, Gerdie Everaert, Freddy Heylen, Dirk Van de gaer, Jozef Plasmans and the participants to the '55th IAES Conference', the 'Euresco Conference on institutions and inequality' and the Pisa Conference on 'Economic Growth and Distribution' for useful comments and suggestions. We would also like to thank Christian Valenduc, Markus Jäntti, Mats Haglund, Antonio Brandolini, Wim Kessels and all others that helped with the dataset. We acknowledge support from the Federal Public Planning Service Science Policy, Interuniversity Attraction Poles Program - Belgian Science Policy [Contract No. P5/21]

3.1. Introduction

Until the mid-1970s the Kuznets curve (Kuznets (1955)) was accepted as an empirical stylised fact. The Kuznets curve describes the relationship between income inequality and per capita income. The relationship between growth and inequality under the Kuznets hypothesis depends on the level of per capita income: once a certain threshold has been crossed (once societies are rich enough), further growth will reduce inequality. In the late 1970s the Kuznets picture was disturbed by a sudden increase (the UK, Germany) or a stagnation (France, Canada) in inequality in some rich countries. These events questioned the universal validity of the Kuznets curve and gave way to a new development in the empirical literature on inequality and economic growth. Although some authors remained faithful to the Kuznets idea and suggested modifications to the basic framework (o.a. Milanovic (1994)), others have questioned the causal linkages between economic growth and inequality. Ever since a closer look at data in the mid-1980s had shown that more inequality was always associated with lower long-run growth, the belief that inequality, rather than growth, is the determining factor in the relationship made way.

The question whether and how inequality is related to economic growth inspired a lot of empirical research over the past decade. In the early 1990s, several authors showed that higher inequality at the beginning of a longer-term period was linked to poor growth (Alesina and Rodrik (1994), Perotti (1994, 1996), Persson and Tabellini (1994)). This resulted in a consensus that inequality worsens growth performances. Gradually the consensus weakened. First, it was argued that the relationship differs between poor and rich countries (Deininger and Squire (1998), Barro (1999)). A negative relationship was found in developing countries, but for richer countries there was no relation at all. Recently a new consensus with a very different content seems to take shape: inequality stimulates economic growth (Forbes (2000), Arjona *et al.* (2001)). Given this evolution, one might conclude that the world has changed drastically over the past decade thereby disturbing well-established economic relations.

However, a closer look shows that not the world but rather the econometric techniques to analyse it have been the subject of major changes. The earliest empirical contribution presented OLS and 2SLS estimates. Next, 3SLS and random and fixed effects panel estimators were used. These were in turn replaced by panel GMM estimators. Still, Durlauf (2001) notes that '*while we have seen remarkable advances in the econometric analysis of many areas of microeconomics and macroeconomics, growth economics has not experienced anything close to such progress*' (p. 65).

We explore the possibilities of a new econometric approach. A short overview of past contributions in the next section clarifies Durlauf's provocative statement and motivates our alternative. In this paper, we analyse the relationship between income inequality and economic growth using a cointegration approach within a VAR model.

The remainder of the paper is structured as follows. We motivate our econometric approach in section 3.2. Section 3.3 looks into two theoretical models of income inequality and economic growth: the complete markets model and the imperfect markets model. A description of the dataset follows in section 3.4. In section 3.5 the results of the cointegration analysis are reported. We try to identify the long run relations and check whether these relations allow us to discriminate between the models presented in section 3.3. Section 3.6 summarizes the most important insights of the paper.

3.2. The econometrics of inequality and growth

There has been a substantial evolution in growth econometrics since the beginning of the 1990s. In the bulk of early empirical work on growth a linear cross-country regression is estimated by means of OLS. Growth is regressed on several (lagged) explanatory variables as GDP, schooling, ... and income inequality (Barro and Sala-I-Martin (1995)). Several shortcomings of this approach have been noted, varying from an omitted variable bias and data heterogeneity to endogeneity.

If an important determinant is omitted in a regression specification, the estimated coefficients of all included variables will be biased. This omitted variable bias is problematic in growth empirics, since the list of factors that can plausibly affect growth seems without limit. Durlauf (2003) discriminates between two kinds of regressors in traditional cross-country regressions: the ones offered by the Solow growth model (population growth, technological change, physical and human capital and savings rates) and those added by the new growth theories. While the former list is fixed, no consensus about the latter exists. As the number of data points available for growth estimates is not that large, a lot of potentially relevant regressors need to be excluded. The omitted variable problem seems hard to overcome as modern growth theories are fundamentally open-ended: one growth theory typically has no bearing on the empirical relevance of another.

Durlauf and Johnson (1995), Canova (1999), Krueger and Lindahl (2000), Kourtellos (2002) and Sonedda (2003) show that the assumption of parameter homogeneity in standard growth analyses is neither supported by the data, nor by theory. Durlauf (2001) argues that *'there is*

nothing in growth theory which would lead one to think that the marginal effect of a change in high school enrolment percentages on the per capita growth of the US should be the same as the effect on a country in sub-Saharan Africa' (p.67). He agrees that this argument is generally applicable in econometrics, but as any parsimonious growth regression will necessarily leave out many factors that from the perspective of economic theory affect the parameters of the included variables, it is particularly salient in the case of cross-country growth. The different 'income inequality – growth' relationship between richer and poorer countries (Deininger and Squire (1998), Barro (1999, 2000)) also illustrates that enforcing one overall relation will necessarily induce poor estimation results¹.

An important element in the heterogeneity debate, is the lower data quality in developing countries (Schultz (1999)). One can imagine that lower income countries systematically underreport inequality due to a failure in data collection. If this is true, low inequality will be linked to a poor growth on the basis of a measurement error. Weede (1997) illustrates the importance of the data selection: he shows that the results of Persson and Tabellini (1994) are neither robust to different data sources, nor to the exclusion of a few original data points. Quah (2000) notes that *'researchers have long known about the biases and omissions in developing-country national income accounts. Comparison of those data with the data of developed countries can be unreliable even when within-country analysis over time for a given economy is perfectly sensible' (p. 5).*

Given the heterogeneity problems, we choose to focus on individual OECD countries. Measurement error is probably less severe in those countries. Moreover it will enable us to evaluate the believed homogeneity within this group of richer countries. Our 'individual country approach' is also compatible with another comment by Durlauf (2001): *'Empirical growth studies virtually always assume that one theory is equally valid for all countries, whereas it is far more natural to think a given theory will explain the growth experience of each country more or less well depending on the country's individual characteristics' (p.69).* Evans (1998) shows that different growth models may characterise the growth experiences of well-educated and poorly-educated countries. Park (1994) notes that although the suggestive empirical results established in the cross-country analyses can provide a useful guide for country studies, the challenge of empirical work is testing the theoretical insights against the economic evolution of individual countries using time series data. Relevant country specific information gets lost amidst the large number of factors affecting growth performance in cross-country studies. Brock and Durlauf (2000) argue that theory and parameter heterogeneity (uncertainty) is of major importance in a policy-relevant empirical analysis of growth.

¹ However, Bleaney and Nishiyama (2004) do not find evidence that the sign of the initial income inequality coefficient differs between rich and poor countries in cross-country regressions.

Endogeneity is a major problem in growth regressions. One can quite easily argue that education stimulates growth, but one can as easily explain why growth influences education decisions. A lot of variables have an impact on the growth performance, but growth in turn influences almost all economic decisions. Using an instrumental variable approach to deal with endogeneity is not straightforward as it is problematic to identify instruments that simultaneously are correlated with the included growth determinants and uncorrelated with the residuals. Durlauf (2001) notes that *'those studies which attempt to use instrumental variables to address regressor endogeneity have not been persuasive in that the choices of instruments have not met the necessary exogeneity requirements for instrument validity'* (p.66). A valid instrument has to be uncorrelated with all growth theories not embodied by the regression. But because so many factors can plausibly influence growth, this condition is virtually impossible to satisfy.

A frequently used alternative solution, explaining subsequent growth by including explanatory variables at the beginning of the period, does not fully solve the endogeneity problem either, as expectations about economic growth will also matter in the decision-making process with respect to schooling, investment, ...

A panel data estimation by fixed effects reduces the omitted variable problem as the country specific factors that are fixed over time are eliminated (Arjano *et al.* (2001), Forbes (2000)). Moreover, it becomes possible to evaluate the effects of changes in inequality, which is more relevant from a policy point of view. But a new problem arises as estimation by fixed or random effects can be biased if the specification that needs to be estimated contains a lagged endogeneous variable (Nickell (1981)). Growth is known to be characterised by a catch-up effect (conditional convergence). Given the dimensions of the panels that are typically being used (e.g., Forbes (2000) uses 6 observations over a period of 25 years for 45 countries), the convergence effect renders the fixed and random effects estimators useless for this kind of analysis. Besides this additional problem, estimation by random or fixed effects does not deal with the endogeneity issue.

An alternative is the use of a first-differenced GMM estimator (the Arellano-Bond estimator) (Forbes (2000)). This kind of estimator eliminates the country specific effects, but the need to identify the appropriate instruments remains. If the first stage relationship between differenced independent variables and lagged level variables is weak, the GMM estimates will be biased towards their fixed-effects counterparts (Stock *et al.* (2002)). In addition, Blundell *et al.* (2000) show that the instruments used in the first-differenced GMM become less informative with series that are highly autoregressive. As inequality series are characterised by a high degree of persistence, there is a substantial risk that the GMM-results have a large finite sample bias. The most fundamental criticism on the use of the

Arellano-Bond estimator is that it is designed for micro data sets, i.e. for a cross section dimension that tends to infinity (Bond (2002)). This condition is gravely violated in the context of growth econometrics.

Given the above problems, how should one proceed? We choose not to resort to the 'classic' cross-country approach but focus on the time dimension in the data. This eliminates problems related to parameter and theory heterogeneity. Given the major endogeneity problems in growth econometrics, a VAR model seems to be a suitable framework. A VAR model also steers clear of a priori restrictions (with respect to stationarity, causality, ...) on the estimates. One exception is the assumption of a linear relationship between the different variables. This assumption is not undisputed (Banerjee and Duflo (2003), Krueger and Lindahl (2000)). As we only include OECD countries in our study, the dispersion of most variables is quite limited. Given this limited range, the linearity assumption seems less controversial.

Hauk and Wacziarg (2004) choose a different approach to the econometric difficulties in growth regressions. Instead of reducing the biases in estimates by improving on the methodology, they evaluate the bias properties (e.g., size and direction) of common estimators in growth regressions.

3.3. The models

We want to discriminate between two models of growth and inequality: the complete markets model (CMM) and the imperfect markets model (IMM) (Perotti (1996)). Next to these models, Perotti (1996) also considers the socio-political instability and the fertility model. We do not integrate these models in our analysis, as we a priori consider them to be less relevant for a sample of OECD countries.

In the CMM each economic agent can fully borrow against the present discounted value of future earnings. High inequality affects investment decisions, as a higher government intervention (more redistributive measures) will be demanded by the population. Redistribution reduces growth through tax distortions and reduced capital accumulation. In the IMM not all planned investment (especially in human capital) can be executed as the poor are credit constrained. Redistribution to the poor relaxes the credit constraint thereby stimulating growth and investment. At the same time redistribution has similar negative effects as in the complete markets model. Although imperfections in the credit market will be especially severe in developing countries, we believe that the IMM has relevance for a sample of more advanced countries. Cameron and Taber (2004) find no evidence of

educational borrowing constraints in the US. But they add that *"their results do not imply that credit market constraints would not exist in the absence of the assortment of private and government programs currently available"* (p. 180). Blöndal et al. (2002) argue that in the OECD *"in the absence of government intervention, investment in human capital is difficult to finance through unsecured personal loans"* (p. 53). By means of micro-economic income data for a number of OECD countries Headey and Muffels (2002) illustrate that poor people invest more in human capital in more generous welfare states.

Both models predict a negative relation between growth and inequality. However the underlying mechanisms that result in this reduced form are different.

We elaborate some more on the testable implications of both theories by means of a basic theoretical model inspired by Bénabou (1996) and Aghion and Howitt (1998). The models we present are deliberately kept simple which implies that some stringent assumptions have to be made. The sole aim of the theoretical elaboration is to motivate the empirical section of the paper. It seems needless to pursue a more 'sophisticated' approach. That is also the rationale for presenting two models that differ with respect to the redistribution system: integrating both systems in one model is possible, but severely complicates the discussion and has little added value. A different theoretical elaboration on both models is provided by Sonedda (2003).

The full mathematical derivations of some expressions can be found in appendix A.

3.3.1. Set up

We consider an overlapping generations model in which n individuals live for two periods. The intertemporal utility, U , of an individual i born at time t is given by

$$U_i^t = \ln c_i^t + \rho \ln d_i^t \quad 0 < \rho < 1 \quad (1)$$

where c^t and d^t denote current and future consumption respectively. The parameter ρ is a measure of time preference. There is only one good in the economy that serves both as capital and consumption good. Production of the future consumption good takes place at time t according to an AK technology

$$y_i^t = \eta k_i^t \quad (2)$$

The parameter η is an efficiency measure, k^i is an individual's education level. In line with the work by Lindbeck (1985, 1988, 1993) and Davis and Henrekson (2004) we assume that efficiency is a decreasing function of the tax rate β . For simplicity and without loss of generality we impose an efficiency loss that is proportional to the tax rate,

$$\eta = (1 - \kappa\beta)\eta^* \quad (3)$$

with η^* the maximum efficiency (with zero tax rate)

$$0 < \kappa \leq 1$$

Using (3) we can rewrite the production function (2) as

$$y_i^j = (1 - \kappa\beta)\eta^* k_i^j \quad (2')$$

An individual's education level is both determined by his own effort and by the basic level of knowledge and skills in the society.

$$k_i^j = (e_i^j)^\delta (A_i)^{1-\alpha} \quad \text{with } 0 < \alpha < 1 \quad (4)$$

where e^i denotes the education level attained by an individual i as a result of his own investment in human capital and A is the basic level of knowledge and skills in the society. To increase his education level, an individual can invest in human capital (h^i). Human capital investment is characterised by decreasing returns. For highly educated people it will take more time, money and effort to further increase their education level (e.g., higher information costs, less qualified teachers, etc.). Bils and Klenow (2000) also require that the human-capital returns to schooling exhibit diminishing returns. Psacharopoulos (1994) provides empirical support for this assumption.

$$e_i^j = (h_i^j)^\gamma \quad \text{with } 0 < \gamma = \frac{\alpha}{\delta} < 1 \quad (5)$$

The accumulation of knowledge and skills follows from past production activities (i.e. the economy is characterized by a learning-by-doing process).

$$A_t = \frac{1}{n} \sum_i y_{t-1}^i \quad (6)$$

We implicitly assume that the link between knowledge and former production is one-to-one. If we allow for depreciation, i.e. only part of average past production results in current knowledge, the main results of the model do not change.

Individuals differ in their initial endowments. An individual's endowment upon birth at time t is given by $\varepsilon_i^j A_t$, with $\varepsilon^j (\geq 0)$ an identically and independently distributed random shock with mean 1 that measures individual i 's access to general knowledge at time of birth.

Individual i can directly 'consume' his initial endowment, i.e. he can employ the efficiency units of labour he is endowed with to produce current consumption according to a linear 'one-for-one' technology

$$w_t^i = \varepsilon_i^j A_t \quad (7)$$

Alternatively, he can invest it into the production of future consumption goods (according to (2'), (4) and (5)).

3.3.2. Complete markets model (CMM)

Current consumption will be equal to the amount of initial endowments augmented with the amount of borrowing (b^i), less the amount of investment in human capital

$$c_t^i = w_t^i + b_t^i - h_t^i \quad (8)$$

We introduce a government that redistributes income (intra-generational transfers). First, *individual* income is submitted to a tax rate β . Next, every individual receives an equal part of the total tax revenue. Hence, everybody receives a fraction β of the *average* income in the society (\bar{y}). We implicitly assume that redistribution occurs at no *direct* cost, but we could easily introduce a deadweight cost by handing out less than the total tax revenue. Note that redistribution will have an indirect cost due to its negative impact on efficiency.

Future consumption equals future production after redistribution, less the debt repayment

$$d_t^i = (1 - \beta)y_t^i + \beta\bar{y}_t - r_t b_t^i \quad (9)$$

with $r (>1)$ the (gross) market interest rate endogenously determined by the loan market clearing condition: the sum of net-borrowings must equal 0.

$$\sum_i b_t^i = 0 \quad (10)$$

Each individual will spread his endowment over consumption in period 1 and production in period 1 (consumption in period 2) as to maximize intertemporal utility (expression (1)). In the case of a perfect capital market, no credit constraints exist. Everybody can borrow freely as long as the capital market is in equilibrium.

An individual's decision then becomes

$$\begin{aligned} \text{Max} \{ \ln c_t^i + \rho \ln d_t^i \} & \quad \text{w.r.t. } b_t^i, h_t^i \\ \text{s.t. expression (10)} & \end{aligned} \quad (11)$$

Or after substitution of (8) and (9) into (11)

$$\begin{aligned} \text{Max} \{ \ln(w_t^i + b_t^i - h_t^i) + \rho \ln((1 - \beta)y_t^i + \beta\bar{y}_t - r_t b_t^i) \} & \quad \text{w.r.t. } b_t^i, h_t^i \\ \text{s.t. expression (10)} & \end{aligned} \quad (11')$$

Some straightforward manipulation of the first order conditions, leads to the following expression for an individual's investment

$$h_t = h_t^i = \frac{\rho\alpha(1 - \beta)}{1 + \rho\alpha(1 - \beta)} A_t \quad (12)$$

Every agent will invest the same amount of capital in the production process (irrespective of his initial endowment). The first derivative with respect to β is negative², so redistribution reduces investment levels

² We also consider the impact of a change in the tax rate on the past accumulation of knowledge and skills (A) because we are interested in long term (steady state) relationship. Hence, we do not treat A as predetermined for the first derivative of (12) with respect to β .

$$\frac{\delta h_t}{\delta \beta} = \frac{-\rho\alpha}{(1+\rho\alpha(1-\beta))^2} A_t + \frac{\rho\alpha(1-\beta)}{1+\rho\alpha(1-\beta)} \left(-\frac{\kappa}{n} \sum_i k_{t-1}^i \right) < 0 \quad (13)$$

Proposition 1a (CMM):

“If agents can borrow freely, redistribution reduces individuals’ optimal investment human capital”.

Next we derive an expression for the steady state growth:

$$g_y = \ln \left(\frac{\sum_i y_t^i}{\sum_i y_{t-1}^i} \right) = \ln \eta^* + \ln(1-\kappa\beta) + \alpha \ln \rho\alpha + \alpha \ln(1-\beta) - \alpha \ln(1+\rho\alpha(1-\beta)) \quad (14)$$

The partial derivative of g with respect to β is

$$\frac{\delta g_y}{\delta \beta} = -\frac{\alpha}{1-\beta} - \frac{\kappa}{1-\kappa\beta} + \frac{\rho\alpha^2}{1+\rho\alpha(1-\beta)} < 0 \quad (15)$$

If there is more redistribution (larger β), growth will slow down. Meltzer and Richard (1981) (see also Bénabou (1996)) have shown that the preferred tax rate by the median voter will depend on the relative position of his income to the mean of the income distribution. The larger the gap between the median and the mean income is (i.e. the more skewed to the left the income distribution is), the higher the preferred tax rate will be. This brings us to proposition 2a:

Proposition 2a (CMM):

“If capital markets are perfect and more inequality leads to more redistribution, then more inequality will hamper growth”.

3.3.3. Imperfect markets model (IMM)

Now, assume that capital markets are absent (b^i is equal to 0 for all individuals) and the government chooses to redistribute income across generations (intergenerational transfers). Current consumption is represented by

$$c_t^i = w_t^i - h_t^i + \beta A_t \quad (16)$$

Each individual receives an equal share of the tax revenue collected from the *previous* generation:

$$\frac{1}{n} \sum_i \beta y_{t-1}^i = \beta A_t \quad (17)$$

Future consumption equals production minus taxes

$$d_t^i = y_t^i (1 - \beta) \quad (18)$$

The individual's maximization problem becomes (substitution of (17) and (18) into(1)):

$$\text{Max} \left\{ \ln(w_t^i - h_t^i + \beta A_t) + \rho \ln(y_t^i (1 - \beta)) \right\} \quad \text{w.r.t. } h_t^i \quad (19)$$

which leads to

$$h_t^i = \frac{\rho\alpha}{1 + \rho\alpha} (\varepsilon_t^i A_t + \beta A_t) = \frac{\rho\alpha}{1 + \rho\alpha} \left(\frac{\eta^*}{n} \sum_i k_{t-1}^i \right) (1 - \kappa\beta) (\varepsilon_t^i + \beta) \quad (20)$$

In contrast to the perfect market case, investment will differ across individuals. The first derivative of this expression with respect to redistribution is equal to

$$\frac{\delta h_t^i}{\delta \beta} = \frac{\rho\alpha}{1 + \rho\alpha} \left(\frac{\eta^*}{n} \sum_i k_{t-1}^i \right) (-\kappa\varepsilon_t^i + 1 - 2\kappa\beta) \quad (21)$$

The sign of expression (21) will be determined by the sign of the last factor:

$$(-\kappa\varepsilon_t^i + 1 - 2\kappa\beta) \begin{cases} > 0 & \text{if } \varepsilon_t^i < \frac{1 - 2\kappa\beta}{\kappa} \\ < 0 & \text{if } \varepsilon_t^i > \frac{1 - 2\kappa\beta}{\kappa} \end{cases}$$

As redistribution relaxes credit constraints, the poorly endowed (ε^i sufficiently low) will invest more. The 'rich' will invest less. The higher κ , the higher the 'cost' of the tax system in terms of 'lost efficiency' and the lower the number of people that will benefit from redistribution. The effect of redistribution on aggregate investment is positive for 'normal' values of κ (see appendix (A7)),

Substituting (20) into (5) gives in terms of the individual education level

$$e_i^j = (h_i^j)^\gamma = \left(\frac{\rho\alpha}{1+\rho\alpha} \right)^\gamma (A_i)^\gamma (\varepsilon_i^j + \beta)^\gamma$$

Total education can then be expressed as:

$$e_i^{tot} = \left(\frac{\rho\alpha}{1+\rho\alpha} \right)^\gamma (A_i)^\gamma \sum_i (\varepsilon_i^j + \beta)^\gamma \quad (22)$$

The effect of redistribution is double: a higher β negatively influences efficiency (and thereby the accumulated knowledge and skills, A) but at the same time leads to an increase of the third factor of expression (22). For normal parameter values the total effect of redistribution will be positive (parallel to the effect on aggregate investment).

The effect of more equality (in the sense of a lower variance of incomes) on total education is univocally positive: x^γ is a concave function of x ($0 < \gamma < 1$), so by Jensen's inequality we know that a more unequal distribution of endowments (larger variance of ε_i^j), for a given amount of redistribution (fixed β), tends to lower total education.

Proposition 1b (IMM):

"If agents cannot borrow and returns to investment in human capital are decreasing, more inequality reduces total education.

If the efficiency loss is not excessive, redistribution has a positive impact on investment in human capital and education levels".

We can again derive an expression for growth

$$g_y = \ln \eta^* + \ln(1 - \kappa\beta) - \ln n + \alpha \ln \left(\frac{\rho\alpha}{1 + \rho\alpha} \right) + \ln \left(\sum_i (\varepsilon_i^j + \beta)^\alpha \right) \quad (23)$$

In contrast to the CMM, the effect of redistribution on growth is ambiguous. On the one hand, a negative effect is still present through the second term of expression (23). But now there is also a positive impact through the fourth term.

Proposition 2b (IMM)::

“If capital markets are absent, redistribution will stimulate growth through the relaxation of credit constraints. The ‘full’ effect of redistribution is less clear”.

Before we can further explore the potential of the VAR-framework to evaluate the empirical validity of the above theoretical models, we need to take a closer look into the data. Given the above results we need to collect time series for income inequality, human capital, economic growth and redistribution.

3.4. The data

We include 9 OECD countries in the empirical investigation. The choice of the countries was somewhat forced upon us due to the limited data availability. Fortunately the sample seems representative for the entire OECD. Firstly, with Canada, France, Italy, the UK and the US, 5 members of the G-7 are present, next to 4 smaller countries: Belgium, Finland, the Netherlands and Sweden. Secondly, the sample contains countries with an extensive social security system (Belgium, Finland and Sweden) as well as countries with a limited one (the UK and the US). Thirdly, also non-EU countries are present. It would be informative to include Germany in the estimates, but we lack sufficient data to do so.

Although we have selected countries based on data availability and data consistency, we admit that measurement error remains a problem, especially with respect to the enrolment and inequality series.

For each country in our sample we collected annual data for income inequality, secondary and tertiary enrolment, economic growth and social security expenditure. More detailed information on the inequality series is provided in chapter 2. The enrolment series are described in appendix B. Graphs of all series are presented in appendix C.

We first motivate the variable choice.

3.4.1. *Income inequality*

Cowell (1995) shows that, compared to the gini coefficient, the generalized entropy indices have some attractive properties (incorporation of inequality preferences, decomposability). However, we mainly use the former in this paper. The reason is twofold. Firstly, the availability of the gini coefficient is far greater than that of any other inequality index. Since we need an annual inequality measure over the longest possible time period, the gini coefficient turns out to be the only viable alternative. The one exception is France for which we lack data on the gini coefficient and instead use the 5% top income share as a proxy for income inequality. Secondly, the use of the gini coefficient is commonplace in empirical applications. Using the gini coefficient allows a comparison of our results with other studies.

A lot of recent empirical work in the field of inequality and economic growth (Banerjee and Duflo (2003), Forbes (2000), Barro (1999)) is based on the Deininger and Squire (DS) data set (Deininger and Squire (1996)). Because the DS data set has substantially increased data comparability, both over countries and time, it has somewhat become the standard for data sources on inequality. Nevertheless, Atkinson and Brandolini (1999) show that there remain important problems, even with the so-called 'high quality' data in the DS data (see also chapter 2).

In its current form, the DS data set is only applicable to cross section and panel data estimates. As we intent to explore a time series approach, we need to compile a suited data set.

In our sample, the gini coefficients for the different countries are not all based on the same income concept, nor were they all collected in the same manner. Some gini coefficients were derived from census data, others were calculated out of income tax data. Both types have some drawbacks. Income inequality measures based on income tax data might underestimate inequality as only those incomes that are high enough to be taxable are included in the calculation. The quality of census data will depend on the representativity of the sample. In chapter 2 we describe the problems related to the measurement of income inequality in more detail (see also Atkinson (2003)).

The testable implications derived in section 3 are based on inequality of incomes before redistribution. Unfortunately income data are scarce and do not always allow for a perfect

test of the theoretical models. Only for Belgium and for France we have data on income before taxes. Even then the available time series are only imperfect proxies for the 'theoretically optimal choice of income', since, among others, the government support for education is not taken into account. Hence, our data series serve as rough proxies for the theoretically optimal inequality concepts. For most countries data availability forces us to use net income or household disposable income. This has implications for the empirical exercise. Firstly, we will less likely detect the link between inequality and redistribution predicted by the CMM. Secondly, as in the IMM redistribution matters for enrolment because it reduces 'post-redistribution-inequality', we should not expect to detect a positive effect of redistribution on enrolment if we do not use it in combination with pre-redistribution inequality in the estimates.

Based on the theoretical models we can also make some reservation with respect to the choice of inequality measure. To test for the CMM, the middle incomes (median voter) should be highlighted. In the IMM the bottom incomes (credit constraints) and the dispersion of incomes matter more. As the gini coefficient is especially sensitive to changes in the middle incomes, it is an a priori acceptable choice in case of the CMM. To test for the IMM one might prefer a head count index.

All of the above remarks also matter if one uses the DS data. Still most panel data studies using the DS data ignore the potential problems and use different income concepts for different countries in a single panel estimation (Knowles (2001)). Rehme (2002) illustrates how mixing measures of gross and net income inequality can blur the estimation results if redistribution negatively affects growth. Some authors correct the data for the differences in income concept, but Atkinson and Brandolini (1999) doubt that these adjustments really solve the problem.

A related issue is that our income definition might be too narrow. Deininger and Squire (1998) argue that it is the asset distribution that really matters for the systematic effect of inequality on growth. They believe that land distribution is to be preferred to income distribution as a proxy for asset distribution.

3.4.2. *Enrolment rates*

The data set includes enrolment rates in secondary and higher education as a proxy for human capital. In the estimates enrolment appears as a proxy for the investment in human

capital as well as for the stock of human capital, which is in line with earlier work (e.g., Perotti (1996)).

Although the theoretical literature does not force us to interpret investment in this narrow way, human capital formation is likely to be most gravely affected due to credit constraints as the collateral is highly insecure. The choice of this narrow investment variable has a drawback: in the absence of credit constraints the significance of the effect of redistribution on investment might be underestimated. This is only a minor problem if investment in human and physical capital are complements rather than substitutes (see also Azariadis and de la Croix (2003)).

We include both secondary and tertiary education in the estimates to capture the full effects of enrolment. It seems likely that the evolution of enrolment in secondary education was a driving force at the beginning of our sample (the 1960s), but that enrolment in tertiary education has gradually taken over this leading role.

There is some arbitrariness in the exact definition of the enrolment rates, which also makes it hard to compare them across countries. Firstly, the education system differs across countries. The study length can differ, as can the age at the time of first enrolment. This is of importance when computing the number of potential students, which is the denominator of the enrolment rate. It is even problematic in countries that have delegated the design of the education system to the constituent regions (the UK and Canada). Secondly, the education system can change over time. Especially the division between primary and secondary education has been subject to major changes in Finland and Sweden. For the number of potential students for tertiary education we look at the first 5 years after the normal end date of secondary education, which is of course a rough approximation. Although the result might not be the best possible measure for the 'average' enrolment rate, we are confident that it is a good proxy for it (and our results are robust to different choices).

For higher education we use total tertiary education as well as university education (again a choice based solely on data availability). Although the levels are clearly different, the correlation between both series is very high (above 95% for all countries). Therefore, we can safely use the term 'higher education' without further distinction and concentrate on the longest and most consistent series.

In appendix B we describe the enrolment data in more detail.

3.4.3. *Economic growth*

We use the 'GDP at market prices, in volume and at local currency' and 'total population' from the OECD Economic Outlook database (1960-2000) (OECD (2002b)). In the estimation

we use the first difference of the logarithm of the ratio of GDP to total population. As there is no cross-section dimension in the estimates, it is not necessary to adjust for purchasing power parity.

3.4.4. *Social security*

As a proxy for redistribution we use social security expenditure. Sinn (1994) and Wigger (2001) document two ways by which an increase in social security expenditure redistributes resources in a society. We use the public and mandatory private social security expenditure as a percentage of GDP. The data are collected from two sources. Our basic series is taken from the OECD Social Expenditure Database (OECD (2002b)). As for most countries this series starts in 1980 (except for Italy (1982), the Netherlands (1995) and Sweden (1993)), we need to extend it backwards (the latest year included is 1997). To this end we use the growth rate of the comparable ILO data series. Although the levels of both series are sometimes quite different, the growth rates of the OECD series and the ILO series in the overlapping period (1980-1993) are very similar. The start date for the ILO series is 1960, but the most recent year available is at best 1993. Missing values were obtained by linear interpolation.

The use of social security expenditure in our set-up can be questioned. Given the theoretical models we need a variable that captures both the redistributive efforts in a society (IMM) and the distortionary effects of fiscal policy (CMM). Our variable choice partially meets both requirements. The ideal variable again depends on the theoretical model we have in mind. For the IMM we want to measure the redistributive efforts that reduce capital constraints with respect to human capital formation. On the one hand, social security expenditure seems too broad a measure as also expenditure related to health, housing, ... is taken into account. Note, however, that relaxing credit constraints in one field, opens up resources for other purposes. On the other hand, not all relevant redistributive efforts are included (for instance study loans might be overlooked). The CMM deals with all types of government expenditure that are redistributive in nature. From this perspective our broad variable choice seems acceptable. However, in the CMM not the magnitude of the expenditure, but the distortionary effects of the related taxation matter. If taxation were fully lump sum, a redistributive fiscal policy would not distort growth. Easterly and Rebelo (1993) and Perotti (1996) propose the average marginal tax rate as an alternative. Sonedda (2003) uses marginal and average tax rates. But tax-based indicators also have their drawbacks as tax revenues are used for multiple purposes besides redistribution (e.g., defence). Moreover, it is hard to obtain a time

series that goes back as far as the 1960s. As it is possible to construct a long enough and consistent annual time series for social security expenditure, and as no alternative variable seems a priori superior, we believe it is an acceptable choice.

3.5. Cointegration analysis

A number of explorative univariate unit root tests (results not shown) indicated that more than half of the time series in the data set potentially display non-stationary behaviour (the tests for the income inequality series are discussed at length in chapter 2). Given these results and the endogeneity problems in growth econometrics, we choose to test for cointegration in a VAR-framework (eg. Johansen and Juselius (1994)).

As the time dimension of our series is rather limited, we cannot use asymptotic theory and need to perform small sample corrections. Given the small time dimension, we only allow for 2 lags in the VAR. The results are only marginally altered by including a third lag. The choice for 2 lags in the VAR results in a 1st order VECM of the following form:

$$\Delta X_t = \mu_0 + \Pi X_{t-1} + \Gamma \Delta X_{t-1} + \varepsilon_t \quad (24)$$

where X_t is a 5×1 vector containing enrolment in secondary and tertiary education, growth, income inequality and social security expenditures in year t and μ_0 a vector of constants. Based on our univariate analysis, we do not restrict the constant to the cointegration space, as some individual series display trending patterns (Franses (2001)). This might not be in line with economic intuition in the longer term, but it is acceptable over the short period we consider.

If not all variables in X_t are (trend) stationary, the matrix π will not be of full rank. If the system is cointegrated, i.e. there exist linear combinations of the non-stationary variables which are stationary, we can rewrite π as the product of two full column rank matrices, $\Pi = \alpha\beta^T$. Both matrices are of dimension $5 \times r$, with r being the number of cointegrating relations. Expression (24) can be rewritten as:

$$\Delta X_t = \mu_0 + \alpha\beta^T X_{t-1} + \Gamma \Delta X_{t-1} + \varepsilon_t \quad (25)$$

The matrix β contains the long run (cointegrating) relationships, the matrix α the short run adjustments towards these long run equilibria.

3.5.1. The cointegrating rank

For the determination of the cointegrating rank we use the trace statistic evaluated against its *adjusted* asymptotic 95% critical value. Johansen (2002) illustrates that the actual probability of rejecting a correct null hypothesis in a finite sample is much larger than the 5% nominal value. In other words: we should use higher critical test values than the asymptotic ones.

Johansen (2002) introduces a correction factor for the trace statistic which leads to a good approximation of these corrected critical values. If we reject the null hypothesis with the asymptotic critical values, we also look at the corrected ones.

Table 1: cointegrating rank (growth, social security, income inequality, secondary and tertiary education)

Country	Rank	Trace	Trace 95% - Corrected values	Conclusion
Belgium Income after taxes (2 lags)	$H_0: r=0; H_1: r>0$	95.22	74.37	2 cointegration relations is not rejected (asymptotic critical values: 2)
	$H_0: r\leq 1; H_1: r>1$	57.39	51.92	
	$H_0: r\leq 2; H_1: r>2$	28.90	29.70	
Belgium Income before taxes (2 lags)	$H_0: r=0; H_1: r>0$	91.67	74.40	2 cointegration relations is not rejected (asymptotic critical values: 2)
	$H_0: r\leq 1; H_1: r>1$	54.19	52.16	
	$H_0: r\leq 2; H_1: r>2$	26.20	29.70	
Canada (2 lags)	$H_0: r=0; H_1: r>0$	115.50	79.53	2 cointegration relations is not rejected (asymptotic critical values: 3)
	$H_0: r\leq 1; H_1: r>1$	72.09	58.39	
	$H_0: r\leq 2; H_1: r>2$	43.15	45.57 (25.80)	
France (2 lags)	$H_0: r=0; H_1: r>0$	90.58	77.27	2 cointegration relations is not rejected (asymptotic critical values: 2)
	$H_0: r\leq 1; H_1: r>1$	59.01	54.31	
	$H_0: r\leq 2; H_1: r>2$	29.20	50.37 (29.70)	
Finland (2 lags)	$H_0: r=0; H_1: r>0$	100.10	75.98	2 cointegration relations is not rejected (asymptotic critical values: 3)
	$H_0: r\leq 1; H_1: r>1$	55.21	52.35	
	$H_0: r\leq 2; H_1: r>2$	33.06	34.15 (29.70)	
Italy (2 lags)	$H_0: r=0; H_1: r>0$	83.10	78.47	1 cointegration relation is not rejected (asymptotic critical values: 1)
	$H_0: r\leq 1; H_1: r>1$	45.47	52.67 (47.20)	
	$H_0: r\leq 2; H_1: r>2$			
Netherlands (2 lags)	$H_0: r=0; H_1: r>0$	101.90	72.36	2 cointegration relations is not rejected (asymptotic critical values:3)
	$H_0: r\leq 1; H_1: r>1$	59.83	59.57	
	$H_0: r\leq 2; H_1: r>2$	30.13	33.14 (29.70)	
Sweden (2 lags)	$H_0: r=0; H_1: r>0$	51.24	50.13	1 cointegration relation is not rejected (asymptotic critical values: 1)
	$H_0: r\leq 1; H_1: r>1$	20.14	26.70	
UK (2 lags)	$H_0: r=0; H_1: r>0$	77.40	74.20	1 cointegration relation is not rejected (asymptotic critical values: 1)
	$H_0: r\leq 1; H_1: r>1$	44.06	47.20	
USA (2 lags)	$H_0: r=0; H_1: r>0$	85.75	75.58	2 cointegration relations is not rejected (asymptotic critical values: 2)
	$H_0: r\leq 1; H_1: r>1$	52.59	50.99	
	$H_0: r\leq 2; H_1: r>2$	23.42	29.70	

Notes: The figures in italic in the fourth column are corrected values. If we do not reject the null hypothesis (H_0) on the basis of the asymptotic value for the trace statistic, we give the asymptotic value (not in italic) as the corrected one can only be higher. If the use of the corrected values changes the conclusion, we also report the asymptotic value (between brackets)

In the table the corrected values are in italic. We report the asymptotic values between brackets if the conclusion about the rank changes due to the correction.

It should be noted that the correction factor is only an approximation of the actual small sample distortion (Johansen *et al.* (2002)). Its reliability also depends on the parameter values. However, throughout the correction appears to be a useful supplement to the 'classic' analysis (Johansen (2002)).

We do not reject two cointegrating relations for all countries, except for Italy, Sweden and the UK. If we do not apply the small sample correction, we do not reject 3 cointegrating relations for Canada, Finland and the Netherlands.

In the presence of multiple cointegrating relations ($r > 1$), the estimates are not unique and directly interpretable. We can identify the long term relationships between the 5 variables by imposing coefficient restrictions and the long term relations (β) and the short run adjustments (α). Based on the identified relationships we try to distinguish between the IMM and the CMM. Note that the number of model results (cf. section 3.3) we can test *simultaneously* is equal to the number of cointegrating relationships.

3.5.2. Identification of the long term relationships

We identify the long run relations by imposing coefficient restrictions on the cointegrating vectors (beta-vector) and the short run adjustments (alpha-vector). The choice of restrictions is based on the testable results of the IMM and CMM. Again, Johansen (2002b) notes that the asymptotic results of the estimates are not accurate enough for small samples. His results indicates that the actual size can be quite distorted (much larger than the nominal size) in small samples. We do not provide a robust correction for this small sample bias, but, based on Johansen's results, a correction factor for the likelihood ratio of between 1.3 and 1.7 seems probable. Therefore, to convincingly reject the restrictions, the p-value should be sufficiently below 5%.

We use the following 2-step testing procedure:

1. Can we identify the long run relations in line with the results of the IMM?

growth = f(enrolment (+); social security(-); ...] (proposition 2b)

enrolment = f(gini (-); social security(+); ...] (proposition 1b)

2. Can we identify the long run relations in line with the results of the CMM?

$$\text{growth} = f[\text{social security}(-); \dots] \quad (\text{proposition 2a})$$

$$\text{social security} = f[\text{gini} (+), \dots] \quad (\text{proposition 2a})$$

Our objective is to check whether it is possible to detect an identification of the long term relations in line with the theoretical models. The identification of multiple cointegrating vectors is not unique and changing the order in which restrictions are imposed can sometimes drastically change the results. Therefore, we systematically explore the different sequences of restrictions. If we find both the correct growth and/or enrolment relation and/or inequality relation in the first step, we can argue that the data support the IMM. But we can not reject the CMM. Next we check whether the implications of the CMM are compatible with the long run relations. We check whether growth is negatively related to social security, and whether social security is positively related to inequality. The latter requirement might be somewhat too strict given the limitations of the dataset. We should not be surprised if we fail to find a significant effect of inequality on social security expenditure if we can not use an inequality measure for income before taxes (i.e. for all countries except Belgium and France). A similar argument holds for the effect of credit constraints on enrolment rates in the IMM. But now the effect should not vanish: higher post redistribution inequality should still result in less enrolment irrespective of the amount of redistribution taking place. Combined with our earlier remark that the investment variable might suit the IMM better than the CMM (cf. supra) the above comments indicate that the tests give a slightly preferential treatment to the IMM, although the gini coefficient better fits the CMM.

Note that the ambition of this empirical exercise is limited: we want to check if the data support the models. Even if the data fit a model perfectly, we can not yet call this model the 'true model'. We only give an indication which (if any) model suits reality better. Also recall that the variables in the estimates are only approximations of the 'theoretical' variables in the models (cf. supra).

In table 2 we present the beta matrix (after imposing restrictions), i.e. the long term relations. The standard errors are between brackets. We impose zero restrictions on the alpha matrix, the short run adjustments, if the coefficients have a wrong sign or if they are highly insignificant. These 'corrections' are obviously somewhat ad hoc but they can be thought of as a kind of robustness check. We do not present the alpha matrix as it is not essential in our analysis.

Table 2: The long term relations

	Growth	Gini	Secondary education	Higher education	Social security
Belgium before taxes					
Enrolment relation	-14.222 (2.573)	5.822 (1.610)	1.000	-3.832 (0.489)	0.000
Growth relation (p-value: 0.942)	1.000	0.000	-0.527 (0.117)	0.000	0.0083 (0.0024)
Belgium after taxes					
Enrolment relation	-13.615 (1.912)	3.564 (0.923)	1.000	-2.481 (0.332)	0.000
Growth relation (p-value: 0.872)	1.000	0.000	-1.049 (0.274)	0.000	0.0137 (0.0059)
Canada					
Enrolment relation	-2.101 (0.268)	6.316 (1.097)	1.000	-1.075 (0.088)	0.000
Growth relation (p-value: 0.319)	1.000	22.765 (5.787)	0.000	0.000	0.0818 (0.0118)
Finland					
Enrolment relation	-2.104 (0.267)	6.300 (1.089)	1.000	-1.077 (0.088)	0.000
Inequality relation (p-value: 0.379)	0.000	1.000	0.000	0.000	0.0033 (0.0004)
France					
Enrolment relation	0.000	2.746 (0.572)	1.000	-0.487 (0.1067)	0.000
Growth relation (p-value: 0.576)	1.000	0.000	0.000	0.000	0.0027 (0.0003)
Growth relation	1.000	0.000	0.000	-0.031 (0.012)	0.0030 (0.0002)
Social security relation (p-value: 0.729)	352.80 (12.42)	-72.88 (23.48)	-18.792 (4.589)	0.000	1.000
The Netherlands					
Growth relation	1.000	0.000	-0.566 (0.073)	0.000	0.0137 (0.0015)
Social security relation (p-value: 0.225)	219.82 (25.96)	-123.74 (15.43)	0.000	0.000	1.000
The USA					
Enrolment relation	-0.316 (0.074)	0.585 (0.104)	1.000	-0.275 (0.029)	0.000
Growth relation (p-value: 0.718)	1.000	0.0000	-0.508 (0.156)	0.000	0.0045 (0.0009)

Notes: Standard errors are reported between brackets.

For Belgium we find an enrolment relationship compatible with the IMM. We can, however, remove income inequality after taxes from the enrolment relation. If we use inequality before taxes, the elimination of the gini coefficient cannot be accepted. Social security is negatively related to growth, which is compatible with both models. Enrolment in secondary education is positively related to growth. The impact of enrolment in higher education on growth is not significant. We can not identify a social security relation in line with the CMM.

Inequality is detrimental to enrolment in Canada. As the data do not confirm the positive link between enrolment and growth, there is only partial support for the IMM. The data neither indicate that inequality induces social security expenditure.

We again find a robust enrolment relation for Finland. It is hard to identify the second cointegration relationship in terms of either the IMM or the CMM. The data do neither fit a long term social security nor a long term growth relationship. The only meaningful interpretation seems to be an inequality relationship: income inequality is reduced by an increase in social security expenditure.

For France we find a long term enrolment relationship in line with the IMM: more inequality reduces enrolment. Different from other countries, economic growth is not significant in the relationship. However, as enrolment rates do not significantly affect growth, the data only partially support the IMM. We also find a negative impact of social security expenditure on economic growth. Next we abandon the enrolment relationship and try to identify a social security relationship instead. We detect a long run relationship in which social security expenditure are positively influenced by inequality and enrolment, and negatively by growth. This is a specification in line with the CMM. Moreover the negative effect of social security on economic growth is still present. Now we also find a (robust) positive link between enrolment in secondary education and growth. Hence, the results for France are mixed.

For Italy there was only 1 cointegration relationship which does neither seem to fit the CMM nor the IMM. Allowing for 2 cointegration relationships does not reduce these identification difficulties.

For the Netherlands we cannot identify an enrolment relation. We do find a growth relation and social security relation consistent with the CMM. More inequality leads to more social security expenditure and social security leads to less growth. However, the results for the Netherlands might be less trustworthy, as our measure of redistribution does not capture its extensive system of students loans (Guille (2000)). The fact that we cannot identify an enrolment relation in line with the IMM might stem from that shortcoming. Credit constraints will be less severe for investment in education.

For Sweden we should again only allow for 1 cointegration relation. We can identify this relation as an enrolment or as a growth relationship (results not shown). If we identify it as an enrolment relation, enrolment is positively related to growth and social security expenditure. Income inequality is not significant. If we choose the second relationship social security expenditure reduce growth. Enrolment is not significant. If we impose homogeneity between countries with respect to the cointegrating rank, i.e. ignore the test results in table 1 and always allow for 2 cointegration relations, we are able to identify a growth relation and a social security relation consistent with the CMM (results not shown). Sweden also has an

extensive system of students loans (Guille (2000)), which might again explain the lack of IMM-compatible relationship in the Swedish data.

Also for the UK we concluded that there was only 1 cointegration relationship. This relationship can be reduced to the stationary behaviour of economic growth³. If we allow for 2 cointegration relations, the second relation can be identified as an enrolment relation in which income inequality reduces enrolment and social security stimulates it (results not shown).

For the US we find an enrolment relation and a growth relation consistent with the IMM. We cannot identify the social security relation implied by the CMM. Carneiro and Heckman (2002) find that currently only a very marginal fraction of the US population is credit constrained. Their results do not necessarily contradict ours as they themselves note that *'the limited role of short run credit constraints in explaining American educational gaps is, no doubt, in part due to the successful operation of policies that were designed to eliminate such constraints'*. We never detect a significant positive relation between social security and enrolment, which is not really surprising given the 'imperfect' inequality measures (cf. supra).

The available data support both the IMM (3 countries) and the CMM (1 country, 2 if we impose a higher cointegrating rank on the Swedish data). Some relations cannot be identified in terms of either the CMM or the IMM. For France the evidence is mixed. Thus different models might be appropriate for different countries.

With some reservation we can draw two more general conclusions: income inequality reduces enrolment in most countries and the direct impact of social security expenditures on growth is negative once we control for investment in education. Both observations strengthen the case of the IMM.

The findings for the Netherlands and Sweden show that it might be useful to select the redistribution variable on a country specific basis. Given their extensive system of students loans it might not be surprising that the results for these countries are more in line with the implications of the CMM. However, one needs an extensive knowledge of the particularities in the social security and education systems of the different countries to further explore this possibility.

³ Although univariate unit root tests do not reject stationarity of the growth rate for most countries, multivariate unit root tests do reject it for all countries but the UK. Univariate unit root tests did not reject stationarity of the inequality index for all countries (cf. chapter 2), but multivariate test lead to a uniform conclusion: stationarity is now clearly rejected. Stationarity of the other variables in the VAR was rejected both in a univariate and in a multivariate framework.

3.5.3. Limitations of the time series approach

The time series approach deals with a number of problems that disturb the reliability of cross section estimates but it is unable to solve all of them. The most important drawback is the limited length of the time series. While there exist longer time series for some of the variables included in the VAR system, most data series (and by extension the entire VAR system) are limited to the 1960s (at best). As economic growth is a long term phenomenon, we would prefer to include more lags in the VAR specification. However, we face a trade-off between theoretical considerations and practicability.

If we do not include sufficient lags, the correlation between the error terms and the lagged variables might differ from zero resulting in inconsistent estimates. We already argued that lagged values of the endogenous variables are not necessarily good instruments because of the importance of expectations about the future for current decision making. So we need to include enough lags to cover a normal 'planning horizon'. Although the appropriate horizon is still a subject of discussion, its length will surely exceed 2 or 3 years.

To evaluate the potential inconsistency of our results due to the presence of autocorrelation in the error terms, we looked at a vector error autocorrelation test. Based on the findings of Doornik (1996) we applied the F-approximation of the Lagrange-multiplier test because of its superior behaviour in small samples. The results indicate that autocorrelation is not much of a problem for the estimates involving Belgium, Canada, Sweden and the UK (we obtain a p-value above 0.7 for each of these countries). For France, Finland and the USA the test values do not reject 'the absence of autocorrelation' at the 10%-level, but they do reject it at the 20%-level for the former and at the 15%-level for the latter two countries. For Italy and the Netherlands the results clearly reject 'no autocorrelation'. On the one hand these test results further strengthen our earlier suspicions of the Dutch data (and results), but on the other hand they deepen our belief in some of the other results.

Due to the short time series, it is hard to check whether the estimated parameters are stable over time. We argued that parameters differ across countries, but they might also depend on the level of development in an individual country. Maddala and Wu (2000) find some evidence of instability over time in growth relationships. But because of the short time period covered in the estimates, this potential 'time heterogeneity' will be much lower than the 'cross country heterogeneity'.

In summary, our results are not unquestionable, but nonetheless the new methodology can serve as a valuable supplement to the techniques currently used in growth econometrics. The time series approach will gradually lead to more robust results as longer time series become available.

3.6. Conclusions

Durlauf (2001) urged growth economists to advance in the field of growth econometrics. The empirical analysis of economic growth is indeed tainted by some very serious flaws and shortcomings. We accepted Durlauf's challenge and this paper is a concise report of an attempt to deal with the problems typical of growth empirics. Based on an overview of the most serious problems noted in the literature, we propose a methodology that deviates in two ways from existing work: firstly, we propose a time series approach instead of a cross section or panel analysis, and secondly, we resort to the Johansen cointegration framework, a methodology that, to our knowledge, has not been applied before in growth econometrics. These 'innovations' deal with heterogeneity and endogeneity problems. However, new problems arise as existing data sets are not suited for this new approach. Even for OECD countries it is hard to collect long and reliable time series for inequality and enrolment. On the one hand, these data problems limit the workability of our method, but on the other hand also open up new perspectives for the future. As longer and more reliable time series will become available for more countries, the trustworthiness of the results will strongly increase. We applied the methodology to the analysis of the relation between income inequality and economic growth.

Keeping the above data reservation in mind, we dare present some cautious conclusions. Both the imperfect market model (Belgium, Canada, the US) and the complete market model (the Netherlands and with some flexibility: France and Sweden) find some support in the data. The fact that the negative effect of inequality on enrolment seems relevant for most countries, enhances the credibility of the IMM. For most countries we detect a negative relation between social security expenditures and economic growth once we control for investment in human capital. This finding is again in line with the implications of the IMM. The support for the CMM is weakened if we take the extensive system of students loans in the Netherlands and Sweden into account. Hence, the data clearly offer more support for the prevalence of the IMM in our sample of OECD countries than for the prevalence of the CMM. Still we cannot completely dismiss the CMM. Our analysis indicates that different models might hold for different countries (even in a subset of rich countries). If this conclusion proves to be robust, it questions the appropriateness of panel estimates for this kind of research and thereby also the validity of previous studies.

3.7. References

- Aghion P., Howitt P. (1998): *Endogenous Growth Theory*, MIT Press, Cambridge
- Alesina A., Rodrik D. (1994): *Distributive Politics and Economic Growth*, Quarterly Journal of Economics, 109, p.465-490
- Arjona R., Ladaïque M., Pearson M. (2001): *Growth, Inequality and Social Protection*, OECD Labour Market and Social Policy Occasional Papers no. 51, 80p.
- Atkinson A.B., Brandolini A. (1999) : *Promise and Pitfalls in the Use of "Secondary" Datasets: Income Inequality in OECD countries*, mimeo, 60p.
- Atkinson A.B. (2003): *Income Inequality in OECD Countries: Data and Explanations*, CESifo Working Paper no. 881, 52p.
- Azariadis C., de la Croix D. (2003) : *Growth or Equality ? Losers and Gainers from Financial Reform*, mimeo, 49p.
- Banerjee A.V., Duflo E. (2003): *Inequality and Growth: What Can the Data Say?*, Journal of Economic Growth, 8, p.267-299
- Barro R.J., Sala-i-Martin X. (1995): *Economic Growth*, McGraw-Hill, New York
- Barro R.J. (1999): *Inequality, Growth and Investment*, NBER Working Paper no. 7038, 37p.
- Barro R.J. (2000): *Inequality and Growth in a Panel of Countries*, Journal of Economic Growth, 5, p.5-32
- Bénabou R. (1996): *Inequality and Growth*, in: Bernanke B.S. and Rotemberg J. (eds.), *NBER Macroeconomics Annual 11*, MIT Press, p.11-92
- Bils M. and Klenow P.J. (2000): *Does Schooling Cause Growth*, The American Economic Review, vol. 90 (5), p.1160-1183
- Bleaney M., Nishiyama A. (2004): *Income Inequality and Growth – Does the Relationship Vary with the Income Level?*, Economics Letters, 84, p.349-355
- Blöndal S., Field S., Girouard N. (2002): *Investment in Human Capital through Upper-Secondary and Tertiary Education*, OECD Economic Studies no. 34, p.41-89
- Blundell R., Bond S., Windmeijer F. (2000): *Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator*, IFS Working Paper no. 00/12, 46p.
- Bond S. (2002): *Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice*, Cemmap Working Paper no. 09/02, 34p.
- Brock W.A., Durlauf S.N. (2000): *Growth Economics and Reality*, NBER Working Paper no. 8041, 48p.
- Cameron S., Taber C. (2004): *Estimation of Educational Borrowing Constraints Using Returns to Schooling*, Journal of Political Economy, 112 (1), p.132-182

- Canova F. (1999): *Testing for Convergence Clubs in Income per Capita: A Predictive Density Approach*, CEPR Discussion Paper no. 2201, 28p.
- Carey D., Tchilinguirian (2000): *Average Effective Tax Rates on Capital, Labour and Consumption*, OECD Economics Department Working Papers no. 258, 54p.
- Carneiro P., Heckman J.J. (2002): *The Evidence on Credit Constraints in Post-Secondary Schooling*, *The Economic Journal*, 112 (8), p.705-734
- Cowell F. (1995): *Measuring Income Inequality (2nd Edition)*, Harvester Wheatsheaf, Hemel Hempstead
- Davis S.J., Henrekson M. (2004): *Tax Effects on Work Activity, Industry Mix and Shadow Economy Size: Evidence from Rich-Country Comparisons*, NBER Working Paper no. 10509, 64p.
- Deininger K., Squire L. (1996): *Measuring Income Inequality: a New Data-Base*, *World Bank Economic Review*, 10 (3), p.565-591
- Deininger K., Squire L. (1998): *New Ways of Looking at Old Issues: Inequality and Growth*, *Journal of Development Economics*, 57, p.259-287
- Doornik J.A. (1996): *Testing Vector Autocorrelation and Heteroscedasticity in Dynamic Models*, URL: www.nuff.ox.ac.uk/users/doornik, Nuffield College
- Durlauf S.N., Johnson P. (1995): *Multiple regimes and cross-country growth behavior*, *Journal of Applied Econometrics*, 10, p.365-384
- Durlauf S.N. (2001): *Manifesto for a Growth Econometrics*, *Journal of Econometrics*, 100, p.65-69
- Durlauf S.N. (2003): *Policy Evaluation and Empirical Growth Research*, Central Bank of Chile Working Paper no. 205, 28p.
- Easterly W., Rebelo S. (1993): *Fiscal Policy and Economic Growth. An Empirical Investigation*, *Journal of Monetary Economics*, 32 (3), p.417-458
- Evans P. (1998): *Using Panel Data to Evaluate Growth Theories*, *International Economic Review*, 39 (2), p.295-306
- Forbes K.J. (2000): *A Reassessment of the Relationship Between Inequality and Growth*, *The American Economic Review*, 90 (4), p.869-887
- Franses P.H. (2001): *How to Deal with Intercept and Trend in Practical Cointegration Analysis?*, *Applied Economics*, 33, p.577-579
- Guille M. (2000): *Students Loans in Europe: An Overview*, PuRE Final Report, 15p.
- Hauk W.R.Jr., Wacziarg R. (2004): *A Monte Carlo Study of Growth Regressions*, NBER Technical Working Paper no. 0296, 51p.
- Headey B., Muffels R. (2002): *Does a Generous Welfare State Generate long-run Efficiency Gains which Counterbalance short-run Losses? Economic Panel Evidence from the U.S., Germany and the Netherlands 1985-96*, Paper presented at the "Conference on Poverty Dynamics", Central European University, Budapest, May 24-26, 26p.

Johansen S., Juselius K. (1994): *Identification of the long-run and the short-run Structure. An Application to the ISLM Model*, Journal of Econometrics, 63, p.7-36

Johansen S. (2002): *A Small Sample Correction for Tests of Cointegrating Rank in the Vector Autoregressive Model*, Econometrica, 70 (5), p.1929-1961

Johansen S. (2002b): *A Small Sample Correction for Tests of Hypotheses on the Cointegrating Vectors*, Journal of Econometrics, 111, p.195-221

Johansen S., Hansen H., Fachi S. (2002): *A Simulation Study of Some Functionals of Random Walk*, mimeo, 22p.

Knowles S. (2001): *Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data*, CREDIT Research Paper no. 01/03, 32p.

Kourtellos A. (2002): *Modelling Coefficient Heterogeneity in Cross-Country Growth Regression Models*, mimeo, 18p.

Krueger A., Lindahl M. (2000): *Education for Growth: Why and for Whom?*, NBER Working Paper no. 7591, 50p.

Kuznets S. (1955): *Economic Growth and Income inequality*, The American Economic Review, 45 (1), p.1-28

Lindbeck A. (1985): *What is Wrong with the West European Economies*, World Economy, 8 (2), p.153-169

Lindbeck A. (1988): *Consequences of the Advanced Welfare State*, World Economy, 11 (1), p.19-37

Lindbeck A. (1993): *Work Disincentives in the Welfare State*, in: Lindbeck A., The Welfare State – The Selected Essays of Assar Lindbeck Vol. II, Edward Elgar, Hants

Maddala G.S., Wu S. (2000): *Cross-country Growth Regressions: Problems of Heterogeneity, Stability and Interpretation*, Applied Economics, 32, p.635-642

Meltzer A.H., Richard S.F. (1981): *A Rational Theory of the Size of Government*, Journal of Political Economy, 89 (5), p.914-927

Milanovic B. (1994): *Determinants of Cross-Country Income Inequality: An 'Augmented' Kuznets' Hypothesis*, World Bank Policy Research Working Paper no. 1246, 31p.

Mitchell B.R. (1998): *International Historical Statistics – Europe 1750 – 1993 (4th Edition)*, Macmillan Reference Ltd., London

Nickell S.J. (1981): *Biases in Dynamic Models with Fixed Effects*, Econometrica, 49 (6), p.1417-1426

OECD (2002): *Education Online Database*, URL: www.oecd.org.

OECD (2002b): *Statistical Compendium on CD-Rom 2002/1*.

Park H. (1994): *Endogenous Growth Theory: Intellectual Appeal and Empirical Shortcomings*, Journal of Economic Perspectives, 8 (1), p.55-72

- Perotti R. (1994): *Income Distribution and Investment*, European Economic Review, 38, p.827-835
- Perotti R. (1996): *Growth, Income Distribution and Democracy: What the Data Say*, Journal of Economic Growth, 1 (2), p.149-187
- Persson T., Tabellini G. (1994): *Is Inequality Harmful for Growth*, The American Economic Review, 84 (3), p.600-621
- Psacharopoulos G. (1994): *Returns to Investment in Education: A Global Update*, World Development, 22 (9), p.1325-1343
- Quah D. (2000): *Cross-Country Growth Comparison: Theory to Empirics*, mimeo, 29p.
- Rehme G. (2002): *(Re-)Distribution of Personal Incomes, Education and Economic Performance Across Countries*, CESifo Working Paper no. 711(5), 45p.
- Schultz T.P. (1999): *Health and Schooling Investments in Africa*, Journal of Economic Perspectives, 13 (3), p.67-88
- Sinn H.W. (1994): *A Theory of the Welfare State*, NBER Working Paper No. 4856, 38p.
- Sonedda D. (2003): *Wealth Inequality, Income Redistribution and Growth in 15 OECD Countries*, mimeo, 35p.
- Stock J., Wright J. , Yogo M. (2002): *A Survey of Weak Instruments and Weak Identification in GMM*, Journal of Business and Economics Statistics, 20 (4), p.518-529
- Weede E. (1997): *Income Inequality, Democracy and Growth Reconsidered*, European Journal of Political Economy, 13, p.751-764
- Wigger B.U. (2001): *Gifs, Bequests, and Growth*, Journal of Macroeconomics, 23 (1), p.121-129

Appendix A: Mathematical derivations
Expression 12:

$$\begin{aligned} \text{Max} \left\{ \ln(w_t^i + b_t^i - h_t^i) + \rho \ln \left((1-\beta)y_t^i + \beta \bar{y}_t - r_t b_t^i \right) \right\} & \quad \text{w.r.t. } b_t^i, h_t^i \\ & \quad \text{s.t. expression (10)} \end{aligned}$$

Maximizing with respect to h :

$$\frac{1}{w_t^i + b_t^i - h_t^i} = \frac{\rho(1-\beta)(1-\kappa\beta)\eta^* \alpha (h_t^i)^{\alpha-1} A_t^{1-\alpha}}{(1-\beta)y_t^i + \beta \bar{y}_t - r_t b_t^i} \quad (\text{A1})$$

Maximizing with respect to b :

$$\frac{1}{w_t^i + b_t^i - h_t^i} = \frac{\rho r_t}{(1-\beta)y_t^i + \beta \bar{y}_t - r_t b_t^i} \quad (\text{A2})$$

Equations (A1) and (A2) reveal the following expression for r :

$$r_t = \eta^* \alpha \left(\frac{A_t}{h_t^i} \right)^{1-\alpha} (1-\kappa\beta)(1-\beta) \quad (\text{A3})$$

which is of course equal to the after tax marginal product of human capital.

From (A3) we can deduce that investment levels have to be equal across individuals ($h_t^i = h_t$). From (A2) we can derive that:

$$(1-\beta)y_t^i + \beta \bar{y}_t - r_t b_t^i = \rho r_t (w_t^i + b_t^i - h_t) \quad (\text{A4})$$

If we aggregate this expression over the entire population (n individuals) and use the restriction that the sum of net-borrowings has to equal 0 (market clearing condition), we get

$$\sum_i \left((1-\beta)y_t^i + \beta \bar{y}_t \right) = \sum_i \rho r_t (w_t^i - h_t)$$

$$\begin{aligned} \Rightarrow n((1-\beta)\bar{y}_i + \beta\bar{y}_i) &= \rho r_i \left(\sum_i w_i^j - nh_i \right) \\ \Rightarrow n\bar{y}_i &= \rho r_i (nA_i - nh_i) \end{aligned} \quad (\text{A5})$$

$$\begin{aligned} \Rightarrow (1-\kappa\beta)\eta^* h_i^\alpha A_i^{1-\alpha} &= \rho\eta^* \alpha A_i^{1-\alpha} h_i^{\alpha-1} (1-\kappa\beta)(1-\beta)(A_i - h_i) \\ \Rightarrow h_i &= \rho\alpha(1-\beta)(A_i - h_i) \end{aligned}$$

which gives expression (12).

Expression 14:

$$\frac{\sum_i y_i^j}{\sum_i y_{i-1}^j} = \frac{(1-\kappa\beta)\eta^* h_i^\alpha A_i^{1-\alpha}}{A_i} = \frac{(1-\kappa\beta)\eta^* \left(\frac{\rho\alpha(1-\beta)}{1+\rho\alpha(1-\beta)} \right)^\alpha A_i^\alpha A_i^{1-\alpha}}{A_i} = \eta^* \left(\frac{\rho\alpha(1-\beta)}{1+\rho\alpha(1-\beta)} \right)^\alpha (1-\kappa\beta)$$

In the first step we use the assumption that there is no population growth. Taking the natural logarithm of this expression results in expression (14).

Expression 20:

$$\text{Max} \left\{ \ln(w_i^j - h_i^j + \beta A_i) + \rho \ln(y_i^j (1-\beta)) \right\} \quad \text{w.r.t. } h_i^j \quad (\text{19})$$

Results in

$$\begin{aligned} \frac{1}{w_i^j - h_i^j + \beta A_i} &= \frac{\rho(1-\beta)\eta^* (1-\kappa\beta) A_i^{1-\alpha} \alpha (h_i^j)^{\alpha-1}}{(1-\beta)\eta^* (1-\kappa\beta) A_i^{1-\alpha} (h_i^j)^\alpha} \Leftrightarrow \frac{1}{w_i^j - h_i^j + \beta A_i} = \frac{\rho\alpha}{h_i^j} \\ \Leftrightarrow h_i^j &= \frac{\rho\alpha}{1+\rho\alpha} (\varepsilon_i^j A_i + \beta A_i) = \frac{\rho\alpha}{1+\rho\alpha} \left(\frac{\eta^*}{n} \sum_i k_{i-1}^j \right) (1-\kappa\beta) (\varepsilon_i^j + \beta) \end{aligned} \quad (\text{20})$$

Total investment is then equal to

$$\begin{aligned}\sum_i h_i^j &= \sum_i \left(\frac{\rho\alpha}{1+\rho\alpha} (\eta^* \bar{k}_{i-1}) (1-\kappa\beta) (\varepsilon_i^j + \beta) \right) = \frac{\rho\alpha}{1+\rho\alpha} (\eta^* \bar{k}_{i-1}) (1-\kappa\beta) \sum_i (\varepsilon_i^j + \beta) \\ &= \frac{\rho\alpha}{1+\rho\alpha} (\eta^* \bar{k}_{i-1}) (1-\kappa\beta) n(1+\beta)\end{aligned}\quad (\text{A6})$$

Redistribution will have a positive effect on total investment in human capital as long as

$$\beta < \frac{1}{2} \left(\frac{1}{\kappa} - 1 \right) \quad (\text{A7})$$

This condition is satisfied for most plausible values of κ and β . For instance, for values of κ below 1/3, the effect is positive independent of the value of β . For values of β comparable to the highest average effective tax rates in the OECD (Carey and Tchilinguirian (2000)), κ can be as high as 1/2. So if the efficiency loss is not too high, redistribution will positively affect total investment.

Expression 23:

$$\sum_i y_i^j = \eta^* (1-\kappa\beta) A_i^{1-\alpha} \sum_i (h_i^j)^\alpha = \eta^* (1-\kappa\beta) \left(\frac{\rho\alpha}{1+\rho\alpha} \right)^\alpha A_i^{1-\alpha} A_i^\alpha \sum_i (\varepsilon_i^j + \beta)^\alpha \quad (\text{A8})$$

$$\Rightarrow \frac{\sum_i y_i^j}{\sum_i y_{i-1}^j} = \frac{\eta^* (1-\kappa\beta) \left(\frac{\rho\alpha}{1+\rho\alpha} \right)^\alpha A_i \sum_i (\varepsilon_i^j + \beta)^\alpha}{n A_i} \quad (\text{A9})$$

Taking the natural logarithm of (A9) gives expression (23).

Appendix B: Enrolment rates (data description)

For Belgium we have a shorter series for enrolment at university (1956 – 1992, Mitchell (1998)). The correlation between this series and tertiary education is very high (98,6%).

For Canada no distinction between students in primary or secondary education is made, which is partly due to the existence of different education systems in the different Canadian regions. However, the sum of years in primary and secondary education is always identical. We suspect that the division is rather arbitrary (i.e. not based on some kind of minimum education programme) and have approximated secondary education by subtracting the population aged 5 to 11 from the original series. As education was obligatory over the period 1960-1998 for children at these ages, the result of the subtraction should be a reasonable proxy for the number of pupils in secondary schools (although officially they might belong to another category). UNESCO reports total enrolment in secondary education in Canada for the period 1980-1995. The ratio of our data and the UNESCO data is systematically between 1,05 and 1,08. So our series might slightly overestimate the number of students in secondary education. However, this error seems to be quite consistent over time. As our estimation result depends on changes in enrolment, we choose not to adjust the series to eliminate the 'error'. Enrolment at university is strongly correlated (.97) with enrolment in tertiary education reported by the OECD (2002) over the period 1985-1997.

The enrolment in secondary education in Finland is again an approximation. First, we added the students in primary and secondary education, and next we subtracted the population aged 7 to 12 from that sum. The rationale for this operation are some clear inconsistencies in the basic series that seem to stem from a changing division between primary and secondary education over the years. Although the levels in the new series are substantially lower than the levels reported by the OECD (2002) over the period 1985-1993, the correlation between the two series is quite high (.93).

For Italy we did not find suitable demographical data before 1970. We extended the population series backwards until 1967 by shifting back age cohorts from 1970. By doing so we ignore the impact of migration and death. The latter will not be very important as we deal with young people. The former might be more relevant. However, if we look at the first comparable period, 1970-1972, there are only very small changes in these age cohorts. Therefore, we can safely assume that the error will be minor. The correlation between university enrolment and tertiary education (OECD) over the period 1985-1995 is again very high (.97).

Secondary enrolment in the Netherlands does not include vocational training. If we only consider university enrolment (WO), the enrolment rate in higher education drops considerably but the correlation between both series is high (.98).

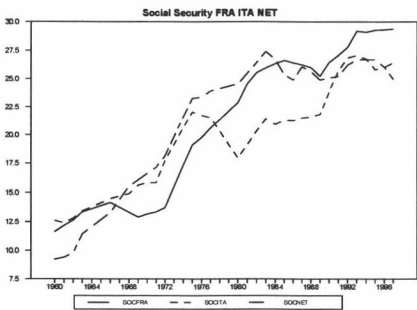
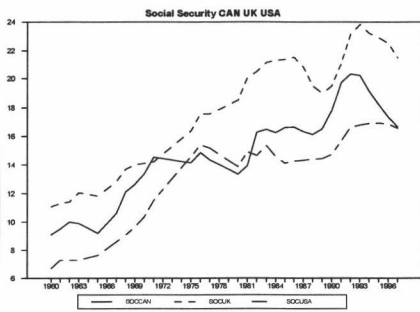
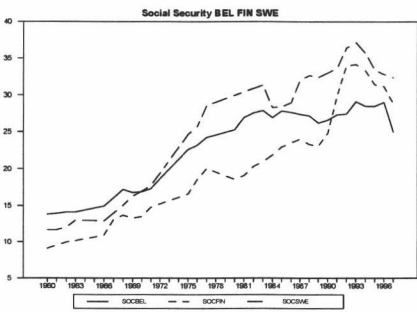
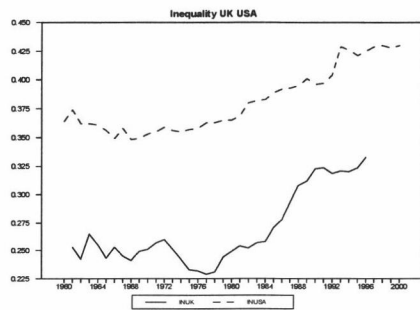
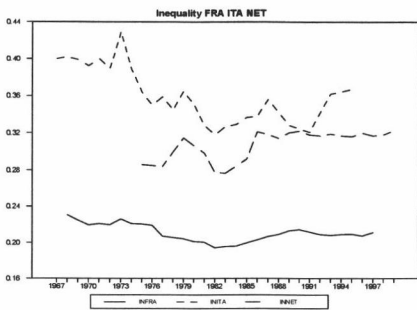
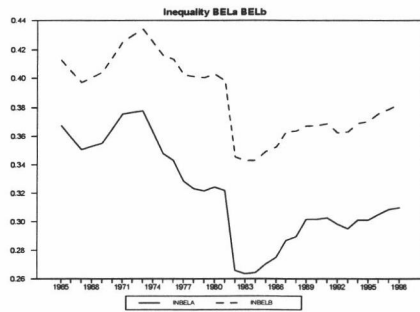
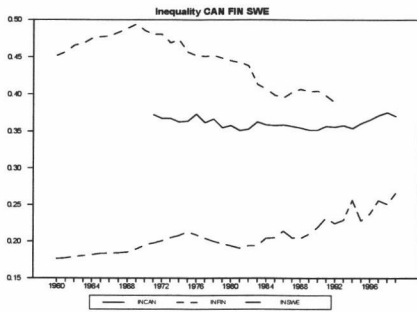
The population statistics for the UK were collected from 3 regional sources: the Office for National Statistics (England and Wales), the General Register Office (Scotland) and Northern Ireland Statistics.

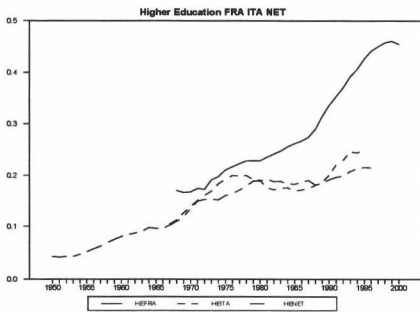
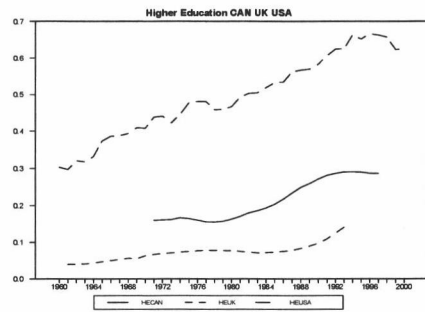
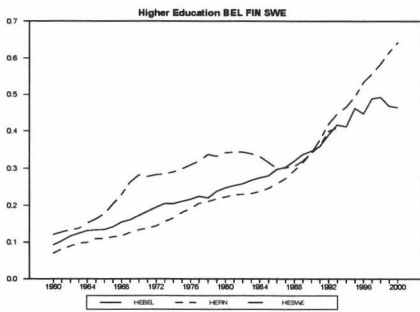
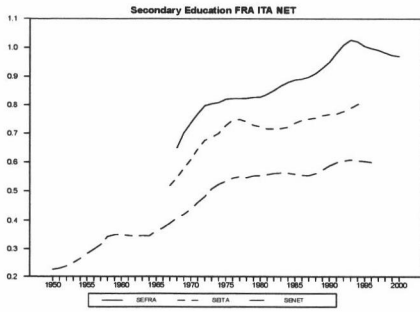
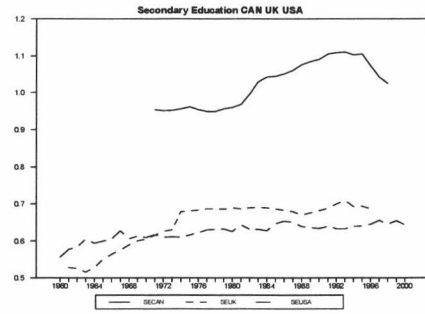
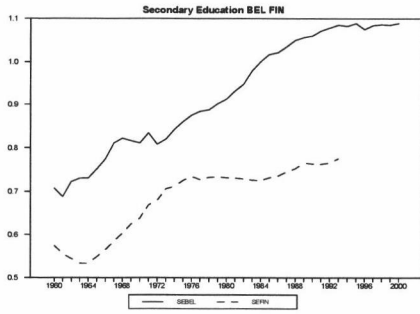
We did not include secondary enrolment in Sweden, as the available series clearly lacked consistency. Different from the Finnish case, it seems hard to cure this problem in a reliable manner.

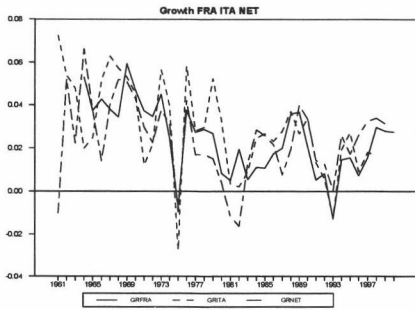
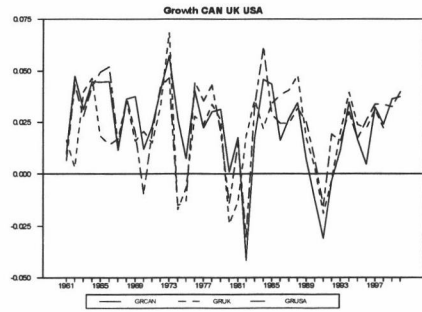
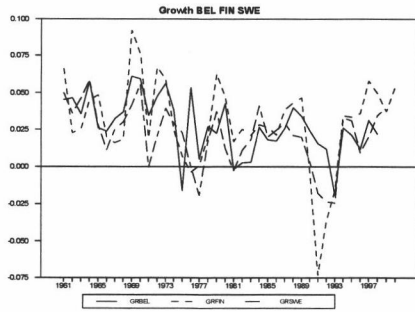
Table B.1: Enrolment rates

Country	Enrolment rate	Source	Period
Belgium	Secondary education / Population aged 12-17	NIS – Statistical Yearbook NIS – Population Statistics	1956 – 2000 1960 – 2000
	Tertiary education / Population aged 18-22	NIS – Statistical Yearbook NIS – Population Statistics	1956 – 2000 1960 – 2000
Canada	Secondary education (approximation) / Population aged 12-17	Statistics Canada	1960 – 1998
	University education / Population aged 18-22	Statistics Canada Statistics Canada	1971 – 2001 1960 – 1997 1971 – 2001
Finland	Secondary education (approximation)/ Population aged 13-18	Mitchell (1998) Statistics Finland	1960 – 1993 1960 – 2001
	Tertiary education / Population aged 19-23	SF – Statistical Yearbook Statistics Finland	1960 – 1993 1960 – 2001
France	Secondary education / Population aged 12-18	Insée – Statistical Yearbook Insée – Statistical Yearbook	1968 – 2000 1960 – 2001
	University education / Population aged 19-23	Insée – Statistical Yearbook Insée – Statistical Yearbook	1968 – 2000 1960 – 2001
Italy	Lower secondary education (scuola media) / Population aged 12-14	ISTAT – Statistical Yearbook Eurostat – New Cronos	1963 – 1995 1970 – 2000
	Higher secondary education (scuole secondari superiori) / Population aged 15-19	ISTAT – Statistical Yearbook Eurostat – New Cronos	1966 – 1995 1970 – 2000
	Secondary education	Combination of lower and higher secondary education	1966 – 1995
	University education/ Population aged 20 – 24	ISTAT – Statistical Yearbook Eurostat – New Cronos	1966 – 1995 1970 – 2000
The Netherlands	Secondary education / Population aged 12 – 17	CBS (MAVO + HAVO + VWO) CBS	1950 – 1996 1950 – 2000
	Tertiary education / Population aged 18 – 22	CBS (WO + HBO) CBS	1950 – 1996 1950 – 2000
The UK	Secondary education / Population aged 12 – 17	ONS ONS, GRO, NISat	1950 – 1999 1961 – 1999
	University education / Population aged 18 - 22	Mitchell (1998) ONS, GRO, NISat	1950 – 1993 1691 – 1999
The USA	High School students / Population aged 12 – 17	USBC USBC	1960 – 2000 1960 – 2000
	College students / Population aged 18 - 22	USBC USBC	1960 – 2000 1960 – 2000
Sweden	Undergraduate students / Population aged 19 – 23	Statistics Sweden Eurostat – New Cronos Statistical Yearbook Sweden	1960 – 2000 1970 – 2001 1960 – 1969

Appendix C: Inequality, social security expenditure, enrolment and growth in graphs







Chapter 4:

Inequality and Growth: From Micro Theory to Macro Empirics

joint work with Dirk Van de gaer

Abstract

To establish the nature of the link between income distribution and economic growth by means of a standard growth regression, one needs to collapse an entire income distribution into a scalar measure of inequality. Due to data shortages macro-economic research has typically been forced to use the gini coefficient for this purpose. Using a simulation set up we check how well different measures of inequality or poverty succeed in detecting the correct relationship. We find that the gini coefficient might not be the worst of choices, but the comparison of the explanatory power of different inequality measures can help to identify the theoretical mechanism through which inequality affects growth.

JEL Classification: D63, O11

We would like to thank Bas van Aarle, David de la Croix, Rafael Doménech, Gerdie Everaert and Freddy Heylen for useful comments and suggestions.

We acknowledge financial support from the Interuniversity Attraction Poles Programme –Belgian Science Policy [Contract No. P5/21].

4.1. Introduction

The link between income inequality and economic growth on a macroeconomic level has frequently been the subject of economic research (Alesina and Rodrik (1994), Perotti (1994, 1996), Persson and Tabellini (1994), Barro (1999), Forbes (2000)). Although a majority of authors agree that inequality negatively affects growth, there are a number of authors who distinctly question this result alluding to flaws in the estimation methodology (e.g., Forbes (2000)). So far no methodology has been entirely convincing. Durlauf (2001) calls for a more advanced approach towards growth econometrics¹. New estimation techniques should explicitly deal with endogeneity, omitted variables bias, parameter heterogeneity and other noted shortcomings in traditional growth regressions.

Next to Durlauf's manifesto for an improved methodology, there is an urgent need for a manifesto for improved inequality data. As no methodology can lead to sensible results in the absence of suitable data, it is a bit premature to impugn the validity of established results on the basis of methodological comments. A thorough investigation of the data required to test the theory seems a more logical first step.

The problems one has to face when measuring income inequality can be divided into two broad categories (see chapter 2).

Firstly, one has to measure income itself. One has to select the relevant income flows (labour income, capital income, etc.) and recipient level (individuals, households, families). One needs to determine at which stage income will be measured (before or after taxes, before or after redistribution, etc.). As tax records do not contain all the necessary information, household surveys come into play. They typically fail to correctly incorporate all income levels: poor households are underrepresented and rich household underreport their incomes. Hence, the risk of systematic measurement error in income data is quite large.

Secondly, one needs to transform income flows into a one-dimensional income inequality measure. Clearly this cannot be done without a loss of information. Many inequality measures can be constructed. As different measures focus on different aspects of the income distribution, the correlation between these measures might be limited. Moreover they can order income vectors differently (Atkinson (1970), Cowell (2000)). Hence, the choice of measure matters for the estimates.

In practice, researchers are constrained by poor data availability. Since its conception in the mid 1990s, the Deininger and Squire dataset (DS, Deininger and Squire (1996)) has somewhat become the standard in macroeconomic empirical work. However, it deals only

¹ For an overview of problems in growth econometrics see chapter 3.

partially with problems of comparability between countries (Atkinson and Brandolini (1999)) and contains only quantile ratios and gini coefficients.

In most "inequality-growth research" the channels through which income inequality influences growth are either not a priori defined (e.g., Forbes (2001)) or one does not (or only partially) base the choice of inequality measure on these channels (e.g., Perotti (1996)). The 'optimal' inequality measure is nonetheless related to the mechanism behind the inequality-growth relationship. If credit constraints prevent poor people from investing in human capital (thereby hampering future growth), income of the poorest after redistribution will matter most. If higher inequality incites voters to demand higher redistribution (and taxes), the pre-redistribution income of the median voter will be the appropriate income concept. By neglecting this, presented estimates often do not actually test what they claim to. A very uneven pre-tax distribution and high taxes can result in the same post-tax inequality as a very even pre-tax distribution in combination with a zero tax rate. However, if taxes affect growth potential, the growth track will be very different.

Since growth regressions are simultaneously infected by all of the above data and econometric problems, it is next to impossible to isolate the bias caused by one specific defect. The aim of this paper is to obtain a better understanding of the importance of the choice of the scalar inequality or poverty measure by means of a simulation exercise. Based on three theoretical models featuring inequality and growth we simulate micro-level income data. These data are used to determine the link between inequality and growth in a simple growth regression. We pose three main questions. Does it matter which measure we use in the estimates? Does the answer to the first question depend on the theoretical model? Can we discriminate between models based on the differences in estimates using different inequality measures?

The advantage of the simulation approach is that we can isolate the impact of the chosen inequality measure and steer clear of all other problems. We have perfect data not infected by measurement error. We have perfect knowledge of the model generating the data such that there are no omitted variables, there is no endogeneity or misspecification problem and we know which measure of inequality should be included in the regression.

Due to a lack of robust empirical findings on the relationship between income inequality and economic growth, no consensus has yet emerged on this issue. Our results add a new element to the debate. They indicate that if the relationship between income inequality and economic growth is captured by the theoretical models we consider, then one should be able

to detect this relationship using a gini coefficient. This strengthens the call for a more advanced approach towards growth econometrics (Durlauf (2001)). Even faced with a scarcity of inequality data one can empirically assess the validity of growth theories as long as econometric shortcomings are overcome.

The remaining sections of the paper are structured as follows. In section 4.2, we algebraically deduce the relationship between growth and inequality for three models in which growth and inequality appear: a complete markets model, a complete markets model with median voter and an imperfect markets model. Section 4.3 presents some specifics with respect to the simulation and the calibration of the models. The list of inequality and poverty measures is presented and motivated in section 4.4. The simulation results are reported and interpreted at length in section 4.5. Finally, we summarize the most important insights of the paper in section 4.6.

4.2. The models

4.2.1. General set up

The models we include in the analysis are based on Aghion and Howitt (1998). We consider an overlapping generations (OLG) model in which n_c individuals live in country c for two periods². The intertemporal utility, U , of an individual i born at time t is given by

$$U_t^i = \ln c_t^i + \rho \ln d_t^i \quad (1)$$

where c_t^i and d_t^i respectively denote current and future consumption. The parameter ρ is negatively related to time preference. The economy consists of one good only that serves both as capital and consumption good. Production of the future good takes place at time t according to an AK technology.

$$y_t^i = \eta_{at}^i (A_t)^{1-\alpha} (h_t^i)^\alpha \quad (2)$$

The parameter η_{at}^i is an individual-specific efficiency measure³, h_t^i is an individual i 's investment in human capital and A_t is the level of knowledge and skills in society. Based on

² In the remaining part of the paper we will omit the country subscript for reasons of notational convenience.

Lindbeck (1985, 1988, 1993) we assume that efficiency is a decreasing function of the tax rate β . For simplicity we impose an efficiency loss that is proportional to the tax rate.

$$\eta_{at}^i = (1 - \kappa\beta)\eta^i \quad 0 < \kappa \leq 1 \quad (3)$$

The maximum level of efficiency, η^i , is reached at a zero tax rate. An individual's production level is determined by his own investment in human capital and by general knowledge and skills in society. The accumulation of knowledge in skills follows from past production activities through a learning-by-doing process.

$$A_t = \frac{1}{n} \sum_i y_{t-1}^i \quad (4)$$

Individuals differ in their initial endowments. An individual's endowment upon birth at time t is given by $\varepsilon_t^i A_t$, with $\varepsilon_t^i (\geq 0)$ an identically and independently distributed random shock with mean one that represents an individual's access to general knowledge. The random endowment shocks cause the initial inequality in the models. The endowment of individual i is given by

$$w_t^i = \varepsilon_t^i A_t \quad (5)$$

An individual can either consume the efficiency units of labour he is endowed with or he can invest them into the production of future consumption goods (according to (2)).

For the remaining part of the paper, we will assume that the individual efficiency levels (the η^i s) are either identical over the entire population or a monotonically increasing function of the random shocks (the ε^i s) (cf. infra). This assumption is not essential for the results but simplifies the derivations substantially.

Current consumption is equal to the amount of initial endowment augmented with the amount of borrowing (b_t^i), less the amount of investment in human capital

$$c_t^i = w_t^i + b_t^i - h_t^i \quad (6)$$

³ The subscript 'at' refers to 'after taxes'. As we will argue below, taxes affect efficiency.

A government can redistribute income. Individual income is submitted to the tax rate β . Next, everybody receives an equal share of total tax revenue. This means that each individual gets a fraction β of average income in the society, \bar{y}_i . Redistribution has an indirect cost due to its negative impact on efficiency.

Future consumption equals future production after redistribution, less debt repayment

$$d_i^j = (1 - \beta)y_i^j + \beta\bar{y}_i - (1 + r_i)b_i^j \quad (7)$$

with $1+r_i$ the (gross) market interest rate, which is endogenously determined by the loan market clearing condition that the sum of net-borrowings must equal 0.

$$\sum_i b_i^j = 0 \quad (8)$$

Each individual will spread his endowment over consumption in period 1 and production in period 1 (consumption in period 2) to maximize intertemporal utility.

Next we complete the model in two different ways. Firstly, we derive the complete markets model with and without median voter. Secondly, we derive the imperfect markets model.

4.2.2. The complete markets model with and without median voter

In the case of a perfect capital market, no credit constraints exist. Everybody can lend or borrow.

Using (6) and (7) into (1), an individual's decision problem becomes

$$\text{Max} \left\{ \ln(w_i^j + b_i^j - h_i^j) + \rho \ln((1 - \beta)y_i^j + \beta\bar{y}_i - (1 + r_i)b_i^j) \right\} \quad (9)$$

$$\text{w.r.t. } b_i^j, h_i^j$$

$$\text{s.t. (8) and (2)}$$

In appendix A we derive the expression for steady state growth in this economy:

$$g_y = \ln(1 - \kappa\beta) + (1 - \alpha) \ln \left(\frac{1}{n} \sum_i (\eta^i)^{1/(1-\alpha)} \right) + \alpha [\ln(\rho\alpha(1 - \beta)) - \ln(\rho\alpha(1 - \beta) + 1)] \quad (10)$$

The efficiency levels are constructed in the following manner

$$\eta^i = (a + b\varepsilon^i)^{1-\alpha} \quad a, b > 0$$

Since we imposed that the mean of the endowments shocks equals one, we get that

$$\frac{1}{n} \sum_i (\eta^i)^{1/(1-\alpha)} = \frac{1}{n} \sum_i (a + b\varepsilon^i) = \frac{1}{n} \left(na + b \sum_i \varepsilon^i \right) = \frac{1}{n} (na + nb)$$

Substituting this expression in (10) results in

$$g_y = \ln(1 - \kappa\beta) + (1 - \alpha) \ln(a + b) + \alpha [\ln(\rho\alpha(1 - \beta)) - \ln(\rho\alpha(1 - \beta) + 1)] \quad (11)$$

Inequality does not appear directly in (11), indicating that inequality can only indirectly influence growth (i.e. if one of the parameters is defined as a function of inequality).

A first model we consider in our simulation is this 'pure' complete markets model. Inequality does not influence growth. Also note that the income of the previous period is absent in the growth expression.

Next we add a link between inequality and growth to the complete markets model. We assume that the tax rate β is determined by a simple majority voting procedure. With single peaked preferences, the outcome of this process is the median ranked blisspoint⁴. It is easy to check that this is the tax rate preferred by the individual with median endowments (and median income). Hence the tax rate is found as the solution to the following maximization problem:

$$\begin{aligned} & \text{Max} \left\{ \ln(w_i^m + b_i^m - h_i^m) + \rho \ln((1 - \beta)y_i^m + \beta \bar{y}_i - (1 + r_i)b_i^m) \right\} \\ & \text{w.r.t. } \beta \end{aligned} \quad (12)$$

⁴ While it is not straightforward to formally derive the sign of the second derivative of utility to taxes, a numerical evaluation indicates that this derivative is consistently negative for all values of the tax rate and varying levels of efficiency and endowment. Thus the necessary condition of single-peaked preferences is satisfied.

where the superscript m refers to the individual with median endowments, our median voter. Contrary to the 'pure' complete markets model, the tax rate now appears as an endogenous variable.

One can show that the tax rate will be increasing in the relative distance from the median pre-tax income to the mean pre-tax income (see also Meltzer and Richard (1981), Bénabou (1996)). Moreover the first derivative of (11) with respect to the tax rate is negative

$$\frac{\delta g_y}{\delta \beta} = \frac{-\kappa}{1-\kappa\beta} + \frac{-\alpha}{\rho\alpha(1-\beta)} + \frac{\alpha}{\rho\alpha(1-\beta)+1} < 0 \quad (13)$$

Thus in the complete markets model with median voter, there is a negative link between growth and inequality.

4.2.3. The imperfect markets model

In the third model we suppress the credit market and assume a zero tax rate. This amounts to the introduction of absolute credit constraints: every agent can only spend his income in both time periods. Expressions (6) and (7) are thereby reduced to:

$$c_t^i = w_t^i - h_t^i \quad (14)$$

$$d_t^i = y_t^i \quad (15)$$

After substitution of (14) and (15) in (1) and using the expressions for endowment (5) and production (2), we get that each individual solves the following maximization problem

$$\text{Max} \left\{ \ln(e_t^i A_t - h_t^i) + \rho \ln \left(\eta (A_t)^{1-\alpha} (h_t^i)^\alpha \right) \right\} \quad (16)$$

$$\text{w.r.t. } h_t^i$$

Note that we assume an identical efficiency level for all agents in the economy ($\eta^i = \eta, \forall i$). In contrast to the complete markets model assuming different efficiency levels does not improve the model and only complicates matters.

Solving (16) results in

$$y_i^j = \eta A_i \left(\frac{\alpha \rho}{1 + \alpha \rho} \right)^\alpha (\varepsilon_i^j)^\alpha \quad (17)$$

Or in terms of the growth rate

$$g_y = \ln \left(\frac{\sum_i y_i^j}{\sum_i y_{i-1}^j} \right) = \ln \left(\frac{\sum_i y_i^j}{n A_i} \right) = \alpha \ln \left(\frac{\rho \alpha}{1 + \rho \alpha} \right) + \ln \eta + \ln \left(\frac{1}{n} \sum_i (\varepsilon_i^j)^\alpha \right) \quad (18)$$

The final term of the above expression contains the individual 'endowment shocks'. As the mean of ε is by definition always equal to one and the function x^α is a concave function of x ($0 < \alpha < 1$), mean preserving spreads of the distribution of ε will decrease the value of the sum. Therefore, the last term in (17) can be interpreted as an equality index.

As we have already argued in the introduction, the relevance of an inequality measure depends on the underlying model. In the complete markets model, the position of the median voter to the mean will matter, in the imperfect markets model the way (in-)equality comes in is through the last term in (18). We will later show this term is closely related to an adjusted Atkinson index.

A priori one would expect that inequality measures will do a better job in detecting the relationship between (in-)equality and growth, the more they are correlated with these concepts.

Also note that the relation between inequality and growth as represented in (18) and the relationship between redistribution (or, given the voting process, between inequality) and growth as represented in (11) is not a linear one. Thus next to the use of the wrong inequality index, non-linearity might further undermine the reliability of the estimates.

4.3. Simulation set-up and calibration

We use the above models to generate individual incomes of the inhabitants and aggregate economic growth for a group of countries. From the individual incomes we derive several inequality measures. We repeat this procedure 1000 times for each inequality-growth model. Next we use the simulated data in a cross-country growth regression.

In the first part of this section we comment on the included countries, the size of their respective populations and GDP level. We clarify how the endowment shocks are generated and motivate the choice of parameter values. We also demonstrate that we can look at income inequality instead of inequality of endowments (which is the driving force in the models). In the second part of the section we amplify on the regression specification and the error term added to the simulated growth rate.

Figure 1 gives a stylised overview of the simulation set-up and the regressions.

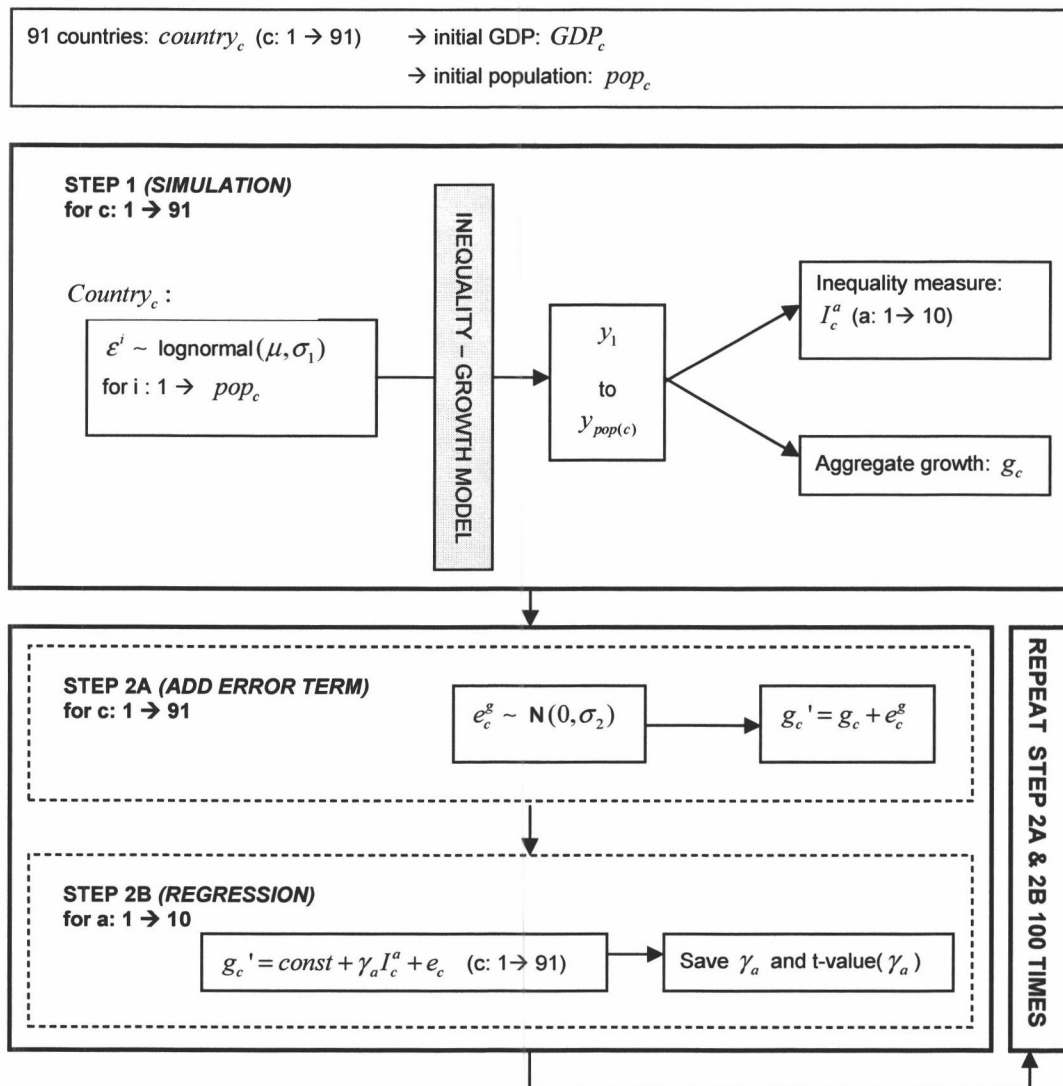
We include 91 countries in the simulation. The selection is based on population statistics provided by the World Bank. First we set the population of the biggest country ("China") equal to 5000. Next, the population of other countries was determined to reflect the real ratio of those countries' population and the Chinese population in 1999. Countries with a rescaled population below 30 were deleted from the sample⁵. The initial GDP per capita of a country in the simulation is its GDP per capita in US\$ of 2001 (PPP).

The 'endowment shocks' of the individuals in a country are randomly drawn from a log normal distribution. A log normal distribution seems to resemble nicely the kernel density estimates of the income distributions for a group of European countries by Papatheodorou *et al.* (2003) and it is generally accepted that it describes the main part of the income distribution quite well -see, e.g., Brown (1991, p.415). The standard deviation of the underlying normal distribution is chosen such that the gini coefficient of the log normal distribution from which each country's endowments are drawn, corresponds to the reported gini coefficient by Deininger and Squire (1996)⁶. Björklund (1993) shows that the value for the gini coefficient based on lifetime income (which better fits our OLG model) is lower due to income mobility over the longer period. His results indicate that compared to a gini coefficient based on annual income the difference amounts to about 25%. Since the Deininger and Squire data are based on annual income, we aim at 75% of the reported gini coefficient. As the population can be quite limited, i.e. small sample errors can be quite large, we check whether the gini coefficient in the simulated sample differs no more than 4 percentage points from the adjusted Deininger and Squire gini coefficient. If this condition is violated, we repeat the draw. Finally, as the mean of the endowment shocks should equal one, we divide each value by the sample mean.

⁵ The smallest country included is Chad, with an actual population of about 7.5 million.

⁶ Aitchison and Brown (1973) give a formula that relates the parameters of the lognormal distribution to its mean and gini coefficient. Note also that for some countries the value of the gini coefficient is not reported (e.g. Burkina Faso). For those countries we base the value on what we believe to be similar countries. See also chapter 2 for a practical guide to inequality measures.

Figure 1: A stylised overview of the simulation and the regressions



We repeat this entire procedure 1000 times for each inequality-growth model. The values reported in the tables are the average values of the coefficients (γ_a) and their t-values.

To check whether our simulation results draw heavily on the choice of the log normal distribution, we repeat the simulations randomly drawing endowment shocks from the frequency table of the US disposable personal incomes in 1991⁷. Ideally one would like to use the real income distributions of all countries included in our set up. However, we lack the data to do so. As we only use one frequency table, the differences in inequality between countries are merely caused by sampling error. Although the endowment shocks of the different countries are drawn from the same initial distribution, the sample distributions of these shocks differ substantially which allows us to identify statistically the effect of differences in inequality on growth.

The parameter related to the rate of time preference, ρ , is set at 0.4, which coincides with a value of about 0.97 on a yearly basis over 30 years. This value is in line with the real-business-cycle literature and is frequently used in the literature on economic growth (e.g., Rillaers (2001)).

For the value of α we do not have a clear reference value. We choose a value of 0.5 attaching equal weights to the level of knowledge and skills in a society and the individual investment effort. With values for α of this magnitude, non linearity problems in the specification of the growth regression are very small.

The κ -parameter, which captures the efficiency loss due to taxation, is fixed at 0.2.

The efficiency levels are chosen to obtain a yearly growth rate of approximately 1.8% (the world growth rate over the past 30 years). This means that an economy grows by 70% over a 30 year period.

In the models income inequality and growth are measured at the same point in time, whereas in empirical work growth is linked to initial levels of inequality. However, in the above models it does not really matter at what stage we measure growth, as our data come from economies that are in steady state. Moreover in real life studies one measures inequality at the beginning of the period to deal with potential endogeneity. In our approach we know that there is no feedback from growth to inequality.

Note also that in the models growth is influenced by inequality of endowments rather than by inequality of incomes. Inequality of endowments is perfectly transformed into inequality of incomes after debt repayments in the complete markets model (with or without median voter) as incomes are a linear transformation of endowments. That is not the case in the imperfect markets model. In that model incomes are also a monotonic transformation of endowments, but the transformation is concave. Hence, inequality of incomes will be lower than inequality

⁷ We derived the income distribution in percentiles from the data provided by Gottschalk and Smeeding (2000) and Atkinson (2003).

of endowments. However, we choose to maintain inequality of incomes as the independent variable in our simulation set up as this better matches existing empirical work. Irrespective of the mechanism behind the inequality growth relationship, most authors have been using income inequality. Some authors have claimed that inequality of incomes is not the best choice. Deininger and Squire (1998) argue that the distribution of assets is more relevant and show that land distribution outperforms income distribution as a proxy for asset distribution. Castelló and Doménech (2002) propose human capital inequality as a proxy for inequality of wealth. To check whether our results depend on this specification error, we repeated the simulation on the basis of endowments. As most results were similar we only report the isolated differences in the discussion.

Once the growth rates have been simulated, we add a randomly drawn error term to them and estimate a cross-country regression of the type

$$g_c = \delta + \gamma * ineq_c + e_c \text{ with } e \sim N(0, \sigma)$$

in which growth is regressed on $ineq_c$, a measure of income inequality. The inclusion of a constant term, δ is motivated by (11) and (18). None of the parameters (ρ, α, \dots) varies over countries. Allowing them to vary, would mean introducing a country specific fixed effect. OLS estimates would then be biased.

The size of the standard deviation of the error term, σ , determines the amount of noise that is added to the data. The value of σ will affect the standard error of γ , which is a measure of the accuracy of the estimator. We first introduce the matrix X

$$X = \begin{pmatrix} 1 & ineq_1 \\ 1 & ineq_2 \\ \vdots & \vdots \\ 1 & ineq_C \end{pmatrix} \text{ (with } C, \text{ the number of countries).}$$

Now we can write the standard error of γ as

$$\sigma_\gamma = \sqrt{\sigma^2 c_{22}} \text{ with } c_{22} \text{ the (2,2) element in } (X'X)^{-1}$$

Increasing the variance of the error term does not affect c_{22} , so the increase is directly translated in a higher standard error which reduces the value of the t-test for the significance of the regression coefficient. It does not change the relative values of the t-test when different inequality measures are used, however. Therefore, we focus on the performance of different inequality and poverty measures, relatively to the performance (t-statistic) of the ideal measure of inequality. The higher this relative t-statistic, the more likely that that coefficient will turn out to be significant in the regression.

We choose the value of the standard deviation such that the theoretically correct inequality measure explains 65% of the variance of the growth rate. This value corresponds with the R^2 s of the cross section growth regressions reported by Barro and Sala-i-Martin (1995). Of course, those authors include various regressors in their set up to capture a complex reality. In our model matters are less complicated as only income inequality matters. We assume that our single explanatory variable explains a similar fraction of the variance of total growth as their robust variables. So by construction the coefficient of the theoretically correct inequality measure should be highly significant. The drawback of raw t-statistics is their dependency on the assumed fraction of the variance explained by the inequality measure. Hence we report the absolute values of the t-statistic and the number of times we found a significant and expected sign in the summary tables mainly for illustrative purposes.

We carry out the 'estimation procedure' 100 times per initial simulation run for each of the 10 included inequality measures. Hence we perform 1.000.000 regressions for each model. We check the robustness of our results in a more limited set up, based on 200 simulation runs (200.000 regressions).

Before presenting and discussing the simulation results, we first elaborate on the included inequality measures in the next section.

4.4. Measures of inequality and poverty

We include 10 different measures of inequality and poverty. We cover a broad range of measures that have been used and discussed in the empirical or theoretical literature on income inequality measurement. We refer to Sen (1973) and Cowell (2000) for a detailed discussion of the inequality measures and their properties.

4.4.1. Inequality measures

The decile ratio (D) is the ratio of the income obtained by the agent that is at the 90th percentile in the income distribution to the income of the person at the 10th percentile.

$$D = \frac{P_{90}}{P_{10}}$$

The gini coefficient (G) is the most frequently used measure in empirical work (see chapter 2). It is particularly sensitive to the middle sections of the income distribution.

$$G = \frac{1}{2n^2 \bar{y}} \sum_i \sum_j |y_i - y_j|, \text{ with } \bar{y} \text{ the mean income}$$

The ratio of median to mean income (M) pops up as the ideal measure of inequality in the complete markets model with a median voter.

$$M = \frac{\text{median}(y)}{\bar{y}}$$

The Atkinson index (A_e) is the most common measure of income inequality based on a classic social welfare function and allows the expression of a wide range of welfare judgements.

$$A_e = 1 - \left(\frac{1}{n} \sum_i \left(\frac{y_i}{\bar{y}} \right)^{1-e} \right)^{1/(1-e)}$$

The parameter e is the parameter of inequality aversion. If e is equal to zero, the income distribution does not matter for social welfare: A_0 always equals zero. If e goes to infinity only the income of the poorest person matters, which is usually labelled the Rawlsian position. We use three values for e : 0.5, 1.5 and 5. The first is not very inequality averse. The second has been put forward as a reasonable approximation of peoples' attitudes with respect to inequality. The third is already quite Rawlsian.

There is a close link between a specific Atkinson index and the perfect measure of inequality for the imperfect markets model. From the definition of A_e , we have that

$$[1 - A_e]^{1-e} = \frac{1}{n} \sum_i \left(\frac{y_i}{y} \right)^{1-e}$$

Taking the logarithm of the left and right hand side of this equation and choosing $\alpha = 1 - e$, it follows that

$$\ln \left(\frac{1}{n} \sum_i \left(\frac{y_i}{y} \right)^\alpha \right) = \alpha \ln [1 - A_{[1-\alpha]}]$$

Now consider the final term of expression (18) and recall that the endowment shocks are on average equal to one,

$$\ln \left(\sum_i (\varepsilon_i')^\alpha \right) = \ln \left(\sum_i \left(\frac{\varepsilon_i'}{\varepsilon} \right)^\alpha \right)$$

Which shows that as y^j varies with ε^j , the adjusted Atkinson index $AA_{1-\alpha} = \ln[1 - A_{1-\alpha}]$ is a perfect measure (of equality) for the imperfect markets model⁸. It must be stressed that this result is very specific to the specific functional forms we used. Moreover, the value of the "inequality aversion parameter" equals $1 - \alpha$ and is no longer a normative, but a positive parameter.

4.4.2. Poverty measures

Next to the above inequality measures we also include some poverty indices in the simulation. To do this, we first have to determine a poverty line, π . According to UNDP data (UNDP (2003)) somewhat more than 10% of the countries has a mean income below the 2\$ a day poverty line. We determine the poverty line at 20% of the mean initial world per capita GDP in order to get a similar amount of countries below the poverty line after growth (which is about 75%). The choice of the poverty line will always be rather arbitrary, but our choice closely resembles existing poverty rates.

The head count index (H) is the ratio of the number of people below the poverty line and the total population

⁸ Remember that any monotonic transformation of an inequality index is also an inequality index. The inequality ordering implied by $AA_{1-\alpha}$ is the same as that implied by $A_{1-\alpha}$.

$$H = \frac{q}{n}, \text{ with } q \text{ the number of people below the poverty line.}$$

The Poverty gap (Q) takes into account the depth of poverty by aggregating the individual poverty gaps $(\pi - y_i)$

$$Q = \frac{1}{q} \sum_i (\pi - y_i) \delta_i \quad \text{where} \quad \begin{cases} \delta_i = 1 & \text{if } y_i < \pi \\ \delta_i = 0 & \text{if } y_i \geq \pi \end{cases}$$

The Foster index (F_α) aggregates normalised individual poverty gaps $((\pi - y_i)/\pi)$, but gives a larger weight to larger gaps. This weight, $\alpha-1$ represents the “inequality aversion” amongst the poor. In the simulation we set α equal to two.

$$F_\alpha = \frac{1}{q} \sum_{k=1}^q \left(\frac{\pi - y_k}{\pi} \right)^\alpha \quad \text{with } \alpha \geq 1$$

The correlations between the inequality and poverty measures derived from the simulated incomes are presented in appendix B.

4.5. Discussion and results

4.5.1. A comment on the use of the log normal distribution

The log normal distribution is completely determined by (a non-linear transformation of) the mean and the variance of the underlying normal distribution (Aitchison and Brown (1973)). This implies that Lorenz curves that are derived from different log normal distributions will never intersect. As the Lorenz curve is independent of the mean, it will be fully determined by the variance. The larger the variance of the log normal distribution, the more convex the Lorenz curve will be. This means that, e.g., different Lorenz curves cannot result in the same gini coefficient. Consequently, a scalar measure should be able to capture the entire distribution. Since the inequality measures we presented in section 4.4 are mutually related in a non-linear way, they should all exhibit an identical explanatory power as long as the empirical specification is flexible enough (e.g., if the inequality measures are included as a

higher order polynomial). In that case, our simulation set up would be unsuitable to answer the questions we posed in the introduction.

However, small sample error is quite important in our simulation. The simulated income distributions differ substantially from the (asymptotic) log normal distributions. We carried out an additional simulation run and computed the number of intersecting Lorenz curves. Out of 4095 pairwise comparisons of Lorenz curves, 542 intersected. Hence, the set up seems appropriate after all.

4.5.2. *The complete markets model without median voter*

In this model income inequality does not affect economic growth. The simulated growth rate does not vary over countries. We detect a significant relationship between the inequality and poverty measure and the growth rate in about 5% of the regressions. This is exactly the 'type I'-error we should find at the 95% significance level in the absence of a relationship between inequality and growth.

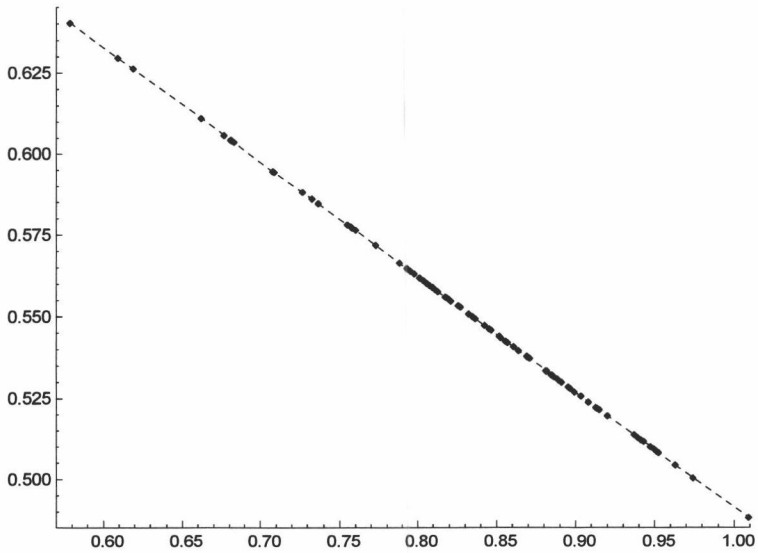
4.5.3. *The complete markets model with median voter*

Graph 1a reveals the linear relationship between the ratio of median to mean income and the tax rate preferred by the median voter. Equation (13) demonstrates that the tax rate negatively influences economic growth. Differentiating this expression with respect to β , we find that

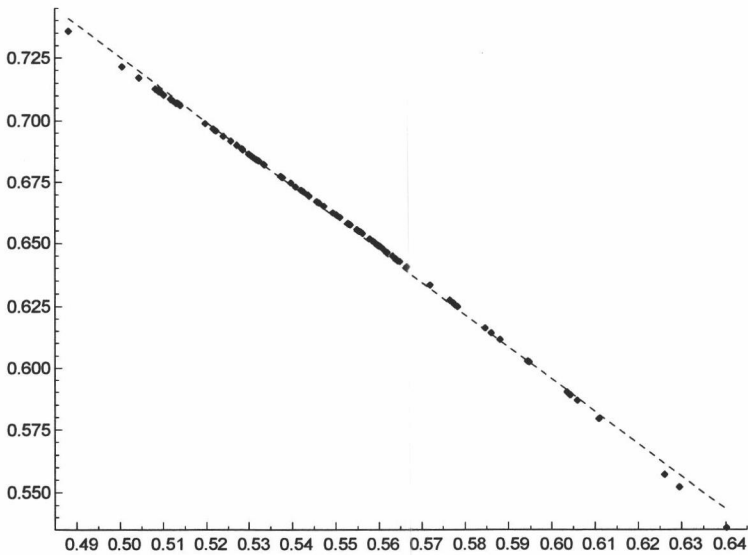
$$\frac{\partial^2 g_y}{\partial \beta^2} = \frac{-\kappa^2}{(1-\kappa\beta)^2} + \frac{-\alpha^2 \rho}{(\rho\alpha(1-\beta))^2} + \frac{\alpha^2 \rho}{(\rho\alpha(1-\beta)+1)^2} < 0$$

Consequently the relationship between the tax rate and economic growth is concave. Graph 1b shows, however, that the deviation from linearity is very small. We thus have a negative relationship between the ratio of median growth to mean income that is almost linear. Therefore, the error one makes by using a linear estimation technique will be fairly small.

Graph 1a: The ratio of median and mean income (x-axis) and the tax rate (y-axis) in the complete markets model with median voter



Graph 1b: The tax rate (x-axis) and the simulated growth rate (y-axis) in the complete markets model with median voter



An overview of the regressions on the simulated data is reported in table 1. We report the average coefficient, the average t-value, the percentage of regressions in which we find the expected sign for the estimated coefficient of the inequality measure (significant at the 95% level) and the t-ratio of the coefficient relative to the t-ratio of the ideal inequality measure.

Recall that the absolute values are determined by the simulation set up. Still they offer some interesting information. Firstly, they illustrate that the measures which we identify as theoretically optimal are indeed very potent under the assumptions made. Secondly, as 'inequality' explains 65% of the growth rate in our set up (by construction), a good proxy for inequality should always be able to detect the correct relationship.

Table 1: Complete markets model with median voter - simulation results (log normal)

	G	D	M	A _{0.5}	A _{1.5}	A ₅	AA _{0.5}	H	Q	F ₂
coef.	-0.036	-0.032	0.040	-0.036	-0.035	-0.031	0.036	-0.018	-0.024	-0.024
t-value	-9.696	-7.934	12.773	-9.755	-9.220	-7.497	9.745	-3.605	-5.190	-5.323
relative t	-0.761	-0.623	1	-0.766	-0.724	-0.589	0.765	-0.283	-0.408	-0.418
% signif.	100	100	100	100	100	100	100	97.3	99.9	99.9

Notes: - The coefficients and t-values in the table are averages.
 - The 'relative t' is the ratio of t-value of the coefficient of an inequality measure and that of the theoretically optimal inequality measure (shaded in grey).
 - We use a one-sided significance test: so if the estimated relationship has a significant but (theoretically) wrong sign, it is not included in '% signif'.

Paired-samples t-tests show that the differences in average t-values are always significant (at the 1%-level)⁹. Hence we can confidently use the relative t-ratio to evaluate the performance of the inequality and poverty measures in the model.

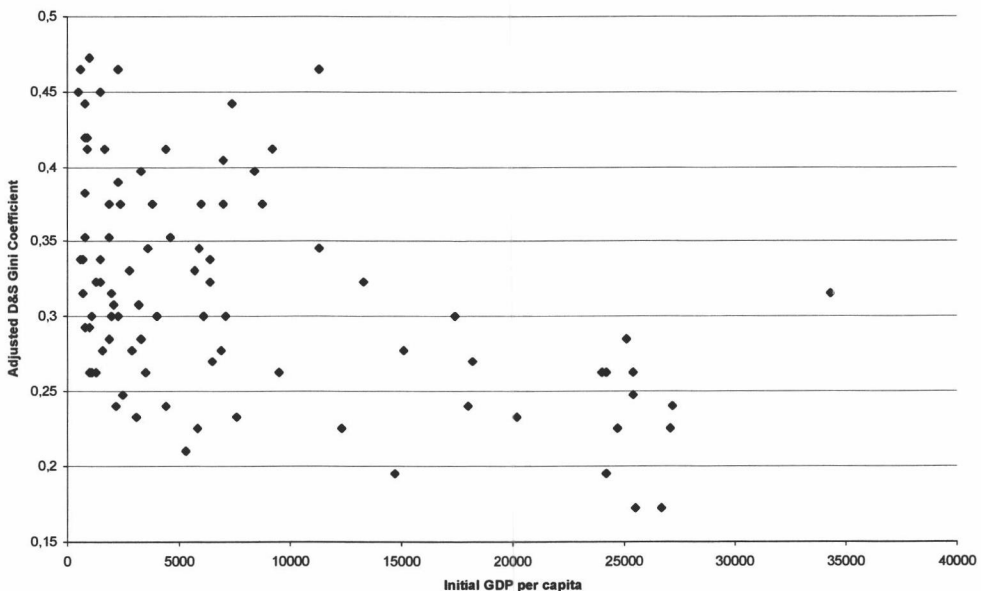
The ratio of median to mean income (M), the measure advanced by the theory, indeed outperforms all other measures. The performance of some other measures is quite reasonable. The relative t-ratio for the gini coefficient (G), two Atkinson indices ($A_{0.5}$, $A_{1.5}$) and the adjusted Atkinson index ($AA_{0.5}$) is above 0.7. For the decile ratio (D) and the third Atkinson index (A_5) it is higher than 0.55. The poverty measures perform less well. Their relative t-ratio stays below 0.45.

While poverty measures perform worse than inequality measures, they still manage to pick up something of what is going on. At first sight, this is surprising. Two channels drive this result. First, given mean income, with the log normal distribution, a higher gini index is always accompanied by a higher level of poverty. Second, with an absolute poverty line,

⁹ The results of the t-tests are available upon request.

initial GDP per capita matters a lot for the value of the poverty indices. As graph 2 demonstrates, the countries with the highest adjusted Deininger and Squire gini coefficients are mostly countries with a low initial GDP per capita (the correlation between initial GDP per capita and the gini coefficient is -0.48). In the countries with the lowest initial GDP per capita a lot of people will have an income below the poverty line. Hence the poverty indices for those countries will be large. Given the fact that we find a strong negative relationship between the gini coefficient and economic growth, the correlation between initial GDP per capita and the gini coefficient helps to explain the relationship between the poverty indices and economic growth.

Graph 2: Initial GDP per capita and the gini coefficient



We repeat the simulation using the US income distribution in 1991 instead of the log normal distribution. We no longer enforce that the gini coefficient of the endowments matches reality. Hence we can check whether our results depend heavily on the distribution of the endowments shocks and assess the argument we provided in the previous paragraph. The results are presented in table 2.

Table 2: Complete markets model with median voter - simulation results (US distribution)

	G	D	M	A _{0.5}	A _{1.5}	A ₅	AA _{0.5}	H	Q	F ₂
coef.	-0.023	-0.002	0.029	-0.022	-0.018	-0.008	0.022	-0.003	-0.001	-0.002
t-value	-7.922	-0.571	13.948	-7.746	-5.608	-2.145	7.760	-0.702	-0.389	-0.413
relative t	-0.568	-0.041	1	-0.555	-0.402	-0.154	0.556	-0.050	-0.028	-0.030
% signif.	100	13.7	100	100	99.9	55.8	100	10.9	7.4	8.2

Notes: See table 1.

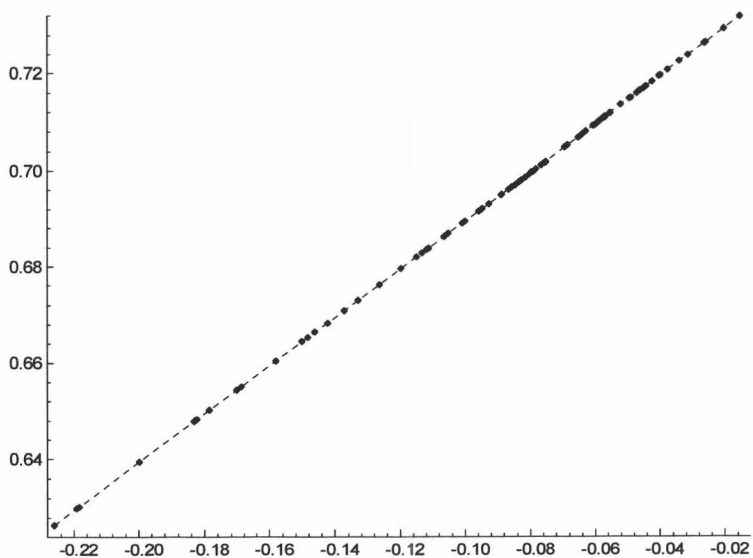
To explain the decrease in relative t-statistics for the non ideal measures, note that the US distribution has more probability mass at the bottom and the top end of the distribution than the log normal distribution (see appendix C). Different measures of inequality react differently to outliers in the distribution (see, e.g., Cowell and Victoria-Feser (1996)). The larger dispersion in the empirical distribution therefore reduces the correlation between different measures of inequality. It makes the non-ideal inequality measures worse proxies for the ideal (and always well performing) measure, such that their estimated coefficient will be biased towards zero, and their t-statistic decreases.

It is striking that the poverty indices and the decile ratio perform much worse. This is due to the fact that the endowment shocks are no longer drawn from log normal distributions and the gini coefficients are no longer forced to match the real world. Hence both channels referred to above disappear. Only the gini coefficient and the (adjusted) Atkinson index with inequality aversion parameter 0.5 have a relative t-ratio of more than 0.5. Of the considered inequality measures, in particular the decile ratio performs much worse and seems more sensitive to the specific form of the income distribution.

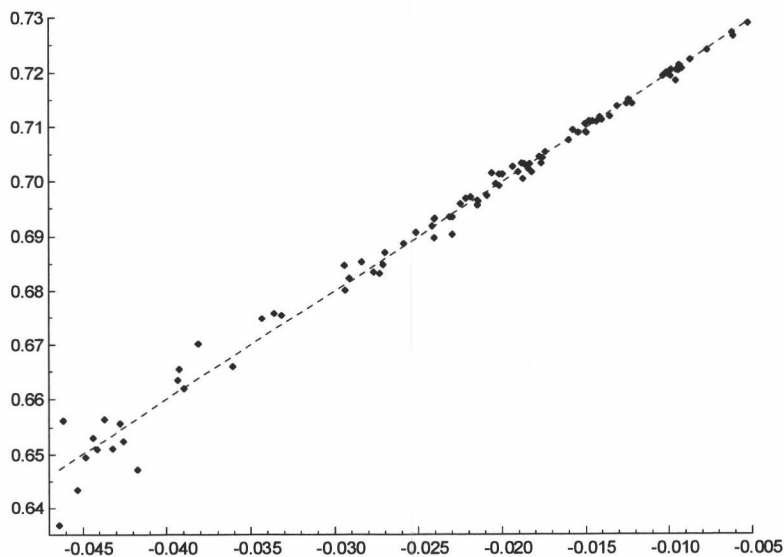
4.5.4. The imperfect markets model

Graph 3a demonstrates the perfect relationship between the adjusted Atkinson index and the growth rate. The plotted adjusted Atkinson index was calculated on the basis of the endowment shocks. As we already mentioned, we choose to look at incomes instead of endowments. As one can immediately deduce from expression (17), incomes are a strictly increasing concave transformation of endowments. Graph 3b, in which we plot the adjusted Atkinson index derived from income data against the growth rate, provides an indication of the error we introduce by looking at incomes rather than endowments. We checked the impact of this error on the simulation results. We do not present the complete results of this robustness check, but only include its main implications in the discussion.

Graph 3a: The adjusted Atkinson index (x-axis) and the simulated growth rate (y-axis) in the imperfect markets model (endowments)



Graph 3b: The adjusted Atkinson index (x-axis) and the simulated growth rate (y-axis) in the imperfect markets model (incomes)



Again paired-samples t-tests confirm that the differences in t-values are highly significant. The theoretically best measure ($AA_{0.5}$) performs best, but it does not outperform the unadjusted Atkinson index. This can easily be explained since the former measure comes close to being a linear transformation of the latter. The non-linearity in the relationship between the $A_{0.5}$ -index and the growth rate is far too small to cloud that relationship when approached by a linear estimation technique. The performance of the other inequality indices (except M), with relative t-ratios above 0.85, is surprisingly good. The ratio of the median to the mean *income* clearly performs worse than in the estimates for the complete markets models. However, if we consider the median to the mean *endowment*, it does much better (relative t-ratio of 0.74). The ratio of median to mean income is the only important victim of the measurement error we introduced by looking at incomes (the performance of some poverty indices is only slightly affected).

The performance of the poverty measures is quite similar to their performance in the complete markets model.

Table 3: Imperfect markets model - simulation results (log normal)

	G	D	M	$A_{0.5}$	$A_{1.5}$	A_5	$AA_{0.5}$	H	Q	F_2
coef.	-0.022	-0.021	0.012	-0.022	-0.022	-0.021	0.022	-0.009	-0.011	-0.016
t-value	-12.071	-10.838	6.492	-12.485	-12.223	-11.047	12.486	-3.295	-3.955	-4.088
relative t	-0.967	-0.868	0.520	-0.999	-0.979	-0.885	1	-0.264	-0.317	-0.327
% signif.	100	100	99.9	100	100	100	100	96.9	99.5	99.5

Notes: See table 1.

If we start from the US distribution we again see that all relative t-ratios decrease. This decrease is most striking for the poverty measures and the decile ratio, for the same reason as given in the complete markets model.

Table 4: Imperfect markets model - simulation results (US distribution)

	G	D	M	$A_{0.5}$	$A_{1.5}$	A_5	$AA_{0.5}$	H	Q	F_2
coef.	-0.013	-0.005	0.007	-0.013	-0.012	-0.008	0.014	-0.000	-0.001	-0.001
t-value	-10.917	-2.863	4.120	-12.105	-10.066	-4.970	12.127	-0.220	-0.319	-0.374
relative t	-0.900	-0.236	0.340	-0.998	-0.830	-0.410	1	-0.018	-0.026	-0.031
% signif.	100	77.6	96.8	100	100	100	100	6.6	8.2	8.4

Notes: See table 1.

4.5.5. *Some further comments*

It does not come as a surprise that poverty measures do not perform well as explanatory variable in the estimates. Although we have included a model in which growth is lowered due to credit constraints, it is not a model that really 'favours' poverty measures. The reason is straightforward: everybody, poor and rich, is affected by credit constraints as there is no credit market in the economy. It seems likely that poverty measures will be more useful in a framework where borrowing is possible but limited by the individual wealth that must serve as collateral. Due to incentive problems, it is not easy to obtain a loan to finance one's investment in human capital, such that children from low income families frequently face borrowing constraints (see, e.g., Becker and Tomes (1986)). In these kind of models the fraction of people below a certain income threshold will determine the impact on the growth rate. The shape of the income distribution above the threshold will only have a limited effect.

In section 4.2 we noted that the relationship between inequality and economic growth was not always linear. Hence using a linear estimation technique (OLS) might introduce an additional error next to the one linked to the imperfectness of the inequality proxy. The simulations indicate that the importance of this error should not be exaggerated in our model specifications. However, it seems likely that in some alternative model specifications non-linearity can be a much bigger issue. Hence its relevance for empirical work can not be dismissed.

4.6. **Conclusions**

How does income inequality influence economic growth? If one wants to provide an empirical answer to this question by means of a standard growth regression, one collapses an entire income distribution into a scalar measure. The literature on income inequality measurement offers several possibilities going from more descriptive measures (e.g., deciles ratio) to theoretically derived ones (e.g., Atkinson index). Apart from these inequality measures, poverty measures might also be informative, e.g., if the credit constraints of the poor are at the heart of the inequality-growth relationship. Does it matter which measure one uses, and if so, how damaging is the impact of using the wrong one?

Since we started from some very crude theoretical models, it might be dangerous to generalize our results. Even after taking into account the limitations of the simulation approach, some carefully-worded conclusions seem valid.

The optimal measure depends on the channel through which the income distribution affects growth. In the complete markets model with median voter the ratio of the median to the mean income performs best. In the imperfect markets model an (adjusted) Atkinson index with a low degree of inequality aversion is better suited. However, our simulation exercise indicates that other measures might also be fit for the job. The gini coefficient and the Atkinson indices with a low degree of inequality aversion turn out to be reasonably significant in our complete markets model with median voter, and are almost as good as the ideal measure in our imperfect markets model. Other inequality and poverty measures are less successful.

If one has data on several inequality measures, it might be possible to discriminate between different models of inequality and growth. If different measures confirm the negative relationship it can be informative to look at the level of significance of their respective coefficients. If the coefficient of the ratio of median to mean income has a higher significance than that of the (adjusted) Atkinson index with a low degree of inequality aversion, that seems to offer support for the relevance of the complete markets model with median voter. If the roles are reversed, the imperfect markets model seems more relevant. Of course, in the absence of an exhaustive list of potentially relevant inequality-growth models, this information can only serve as tentative support.

In practice, one typically has only a limited number of inequality measures available for empirical work. Often the choice is limited to the gini coefficient and the decile ratio. Our simulations support the current practice of using the gini in that case, but should also encourage researchers to construct other inequality indices. While the gini is quite able to recover the relationship between inequality and growth, to discriminate between alternative theories, other measures can be helpful.

The relative weights one attaches to the main conclusions will depend on how one evaluates the comment concerning the use of log normal distribution. If the simulation set up gives undue advantage to the inequality measures that are not theoretically optimal, it will be easier to discriminate between alternative models on the basis of inequality indices in reality. If, on the other hand, the simulation set up reflects reality quite well, the results should reassure empirical researches faced with a scarcity of inequality data.

4.7. References

- Aghion P., Howitt P. (1998): *Endogenous Growth Theory*, MIT Press, Cambridge
- Aitchison J., Brown J.A.C. (1973): *The Lognormal Distribution (4th Edition)*, Cambridge University Press, Cambridge
- Alesina A., Rodrik D. (1994): *Distributive Politics and Economic Growth*, Quarterly Journal of Economics, 109, p.465-490
- Atkinson A.B. (1970): *On the Measurement of Inequality*, Journal of Economic Theory, 2, p. 244-263
- Atkinson A.B. (2003): *Top Incomes in the United Kingdom over the Twentieth Century*, mimeo, 55p.
- Atkinson A.B., Brandolini A. (1999) : *Promise and Pitfalls in the Use of "Secondary" Datasets: Income Inequality in OECD countries*, mimeo, 60p.
- Barro R.J. (1999): *Inequality, Growth and Investment*, NBER Working Paper no. 7038, 37p.
- Barro R.J., Sala-i-Martin X. (1995): *Economic Growth*, McGraw-Hill, New York
- Becker G.S., Tomes N. (1986): *Human Capital and the Rise and Fall of Families*, Journal of Labour Economics, 4, p.1-47
- Bénabou R. (1996): *Inequality and Growth*, in: Bernanke B.S. and Rotemberg J. (eds.), *NBER Macroeconomics Annual 11*, MIT Press, p.11-92
- Björklund A. (1993): *A Comparison Between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89*, Review of Income and Wealth, 39(4), p.377-386
- Brown, H.P. (1991), *Egalitarianism and the Generation of Inequality*, Oxford University Press, Oxford
- Castelló A., Doménech R. (2002): *Human Capital and Economic Growth: Some New Evidence*, The Economic Journal, 112, p.187-200
- Cowell, F.A. (2000), *Measurement of Inequality*, in: Handbook of Income Distribution Vol. 1 (eds. Atkinson A.B., Bourguignon F.), Elsevier, Amsterdam, p.87-166
- Cowell F.A., Victoria-Feser M.P. (1996), *Robustness Properties of Inequality Measures*, Econometrica, 64, p.77-101
- Deininger K., Squire L. (1996): *Measuring Income Inequality: a New Data-Base*, World Bank Economic Review, 10 (3), p.565-591
- Deininger K., Squire L. (1998): *New Ways of Looking at Old Issues: Inequality and Growth*, Journal of Development Economics, 57, p.259-287
- Durlauf S.N. (2001): *Manifesto for a Growth Econometrics*, Journal of Econometrics, 100, p.65-69
- Forbes K.J. (2000): *A Reassessment of the Relationship Between Inequality and Growth*, The American Economic Review, 90 (4), p.869-887

Gobbin N., Rayp G. (2004a): *Income Inequality Data in Growth Empirics: From Cross-Sections to Time series*, UGent-FEB Working Paper no. 252, 31p.

Gobbin N., Rayp G. (2004b): *Inequality and Growth: Does Time Change Anything*, UGent-FEB Working Paper no. 230, 32p.

Gottshalk P., Smeeding T.M. (2000): *Empirical Evidence on Income Inequality in Industrialized Countries*, in: Handbook of Income Distribution Vol. 1 (eds. Atkinson A.B., Bourguignon F.), Elsevier, Amsterdam, p.261-307

Lindbeck A. (1985): *What is Wrong with the West European Economies*, World Economy, 8 (2), p.153-169

Lindbeck A. (1988): *Consequences of the Advanced Welfare State*, World Economy, 11 (1), p.19-37

Lindbeck A. (1993): *Work Disincentives in the Welfare State*, in: Lindbeck A., The Welfare State – The Selected Essays of Assar Lindbeck Vol. II, Edward Elgar, Hants

Meltzer A.H., Richard S.F. (1981): *A Rational Theory of the Size of Government*, Journal of Political Economy, 89 (5), p.914-927

Papathodorou C., Peristera P., Kostaki A. (2003): *Assessing Income Distribution using Kernel Estimates: A Comparative Study in Five European Countries*, paper presented at the '17th Annual Congress of the European Society for Population Economics', 13-15 June, New York City

Perotti R. (1994): *Income Distribution and Investment*, European Economic Review, 38, p.827-835

Perotti R. (1996): *Growth, Income Distribution and Democracy: What the Data Say*, Journal of Economic Growth, 1 (2), p.149-187

Persson T., Tabellini G. (1994): *Is Inequality Harmful for Growth*, The American Economic Review, 84 (3), p.600-621

Rillaers A. (2001): *Education and Income Inequality: The Role of a Social Protection System*, Journal of Population Economics, 14, p.425-443

Sen A. (1973): *On Economic Inequality – The Radcliffe Lectures Delivered in the University of Warwick 1972*, Clarendon Press, Oxford

UNDP (2003): *Human Development Report 2003*, Oxford University Press, New York/Oxford

APPENDIX A: Mathematical derivations
Complete markets model

$$\text{Max} \left\{ \ln(w_t^i + b_t^i - h_t^i) + \rho \ln((1 - \beta)y_t^i + \beta \bar{y}_t - (1 + r_t)b_t^i) \right\}$$

First derivative w.r.t. h^i

$$\frac{-1}{w_t^i + b_t^i - h_t^i} + \rho \frac{(1 - \beta)\alpha(h_t^i)^{\alpha-1} (A_t)^{1-\alpha} (1 - \kappa\beta)\eta^i}{(1 - \beta)y_t^i + \beta \bar{y}_t - (1 + r_t)b_t^i} = 0 \quad (\text{A1})$$

First derivative w.r.t. b^i

$$\frac{1}{w_t^i + b_t^i - h_t^i} - \rho \frac{(1 + r_t)}{(1 - \beta)y_t^i + \beta \bar{y}_t - (1 + r_t)b_t^i} = 0 \quad (\text{A2})$$

From (A1) and (A2) we know that

$$(1 + r_t) = (1 - \beta)\alpha(h_t^i)^{\alpha-1} (A_t)^{1-\alpha} (1 - \kappa\beta)\eta^i$$

From which we can derive that

$$h_t^i = A_t \left(\frac{\alpha(1 - \kappa\beta)(1 - \beta)\eta^i}{(1 + r_t)} \right)^{1/(1-\alpha)} \Rightarrow \sum_i h_t^i = A_t \left(\frac{\alpha(1 - \kappa\beta)(1 - \beta)}{(1 + r_t)} \right)^{1/(1-\alpha)} \sum_i (\eta^i)^{1/(1-\alpha)} \quad (\text{A3})$$

Substituting (A3) into the production function (2) and taking the sum over all individuals results in

$$\sum_i y_t^i = (1 - \kappa\beta)^{1/(1-\alpha)} A_t \left(\frac{\alpha(1 - \beta)}{1 + r_t} \right)^{\alpha/(1-\alpha)} \sum_i (\eta^i)^{1/(1-\alpha)} \quad (\text{A4})$$

From (A2) we can derive an expression for $(1 + r_t)$

$$\rho(1 + r_t)(w_t^i + b_t^i - h_t^i) = (1 - \beta)y_t^i + \beta \bar{y}_t - (1 + r_t)b_t^i$$

Taking the sum of this expression over all individuals, using the loan market clearing condition (8) and expression (A3), we get

$$nA_t(1+r_t) - (1+r_t)^{-\alpha/(1-\alpha)} A_t [\alpha(1-\kappa\beta)(1-\beta)]^{1/(1-\alpha)} \sum_i (\eta^i)^{1/(1-\alpha)} = \frac{1}{\rho} \sum_i y_i^j$$

Substituting (A4) into the right hand side, results in an expression for $(1+r_t)$

$$(1+r_t) = \left(\frac{1}{n}\right)^{(1-\alpha)} (1-\kappa\beta) \left(\sum_i (\eta^i)^{1/(1-\alpha)}\right)^{1-\alpha} \left[[\alpha(1-\beta)]^{1/(1-\alpha)} + \frac{1}{\rho} [\alpha(1-\beta)]^{\alpha/(1-\alpha)} \right]^{1-\alpha} \quad (A5)$$

Next we can substitute (A5) into (A4) to obtain an expression for total income. After some further manipulation we obtain:

$$\sum_i y_i^j = (1-\kappa\beta) A_t \left(\sum_j (\eta^j)^{1/(1-\alpha)}\right)^{-\alpha} \left(\frac{1}{n}\right)^{-\alpha} (\rho\alpha(1-\beta))^\alpha (\rho\alpha(1-\beta)+1)^{-\alpha} \sum_i (\eta^i)^{1/(1-\alpha)}$$

We divide total income by total production of the previous year (nA_{t-1} , see equation (4)) and take the logarithm to get an expression for the steady state growth

$$g_y = \ln \left(\frac{\sum_i y_i}{\sum_i y_{t-1}} \right) = \ln(1-\kappa\beta) + (1-\alpha) \ln \left(\frac{1}{n} \sum_i (\eta^i)^{1/(1-\alpha)} \right) + \alpha [\ln(\rho\alpha(1-\beta)) - \ln(\rho\alpha(1-\beta)+1)] \quad (A6)$$

Imperfect markets model

$$\text{Max} \left\{ \ln(w_t^i - h_t^i) + \rho \ln \left(\eta(A_t)^{1-\alpha} (h_t^i)^\alpha \right) \right\}$$

First derivative w.r.t. h^i

$$\frac{-1}{w_t^i - h_t^i} + \rho \frac{\eta \alpha (h_t^i)^{\alpha-1} (A_t)^{1-\alpha}}{\eta (h_t^i)^\alpha (A_t)^{1-\alpha}} = 0 \Rightarrow \frac{-1}{w_t^i - h_t^i} + \frac{\alpha \rho}{h_t^i} = 0$$

$$\Rightarrow h_t^i = \alpha \rho (w_t^i - h_t^i) \Rightarrow h_t^i = \frac{\alpha \rho}{1 + \alpha \rho} (w_t^i) \Rightarrow h_t^i = \frac{\alpha \rho}{1 + \alpha \rho} (\varepsilon_t^i A_t) \quad (\text{A7})$$

Substituting (A7) into the production function (2) (with $\eta^i = \eta$) results in

$$y_t^i = \eta (A_t)^{1-\alpha} \left(\frac{\alpha \rho}{1 + \alpha \rho} \right)^\alpha (\varepsilon_t^i A_t)^\alpha \Rightarrow y_t^i = \eta A_t \left(\frac{\alpha \rho}{1 + \alpha \rho} \right)^\alpha (\varepsilon_t^i)^\alpha \quad (\text{A8})$$

APPENDIX B:

Table B1: Correlations between inequality measures (Complete markets model with median voter - log normal distribution)

	G	D	M	A _{0.5}	A _{1.5}	A ₅	AA _{0.5}	H	Q	F ₂
G	1	0.884	-0.888	0.988	0.984	0.930	-0.983	0.475	0.673	0.670
D		1	-0.780	0.910	0.922	0.821	-0.913	0.456	0.635	0.671
M			1	-0.885	-0.861	-0.769	0.884	-0.431	-0.586	-0.594
A _{0.5}				1	0.991	0.912	-0.999	0.484	0.679	0.695
A _{1.5}					1	0.943	-0.989	0.483	0.693	0.707
A ₅						1	-0.906	0.424	0.683	0.670
AA _{0.5}							1	-0.486	-0.678	-0.697
H								1	0.805	0.841
Q									1	0.957
F ₂										1

Table B2: Correlations between inequality measures (Imperfect markets model with median voter - log normal distribution)

	G	D	M	A _{0.5}	A _{1.5}	A ₅	AA _{0.5}	H	Q	F ₂
G	1	0.955	-0.679	0.989	0.988	0.963	-0.988	0.395	0.464	0.470
D		1	-0.624	0.959	0.963	0.937	-0.959	0.382	0.463	0.473
M			1	-0.679	-0.657	-0.586	0.679	-0.289	-0.318	-0.324
A _{0.5}				1	0.998	0.968	-0.999	0.407	0.486	0.500
A _{1.5}					1	0.980	-0.998	0.405	0.487	0.502
A ₅						1	-0.967	0.386	0.479	0.489
AA _{0.5}							1	-0.408	-0.486	-0.501
H								1	0.917	0.916
Q									1	0.958
F ₂										1

APPENDIX C:

The cumulative distribution functions of the log normal distribution (corresponding to the adjusted D&S gini coefficient for the US) and the US empirical income distribution

