norden

NORDIC ECONOMIC POLICY REVIEW

Introduction
Torben M. Andersen and Jesper Roine

The Nordic welfare model in an open European labor market
Bernt Bratsberg and Knut Roed

Future Pathways for Labour Market Policy: Including the Excluded
Michael Svarer and Michael Rosholm

Economics of Innovation Policy
Tuomas Takalo and Otto Toivanen

Taxing mobile capital and profits: The Nordic Welfare States
Guttorm Schjelderup

Nordic family policy and maternal employment
Julian V. Johnsen and Katrine V. Løken

Education and equality of opportunity: What have we learned from educational reforms?
Helena Holmlund

Retirement and Health in the Nordic Welfare State
Nabanita Datta Gupta and Bent Jesper Christensen

The future of Welfare services: How worried should we be about Wagner, Baumol and Ageing?
Andreas Bergh

Ethnic fractionalization and the demand for redistribution – Potential implications for the Nordic model
Johanna Mollerstrom

The social upper classes under Social Democracy
Kalle Moene

NUMBER 2 / 2015


norden

Managing Editor
Professor Torben M. Andersen, Department of Economics,
University of Aarhus, Denmark.

Special Editors for this volume
Professor Torben M. Andersen, Department of Economics, Aarhus University and
Professor Jesper Roine, Stockholm Institute of Transition Economics, Stockholm
School of Economics.

Papers published in this volume were presented at the conference "Whither the Nordic Welfare Model" hosted by the Ministry of Finance, Finland, October 2015. This project was financed as part of the Danish presidency flagship program "The Nordic 'supermodel'?"

Nordic Economic Policy Review

Whither the Nordic Welfare Model?

Torben M. Andersen, Jesper Roine, Bernt Bratsberg, Knut Roed, Michael Svarer, Michael Rosholm, Tuomas Takalo, Otto Toivanen, Guttorm Schjelderup, Julian V. Johnsen, Katrine V. Løken, Helena Holmlund, Nabanita Datta Gupta, Bent Jesper Christensen, Andreas Bergh, Johanna Mollerstrom and Kalle Moene

TemaNord 2016:503

Nordic Economic Policy Review
Whither the Nordic Welfare Model?
Torben M. Andersen, Jesper Roine, Bernt Bratsberg, Knut Roed, Michael Svarer, Michael Rosholm, Tuomas Takalo, Otto Toivanen, Guttorm Schjelderup, Julian V. Johnsen, Katrine V. Løken, Helena Holmlund, Nabanita Datta Gupta, Bent Jesper Christensen, Andreas Bergh, Johanna Mollerstrom and Kalle Moene

ISBN 978-92-893-4450-0 (PRINT)
ISBN 978-92-893-4452-4 (PDF)
ISBN 978-92-893-4451-7 (EPUB)
http://dx.doi.org/10.6027/TN2016-503
TemaNord 2016:503
ISSN 0908-6692
© Nordic Council of Ministers 2016
Layout: Hanne Lebech
Cover photo: PUB, UNIT/NCM

Print: Rosendahls-Schultz Grafisk
Printed in Denmark



This publication has been published with financial support by the Nordic Council of Ministers. However, the contents of this publication do not necessarily reflect the views, policies or recommendations of the Nordic Council of Ministers.

www.norden.org/nordpub

Nordic co-operation
Nordic co-operation is one of the world's most extensive forms of regional collaboration, involving Denmark, Finland, Iceland, Norway, Sweden, and the Faroe Islands, Greenland, and Åland.

Nordic co-operation has firm traditions in politics, the economy, and culture. It plays an important role in European and international collaboration, and aims at creating a strong Nordic community in a strong Europe.

Nordic co-operation seeks to safeguard Nordic and regional interests and principles in the global community. Common Nordic values help the region solidify its position as one of the world's most innovative and competitive.

Nordic Council of Ministers
Ved Stranden 18
DK-1061 Copenhagen K
Phone (+45) 3396 0200

www.norden.org

Contents

1. Introduction 7
Torben M. Andersen and Jesper Roine

2. The Nordic welfare model in an open European labor market 19
Bernt Bratsberg and Knut Røed

3. Future Pathways for Labour Market Policy: Including the Excluded 43
Michael Rosholm and Michael Svarer

4. Economics of Innovation Policy 65
Tuomas Takalo and Otto Toivanen

5. Taxing mobile capital and profits: The Nordic Welfare States 91
Guttorm Schjelderup

6. Nordic family policy and maternal employment 115
Julian V. Johnsen and Katrine V. Løken

7. Education and equality of opportunity: What have we learned from educational reforms? 133
Helena Holmlund

8. Retirement and Health in the Nordic Welfare State 171
Nabanita Datta Gupta and Bent Jesper Christensen

9. The future of Welfare services: How worried should we be about Wagner, Baumol and Ageing? 197
Andreas Bergh

10. Ethnic fractionalization and the demand for redistribution – Potential implications for the Nordic model 219
J*ohanna Mollerstrom*

11. The Social Upper Class under Social Democracy 245
Kalle Moene

1. Introduction

1.1 Whither the Nordic Welfare Model?

Torben M. Andersen, Department of Economics, Aarhus University and Jesper Roine, Stockholm Institute of Transition Economics, Stockholm School of Economics

The Nordic Welfare Model frequently attracts international attention and is by many seen as a social model to be inspired by or even to be copied. In recent years, the "Nordic Way" has been a topic for discussion at the World Economic Forum and it has even appeared on the cover of *The Economist* under the heading "The next supermodel".[1] Somewhat paradoxically, however, the debate in the Nordic countries often features skepticism on the future of the model. Does this reflect a timely concern voiced by those best placed to see what is going on, or are the doubts on the contrary a result of model-hypochondria?

A first caveat – or perhaps part of the answer – has to do with the meaning of the very concept "Nordic welfare model". Is it really meaningful to talk about the existence of such a thing? Is the term well-defined given how large the differences are between the Nordic countries and given the major policy shifts in the past decades? The answer clearly depends on what one includes in the meaning of the model concept. If one thinks that it is associated with a certain set of specific policies or certain levels of tax rates or benefits, then clearly the concept is questionable. These things have indeed changed over time and are also different across the Nordic countries. For example, unemployment in-

[1] The Economist, 2 February 2013, http://www.economist.com/news/leaders/21571136-politicians-both-right-and-left-could-learn-nordic-countries-next-supermodel

surance is voluntary in Denmark, Finland and Sweden, but mandatory in Norway. Pension systems are fundamentally different spanning from a large role to funded occupational pensions in Denmark to a notionally defined contribution scheme in Sweden. While tax burdens are high in the Nordic countries (except for Iceland), the tax structure differs with Denmark having the larger share of tax revenue accruing from direct income taxes and value added tax, while Sweden raises much more tax revenue from social contributions.

However, in a longer perspective such a focus on certain policies would seem odd. If one were to look at reforms and levels of tax rates and benefits over the whole history over which the concept of a "Nordic model" has been identified and discussed, these have varied a lot. Indeed, continuous change has been a distinguishing feature of the model, and the changes over the past decades are not in any obvious way larger in magnitude than those in the preceding decades.

If one instead identifies the Nordic model as being concerned with a number of broader principles and goals in terms of outcomes, the concept becomes more well-defined. What matters then are the overall objectives and the overall design of the package. Here the complementarity between policies and institutions is crucial. It is not the ingredients, but the overall packaging, which makes a difference in terms of final outcomes. With this kind of perspective it also becomes clear that the naïve "copy and paste" perspective often taken in comparative policy discussions focusing on a single or few policy instruments is misleading since it overlooks the complementarities between the different policy elements. From this point of view the Nordic model should not be defined or assessed in terms of specific policy instruments, what matters is the overarching objectives. Goals – such as equal opportunities in life regardless of family background, the eradication of poverty, gender equality, the lowering of income inequality, etc. – as well as some principles – such as individually based universal rights to things such as health care and education, well-organized labour markets, etc. – have

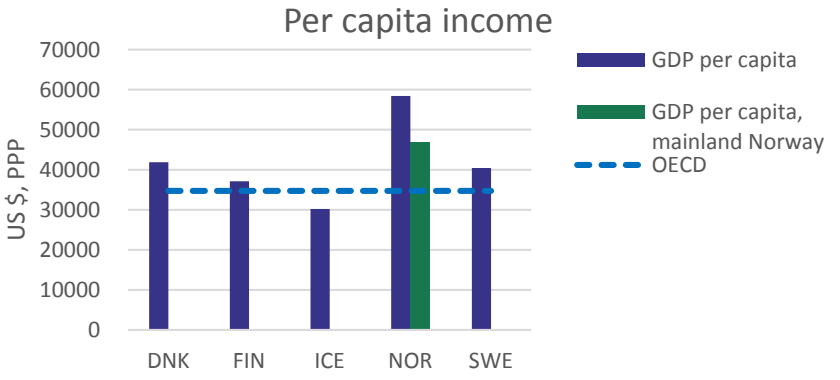
remained largely stable, while the specific policies and instruments to reach them differ across time and countries.[2]

In terms of economic performance the Nordic countries, like most others, have seen good and bad periods, but the Welfare Model – defined in terms of its goals rather than a fixed set of policies – has proved resilient. The Nordic countries stand out today as they did decades ago as being countries with comparatively high living standard and a relatively equal distribution of income. In the jargon of economics, the Nordics seem to have found a way to balance concerns for efficiency and equity. The public sector is large, hence the tax burden is high, and yet the Nordics rank in the top for various indicators of economic performance and competitiveness.[3] Figure 1a–c depicts a few select indicators often used to compare countries along the efficiency and equity dimension. The Nordic countries are high income countries, and have high employment rates, especially for women. Income inequality and poverty is low in international comparison.

[2] Of course these things are (and have been) debated (see e.g. Andersen, Roine and Sundén (2014), Chapter 2, for an overview of different views of the Nordic welfare state). The main point here is to emphasize that the model should be understood in terms of broad goals rather than in terms of specific policies.

[3] In the most recent version of the Global Competitiveness Index 2014–2015 ranking Finland placed 4, Sweden 10, Norway 11, Denmark 13 and Iceland 30, out of 144 countries.

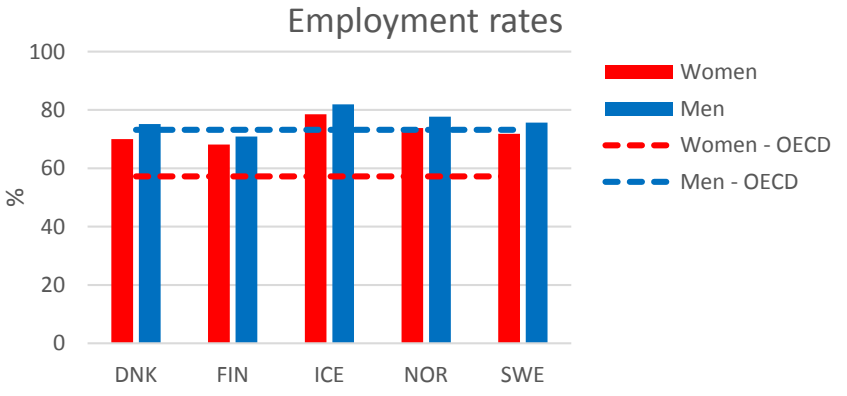
Figure 1a: Performance indicators for Nordic countries – Per capita income



Note: Income per capita is measured in USD PPP corrected. Norway-mainland is GDP corrected for the importance of off-shore oil and gas extraction, 2010.

Source: Data from www.oecd-ilbrary.org

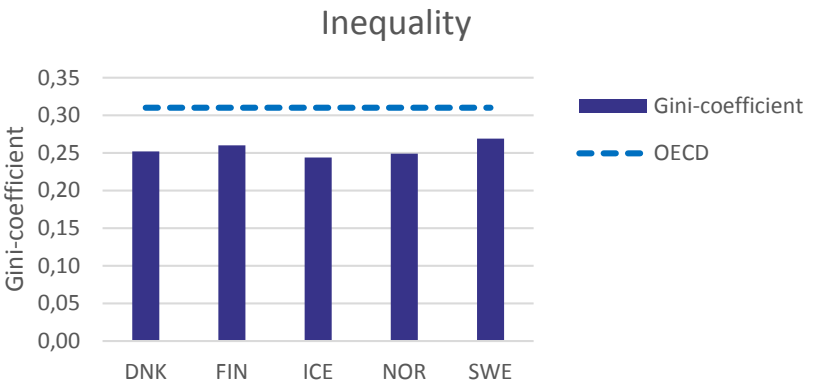
Figure 1b: Performance indicators for Nordic countries – Employment rates



Note: Employment rates for the age group 15–64, 2011.

Source: Data from www.oecd-ilbrary.org

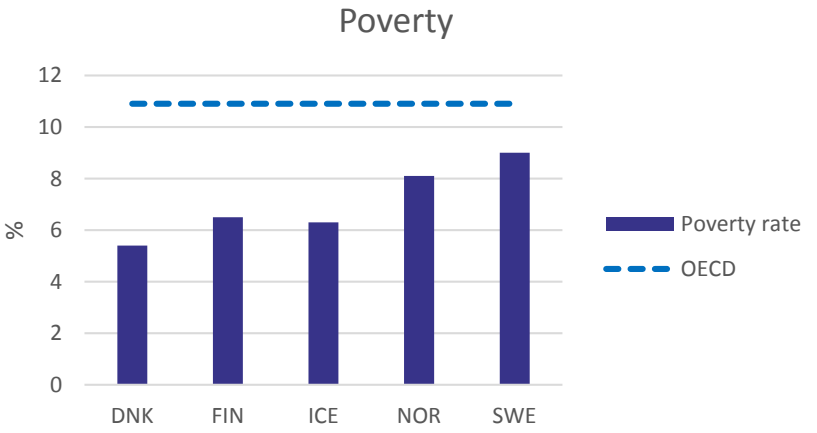
Figure 1c: Performance indicators for Nordic countries – Inequality



Note: Gini-coefficient defined over equivalised disposable income 2010.

Source: data from www.oecd-ilbrary.org

Figure 1d: Performance indicators for Nordic countries – Poverty

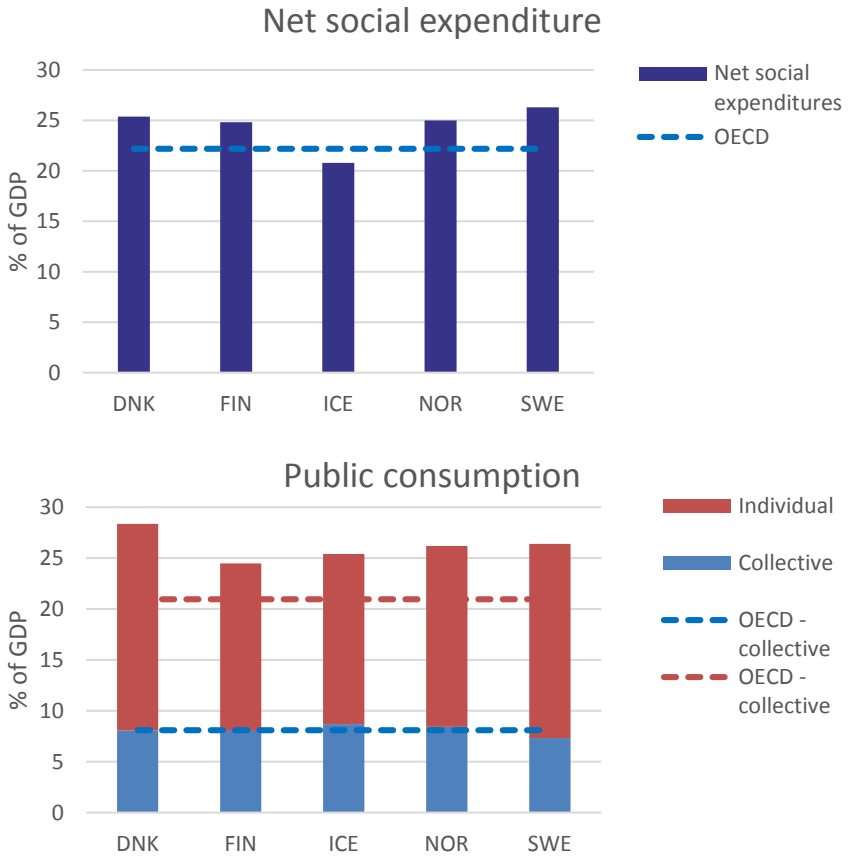


Note: Poverty measured as the share of individuals with equivalised income below 50% of median income.

Source: Data from www.oecd-ilbrary.org

Taxes are, no doubt, high in the Nordic countries, and taxes distort economic incentives. However, the effects of taxes cannot be seen independently of what taxes are financing The two broad expenditure types relate to the social safety net and provision of welfare services, cf. Figure 2. The social safety net plays an important distributional role but it is also provides insurance. The latter may have a direct welfare effect but also be conducive to flexibility and ensures that the costs of changes at the level of society are not fully carried by specific individuals. Welfare services include education, health and care. They are provided universally and at contemporary standards and meeting the requirements of most people. Welfare services are important from a distributional point of view, and in terms of ensuring equal opportunity. Clearly, these activities are also important for labour supply along both the quantitative and qualitative dimension. As examples, day care – which is also associated with other values in relation to family policy and social integration – promote labour supply, especially for women. Education is obviously associated with productivity but is also associated with e.g. later retirement. The complicated interrelation between the effects of taxes and welfare spending underlines the need to continuously recalibrate policies to find the right balance between concerns for efficiency and equity alongside various changes in society.

Figure 2: Public sector activities: social expenditures and public consumption



Note: Net social expenditures correct gross social expenditures for taxes on transfers to make data comparable between countries, where in some transfers are taxable income and in others they are not taxable income, see Adema *et al.* (2011). Public consumption is split between traditional collective expenditures, and expenditures on activities which can be attributed to specific individuals (welfare services). Data applies to 2011.

Source: www.oecd-ilibrary.org

Discussions on the future of the welfare model often has "competitiveness" in the broad sense of the term as a common denominator. Can the Nordic countries remain competitive and thereby sustain a high living standards? Often this debate has the implicit premise that to remain

competitive we cannot deviate too much from other countries, and therefore the Nordic model is particularly vulnerable.

The concern for competitiveness is not new to the small and open economics of the Nordics. It has always been an overriding concern. But "being competitive" is not tantamount to "being alike" and implying that all social models have to converge. This view has no support in, for instance, trade theory stressing the importance of differentiation and comparative advantages. There is also an increasing understanding that different social structures and institutions can be a source of comparative advantages. A recent literature levy a critique on traditional analyses for having a too one-sided focus on identifying the optimal institutional setting, see e.g. Nunn and Trefler (2013). There is no specific institutional setting which is optimal. The reason is that various institutional arrangements have pros and cons, which may be a source of comparative advantage. Countries with flexible employment protection legislation and generous unemployment insurance may have a comparative advantage in industries with substantial short-term variation in demand and thus production, while countries with more strict employment protection legislation and less generous unemployment insurance may have a comparative advantage in production of commodities with less variability. As an example of this Cuñat and Melitz (2012) find in a cross-country study empirical support that countries with more flexible labour markets have a higher degree of specialization in sectors more frequently exposed to sector-specific shocks. This may be interpreted in the sense that the nature of shocks or needs for adjustment to some extent is endogenous, meaning that countries (or rather its companies in the private sector) specialize in the activities for which their particular institutional setting has a comparative advantage. This type of research is still in its infancy, but it is highly suggestive of why different institutional settings (welfare regimes) survive. The important lesson – repeating basic insights from trade theory – is that competitiveness is a question about comparative advantages.

Past performance is important, but the pertinent question is whether the Nordic Welfare Model is robust and resilient to various challenges including changing demographics, globalization, new technologies and environmental changes?

Society is undergoing various changes, some small, some larger, some come gradually and others in clusters. In the debate they are often labelled challenges, signalling that policy initiatives are required. Whether they are a challenge or a threat to the Nordic model as such is a different issue. To take an example, the issue of ageing is undoubtedly a very important policy challenge. The age composition of populations are changing, not least because longevity goes up (and is in itself associated with huge welfare gains). Clearly, the social contract has to be adopted to such a change. Although this may be politically difficult, it is not difficult in a technical sense, and a solution is clearly feasible without changing any fundamental properties of the model. One solution may be to increase retirement ages alongside increases in longevity and thereby ensure that the fraction of life spent in the labor market remains unchanged. It is not possible in any meaningful way to interpret this as a change in the basic principle of the model.

Another much discussed area is that of automation and digitalization and the impact this will have on all aspects of the economy, in particular on the future of work. This is a vast debate with many dimensions but it is interesting to note that some aspects of the challenge and some suggested solutions turn out to be much aligned with basic principles of the Nordic model. Predictions such as the need for continuous education throughout life and the need for individuals to be able to hold several jobs over a life-time, in fact, place demands on policy similar to those of a small open economy in an increasingly globalized world. Ideas such as "protect individuals, not firms" and "make sure the workforce has continuous possibilities to educate and re-educate themselves to meet new challenges" are not less familiar to the Nordic model than to other countries, rather the opposite. For sure, policy will have to change to adapt to new realities, but again, it is not obvious that these challenges make the model obsolete.[4]

Some challenges may be related to properties of the model. Immigration of unskilled or low-skilled individuals may be a particular prob-

[4] Just as an example, in a recent issue of Foreign Affairs Colin and Palier (2015) outline some challanges in the "digital age" and arrive at the conclusion that aspects of the Nordic model are, in fact, better suited to "fostering a more fluid and entrepreneurial economy" than many of the alternatives discussed.

lem in societies with high minimum wages (no working poor), high qualification requirements for jobs, high employment rates for both genders etc.

In other cases solutions may also be model specific, as e.g. in relation to provision of welfare services like education, health and care, which are essential elements of the welfare state. Such welfare services may be subject to both Baumol's cost disease (relative costs increases over time since productivity increases are typically lower than for manufactured products) or Wagner effect (increasing demand for service alongside improvements in material living standard). While the drivers are universal, the solution is model specific, since the Nordic countries have opted for a larger public role in the provision of services

Society is undergoing large changes – as it also has in the past – not least those arising in the intersection between globalization and technological change which changes modes of production which on impact creates both winners and losers in the labour market. The derived effects also include new forms of employment, less stable employment relations etc. Left on its own this may be a source of increasing inequality. This raises questions for traditional distribution policies running via taxes and the social safety net, but also for the possibilities to actively counteract these changes via education, labour market and social policies. Maintaining a high employment level is both a value in itself related to social inclusion and equality, but the financial viability of the welfare model also depends on maintaining a high employment level.

These changes also have wider effects on the interface between markets, civil society and the family. Changed employment relations and demands in the labour market may affect the possibilities the individual has in balancing work-life and family-life. Families may also undergo changes (divorces) and there is an increasing trend in single-families.

Perhaps the most important aspect of the Nordic model should not be found in economic details but rather in the political economy sphere. It is sometimes denoted the "consensus" tradition which permeates industrial relations and politics. The political capital is large and this is reflected in an ability to undertake reforms. Rather wide ranking reforms of pension and retirement schemes have been implemented

smoothly in e.g. Denmark and Sweden, and they are among the few countries who have addressed the ageing problem.

In discussing changes or challenges, it is also worth pointing out that many of these are common to most countries or global in their nature. The need for changes and adjustments should therefore be seen in the perspective of the changes needed in other countries. It is not clear that the challenges are posing a larger problem for the Nordic countries. To list just a few, the US is facing a problem of steeply increasing inequality and segregation. Southern European countries experience an outflow of well-educated young and strong protests and retrenchment of reforms to address the ageing problem.

Globalization and technological changes are associated with collective gains but an unequal distribution of gains and losses. Welfare arrangement may contribute to compensate the losers and (re) distribute the gains, which in turn may be conducive to reforms. Clearly there is a hen-and-egg issue in the interdependence between welfare arrangements and ability to reform – the welfare state may create an environment in which it is easier to undertake reforms, but the reforms are also crucial for the development and viability of the model. How this relationship has been established and developed historically is in itself and interesting question, but beyond this volume to consider.

Rather than looking backward at past performance it is important in due time to consider changes in society and discuss how to address them. This volume addresses some topical issues on the future of the Nordic welfare model. It is, of course, by no means exhaustive, but instead covers a number of broad issues outlining what recent research has to say on them. Each paper is relatively short, given the width of each topic, and the titles of each contribution explains what the covered topic is, so rather than trying to summaries the contributions we hope that they all, together or individually will contribute to a better and more informed discussion about the future challenges, reform needs but also possibilities of the Nordic model.

References

Adema, W., Fron, P., & Ladaique, M. (2011). Is the European Welfare State Really More Expensive?: Indicators on Social Spending, 1980–2012; and a Manual to the OECD Social Expenditure Database (SOCX), *OECD Social, Employment and Migration Working Papers*, No. 124, OECD Publishing. http://dx.doi.org/10.1787/5kg2d2d4pbf0-en

Andersen, T., Roine, J., & Sundén, A. (2014). Hur får vi råd med välfärdsstaten? *SNS Konjunkturrådsrapport 2014,* SNS förlag, Stockholm.

Colin, N. & Palier, B. (2015). The next safety net: Social Policy for a digital age. *Foreign Affairs* 94(4): 29–33.

Cuñat, A. & Melitz, M. (2012). Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage. *Journal of the European Economic Association,* 10(2): 225–254. http://dx.doi.org/10.1111/j.1542-4774.2011.01038.x

Nunn, N. & Trefler. D. (2014). Domestic Institutions as a Source of Comparative Advantage, In G. Gopinath, E. Helpman & K. Rogoff (Ed.) *Handbook of International Economics*,. Volume 4, chapter 5: 263–315. http://dx.doi.org/10.1016/b978-0-444-54314-1.00005-7

2. The Nordic welfare model in an open European labor market[1]

Bernt Bratsberg, The Ragnar Frisch Centre of Economic Research, bernt.bratsberg@frisch.uio.no and Knut Røed, The Ragnar Frisch Centre of Economic Research, knut.roed@frisch.uio.no

Abstract

Is it possible to sustain an ambitious and redistributive Nordic welfare state in a Europe with open borders? Drawing on longitudinal administrative records spanning four decades, we first present discouraging historical evidence showing that labor migrants from low-income source countries tend to have unstable employment careers with marked overrepresentation in welfare programs. This pattern extends to post-accession labor migrants from Eastern Europe, who quickly experience high rates of unemployment. The article discusses possible avenues for making the welfare state "migration robust". We argue that there are alternatives to reclosing borders and/or cutting down welfare state ambitions, and recommend policies based on strengthening of activity requirements in social insurance programs, raising minimum job standards, and substitution of work-oriented services for cash-based family allowances.

---

[1] We are grateful to Gregory Clayes, participants at the 2015 NEPR conference in Helsinki, and the editors for valuable comments. The article draws on research funded by the Norwegian Research Council (projects "Work Life Challenges – workforce management and worker involvement solutions" and "European Strains") and is part of the research activities of the Centre of Equality, Social Organization, and Performance, University of Oslo. Data made available by Statistics Norway have been essential for this research.

2.1 Introduction

The recent enlargements of the EU Single Market represent new opportunities for growth and prosperity. The liberalized migration regime that followed allows workers to flow more easily toward their best potential use (Kahanec *et al*., 2014; see also Clemens, 2011 and Kennan, 2012). Improved labor mobility removes bottlenecks in production processes and facilitates dynamic cushioning of regional cyclical fluctuations. Left unhindered, the open-border policy is a powerful tool for raising and equalizing living standards across Europe, and thus for promoting economic and social cohesion. However, at least in the short and intermediate terms, the integrated labor market also presents some political and economic challenges. Large cross-country differences in labor productivity, wages, and social insurance standards may trigger migration flows that place pressure on present welfare state institutions (Sinn, 2002; Kvist, 2004). In particular, the fact that social insurance benefits in the Nordic countries by far exceed typical wages in most accession countries may distort migration flows and weaken labor migrants' incentives to remain in productive employment over the long haul. As we show below, prior empirical evidence points to a considerable risk that labor immigrants from low-income countries fail to establish lifecycle employment and earnings careers on par with natives, but instead exhibit high rates of premature labor market exit and welfare uptake.[2] Beyond their mere fiscal implications, such processes may well lead to economic marginalization of minority groups and, as discussed by Mollerstrom (this volume), reinforce any decline in popular support for redistribution linked to growing demographic heterogeneity as many natives will perceive it as less likely that they will be on the receiving end. This brings to the fore questions of how labor market institutions and social insurance systems should adapt in order to reap the full benefits – and avoid the perils – of an integrated European labor market.

[2] A large literature studies welfare uptake among immigrants and natives across Europe; see Nannestad (2004) for a review and Boeri (2010) and Barrett and Maitre (2013) for recent cross-country comparisons.

Based on Norwegian administrative register data, this paper first reexamines past experiences with labor immigration. Labor immigration benefits the welfare state in the short run through its immediate expansion of labor supply. But, the impacts on the fiscal sustainability of the welfare state also depend on the migrants' long-term integration in the labor market and their rate of return migration. Our brief review of prior studies, paired with new evidence on labor market outcomes of recent European labor migrants, indicates grounds for concern: Labor immigrants from countries with low economic development have more unstable employment patterns, and face a much higher probability of becoming reliant on social insurance transfers, than natives. We move on to discuss mechanisms that can explain these patterns, such as vulnerability to cyclical fluctuations; lack of language skills needed to adapt to new jobs/occupations in response to structural change; high effective replacement ratios in the social insurance system; and employer incentives to recruit low-skilled immigrant workers to jobs with low wages and poor working conditions. Finally, we discuss some policy options. We argue that a reintroduction of migration barriers is not the way to move forward. Instead we recommend policies aimed at making the Nordic welfare model more "migration robust": First, by establishing (or raising) minimum standards/wages in the labor market in order to prevent social dumping at the tax payers' expense, and, second, to make the social insurance system more participation oriented – essentially by substituting job offers and/or various forms of activation for pure cash transfers.

2.2 Experiences prior to the enlargements of the European labor market

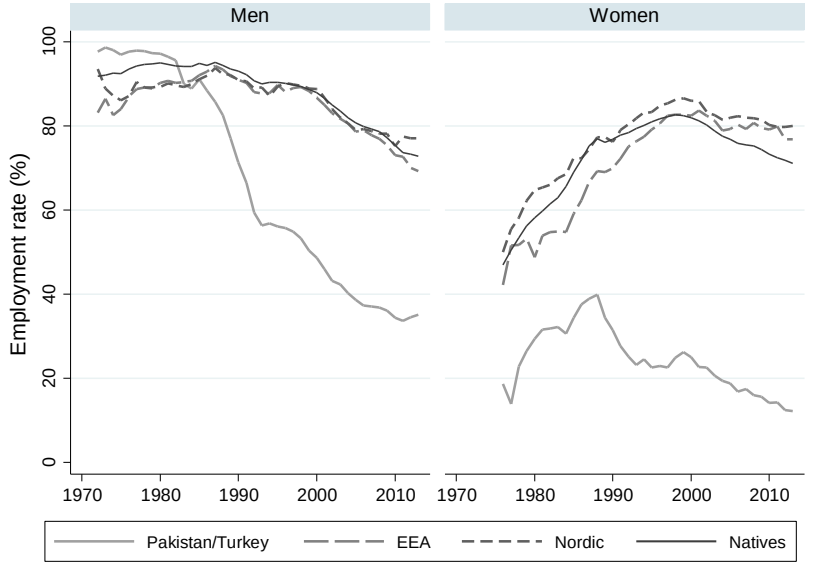
Between 1975 and 2004, work-based immigration to Norway from outside the Nordic region (the EEA area since 1994) was subject to strict regulations. Hence, in order to examine the long-term labor market performance of labor immigrants from low-income countries, we have to go back to the waves that arrived just prior to the 1975 legislation. Although this obviously raises questions about comparability with today's migrants and today's labor market institutions, the exercise has the ad-

vantage that it facilitates assessments of immigrant performance over four decades.

Bratsberg *et al*. (2010; 2014) have examined the lifecycle employment and earnings patterns of these early cohorts of labor migrants to Norway. Their analyses distinguish between immigrants from countries with similar earnings levels and living standards to those of Norway (i.e., Western Europe) and immigrants from countries with considerably lower earnings and living standards (e.g., Pakistan and Turkey). A key finding of these studies is that whereas labor immigrants from Western Europe had employment and earnings patterns similar to those of natives, the labor immigrants from low-income countries had a disproportional tendency to drop out of labor market after 10–15 years of employment.

The left-hand panel of Figure 1 summarizes and updates some of the main findings of the prior studies by showing the annual employment rates of male immigrants who entered Norway during the early 1970s, and then remained in the country until 2013. Similarly, the right-hand panel shows employment rates for female immigrants who entered during the late 1970s (with the delayed entry period explained by the fact that very few women from low-income countries arrived during the early 1970s; the majority of those from the late 1970s being spouses of the male cohorts of the early 1970s). Here, we distinguish between three different regions of origin that dominated labor-related migration to Norway during this period: i) Pakistan and Turkey, ii) the Nordic countries, and iii) the rest of Western Europe (denoted EEA in the figure). In addition, we show employment rates for a comparison group of natives, stratified to have the exact same age composition as the three immigrant groups put together. Since the age compositions of these groups were roughly the same (with the typical immigrant being 24–25 years of age at the time of arrival), we can compare the dynamic employment patterns directly. It is clear that lifecycle employment was *much* lower for labor immigrants from Pakistan and Turkey than for immigrants from Western Europe and natives. Employment levels tended to be high during the first years in Norway, but after around 10–15 years of residency, they started to drop precipitously. Immigrants from the Nordic countries and the rest of Europe, on the other hand, had employment patterns very similar to natives.

Figure 1: Annual employment rates 1972–2013, conditional on continued residency in Norway – Immigrants from the 1970s and native comparison groups

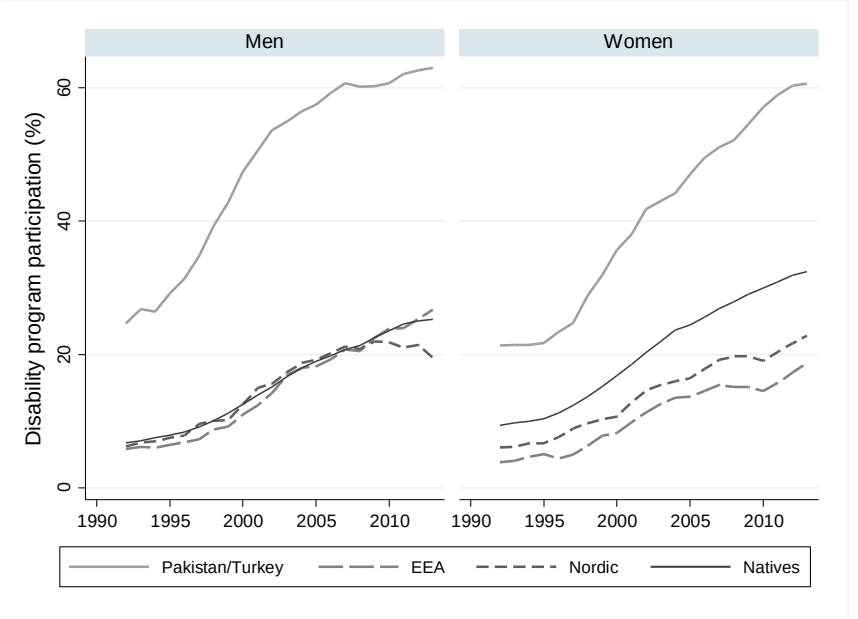


Note: Employment is defined as having annual labor or business earnings exceeding the base amount of the social insurance system (currently NOK 90,068). The figure shows annual averages for those aged 25–64. The "EEA" group consists of immigrants from the following countries (with sample share in parenthesis); The United Kingdom (47), Germany (14), France (12), the Netherlands (10), Spain (4), Switzerland (4), Italy (3), Austria (2), Belgium (2), Ireland (1), Portugal (1), Greece (1), and Luxembourg (0).

Figure 2 illustrates how the low employment rates among immigrants from Pakistan and Turkey translate into correspondingly high participation rates in disability insurance programs. These data are available from 1992 only, but we see that already at this point more than 20% of the Pakistani and Turkish immigrants who arrived 15–20 years earlier had become recipients of disability insurance benefits. By the end of our observation period, more than 60% of the labor migrants still in Norway, as well as their spouses, had become disability insurance claimants, compared to around 20% of the immigrants from high-income countries.

The reason why we focus on disability insurance benefits here, and not, say, unemployment benefits or social assistance, is that the disability program had already become the major social insurance program for the 1970s cohort of labor immigrants when the data series for social insurance transfers start in 1992. Likewise, the unemployment insurance register data are first available in 1989 and we are unable to study the dynamic transitions between employment, unemployment insurance, and disability program enrollment during the critical downturn of the late 1980s. Unemployment insurance was presumably an important income source for many of the immigrants during the slump, but since these benefits are subject to time limits (currently a maximum of two years) the welfare state had to find other ways to ensure lasting solutions for persons who persistently failed to find new employment. Existing empirical evidence (Rege *et al*., 2009; Bratsberg *et al*., 2013) shows that there is a large "grey area" between unemployment and disability insurance programs: Job loss raises the probability of becoming a disability insurance claimant considerably, and Bratsberg *et al*. (2013) argue that disability insurance is sometimes unemployment insurance in disguise. For the cohorts under study, the underlying data indeed show a strong correlation between unemployment and subsequent disability program enrollment: For the men who first entered disability in 1993, unemployment insurance rates in 1989 were twice those of men who did not enter disability insurance.

Figure 2: Disability program participation 1992–2013, conditional on continued residency in Norway. Immigrants from the 1970s and native comparison groups



Note: Figure entries are restricted to those aged 25–64.

Given that we examine a wave of immigrants who arrived during the 1970s, it is now possible to assess their patterns of employment and earnings over their whole potential working lives. On average, the male labor immigrants from Pakistan and Turkey were employed 61.9% of all years between the ages of 25 and 66. To put this number into perspective, we have computed the corresponding number for native men of the same birth cohorts, who were employed 85.5% of their potential working lives. Examining annual earnings instead, we find that the labor immigrants from Pakistan and Turkey on average earned 177,791 NOK per year (not conditional on employment, and inflated to 2012 currency), whereas the comparable group of native men earned 328,464 NOK. Hence, the labor immigrants' lifetime earnings were on average only 54% of those of native men of the same birth cohorts.

Moving on to the spouses of these labor immigrants, we find that the women were employed in only 21.9% of all years between the ages of 25 and 66, with average annual earnings as low as 43,737 NOK. By comparison, native females of the same birth cohorts were employed 68.0% of their potential working-life years, with average earnings equal to 176,772 NOK. Hence, the lifetime earnings of the cohort of immigrant women were only 25% of those of native-born women.

Lower lifetime earnings than natives do not necessarily indicate that immigrants represent a fiscal burden for the welfare state, however. In order to provide a broader assessment of the fiscal consequences of migration, one has to include contributions and expenditures over the complete lifecycle, taking into account that tax payers do not have to pay the costs of child care and education before immigrants arrive and that some immigrants will spend the last – and the most cost intensive in terms of health care – years of their life in their country of origin. Hence, the break-even point of direct taxes paid versus benefits received likely involves lower lifetime labor earnings for immigrants than for natives (Preston, 2014).

Nevertheless, it is of paramount importance to understand why the immigrants from Pakistan and Turkey performed so poorly in the Norwegian labor market over the long term. Unfortunately, we are not able to provide complete evidence-based answers to this question. We know that business cycle fluctuations played an important role, as many of the immigrants lost their foothold in the labor market around the major cyclical downturns in the early and late 1980s; see Bratsberg *et al*. (2010). A large fraction was originally recruited to declining (and, to some extent, dying) industries and they did not possess the human capital and language skills typically required by the new and growing industries. Thus, dependency on temporary social insurance became prevalent. Since social insurance benefits are more generous for persons with children and dependent spouses, many of the immigrants experienced that social insurance gave as high, and in some cases even higher, family income than fulltime work (Bratsberg *et al*., 2010). This situation might have undermined incentives to provide the effort required for regaining a foothold in the labor market, thereby transforming temporary insurances into permanent ones.

The presence of a relatively comprehensive social security net in Norway, combined with large differences in living standards between Norway and the source country, further weakened the incentives for return migration among the labor migrants from low-income countries, even in cases were new employment could not be found in Norway. This illustrates an important asymmetry in labor-motivated migration patterns between countries with very different levels of development: Whereas high labor demand during economic booms in the wealthy country will attract workers from poor countries, there is no reason to believe that a subsequent economic bust will set the migration flows in reverse. At this point, there is a significant difference between labor migration flows across countries with similar and countries with very different living standards. The discrepancy also shows up in our data: While as many as three in four of the 1970s immigrants from the Nordic and other Western European countries had left Norway by 2013, this was the case for only one in four of the Pakistani and Turkish immigrants – despite the latter group's *much* poorer performance in the Norwegian labor market.

For those who did stay in Norway, the long-term labor market performance of immigrants from the Nordic and other Western European countries is actually a completely different story: As Figures 1 and 2 showed, their lifecycle employment patterns are hardly distinguishable from those of natives, and for females, participation in disability insurance programs is even considerably lower than for similarly aged native women.

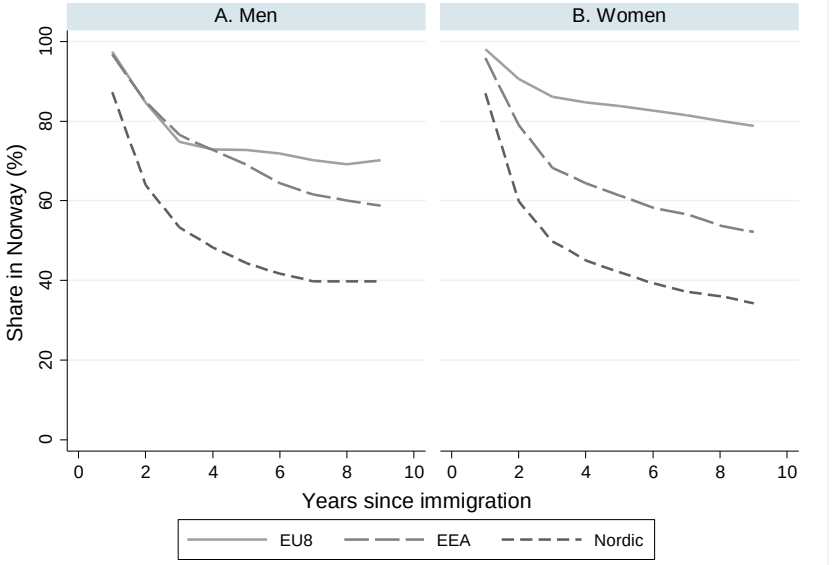
2.3 Experiences since the expansions of the European labor market

So, for the issue of how the new European flows of labor migrants will affect the long-term fiscal sustainability of the Nordic welfare states, a key question is whether the eastwards extensions of the European labor market will lead to migration flows and long-term employment patterns that resemble the 1970s experiences with migration from poor or from rich source countries. Geographical and cultural closeness suggest that past immigration from Europe is the most relevant

reference. However, since the new European labor market covers countries with large differences in economic development, and with very different labor market and social insurance institutions, the answer to the question is not obvious.

Given that the first eastwards extension of the European labor market took place only 11 years ago (in 2004), it is obviously too early to paint a complete picture of the lifecycle employment and earnings patterns of the new immigrants. What we can do, however, is to examine economic outcomes over a 10-year period after entry. To do this, we look at three groups of recent European immigrants to Norway: i) those from the nearby Nordic countries, ii) those from other countries in Western Europe (for simplicity denoted EEA), and iii) those from the 2004 accession countries in Eastern Europe (denoted EU8). Specifically, we examine immigrants from these regions who arrived in Norway between 2005 and 2008 and were 17 to 46 years of age at the time of entry. Figure 3 first shows how long these migrants stayed in Norway. While fewer than 40% of the Nordic immigrants remained in Norway 10 years after arrival, this was the case for 70–80% of the migrants from EU8, and around 50–60% of the migrants from other countries in Western Europe, depending on gender. In other words, the return migration patterns of the new immigrants from the accession countries in Eastern Europe are closer to those we saw among 1970s immigrants from low-income countries than from countries in Western Europe.

Figure 3: Share of immigrants residing in Norway, by years since entry and region of origin

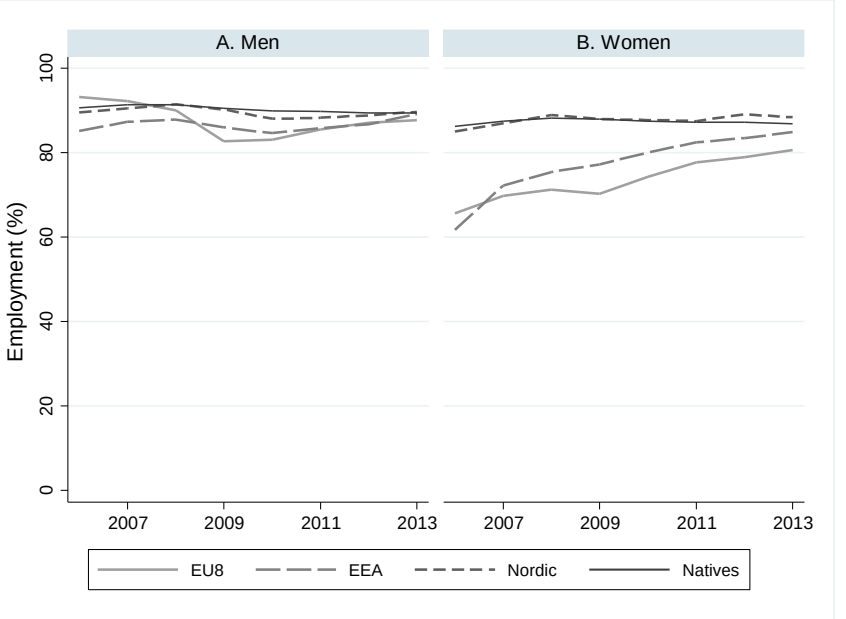


Note: Immigrant populations consist of 2005–2008 arrivals age 17–46 at entry. The EU8 group consists of immigrants from the following countries (with sample share in parenthesis); Poland (75), Lithuania (14), Slovakia (3), Latvia (3), Estonia (2), Hungary (1), Czech Republic (1), and Slovenia (0); and the "EEA" group of Germany (50), the United Kingdom (14), the Netherlands (10), France (9), Spain (4), Italy (4), Portugal (2), Austria (2), Switzerland (1), Belgium (1), Greece (1), Ireland (1), Liechtenstein (0), Luxembourg (0), Andorra (0), and San Marino (0).

Figure 4 displays, separately by gender, annual employment rates for each of the three immigrant groups. The employment figures are computed for persons aged 25 or more, conditional on continued residency in Norway (at the end of each calendar year), and also conditional on not being enrolled in education during the year. Again, we add a native comparison group with the same age composition as the various immigrant groups combined. (There is some variation in age across the three immigrant groups. To illustrate, among those in Norway at the end of 2010 the average age was 34.5 for the EU8 group, 35.7 for the EEA group, and 31.6 for those from the Nordic countries. When we reweight

the native data to reflect the age distribution of each immigrant group, we uncover however only minor differences from the overall native employment numbers displayed in Figure 4.) The figure shows that employment rates for Nordic immigrants again tend to be high and indistinguishable from those of natives. Male immigrants from Eastern Europe also had very high employment rates in 2007 and 2008, but their employment rates apparently took a serious blow during the financial crisis in 2009–2010. In 2013, however, their employment was almost back to native levels. Male immigrants from other Western European countries have consistently had employment rates somewhat below native levels. For female immigrants from both Western and Eastern Europe, we see patterns of relatively low employment rates the first years after immigration, followed by a gradual convergence toward native levels. The underlying data reveal that many of the female non-Nordic immigrants were admitted as family immigrants (44% of the EU8 women and 39% of the EEA women, compared to only 1% and 7% of the male groups). The data also reveal that the convergence of female employment rates over time foremost is attributable to strong employment growth for those who entered on a family visa. A second point to note here is that, while Figure 1 showed similar employment profiles for natives and the 1970s immigrants from both the Nordic countries and other countries in Western Europe, the recent data indicate some differences between these groups. In particular, the recent Western European immigrants from outside the Nordic countries exhibit lower return migration rates and lag somewhat behind the Nordic group in the labor market. Plausible explanations for these patterns relate to the stronger links between admission and job offers for labor immigrants from outside the Nordic countries during the 1970s, with a greater emphasis on skills considered to be needed in the Norwegian labor market, along with the much stronger economic growth in Norway compared to the United Kingdom and continental Europe between 1970 and 2010. Hence, it is probable that the 1970s' cohorts of immigrants from Western Europe were particularly favorably selected in terms of their job opportunities in Norway.

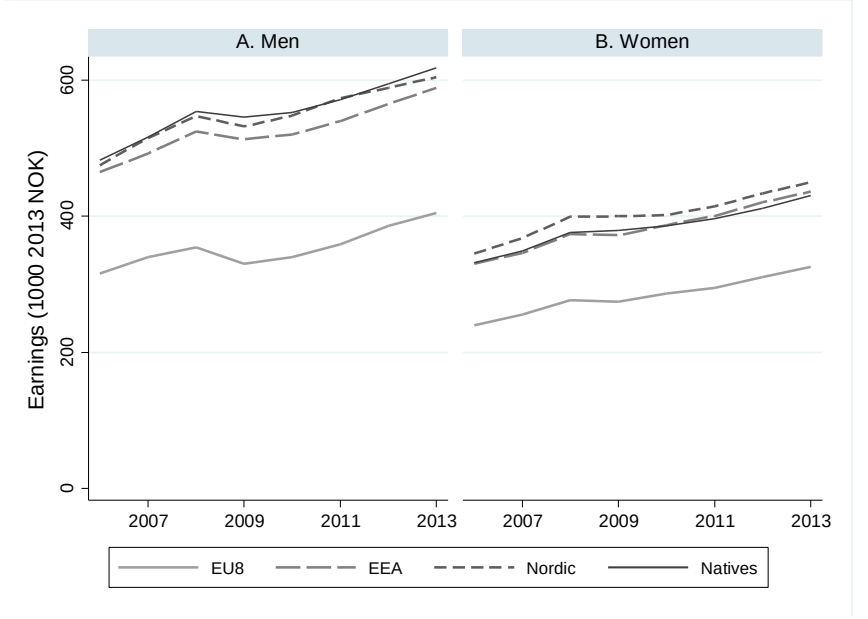
Figure 4: Employment rates 2006–2013, immigrants arriving in 2005–2008 and native comparison groups



Note: Employment is defined as having annual labor or business earnings exceeding the base amount of the social insurance system (in 2013 NOK 84,204). The data include persons at least 25 years of age, not enrolled in education, and with residency in Norway at the end of the calendar year.

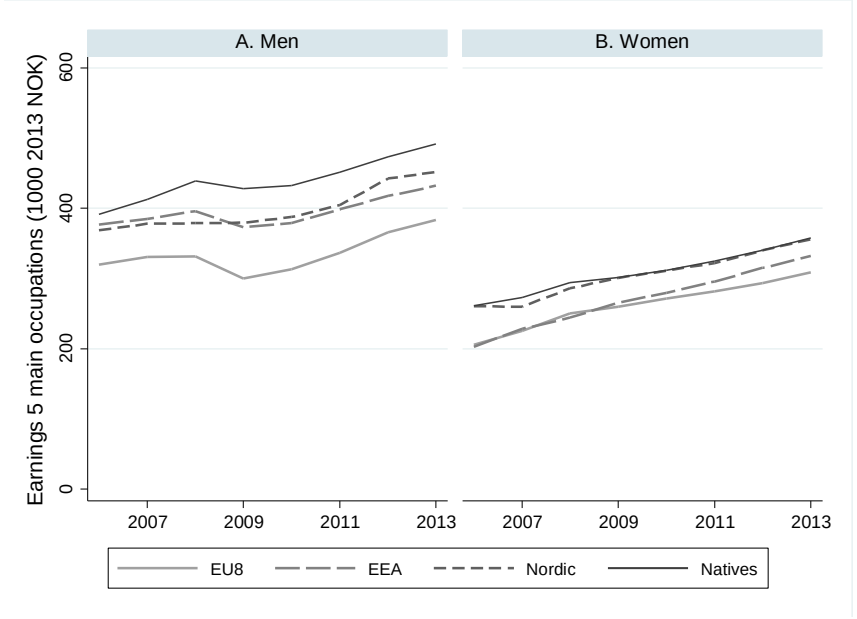
Figure 5 shows average annual earnings for those who were employed each year. It is clear that labor earnings tend to be much lower for Eastern European immigrants than for all the other groups. Moreover, the figure gives no indication of the assimilation effects typically found elsewhere in the literature, whereby immigrant earnings grow more rapidly than those of natives during the first years in the host country. Instead, the earnings gap between natives and EU8 immigrants remained constant over the eight-year period considered, with the 2013 earnings of male EU8 immigrants 34% below those of native men and the earnings of female EU8 immigrants 24% below those of native women.

Figure 5: Annual labor and business earnings 2007–2013, conditional on employment. Immigrants arriving in 2005–2008 and native comparison groups



One possible explanation for the relatively low earnings of EU8 immigrants is that they were recruited into occupations with particularly low wages. If we focus on immigrants and natives employed in the major immigrant occupations, the earnings gaps are reduced considerably; see Figure 6. The earnings differential between EU8 immigrants and natives remains significant, however, and again there is no indication of assimilation effects on immigrant earnings. When we compare natives and immigrants from accession countries who in 2008 worked in the five most common immigrant occupations, the earnings gap in 2013 stood at 22% for males and 13% for women.

Figure 6: Annual labor and business earnings 2007–2013, conditional on employment in one of the five main 2008 immigrant occupations



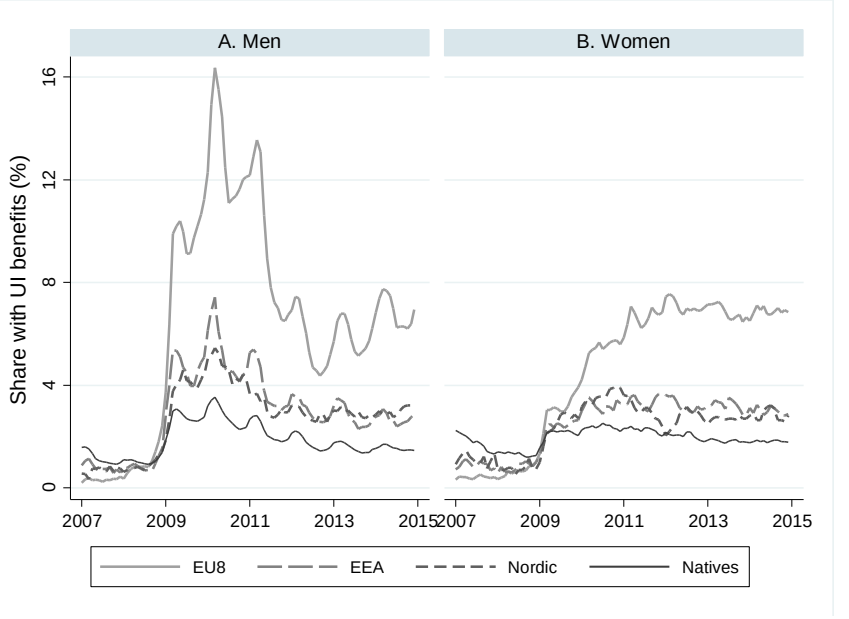
Note: The five main occupations and their employment share in the immigrant data are, for men, carpenter (11), clerical (10), construction laborer (6), cabinet maker (6), and brick layer (4), and, for women, cleaner (18), child care (6), sales (6), clerical (6), and waiter (5).

Even though most of the new Eastern European immigrants managed to maintain a foothold in the Norwegian labor market through the financial crisis, a relatively large fraction also claimed unemployment insurance (UI) benefits. Figure 7 shows the uptake of UI benefits month by month for immigrants still residing in Norway. For men, there was a huge spike in benefit claims around the financial crisis, starting late 2008, particularly for men from Eastern Europe. The EU8 claimant rates came back down around 2012, but have remained at much higher levels than those of other European immigrants and natives. And since 2012 they have again displayed an increasing pattern. For Eastern European women, we have seen a more monotonous increase in UI claimant rates after the financial crisis, albeit with slower

growth since 2012. Of particular concern is that the UI claimant rates of the EU8 group seem to have stabilized at very high levels when compared to natives. In fact, in December 2014 (the last entry in Figure 7) the UI claimant rate among men from accession countries who had immigrated to Norway in 2005–2008 was five times that of native men of the same age, and the claimant rate among women four times that of similarly aged native women.[3] Even immigrants from the Nordic and other Western European countries have had higher UI claimant rates than natives in the aftermath of the financial crisis, but at much lower levels than those of the EU8 immigrants.

[3] As with earnings (see Figures 5 and 6), immigrant-native differences in UI uptake are smaller when we consider workers in the same occupation. To illustrate, when we restrict the sample to those in the major immigrant occupations used in Figure 6, the January 2011 uptake rate among male EU8 immigrants was 2.8 times that of native men (14.8 vs. 5.2%) and the uptake rate of EU8 women twice that of native women (7.7 vs. 3.8%). Bratsberg *et al*. (2014) study overall UI uptake during the financial crisis and find that differences in age, tenure, industry and occupation account for 40% of the observed difference in uptake between immigrant men from Eastern Europe and native men and 60% of the observed difference in UI uptake among women.

Figure 7: Monthly unemployment insurance claimant rates 2007–2014, immigrants arriving in 2005–2008 and native comparison groups



Note: The data include persons at least 25 years of age who are not enrolled in education and with residency in Norway at the end of each calendar year.

2.4 Structural challenges

While it is too early to draw any firm conclusions regarding lifecycle employment in Norway, we do see some discouraging medium and long-term labor market performance patterns among the post-enlargement immigrants from lower-income countries in Europe. Why do we apparently fail to achieve full labor market integration on par with natives? We will focus here on three interrelated explanations.

The first is that immigrants often are recruited to jobs with low general skills requirements, and, in particular, to jobs where Norwegian language skills is not a key ingredient. These jobs are often temporary in nature and disproportionally found in cyclically sensitive industries

such as construction. And since most firms practice a last-in-first-out principle in case of downsizing, immigrant workers generally have a high risk of becoming unemployed. Once unemployed, the lack of general qualifications and language skills obviously become a serious handicap in attempts to find other types of work.

Second, those who become unemployed are in most cases entitled to unemployment insurance (UI). In principle, UI entitlements are fully transferable within Europe. Among labor immigrant in Norway, entitlement will in any case follow if their labor earnings during the prior calendar year exceeded 1.5 times the base amount of the social security system (in 2013, earnings above NOK 126,000 – or one third of the average earnings of male EU8 immigrants depicted in Figure 5). For workers from, say, Poland or Lithuania, Norwegian UI benefits will typically exceed earnings in the home country by an order of magnitude. In Table 1, we report average monthly UI benefits and wages in the home country, collected from the OECD Social and Welfare Statistics database (see columns I and II). These data show that typical Norwegian UI benefits are 7 to 15 times average UI benefits – and 4 to 5 times average wages – in Eastern Europe. Because the preceding section showed that Eastern European migrants earn less than natives and other migrants, we have also computed monthly UI benefits for those who actually claimed benefits in Norway in 2010 (see column III) as well as monthly pay among wage earners (column IV) and average monthly labor earnings for those employed during the year (column V). Although both benefits and wages of Eastern European migrants fall below those of the other groups considered, they remain much higher than UI benefits and wages at home. Hence, incentives for returning home to look for employment there are weak. A probable consequence is that many immigrants from accession countries remain registered as unemployed in Norway, despite being poorly qualified for new employment. For some of the unemployed, it will also be tempting to bring the insurance money back to the home country, where costs of living may be less than half of those in Norway, and thereby obtain a higher standard of living than what even a fulltime job could deliver in Norway. The rules of the UI program allow for such stays within the European labor market for a period of up to three months, but the absence of border controls be-

tween European countries obviously implies that it can be done to a much larger extent in practice.[4] Such opportunities may undermine incentives for active job search in Norway, and raise the reservation wage of the unemployed.

Table 1: Unemployment benefits and average earnings at home and in Norway, immigrants from the Baltic States, Poland, and the Nordic countries, 2010

	(I)	(II)	(III)	(IV)	(V)
	At home		In Norway		
	Monthly UI benefits	Monthly wage income if employed	Monthly UI benefits among claimants	Monthly pay, wage earners	Monthly earnings
Estonia	405	809	1,891	3,344	3,043
Latvia	411	684	1,686	3,282	3,200
Lithuania	188	561	1,579	3,224	3,091
Poland	223	754	1,689	3,353	3,134
Denmark	2,188	4,208	2,095	5,188	5,120
Finland	1,584	3,283	2,186	4,141	4,005
Iceland	1,547	2,793	1,900	4,599	4,429
Norway	2,948	4,916	1,929	4,491	4,500
Sweden	1,545	3,217	2,067	4,432	4,308

Note: Benefits and wages are converted to Euros using average exchange rate for 2010. Source of entries in columns (I) and (II) is OECD iLibrary, OECD Social and Welfare Statistics. Entries in columns (III)–(V) are authors' calculations based on the register data for the immigrant and native samples used in Figures 4 and 5.

Third, since European legislation implies that welfare state entitlements are transferred to the country of employment, a job in, say, Norway, not only grants membership in the Norwegian social insurance system, but also entails eligibility to various family allowances. For families with children, this entails that a job in Norway may be attractive even if the

[4] Although the higher costs of living in Norway will mitigate some of the cross-country differences uncovered in Table 1, the possibility of exporting benefits justifies comparisons without adjustments for purchasing power parity.

offered wage is extremely low. For example, the Norwegian cash-for-care subsidy for a one-year old child now amounts to NOK 6,000 per month, which adjusted to the 2010 wage levels and exchange rates used in Table 1 corresponds to EUR 629, or around 80% of average earnings in Poland. Such features give employers and prospective immigrant employees incentives to agree on very low wages and poor working conditions. While this can be a win-win situation for the employer and the immigrant worker – at least in the short run – it may stimulate the creation of poor jobs with high subsequent unemployment or disability risk and substantial (expected) costs for the welfare state.

2.5 A more robust welfare model

There is now an ongoing policy debate in several European countries about reestablishing elements of the previously existing migration barriers; either by making eligibility of economic transfers from a particular country conditional on past social security contributions to that same country (i.e., limit the transferability of eligibility), or by adjusting benefits with a country specific cost-of-living index when they are exported to another country. The latter would mean, for example, that Norwegian UI benefits paid out in Poland would be cut by more than one half compared to the current level.

Requiring a country-specific contribution period before benefits can be claimed may reduce some of the incentive distortions discussed in the previous section, particularly those related to creation and acceptance of very poor and short-lived jobs. However, this potential advantage must be balanced against the disadvantage of also reducing welfare-enhancing labor mobility within Europe. Further, introducing cost-of-living adjustments to social insurance payments appears to us to be a "dead end". Given that people can travel freely across Europe without notifying authorities of their whereabouts at each point in time, the scope for circumventing downwards cost adjustments appear almost limitless (unless draconian control measures are implemented).

In any case, we will argue that a strategy designed to strengthen the sustainability of the Nordic welfare model primarily must consist of pol-

icies aimed at making the welfare system *robust* to the existence of an open European labor market, and not of policies aimed at reversing it.
How can this be achieved in practice? One obvious policy option is to scale down on social insurance programs and thus make the welfare state less generous for everyone. Such a policy might trigger a "race to the bottom", as it seeks to ensure that a country's own welfare state provisions are not too generous relative to those in other countries. Thriftier social insurance would improve work incentives for natives as well as immigrants, and reduce migration distortions generated by cross-country differences in social insurance systems. However, the relatively generous social insurance programs in the Nordic countries are in place for a reason. They reflect voter preferences for a low-risk society with sound insurance arrangements in case of sickness, disability, or involuntary unemployment. They also reflect preferences for a relatively egalitarian society with little poverty. Viewed as a whole, the "Nordic model" has been successful in delivering economic growth and high standards of living for the vast majority of its citizens.

Is there an alternative way? We will argue that there is, but obviously not without its pros and cons. We will sketch a policy based on three pillars:

A first pillar consists of ensuring minimum standards in the labor market, including a minimum wage and possibly limitations on the use of temporary contracts. In the absence of such minimum standards, employers will have an incentive to recruit foreign workers with a high expected future income flow from the welfare state as such workers are willing to accept lower wages, ceteris paribus. This may result in an "adverse selection" of foreign workers (from a fiscal point of view), and also imply particularly high social insurance replacements among those who do arrive, as the progressive nature of social insurance entails that the replacement ratio declines with earnings. In a worst case scenario, firms could repeatedly recruit new immigrant workers to temporary and poorly paid jobs, based on the premise that their "real pay" would come from the welfare state. Hence, a benefit of minimum standards is that they may remove externalities arising from the fact that a job contract in, say, Norway, not only commits the employer, but also the Norwegian welfare

state, to future payments. Another benefit of minimum standards is that they will reduce average replacement ratios of welfare benefits.

A second pillar consists of transforming the social insurance system from being based on passivating and easily exportable cash transfers to being based on active participation. There is now ample empirical evidence showing that the moral hazard problems in social insurance can be considerably reduced by offering income *through active participation* rather than cash benefits only; see Røed (2012) for a recent overview and discussion of the literature. Participation-based social insurance further leads to more favorable selection by attracting persons who wish to work and contribute while deterring persons with low work-morale. In relation to longer term spells of both unemployment and sick leave/disability, active participation entails that social insurance payments to a larger extent become contingent on participation in job search, training, communal work, or vocational/medical rehabilitation programs. In particular, with active participation the design of temporary and permanent disability insurance programs will encourage and support the use of remaining (partial) work capacity, if necessary through the establishment of sheltered employment. A job offer is obviously also more place-bound than a cash transfer, and cannot readily be exported to a home country with lower costs of living.

The third pillar consists of transforming family allowances from being based on cash transfers to being based on the supply of free/affordable family-friendly and work-oriented services. For example, instead of offering (exportable) cash-for-care subsidies, the welfare state can offer high-quality child care directly.

Now, all of these policies also have some downsides. Higher minimum standards in the labor market distort the price-adjustment mechanism in the labor market and may raise unemployment among low-skilled workers. A more activity-oriented social insurance system may come with high administration costs and may require a large number of sheltered workplaces adapted for persons with reduced work capacity. Offering publicly provided childcare instead of cash transfers reduces the families' freedom of choice. Finding the optimal policy inevitably involves some tradeoffs. It is about balancing conflicting arguments. The point we wish to emphasize here is that policy makers actually have a

range of options. There are viable alternatives to the reintroduction of migration barriers and to benefit-cutting competition. We will also argue that the extension of the common European labor market to include countries with lower economic development implies, ceteris paribus, that the optimal balancing points shift toward higher employment standards and more place-bound social insurance and family support programs. Hence, if the policy was close to its optimum prior to the enlargements of the European labor market, it probably needs considerable adjustment now.

2.6 Conclusion

The recent enlargements of the European labor market represent a considerable challenge for welfare state economies. Provided that we wish to preserve the freedom of movement across European national borders, welfare state institutions in rich countries need to adapt. We have argued that business as usual is not a viable option, and that we either need to scale down on income insurance and family support programs (a race to the bottom), or make welfare state institutions more migration robust by i) raising minimum standards in the labor market, ii) making income insurance programs more participation oriented, and iii) substitute place-bound services like free childcare for exportable cash transfers in family support programs.

References

Barrett, A., & B. Maître (2013). Immigrant welfare receipt across Europe, *International Journal of Manpower* 34(1), 8–23. http://dx.doi.org/10.1108/01437721311319629

Boeri, T. (2010). Immigration to the Land of Redistribution, *Economica* 77(308), 651–87. http://dx.doi.org/10.1108/01437721311319629

Bratsberg, B., Fevang, E., & Røed, K. (2013). Job Loss and Disability Insurance, *Labour Economics* 24, 137–150. http://dx.doi.org/10.1016/j.labeco.2013.08.004

Bratsberg, B., Raaum, O., & Røed, K. (2010). When minority Labor Migrants Meet the Welfare State, *Journal of Labor Economics* 28(3), 633–676. http://dx.doi.org/10.1086/650546

Bratsberg, B., Raaum, O., & Røed, K. (2014). Immigrants, Labour Market Performance, and Social Insurance, *The Economic Journal* 124(580), F644–F683. http://dx.doi.org/10.1111/ecoj.12182

Clemens, M.A. (2011). Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?, *Journal of Economic Perspectives* 25(3), 83–106. http://dx.doi.org/10.1257/jep.25.3.83

Kahanec, M., Pytlikova, M., & Zimmermann, K.F. (2014). The Free Movement in an Enlarged European Union: Institutional Underpinnings of Economic Adjustment, IZA DP No. 8456.

Kennan, J. (2012). Open Borders, *Review of Economic Dynamics* 16(2), L1–L13. http://dx.doi.org/10.1016/j.red.2012.08.003

Kvist, J. (2004), Does EU enlargement start a race to the bottom? Strategic interaction among EU member states in social policy, *Journal of European Social Policy* 14, 301–318. http://dx.doi.org/10.1177/0958928704044625

Mollerstrom, J. (2015). Ethnic fractionalization and the demand for redistribution – Potential implications for the Nordic model, *Nordic Economic Policy Review* (this volume).

Nannestad, P. (2004). Immigration as a Challenge to the Danish Welfare State? *European Journal of Political Economy* 20(3), 755–67. http://dx.doi.org/10.1016/j.ejpoleco.2004.03.003

Preston, I. (2014), The Effect of Immigration on Public Finances, *Economic Journal* 124 (580), F569–592. http://dx.doi.org/10.1111/ecoj.12180

Rege, M., Telle, K., & Votruba, M. (2009). The Effect of Plant Downsizing on Disability Pension Utilization, *Journal of the European Economic Association 7*(5), 754–785. *http://dx.doi.org/10.1162/JEEA.2009.7.4.754*

Røed, K. (2012). Active Social Insurance. *IZA Journal of Labor Policy* 1:8 (2012) http://dx.doi.org/doi:10.1186/2193-9004-1-8

Sinn, H-W. (2002). EU enlargement and the future of the welfare state, *Scottish Journal of Political Economy* 49, 104–115. http://dx.doi.org/10.1111/1467-9485.00223

3. Future Pathways for Labour Market Policy: Including the Excluded

Michael Rosholm and Michael Svarer, Department of Economics and Business Economics, Aarhus University

Abstract

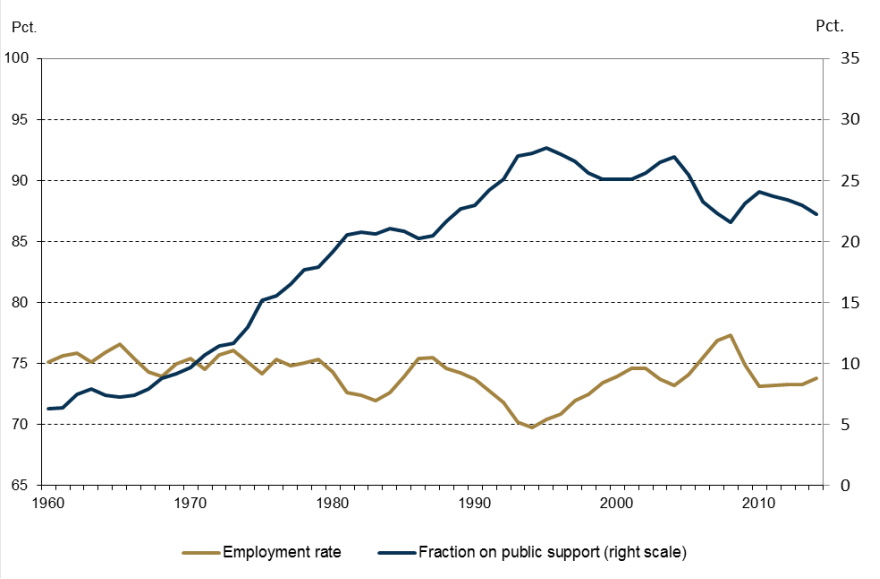
It is a challenge for the Nordic countries to increase the employment rate for people on the edge of the labour market. We discuss the potential for reducing exclusion from employment with a special focus on Denmark. A series of labour market reforms in the recent decades have reduced the number of people on public income transfers and are expected to have a further positive impact on the employment rate in the years to come when the reforms are fully implemented. The main focus of the reforms have been on increasing economic incentives to work for people in or close to the labour market and on fine tuning active labour market policy for unemployed in the unemployment insurance system. A remaining challenge for current labour market policy is to increase the participation rate for those who are at risk of getting excluded from the labour market. We present new evidence on active labour policies for people on the edge of the labour market and discuss the potential for increasing employment rates for this particular group of unemployed.

3.1 Introduction

In comparison to most other OECD countries, the Nordic labour markets are characterized by high participation rates. Still, the fraction of people on the edge of the labour market who are either on sickness benefits, social assistance benefits or disability pensions remains high. It is, therefore, a major economic political challenge to support inclusion of people who currently are more or less excluded from the labour market. This article will concentrate on the Danish case and will focus on the role of active labour market policy to stimulate the supply side of the labour market. Alternative measures like in-work tax credits for vulnerable groups in the labour market or firm-based policies to encompass individuals with low working capacity into firms are not considered. Although the focus is on Denmark and narrowed to a particular set of policy instruments, the main conclusions and empirical findings are clearly relevant for the other Nordic countries as well since they face similar challenges and also rely heavily on the use of active labour market policies.

The employment rate in Denmark has been remarkably stable in the past 50 years. Figure 1 shows that close to 75% of the population aged 18–64 were employed in 1960, and the same is true today. The gender composition of the employed has changed remarkably, though. From the mid-1960s and onwards, the female labour force participation rate increased significantly. Despite this huge inflow of female workers, the overall employment rate barely changed. The implication is, of course, a similar decline in the male employment rate.

Figure 1: Employment rate and fraction on public support, age 18–64



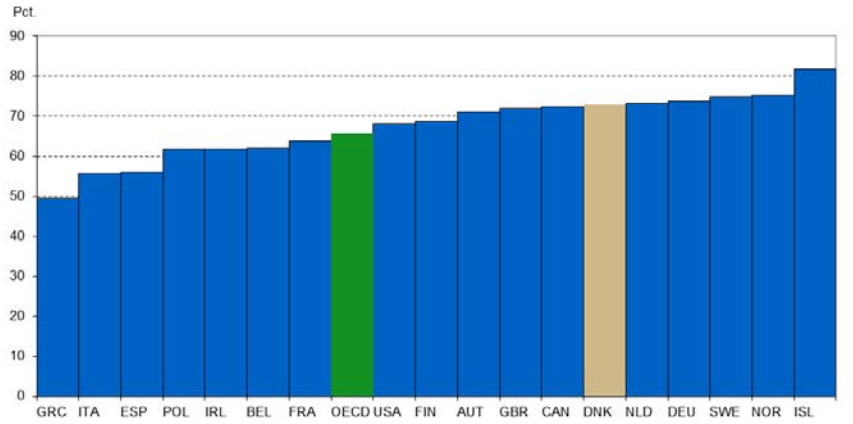
Source: Statistics Denmark, ADAM data bank.

At the same time the number of people receiving some kind of public transfer exhibited an almost exponential growth pattern in the 1970s and 1980s. In the early 1990s the number of people on public support had increased from 200,000 in 1960, corresponding to 6% of the population aged 18–64, to almost 1 million, representing almost 27% of the population in the age group 18–64. Since then a series of labour market reforms in combination with more favourable macroeconomic conditions have brought the fraction in the age group 18–64 receiving public transfers down to around 22%. The drop in the number of people on public support since 1993 has been mirrored in the employment rate, which has risen from 70 to close to 75%. An interesting point to note is that almost the full population in working ages in Denmark is either working or receiving some type of public transfer. In contrast to the vast majority of other countries, the fraction of people not working is entitled to some kind of public transfer.

Currently, around 800,000 Danes between 18–64 years old receive some type of public transfer. Approximately 100,000 of these are expected to become employed within the next 5 years or so, when the business cycle normalizes, and when the impacts of recent reforms in the unemployment insurance system and early retirement systems are realized (Danish Economic Council, 2015). If this assessment turns out to be correct, the employment rate will break the 50 years old glass ceiling and reach a level around 78%.

As shown in Figure 2, the employment rate in Denmark is relatively high in an international comparison, and is only marginally lower than the other Nordic countries, with the exception of Iceland, which has a remarkably high employment rate.

Figure 2: Employment rate in selected countries, 2014



Note: In per cent of population between 15–74 years old.

Source: OECD.

In any event, with the expected increase in the employment rate over the next five years, it will be a challenge for Danish labour market policy in the future to increase the employment rate even further. The structural unemployment rate is at an all-time low, and it is difficult to perceive policies to lower it even further.

There are several reasons why persons of working ages do not work. A common theme of the recent reforms in Denmark related to unemployment insurance, social assistance and early retirement has been to increase incentives to work by reducing the generosity of and in particular the eligibility for public transfers. The assessment of e.g. the Danish Economic Council (Danish Economic Council, 2015) is that these reforms have had a positive impact on employment rates.

Another strategy in order to increase employment rates is to improve the employability and qualifications of those receiving public support. Active labour market policies constitute an important supporting pillar of the flexible labour markets of Denmark and several other European countries. In Denmark, these policies have contributed to lowering the gross unemployment rate to remarkably low levels since the early 1990s (see e.g. Andersen & Svarer, 2007). The active labour market policy tools have consisted of traditional measures like classroom training, job search assistance, employment subsidies and job training and have been used quite intensively to lower unemployment rates (see e.g. Andersen & Svarer, 2012).

Whereas benefits cuts and active labour market policies have been used intensively in the last 30 years for those close to the labour market, but not in jobs, less has been done to help individuals who face a longer route into jobs – the excluded or those at risk of exclusion.

In recent years more focus has been devoted to helping unemployed with a high risk of being excluded from the labour market into employment. One important reason for this is that this group is fairly large compared to the structurally unemployed. Another reason is, of course, that a life in social and labour market exclusion is costly, not only to the individual itself, but also to society. It is obvious that policies aimed at individuals on the edge of the labour market have a lower probability of generating higher aggregate employment in the short run than e.g. increasing the early retirement age from 63 to 64 or reducing unemployment benefit periods from 4 to 2 years, and they should therefore be assessed accordingly. On the other hand, for each person among the excluded who finds stable employment, the gain to society is relatively large. This is in particular the case if the policy manages to bring down the inflow into public transfer for young people.

In this article we present a number of recent Danish measures that have been introduced to prevent and combat exclusion for the more vulnerable unemployed, and we provide a brief evaluation and discussion of the potential of including the excluded in the future labour market.

3.2 Active labour market policies for unemployed on the edge of the labour market

It can be argued that for this group, the best way to prevent a life in social exclusion is to provide sufficient capabilities to undertake a qualifying education, as education seems to offer at least partial insurance against a life on the social margin.

There is a large literature (see e.g. Elango *et al*. (2015) for an overview) on how early childhood interventions in nurseries, kindergartens and schools may prevent social exclusion, but this is not the topic of the present paper. We look at policies designed to assist adolescents and adults at risk of exclusion or already excluded.

One important issue is the dropouts from the youth educational system. Dropout rates are fairly low in high schools, but in vocational schools (which provide qualifying educations), the gross dropout rate is around 50%. However, around 20 percentage points of these eventually complete another qualifying education, leaving a net dropout rate of 30%, which is remarkably high and has not changed much in the past 15 years.

Many of the young persons who drop out end up receiving social assistance, and various policies have been attempted to help prepare them for the youth education system, but so far none have been successful. A recent analysis from Rangvid *et al.* (2015) use propensity score matching and IV methods to show that amongst all the preparatory measures used to help young people prepare for the youth education system, only one (production schools) might have positive impacts. However, even this result is not robust (IV-estimation shows positive effects, while matching shows negative effects), and most of the other measures show significant negative impacts. Moreover, all these measures are quite costly. Hence, it would seem that there is room for improvement.

Recently, two interventions have been introduced and tested with the explicit aim of helping this group of at-risk youth; a "bridge building" intervention and a mentoring intervention. In the following, we briefly present some of the main results from the evaluations of these interventions.

3.2.1 Case: Building bridges for young unemployed

In 2013 the Danish labour market authority initiated an intervention aimed at young people without a qualifying education who received social assistance. The focus was especially on those perceived to have some type of disadvantage (academically or socially). The main aim was to help these young persons into education and to ensure that they completed the education they had enrolled in. The intervention was multi-facetted and took place at an ordinary educational institution, typically a vocational school, hence the term "bridge-building". The young persons would meet there each day, attend classes in reading, writing and mathematics, learn to be there on time every day (an important part of the intervention), learn to feel comfortable at the school, visit other educational institutions, and so on. Each person would be assigned a personal mentor, and there would be one (and only one) contact person to the municipal system during the intervention period.

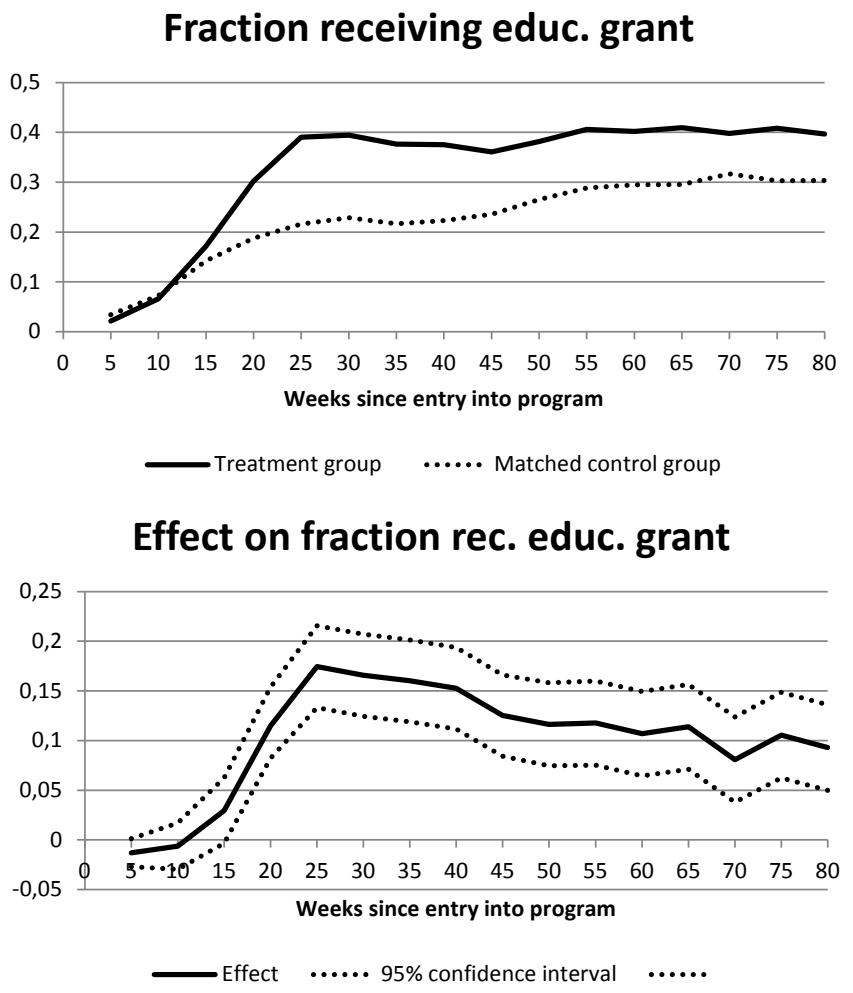
A total of 2,600 young persons participated in the programme, spread over 12 different educational institutions located all over Denmark. The intervention was based on voluntary participation both by the educational institutions and by the unemployed, so there is no exogenous variation in participation that can be exploited to evaluate the effect of the intervention. To obtain a measure of how the intervention has effected educational enrolment, Rosholm & Svarer (2015) use a matching estimator strategy. Based on rich information from administrative registers, grades from compulsory school, labour market and educational histories, socioeconomic status, age, gender and ethnicity, marital/cohabitation status, health status (including information on mental health), drug and alcohol abuse, and a subjective case worker

evaluation of "readiness for education", each participant is matched with up to 10 control persons.[1]

Figure 3 shows the estimated effects of the bridge building intervention on the fraction in education. The left hand panel shows the fraction receiving educational support in the treatment group and a matched control group. In the matched control group, approximately 20% are in education 26 weeks after the start of the intervention, while the treatment group has nearly 40% in education after 26 weeks. After 1½ years, the fraction in education has increased to 30% in the matched control group, while it is still 40% in the treatment group. The figure on the right hand side shows the estimated impact of the intervention as well as the 95% confidence interval.

[1] For more information on this study, see Rosholm & Svarer (2015).

Figure 3: Effects on education

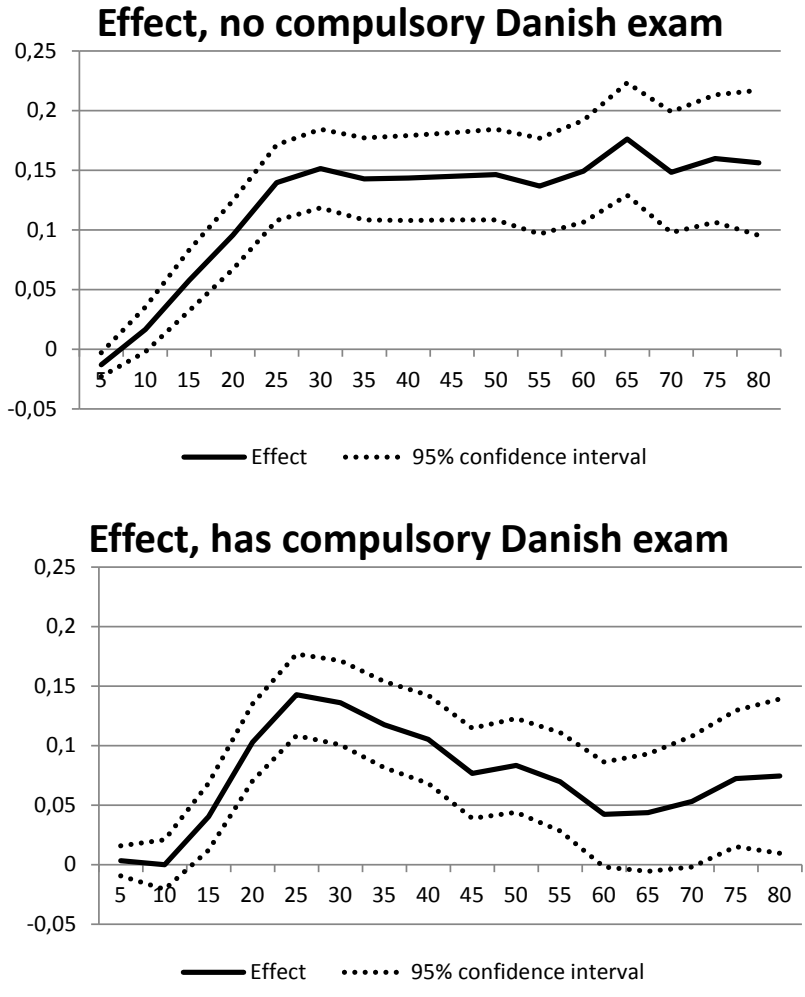


The figure 3a shows the fraction of the treatment- and control group enrolled in education at a particular time measured since the moment they entered the bridge building project. The figure 3b the average treatment effect of the bridge building project from a matching analysis. For more details see Rosholm & Svarer (2015).

First graph in Figure 3 shows that after 26 weeks, the fraction of young persons who are in the educational system has increased by 17 percentage points relative to the matched control group – around 80% in relative terms. 1½ years after programme start, the effect has declined to 10 percentage points, which is still a 25% increase in the fraction in the educational system. Moreover, this effect is statistically significant. Additional analyses show that the fraction completing the basic semester in the vocational schools increases from 5 to slightly above 10%, again statistically significant, and there is also a subsequent positive impact of entering the main part of the vocational track.

Although the individuals in the intervention have relatively poor qualifications, additional analyses show that the intervention seems to work even better for the weaker among these young persons. Figure 4 shows the effects for those who have a grade in Danish from compulsory school leaving exams at grade 9, and for those who do not. Note that for those who do not have such an exam (47% of the treatment group), the effect is 15 percentage points after 1½ years, while it is only 5–10% for those who have such a grade. Although the difference is not statistically significant, it is a notable pattern, since many interventions are often less effective when aimed at weaker groups in the labour market. Similarly, we find larger effects for young persons who have spent more time on social assistance during the past year.

Figure 4: Effects by school leaving exam (or not) in Danish



Note: The figures show the average treatment effect of the bridge building project from a matching analysis for people who do not have an exam in Danish from compulsory school (left) and for those who have an exam in Danish from compulsory school (right). For more details see Rosholm & Svarer (2015).

In terms of cost considerations, the per person cost of the bridge building intervention was DKK 45,000. This is a fairly large upfront investment. However, compared to a lifetime of public income support and associated costs, it may be worthwhile if the substantial effects we find can be sustained in the long run.

All in all, the findings from the bridge building project suggest that such an intervention aimed at helping young vulnerable person into vocational education has increased enrolment into vocational education. The real test for the programme is whether the increased enrolment in education also leads to more completed education and increased employment rates for this group. Future analysis will reveal whether this is the case.

3.2.2 Case: Mentoring of young people on social assistance

In 2012, the Danish labour market authority initiated an intervention that used mentors for young individuals without a qualifying education who were considered to be at risk of having difficulties in the educational system.

In each of 13 job centres, 200 young persons were randomized into either a treatment or a control group. The control group would receive treatment as usual, which would be regular meetings with case workers (every 13 weeks), and occasional participation in some of the active measures aimed at this particular group, which would often be the set of interventions found to be not effective by Rangvid *et al*. (2015).

The treatment group would receive the treatment as usual plus an offer of a mentor for 52 weeks. It was voluntary to accept the mentor, and 92% of the treatment group were assigned a mentor. The mentor is employed by the job centre to support the activities agreed upon between the job centre and the young person in the job- and education plan. The mentor then follows the young person until he or she starts an education or a job, or until a maximum of 12 months has elapsed. The mentor may focus on helping the young person finding a job or an educational institution, but he may also help with more personal tasks, which were not specified in advance. This could be anything from get-

ting up in the morning and attending meetings and activation programmes to helping reach an agreement with the bank on debt repayments, meeting a general physician or a psychologist, or other personal problems the young persons might have.

Of the 1,299 in the treatment group, 1,193 individuals, corresponding to 92% of the treatment group, participated in at least part of the intervention.[2]

Figure 5 below shows the average number of minutes per week each mentee spends with the mentor.

Figure 5: Average weekly time spent with a mentor

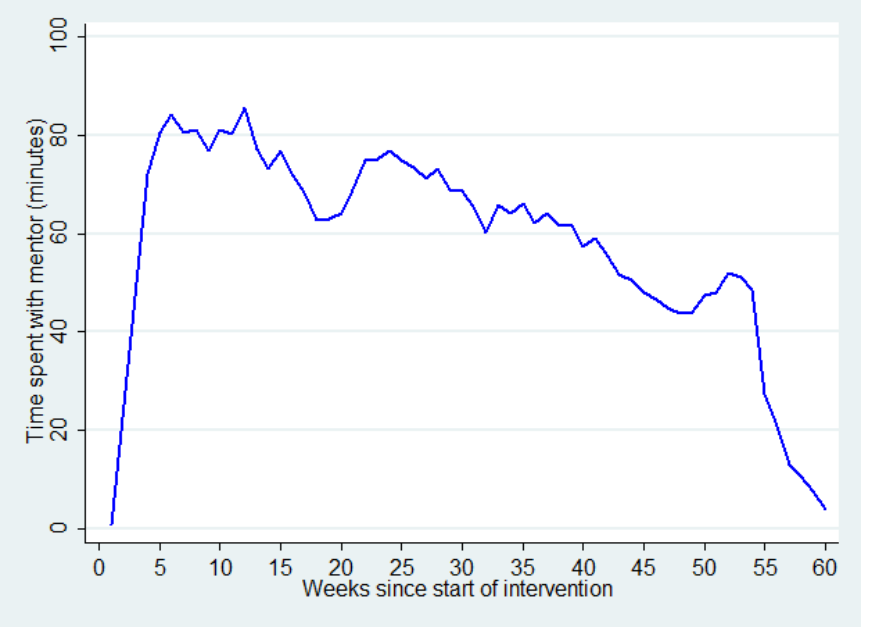
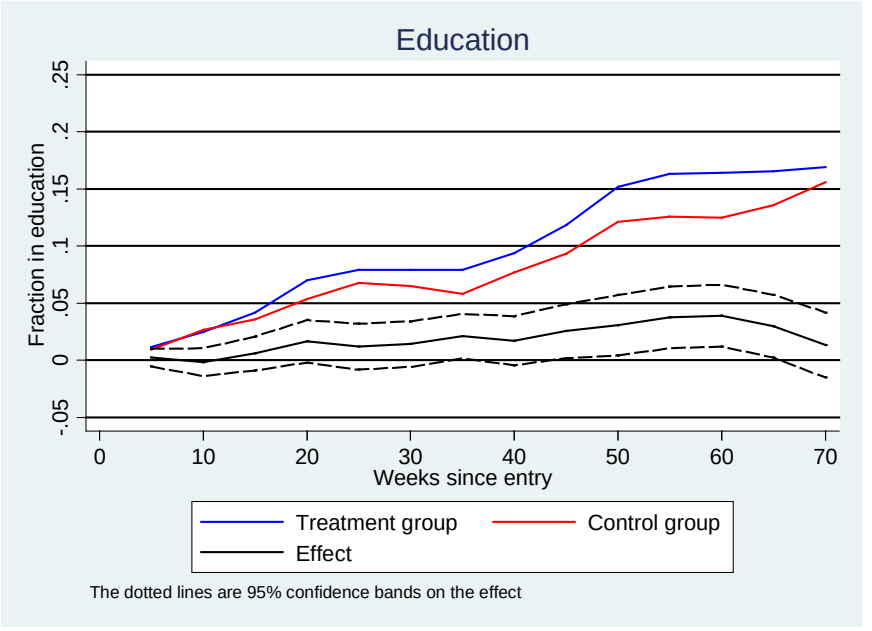


Figure 6 shows the impact of the programme on the main outcome of interest, namely the fraction entering the educational system.

[2] For more details on the mentoring intervention, see Svarer *et al.* (2014).

Figure 6: Effect on education (receiving study support)



The first thing to note is that, in the control group, only around 12% are in education one year after being assigned a mentor, while the fraction in education in the treatment group is around 16%. The effect of the programme is depicted by the solid black line, which shows that there is indeed a 4 percentage points effect of the programme, and this effect is statistically significant after slightly less than a year, although it does tend to disappear again towards the end of the observation period. In Figure 7a–c we show the effect of being assigned a mentor on different sub groups, namely those who have no compulsory Danish exam, those who have a grade below 4 (the median in the sample among those who have a grade), and those who have grade 4 or above. Being assigned a mentor appears to be effective for youth with low grades, but not for those with higher grades or those without a compulsory school leaving exam. This is in contrast to the results for the bridge building intervention, which appeared to be more effective for youth without a compulsory school leaving grade in Danish.

Figure 7a: The effect of the mentoring intervention, by grade in Danish in compulsory school – Education, Danish GPA≥4

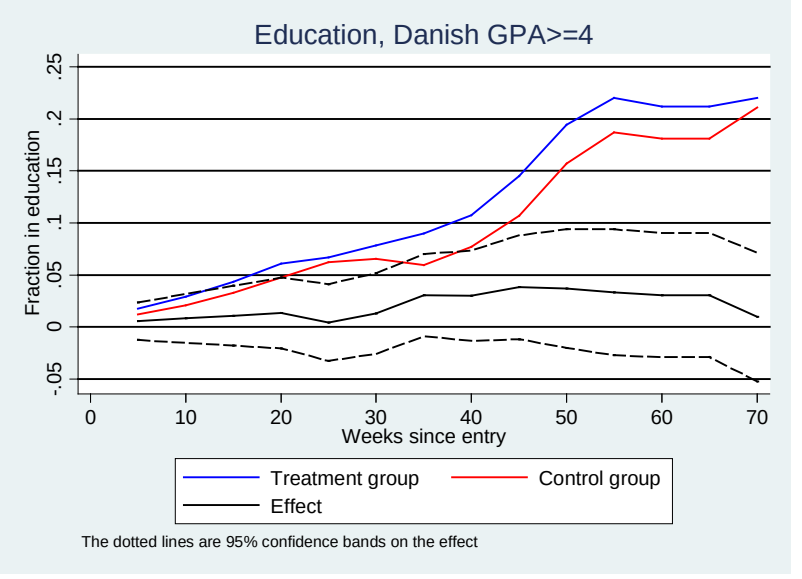


Figure 7b: The effect of the mentoring intervention, by grade in Danish in compulsory school – Education, Danish GPA<4

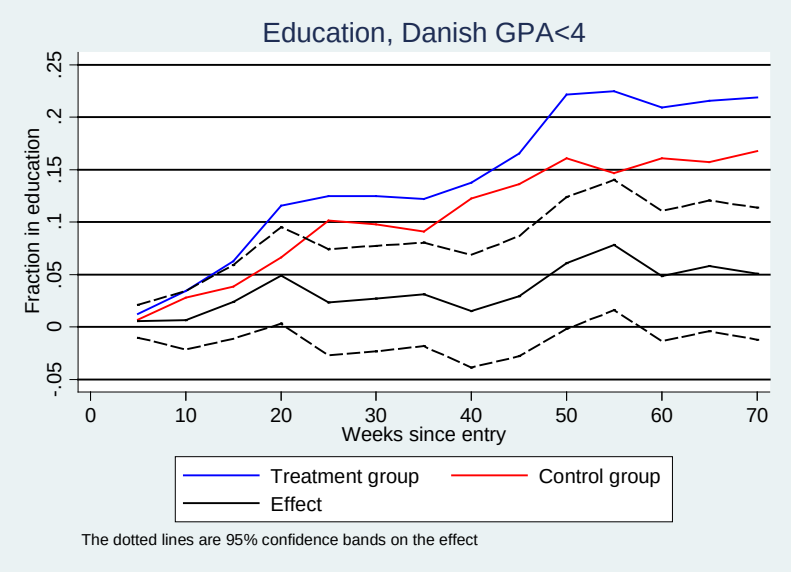
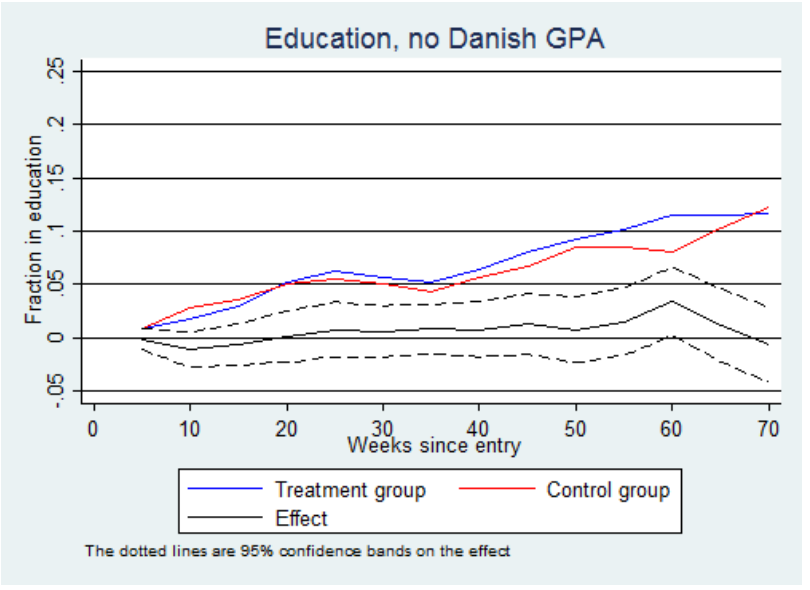
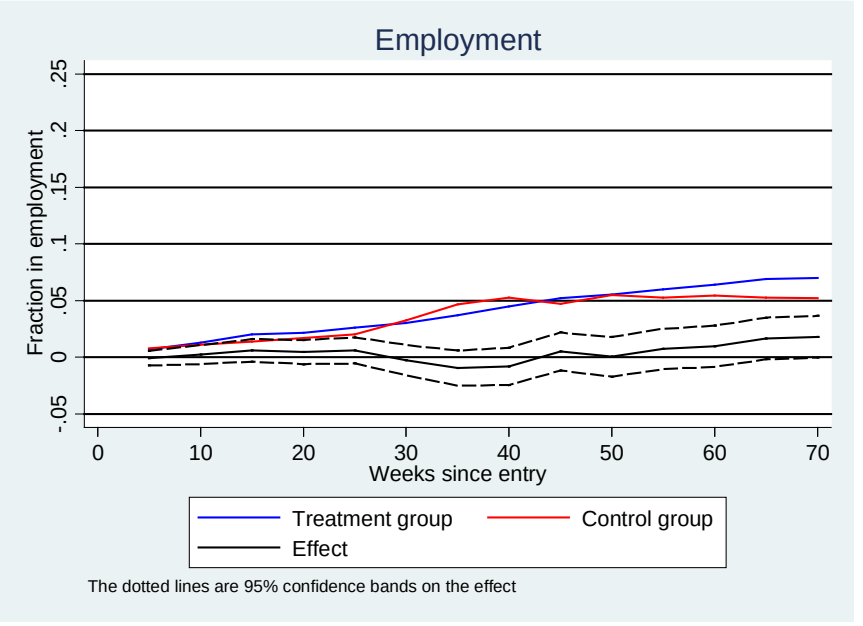


Figure 7c: The effect of the mentoring intervention, by grade in Danish in compulsory school



Finally, in Figure 8 the impact of the mentor intervention on the fraction in employment is shown. Around 5% in the control group are in employment after a year. In the treatment group, this is slightly larger, growing to 7% after 70 weeks. This implies a small but borderline statistically significant effect on employment rates.

Figure 8: Effect of the mentoring intervention on employment



The mentoring intervention suggests that there is a potential to increase educational attainment for a group of vulnerable young unemployed. Future assessments will tell whether the increased enrolment into education results in more completed education and more stable employment paths. At this stage it will also make more sense to conduct a proper cost-benefit analysis of the intervention.

For the mentoring intervention, the price per participant was around DKK 25,000, which is considerably cheaper than the bridge building intervention which also included mentoring. Still, the bridge building intervention yielded impacts that were three times larger than the mentoring intervention, so it may be that the former is still more cost effective.

3.2.3 Case: Measuring Employability

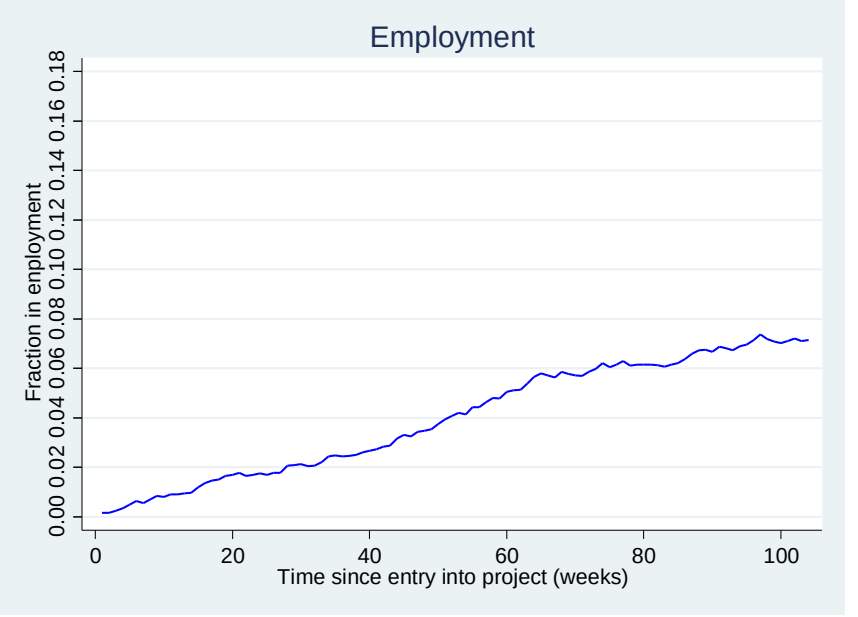
A final important policy issue is that helping the group of adult individuals at risk of social exclusion takes time as it is a group with complex problems and issues. According to the Expert Group (2015), they have very little education, very little working experience, 36% of them use psycho-pharmaca, they have various physical health problems, 17% of them were placed in foster care as children, they have severe debt issues, and so on. Hence, improving their employability requires that they progress in several dimensions, and we do not know which of these are more important, or if progress in some dimensions require progress in other dimensions.

A recent research project (see Rosholm (2015) for details) tries to measure the progress of persons at risk of exclusion by use of repeated surveys. These surveys are carried out as part of the meetings taking place in job centres during meetings with case workers taking place approximately every 13 weeks. Both the case worker and the client is asked to score the client on issues such as access to a network for job search, ability to cooperate, self-confidence, ability to take initiative, extroversion, ability to cope with everyday issues, health and health coping, reservation wages, job search strategies, subjective employability assessments, etc.

The aim of the employability indicator project is to assess whether improvements on these indicators can predict subsequent employment. The problem is that at the outset of this project in late 2013 and early 2014, half of the participants – not immediately employable individuals receiving social assistance – had been on public income assistance uninterrupted at least since the beginning of 2008; that is, for six years. 90% of them had been employed less than 20% of the time in the same period. Hence, their prospects are meagre, and it is very difficult to judge the quality of active labour market policies on an outcome such as employment since they are very far away from employment. Hence, there is a perceived need for intermediate measures that can point towards (predict) employment, but that may be easier to affect in the short run.

Figure 9 shows the fraction in employment among the participants in the project.

Figure 9: Fraction employed in the employability indicator project



Only around 7% of the participants in this project – who do not receive anything beyond the "treatment as usual" apart from their responding to the employability survey at the regular meetings with case workers in the job centre – were employed two years after entry into the programme. This also demonstrates that they are indeed hard to place. Still, the figure hides the fact that 14% of the participants had actually held a job at some point during the two years after entry into the programme.

The project also asks questions about job search behaviour, specifically about the types of job search channels used. There are a total of six different job search channels that can be specified: 1) responding to newspaper adds, 2) via the internet, 3) unprompted applications, 4) using ones network, 5) via job training sites, and 6) via temporary help agencies. If we regress the number of search channels used on a set of individual specific variables in a linear regression model, we get an R-squared of 3%. When we include the employability indicators (initial levels as well as subsequent changes), the R-squared increases to 26%.

Hence, these indicators are certainly highly predictive of commencing job search. In particular, questions on health coping strategies, self-confidence, knowledge about the labour market, purposefulness, everyday coping, cooperation, initiative, and the reservation wage affect the number of search channels used.

Next, when we estimate a linear probability model of an individual being employed in a given week on a set of individual background variables, the R-squared is below 3%. Once the information on employability indicators and job search channels is included, the R-squared increases to 8%. This is not quite as dramatic an increase as we found for job search, but still, we almost triple the explained variation in the model.

The model further reveals that, not surprisingly, starting to search is a very important predictor of eventually finding employment. Moreover, informal channels are – for this particular group – the most effective. In particular, search via job training sites, temp agencies, and unprompted applications are the ones that are significantly associated with finding employment.

Using employability indicators has certain implications for labour market policy. First of all, it enables us to investigate whether different active labour market policies are effective in creating employability in the important dimensions. Moreover, in time, it may enable case workers and clients to be much more specific about the aims of the policies employed, and to design policies to create progress in certain relevant dimensions. It may thus become easier to tailor active policies to this group of very vulnerable – and highly heterogeneous – individuals, for whom we have yet to "crack the nut".

3.3 Conclusion and discussion

As labour market policies become more and more successful, partly due to continual use of the existing evidence regarding their effectiveness, it may become increasingly difficult to reap the fruits of additional refinements in the same policies in terms of reducing e.g. the structural unemployment rates.

Hence, there is a case to take a closer look at policies aimed at individuals at the edge of the labour market; the excluded and those at risk of exclusion. If we can prevent some of the young persons at risk of exclusion from becoming excluded, even a fairly large investment in adequate policies may pay off in the longer term. In addition, if we can become better at helping those who have become excluded back into the labour market through a better understanding of their problem, then we can better tailor policies to help them that may also pay off.

We have discussed a couple of interventions aimed at preventing exclusion with the potential to improve active labour market policies aimed at youth as well as an employability project aimed at better understanding how to create progress for those at the edge of the labour market. This type of policies may become more important in the future if we want to continue to strive to include as many as possible and hence sustain the welfare state in the longer term.

References

Andersen, T. & Svarer, M. (2007). Flexicurity – Labour Market Performance in Denmark, *CESifo Economic Studies*, 53 (3), 389–429. http://dx.doi.org/10.1093/cesifo/ifm015

Andersen, T. & Svarer, M. (2012). *Active labour market policies in a recession*. IZA Journal of Labor Policy.

De Økonomiske Råd (2015). *Dansk Økonomi*. Efterår 2015.

Elango, S., García, J. L., Heckman, J. J., & Hojman, A. P. (2015). Early Childhood Education, forthcoming in *Means-Tested Transfer Programs in the United States*, Volume II, edited by Robert A. Moffitt.

Ekspertgruppen (2015). Nye *veje til job – for borgere i udkanten af arbejdsmarkedet, Ekspertgruppen om udredning af den aktive beskæftigelsesindsats*. Report to the Ministry of Employment.

Rangvid, B. S., Jensen, V. M., & Nielsen, S. S. (2015). *Forberedende tilbud og overgang til ungdomsuddannelse*. SFI-rapport 15:14.

Rosholm, M. (2015). The employability indicator project – some first results, Draft, Aarhus University.

Rosholm, M. & Svarer, M. (2015). *Kvantitativ evaluering af brobygningsindsatsen, report to the Danish Labour Market Board* (http://star.dk/da/Om-STAR/Publikationer/2015/11/Effektmaaling-af-videnspiloten-Brobygning-til-Uddannelse.aspx).

Svarer, M., Rosholm, M., Havn, L. & Høeberg, L. (2014). *Evaluering af mentorindsats til unge uden uddannelse og job, report to the Danish Labour Market Board* (http://star.dk/da/Tal-og-analyser/Ny-viden/Kontrollerede-forsoeg/Mentorforsoeget.aspx).

4. Economics of Innovation Policy[1]

Tuomas Takalo, Hanken School of Economics and HECER and Otto Toivanen, KU Leuven, Aalto University and CEPR

Abstract

We argue that the design of innovation policy in the Nordic countries should better acknowledge i) the uncertainty related to outcomes of innovative activities, ii) the benefits of agglomeration, iii) the effects of being small open economies, and iv) the impact of digitization. All these call for a predictable institutional environment that allows research-resources to agglomerate through a bottom-up process and to flow to their best, often unexpected, uses. Indirect innovation policies such as e.g. basic research, education, competition policy, and financial and labor market regulations may be more important than direct innovation policies such as intellectual property and government support for private R&D, especially in small open economies where benefits from direct support of private R&D and strong domestic intellectual property rights are low.

---

[1] Tuomas Takalo, Hanken School of Economics, E-mail: tuomas.takalo@gmail.com. Otto Toivanen, KU Leuven and CEPR, E-mail: otto.toivanen@kuleuven.be. We thank Ari Hyytinen, Markku Stenborg, Matti Pohjola, Jesper Roine, Petri Rouvinen, and the seminar participants at the NPER conference in Helsinki for useful discussion and comments. Toivanen thanks the Bank of Finland for hospitality. The usual caveat applies.

4.1 Introduction

The key driver of economic growth is innovation (see, e.g., Aghion and Howitt, 2009). This consensus on the policy goal – to foster innovation – has not lead to agreement as to the means to achieve it. Recent books written by academics for the wider audience illustrate the large variation in policy advice: Lerner (2009a) and Acemoglu and Robinson (2013) argue that governments should focus on creating the right institutional environment for the private sector to work. Mazzucato (2013) and Atkinson (2015) make strongly the case that governments should take an active role in choosing the direction of research, development and innovation activities. The objective of this article is to discuss what economic research suggests as to what innovation policy should look like and what role the government should take.

At the heart of the economic approach to innovation policy is the concept of market failure which creates a wedge between social and private returns to innovative activity.[2] The main market failure in the area of innovation is the imperfect appropriability of the returns to research and development (R&D) investments, as innovative firms and individuals cannot capture all benefits that their innovations provide, but share them with consumers and other firms and users (Nelson 1959 and Arrow, 1962).[3] Financial market imperfections in relation to the funding of R&D investments are often mentioned as another important market failure (see Hall and Lerner 2010 and Kerr and Nanda 2014 for surveys). As a result of these market failures, the private sector is likely to invest too little in R&D activities.[4] Roughly speaking, the private sec-

[2] Some scholars such as Nelson (2009) and Mazzucato (2013) forcefully argue that market failure allows a too narrow role for the government, and advocate the systems of innovation approach instead. We primarily view this difference as a semantic one. For example, in the cases that Mazzucato (2013) brings forth to argue for an active government, the government acts to correct market failures such as missing markets, imperfect competition, imperfect information and other systemic problems that are not solved by market forces.

[3] We will use the words "R&D", "invention", and "innovation" almost interchangeably albeit they do involve subtle but important differences. See, e.g., Carlino and Kerr (2015) for a discussion.

[4] R&D projects may also generate negative social externalities (e.g., competition at the marketplace may lead to business stealing and duplication of R&D costs). While in theory these adverse effects of R&D investments could result in overinvestment in R&D, in practice underinvestment due to imperfect appropriability and financial market imperfections is a much more likely outcome. For example, Jones and Williams (1998)

tor should take care of activities where social welfare mainly consists of private profits, and the government should provide those activities with high social returns but low or non-existing private profits, and in the possibly large grey area in-between, the government may design policies that complement innovation in the private sector and steer the private sector to choose actions that are closer to the social optimum.

The starting point of our analysis is the fundamental challenge of innovation policy: how to encourage the development of new innovations while achieving the potentially conflicting goal of ensuring maximal diffusion of those innovations? We stress four features that shape innovative activity and the government's role in it. First, there is considerable uncertainty as to who will succeed in research and in commercializing that research and when. The endemic informational problems in innovation create scope for both positive and negative unintended consequences of government policies.[5]

The second feature which we stress is agglomeration. Evidence (e.g., Jaffe,1989, Cowan and Zinovyeva, 2013, and Carlino and Kerr, 2015) suggests that agglomeration of innovative activities leads to higher productivity, and should thus be encouraged. In our view the best option to foster agglomeration is to invest in high-quality basic research and to build an institutional environment that channels, in a bottom-up manner, human and financial capital to those geographic and intellectual areas that show signs of success.

Our third key feature is the universal good nature of knowledge.[6] This is a two-edged sword for the Nordic countries: On the one hand, it means that the Nordic countries should actively suck in new knowledge generated by the more than 99% of human population living elsewhere.

---

estimate that the socially optimal level of R&D in the US would be 2–4 times the actual one, despite all the policies of promoting innovation that are already in place.

[5] Holmstrom and Myerson (1983) provide an important analysis of how incomplete information affects the social planner's problem of which policy to choose.

[6] Admittedly there is evidence that knowledge spillovers are still to some extent local (which provides a rationale for favoring agglomeration within countries, as discussed in the previous paragraph). At the same time, there is plenty of evidence of increasingly strong international knowledge flows (see, e.g., Griffith *et al*. 2011). As an early example, the first Finnish telephone company was established in Helsinki in 1877, only in a year after Bell got his patent on the telephone in the US.

On the other hand, this means that a large part of the wedge between social and private welfare, i.e., the very basis for an active government role in supporting innovation, disappears.[7] Almost without exception, the existing literature on innovation policy takes a "large country" approach. Some policy conclusions, however, may change markedly when a small open economy approach is adopted.

Fourth, the design of innovation policies should take into account that we are only in the early phases of digitization that is increasing international knowledge flows and bringing other large but unknown changes to us. The best way to prepare for the future is to provide a sound institutional structure that allows the economy to adjust. This calls for increased flexibility at all levels of the institutional set-up, and especially in education.[8]

This takes us to the main point of this article: the most important innovation policies are likely to be "non-innovation" policies that determine the institutional environment for innovation but are not directly aimed at promoting innovation. Education, basic research, financial and labor market regulations, competition and regional policies, and bankruptcy laws are examples of "indirect innovation" policies that may affect innovation more than direct innovation policies such as intellectual property, and government support for private R&D.

The remainder of the paper follows the above themes, In Section 1 we discuss direct innovation policies. Section 2 is devoted to indirect innovation policies. We offer conclusions in Section 3.

[7] A large part of the wedge between social and private welfare is consumer surplus. In the case of, say, Astra pharmaceuticals, most of the consumer surplus generated by Astra's new drugs resides somewhere else but in Sweden and should be ignored when designing an innovation policy that maximizes the social welfare in Sweden. Also, technological spillovers contributing to the welfare wedge partially flow abroad.

[8] For research on the impact of digitization, see, e.g., Greenstein *et al*. (2013) and Goldfarb *et al*. (2015). Brynjolfsson and McAfee (2015) and Bessen (2015) popularize this research.

4.2 Direct innovation policies

4.2.1 Intellectual property rights

Intellectual property has many facets that have been extensively analyzed (see, e.g., Menell and Scotchmer, 2007, for a survey). Intellectual property attempts to solve the fundamental tradeoff of innovation policy by legal means, as it confers an innovator a temporary exclusive right to her innovation. This right provides a possibility to monetize innovation and thereby enhances the incentives to innovate. After the right expires, the innovation and protected knowledge becomes freely usable. The basic disadvantages of intellectual property right are the reduced consumer surplus and technological spillovers that follow when the property right is in force. Basic economic theory (see, e.g., Takalo,2001, for a summary) suggests that as a result of these trade-offs, there should be an inverse-U shaped relationship between social welfare and the strength of intellectual property protection.

Somewhat puzzlingly, however, to date there is little evidence that stronger intellectual property generates more innovation (see. e.g., Boldrin and Levine, 2008 and Lerner, 2009b). As a necessary condition for a welfare improving intellectual property policy is that it enhances incentives to innovate, this suggests that weaker intellectual property rights would be optimal.

Over the recent decades economic research of intellectual property has focused on cumulative innovation, which has produced a more nuanced view of the intellectual property system. On the positive side, the intellectual property system has created a market for knowledge (for evidence, see, Branstetter *et al*., 2006, Serrano, 2010 and Galasso *et al*., 2013) that in some circumstances may have facilitated knowledge transfers and financing of innovations. But the literature has documented another major draw-back of the intellectual property system: the boundaries of intellectual property rights are inherently imprecise and are ultimately defined by courts. From an innovator's point of view this leads to a threat of intellectual property disputes, which acts like as a tax on innovation. As a result, the basic theoretical result of the positive effect of stronger intellectual property on innovation may be over-

turned when innovation is cumulative and boundaries of intellectual property imprecise (see, e.g., Bessen and Maskin, 2009), potentially explaining the puzzling empirical results.[9]

Even when these more complex effects are acknowledged, stronger intellectual property rights are hardly welfare improving. If anything, recent empirical research suggests that social costs related to imprecise boundaries of intellectual property rights are rising and, at least in the US, may exceed the social benefits of the intellectual property system (Jaffe and Lerner, 2004, Boldrin and Levine, 2008, Bessen and Meurer, 2008, and Turner *et al*., 2013).[10]

For a small open economy, an optimal intellectual property system would probably warrant strong intellectual property rights in the rest of the world but weak intellectual property rights at home (Scotchmer 2004a). This would allow the country's own citizens and firms to use and experiment with innovations developed elsewhere more easily but exporting firms would nonetheless have incentives thanks to strong intellectual property rights abroad. The drawbacks of the strong intellectual property rights would be borne by citizens and firms abroad.

4.2.2 Government funding of private R&D

Public funding of private R&D through subsidies, soft loans, and tax incentives is a widely used policy tool. OECD countries spent almost USD 50 billion of taxpayers' money on supporting private R&D in 2013.[11] Governments have also adopted more tools over time, especial-

[9] Some recent empirical papers attempt to test the predictions of the basic theory of intellectual property, and the theory of intellectual property with cumulative innovation separately: See Izhak *et al*. (2015) for the basic theory, and Williams (2013), Sampat and Williams (2015), and Galasso and Schankerman (2015) for cumulative innovation. The findings of these studies support those of the earlier ones: A positive causal effect of stronger intellectual property on innovation is difficult to come by.

[10] An important exception is Aghion *et al*. (2014) who show that countries with stronger intellectual property regimes may benefit more from reforms that enhance competition in the marketplace.

[11] We arrive at this figure by multiplying Business Enterprise R&D (BERD) measures in 2010 PPP USD by the percentage of BERD financed by government, obtained from OECD Main Science and Technology Indicators www-site (accessed 16 September 2015). The same figure for the 5 Nordic countries was a little over USD 1billion.

ly introducing R&D tax credits (e.g., Finland introduced tax credits for 2013–2014, and Sweden in 2014).

The basic mechanism of most of these support schemes is similar in that the government pays some fraction of the marginal cost of R&D.[12] Lowering the marginal cost means that a supported firm invests more, at least partially closing the gap between the privately and socially optimal levels of R&D. There is also a hope that additional finance by the government would attract new firms to start R&D, but recent research shows that existing policies merely lowering marginal costs of R&D are not effective policy tools to this end (see Czarnitzki *et al*., 2015 and Lach *et al*., 2015). Extrapolating the results from the literature on corporate taxation (e.g., Devereux and Griffith, 1998), it is likely that average R&D cost, and not the marginal one, is what matters for the firms' discrete decisions on whether to start investing in R&D or not.

These financial support policies have also important differences. First, subsidies can be tailored for each project for which the government receives an application (for research that makes use of this, see Takalo *et al*., 2013a), whereas every eligible firm can make a claim for tax credits.[13] One thus needs to trade off the propensity of firms to apply and receive support with the government's ability to tailor the support to the particular project. The application process for subsidies also means that the government may become a focal point for information on emerging agglomeration patterns.

Second, tax incentives in their purest form only work for firms that are profitable and pay taxes. This severely hampers their effectiveness in encouraging start-up innovation. Many countries like Norway and the Netherlands have therefore resorted to "subsidy-like" tax incentives where the R&D-performing firm gets what amounts to a discount on labor-related social costs and taxes. A further problem with tax credits, especially if they have a cap, is that a large part of government expendi-

[12] This is the case e.g. in all the European countries whose schemes we are familiar with.

[13] In several European countries the probability of applying for an R&D subsidy is usually below 10%, and below 20% even for R&D performing firms (see Czarnitzki *et al*. 2014). One should however note that the uptake of R&D tax credits is not universal either. Busom *et al*. (2012) report a usage rate of less than 50% in Spain for R&D – performing (i.e., eligible) firms, and in the Netherlands the usage percentage is round 80% for firms with > 10 employees and round 40% for smaller firms (Verhoeken *et al*. 2012).

ture consists of transfers to firms investing beyond the cap, with no incentive effect.[14] For this reason some countries (e.g., US) give tax credits on incremental R&D. This in turn distorts firms' investment decisions over time.

Given the amounts of tax euros channeled to private sector R&D through these policy tools, it is no surprise that a vast empirical literature studying their treatment effects exists.[15] Takalo *et al*. (2013b) emphasize that the extent to which government support increases private R&D do not directly map into social benefits. The reason is that a firm equates the private benefits of R&D with the marginal cost of R&D, but ignores consumer surplus and knowledge spillovers. For example, a small increase in an investment in an R&D project creating large consumer surplus and spillovers may be socially much more beneficial than a large increase in R&D in a project with small (but still positive) consumer surplus or spillovers.

In small open economies, one should pay attention to the share of consumer surplus and spillovers flowing outside the borders where they do not benefit the local tax payers (Conti 2015 and Czarnitzki *et al*., 2015). While existing policies typically impose restrictions on offshoring of government funded projects, the open-economy view could call for more radical changes in policy-thinking. For example, if the outflows of consumer surplus and spillovers constitute a large share of the welfare effects of R&D beyond private profits, private R&D without support may be close to the socially optimal level from a national point of view. For another example, while the standard theory suggests that R&D projects waiving (strong) intellectual property should be prioritized when granting R&D subsidies, in a small open economy the argument is weaker in the case of exporting firms. These open-economy considerations also suggest that the benefits from international coordination of

[14] For example, both the Finnish R&D tax credit scheme (that was in place 2013–2014) and the Swedish one introduced in 2014 have such a cap.

[15] For literature surveys on the effects of R&D subsidies, see David *et al*. (2000), Klette *et al*. (2000), Garciá-Quevedo (2004), Cerulli (2010), and Zúñica-Vicente *et al*. (2014), and on the effects of R&D tax incentives, see Hall and van Reenen (2000), Mohnen and Lokshin (2010), and European Commission (2013).

R&D support policies could be large (see Czarnitzki *et al*., 2015 for a welfare comparison of national versus EU-wide support policies).

4.2.3 Other innovation policy tools

Prizes and contests are an old way of supporting innovation (see Scotchmer, 2004b) but over the past century they have been relatively little used. Using Maurer and Scotchmer's (2004) classification of prize types, *targeted prizes* are posted ex ante by a sponsor (e.g., a public agency) who has identified a problem to be solved. The prize is awarded to the first entity that solves the problem. For example, the Clay Mathematics Institute announced in 2000 a USD 1,000,000 prize for the first solution for each of seven unsolved mathematical problems.

Blue-sky prizes are awarded ex post for innovations that the sponsor considers valuable. A blue-sky prize could be granted in an ad hoc manner each time the sponsor observes a particularly valuable innovation, or the sponsor can commit to grant the prize. The Nobel Prize is the most well-known example of blue-sky prizes, and the Finnish Millenium Technology Prize another. The incentive of effects of blue-sky prizes are probably quite small, and they should be seen more as a marketing tool.

In contrast, targeted prizes could constitute an efficient innovation policy tool. If the rewarded solution is put in the public domain for free use, the prizes completely solve the ex post problem of diffusion of innovations. The problem with targeted prizes is that the sponsor should know ex ante what should be invented.

Setting up contests for targeted prizes helps to aggregate information from innovators, as the sponsor can compare the proposals. Modern information and communication technologies have enabled both the public and the private sector to set up innovation prize platforms (such as Challenge.gov) where not only solutions but also problems are posted. Such crowdsourcing, another manifestation of the changes brought by digitization, provides a new avenue to identify the right problems for prizes and set up contests.

Another tricky task with prizes is to make sure that they reflect the social value of innovations so that they are of proper size. Estimating a proper size for a prize is difficult since this not only depends on the val-

ue of an innovation but also the costs of creating it. Kremer (1998) proposes an interesting public patent-buyout solution to the problem of eliciting information: The patent authority could auction a patent right and use information revealed by bids so as to give an appropriate reward to the patent applicant. To preserve incentives in the auction, a patent grant should de facto be granted with a small probability, otherwise the invention could be put in the public domain. Shavell and Van Ypersle (2001) propose a simpler, but less perfect, mechanism to relate the size of prize to the value of innovation, reminiscent of the royalty-based licensing fees.

Being monetary rewards, prizes are vulnerable to misuse and ex post opportunism (e.g., once the problem is solved, why should the sponsor give the reward).[16] Furthermore, contests inherently involve duplication of R&D costs when the participants race against each other to obtain the prize.

Nonetheless, targeted prizes provide an underused tool of innovation policy. For example, there are numerous diseases that are more prevalent mostly in the Nordic countries. Posting a correctly designed prize would be a simple means to complement (the small) market incentives.

Public procurement and production also provide tools for innovation policy. Governments can provide services to complement private sector innovation, work in partnerships with private entities, buy innovations from private contractors, or directly produce innovations themselves. Such public procurement and production of innovations and complementary services have been widely used thorough the economic history (see, e.g., Scotchmer, 2004b and Mazzucato, 2013), but still may have some untapped potential for innovation policy (Edler and Georghiou, 2007).

In theory, some public innovation support services, direct public production and procurement share the benefits and costs with targeted prizes. On the one hand, the ex ante incentives to innovate can be inefficient, since the decision of what to invent and what information to produce is made by the government. On the other hand, nothing prevents efficient diffusion of innovations ex post. However, a part of public procurement

[16] A classic example of these problems is the Longitude prize (see., e.g, Sobel, 1995).

and production is concentrated on nationally strategic sectors such as defense with the purpose of *minimizing* the diffusion of research results.

Promotion of *research joint ventures* (RJVs) and other forms of R&D cooperation is a widely used tool of innovation policy in industrialized countries. RJVs allow participating firms to internalize technological spillovers and thereby they should enhance R&D efforts. Therefore, RJVs are, for example, prioritized in subsidy allocation decisions in several countries, and constitute a block exemption under the EU competition law. There is some evidence (e.g., Branstetter and Sakakibara, 2002) that RJVs have the stated beneficial effects in enhancing spillovers and R&D efforts. There is however also evidence that RJVs are primarily motivated by cost sharing (Röller *et al*., 2007) and lead to product market collusion (e.g., Hellman and Sovinsky, 2010 and Duso *et al*., 2014).

4.3 Indirect innovation policies

4.3.1 Education

There is rather little robust empirical evidence on the relation between education and innovation.[17] One exception is Toivanen and Väänänen (2015) who find a positive causal impact of education on invention. This suggests that indeed, a policy reaction to Jones's (2005) advice of "having more inventors in order to become richer" as a society is to increase investments in (engineering) higher education.

A key insight from innovation research is the skewed distribution of innovative outcomes, with a low median but a high mean value of innovations (e.g. Pakes, 1986 and Lanjouw, 1998). To us, this seems to call for an education system that generates a wide-skill base and allows different skills to be combined in possibly unexpected ways, i.e., an education system that encourages individuals to acquire a variety of skills and allows individuals with specialized skills to easily match with each other.

[17] There is a very large literature on the causal effect of education on individual wages.

It is well known that innovative activity is concentrated geographically and that high-quality universities play a central role in this agglomeration process (see Audretch and Feldman, 1996, for a seminal paper, and Carlino and Kerr, 2015, for a survey of the empirical evidence). Top universities contribute to the agglomeration of innovative activity in many ways. One important channel is the supply of educated individuals on which innovative activity depends: For example, Moretti (2004) finds a 0.5 percentage point increase in the plant-level productivity as the consequence of a 1 percentage increase in the share of college graduates in the population of a metropolitan area in the US.

As small open economies, the Nordic countries greatly benefit from the knowledge and innovations created elsewhere. While innovation continues to exhibit locational economies of scale also in future, digitization and modern ICT are making knowledge flows less dependent on geography (Griffith *et al*., 2011), suggesting a crucial role for education in enhancing absorptive capacity of the countries.

4.3.2 Basic research

There is plenty of anecdotal evidence of successful private sector innovations that are based on research in government funded laboratories and universities, often without a direct commercial objective in mind (see, e.g., Mazzucato, 2013). But just as in the case of education, there is little in terms of rigorous causal evidence.[18] Basic research done at high-quality universities is a source of significant local knowledge spillovers to the private sector (e.g., Jaffe, 1989, Breschi *et al*., 2006, and Carlino and Kerr, 2015). As innovative firms seek to benefit from these spillovers, they locate close to universities (e.g., Jaffe, 1989, Anselin *et al*., 1997, and Abramovsky *et al*., 2007). This forms another important channel through which universities contribute to agglomeration of innovation (Carlino and Kerr, 2015). In small open economies in particu-

[18] Sveikauskas (2007) offers a survey of the scant literature, and Hausmann (2012) and Akcigit *et al*. (2014) recent contributions.

lar, one should not discount the importance of high-quality basic research as a pull-factor of foreign R&D (e.g. Belderbos *et al*., 2014).

Any government needs to make decisions on how to allocate the resources devoted to basic research. Despite difficulties created by incomplete information, the government may well be in a position to make high-level decisions regarding allocation of resources across different fields of basic research (e.g., health vs. environment). But it should delegate resource allocation decision-making within research fields to its leading experts and allow, through that same system, reallocation across fields as a function of outcomes. Such a bottom-up approach would hopefully lead to a limited number of large, active research centres within each field that would compete against each other for top researches and funds. This should not only improve the quality of basic research but also seed up commercialization of that research (Goldfarb and Henrekson, 2003).

4.3.3 Taxation

The principles of good (corporate) taxation (see Mirrlees *et al*., 2011, pp. 22) minimizes negative effects on welfare and economic efficiency, has low administrative costs; is distributionally fair, and transparent. In cases where production or consumption of goods and services is associated with large externalities, it is theoretically justified to make exceptions to these principles. However, in practice corporate tax incentive schemes tend to become complex and unpredictable and increase tax planning and avoidance (see Mirrlees *et al*., 2011).[19] If tax incentives are used to as an innovation policy tool they should be simple, and focused on innovation or their financing incentives directly. As also concluded by the European Commission (2013), it is therefore much easier to justify, say, R&D tax credits rather than, say, IPR boxes from an innovation policy point of view.

[19] For example, in Finland corporate taxation changes almost annually (e.g., R&D tax credits were in force in 2013–2014, and business angel tax relief was introduced for years 2013–2015).

Just as there is evidence of countries competing in terms of the level of corporate taxation (Devereux *et al*., 2008), they are also likely to use various R&D incentives for the same purpose. In particular, competition for intellectual property revenues is tempting since intangible assets are relatively easy to reallocate from one location to another based on tax considerations (see, e.g., European Commission, 2013 and Griffith *et al*., 2014). In our view, introduction of IPR boxes at best amounts to a Prisoner's Dilemma – game among countries where the detrimental Nash equilibrium should be avoided by international cooperation.

In contrast, tax competition for innovative corporations and individual inventors might be beneficial for the Nordic countries: Danish evidence (Kleven *et al*., 2014) suggests that small open economies with relatively homogenous populations may benefit substantially from tax schemes that give temporary preferential treatment to foreign high-skilled individuals. Akcigit *et al*. (2014) find that top-inventors are sensitive to top income tax rates in choosing where to locate. Taxation of individual inventors should also affect their incentives and individuals' career choices.

Similarly, the effects of (average) corporate taxation are larger at the extensive margin than at the intensive margin: the possibility to make money is one of the key drivers of (high-growth) entrepreneurship (Lerner, 2009a and Isenberg, 2013). The example of earlier successful entrepreneurs and their role as business angels are vital in the creation of a culture of entrepreneurship and risk-taking. However, a large gap between corporate and personal tax rates is conducive for tax planning and avoidance efforts, and successful entrepreneurs and associated capital gains will almost by definition increase income inequality.

4.3.4 Other indirect innovation policies

Besides the policies listed above there is a variety regulatory policies that have a significant impact on innovative activity. We discuss briefly here some selected regulatory policies.

Competition policy is an important part of an innovation infrastructure (Shapiro, 2002, Encaoua and Hollander, 2002, and Segal and Whinston, 2007). According to an extensive literature, there appears to be an

inverse-U relationship between market structure and innovation activity created by two opposing forces: On the one hand, competition is bad for innovation since it reduces the returns to successful innovation; on the other hand, competition is conducive for innovations since it forces the firms to innovate so as to escape competition.[20] This suggests that liberalization of protected and regulated industries might promote innovation. Intensified competition in an upstream industry may also increase innovation in a downstream industry. For example, liberalization of financial services sector not only generated frantic innovation in the industry itself but also in the real sector (Amore *et al*., 2013 and Chava *et al*., 2013).[21]

Trade policy matters for innovation for several reasons. In particular, countries that are open to trade will reap a larger part of international knowledge spillovers and the potentially greatest benefit of innovation investments made elsewhere: new goods and services. While this is uncontroversial, we need to understand much better what shapes international knowledge flows. For example, cultural aspects such as ethnicity may importantly shape international knowledge flows (Kerr, 2007). Trade also increases competition, thereby possibly changing incentives to innovate (Bloom *et al*., 2015).

In general, the beneficial effects of enhanced competition and trade openness on innovation appear to be the largest in countries like the Nordic ones where firms are closer to technological frontier and where corruption does not distort competition (Dabla-Norris *et al*., 2013 and Aghion *et al*., 2013).

From the innovation policy point of view, well-functioning *labor markets* would encourage risk taking and reallocate labor from declining industries and regions to rising ones. Also the efficiency of direct innovation policy tools may depend on the functioning of labor markets. For example, R&D subsidies and tax credit may affect only the wages of

[20] The classic references are Kamien and Schwartz (1975) and Aghion *et al*. (2005). Kilponen and Santavirta (2007) document the existence of the inverse-U relationship in Finland. However, Hashmi (2013) finds a negative relationship between the intensity of competition and innovation in the US.

[21] Some financial innovations clearly generated negative externalities.

R&D personnel if the supply of R&D personnel is inflexible. (e.g., Goolsbee, 1998 and Wolff and Reinthaler, 2008).

Unfortunately the empirical literature on the relation between labor market regulations and innovation is rather unsettled. On the one hand, the Danish-type flexicurity with relatively weak employment protection but relatively high unemployment benefits might be particularly conducive for start-up formation and radical innovation but, on the other hand, weak employment protection may deteriorate employees' incentives to innovate in established corporations (see, e.g., Acharya *et al*., 2013, Bozkaya and Kerr, 2013, and Griffith and Macartney, 2014 for different results).

As mentioned in the introduction, *financial market* imperfections constitute an important rationale for an active innovation policy. R&D activities are inherently opaque, human capital intensive, and involve soft information. As a result, innovative start-ups have difficulties to access to outside finance due to informational asymmetries and lack of collateralizable assets (Hall and Lerner, 2010 and Kerr and Nanda, 2014).

It is notoriously difficult to identify the existence of such financial constraints (see Hall and Lerner, 2010 for various empirical strategies): The fact that some firms suffer from lack of finance may just indicate the financial markets work as they should, and are denying funding of bad projects. Furthermore, even in theory it is difficult to identify the right policy response to these financial market imperfections: Informational asymmetries may even lead to *overfinancing*, which would call for a punitive taxation of start-up finance (e.g., de Meza and Webb, 1987, Boadway and Keen, 2005, and Takalo and Toivanen, 2013).[22] Despite these challenges, two broad conclusions emerge. First, bank lending remains an important source of outside finance, even for start-ups (Robb and Robinson, 2014, Kerr and Nanda, 2014). Bank lending and associated credit constraints are also procyclical (Aghion *et al*., 2012).

Second, the evidence suggests that private sector equity investing is conducive for innovation (Hall and Lerner, 2010 and Kaplan and Lerner,

[22] As illustrated by the dot-com boom and bust at the turn of the millennium, and ongoing financial and economic crisis that begun from the US subprime mortgage markets, this kind of over-financing is not just a theoretical curiosity, and may have severe macroeconomic consequences.

2010). Equity investors have both incentives and human capital for ex ante screening, interim monitoring and value-enhancing advice Furthermore, because innovative investments are complex and risky, optimal financing contracts become complex, too: Investors need to have both a share of upside returns in case of a success and control rights in case of a failure (Kaplan and Strömberg, 2003). Whether private sector equity financing markets work efficiently or not appear to matter more for countries close to technological frontier, such as the Nordic countries (Aghion and Mayer-Foulkes, 2005 and Dabla-Norris *et al*., 2013).

Based on these conclusions, there seems to be a case for policies that improve early-stage equity financing in the Nordic countries. But the right policy is hardly based on public equity investing in commercial projects. Rather one should create the right environment for private sector equity investors. More generally, if there is need for public innovation finance beyond R&D subsidies, the public sector should not mimic private innovation finance but invest differently, operating when liquidity in financial markets dries up and focusing on projects where the ratio of social returns to private returns is high.

Besides the many issues discussed above (e.g., taxation, education, basic research, and labour markets), the legal environment matters for private sector investors, For example, Hyytinen *et al*. (2003) show how a strengthening of the Finnish investor protection legislation enhanced the role of equity finance in the Finnish corporate finance environment.

But again, identifying the right policies to improve legislation is not easy. For example, while a lenient *bankruptcy legislation* clearly encourages entrepreneurial risk-taking by reducing the cost of failure, it also discourages financing of entrepreneurship. The evidence on which of the two opposing effects dominates remains inconclusive (see, e.g., Acharya and Subramanian 2009, and Cerqueiro *et al*. 2014, for conflicting results) and likely depend on the institutional context. There is little research on the Nordic countries regarding the matter, but Koskinen *et al*. (2007) find that weakening of strong creditor rights in corporate bankruptcy in Finland boosted corporate investments and firm valuations.

4.4 Conclusions

The wide consensus that innovation is important for economic growth and thus human well-being is based on a solid theoretical and empirical basis. The theoretical basis for innovation policy is also solid: Because of consumer surplus and technological spillovers which are not captured by innovating firms and individuals, there is too little innovative activity in the private sector. Unfortunately the empirical knowledge of the efficacy of different innovation related policies remains controversial.

The central feature of innovation is uncertainty, and sound innovation policies acknowledge the limited ability of even the best-informed agents to make good choices in tomorrow's increasingly digitized, interconnected world. To us, this clearly suggests an emphasis for a bottom-up approach, rather than vice versa, where resources flow to those sectors and regions that show signs of success. It may well be that the best governmental innovation policies are the least headline-grabbing ones, focusing on building the right infrastructure for better informed agents with stronger incentives, be they academic researchers or corporate inventors.

But in building the better innovation infrastructure the governments should be bold. The Nordic countries stand to continue to do well but only if they maintain and improve a good basic education system, high-quality universities, and an open, competitive and sufficiently flexible environment that enables experimentation and growth of those who succeed in innovation. Sparse public resources can be used more efficiently if the local economies of scale in innovation are better recognized. Since a large part of consumer surplus and knowledge spillovers generated by innovations coming from a Nordic country almost by definition reside elsewhere, the role for more targeted national innovation policies may be more limited than has been thought previously: as examples, domestic intellectual property protection might optimally be weak instead of strong, and direct support to private R&D should be targeted to those projects generating high domestic spillovers, including consumer surplus, rather than those projects aiming at conquering the world.

References

Abramovsky, L., Harrison, R. & Simpson, H. (2007). University Research and the Location of Business R&D, *Economic Journal*, 117, C114-C141. http://dx.doi.org/10.1111/j.1468-0297.2007.02038.x

Acemoglu, D. & Robinson, J. (2012). *Why Nations Fail?* Deckle Edge.

Acharya, V., a& Subramarian, R. K. (2013). Bankruptcy Codes and Innovation. *Review of Financial Studies*, 22, 4949–4988. http://dx.doi.org/10.1093/rfs/hhp019

Acharya, V., Baghai-Wadji, R. & Subramarian, K. (2013). Labor Laws and Innovation. *Journal of Law and Economics*, 53, 997–1037. http://dx.doi.org/10.1086/674106

Aghion, P., Akcigit, U., & Howitt, P. (2013). *What Do We Learn From Schumpeterian Growth Theory?* NBER Working Paper No. 18824.

Aghion, P., Askenazy, P., Berman, N., Cette, G., & Eymard, L. (2012). Credit Constraints and the Cyclicality of R&D Investment: Evidence from France. *Journal of European Economic Association*, 10, 1001–1024. http://dx.doi.org/10.1111/j.1542-4774.2012.01093.x

Aghion, P., Bloom, N. Blundell, R., Griffith, R. & Howitt, P. (2005). Competition and Innovation: An Inverted-U relationship. *Quarterly Journal of Economics* 120: 701–728.

Aghion, P. & Howitt, P. (2009). *The Economics of Growth*. Cambridge, MA: MIT Press.

Aghion, P., Howitt, P., & Prantl, S. (2014). *Patent Rights, Product Market Reforms, and Innovation*. mimeo.

Aghion, P. & Mayer-Foulkes, D. (2005). The Effect of Financial Development on Convergence: Theory and Evidence. *Quarterly Journal of Economics*, 120, 173–222.

Akcigit, U., Baslandze, S. & Stantcheva, S. (2014). *Taxation and International Mobility of Inventors*. NBER wp. 21024.

Akcigit, U., Hanley, D., & Serrano-Velarde, N. (2014). *Back to Basics: Basic Research Spillovers, Innovation Policy and Growth*. NBER wp. 19473.

Amore, M., Schneider, C. & Zaldokas, A. (2013). Credit Supply and Corporate Innovations. *Journal of Financial Economics* 109, 835–855. http://dx.doi.org/10.1016/j.jfineco.2013.04.006

Anselin, L., Varga, A., & Zoltan, A. (1997). Local Geographic Spillovers between University Research and High Technology Education. *Journal of Urban Economics*, 42, 442–448. http://dx.doi.org/10.1006/juec.1997.2032

Arrow, K. .J. (1962). Economic Welfare and the Allocation of Resources for Invention. In Nelson, R. R. (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609–625. Princeton, N.J.: Princeton University Press.

Atkinson, A. B. (2015). *Inequality: What Can Be Done?* Cambridge, MA: Harvard University Press. http://dx.doi.org/10.4159/9780674287013

Audretsch, D. & Feldman, M. (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, 86, 630–640.

Belderbos, R., Van Roy, V., Leten, B., & Thijs, B. (2014). Academic Research Strengths and Multinational Firms' Foreign R&D Location Decisions: Evidence from R&D Investments in European Regions. *Environment and Planning A.*, 46, 9290–042. http://dx.doi.org/10.1068/a45536

Bessen, J. (2015). *Learning by Doing: The Real Connection between Innovation, Wages, and Wealth.* New Haven, CT: Yale University Press.

Bessen, J. & Maskin, E. (2009). Sequential Innovation, Patents and Imitation, *RAND Journal of Economics*, 40, 611–635. http://dx.doi.org/10.1111/j.1756-2171.2009.00081.x

Bessen, J. & Meurer, M. J. (2008). *Patent Failure: How Judges, Bureaucrats and Lawyers Put Innovators at Risk*. Princeton, NJ: Princeton University Press.

Bloom, N., Draca, M., & Van Reenen, J. (2015). Trade Induced Technical Change: The Impact of Chinese Imports on Innovation, Diffusion and Productivity. *Review of Economic Studies*, forthcoming.

Boadway, R. & Keen, M, (2006). Financing and Taxing New Firms under Asymmetric Information. *FinanzArchiv*, 62, 471–502. http://dx.doi.org/10.1628/001522106X172661

Boldrin, M. & Levine, D. K. (2013). The Case against Patents. *Journal of Economic Perspectives* 27: 3–22. http://dx.doi.org/10.1257/jep.27.1.3

Bozkaya, A., & Kerr, W. R. (2013). Labor Regulations and European Venture Capital. *Journal of Economics & Management Strategy* 23, 776–810. http://dx.doi.org/10.1111/jems.12077

Branstetter, L., Fisman, R., & Foley, F. (2006). Do Stronger Intellectual Property Rights Increase International Technology Transfer? Empirical Evidence from U.S. Firm-Level Panel Data. *Quarterly Journal of Economics*, 121, 321–349. http://dx.doi.org/10.1093/qje/121.1.321

Branstetter, L., & Sakakibara, M. (2002). When Do Research Consortio Work Well and Why? Evidence from Japanese Panel Data. *American Economic Review*, 92, 143–159. http://dx.doi.org/10.1257/000282802760015649

Breschi, S., Tarasconi, G., Catalini, C., Novella, L., Guatta, P., & Johnson, H. (2006). Highly Cited Patents, Highly Cited Publications, and Research Networks, mimeo.

Brynjolfsson, E., & McAfee, A. (2015). *The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies*.

Busom, I., B. Corchuelo, & Martinez Ros, E. (2012). Tax Incentives or Subsidies for R&D. UNU-MERIT Working Paper Series 2012–056. http://dx.doi.org/10.2139/ssrn.2330222

Carlino, G., & Kerr, W. R. (2015). Agglomeration and Innovation. In G. Duranton, V. Henderson, and W. C. Strange (eds.), *Handbook of Regional and Urban Economics*, 5, 349–404. Amsterdam: Elsevier. http://dx.doi.org/10.1016/b978-0-444-59517-1.00006-4

Chava, S., Oettl, A., Subramanian, A., & Subramanian, K. (2013). Banking Deregulation and Innovation. *Journal of Financial Economics* 109: 759–774. http://dx.doi.org/10.1016/j.jfineco.2013.03.015

Czarnitzki, D., E. Huergo, M. Köhler, P. Mohnen, S. Pacher, T. Takalo, & Toivanen, O. (2015). *Welfare Effects of European R&D Support Policies*. mimeo.

Cerqueiro, G. D. Hegde, M. F. Penas, & Seamans, R. C. (2014). *Debtor Rights, Credit Supply, and Innovation*. TILEC Discussion Paper No. 2014–011.

Cerulli, G. (2010). Modelling and measuring the effect of public subsidies on business R&D: a critical review of the economic literature. *Economic Record* 86, pp. 421–449.

Conti, A. (2015). *Evaluating Government R&D Grants to Startups: The Case of a Small Open Economy*. mimeo.

Cowan, R. & Zinovyeva, N. (2013). University Effect on Regional Innovation. *Research Policy*, 42, 788–800. http://dx.doi.org/10.1016/j.respol.2012.10.001

Dabla-Norris, E., Ho, G., & Kyobe, A. (2013). *Reforms and Distance to Frontier*. IMF Technical Note to Staff Discussion Note 13/08. Washington, D.C.: International Monetary Fund.

David, B. Hall, & Toole, A. (2000). Is public R&D a complement or a substitute for private R&D? A review of the econometric evidence. *Research Policy* 29, pp. 497–529. http://dx.doi.org/10.1016/S0048-7333(99)00087-6

De Meza, D. & Webb, D. C. (1987). Too Much Investment: A Problem of Asymmetric Information. *Quarterly Journal of Economics*, 102, 281–292. http://dx.doi.org/10.2307/1885064

Devereux, M. & Griffith, R. (1998). Taxes and the Location of Production: Evidence from a Panel of US Multinationals. *Journal of Public Economics*, 68, 335–367. http://dx.doi.org/10.1016/S0047-2727(98)00014-0

Devereux, M., Lockwood, B., & Redoano, M. (2008). Do Countries Compete over Corporate Tax Rates? *Journal of Public Economics*, 92, 1210–1235. http://dx.doi.org/10.1016/j.jpubeco.2007.09.005

Duso, T., Röller, L.-H., & Seldeslachts, J. (2014). Collusion through Joint R&D: An Empirical Assessment. *Review of Economics and Statistics*, 96, 349–370. http://dx.doi.org/10.1162/REST_a_00367

Edler, J. & Georghiou, L. (2007). Public Procurement and Innovation – Resurrecting the Demand Side. *Research Policy*, 36, 949–963. http://dx.doi.org/10.1016/j.respol.2007.03.003

Encaoua, D. & Hollander, A. (2002). Competition Policy and Innovation. *Oxford Review of Economic Policy*, 18, 63–79. http://dx.doi.org/10.1093/oxrep/18.1.63

European Commission (2013). *A Study on R&D Tax Incentives*. Final Report. TAXUD/2013/DE/315. FWC No. TAXUD/2010/CC/104.

Galasso, A. & Schankerman, M. (2015). Patents and Cumulative Innovation: Causal Evidence from the Courts. *Quarterly Journal of Economics*, 130, 317–369. http://dx.doi.org/10.1093/qje/qju029

Galasso, A., Schankerman, M., & Serrano, C. J. (2013). Trading and Enforcing Patent Rights. *RAND Journal of Economics*, 44, 275–312. http://dx.doi.org/10.1111/1756-2171.12020

Galasso, A. & Schankerman, M. (2015). Patents and Cumulative Innovation: Causal Evidence from the Courts. *Quarterly Journal of Economics*, 317–269. http://dx.doi.org/10.1093/qje/qju029

Garciá-Quevedo, J. (2004). Do public subsidies complement business R&D? A meta-analysis of the econometric evidence. *KYKLOS* 57, pp. 87–102. http://dx.doi.org/10.1111/j.0023-5962.2004.00244.x

Goldfarb, B. & Henrekson, M. (2003). Bottom-up versus Top-down Policies towards the Commercialization of University Intellectual Property. *Research Policy*, 32, 639–658. http://dx.doi.org/10.1016/S0048-7333(02)00034-3

Goldfarb, A., Greenstein, S., & Tucker, C. (eds.), (2015). *The Economic Analysis of the Digital Economy*. Chicago, IL: Chicago University Press. Edward Elgar Publishing. http://dx.doi.org/10.7208/chicago/9780226206981.001.0001

Goolsbee, A. (1998) Does Government R&D Policy Mainly Benefit Scientists and Engineers? *American Economic Review*, 88, 298–302.

Greenstein, S., Goldfarb, A., & Tucker, C. (eds.), (2013) *The Economics of Digitization*. Cheltenham, UK: Edward Elgar Publishing. http://dx.doi.org/10.4337/9781784710408

Griffith, R., Lee, S. & Van Reenen, J. (2011). Is Distance Dying at last? Falling Home Bias in Fixed Effects Models of Patent Citations. *Quantitative Economics* 2, 211–249. http://personalpages.manchester.ac.uk/staff/rachel.griffith/PublishedPapers/GriffithLeeVanReenen2011.pdf, http://dx.doi.org/10.3982/QE59

Griffith, R., & Macartney, G. (2014). Employment Protection Legislation, Multinational Firms and Innovation. *Review of Economics and Statistics*, 96, 135–150. http://personalpages.manchester.ac.uk/staff/rachel.griffith/PublishedPapers/GriffithMacartneyREStat.pdf, http://dx.doi.org/10.1162/REST_a_00348

Griffith, R., H. Miller & O'Connell, M. (2014). Ownership of intellectual property and corporate taxation. *Journal of Public Economics*, vol. 112, pp. 12–23. http://dx.doi.org/10.1016/j.jpubeco.2014.01.009

Hall, B. & Lerner, J. (2010). The Financing of R&D and Innovation. In B. Hall & N. Rosenberg (eds.) *Handbook of the Economics of Innovation*, 609–639. Amsterdam, Elsevier. http://dx.doi.org/10.1016/S0169-7218(10)01014-2

Hall, B. & Van Reenen, J. (2000). How effective are fiscal incentives for R&D? A review of the evidence. *Research Policy* 29, pp. 449–469. http://dx.doi.org/10.1016/S0048-7333(99)00085-2

Hashmi, A. R. (2013). Competition and Innovation: The Inverted-U Relationship Revised. *Review of Economics and Statistics,* forthcoming. http://dx.doi.org/10.1162/REST_a_00364

Hausmann, N. (2012). Un*iversity Innovation, Local Economic Growth, and Innovation*. US Census Bureau Center for Economics Studies Paper No. CES-WP-12-1.

Helland, E. & Sovinsky (Goeree) M. (2010). Do Research Joint Ventures Serve a Collusive Function, mimeo. UZH.

Holmstrom, B. & Myerson, R. (1983). Efficient and Durable Decision Rules with Incomplete Information. *Econometrica*, 51, 1799–1820. http://dx.doi.org/10.2307/1912117

Hyytinen, A., Kuosa, I., & Takalo, T., (2003). Law or Finance: Evidence from Finland. *European Journal of Law and Economics*, 16, 59–89. http://dx.doi.org/10.1023/A:1023932325783

Isenberg, D. (2013). *Worthless, Impossible and Stupid: How Contrarian Entrepreneurs Create and Capture Extraordinary Value*. Harvard Business School Publishing, Boston, MA.

Izhak, O., Saxell, T., & Takalo, T. (2015). *Patent Duration, Breadth, and Costly Imitation: Evidence from the US Pharmaceutical Market*. mimeo.

Jaffe, A. (1989). Real Effects of Academic Research. *American Economic Review*, 79, 957–970.

Jaffe, A. B. & Lerner, J. (2004). *Innovations and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It.* Princeton, NJ: Princeton University Press.

Jones, C. (2005). Growth & Ideas, ch. 16 in Handbook of Economic Growth 1B, Aghion P. & S. *Durlauf (eds.), Elsevier B.V.* http://dx.doi.org/10.1016/S1574-0684(05)01016-6

Jones, C. & J. Williams, (1998). Measuring the Social Return to R&D. *Quarterly Journal of Economics*, 113, 1119–1135. http://dx.doi.org/10.1162/003355398555856

Kamien, M. I. & Schwartz, N. L. (1975). Market structure and innovation: A survey. *Journal of Economic Literature* 13: 1–37.

Kaplan, S. N. & Lerner, J. (2010). It Ain't Broke: The Past, Present, and Future of Venture Capital. *Journal of Applied Corporate Finance*, 22, 1–12.

Kaplan, S. N. & Strömberg, P. (2003). Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts. *Review of Economic Studies*, 70, 281–315. http://dx.doi.org/10.1111/j.1745-6622.2010.00272.x

Kerr, W. R. (2007). Ethnic Scientific Communities and International Technology Diffusion. *Review of Economics and Statistics*, 90, 518–537. http://www.people.hbs.edu/wkerr/KerrRestat08_EthProd.pdf, http://dx.doi.org/10.1162/rest.90.3.518

Kerr, W. R, & Nanda, R. (2014). Financing Innovation. Forthcoming in *Annual Review of Financial Economics*. http://dx.doi.org/10.2139/ssrn.2519572

Klette, T. J., Møen, J., & Griliches, Z. (2000). Do Subsidies to Commercial R&D Reduce Market Failures? Microeconomic Evaluation Studies. *Research Policy* 29 (4–5): 471–495. http://dx.doi.org/10.1016/S0048-7333(99)00086-4

Kleven, H., Landais, C., Saez. E. & Schultz, E. (2014). Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark. *Quarterly Journal of Economics*, 333–278. http://dx.doi.org/10.1093/qje/qjt033

Koskinen, Y., Korkeamäki, T., & Takalo, T. (2007). Phoenix Rising: Legal Reforms and Changes in Valuations in Finland during the Economic Crisis. *Journal of Financial Stability* 3, 33–58. http://dx.doi.org/10.1016/j.jfs.2007.02.002

Kremer, M. (1998). Patent buyouts: A Mechanism for Encouraging Innovation. *Quarterly Journal of Economics* 113, 1137–1167. http://dx.doi.org/10.1162/003355398555865

Lach, S., Neeman, Z., & Schankerman, M. (2015). Government Funding for R&D Start-Ups, mimeo.

Lanjouw, J. (1998). Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Values. *Review of Economic Studies*, 65, 671–710. http://dx.doi.org/10.1111/1467-937X.00064

Lerner, J. (2009a). *Boulevard of Broken Dreams. Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed – and What to Do about It*? Princeton University Press, Princeton, N.J.

Lerner, J. (2009b). The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues. *American Economic Review (Papers & Proceedings)*, 99, 343–348.

Mazzucato, M. (2013). *The Entrepreneurial State*, Anthem Press.

Maurer, S.M. & Scotchmer, S. (2004). Procuring Knowledge. *Advances in the Study of Entrepreneurship, Innovation and Growt*h, 15, 1–31, edited by G. Libecap. Amsterdam: Elsevier. http://dx.doi.org/10.1016/s1048-4736(04)01501-2

Menell, P. S. & Scotchmer, S. (2007). Intellectual Property. In M. Polinsky and S. Shavell (eds.), *Handbook of Law and Economics*. Amsterdam: Elsevier.

Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Choate, R., Gammie, M., Johnson, P., Myles, G., & Poterba, J. (2011). *Tax by Design*, Oxford University Press, ISBN 978-0-19-955374-7.

Mohnen, P. & Lokshin, B. (2010). What does it take for an R&D tax incentive policy to be effective? In Ghosal, V. (ed.), *Reforming Rules and Regulations*, pp. 33–58. Cambridge, MA: MIT Press. http://dx.doi.org/10.7551/mitpress/9780262014687.003.0002

Moretti, E. (2004). Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. *American Economic Review*, Vol. 94, No. 3, 656–690. http://dx.doi.org/10.1257/0002828041464623

Nelson, R.R. (2009). Building Effective "Innovation Systems" versus Dealing with "Market Failures" as Ways of Thinking about Technology Policy. In Dominique Foray (ed.), *The New Economics of Innovation Policy*, Edward Elgar: Cheltenham, UK: http://dx.doi.org/10.4337/9781848449169.00007

Nelson, R.R. (1959). The simple economics of basic scientific research. *Journal of Political Economy* 49, pp. 297–306. http://dx.doi.org/10.1086/258177

OECD, (2011). *OECD Science, Technology and Industry Scoreboard 2011*. OECD Publishing. http://dx.doi.org/10.1787/sti_scoreboard-2011-en

Pakes, A. (1986). Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica*, 54, 755–784.

Rob, A. & Robinson, D. (2012). The Capital Structure Decisions of New Firms. *Review of Financial Studies*, 27, 153–179. http://dx.doi.org/10.1093/rfs/hhs072

Röller, L.-H., Siebert, R., & Tombak, M. (2007). Why Firms Form (or Do not Form) RJVs? *Economic Journal*, 117, 1122–1144. http://dx.doi.org/10.1111/j.1468-0297.2007.02069.x

Sampat, B. & Williams, H. (2015). How Do Patents Affect Follow on Innovation? Evidence from the Human Genome. NBER Working Paper No. 21666.

Schankerman, M. & Serrano, C. (2010). The Dynamics of the Transfer and Renewal of Patents. *RAND Journal of Economics*, 41, 686–708. http://dx.doi.org/10.1111/j.1756-2171.2010.00117.x

Scotchmer, S. (2004a). The Political Economy of Intellectual Property Treaties. *Journal of Law, Economics, and Organizations* 20, 415–437

Scotchmer, S. (2004b). *Innovation and Incentives*. Cambridge, MA: MIT Press.

Segal, I. & Whinston. M. D. (2007) Antitrust in Innovative Industries. *American Economic Review*, 97, 1703–1730. http://dx.doi.org/10.1257/aer.97.5.1703

Sobel, D. (1995). *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*. New York, NY: Walker.

Sveikauskas, L. (2007). R*&D and Productivity Growth: A Review of the Literature*. Bureau of Labor Statistics Working Paper 408.

Shapiro, C. (2002). *Competition Policy and Innovation*. OECD Science, Technology and Industry Working Papers, 2002/11, OECD Publishing.

Shavell, S. & T. van Ypserele (2001). Rewards versus Intellectual Property Rights. *Journal of Law and Economics*, 44, 525–547. http://dx.doi.org/10.1086/322811

Serrano, C. J. (2010). The Dynamics of the Transfer and Renewal of Patents. *RAND Journal of Economics*, 41, 686–708. http://dx.doi.org/10.1111/j.1756-2171.2010.00117.x

Takalo, T. (2001). On the Optimal Patent Policy, *Finnish Economic Papers*, 14, 33–40. Takalo, T., Tanayama, T. and Toivanen, O. (2013a). Estimating the benefits of targeted R&D subsidies, *The Review of Economics and Statistics*, 95, 55–272.

Takalo, T., Tanayama, T., & Toivanen, O. (2013b). Market Failures and the Additionality Effects of Public Support to Private R&D: Theory and Empirical Implications. *International Journal of Industrial Organization.* 31, 634–642. http://dx.doi.org/10.1016/j.ijindorg.2013.02.002

Takalo, T. & Toivanen, O. (2012). Entrepreneurship, Financiership, and Selection, *Scandinavian Journal of Economics*, 114, 601–628. http://dx.doi.org/10.1111/j.1467-9442.2011.01685.x

Toivanen, O. & Väänänen, L. (2015). Education and Invention, *Review of Economics and Statistics*, forthcoming. http://dx.doi.org/10.1162/REST_a_00520

Turner, J. L., Bessen, J., Neuhäusler, P., & Williams, J. W. (2013). The Costs and Benefits of the United States Patents, Boston University School of Law, *Law and Economics Research Papers* No. 13–2014.

Verhoeven, W., van Stel, A., & Timmermans, N. (2012). *Evaluate WBSO 2006-2010: Effecten*. Doelgroepbereik en uitvoering. https://www.rijksoverheid.nl/documenten/rapporten/2012/04/02/hoofdrapport-evaluatie-wbso-2006-2010

Williams, H. (2013). Intellectual Property Rights and Innovation: Evidence from the Human Genome, *Journal of Political Economy*, 121, 1–27. http://dx.doi.org/10.1086/669706

Wolff, G. B. & Reinthaler, V. (2008). The effectiveness of subsidies revisited: Accounting for wage and employment effects in business R&D, *Research Policy*, 37, 1403–1412. http://dx.doi.org/10.1016/j.respol.2008.04.023

Zúñica-Vicente, J., Alonso-Borrego, C., Forcadell, F., & Galan, J. (2014). Assessing the Effect of Public Subsidies on Firm R&D Investment: a Survey. *Journal of Economic Surveys*, 18, 36–67. http://dx.doi.org/10.1111/j.1467-6419.2012.00738.x

5. Taxing mobile capital and profits: The Nordic Welfare States

Guttorm Schjelderup,[1] Norwegian School of Economics and Norwegian Center for Taxation, e-mail: guttorm.schjelderup@nhh.no

Summary

This paper discusses trends in capital taxation and the role of the corporate tax rate in a welfare state. It provides a summary of the tax competition literature with special application to capital taxation in small versus large countries. A main finding from this literature is that small countries set lower taxes on capital than large countries. In line with this prediction the paper shows that the Nordic countries undertook tax reforms in the 1990s, which lead to lower ratios of statutory corporate to wage taxes than in most OECD countries. The second part of the paper is devoted to tax base erosion by multinationals and how to combat it. Finally, the paper offers some concluding remarks on redistribution and the pressures of tax competition.

[1] I am grateful for comments by Rolf Aarberge, Jarle Møen, Jørn Rattsø and Dirk Schindler. Financial support from the Norwegian Tax Administration and the Norwegian Research Council is gratefully acknowledged.

5.1 Introduction

The Nordic countries have traditionally been characterized by an extensive welfare state, a homogenous population and labour force, and redistributive taxation. This has changed in recent years. Increased immigration, an aging population, and competition for capital among countries have put pressures on public finances and the welfare state. These changes can be attributed to the globalization process whereby national economies become more integrated. Economic integration takes place in terms of increasing factor mobility, in particular mobility of capital, and rising volumes of trade in goods and services. Globalization has costs and benefits. On the one hand, globalization leads to a more efficient allocation of worldwide resources and thus to higher output and growth. On the other hand, globalization and free capital mobility disrupts employment patterns, makes incomes more volatile, and threatens the government's ability to redistribute income and to provide public services.

An argument frequently used by political lobby groups is that with free capital mobility corporations shouldn't be taxed at all and that taxing investment income is actually bad for workers. The argument is that if you cut taxes on investment income, more investment is encouraged. More investment means people have more equipment and technology to work with, which should increase the productivity of labour and thus wages and economic growth. Put differently, a tax on mobile capital would lead to an outflow of capital that would cause wages to fall; effectively shifting the full burden of the tax on capital onto workers. It is then better to tax workers directly and levy a zero tax on capital.

The argument above relies on strong assumptions, among them that labour is immobile and cannot evade taxation, that there are no country specific rents, and that domestic firms are not owned in part by foreigners. If domestic firms, say, are partly owned by foreigners, taxing capital would imply that some of the tax burden is shifted onto foreigners and that part of the welfare state is then financed by foreigners. This alone may warrant a positive tax on investment capital (see Huizinga and Nielsen 1997). Industrial agglomeration also modifies the zero-tax results. If industrial agglomeration is concentrated in one single country, a

government may, through a positive source tax on capital, be able to exploit the locational rent created by agglomeration forces and thus increase welfare.[2] The zero tax on capital result is also difficult to confirm empirically. Yagan (2014), for example, studies the effects of the 2003 dividend tax cut in the US. He finds that it caused zero change in corporate investment in U.S. unlisted firms and that it had no impact on employee compensation. It did, however, have an immediate impact on financial pay-outs to shareholders.[3] Alstadsæter *et al*. (2015) use Swedish panel data for unlisted firms and find that the Swedish 2006 dividend tax cut did not affect aggregate investment but that it affected the allocation of corporate investment. In particular, they find that relative to cash-rich firms, cash-constrained firms increased their investments after the dividend tax cut.[4]

In fact, there are good reasons to tax capital income at the corporate level. An important reason is that the corporate tax plays an essential withholding function, acting as a "backstop" to the personal income tax. If a country abolished the corporate tax rate, wealthy individuals in particular would be given an incentive to reclassify their labour earnings as corporate income, typically using offshore corporate structures and escape tax. The corporate tax might also be needed to avoid excessive income shifting between labour income and capital income. Finally, the corporate tax also acts as a withholding tax on equity income earned by non-resident shareholders, who might otherwise escape taxation in the source country.

Countries throughout the world have reduced their corporate taxes in an effort to attract or retain corporate investments. The Nordic countries have pioneered what is commonly known as the dual income tax (DIT). It combines a flat tax rate on capital income with progressive taxation of labour income. One of the arguments in favor of the DIT is that it allows policy makers to lower the corporate tax rate to reduce the

---

[2] See Kind *et al*. (2000) on industrial clusters, economic rents and tax policy.

[3] See also Serrato and Zidar (2014) for similar findings on corporate tax cuts.

[4] Very little is known about the effect a dividend cut has on investments by listed firms. The new view of dividend taxation assumes that investments are funded by retained earnings rather than new equity and suggests that listed firms should not be affected by a dividend tax cut (see Auerbach and Hasset 2002).

risk of capital flight, whilst at the same time tax distributed dividends to personal shareholders.

In the continuation I discuss globalization and capital taxation. Tax competition is putting pressure on capital taxes and makes redistribution more difficult and may affect the size and structure of the welfare state and increase income inequality. If what the Nordic countries looked like in the past was the reason for their success, then the future may seem bleaker.

5.2 Globalization, tax competition and trends in capital taxation

The term tax competition is used in the literature to describe how capital taxes are set by independent governments that do not cooperate, and the effect of tax setting on national tax bases. The early contributions consider a country with many identical regions each playing host to competitive firms producing a single output by means of a nationally fixed stock of mobile capital and an immobile factor fixed in supply. The latter could be interpreted as land or labour, and may give rise to pure profits. It is assumed that each region's supply of a public good is financed entirely by a tax on capital employed within its borders (source tax). Tax policy affects the distribution of the country's (world) capital stock.

A fundamental insight is that a rise in the capital tax rate of one region benefits other regions by increasing their capital supplies and, hence, their revenues. Put differently, a tax increase in one region causes a positive externality for other regions. However, the government in each region neglects these externalities since it is only concerned with the welfare of its own residents. The end result is that taxes are set too low resulting in underprovision of public goods. An increase in all tax rates at the same time by a small amount would increase public goods supplies and welfare in all regions.[5]

[5] See e.g. Zodrow and Mieszkowski (1986) and Wilson (1986).

Later amendments to the early theories of tax competition have for example allowed countries to use expenditure levels of public input goods, and multiple tax instruments as strategic variables. These expansions still show that there is still a negative externality from competition over capital that puts pressure on tax rates and the financing of the welfare state.[6]

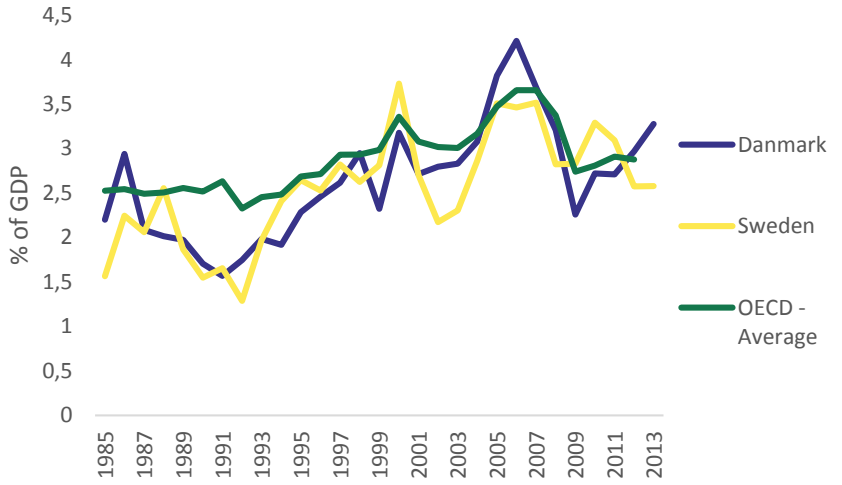
If the literature on tax competition was correct in its prediction one should observe falling tax rates on capital, since capital arguably is the most mobile tax base. Indeed, statutory corporate tax rates in the OECD have fallen substantially since the liberalization of foreign exchange regulations in the mid-80s. The average (unweighted) corporate tax rate in the OECD fell from 48.2% in 1985 to 24.8% in 2015. The same drop is evident in the EU countries where the EU average in 2015 is 22.2%. However, the development in statutory corporate tax rates alone is not a good indicator of whether tax competition reduces tax revenue from the corporate tax. A better indicator is the development in corporate tax revenue. As seen from Figure 1, corporate tax revenue as a share of GDP does not show a drop. There is country variation, but quite a number of OECD countries show a weakly rising trend over time. This is not what theories of tax competition predict. Figure 1 also shows that the corporate tax is a modest revenue raiser, which begs the question why one should worry about corporate tax competition. One reason for why policymakers should worry is that the corporate tax rate is a backstop for the personal tax rate. The latter is a major revenue raiser in all OECD countries.

One reason for the rise in tax revenue is that corporate tax rate reductions have been accompanied by base broadening policies in most countries, for example, by limitations to interest tax deductibility through thin capitalization rules, reduced investment credits and less favourable depreciation allowances. Furthermore, a growing degree of incorporation may also explain part of the broadening of the corporate tax base. Finally,

[6] See e.g. Wildasin (1989), Bucovetsky and Wilson (1991) and Bjorvatn and Schjelderup (2002). For a survey see Wilson (1999) and Wildasin and Wilson (2004).

the reduction in corporate tax rates may have encouraged a shift of income from the personal towards the corporate tax base.[7]

Figure 1: Corporate tax as share of GDP 1955–2013



Source: OECD Tax Database 2015.

Some of the findings in the tax competition literature are particularly relevant for small open economies such as the Nordic countries. The literature models differences in size by assuming that each household owns a unit of capital, but that regions differ by population size (only). Per capita levels are therefore the same in each region, and imply that capital will not move between regions unless taxes differ. In this setting, the literature finds that the small region has an incentive to underbid the large country. Since the large region is a large demander in the international capital market, a reduction in its tax rate (t) will increase the after-tax return to capital (r) substantially. The movement of capital across regions depends on the cost of capital, that is, $r+t$, which means that a reduction in t has a modest effect on the cost of capital. As a result, the large country has weak incentives to bid for capital. In contrast, the small country can-

[7] Alstadsæther and Thoresen (2009) study income shifting between personal and corporate tax bases in Norway for the period after the 1992 tax reform. They find significant evidence for income shifting.

not affect the after-tax return on capital and a reduction in its tax rate will therefore lower the cost of capital by a large amount.

An interesting result in the tax competition literature is that if the small country is sufficiently small, it sets a lower tax rate than the large country and attracts an over proportional share of the total capital stock. By doing so it "wins" the competition for capital in that it can obtain higher per capita utility in equilibrium (Bucovetsky (1991) and Wilson (1991)). The result has interesting policy implications, as small countries may do relatively well in an integrating world. It should be noted, however, that the outcome is still inefficient in the sense that tax rates would be too low and the provision of public goods less than in the absence of competition.

There are other ways of analyzing country size than by population size. Haufler and Wooton (1999) model a multinational firm that wants to sell its goods in both a small and a large country. The large country has more consumers. The firm faces a locational choice: it can only locate and produce its good in one country. They show that firms prefer to locate in large countries (large market size) in order to save transport costs on exports (it is better to export small amounts of goods than large amounts when there are transport costs). As a consequence, the large country can utilize its market size and set a higher capital tax rate. As in the models by Bucovetsky (1991) and Wilson (1991), the outcome of tax competition is that the large country sets a higher tax rate.

In order to put these theories to the test one can group countries according to their size. For example, the G7 countries had an average (unweighted) statutory corporate tax rate of 38.7% in 2000, and this rate had fallen to 31% in 2014.[8] In contrast, the Nordic countries had an average statutory corporate tax rate of 29.4% in 2000 and 22.7% in 2014. The difference in corporate tax rates among these large and small countries is quite telling. The Nordic countries have substantially lower corporate tax rates than the G7 countries.

[8] G7 countries are USA, Japan, Germany, UK, France, Australia and Italy. The Nordic countries are Denmark, Iceland, Finland, Norway and Sweden.

The Nordic countries implemented major tax reforms in the early 1990s partly in response to the pressures of globalization. Klemm *et al*. (2009) study the relationship between wage taxes and corporate taxes. They document that on average, European Union (EU) member states have reduced their reliance on capital taxes and increased the share of labour taxes in total tax revenues during the past 30 years. They also find that the policy responses have been rather diverse. In 2004, the classical welfare states in Scandinavia and continental Europe had lower ratios of statutory corporate to wage taxes than the Anglo-Saxon countries (except Ireland). In 2004, the corporate tax rate was only 63% of the wage tax rate for an average worker in Sweden, but 171% of the wage tax rate in the United States. Such differences are in striking contrast to the common perception that social democratic governments (as in Scandinavia and continental Europe) share a higher preference for redistribution, as compared to more conservative and free market oriented types of governments.

5.3 The political economy of tax competition

In the studies mentioned above an underlying assumption is that politicians conduct policies that are to the best for society as a whole by maximizing the sum of individuals' utilities (Benevolence). A different perspective on tax competition is taken by the public choice literature, which challenges the notion that competition to attract capital is harmful. The basic idea is that competition reduces the rent-seeking activities of government officials and may force a more efficient use of public funds. The literature can be divided into two categories.

The first category does not take into account electoral systems or re-election concerns, but assumes that governments are partly benevolent and partly Leviathan. Hence, government officials are concerned in part with maximizing the public sector by diverting some tax revenue for own consumption. This strand of the literature finds that the outcome of tax competition on tax rates, public expenditures, and welfare

depends on an assessment of the relative strength of Leviathan versus Benevolence.[9] In the context of the Nordic countries, the discussion of Leviathan versus Benevolence could be related to politicians' receiving campaign funding and retirement positions from wealthy individuals and special interest groups in return for favorable policy.

The second part of the literature models tax competition in the presence of voting and there are different approaches to how this is done. Persson and Tabellini (1992) study a two-country model where each government levies a source tax on mobile capital to finance government transfers. A fall in the cost of investing abroad (i.e. increasing competition) puts downward pressure on tax rates. At the same time, however, there is a second, political effect in place since policy is chosen by a policymaker who represents the preferences of the median voter. Tax competition is shown to make the median voter select a more leftist government, whose distributional preferences call for higher taxes on capital, and this partly mitigates the tendency of tax competition to lower taxes on capital.

Biglaser and Mezzetti (1997) study how regions compete to attract large firms. Their starting point is the observation that some US states seem to offer "tax packages" to firms that often exceed the "economic value" of the firm's instate investment project. They assume that when preparing a bid, legislators take into account both the public's interest and the bid's impact on their probability of re-election. The competition among regions follows the rules of an English auction. Since politicians value their re-election, their bid for investments is distorted away from the value of the project to voters and may result in an inefficient location of firms in the sense that legislators give away too much of the taxpayers' money in order to attract firms.

Janeba and Schjelderup (2009) compare the outcomes of increasing capital tax competition under presidential–congressional and parliamentary democracies, in a setting where politicians value rents and reelection to office. In their model, a presidential-congressional system features shifting majorities in the legislature that are issue dependent

[9] See, for example, Brennan and Buchanan (1980); Edwards and Keen (1996); Rauscher (1998).

(here the revenue and expenditure sides of the government budget). The majority that passes tax policy may differ from the majority passing the expenditure allocation. Thus shifting majorities limit the possibility of rent-seeking and increases accountability of elected policy makers. By contrast, in a parliamentary democracy a cohesive majority passes the entire budget in one vote. In a closed economy, the cohesive majority in a parliamentary regime tends to deliver more public goods than under a presidential system because it appeals to voters from all supporting legislators' districts. Yet, the system has also a negative consequence because the majority coalition is powerful and therefore tends to extract more rents. They find that tax competition among presidential–congressional democracies is typically welfare improving, while harmful among parliamentary democracies if under the latter, public goods are sufficiently valued. The results hold when politicians seek re-election because of exogenous benefits of holding office. By contrast, when politicians hold office only to extract rents, tax competition is harmful if politicians are sufficiently patient.

5.4 Profit shifting and multinationals

One of the most pronounced characteristics the last 30 years is the growth in foreign direct investments (FDI) and thus the rising importance of multinationals. Growth rates have been between 10 and 20% annually and an increasing share of trade worldwide is between affiliates of multinational firms. The rising importance of multinationals has gone hand in hand with industrialized countries reforming their corporate income tax policies in order to attract investment. Statutory rates in the OECD, for example, have fallen from an average of 50% in the early 1980s to 25% in 2014. Despite falling tax rates, multinationals have come under fire for siphoning off profits into tax havens. Corporations have responded by saying that their objective is to reduce their worldwide taxes consistent with national laws in order to maximize post-tax global profit. This has prompted governments around the world to overhaul their tax systems and the OECD to launch its BEPS project (OECD, 2013).

The OECD (2013) report on *Base erosion and profit shifting* (BEPS) identifies transfer pricing and debt shifting (thin capitalization) as major reasons for the tax-revenue drain in high-tax countries. Both strategies are regulated by the OECD's arm's length standard, which states that transfer prices should reflect market prices chosen by unrelated parties engaged in similar trades under similar circumstances (Eden, 1998; OECD, 2010, art. 9). Such pricing, however, may be difficult to enforce because of the lack of market parallels, multinationals' use of tax havens, and lack of disclosure of either earnings worldwide or pricing methods.

Multinationals in effect report income by choosing prices on intra-firm trade. By selecting to over-invoice (under-invoice) sales to affiliates in high-tax (low-tax) countries, multinationals can shift profits to low-tax countries and thus save taxes. For instance, royalties for using a brand name or a patent do not have an obvious market parallel; hence, multinationals have considerable discretion in setting prices on such transactions. There is clearly a grey area between strictly legal tax planning and illegal tax evasion. Multinationals may voluntarily or involuntarily cross this line. Furthermore, in some cases, the deviation from the true price of a good or service is so small that the tax authorities would not bother with it, but if the transaction volume is large, substantial amounts of profits can be shifted.

There is substantial evidence of profit shifting by multinational across countries. Pak and Zdanowicz (2001) find that the volume of profit shifting in U.S. multinationals was equal to 18% of total reported corporate profits in 2000. Bartelsman and Beetsma (2003) study OECD data and point out that 65% to 87% of the (potential) additional tax revenue, stemming from a unilateral tax increase, is lost due to profit shifting by transfer pricing. The literature on profit shifting by abusive transfer prices indicates that it is differences in statutory tax rates that provide profit shifting incentives.[10]

[10] Evidence for transfer pricing in the U.S. is given in Clausing (2003) and Bernard *et al*. (2006); for Norway in Langli and Saudagaran (2004); for Germany in Weichenrieder (2009). Oyelere and Emmanuel (1998) show that foreign-owned affiliates in the UK are characterized by lower profits but higher dividend distributions (than UK-controlled firms). Evidence for transfer pricing in European multinationals is given in Dharmapala and Riedel (2013).

There are few empirical studies on Scandinavian data. Langli and Saudagaran (2004) compare the profitability of Norwegian-owned and foreign-owned companies in manufacturing and trade in the years 1993 to 1996. They find that foreign-owned enterprises have a profit margin 2.6 percentage points lower than Norwegian-owned enterprises. This is consistent with a net shifting of profits out of Norway by foreign-owned enterprises. Balsvik *et al*. (2009) expand the data set used by Langli and Saudagaran (2004). They find that multinational corporations shift profits both out of Norway and into Norway but that the net flow is out of Norway. The loss in tax revenue is estimated to be in the order of 30% of the potential tax revenue from foreign multinational enterprises. Another finding in this study is that multinational enterprises in Norway have a profit margin of 1.5 to 4 percentage points lower than comparable domestic enterprises. Their findings, then, are consistent with the findings of Langli and Saudagaran (2004).

The fact that multinationals pay less tax than national firms is not only due to abusive transfer prices. Multinationals can also structure their financing arrangements to minimize tax. The capitalization of a company has an impact on the amount of profit a company reports for tax purposes. Since interest is tax deductible in most countries, a high level of debt, and thus the amount of interest it pays, reduces taxable profit. Lending and borrowing arrangements can be structured so that affiliates in high-tax countries have "too much debt" (thin capitalization) and where the set-up is that interest is received by an affiliate in a jurisdiction that does not tax interest income.

Tax motivated profit shifting by multinationals erodes national tax bases and constitutes a serious risk to tax revenues, tax sovereignty and tax fairness. It also means that multinationals have a competitive edge that has implications for competition in markets and in the long run, the ownership structure in industries. It also implies that the multinationality in the tax base rises and thus that the tax sensitivity of the corporate tax base goes up. The latter may limit the scope for corporate taxes and increase the excess burden from taxation due to a narrower tax base. This prompts the question of what countries can do to reduce this problem.

5.5 Curbing base erosion

The main challenge for tax authorities is to figure out whether intra-firm transactions across borders satisfy the arm's length principle (ALP). This means leverage and the prices on intra-firm transactions correspond to what two independent entities would have agreed on. This is often difficult, since some goods especially intangibles have no obvious market parallel. Loans are also firm and project specific, and it can be a challenge to assess the terms a third party lender would be willing to enter into. The OECD identifies five factors that determine comparability: the functions performed by the parties, the contractual terms, the economic circumstances, the characteristics of the property or service transferred, and the business strategies pursued by the parties (OECD 2010). To carry through an evaluation based on these criteria is costly and very difficult. For such reasons some have proposed abandoning the principle of separate accounting (SA) that most countries rely on to determine profits, but rather to consolidate all profits into a worldwide singe measure and then apportion taxable income to countries based on activity weights of each firm.

There is much to be said about a transition to formula apportionment (FA), but it requires, among other things, political cohesion to agree on uniform apportionment weights. Nielsen *et al.* (2010) develop a theoretical model that compares basic properties of FA to SA. The focal point of their analysis is how changes in tax rates affect capital formation, input choice, and transfer pricing, as well as on spillovers on tax revenue in other countries. A significant difference between the two tax principles is that the SA system is based on reported income whilst taxation under the FA system is based on reported activity. They show that these fundamental characteristics introduce different tax spillovers across countries under the two tax systems, which makes it impossible to unambiguously favour one system over the other. Nielsen *et al.* (2010) find that the relative strength of tax spillovers under the two regimes depends on how costly it is for multinational enterprises to undertake transfer pricing, and how much pure profit the MNEs generate. These considerations also determine whether SA or FA implies the higher level of tax in a non-

cooperative equilibrium, and in the end which of the two schemes is preferable from an international perspective.

Under the SA system, the amount of interest that an affiliate of a multinational can deduct is determined by the rate of interest applied to its debt and the amount of its debt. Countries limit tax-induced income shifting via the transfer price by auditing a firm's transfer price to make sure that the interest rate is in line with what ALP. This involves considering the specific attributes of the company in determining the amount of debt that the company would be able to obtain from independent lenders. There are clear disadvantages to using the ALP on the interest rate, because it requires skills, resources, and specialization to establish what a third party would lend. For such reasons many countries rely on ratio approaches.

Under a ratio approach, also often referred to as a safe harbor rule or a thin capitalization rule, the amount of debt on which interest may be deducted for tax purposes is set by a pre-determined ratio. The exact definitions of the debt measure in the numerator of the ratio and of assets or equity in its denominator vary across countries, but the most common rule is either to use a ratio based on total debt-to-equity or internal (corporate group) debt-to-equity. The empirical literature on the effect of different types of thin capitalization rules on the firm's financial structure encompasses both US and European multinationals. It concludes that thin capitalization rules have a substantial effect on both internal and external leverage.[11]

Despite that empirical studies show that the ratio rules have an impact on multinationals' ability to shift profit by debt, there is a growing perception that these rules are not effective. A small but growing group of countries have therefore implemented what is called earnings stripping rules.[12] These rules operate to restrict interest deductions that ex-

[11] Buettner *et al.* (2012) study foreign affiliates of German multinationals whereas Blouin *et al.* (2014) investigate how thin capitalization rules worldwide affect the capital structure of foreign affiliates of US multinational firms. Both find that thin capitalization rules affect the leverage in multinationals.

[12] A handful of countries use both safe harbour rules and earnings stripping rules, either simultaneously or they impose a marginal earnings stripping requirement that applies only if the safe harbour limit is exceeded. The countries that fall into the first category are Denmark and Japan. Bulgaria, France and the US impose an earnings stripping rule only if the safe ratio rule is exceeded.

ceed a certain threshold, such as a percentage of EBITDA or EBIT.[13] Finland, Germany, Italy, Norway, Portugal, and Spain have implemented such rules. The Norwegian Tax Committee proposed in 2014 (NOU 2014:13) that both internal and external debt should be embedded in the earnings stripping rule and that EBIT was a better measure. They argued that external debt could be used to shift profit and that firms' leverage was too high due to the deductibility of interest. The latter implies too high risk premiums and too little investments.[14]

There is a literature that discusses which rule, ratio or earnings stripping, should be preferred if the aim is to maximize national income. In a recent paper this question is answered by Gresik *et al*. (2015b). They develop a general equilibrium framework with both labour and capital that allows them to analyze the variation in thin capitalization rules observed in practice. Their model embeds thin capitalization and transfer pricing behaviour of multinationals. They show that the policy that maximizes the host country's national income is an earnings stripping rule without a safe harbour rule.

5.6 FDI, tax havens and multinationals: A bane or a boon for a host country?

An interesting question is whether attracting foreign direct investment (FDI) from multinationals is a bane or a boon for a host country given the ability multinationals have to shift income. Multinationals often use tax haven conduit companies to shift income. Hines (2010) argues that although the tax avoidance opportunities presented by tax havens may reduce revenues in high-tax jurisdictions, they may have offsetting effects on FDI that are attractive to the same governments. If governments cannot distinguish between mobile and immobile investments, tax havens permit governments to subject immobile investments to higher taxes than mobile investments. Hong and Smart (2010) demon-

[13] EBITDA= Earnings Before Interest, Tax, and Depreciation Allowances. EBIT = Earnings Before Interest and Tax.

[14] See Sørensen (2014) for an analysis.

strate this effect in a model where multinationals shift profit by debt from a tax haven affiliate. They show that providing a tax deduction for interest payments on subsidiary debt allows host countries to maintain or even increase high business tax rates, and to attract more mobile investments from multinationals because the tax deductibility of interest reduces the firm's after-tax cost of capital. The end outcome is higher host welfare.

Slemrod and Wilson (2009) model tax havens that are "parasitic" on the tax revenues of non-haven countries in that they sell concealment services to taxpayers in non-havens. Non-haven countries must expend real resources to prevent tax base erosion. They show that tax havens increase the social costs that a country incurs when it increases its tax on capital. This aggravates the tax competition problem and results in lower welfare.

Gresik *et al*. (2015a) use the model by Hong and Smart but also include transfer pricing in the model. They allow the host country to decide on the corporate tax rate and thin capitalization rules (equity-debt-ratio) that may limit profit shifting by excessive interest deductions. In their model the multinational firm has a financing subsidiary located in a tax haven and an operational subsidiary in a high-tax country. The multinational can shift profit to the tax haven affiliate by the level of internal debt and the interest rate (transfer price) it charges on intra-company loans.

Which countries benefit or lose from attracting FDI depends in general on country characteristics. Developed countries have better institutional quality than emerging or developing countries in the sense that their tax systems make it more costly for multinationals to engage in aggressive tax-induced transfer pricing. They also have a low cost of capital, high rents for domestic entrepreneurs, and a moderate to high capital share in multinational production relative to emerging and developing countries. Gresik *et al*. (2015a) show that these differences matter. Developed countries can benefit from FDI and that a welfare maximum exists with an optimal corporate tax rate and a thin capitalization rule that are largely in line with average current tax rates and thin capitalization rules in the OECD. Developing countries, however, do not stand to benefit from policies that attract positive FDI. While per-

missive thin capitalization limits may be needed in developing countries to attract FDI, the amount of debt financing allowed by such permissive rules may facilitate aggressive transfer pricing that can result in lower welfare. The optimal tax policy for developing countries is to effectively eliminate the tax benefits of debt financing and only tax domestic firms.

Those who advocate the usefulness of tax havens as conduits lack a fully convincing explanation for why governments need to use tax havens to discriminate between mobile and immobile tax bases, rather than designing their tax systems to achieve this discrimination at lower costs.[15] Is the inability to conduct rational tax policy due to failures in the political system and/or lobbying, say? Schjelderup (2015) emphasizes the secrecy aspects of tax havens and argues that tax havens extend beyond just profit shifting activities and that to fully assess the welfare effects of them we need to fully assess the full range of their activities. Nevertheless, the lesson from a policy perspective from these studies is that the tax authorities must have resources and tools to secure tax compliance at their disposal. If not multinational investments may be a bane.

5.7 Some concluding comments

In this paper I discuss the challenges of taxing capital for small open economies. Although the corporate tax share of GDP in most countries is only around 3–4%, it is an important tax because it acts as a "backstop" for the personal tax rate. Wealthy individuals in particular would be given an incentive to reclassify their labour income as corporate income typically using offshore corporate structures to escape tax. The pressures of tax competition are exacerbated by tax planning and income shifting to low tax countries by multinationals. Studies show that multinationals pay less tax than domestic firms and this may give them a competitive edge over domestic firms. The long term effects may be changes in ownership structure that affect competition in markets and

[15] Wilson (2015) points this out in a survey of the literature on tax havens and tax competition.

make the corporate tax base more tax sensitive. Profit shifting is undertaken through transfer pricing and thin capitalization (excessive debt). Recent studies show that rules that restrict interest deductions that exceed a certain threshold such as a percentage of EBIT or EBITDA are the best defence against the use of debt to shift profit. Source taxes on royalty payments that are another effective defence mechanism. Yet establishing arm's length prices in transactions between affiliates of multinationals is a problem that will rid tax authorities also in the future.

I have not discussed taxes that fall on the capital stock such as property taxes, the wealth tax and the inheritance tax. Concerning the two latter taxes, the Nordic welfare states set themselves apart from larger countries. Table 1 provides and overview of the wealth and inheritance tax in the Nordic countries versus some large countries.

Table 1: Wealth and inheritance taxes

	Wealth Tax	Inheritance tax
Denmark	NO	15%–36.15%
Finland[16]	NO	10%
Iceland	NO	10%
Norway	YES	NO
Sweden	NO	NO
United States*	NO	40%
United Kingdom*	NO	40%
France*	YES	45%
Germany*	NO	30%

* The U.S. inheritance tax has an exemption of USD 5,430,000 in 2015. This is considerably larger than the exemptions in France (USD 105,945), Germany (USD 423,782), and the UK (USD 488,280).

Source: Center For Federal Tax Policy, 2015.

It is interesting to note that the Nordic countries seem to have gone further in terms of abolishing redistributive capital taxes than countries traditionally associated with polices much less tuned to redistribution. Aaberge and Atkinson (2010) shown how income inequality at the top of the distribution has increased in Anglo-Saxon countries, whereas the same rise in top income shares was not experienced by Continental Eu-

[16] Finland had a wealth tax until 2006 and a temporary wealth tax reintroduced in 2010, for four years.

ropean countries. They find that the Norwegian and Swedish experience over the twentieth century is similar to the Anglo-Saxon countries in that top shares, and the concentration among top incomes have first fallen and then risen. Norway differs from Sweden in that that the top shares rose more sharply in the period 1990–2006. Between 1980 and 2004, for example, the share of the top 1% more than doubled in Norway, but rose less than half in Sweden.

Several explanations have been put forward to explain why Norway sets itself apart. The implementation of the 1992 tax reform abolished the dividend tax and lead to a sharp increase in dividends and capital gains among the richest in Norway. Capital taxation in Sweden was less favourable. Substantial oil production in Norway started some 15 years before the rise in inequality, but could still be an explanatory factor due to constrained cash in this sector in the initial phase of production. Capital market reforms with liberalization of interest rates and an upturn in business cycles are also important factors that are hard to disentangle, but they certainly played a role.

Capital taxation also affects income mobility, and concerns about rising inequality have often been countered by constant changes in the composition of top income earners. If so, the rise in top incomes may not translate into "economic power". Aaberge *et al*. (2013) study who enters and leaves the top income groups in Norway in the period 1967–2011. Their main conclusion is that despite large changes in top income mobility over the last four decades, the magnitude of the effect of the changes in mobility on the income shares was moderate.

An interesting question is how voters will respond to rising inequality. Standard neoclassical theory predicts that inequality and the size of behavioural responses determines redistributive preferences (Meltzer & Richard, 1981). Following this literature one would expect voters to respond to lower capital taxes and increased inequality by demanding redistributive measures. A major concern, however, is that the public is misinformed about income inequality (Bartels, 2005; Slemrod, 2006). Such misconception may explain why there has been so little redistribution in the US, and that the political response to rising inequality in the US has been to further decrease capital taxes by reducing the top marginal tax rate from 75% in 1970 to 35% in 2012 (see IRS, 2014).

Gilens and Page (2014) study American politics and use data from 1981–2002 that has been collected with the purpose for estimating the influence upon public policy of poor citizens, "affluent" citizens, and those in the middle of the income distribution. A central message that emerges from their study is that: "...economic elites and organized groups representing business interests have substantial independent impacts on U.S. government policy, while mass-based interest groups and average citizens have little or no independent influence". Their study indicates that the majority does not rule, and that when the majority of citizens disagree with economic elites or their organized interest, they generally lose. In the Nordic countries, the economic elites and commercial interests have been strong drivers in reducing capital tax rates. In particular, this pertains to the abolishment of the wealth and the inheritance tax. But it is not clear in these specific cases what the preferences of the majority of voters were. A better example of conflict between the majority of voters and the economic elite is the increase in property taxes. In Norway, for example, "affluent voters" and organized interest from the business community have been advocates of a transition from the wealth tax to higher property taxes. The median voter, in contrast, is strongly against property taxation. Property taxes have certainly gone up through higher valuation of housing values, and the wealth tax has been reduced.

Top income earners in the Scandinavian countries mainly derive their income from capital. Competition among countries to attract mobile capital is a persistent phenomenon and will be a driver towards still lower taxes on mobile capital. A major change from the past, then, is less ability to redistribute, increasing income inequality, and rising immigration from poor countries. In sum these forces may affect trust between members of society. The level of trust is positively linked to economic growth.[17] Herein lies a major challenge for the Nordic welfare states

[17] For a survey of the literature on trust and economic growth see Christian Bjørnskov (2012).

References

Aaberge, R. & Atkinson, A.B. (2010). Top incomes in Norway, in Atkinson, A.B. and Piketty, T. (Eds.): *Top Incomes A Global Perspective*, Oxford University Press, Oxford, 448–482, SSB.

Aaberge, R,. Atkinson, A.B., & Modalsli, J. (2013). The ins and outs of top income mobility. *SSB Discussion paper* No. 762.

Alstadsæther, A. & Thoresen, T.O. (2009). Shifts in organizational form under a dual income tax system, *International Tax and Public Finance 66*(4), 384–418.

Alstadsæter, A., Jacob, M. & Michaely, R. (2015). Do dividend taxes affect corporate investment. *Journal of Public Economics* (forthcoming). http://dx.doi.org/10.1016/j.jpubeco.2015.05.001

Auerbach, Alan J., & Kevin A. Hassett (2002). On the marginal source of investment funds. *Journal of Public Economics*, 87, 205–232. http://dx.doi.org/10.1016/S0047-2727(01)00172-4

Balsvik, R., Jensen, S., Møen, J. & Tropina, J. (2009). *Kunnskapsstatus for hva økonomisk forskning har avdekket om flernasjonale selskapers internprising i Norge*. SNF Rapport 11/09.

Bartels, L. M. (2005). Homer gets a tax cut: Inequality and public policy in the American mind. *Perspectives on Politics*, 3 (01), 15–31. http://dx.doi.org/10.1017/s1537592705050036

Bartelsman, E.J. & Beetsma, R.M.W.J. (2003). Why pay more? Corporate tax avoidance through transfer-pricing in OECD Countries. *Journal of Public Economics 87*, 2225–2252. http://dx.doi.org/10.1016/S0047-2727(02)00018-X

Bernard, A.B., Jensen, J.B. & Schott, P.K. (2006). NBER Working Paper No. 12493.

Bjørnskov, C. (2012). How does social trust affect economic growth. *Southern Economic Journal 78*(4), 1346–1368. http://dx.doi.org/10.4284/0038-4038-78.4.1346

Biglaser, G. & C. Mezzetti, (1997). Politicians' decision making with re-election concerns. *Journal of Public Economics 66*, 425–447. http://dx.doi.org/10.1016/S0047-2727(97)00032-7

Blouin, J., Huizinga, H. P., Laeven, L. & Nicodeme, G. (2014). Thin Capitalization Rules and Multinational Firm Capital Structure. *CentER Discussion Paper 007*, Tilburg.

Buettner, T., Overesch, M., Schreiber, U. & Wamser, G. (2012). The impact of thin-capitalization rules on the capital structure of multinational firms, *Journal of Public Economics 96*, 930–938. http://dx.doi.org/10.1016/j.jpubeco.2012.06.008

Brennan, G. & Buchanan, J. (1980). *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge University Press.

Bucovetsky, S. (1991). Asymmetric tax competition. *Journal of Urban Economics* 30, 167–181. http://dx.doi.org/10.1016/0094-1190(91)90034-5

Bucovetsky, S. & Wilson, J. D. (1991). T*ax competition with two instruments, Regional Science and Urban Economics 96*, 329–341. http://dx.doi.org/10.1016/0166-0462(91)90062-r

Clausing, K. (2003). Tax-motivated transfer-pricing on Intra-firm Trade. *Journal of Public Economics 87*, 2207–2223. http://dx.doi.org/10.1016/S0047-2727(02)00015-4

Dharmapala, D. & Riedel, N. (2013). Earnings shocks and tax-motivated income-shifting: Evidence from European multinationals. *Journal of Public Economics 97*, 95–107. http://dx.doi.org/10.1016/j.jpubeco.2012.08.004

Eden, L. (1998). *Taxing Multinationals: Transfer Pricing and Corporate Income Taxation in North America*. University of Toronto Press, Toronto.

Edwards, J. & Keen, M. (1996). Tax Competition and Leviathan, *European Economic Review* 44, 113–34. http://dx.doi.org/10.1016/0014-2921(95)00057-7

Gilens, M. & Page B.I. (2014). Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens. *Perspectives on Politics*, Vol. 12(03), pp 564–581. http://dx.doi.org/10.1017/S1537592714001595

Gresik, T., Schindler, D. & Schjelderup, G. (2015a). *The Effect of Tax Havens on Host Country Welfare*. CESifo Working Paper Series 5314, CESifo Group Munich.

Gresik, T., Schindler, D. & Schjelderup, G. (2015b). *Immobilizing profit shifting: Is it safe to strip in the harbor*. Manuscript University of Notre Dame.

Haufler, A. & Wooton, I. (1999). Country size and tax competition for foreign direct investment. *Journal of Public Economics 71*, 121–139. http://dx.doi.org/10.1016/S0047-2727(98)00055-3

Hines, J.R. (2010). Treasure Islands. *Journal of Economic Perspectives* 24, 103–126. http://dx.doi.org/10.1257/jep.24.4.103

Hong, Q. & Smart, M. (2010). In praise of tax havens: International tax planning and foreign direct investment. *European Economic Review 54*:82–95. http://dx.doi.org/10.1016/j.euroecorev.2009.06.006

Huizinga, H. & Nielsen, S.B. (1997). Capital income and profit taxation with foreign ownership of firms. *Journal of International Economics*, 1997, 88–106. http://dx.doi.org/10.1016/s0022-1996(96)01449-3

IRS. (2014). *SOI Tax Stats – Historical Data Tables, Table 23*. See http://www.irs.gov/uac/SOI-Tax-Stats-Historical-Data-Tables

Janeba, E. & Schjelderup, G. (2009). The Welfare Effects of Tax Competition Reconsidered. Politicians and Political Institutions. *Economic Journal* 119, 1–19. http://dx.doi.org/10.1111/j.1468-0297.2009.02263.x

Keen, M. & Marchand, M. (1997). Fiscal competition and the pattern of public spending. *Journal of Public Economics 63*, 33–53. http://dx.doi.org/10.1016/S0047-2727(97)00035-2

Kind, H.J., Midelfart Knarvik, K.H., & Schjelderup, G. (2000). Competing for capital in a lumpy world. *Journal of Public Economics 78*, 253–274. http://dx.doi.org/10.1016/S0047-2727(99)00119-X

Klemm, A., Haufler, A. & Schjelderup, G. (2009). Economic integration and the relationship between profit and wage taxes. *Public Choice 138*, 423–446. http://dx.doi.org/10.1007/s11127-008-9368-9

Langli J.C. & Saudagaran, S.M. (2004). Taxable Income Differences Between Foreign and Domestic Controlled Corporations in Norway. *European Accounting Review 13*(4), 713–41. http://dx.doi.org/10.1080/0963818042000237115

Meltzer, A. H. & Richard, S. F. (1981). A Rational Theory of the Size of Government. *The Journal of Political Economy 89* (5), 914–927. http://dx.doi.org/10.1086/261013

Nielsen, S.B., Raimondos-Møller, P., & Schjelderup, G. (2010). Company taxation and tax spillovers: Separate Accounting versus Formula Apportionment. *European Economic Review 54*, 221–232. http://dx.doi.org/10.1016/j.euroecorev.2009.06.005

NOU (2014). 13. *Kapitalbeskatning i en internasjonal økonomi. Norges Offentlige Utredninger*. Oslo.

OECD (2010). *OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations 2010*. OECD Publishing, Paris.

OECD (2013). *Action Plan on Base Erosion and Profit Shifting*. OECD Publishing, Paris.

Oyelere, P. B. & Emmanuel, C.R. (1998). International Transfer Pricing and Income Shifting: Evidence from the UK. *European Accounting Review 7*, 623–635. http://dx.doi.org/10.1080/096381898336222

Pak, S.J. & Zdanowicz, J.S., (2001). *Executive Summary of "US Trade with the World: An Estimate of 2000 Lost US Federal Income Tax Revenues due to Over-Invoiced Imports and Under-Invoiced Exports*". Mimeo, Florida International University.

Persson, T. & G. Tabellini, (1992). The politics of 1992: Fiscal policy and European integration. *Review of Economic Studies 59*, 689–701. http://dx.doi.org/10.2307/2297993

Rauscher, M. (1998). Leviathan and Competition Among Jurisdictions: The Case of benefit taxation. *Journal of Urban Economics 44*, 59–67. http://dx.doi.org/10.1006/juec.1997.2060

Schjelderup, G. (2015). *Secrecy Jurisdictions, International Tax and Public Finance*, forthcoming.

Serrato, J. C. S. & Zidar, O. (2014). *Who benefits from state corporate tax cuts? A local labor approach with heterogeneous firms*. NBER Working paper 20289.

Slemrod, J. (2006). The role of misconceptions in support for regressive tax reform. *National Tax Journal*, 57–75. http://dx.doi.org/10.17310/ntj.2006.1.03

Slemrod, J. & Wilson, J.D. (2009). Tax Competition with Parasitic Tax Havens. *Journal of Public Economics* 93, 1261–70. http://dx.doi.org/10.1016/j.jpubeco.2009.08.004

Sørensen, P. B. (2014). *Taxation and the Optimal Constraint on Corporate Debt Finance*. ESifo Working Paper No. 5101.

Weichenrieder, A.J. (2009). Profit shifting in the EU: Evidence from Germany. *International Tax and Public Finance 16*, 281–297. http://dx.doi.org/10.1007/s10797-008-9068-x

Wildasin, D.E. (1989). Nash equilibria in models of fiscal competition. *Journal of Public Economics* 35, 299–240.

Wildasin, D.E. (1991). Some rudimentary "duopolity" theory. *Regional Science and Urban Economics 21*, 193–212. http://dx.doi.org/10.1016/S0047-2727(03)00057-4

Wildasin, D.E. & Wilson, J.D. (2004). Tax competition: bane or boon? *Journal of Public Economics 88*, 1065–1091. http://dx.doi.org/10.1016/S0047-2727(03)00057-4

Wilson, J.D. (1986). A theory of interregional tax competition, *Journal of Urban Economics* 19, 296–315. http://dx.doi.org/10.1016/0094-1190(86)90045-8

Wilson, J.D. (1999). Theories of tax competition, *National Tax Journal 52*, 269–304.

Wilson, J.D. (1991). Tax competition with interregional differences in factor endowments. *Regional Science and Urban Economics 21*, 423–451. http://dx.doi.org/10.1016/0166-0462(91)90066-V

Wilson J.D. (2015). Tax havens in a world of competing countries. CESifo Dice Report 12 (4). Zodrow, G. and Mieszkowski, P. (1986). Pigou, Tiebout, property taxation and the underprovision of local public goods. *Journal of Urban Economics 19*, 356–370.

Yagan, D. (2001). *Capital tax reform and the real economy: The effects of the 2003 dividend tax cut*. NBER working paper 21003.

6. Nordic family policy and maternal employment

Julian V. Johnsen, Department of Economics, University of Bergen, email: julian.johnsen@uib.no and Katrine V. Løken, Department of Economics, University of Bergen, email: katrine.loken@uib.no

Abstract

We survey the quasi-experimental literature to evaluate the effectiveness of family policies in terms of increasing maternal employment. The findings are mixed but mostly follow the theoretical predictions, although the effects are smaller than expected. While parental leave policies do not affect maternal employment, cash-for-care benefits decreases maternal employment in the short and long run. Subsidized child care can increase maternal employment, especially for mothers of younger children. When evaluating the overall effect of the "package" of Nordic family policies using cross-country evidence, we find that the remaining gender gaps in the labor market are concentrated at the top of the career ladder. To increase the share of female managers and senior officials, the current system of family policies should be made more flexible.

6.1 Introduction

Because women tend to be the main caregivers for children, having children is associated with an increase in the gender gap in the labor market. The wage gap between women and men with children is much larger than the wage gap between women and men without children (Bertrand *et al*., 2010). In Sweden, the within-couple wage gap between husband and wife more than doubles after the couple becomes parents (Angelov *et al*., 2013). A recent paper provides causal evidence that in Denmark having a child has a large and long-lasting negative effect on female labor supply, especially in terms of employment rates (Lundborg *et al*., 2014). Economists explain this pattern with comparative advantages leading to specialization in the household (Becker, 1981). If women have a comparative advantage in terms of child care, having children increases the intra-household specialization, leading to less gender equality in the labor market.

In general, family policies aim to aid the reconciliation of work and family life, making it easier for parents to work (increasing the maternal labor supply) and for workers to become parents (increasing fertility).[1] This paper focuses on the former, the impact of family policies on maternal labor supply. This is not to say that the potential fertility effects of family policies are unimportant. The effects are mainly left out because of the limited scope of the paper and for the simple reason that there is little hard evidence to review on the fertility effects of family policies.[2]

Boosting female employment and gender equality has been a key goal of Nordic family policies (Björnberg, 2013). Several comparative studies have shown that the Nordic countries share key similarities in family policies and that the Nordic approach to family policies differs from that of other European countries (e.g., Ellingsæter, 2006; Bradshaw and Hatland, 2007). Two key policies that the Nordic countries

[1] Another important goal of family policies, which is not discussed here, is to promote child welfare and development.

[2] In theory, most family policies involve a transfer of resources from people without children to people with children and thus could increase fertility by increasing the incentives to have children.

share are a long and generous paid parental leave and high coverage rates of subsidized pre-school child care.

To discuss the potential effects of Nordic family policies on maternal employment, it is important to distinguish between policies that promote home care versus policies that promote the use of formal child care. Parental leave programs and cash-for-care benefits promote home care. In theory, home care leads to increased intra-household specialization and thus decreases maternal employment.[3] Promoting the use of formal child care through subsidies should work in the other direction, potentially increasing maternal employment by decreasing intra-household specialization.

In this paper, we have conducted a partial review of the effectiveness of the main Nordic family policies in increasing gender equality in the labor market. At first glance, the "package" of family policies seems to have been effective: Many of the family policies were implemented or expanded in the 1970s, which is the same period that the Nordic countries saw the largest increase in female employment rates. However, there is a basic "chicken and the egg" problem with this interpretation: Do Nordic countries have high gender equality in the labor market because of their generous family policies, or has the high share of working women resulted in increased political demand for generous family policies?[4] This question cannot be easily answered. In addition, the different policies should in theory have opposite effects on maternal employment depending on whether the policies promote home care or the use of formal child care. To assess the causal effect of Nordic family policies, the literature review focuses on studies that used quasi-experimental designs. These designs used reforms and other "natural" variations to identify the causal effects of one variable on another. The benefit of quasi-experimental studies is that they deliver clear-cut results on the immediate effects of the policy reforms studied.

---

[3] Most parental leave programs include job protection, allowing mothers to return to their previous job after parental leave. Shorter parental leaves with job protection could increase maternal employment by providing job stability and the legal right to continuous employment.

[4] Even if increased female labor force participation caused demand for family policies (instead of generous family policies driving female labor force participation), figuring out the causal effect of the demanded family policy would remain an important policy question.

The drawback of quasi-experimental studies is that they mainly provide insights into the immediate individual effects of policy reforms. Some policies might work in the long run by changing norms related to family and work. Family policies could also have important general equilibrium effects that are hard to study in a quasi-experimental setting. Policies promoting home care incentivize women to take longer career breaks after child care, while promoting child care works in the other direction. Because employers do not know which women will have children, longer parental leaves create a disincentive to hire women if promoters dislike long career breaks. Promoting the use of formal child care or reserving part of parental leave for the exclusive use of fathers alleviates these disincentives for employers. These potential general equilibrium effects cannot easily be tested in a causal framework, and we are reduced to speculation based on cross-country evidence. In general, countries with more generous family policies have higher maternal employment but fewer female managers and senior officials. This has led to speculation that Nordic family policies have created a glass ceiling, incentivizing women to enter the labor market but stopping them from climbing the career ladder (Datta Gupta *et al*., 2008).

The rest of the paper proceeds as follows. Section 1 presents the details of Nordic family policies, first for parental leave and cash-for-care and second for subsidized child care. Section 2 provides a partial review of the literature on the causal effect of Nordic family policies on maternal employment, distinguishing between policies that promote home care and policies that promote child care. Section 3 summarizes the conclusions from the literature review and provides suggestions for how future policies can close the remaining gender gaps in the Nordic labor markets. This section also draws on cross-country evidence to fill the gaps unanswered by the quasi-experimental evidence.

6.2 Nordic family policies

This section presents details on the Nordic family policies of parental leave, cash-for-care, and subsidized child care. The Nordic approach to family policies, combining generous parental leave with extensive provision of subsidized child care, is expensive. The Nordic countries outspend all other countries when it comes to parental leave and child care support. These expenditures contribute to the high tax pressure in the Nordic countries. According to the Organization for Economic Co-operation and Development (OECD), in 2011 the total tax revenue as a percentage of the gross domestic product (GDP) was 47.7% in Denmark and 44.2% in Sweden, which makes them the countries with the highest tax pressure in the world. Finland follows with 43.7%, then Norway with 42.5%, and finally Iceland with 36.0%. The unweighted average for the OECD countries is 34.1%.

6.2.1 Parental leave programs and cash-for-care

Parental leave policies provide parents with a job leave of a certain length, in order for them to take care of infant children. Parental leave can be assigned to the mother (maternity leave) or the father (paternity leave) or be available to either parent (parental leave). Often, these policies include a form of job protection, in that the parent on leave is entitled to return to her previous job after her leave expires. Parental leave can either be paid or unpaid. Paid parental leave provides either a certain percentage of income replacement or a universal flat rate of benefits. Parental leave policies of some form exist in all OECD countries.

In addition to parental leave, several Nordic countries have implemented cash-for-care programs. This policy is aimed at smoothing the transition between parental leave and formal child care, giving families the opportunity to stay home after the parental leave expires. Cash-for-care policies are in reality extensions of parental leave, although with a fixed benefit for staying at home instead of replacing a certain percentage of income. This makes cash-for-care benefits more beneficial for parents with lower incomes. The take-up of these programs is relatively low, except for Finland. Column 1 of Table 1 shows the total weeks of

paid leave available to mothers in Nordic countries and selected comparison countries. The numbers include maternity leave, parental leave, and total weeks of eligibility for cash-for-care benefits. The numbers are scaled by the replacement rates, meaning that the numbers show the number of full-rate equivalent weeks available to mothers.

The take-up of parental leave and cash-for-care benefits is gendered, with many countries providing mostly maternity leave. In general, replacing maternity leave with paternity leave (available to either parent) does not change the gendered take-up (Datta Gupta *et al.*, 2008). However, the introduction of paternity leave quotas has induced more men to stay at home during the child's first year. The Nordic countries, with the exception of Denmark, have been the forerunners in implementing paternity leave quota policies and have paternity leave quotas larger than the OECD average. The number of full-rate equivalent weeks of leave available to fathers in the Nordic countries and selected comparison countries can be seen in column 2 of Table 1.

The combination of long parental leave periods and high replacement rates means that the Nordic countries spend more on parental leave per child than the average OECD country, as can be seen in column 3 of Table 1. The Nordic system of paid parental leave includes job protection. Although there are some small differences in the length and replacement rates across the Nordic countries, most countries provide more generous programs than the OECD average. Evidence from one country is therefore likely to be externally valid for all Nordic countries.

6.2.2 Subsidized child care

The Nordic countries, with the exception of Finland, have high formal child care enrollment rates. For 3- to 5-year-olds, Denmark, Iceland, Norway, and Sweden have close to 100% enrollment rates (Finland has 74%) compared to the OECD average of 82%, as seen in column 5 of Table 1. Finland's lower enrollment rates are probably driven by the cash-for-care program that can be used until the child is 3 years old. The enrollment rates for 0- to 2-year-olds are lower in the Nordic countries, largely because of the long parental leave. Still, the Danish, Icelandic, Norwegian, and Swedish enrollment rates for 0- to 2 year olds are high-

er than the OECD average, as shown in column 4 of Table 1. The Nordic child care model is based on day care, meaning that children are cared for only during standard weekday work hours. As a result, publicly provided child care is typically inflexible and does not cater to child care needs at irregular hours.

The Nordic countries have high enrollment rates in formal child care because of extensive subsidies. In terms of public expenditure per child on child care and pre-primary education, the Nordic countries far outspend the OECD average and the comparison countries, as seen in column 6 of Table 1. Even though Finland has lower child care enrollment rates compared to the other Nordic countries, Finland still spend similar amounts on child care support. This is because the measure of child care support includes cash-for-care benefits, which in Finland are generous and can be claimed up until the child is 3 years old.[5]

6.3 The causal effects of Nordic family policy on gender equality

In this section, we review studies on the effects of paid parental leave, cash-for-care programs, and subsidized child care on maternal employment. The review is based on selected studies using Nordic data and quasi-experimental methods aiming at identifying causal effects.

6.3.1 Effects of paid parental leave and cash-for-care benefits

Even in the absence of a formal right to parental leave and job protection, mothers are likely to take leaves around childbirth and the first weeks or months after. The existence of a certain amount of parental leave combined with job protection could in theory increase maternal employment, because the leave smoothens the career break of the mother around

[5] With the exception of Finland, when evaluating the effects of subsidized child care on the gender gap, the other Nordic countries have very similar systems, and the external validity of single country studies are likely to be high for the other Nordic countries.

childbirth and allows her to return to her previous employment. Several studies show this to be the case: Job protection can increase the job continuity and labor force participation of women after childbirth (e.g., Waldfogel, 1998). The right to maternity leave and legal protection against job loss ensure that women can return to employment after childbirth and enhance gender equality in the labor market.[6]

However, policies that promote longer periods of home-based care can in theory have a negative impact on maternal labor supply. The longer the parental leave, the longer the mother's job disruption. In addition, longer leave periods might increase the degree of intra-household specialization, with the mother's household work increasing. The Nordic countries provide long periods of job-protected parental leave, and because the replacement rates are high, most women are incentivized to use the entire length of the leave. If mothers have long periods outside the labor market, this can negatively impact their careers in the long run.

Despite these theoretical predictions, the empirical results show that the length of parental leave has no causal effect on maternal labor supply. Carneiro *et al*. (2015) and Dahl *et al*. (2015) show that the different reforms that increased parental leave in Norway had no effect on long-run maternal labor supply. Similarly, Liu and Skans (2010) find no effect on mothers' earnings after an increase in parental leave in Sweden. These results are confirmed in studies from non-Nordic countries, such as Austria (e.g., Lalive *et al*., 2014) and Germany (Schönberg and Ludsteck, 2014). It seems as if you already participate in the labor market you are not hurt by taking a long leave. This could be because the take-up rates of parental leave are close to 100% for Nordic mothers, a sign that there is little stigma against women taking long parental leaves.

Similar to parental leave, cash-for-care benefits are a family policy promoting the use of home care. Because cash-for-care benefits last for a much longer period than parental leave (in Finland, up until the child is 3 years old), the potential negative effect on maternal employment is

[6] In theory, a minimum amount of parental leave with job protection could increase fertility, because working women can have children without fear of losing their employment. We do not know of any studies that test this prediction with a causal design.

larger. In addition, because the take-up rates are much lower than for parental leave, employers could discriminate against mothers who take up cash-for-care benefits. For these policies, empirical studies have confirmed the negative theoretical effects. Cash-for-care delay mothers' reentry in the labor market (e.g., Schøne, 2004, for Norway and Kosonen, 2014, for Finland). In addition, Drange and Rege (2013) find that mothers have lower labor market earnings in the years after the cash-for-care program runs out. However, when the child starts school at age 6, the effects are no longer present. The reason is that cash-for-care programs do not make mothers exit the labor market altogether. The negative earnings effect in the medium run is driven by affected mothers working part time instead of full time. When looking at sub-groups, the negative effect of cash-for-care policies on maternal employment is much larger for non-Western immigrant mothers (Naz, 2010). In summary, the literature consistently shows that cash-for-care programs have a clear negative effect on the gender gap in the labor market.

A potentially important effect of policies that promote home-based care is related to incentives for employers. If employers dislike longer job leave periods, then generous parental leave policies can affect the employment opportunities of not just mothers but also all women of child-bearing age. If employers cannot know up front which women are planning to have children, longer parental leaves create incentives against hiring all women of childbearing age.[7] This disincentive could decrease women's chances of landing a first job or being promoted, especially in certain occupations. Therefore, although long periods of maternity leave can enable women to combine work and family life, long leaves might harm women's labor market opportunities in the long run, by reinforcing the incentives for women to specialize in the household sector. Such general equilibrium effects are potentially important, but it is difficult to verify their existence empirically.

A theoretical solution to the problem of parental leave creating disincentives to hire women is to reserve parts of parental leave for fa-

[7] This is an example of statistical discrimination, where employers discriminate against women of childbearing age because they more often have longer periods away from work. Statistical discrimination is normally illegal, but in reality, such laws are hard to enforce.

thers. If parental leave is split into two equal quotas, one for the exclusive use of the mother and one for the exclusive use of the father, then there would be no disincentive effect for employers to hire women. In addition, paternity leave quotas can potentially affect fatherhood norms and increase fathers' long-term involvement in the child's upbringing. Both theoretical predictions are unfortunately difficult to test in a causal framework.

In contrast to many other countries, the Nordic paternity leave quotas have been effective in inducing men to take parental leave. The full effect of the policy takes time to accumulate and is reinforced by peer effects (Dahl *et al*., 2014), perhaps suggestive that paternity leave quotas have slowly changed underlying gender norms. Despite the success of paternity leave quotas in inducing men to provide home care, there is scant causal evidence that paternity leave quotas have long-run impacts on the employment of mothers and fathers. Paternity leave quotas do not affect intra-household specialization and gender gaps in wages and employment in Sweden (Ekberg *et al*., 2013). In Norway, the findings are more mixed, with one study confirming the Swedish results (Cools *et al*., 2015), and another finding that paternity leave decreased fathers' long-run earnings, possibly by increasing fathers' long-term involvement in the child's upbringing (Rege and Solli, 2013).[8]

The general finding that paternity leave quotas had little effect on gender equality in the labor market is perhaps not surprising in light of the literature that has found no effect of extending the length of parental leave for mothers: If there is no immediate harm to women's employment by extending parental leave, then there are no short-run harmful employment effects to alleviate with paternity leave quotas. However, paternity leave quotas could have general equilibrium effects by decreasing the incentives for employers to statistically discriminate against women of childbearing age. This effect is theoretically reasonable but hard to measure in a causal framework.

[8] The study measures long-run earnings up until the child is 5 years old.

6.3.2 Effects of subsidized child care

Without affordable child care, some (predominantly) women face a choice between having children and taking care of them instead of working, or working and not having children. In theory, subsidizing child care could therefore increase fertility (of women who had previously decided to work and remain childless), or the maternal labor supply (of mothers who previously stayed at home with her children), or both. On the individual level, subsidizing the use of formal child care should in theory increase maternal labor supply, by making it more attractive for mothers to work when their children are pre-school age. Additionally, the disincentives for employers to hire women decrease because increased use of child care instead of home care will shorten the length of mothers' career interruptions after childbirth.

The empirical findings on the effect of subsidized child care on maternal employment indicate that the institutional setting plays an important role. An often-cited study by Havnes and Mogstad (2011) finds that the expansion of formal child care in Norway in the 1970s did not increase maternal employment. Instead, the expansion led to a crowding out of the use of informal child care,[9] suggesting a significant net cost of the child care subsidies. However, there are important differences between the institutional setting of the 1970s and today. Most importantly, there has been a reduction in the availability of informal child care and an increase in child care slots for the youngest children. A more recent study by Andresen and Havnes (2015) finds that an expansion of subsidized child care for 1 to 2 year olds during the 2000s in Norway increased the employment of the affected mothers. The findings from studies on non-Nordic countries are mixed, although many studies have shown positive effects (e.g., Bauernschuster and Schlotter, 2015).

Regarding the cost of formal child care, the findings are also mixed. Black *et al*. (2014) find that reduced child care costs in Norway did not increase the labor supply of affected mothers. However, the study involved only the mothers of 5 year olds due to data limitations. Similarly,

[9] Informal child care arrangements are often performed by grandparents but can also involve the use of nannies or babysitters, although these latter arrangements are uncommon in Nordic countries today.

Lundin *et al*. (2008) find no effect of reduced child care costs on the maternal labor supply in Sweden. However, Simonsen (2010) shows that in Denmark increases in the cost of child care have a significant negative effect on the maternal labor supply. However, the effect is only in the short run during the first 12 months after childbirth. Denmark differs from the other Nordic countries in that the take-up of child care during the first 12 months is significantly higher. This again shows that the institutional setting seems to matter for the effectiveness of subsidizing child care in increasing the maternal labor supply.

6.4 Summary and suggestions for future policy[10]

The Nordic countries have been at the forefront when it comes to implementing family policies that are often seen as enabling the reconciliation of work and family life (Datta Gupta *et al*., 2008). The combination of working and having children is relatively common in the Nordic countries. Family policies could be important for keeping fertility levels up, and the Nordic countries have avoided the dramatic fall in fertility rates that has occurred in Southern Europe during the last few decades.[11]

A clear goal of the Nordic family policies has been to promote gender equality in the labor market. However, theoretically, policies that promote the use of home care could decrease maternal employment while policies that promote the use of child care could increase maternal employment. The empirical literature tends to confirm the theoreti-

[10] Our policy suggestions are aimed at increasing gender equality through increasing maternal employment. As mentioned, family policies also have other important goals such as increasing fertility and promoting child welfare and development. The effect of family policies on these outcomes, not discussed here, should also be considered for policy reforms.

[11] Whether the higher fertility rates of the Nordic countries are caused by their family policies is hard to answer. An ideal experiment would be to compare the fertility choices of women who live in an institutional setting with no parental leave to the fertility choices of women who live in an institutional setting with generous parental leave. This can be done only by comparing women who live in different countries or in different time periods, making it nearly impossible to separate the effect of generous parental leave from other variations across time periods and countries. A counter-argument to family policies being the main driver of high fertility in the Nordic countries is that the fertility rate in the US, a setting without generous family policies, is as high as in the Nordic countries. More research is needed to understand the role of family policies for fertility decisions.

cal predictions, although the estimated effects are perhaps smaller than expected. The findings from the literature focusing on the causal effects of pro-home-care policies can be summarized as follows: First, job protection around childbirth promotes job stability and maternal employment. Second, expanding the length of maternal parental leave has not decreased the long-term maternal labor supply. Third, the introduction of paternity leave quotas has induced men to take parental leave, but there is little causal evidence that paternity leave has changed intra-household specialization. Finally, cash-for-care policies are harmful to gender equality in the labor market. In summary, dropping cash-for-care programs seems the obvious method for changing home care policies to promote more gender equality in the labor market.

The effectiveness of subsidized child care in increasing maternal employment depends on the institutional setting. Policies that promote the use of child care can be effective when little informal child care is available. This is especially the case when child care is provided to groups who previously had little access to formal and informal child care. For Norway and Sweden, at least, mothers of 1 to 2 year olds are "caught in between" the parental leave that expires after approximately 1 year and the fact many children do not get a spot in formal child care when they turn 1. Expanding child care to these mothers seems to be effective in increasing their labor supply.

Moving on from the causal evidence available, when looking at the cross-country evidence, the Nordic countries in general do well in terms of gender equality in the labor market. Since the 1970s and 1980s, the employment rates for Nordic women have approached those of men, and the Nordic countries are close to full gender equality when it comes to the labor force participation rates of men and women. The employment rate for women with young children is far higher in the Nordic countries compared to the OECD average, as shown in columns 7 and 8 of Table 1.

Another interesting point to note from columns 7 and 8 of Table 1 is that Finland has low employment rates for mothers of children aged less than 3, a pattern that fits the consistent causal evidence that cash-for-care benefits decrease maternal employment. Dropping the cash-for-care policies in favor of earlier access to subsidized child care will

increase gender equality in the labor market. Finland has most to gain from this policy since this country has the most generous cash-for-care program, least generous subsidized child care, and lowest employment rates for mothers with children younger than 3 years old. The other Nordic countries have less to gain from changing current family policies for the overall gender employment gap; however, the employment rates for mothers with 1 to 2 year olds could likely be slightly increased by dropping cash-for-care policies in favor of earlier access to subsidized child care.[12]

The remaining gender gap in the Nordic countries seems to be concentrated at the top: In terms of the gender gap among senior managers and officials, the World Economic Forum (2013) ranked Denmark 72nd, Finland 68th, Iceland 22nd, Norway 58th, and Sweden 44th. Arulampalam *et al*. (2007) show that countries with more generous family policies have larger gender wage gaps at the higher end of the income distribution. This has led to the theory that generous family policies might create a glass ceiling, meaning that very few women enter management jobs. Albrecht *et al*. (2015) argue that what is different about the Nordic countries are the constraints that women face. The authors show relationships in Sweden between the glass ceiling and the take-up of parental leave that are consistent with generous parental leave being an explanation. Datta Gupta *et al*. (2008) argue that longer job leaves are more harmful to the careers of highly educated women, and women in the private sector. However, it is very hard to distinguish the separate effect of generous parental leave from other potential explanations.[13]

To close the remaining gender gap at the top of the career ladder, we argue that the focus should be on making the existing family policies more flexible. Public child care typically offers care only during regular work hours, which is probably insufficient to cater to the needs of

[12] Take-up of cash-for-care programs is higher among non-Western immigrants, and the negative effect on maternal employment is higher for this group (Naz, 2010). Dropping cash-for-care programs in favor of increased access to subsidized child care could therefore also be seen as a tool to promote social integration of non-Western immigrant women.

[13] Other potential reasons for the glass ceiling effects could be statistical discrimination, gender differences in career orientation, or a limited market for household services (that women in top jobs may need to succeed) because of the compressed wage structure in the Nordic countries.

women in jobs that demand regular overtime. The statistics seem to indicate that this is a potential problem: Women in Nordic countries select part-time work and work less overtime, fewer long hours, and fewer evenings compared to men (Plantenga and Remery, 2010). There is also suggestive evidence that providing more flexible child care arrangements can boost the full-time labor supply of women. Johnsen (2015) finds that grandparents' retirement can increase adult daughters' labor supply through providing child care. If grandparents' child care can increase maternal labor supply in Norway, a country with high coverage rates of formal child care, the effect is likely to be driven by child care during irregular hours. To achieve more flexibility in the system, potential policy reforms include admitting children to formal child care twice a year instead of once a year, and increasing the options for women to substitute parts of the parental leave for subsidized child care. Another option is to subsidize informal child care at irregular hours (nannies at home, pick-up-services at day care), or other domestic services. Hallden and Stenberg (2013) evaluate a policy in Sweden that gives a 50% tax deduction on domestic services. They find that the policy increased the labor supply of married women. The authors cannot isolate the effect on career-oriented women; however, if anything, the increase in labor supply was similar for women of various socioeconomic backgrounds. Thus, this type of policy might help reduce the overall gender gap in the labor market (for example, pushing mothers from part-time to full-time work), as well as increase the share of female managers and top earners.

References

Albrecht, J., Skogman Thoursie, P., & Vroman, S. (2015). Parental leave and the glass ceiling in Sweden. *Research in Labor Economics*, 41, 89–114.

Andresen, M.E., & Havnes, T. (2014). *Women and children first? Labor market effects of universal child care for toddlers*. Working Paper.

Angelov, N., Johansson, P., & Lindahl, E. (2013). *Is the persistent gender gap in income and wages due to unequal family responsibilities?* IZA Discussion Papers 7181.

Arulampalam, W., Booth, A.L., & Bryan, M.L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review*, 60, 163–186.

Bauernschuster, S., & Schlotter, M. (2015). Public child care and mothers' labor supply – Evidence from two quasi-experiments. *Journal of Public Economics*, 123, 1–16.

Becker, G. (1981). *A Treatise on the Family*, Harvard University Press, Cambridge, Mass.

Bertrand, M., Goldin, C., & Katz, L.F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2, 228–255.

Björnberg, U. (2013). Introduction. In U. Björnberg and M.H. Ottosen (eds.), *Challenges for Future Family Policies in the Nordic Countries*, SFI, Copenhagen.

Black, S.E., Devereux, P.J., Løken, K.V., & Salvanes, K.G. (2014). Care or cash? The effect of child care subsidies on student performance. *Review of Economics and Statistics*, 96, 824–837.

Bradshaw, J., & Hatland A. (2007). *Social Policy, Employment and Family Change in Comparative Perspective*, Edgar Elgar, London.

Carneiro, P., Løken, K. V., & Salvanes, K. (2015). A flying start? Maternity leave and long run outcomes of children. *Journal of Political Economy*, 123, 365–412.

Cools, S., Fiva, J.H., & Kirkebøen, L.J. (2015). Causal effects of paternity leave on children and parents. *Scandinavian Journal of Economics*, 117, 801–828.

Dahl, G., Løken, K.V., & Mogstad, M. (2014). Peer effects in program participation. *American Economic Review*, 104, 2049–2074. http://www.nber.org/papers/w18198.pdf?new_window=1

Dahl, G., Løken, K.V., Mogstad, M., & Salvanes, K. (2015). What is the case for paid maternity leave? *Review of Economics and Statistics*, forthcoming.

Datta Gupta, N., Smith, N., & Verner, M. (2008). The impact of Nordic countries' family friendly policies on employment, wages, and children. *Review of Economics of the Household*, 6, 65–89.

Drange, N., & Rege, M. (2013). Trapped at home: The effects of mothers' temporary labor market exits on their subsequent work career, *Labour Economics*, 24, 125–136.

Ekberg, J., Eriksson, R., & Friebel, G. (2013). Parental leave – A policy evaluation of the Swedish "Daddy-Month" reform. *Journal of Public Economics*, 97, 131–143.

Ellingsæter, A.L. (2006). The Norwegian childcare regime and its paradoxes, in A.L. Ellingsæter and A. Leira (eds.), *Politicising Parenthood in Scandinavia. Gender Relations in Welfare States*, The Policy Press, Bristol.

Halldén, K., & Stenberg, A. (2013). *The relationship between hours of domestic services and female earnings: Panel register data evidence from a reform*. IZA Discussion Papers 7649.

Havnes, T., & Mogstad, M. (2011). Money for nothing? Universal child care and maternal Employment. *Journal of Public Economics*, 95, 1455–1465.

Johnsen, J.V. (2015). *Work and Families: Four Essays in Applied Microeconomics*. University of Bergen, Bergen.

Kosonen, T. (2014). To work or not to work? The effect of childcare subsidies on the labour supply of parents. *BE Journal of Economic Analysis and & Policy*, 14, 817–848.

Lalive, R., Schlosser, A., Steinhauer, A., & Zweimüller, J. (2014). Parental leave and mothers' careers: The relative importance of job protection and cash benefits. *Review of Economic Studies*, 81, 219–265.

Liu, Q., & Skans, O.N. (2010). The duration of paid parental leave and children's scholastic performance, *BE Journal of Economic Analysis & Policy*, 10.

Lundborg, P., Plug, E., & Rasmussen, A.W. (2014). *Fertility effects on female labor supply: IV evidence from IVF treatments*. Working Paper.

Lundin, D., Mörk, E., & Öckert, B. (2008). How far can reduced childcare prices push female labour supply? *Labour Economics*, 15, 647–659.

Naz, G. (2010). Effect of a family policy reform on immigrants' labor supply and earnings. *Labour*, 24, 74–92.

Nielsen, H.S., Simonsen, M. & Verner, M. (2004). Does the gap in family friendly policies drive the family gap? *Scandinavian Journal of Economics,* 106, 721–744.

Plantenga, F., & Remery, C. (2010). *Flexible Working Time Arrangements and Gender Equality*. Office for Official Publications of the European Communities, Luxembourg.

Rege, M., & Solli, I.F. (2013). The impact of paternity leave on fathers' future earnings. *Demography*, 50, 2255–2277.

Schönberg, U., & Ludsteck, J. (2014). Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. *Journal of Labor Economics*, 32, 469–505.

Schøne, P. (2004). Labour supply effects of a cash-for-care subsidy. *Journal of Population Economics*, 17, 703–727.

Simonsen, M. (2010). Price of high-quality daycare and female employment, *Scandinavian Journal of Economics*, 112, 570–594.

Waldfogel, J. (1998). The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics*, 16, 505–545.

World Economic Forum (2013). *The Global Gender Gap Report 2013*, World Economic Forum, Geneva.

Table 1: Indicators of family policies and gender equality in the Nordic countries compared to the OECD average and a selection of comparison countries

	Full-rate equivalent total paid weeks of parental leave		Public expenditure on maternity and parental leaves per child born in USD (PPP converted)	Pre-school enrollment rates		Expenditure on childcare support and pre-primary education per child in USD (PPP converted)	Maternal employment rates		Gender gap in full-time equivalent employment rates (15–64)	Female share of managers divided by female share of total employment
	Available to mothers	Reserved for fathers		0–2 year olds	3–5 year olds		Youngest child 0–2	Youngest child 3–5		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Denmark	26.7	1.1	21,803	67.0	97.7	10,119	75.3	80.1	12.1	29.1
Finland	42.6	6.4	25,346	27.7	74.0	9,538	49.7	84.1	8.5	30.2
Iceland	16.8	8.4	13,444	58.0	96.8	10,323	-	84.8***	22.0	39.4
Norway	36.6	12.7	31,735	54.3	96.5	10,553	-	80.1**	13.4	34.6
Sweden	38.1	7.6	26,139	47.3	94.0	9,555	71.9*	81.3*	10.6	37.3
Germany	34.7	4.1	10,577	29.3	94.6	4,399	52.3	71.5	24.6	32.5
UK	12.1	0.4	9,076	35.1	96.3	7,818	58.5	62.7	25.5	38.6
US	0.0	0.0	-	-	65.7	5,454	55.5	61.7	-	50.6
OECD average	27.6	4.2	12,435	34.9	82.0	6,140	52.9	66.3	22.4	36.4
Year	2014	2014	2011	2013	2012	2011	2013	2013	2013	2013

* Data from 2009.

** Employment rate for all women aged 25–39.

*** At least one child aged 0–14, data from 2002.

Source: All data from the OECD online family database: http://www.oecd.org/els/family/database.htm

7. Education and equality of opportunity: What have we learned from educational reforms[1]

Helena Holmlund, Institute for Evaluation of Labour Market and Education Policy (IFAU)

Abstract

Equality of opportunity has been one of the central ideas governing education policy in the Nordic welfare state. This paper takes its starting point in the shared history of educational reform in the Nordic countries, and presents evidence that the *comprehensive school reforms* that implied a shift from selective two-tier schooling systems to unified compulsory schools were beneficial for equality of opportunity. This evidence is compared to a *choice and voucher reform* that in the 1990s introduced pedagogical as well as organizational variety in the education system in Sweden. The Swedish choice reform is unique in an international perspective, and has reshaped the education sector dramatically as a growing number of pupils attend non-public independent schools. The current education debate shows a widespread concern that the introduction of choice has led to a backlash for equality of opportunity. However, recent research finds no indication that parental background has become more important

[1] I am grateful to Anders Björklund, Bertil Holmlund, Kristian Koerselman, Jesper Roine and conference participants at the NEPR conference in Helsinki, October 2015, for valuable comments.

in explaining pupil outcomes. The Swedish education system nevertheless faces a number of challenges if it is to level the playing field and create equal opportunities for all pupils: schools are becoming increasingly more segregated, much as a consequence of immigration, and disadvantaged pupils are less likely to exercise school choice compared to their more advantaged peers. This development calls for a debate on policies that aim to limit school segregation.

7.1 Introduction

There is a widely shared belief that all children, regardless of family background, should face equal life chances and have equal opportunities to succeed in life. Family background, in a broad sense, can be referred to as background factors not chosen by the individual, and inequalities in outcomes that are related to family background are therefore considered unfair. Despite this belief, parental background remains a very strong predictor for children's educational attainment and for success in the labour market in general. For example, in PISA (Programme for International Student Assessment), all surveyed countries show a strong socioeconomic gradient with respect to students' results (National Agency for Education, 2013) and correlations between parents' and children's years of schooling range from about 0.20 in the Nordic countries to about 0.5 in the U.S. (Holmlund *et al.*, 2011). Because of these strong correlations, policy makers often turn to the education system as a tool to level the playing field and create equal opportunities for children from different backgrounds.

The Nordic countries have a shared history of educational reform, including the expansion of publicly subsidized pre-school and the school reforms that gradually expanded and re-modelled compulsory education. One of the explicit aims of these reforms was to enhance equality of opportunity and increase social mobility, and as such, these policies played an important role in the construction of the welfare state. Research shows that in the Nordic countries, the relationship between parents' and children's socio-economic status is weaker than in for example the U.S. (Björklund and Jäntti, 1997; Björklund *et al.*, 2002).

Studies in this field typically estimate the association between parents' and children's education or earnings, and the stronger the association, the more important is family background for an individual's life outcomes, and the lower is social (or intergenerational) mobility. There are many different reasons why social mobility has been higher in the Nordic countries compared to many other nations, but the strong emphasis on equality of opportunity in education is one potential candidate that may explain this pattern.[2]

From a theoretical perspective, there are several reasons for which the design of the education system can be important for the intergenerational persistence in, e.g., educational outcomes. Public investment in education may help poor families to overcome credit constraints, and give all children access to education throughout the education system. Public policy can also affect intergenerational persistence through its interplay with educational choices (Björklund and Salvanes, 2011). If families of different socio-economic status hold different information about school quality and future returns to education, or have different preferences for their children's education, elements of choice through early tracking or choice of school may imply a stronger intergenerational relationship compared to a comprehensive system with limited options to choose between different tracks, pedagogical orientations or schools. In the latter case, the degree of parental influence over educational choices will be limited.

The Nordic countries all reformed their compulsory school systems throughout the 1950s–1970s. Common features were to extend the length of compulsory education, and to postpone the differentiation of pupils into different educational tracks (typically vocational and academic tracks) to a higher age. Since the early 1990s however, school reforms in Sweden have introduced new elements to the education system, which distinctly differentiates Sweden from most other countries.[3]

---

[2] In this paper the term "social mobility" is used broadly, as a synonym to "intergenerational mobility", and refers to mobility in income and educational outcomes. In the sociology literature, "social mobility" relates more closely to the concept of social class (see Erikson and Goldthorpe, 1992).

[3] Sweden is one of few countries in the world with a generalized voucher system and independent voucher schools face relatively loose regulations. In the Nordic countries, Denmark has a long tradition of independent schools (around 13% of all compulsory school pupils attend an independent school), which are often

In the 1980s, critics had started to question the long-lasting social-democratic influence over the school and education system, and in particular advocated more diversity in educational provision, and the right for parents to exert influence over their children's education (National Agency for Education, 2003). This debate eventually led to the passing of a series of educational reforms in Sweden. In the early 1990s, a choice and voucher reform implied that private schools were granted public funding, and greater opportunities for pupils to opt out of their assigned local public schools to either attend a voucher-funded private (or "independent") school, or a public school outside of their catchment area. During the same time period, the public school system was decentralized and the previous state control was replaced by a municipality (i.e. local authority) maintained system, with the purpose to let local priorities play a larger role in e.g., resource allocation.

Since the introduction of universal vouchers, many new independent voucher-funded schools have opened in Sweden. As of 2014, about 14% of all compulsory school pupils attend an independent school.[4] These schools represent various pedagogical ideas and religious affiliations, but it is also common that for-profit corporations run schools with a general profile.

There were two main arguments for school choice and diversity in educational provision put forward in the debate: parents should have a right to choose their child's school, and competition between schools and different pedagogical ideas should increase efficiency in the school system and improve educational outcomes (Government bill 1991/92:95). Opponents of the reforms, on the other hand, argued that a choice-based system with independent schools would lead to segregation and inequality, and that the reforms were at odds with the ideals of an inclusive education system providing equal education to all pupils, regardless of family background (Opposition bill 1991/92: Ub62). In the

---

run by parent co-ops with religious or pedagogical profiles. Ownership regulations are different than in the Swedish system and imply that schools cannot be run as for-profit corporations. In Finland and Norway, independent schools are rare and only about 2% of compulsory school level pupils attend such a school (SOU 2013:56).

[4] Calculation based on statistics available at www.skolverket.se

current education debate, 25 years after the passing of the reform, equity concerns are again on the policy agenda: between-school inequality is rising both in terms of test scores and parental background (Holmlund *et al*., 2014). In tandem, Sweden's scores in PISA have declined sharply and this empirical correlation has brought additional light on the topic of choice, segregation and educational performance.

The purpose of this paper is twofold. The first purpose is to summarize the existing evidence on the role of educational reforms for intergenerational mobility in the Nordic countries. Specifically, the paper will compare *comprehensive schooling reforms* that took place in the 1950s–1970s, to the *choice and voucher reform* in Sweden in the 1990s.[5] Importantly, the latter reform has affected more recent cohorts that have not yet completed their education and entered the labour market. Research on intergenerational mobility in relation to this reform has therefore used age 16 school performance to proxy for intergenerational mobility. That is, using intermediate outcomes that are highly correlated with long-run education and labour market outcomes, it is possible to shed light on more recent trends in mobility.

The second purpose is to identify and discuss future challenges to the ideal of equality of opportunity in education, in light of recent developments in Swedish society. Reconciling the literature, the main conclusions of the paper are that the comprehensive school reforms likely have improved social mobility, while the more recent school choice reform in Sweden does not seem to have affected short-run outcomes that proxy for social mobility. The Swedish education system nevertheless faces a number of challenges if it is to level the playing field and create equal opportunities for all pupils: school segregation has been on the rising for a long time, much as a consequence of immigration, and disadvantaged pupils are less likely to exercise school choice compared to their more advantaged peers.

The remainder of the paper is structured as follows: section 2 discusses some theoretical predictions relating to intergenerational mobil-

[5] This paper will not cover the literature on the expansion of subsidized pre-school education in the Nordic countries. Evidence from Norway shows that provision of subsidized pre-school has increased intergenerational mobility (Havnes and Mogstad, 2011).

ity and the workings of the school system; section 3 summarizes the evidence from the compulsory schooling reforms in the Nordic countries; section 4 moves on to more recent evidence on the choice and voucher reform in Sweden and section 5 outlines future challenges to equality of opportunity in the Swedish school system. Finally, section 6 concludes with a short summary and discussion.

7.2 Theoretical considerations related to the intergenerational transmission of human capital

Economic models of human capital take their starting point in a model of investment, originating from Becker and Tomes (1979). In such a model, a child's human capital, defined broadly as for example education, and cognitive and non-cognitive skills, will be determined by parental inputs (both in terms of nature and nurture) and by public investment in education. In this framework, one would expect a positive correlation between child's and parent's human capital (or income), in part because of the genetic transmission of abilities, and in part because credit constraints might imply that well-off families can afford to invest more in their offspring's education than poor families. Parental investment in human capital should be interpreted not only as formal education, but also as parents' own time spent with the child, the quality of time spent together, and their parenting skills.

The optimal parental investment will depend on factors such as the return to human capital, the degree of public investment, and the institutional set-up of the school system – and therefore these factors may in turn affect the strength of the intergenerational associations. As an example, a model by Solon (2004) shows that intergenerational income mobility is decreasing in the returns to human capital and increasing in the degree of public investment in education.

Public investment in education is clearly a policy that relaxes credit constraints for poor families and the theoretical prediction is therefore that it should lower intergenerational persistence in economic outcomes. But how can institutional arrangements such as ability tracking,

comprehensive vs. selective systems, and school choice affect the intergenerational link and social mobility?[6]

First, consider the use of ability tracking within a school, i.e., pupils are sorted into different classes based on their previous school performance. As suggested by Betts (2011), in a system without tracking, affluent parents might invest in private tutoring on top of public education, in order to obtain good results and gain access to the best universities. In a system with ability tracking, on the other hand, well-off parents might see less need for extra tutoring if they consider placement of their children in a high track as a substitute for private investment.

Second, how does parental background influence educational choice in a selective two-tier system compared to a comprehensive system? Early differentiation is argued to increase the importance of parental characteristics for educational choice. The earlier the choice is made, the less accurate is information about pupil ability, which tends to put weight on parental socioeconomic background rather than pupil ability in the educational decision (Björklund and Salvanes, 2011). Moreover, the younger is the child, the stronger is the influence of parental, rather than the pupils' own, preferences for education. The prediction is therefore that postponing differentiation into academic and vocational tracks to a later age will limit the importance of family background for educational choice, and therefore be beneficial for social mobility.

Third, school choice policies may also be of importance for intergenerational mobility. On the one hand, if parental preferences for education vary by family background, or if families have different access to information about school quality and different abilities to interpret such information, it is possible that choice leads to a segregated school system where the type or quality of the school attended is correlated with family background. On the other hand, it can be argued that school choice provides an opportunity for disadvantaged children to "escape" low-performing schools in poor neighbourhoods and the link between school quality and family background should therefore weaken.

---

[6] This question highlights the potential trade-off between efficiency and inequality, where proponents of tracking argue that tracking increases efficiency, while opponents point to increased inequality as a result of separating pupils into different tracks or schools (Betts, 2011).

Most empirical studies that focus on the decision of opting out from assigned schools find that advantaged families are more prone to opting out (see Böhlmark *et al*., 2015 for a summary), which implies that increased sorting seems to be the dominating factor. In such systems, children with a favourable background will access the best schools, which will strengthen the association between parents' and offspring's outcomes. School choice within the publicly funded school system can thus lead to sorting, but it can also limit the demand for highly selective tuition-charging private schools (Epple *et al*., 2002). High demand for such schools could potentially lead to sorting between the public and private sectors, and imply very strong correlations between family background and children's educational and labour market outcomes.

All in all, summing up the theoretical predictions above, it can be argued that a comprehensive system without the possibility of opting out of assigned schools will tend to limit the influence of parental background for educational outcomes. However, in such a system, parental preferences for high-quality selective education or for a specific peer group composition might open up for a market of tuition charging private schools to which only affluent families have access. This scenario predicts that family background will have a strong influence on educational success. Which theoretical mechanism that dominates will ultimately be an empirical question.

7.3 Compulsory schooling reforms, tracking and intergenerational persistence

7.3.1 The compulsory schooling reforms

In the 1940s and 1950s, children in the Nordic countries commonly started school at the age of 7 and went to a common primary school for 4–6 years, after which they were split into two tracks: remaining in

primary school or attending a general lower secondary school.[7] Pupils who remained in primary school completed 7–8 years of schooling, which at the time was the compulsory minimum. Admission to lower secondary school was typically based on school grades, teacher assessments and entrance examinations. Attending lower secondary school implied leaving school after 9–10 years, if not continuing further at the upper secondary level. This two-track selective system differentiated pupils at the age of 10–12, into paths that were distinctly different: the secondary school path had a more academically oriented curriculum and prepared for future academic studies, while the primary school path was followed by vocational training and did not grant access to academic studies.

This educational model was reformed and replaced by a comprehensive school with a similar curriculum for all pupils from 1st to 9th grade.[8] One explicit motivation for these reforms in the Nordic countries was to give all children, irrespective of family background, the same basic education, that is, to provide equal educational opportunities for all. Early differentiation was considered a disadvantage for children from low-educated households who disproportionally remained in primary school, and the reforms aimed at raising these children to a higher academic level.[9]

One common feature of the reforms was that they were rolled out gradually across different regions. In Sweden, the reform started at a small scale in 1949 and was implemented throughout the 1950s. In Norway, the reform period spanned 1960–1972. In Finland, the comprehensive school was introduced in the 1970s, but at this point in time the length of schooling had already been harmonized for different

[7] This section builds on Meghir and Palme (2005), Arendt (2005), Pekkarinen *et al*. (2009) and Aakvik *et al*. (2010).

[8] In Sweden, some tracking initially remained within the comprehensive system, but was abolished in 1969 (National Agency for Education, 1969).

[9] Equality of opportunity was an explicit motivation for the reforms, but also other motivations are worth highlighting. In Sweden for example, demand for education beyond the compulsory level was growing among the baby boom cohorts born in the 1940s, and the education level was increasing in the population already before the reform was rolled out. The compulsory school reform was therefore also a means to meet increased demand for education in the population (see Holmlund, 2007).

tracks, and as such the reform served mainly to merge the two-tier system into a comprehensive school system. The gradual implementation has been of great importance to researchers, who have been able to evaluate the reforms by comparing individuals who went to school in different regions at different points in time. The exception is Denmark, where reforms were implemented throughout the country at a single point in time: In 1958 early sorting was limited, and in 1975 the two track-system was fully abandoned and the years of compulsory education increased from 7 to 9. The next section summarizes some of the findings from this research.

7.4 Causal effects of compulsory schooling reforms on intergenerational mobility

Although there are some theoretical predictions that relate features of the education system, such as ability tracking, to social mobility, it is a challenging task to provide empirical evidence on this topic. As already mentioned, comparisons of intergenerational income mobility in different countries have shown that mobility is higher in the Nordic countries than in continental Europe and in the U.S. and the U.K. (Björklund and Jäntti, 1997, Blanden, 2013). This pattern is often interpreted as a success of the Nordic welfare state, but cross-country comparisons are not enough to understand the underlying mechanisms of specific policies, and may also be explained by other cross-country differences. When it comes to education policy and ability tracking specifically, most cross-country comparisons show that tracking increases inequalities and the role of family background for educational outcomes (Amermüller, 2005, Hanushek and Wossman, 2006, Brunello and Checci, 2007).[10] These studies all share the limitations of a cross-country comparison: it is not obvious that the analysis can account for all the "other" relevant factors that explain differences in outcomes between countries.

[10] See Waldinger (2007) for an exception.

More recently, experimental designs have been used to compare the outcomes of ability-grouped vs. mixed-group pupils. Evidence from primary school children in Kenya (Duflo *et al*., 2011) and university students in the Netherlands (Booij *et al*., 2015) shows that tracking can have positive effects on outcomes both for low and high ability students, and that the positive tracking effects for low ability students can be higher than the positive peer effects that would be incurred from a mixed peer group.

While the cross-country variation is potentially unsuitable for causal inference, the recent experiments are more compelling when it comes to internal validity of the estimates. Nevertheless, their results are not always generalizeable to a system-wide context and they provide limited evidence on the effects of reforms at the institutional level. Instead, a large body of research based on the comprehensive school reforms in the Nordic countries has provided evidence on the causal effects of a system-wide change to the education system on a range of outcomes. These within-country comparisons exploit the gradual roll-out and are therefore able to isolate the effects of the reforms in a setting where it is unlikely that other factors confound the results.

The first paper to exploit the gradual implementation in order to study the effects of a comprehensive school reform in the Nordics was Meghir and Palme's (2005) influential study on the Swedish reform.[11] The main findings from their paper are that the reform increased years of schooling and lifetime earnings for children with low-skilled fathers, while there is a tendency to find negative earnings effects for children whose fathers were high-skilled. These results imply that the reform contributed to reducing inequalities in labour market outcomes by family background, and as such, one of Meghir and Palme's conclusions is that the reform increased intergenerational mobility.

The Finnish reform, explored in two papers by Pekkarinen *et al*. (2009; 2013) provides evidence in line with the study of the Swedish case. Pekkarinen *et al*. (2009) estimate the intergenerational income

[11] Erikson (1996) had previously studied equality of opportunity in the light of the Swedish reform, and found that the introduction of the comprehensive school coincided with reduced inequality in education.

elasticity, which is the regression coefficient from a regression of offspring's log life time income on parents' log lifetime income. A high value indicates strong income persistence between generations (i.e., low social mobility) and a low value indicates high social mobility. They find that the reform reduced the intergenerational income elasticity by 23%, from 0.30 to 0.23. This is a sizeable effect and is backed up further by the 2013 study which finds that the reform improved the cognitive skills of boys with low educated parents, while not affecting boys whose parents had some education above compulsory level.[12]

Estimates for Norway are in line with the evidence presented above, and lend further support to the conclusion that the comprehensive school reforms contributed to increasing intergenerational mobility. Aavik *et al*. (2010) find that the effect of family background on educational attainment was lowered as a result of the Norwegian reform. The Danish reform, which was instituted throughout the country at a single point in time, does not allow for an evaluation design similar to that of the studies described above. As a consequence there is no published research that can complete the picture with estimates from Denmark.

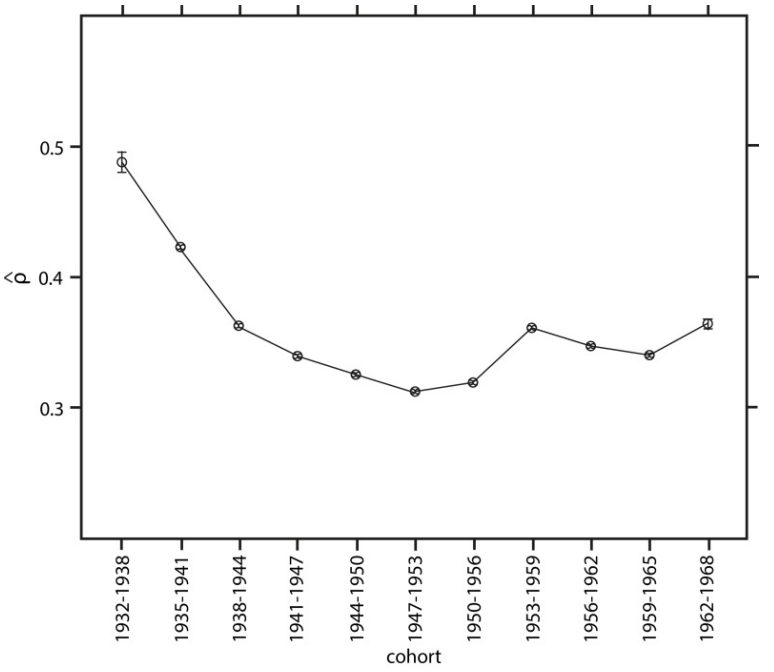
Before concluding this section, it is worth noting that these studies build on regression models that relate children's education and earnings to their parents' education and earnings. While these observed measures of socio-economic status are highly relevant, alternative empirical methods provide additional evidence to support the findings presented above. Björklund *et al*. (2009) make use of "sibling correlations" to capture the importance of family and community background for children's labour market success. The sibling correlation estimates how much of the variation in outcomes, in this case life time income, that can be attributed to the family in *a broad sense*. If siblings are very similar (the sibling correlation is high), shared family background factors, such as socio-economic status, marital stability and parenting skills, as well as shared community factors including the school and the neighbourhood, are important for children's outcomes. If the sibling

[12] Similar estimations of the reform effect on the intergenerational income elasticity in Sweden can be found in Holmlund (2008), who finds that the Swedish reform reduced the elasticity by 12%.

correlation is low and siblings are not alike in terms of outcomes, family and community background seems to be a less important explanation. The sibling correlation, by virtue of capturing all sorts of shared background factors – also factors that cannot be observed in the data – is considered an important complement to the more traditional regression-based measures that link parents to their children. Figure 1 shows the brother correlation in income, as estimated in Björklund *et al*. (2009).[13] The figure displays a sharp decline in the brother correlation, from 0.49 for cohorts born 1932–1938 to 0.32 for cohorts born around 1950, thus suggesting that family and other shared background factors became less important determinants of long-run income in adulthood. In additional analyses, it is shown that the decline can be explained by changes in the distribution of years of education. While the early decline pre-dates the comprehensive school reform, these results may reflect earlier expansions of compulsory education (see Fisher *et al*., 2013 for a description of earlier reforms) and are suggestive of the importance of the education system for intergenerational mobility.

[13] Figure 1 is based on full biological brothers born at most seven years apart. One concern with comparing the sibling correlation over time is that it might be sensitive to trends in child spacing and family structure. Björklund and Jäntti (2012) find that sibling similarities are not much affected by age differences of siblings in the sample and they therefore conclude that shared permanent factors are most important in explaining the sibling correlation. The sibling correlations in Figure 3 are based on siblings with the same mother (i.e. also half siblings on the maternal side) born at least three years apart. The conclusions from Figure 3 are robust also to alternative specifications including only full biological siblings.

Figure 1: Brother correlations in income in Sweden



Source: Björklund *et al*. (2009).

In the next section, indicators of social mobility will be related to more recent reforms to the education system in Sweden.

7.5 Parental choice and the importance of family background for pupil achievement

The research presented in the previous section shows that the shift from a selective to a comprehensive system, as manifested by the reforms in the Nordic countries in the 1950s–1970s, reduced inequalities in educational attainment and in life time income by family background. In this section, these results will be contrasted with a more recent re-

form in Sweden that in the early 1990s introduced universal school vouchers and generalized school choice.[14]

7.5.1 The choice and voucher reform

In 1992, Sweden introduced a nation-wide voucher system, which facilitated public funding of private schools, and introduced school choice. In the pre-reform period, with institutions dating back to the comprehensive system described above, almost all pupils attended the local school in their catchment area; less than one percent of all pupils attended a private school. The reform was based on two main elements. First, privately maintained "independent" schools receive public funding through school vouchers after having gained approval by the Swedish National Agency of Education, NAE. Pupils' home municipalities have to provide independent schools with a grant, based on the average per-pupil expenditure in the public school system, for each pupil who chooses to enrol in an independent school. This new law has given rise to a large number of new schools, whose existence depends solely on funding through vouchers. The number of independent schools (at the compulsory level) has increased, from about 170 registered independent schools in 1993 to 800 in 2014. In 2014, about 14% of all compulsory school pupils attended an independent school. Independent schools are open to all pupils: by law they are not allowed to charge tuition fees on top of the voucher, nor can they select pupils by ability or family background. If an independent school is oversubscribed, three selection criteria for admission are allowed: proximity to the school; waiting list (by date of application); and priority for children whose older siblings are already enrolled in the school.[15] Second, the voucher reform also

[14] For more details on the reform and the current institutional setting, see SOU 2013:56.

[15] Independent schools were initially allowed to charge moderate tuition fees, but the right to charge fees was abolished in 1997. The funding rules have varied over time: at the outset of the reform, the voucher should amount to 85% of per-pupil expenditure in the public schools in the municipality. After 1997, the voucher to independent schools should be determined "on the same grounds as to public schools" (Government bill 1994/95:157). Independent schools can have different types of governing bodies, for example non-profit foundations and for-profit corporations. A majority of pupils attend schools with a general profile, belonging to a for-profit corporation (Swedish Association of Independent Schools 2015).

introduced choice between public schools, although maintaining priority for pupils residing close to the public school. Slots are first allocated to pupils within the public schools' catchment areas, after which pupils from other areas can be granted admission.

School choice and competition cannot explain the sharp decline in Swedish PISA scores (Böhlmark and Lindahl, 2015, Wondratschek *et al*., 2013). The current Swedish education debate is nevertheless strongly concerned that inequalities in educational outcomes are on the rising, and it is often believed that the importance of family background for educational success has increased as a result of the choice and voucher reform (National Agency for Education, 2009). The rationale for these concerns is that the possibility to exert school choice and the probability of gaining access to the preferred school might be related to a pupils' family background, as discussed in Section 2. Even though the institutional set-up is designed with the purpose of giving equal access to children from different family backgrounds, socio-economically advantaged families might have better information about school quality and about the waiting list principle at oversubscribed schools, which in turn implies a higher probability of accessing a popular, high-quality, school.

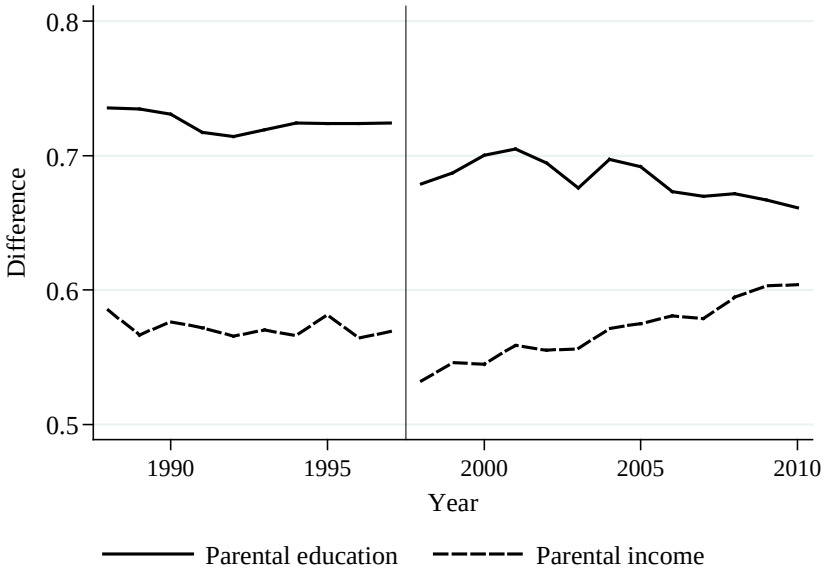
While research based on the Nordic comprehensive school reforms has focused on the relationship between family background and adult outcomes, such as completed education (years of schooling), earnings or income, recent studies of the Swedish choice and voucher reform have limited the analysis to intermediate outcomes such as grade point averages (GPA) and test scores at age 16. The rationale for this is that the cohorts that are included in the analyses have been too young for long-term outcomes to be relevant (cohorts born 1972–1993). Since school results at age 16 are good predictors of long-term labour market outcomes, the results can nevertheless serve as a proxy for intergenerational mobility in the cohorts affected by the choice reform. The next section presents evidence that relates indicators of intergenerational mobility to the choice and voucher reform.

7.5.2 Parental background and school performance

The first study to assess the family background gradient in school performance, in the wake of the choice reform, was Björklund *et al*. (2003). Since then, a number of studies have followed suit and reported results on this topic (see for example Gustafsson and Yang-Hansen, 2009, Fredriksson and Vlachos, 2011, National Agency for Education, 2012, Böhlmark and Holmlund, 2012). More recently, Holmlund *et al*. (2014) provide a range of results, based on various data sources, which describe how the importance of family background for school performance at age 16 has evolved in the period 1988–2009. The main results are presented in Figures 2 and 3.

Figure 2 shows the GPA difference, expressed in standard deviations, between pupils characterized by high and low parental education and income, respectively. The vertical line in 1988 reflects a grading reform. First, the graph reveals that there are sizeable differences in terms of school performance for pupils of different background. The difference between children of high and low educated parents is about 0.7 standard deviations; slightly below 0.6 comparing high and low income, and the differences remain stable until 1998. The new grading system implied a shift towards smaller differences between groups: family background has a lower explanatory power of school performance in the new system. The GPA difference by parental education turns out to be relatively stable over time also after the new grading system was introduced: if anything, the gap between pupils with high and low educated parents is shrinking. On the contrary, differences by parental income are increasing throughout 1998–2009. This pattern can be the result of an increasing role for parental income in determining school performance, but it cannot be excluded that the result is explained by compositional effects that are linked to increased income inequality in the population. The group defined as high income (the top 25% of the income distribution) is gradually becoming richer during the study period, which can in itself affect the GPA difference presented in Figure 2.

Figure 2: Differences in age-16 GPA by family background



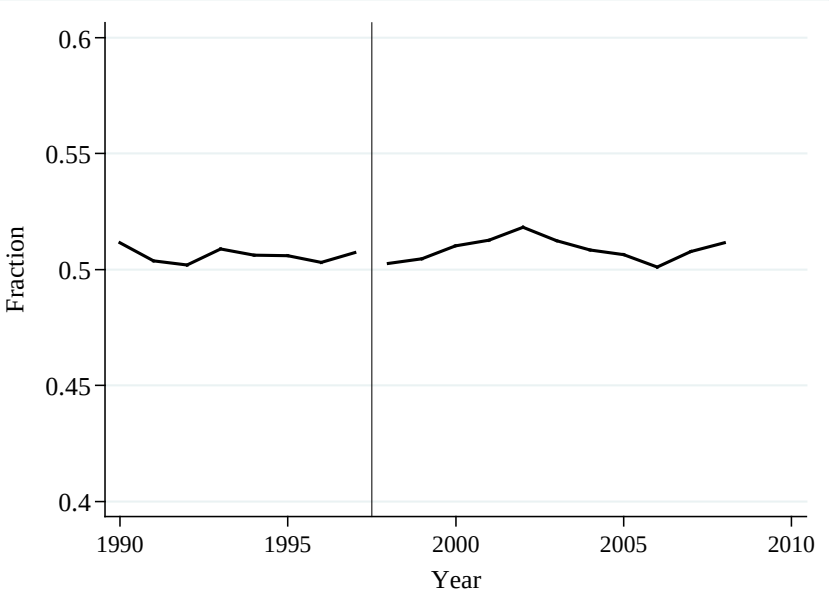
Notes: The figure shows differences in the GPA (based on Swedish, English and mathematics) between pupils with different family background. The lines show the differences between pupils characterized by high/low parental education or income. Highly educated parents are defined as at least one parent belonging to the top 25% of the education distribution (by parents' birth year and gender); the rest are defined as low educated. High income parents are defined as at least one parent belonging to the top 25% of the income distribution (by parents' birth year and gender), when income is measured in the age range 35–45. Remaining parents are defined as low income parents. The vertical line indicates a change in the grading system.

Source: Holmlund *et al*. (2014).

Next, Figure 3 presents sibling correlations in GPA at age 16. The correlations reveal that family and community background factors shared by siblings can explain about 50% of the variation in school performance, as measured by the GPA. Thus, siblings' shared background explains a large fraction of the variation, especially compared to the school which

explained 11% of the variation in 2009.[16] As in the previous graph, the shift to a new grading regime is demonstrated as a break in the time series. Importantly, there appears to be no trend in the sibling correlation over time, which means that family background factors have become neither more nor less important in shaping pupils' age-16 GPA.

Figure 3: Sibling correlations in age-16 GPA



Notes: The figure shows sibling correlations (fraction of variance explained by the family) in 9th grade GPA. Calculations are based on siblings born at most three years apart. The vertical line indicates a change in the grading system.

Source: Holmlund *et al*. (2014).

The key findings from Figures 2 and 3 are that family background factors are very important for school performance – but their importance is remarkably stable over time. While there is some evidence indicating that differences are increasing by parental income, the other measures

[16] See Holmlund *et al*. (2014) for variance decompositions of GPA at the municipality, school and family levels.

adopted do not confirm the notion that family background has become a stronger determinant of educational success in Sweden.[17] Based on these time-series from 1988–2009, it seems like the choice-based system introduced in the 1990s has not affected intergenerational mobility, at least not in the short run. Admittedly, this conclusion is based on the observation of a time series, and does not rely on a solid identification strategy as was the case for the comprehensive school reforms. It cannot be ruled out that introducing choice has increased the social gradient in school performance, but other changes in society may have had effects going in the opposite direction and therefore cancelled out the effects of choice (see Björklund *et al*., 2003).

7.6 The Swedish model: lessons learned and challenges for the future

The previous section provided evidence indicating that the growing sector of independent schools and the possibility to opt out of the local public school, have not strengthened the link between parental background and children's outcomes among cohorts that left compulsory education between 1988 and 2009. There are however some striking empirical features, related to the choice and voucher reform, but also to other changes to the Swedish society, that constitute challenges to the future of the school system and to equality of opportunity. In this section, a few of these challenges and their related policies will be discussed.

[17] Holmlund *et al*. (2014) show that these conclusions are robust also to using alternative sources of data. Analyses based on PISA data are somewhat inconclusive as to whether parental background has become more important for pupils' results in Sweden. The PISA index for socio-economic status has become a stronger predictor for reading performance between 2000 and 2009, while the same index has not become more important for mathematics between 2003 and 2012 (National Agency for Education 2013).

7.6.1 Choice and sorting to schools

Swedish independent schools cannot charge tuition fees, nor can they cream-skim the best pupils.[18] However, there is scope for independent schools to influence the pool of applicants indirectly; either by locating in advantaged areas, or by targeting information to specific groups of parents. In addition, anecdotal evidence indicates that independent schools do not always abide by the rules and reject low-performing students with disruptive behaviour (SVT 2013). On the demand side, preferences for schooling and access to information about school quality might differ by family background and affect the decision to opt out of the local public school. The waiting list principle to admit pupils to oversubscribed schools might also discriminate against groups of pupils with less information about how the school choice system works. These mechanisms typically imply that school choice might increase school segregation.

Who are the pupils that opt out of their local public school to attend an independent school? Table 1 shows that native children, as well as children with high-income and highly educated parents are more likely to attend an independent school, compared to immigrants and children whose parents have lower income and education.

[18] Independent schools cannot charge tuition fees, but schools are allowed to receive donations from parents.

Table 1: Fraction of 9th grade pupils in different demographic groups attending an independent school in 2009

All	10.50
Immigrant	8.45
Native	10.65
Immigrant background	11.28
Swedish background	10.36
High parental education	15.28
Low parental education	6.07
High parental income	16.37
Low parental income	9.27

Note: High parental education refers to at least one parent with a three year university degree. Low parental education is defined as both parents holding only compulsory education. High vs. low parental income is defined as having family income (defined during school age) belonging to the top 20 or bottom 20 percentiles of the income distribution.

Source: Own calculations based on the 9th grade register matched with parental background information.

Figure 4 (adapted from Holmlund *et al*., 2014) adds to this picture by presenting odds ratios of the probability of attending an independent school, for pupils of different family background, living in different types of neighbourhoods. Neighbourhoods (defined as catchment areas) are characterized as advantaged or disadvantaged based on an index of family background that takes into account parental education, parental income, and migration history. Odds ratios are calculated by comparing only pupils living in the same catchment area, who should be expected to face the same supply of schools to choose from at a reasonable distance. An odds ratio equal to one means that two groups of students, e.g., immigrants and natives, have the same probability to attend an independent school, while an odds ratio below (above) one implies a lower (higher) probability. The figure shows that foreign background pupils have a much lower probability (odds ratio 0.55) than natives to attend an independent school, if they live in a disadvantaged catchment area. That is, children with Swedish background are more likely than immigrant background children to attend independent schools if they live in poor neighbourhoods – and strikingly, the figure also shows that this is not the case in catchment areas with a more favourable demographic composition, where pupils of Swedish background have a much

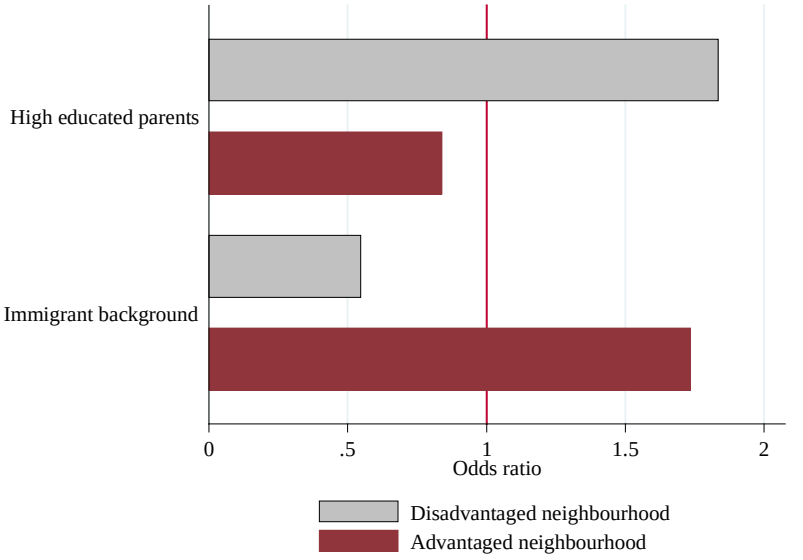
lower probability to attend independent schools. Similarly, children with highly educated parents have a much higher probability (odds ratio 1.83) of attending an independent school compared to children whose parents have lower education, if they live in a catchment area with high proportions of disadvantaged background peers. Instead, the probabilities are reversed for children of different educational background in advantaged catchment areas.

How can these sorting patterns be interpreted? The demographic composition of the catchment area can be seen as a proxy for the expected peer group in the local public school. In disadvantaged areas, pupils with Swedish background or with highly educated parents are more likely to leave the public school to attend an independent school than pupils with immigrant background and low educated parents. In contrast, in socially strong areas, the same pupils have a lower likelihood of attending an independent school than disadvantaged pupils. Advantaged pupils who live in neighbourhoods with similar peers remain in the public school and even attend the public school to a larger extent than more disadvantaged children do.

To sum up, the odds ratios in Figure 4 show that pupils with Swedish background and pupils whose parents have high education sort themselves into school types (public or independent) depending on the demographic composition of the catchment area which they belong to. If the demographic composition is favourable, they are likely to remain in the public school, while they are more likely to opt out and attend an independent school if they expect to find a more disadvantaged peer group in the local public school. To some extent, these sorting patterns reveal that peer group considerations are important when school choice is exercised.[19]

[19] See also Andersson *et al.* (2012) for similar findings.

Figure 4: Probabilities of attending an independent school, by parental background and neighbourhood demographics



Notes: Neighbourhoods are defined as advantaged (disadvantaged) if they are above (below) the median in the pupil-weighted distribution of the family background index. The family background index is a measure of expected GPA given by parental background, and is calculated by predicting GPA using the regression coefficients from a regression of GPA on family background factors such as parental education, parental income and immigrant status. High education parents are defined as at least one parent with a three-year university degree. Pupils with foreign background are born abroad to two foreign-born parents, or born in Sweden to two foreign-born parents. Odds ratios are computed by comparing pupils living in the same catchment area.

Source: Holmlund *et al*. (2014).

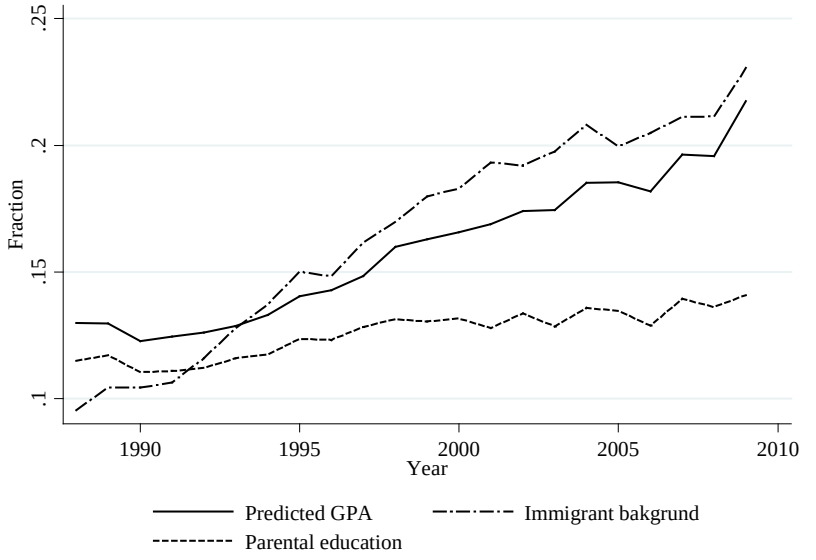
The sorting to independent schools as displayed in Table 1 and Figure 4 suggests that school choice is related to school segregation. The next section discusses the evolution of school segregation and its causes in more detail.

7.6.2 School segregation

Starting in the mid-1990s, school segregation has increased gradually in Sweden. Figure 5 illustrates this development with one indicator of segregation: the between-school variation in various measures of family background.[20] First, the solid line shows the between-school variation in the family background index "expected GPA", that is, the GPA predicted by a set of family background characteristics such as parental education, parental earnings and migration history. It is clear from the graph that in terms of this combined measure of family background factors, schools have become more segregated throughout the period 1988–2009. Next, by studying segregation by different family background characteristics separately, it turns out that segregation between pupils of Swedish or immigrant background has increased the most, while the increase in segregation by education background is less dramatic.

[20] This pattern is robust also using alternative segregation indices; see for example Böhlmark *et al*. (2015).

Figure 5: Between-school variation in family background



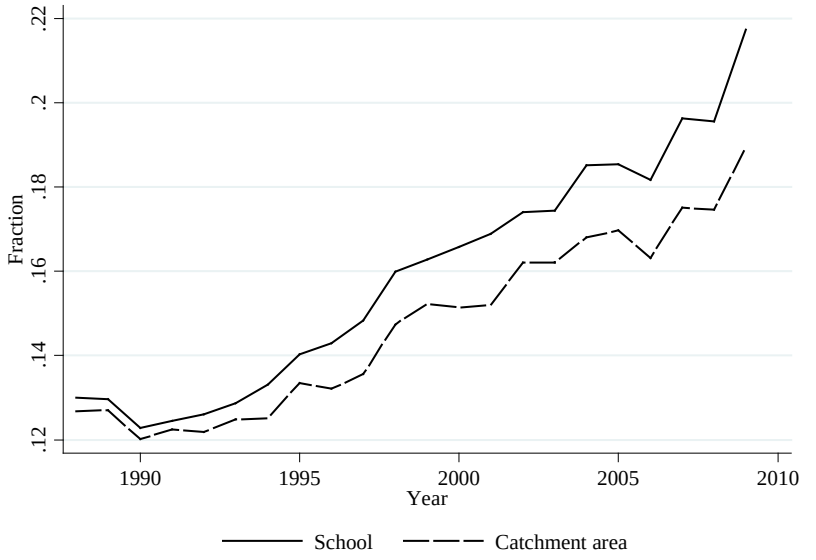
Note: The between-school variation is calculated annually using ANOVA. Parental education refers to the parental average years of schooling. Pupils with immigrant background are born abroad to two foreign-born parents, or born in Sweden to two foreign-born parents. GPA is predicted by parental background factors: education, income and migration background.

Source: Adapted from Holmlund *et al*. (2014).

Most children attend a school in their local catchment area, which means that school segregation is likely to be explained by residential segregation. However, segregation between schools took off in the early 1990s and has continued to increase as the number of pupils in independent schools has been growing. This empirical correlation has brought attention to the question of whether choice exacerbates school segregation over and above the segregation that is given by residential segregation. To bring clarity on this topic, Böhlmark *et al*. (2015) and Holmlund *et al*. (2014) study the association between the fraction of pupils opting out to independent schools, and various indicators of school segregation at the municipality level in Sweden. Their key finding is that the main contributor to school segregation is residential seg-

regation, but segregation between schools has increased more than what is predicted by changes in residential sorting in municipalities where choice has become more prevalent. Figure 6 illustrates this result by graphing the between-school and between-neighbourhood (defined as catchment areas) variation in the family background index "expected GPA" over time. Around 1990, before the choice reform when almost all pupils attended the local public school, segregation between schools and residential neighbourhoods was virtually identical. Over time, both residential segregation and school segregation has increased, but as the number of pupils enrolled in independent schools is increasing, segregation between schools is increasing more than segregation between neighbourhoods. Nevertheless, residential segregation remains the main explanation to the development over time, and as seen in Figure 5, segregation between children with different migration histories has increased the most. Holmlund *et al*. (2014) show that municipalities and residential neighbourhoods with a large immigrant population host disproportionate numbers of new immigrants, as the share of immigrants in the population is growing. This has undoubtedly increased residential segregation in the country, and has naturally also affected school segregation.

Figure 6: Between-school and between-catchment area variation in family background (predicted GPA)



Note: The between-school variation is calculated annually using ANOVA. GPA is predicted by parental background factors: education, income and migration background.

Source: Holmlund *et al*. (2014)

7.6.3 Policy considerations of choice and sorting

This section discusses policy considerations related to the social gradient in enrolment at independent schools and to increasing school segregation, in relation to equality of opportunity. Its purpose is to highlight policy areas that are relevant for the future development of the school system with regards to segregation and sorting.

As Section 5.1 has shown, the choice to opt out is not equally distributed across different socioeconomic groups. This has not implied large changes to school segregation at the aggregate level, but the social gradient in independent school attendance is nevertheless relevant from the perspective of equality of opportunity. This social gradient can be the result of differences in preferences, differences in access to in-

formation (and abilities to interpret information), and may also result from the waiting list principle which benefits parents who are forward-looking, well informed and do not move home. To date, there is no research using data on parents' revealed preferences for schools – most research builds on the school attended by the pupils. As a consequence, we know little about whether the sorting to independent schools is due to families ranking school differently (i.e., differences in preferences and/or information), or due to the assignment mechanism (i.e., families rank schools similarly, but the waiting list leads to sorting). However, from the perspective of equal life chances – and the notion that children do not choose their parents – none of the above explanations for the social gradient are justifiable. Even in the extreme case where there are no quality differences between schools and the sorting of pupils has no consequences for pupils' learning, one might still argue that families derive utility from the consumption value of school choice, and that the allocation of this consumption value across families is a policy issue.

Should policy makers be worried about increasing school segregation? To answer this question, it is important to consider what consequences segregation may have for pupil achievement – on average and for different groups of pupils – and for society at large. Taken together, the evidence in Böhlmark and Lindahl (2015) and Böhlmark *et al*. (2015) indicate that if segregation induced by choice has had negative effects on pupil performance on average – these have been offset by positive effects from school competition, indicating that the net effect of the choice reform on pupil performance is slightly positive. But because of residential sorting, school segregation has increased much more than what can be explained by the choice reform, and it is therefore important to consider its potential effects for the education system, and for society at large.

First, how does growing up in a segregated neighbourhood affect children's life chances? How does a segregated school affect the learning environment, how does it affect inequality in pupils' outcomes and are there effects on other life outcomes, such as criminal involvement? As already mentioned in Section 3.2, recent experimental evidence has shown that sorting pupils by ability can improve outcomes, and that the positive effect of tracking can be larger than the positive peer-effect for

low-ability students in a mixed peer group. However, it is unclear to what extent the evidence from these studies can be generalised to the Swedish compulsory education system and to a setting where pupils are segregated by family background, into different schools. Grouping pupils by ability is not the same as grouping by socioeconomic background, and within-school ability sorting is different than between-school segregation (see Betts (2011) for a discussion on the differences between tracking within and between schools).

Research focusing on segregation per se is potentially more informative for the Swedish case. Billings *et al*. (2014) present evidence that increased racial segregation in a U.S. school district, as a result of ending a de-segregating busing scheme, led to lower test scores for both white and minority pupils who were assigned to schools with larger numbers of minority pupils. In addition, the study finds that while compensatory resource allocation was able to efficiently remedy the negative effects on test scores, crime increased substantially among minority males who were assigned to schools with large proportions of other minority students. These findings align with earlier studies on peer racial composition and school segregation (see for example Deming, 2011). The key insights from these studies is that high concentrations of disadvantaged youth in segregated schools might increase test score inequality between children of different family backgrounds, and affect behaviour also in other dimensions such as crime. The conclusions are further corroborated by evidence from the Moving to opportunity-programme (MTO) that offered randomly selected families, living in disadvantaged neighbourhoods in the U.S., financial help to move to a better neighbourhood. Children who were able to move out of poor neighbourhoods at a young age turn out to be more likely to attend college and to live in better residential areas as adults, compared to children who were not able to move (Chetty *et al*., *forthcoming*).

The conclusions from the U.S. literature on racial segregation are suggestive of the consequences of segregation also in other settings, but the Swedish case of segregation between refugee immigrants and natives also has its unique features. Refugee families might have gone through traumatic events in their home countries; they come from many different nations, speak different languages, have varying educa-

tional histories and have arrived in Sweden at different times. Research on the outcomes of refugee children in Sweden shows that growing up in a neighbourhood with a large immigrant population increases the probability of engaging in crime, while there are no negative effects on school performance at age 16 (Grönqvist *et al*., 2015). As such, the findings of the effects of segregation in Sweden are in line with those of Billings *et al*. (2014). Neighbourhood exposure to refugee peers does not seem to be negative in all respects, however. Åslund *et al*. (2011) find that a larger ethnic community, that is, a larger community of individuals from the same country (or region) of origin, is beneficial for school performance. In addition, the positive effect of the ethnic community is increasing the more highly skilled are its members. One way of interpreting these results is therefore that high concentrations of the most disadvantaged groups in general is detrimental and implies increased risks for criminal activity, while sorting specifically by ethnic groups might be beneficial for refugees.[21]

Second, moving beyond the effects of segregation on individual outcomes, there is also a general concern that social cohesion in society is adversely affected if interactions between children from different backgrounds are limited (Levin 1998). This argument is discussed further in Blomqvist and Rothstein (2000) who argue that that integration of different groups is beneficial to foster tolerance and solidarity, and to discourage discrimination, inequality and violence in society.

To sum up, the research presented above has shown that sorting of pupils by family background, with large concentrations of the disadvantaged (minority or immigrant) group in some schools or neighbourhoods increases crime. There is thus cause for concern that a high level of segregation can have detrimental effects for individuals as well as for society.

If the policy aim is to provide equal access to the independent school sector for all children regardless of their family background, and

[21] Segregation and differentiation by ability may also affect what pupils learn in school through other mechanisms. For example, if achievement targets and teachers' expectations are adjusted to the group average, differentiation may increase inequality in educational performance (Figlio and Stone 2004; Bonnesrönning 2008).

to limit school segregation in order to avoid highly disadvantaged clusters, what are the consequences for policy?

To begin with, it is a challenge to design a school choice mechanism that allows for parental choice but at the same time limits sorting and quality differences across schools. The Swedish model has explicitly aimed at equal access for all pupils: independent schools cannot charge tuition fees, nor can they cherry-pick the best pupils. However, an international outlook shows that there are other models than using date of application, as in Sweden, to allocate pupils to oversubscribed schools. In this regard, there is room for a policy debate that opens up to alternatives. International examples show that school choice can be combined with re-distributive vouchers, such as in the Netherlands, and that the mechanism to assign pupils to schools can include minority quotas or lotteries, such as in the U.S.

Next, the overall trends in residential segregation show that school segregation must be tackled not only by providing equal access to independent schools. Examples of de-segregation policies in the U.S. that aim to counterbalance residential segregation include busing and minority quotas. Recent experiences of failing schools in disadvantaged areas in Sweden have led to similar actions: the closing of a school in Rosengård (Malmö) and busing of its pupils to a number of other schools in the city. Ultimately, residential segregation, refugee placements and catchment area boundaries constitute policy areas that are all relevant for school segregation. It should be noted that if the policy objective is to limit sorting and to create more heterogeneity by parental background, alternatives for admitting pupils to public schools must also be discussed.

7.7 Discussion

This paper has summarised evidence from the Nordic countries on the role of the education system for equality of opportunity. The comprehensive school reforms that were rolled out in the Nordic countries in the 1950s–1970s aimed to reduce inequalities in educational outcomes, and the research evidence shows that the reforms were successful in this regard: parental background became less important for labour market outcomes for the cohorts that went to the comprehensive school.

More recently, Sweden has undertaken reforms to introduce parental choice and a universal voucher scheme, and as a result many new independent voucher-funded schools have opened. These schools have pedagogical, religious or general profiles, and can be run as a non-profit foundation or a for-profit corporation. While the comprehensive school reforms aimed to unify the curriculum and the length of compulsory schooling, the choice and voucher reform can be considered as a step towards more heterogeneity in educational provision. Although this heterogeneity does not resemble the selective, early tracking system pre-dating the comprehensive school, it will potentially imply that pupils are exposed to different "types" of schools and to differences in pedagogical profiles and school governance. It has been hypothesised that this educational variety, in combination with parental choice, provides a setting where parental background will become more important for school performance. The recent studies on this topic show that there is little evidence to support this argument: family background factors remain strong predictors for school performance, but have not become more important as school choice has grown more common. Using school performance as a proxy for future labour market outcomes, there is no indication that social mobility is declining for the cohorts that left compulsory education up to 15 years after the choice reform was introduced. In addition, there is no evidence supporting the notion that school choice and competition can explain the plummeting Swedish PISA scores (Böhlmark and Lindahl, 2015, Wondratschek *et al*., 2013).

Education policy with the aim to promote equality of opportunity is nevertheless facing challenges for the future. The Swedish example shows that segregation has been increasing for a long time, and current

migration flows are highlighting the importance of policies that limit high concentrations of socioeconomically disadvantaged or foreign-born pupils in schools. It is also evident that access to independent schools is not equal for pupils of different family background. There is a social gradient in school choice. The comprehensive school abandoned ability tracking, and today's independent schools are not allowed to admit pupils based on ability tests. Selection on ability is thus considered as an unfair mechanism that creates inequality. From an equality of opportunity perspective however, one can argue that (de facto) selection on parental background is equally unfair, and it is not clear why one type of selection is considered acceptable and the other one not.

References

Aakvik, A., Salvanes, K. G. & Vaage, K. (2010). Measuring heterogeneity in the returns to education using an education reform. *European Economic Review*, 54, pp. 483–500. http://dx.doi.org/10.1016/j.euroecorev.2009.09.001

Amermüller, A. (2005). Educational opportunities and the role of institutions. *Centre for European Economic Research*, Discussion paper 05–44.

Andersson, E., Malmberg, B., & Östh, J. (2012). Travel-to-school distances in Sweden 2000–2006: Changing school geography with equality implications. *Journal of Transport Geography*, 23, pp. 35–13. http://dx.doi.org/10.1016/j.jtrangeo.2012.03.022

Arendt, J. N. (2005). Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review*, 25, pp. 149–160. http://dx.doi.org/10.1016/j.econedurev.2004.04.008

Åslund, O., Edin, P.-A., Fredriksson, P. & Grönqvist, H. (2011). Peers, neighbourhoods, and immigrant student achievement: Evidence from a placement policy. *American Economic Journal: Applied Economics*, 3, pp. 67–95. http://dx.doi.org/10.1257/app.3.2.67

Becker, G. S. & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6), pp. 1153–1189. http://dx.doi.org/10.1086/260831

Betts, J. R. (2011). The economics of tracking in education. In E. A. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, vol. 3, Elseveier. http://dx.doi.org/10.1016/b978-0-444-53429-3.00007-7

Billings, S. B., Deming, D. J., & Rockoff, J. E. (2014). School segregation, educational attainment and crime: Evidence from the end of busing in Charlotte-Mecklenburg. *Quarterly Journal of Economics*, 129(1), pp. 435–476. http://dx.doi.org/10.1093/qje/qjt026

Björklund, A. & Jäntti, M. (1997). Intergenerational income mobility in Sweden compared to the United States. *American Economic Review*, 87(5), pp. 1009–1018.

Björklund, A., Eriksson, T., Jäntti, M., Raaum, O., & Österbacka, E. (2002). Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. *Journal of Population Economics*, 15, pp. 757–772. http://dx.doi.org/10.1007/s001480100095

Björklund, A., Lindahl M., & Sund, K. (2003). Family background and school performance during a turbulent era of school reforms. *Swedish Economic Policy Review*, 10, pp. 111–136.

Björklund, A. Jäntti, M., & Lindquist, M. J. (2009). Family background and income during the rise of the welfare state: Brother correlations in income for Swedish men born 1932–1968. *Journal of Public Economics*, 93, pp. 671–680. http://dx.doi.org/10.1016/j.jpubeco.2009.02.006

Björklund, A. & Salvanes, K. G. (2011). Education and family background: Mechanisms and policies. In Hanushek E. A., S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, vol. 3, Elseveier. http://dx.doi.org/10.1016/b978-0-444-53429-3.00003-x

Björklund, A. & Jäntti, M. (2012). How important is family background for labour-economic outcomes? *Labour Economics* 19, pp. 465–474.

Blanden, J. (2013). Cross-country rankings in intergenerational mobility: A comparison of approaches from Economics and Sociology. *Journal of Economic Surveys*, 27(1). pp. 38–73. http://dx.doi.org/10.1111/j.1467-6419.2011.00690.x

Blomqvist, P. & B. Rothstein (2000). *Välfärdsstatens nya ansikte – demokrati och marknadsreformer inom den offentliga sektorn*. Agora förlag, Stockholm.

Bonesrönning, H. (2008). The effect of grading practices on gender differences in academic performance. *Bulletin of Economic Research*, 60(3), pp. 245–264. http://dx.doi.org/10.1111/j.1467-8586.2008.00278.x

Booij, A. S., Leuven, E., & Oosterbeek, H. (2015). Ability peer effects in university: Evidence from a randomized experiment. *IZA Working paper No. 8769*.

Brunello, G. & Checchi, D. D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy*, 22(52), pp. 781–861. http://dx.doi.org/10.1111/j.1468-0327.2007.00189.x

Böhlmark, A. & Holmlund, H. (2012). *Lika möjligheter? Familjebakgrund och skolprestationer 1988–2010*. IFAU report 2012:14.

Böhlmark, A., H. Holmlund & Lindahl, M. (2015). *School choice and segregation: Evidence from Sweden*. IFAU working paper 2015:8.

Böhlmark, A. & Lindahl, M. (2015). Independent schools and long-run educational outcomes – evidence from Sweden's large scale voucher reform. *Economica* 82, pp. 508–551. http://dx.doi.org/10.1111/ecca.12130

Chetty, R., Hendren, N., & Katz, L. (forthcoming). The effects of exposure to better neighbourhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review.*

Deming, D. J. (2011). Better schools, less crime? *Quarterly Journal of Economics*, 126, pp. 2063–2115. http://dx.doi.org/10.1093/qje/qjr036

Duflo, E., P. Dupas & Kremer, M. (2011). Peer effects, teacher incentives and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101, pp. 1739–1774. http://dx.doi.org/10.1257/aer.101.5.1739

Epple, D., Newlon, E. & Romano, R. (2002). Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics*, 83, pp. 1–48. http://dx.doi.org/10.1016/S0047-2727(00)00175-4

Erikson, R. & Goldthorpe, J.H. (1992). *The constant flux – A study of class mobility in industrial societies*, Clarendon, Oxford, England.

Erikson, R. (1996). Explaining Change in Educational Inequality – Economic Security and School Reforms. In Robert Erikson and Jan O. Jonsson (eds.), *Can Education Be Equalized?*, Westview Press.

Figlio, D. N. & Lucas, M. E., (2004). Do high grading standards affect student performance? *Journal of Public Economics*, 88(9–10), pp. 1815–1834. http://dx.doi.org/10.1016/S0047-2727(03)00039-2

Fisher, M., Karlsson, M., & Nilsson, T. (2013). Effects of compulsory schooling on mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health*, 10, pp. 3596–3618. http://dx.doi.org/10.3390/ijerph10083596

Fredriksson, P. & Vlachos, J. (2011). *Reformer och resultat: Kommer regeringens utbildningsreformer att ha någon betydelse?* Report to the Swedish Fiscal Policy Council 2013/3.

Government Bill 1991/92:95. *Om valfrihet och fristående skolor*.

Government Bill 1994/95:157. *Ändring av bidragsbestämmelserna för fristående skolor godkända för vanlig skolplikt m.m.*

Grönqvist, H., Niknami, S., & Robling, P.-O. (2015). Childhood exposure to segregation and long-run criminal involvement. SOFI Working paper 1/2015.

Gustafsson, J. E. & Yang-Hansen, K. (2009). *Resultatförändringar i svensk grundskola. Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer*. Stockholm, Skolverket.

Hanushek, E. A. and Woessmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal*, 116, pp. C63–C76. http://dx.doi.org/10.1111/j.1468-0297.2006.01076.x

Havnes, T. & Mogstad, M. (2011). No child left behind: Subsidized child care and children's long run outcomes. *American Economic Journal: Economic Policy* 3, pp. 97–129. http://dx.doi.org/10.1257/pol.3.2.97

Holmlund, H. (2007). *A researcher's guide to the Swedish compulsory school reform*. SOFI Working paper 9/2007.

Holmlund, H. (2008). *Intergenerational mobility and assortative mating: Effects of an educational reform*. CEE Discussion paper 91.

Holmlund, H., M. Lindahl and E. Plug (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, 49(3). pp. 615–651. http://dx.doi.org/10.1257/jel.49.3.615

Holmlund, H., Hägglund, J., Lindahl, E., Martinson, S., Sjögren, A., Vikman, U., & Öckert, B. (2014). *Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola*. IFAU report 2014:25.

Levin, H. (1988). Educational vouchers: Effectiveness, choice and costs. *Journal of Policy Analysis and Management*, 17(3), pp. 373–392. http://dx.doi.org/10.1002/(SICI)1520-6688(199822)17:3%3C373::AID-PAM1%3E3.0.CO;2-D

Meghir, C. & Palme, M. (2005). Educational reform, ability and family background. *American Economic Review*, 95:1, pp. 414–424. http://dx.doi.org/10.1257/0002828053828671

National Agency for Education (Skolöverstyrelsen) (1969). *Läroplan för grundskolan Lgr69*. Skolöverstyrelsen, Stockholm.

National Agency for Education (Skolverket) (2003). *Valfrihet och dess effekter inom skolområdet*. Report 230.

National Agency for Education (Skolverket) (2009). *Vad påverkar resultaten i svensk grundskola?* Skolverket, Stockholm.

National Agency for Education (Skolverket) (2012). *Likvärdig utbildning i svensk grundskola? – En kvantitativ analys av likvärdighet över tid*. Report 374.

National Agency for Education (Skolverket) (2013). *PISA 2012. 15-åringars kunskaper i matematik, läsförståelse och naturvetenskap*. Report 398.

Opposition bill 1991/92: Ub62. *Motion med anledning av prop. 1991/92:95 Om valfrihet och fristående skolor.*

Pekkarinen, T., Uusitalo, R. & Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, 93, pp. 965–973. http://dx.doi.org/10.1016/j.jpubeco.2009.04.006

Pekkarinen, T., Uusitalo, R. & Kerr, S. (2013). School tracking and the development of cognitive skills. *Journal of Labor Economics*, 31:3, pp. 577–602. http://dx.doi.org/10.1086/669493

Solon, G. (2004). A model of intergenerational mobility variation over time and place. In M. Corak (ed.), *Generational income mobility in North America and Europe*. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/cbo9780511492549.003

SOU 2013:56. *Friskolorna i samhället.* Fritzes, Stockholm.

SVT (2013). *Friskolor väljer bort besvärliga elever*. Downloaded from http://www.svt.se/ug/friskolor-valjer-bort-besvarliga-elever. Published 28 October 2013. Downloaded 28 January 2013.

Swedish Association of Independent Schools (Friskolornas riksförbund) (2015). *Fakta om friskolor.*

Waldinger, F. (2007). *Does ability tracking exacerbate the role of family background for students' test scores?* Unpublished manuscript.

Wondratschek, V., Edmark, K., & Frölich, M. (2013). The short and long-term effects of school choice on student outcomes – Evidence from a school choice reform in Sweden. *Annals of Economics and Statistics*. 111–112, pp. 71–102.

8. Retirement and Health in the Nordic Welfare State[1]

Nabanita Datta Gupta, Department of Economics and Business Economics, Aarhus University, ndg@econ.au.dk and Bent Jesper Christensen, Department of Economics and Business Economics, Aarhus University, bjchristensen@econ.au.dk

Summary

In Nordic welfare state countries, a multitude of exit options with relatively high replacement rates are available for workers approaching retirement age. We review the recent literature on retirement and health in Nordic countries, with main focus on Denmark, over a period characterized by reforms of the welfare system, paying special attention to eligibility to early retirement programs, disability pension, and wage-subsidized jobs. We collect evidence on two main questions: 1) Are there substantial moral hazard effects (ex ante and ex post) in claiming various retirement and health related social benefits in the Nordic Welfare State? 2) Can early retirement lead to (mental) health deterioration? The analyses reported on make use of comprehensive register data on individuals, linking their labor market characteristics to medical diagnoses.

[1] The authors gratefully acknowledge comments and suggestions from Tarmo Valkonen (the discussant) and seminar participants at the 2015 Nordic Economic Policy Review Conference in Helsinki, October, 2015, and from an anonymous referee.

8.1 Introduction

In Nordic welfare state countries, a multitude of retirement options exist for older workers which potentially reinforce the effect of health deterioration on retirement. The existence of various early retirement and disability exit options coupled with universal health care regime and low out-of-pocket medical expenditures may imply that older workers when faced with a health shock have a strong incentive to reduce their labor force participation. Even if an individual remains able to work, labor market exit may be a tempting choice if benefits are sufficiently generous relative to earnings from continued work, particularly when figuring in utility from increased leisure time. A moral hazard problem may arise, with the worker exaggerating the severity of the health condition in order to obtain benefits, failing to report on recovery from an earlier condition leading to benefits, or before health problems arise engaging in risky behavior and life style choices increasing the risk of a health shock, knowing the generosity of the system in place. Further, moral hazard may be affected by social norms (when is it acceptable to opt out?) and justification effects (a health issue may be claimed as an excuse for exit).

Fuelled by concerns about such over-exploitation of the system, Nordic governments over the past two decades have strengthened incentives to retain older workers, made disability screening procedure more stringent, expanded wage subsidized jobs substantially, and started gradually shortening the early retirement period. Some have viewed the reforms as problematic, but if the moral hazard problem is substantial, the reforms are probably justified, schemes having been on the generous side in the past. Therefore, the first aim of this paper is to shed light on whether there are significant moral hazard effects in claiming various retirement and health related social benefits in the Nordic Welfare Model.

An additional rationale for retaining older workers is the continuing increase in longevity. As the population ages, dealing with mental deterioration will add substantially to health care costs. Indeed, it may be that withdrawal from work can lead to deterioration of mental health. If staying longer on the labor market is a way to stave off cognitive de-

cline, there are additional benefits to reforms increasing statutory retirement ages, aside from the need to cope with aging and demographic change per se. Therefore, the second aim of this paper is to investigate whether there are negative health effects of early retirement, in terms of deterioration of cognitive capacity.

We review the literature on health and retirement in the Nordic welfare state, with main focus on Denmark, over a time period characterized by reforms of the welfare system. We pay special attention to eligibility to early exit programs, disability pension, and wage subsidized jobs. The analyses reported on in this paper make use of comprehensive register data on individuals, linking their labor market characteristics to medical diagnoses, in contrast to many studies in this area which are limited to using self-reported health and survey data.

In what follows, we describe in Section 1 the numerous Nordic welfare state institutions relevant for retirement and health. In Section 2, we review the evidence of moral hazard induced retirement in Nordic countries, with a particular focus on Denmark. In Section 3, we use Danish register data to conduct new analyses of the effect of retirement on diagnosed cognitive decline. In Section 4, we report on recent policy initiatives undertaken in Sweden and Denmark to encourage individuals with some remaining working capacity to remain on the labor market. Finally, in Section 5, we offer a brief conclusion.

8.2 Welfare state institutions that insure against health-related loss of working capacity

The Nordic welfare states share in common an elaborate tax-financed social safety net in the event of a health shock. When a worker suffers health-related loss of working capacity, depending on whether the loss is temporary or permanent, the relevant compensatory institutions are sickness absence, wage subsidy programs, disability pension, and, to some extent, early retirement. Furthermore, health services are provided in a setting of universal health insurance and largely public health care delivery decentralized at the level of the region, but

moving towards greater consolidation following the economic crisis (Saltman *et al*., 2012).

Sickness absence programs compensate employers for the period the worker is unable to perform due to injury/illness. All Nordic countries offer quite generous sick leave with maximum period of benefit varying between a year (Norway, Finland, and Denmark) to 2.5 years (Sweden) (Thorsen *et al*., 2015). There is some co-insurance element in that employers are responsible for the compensation for a part of this period while the state and local authorities cover the rest. It is usually the municipalities within the Nordic system that are responsible for administering benefit programs, following up on health assessments, and initiating various return-to-work (RTW) measures. In Denmark, the municipalities are responsible for vocational rehabilitation programs including workplace accommodation, working aids, counselling, re-education/re-training offers, etc. (see Bach *et al*., 2007). Harmonized data from the European Union Labor Force Surveys show that long-term sickness absence is high in Sweden and Norway, but low in Denmark and Iceland, with Finland in between, while the pattern is reversed for short-term sickness absence (Thorsen *et al*., 2015).

In recent years, all the Nordics have implemented a series of measures to strengthen RTW of the sick-listed, including workability assessments, more frequent follow-up assessments, and partial sick leave. Particularly the latter seems to be related to increased workability of persons with disability (Thorsen *et al*., 2015). In Denmark, if ordinary return to work is not possible, workers can return to wage-subsidized jobs (Flexjobs) involving flexible working conditions, shorter hours, etc., where employers are compensated 50–67% of the wage depending on the extent of working capacity loss. The introduction of Flexjob in Denmark has been shown to improve the employability of disabled workers in the 35–44 years age group by 10.5–12.5 percentage points (Datta Gupta and Larsen, 2010a).

The increasing efforts on getting sick-listed workers back to work as quickly as possible may not be optimal from a welfare point of view if it means that more and more workers report to work while still sick or in pain, as this increases the disutility from work. On the other hand, some degree of presenteeism (reporting to work while sick or in pain) could

be beneficial. Randomized trials from Finland and Norway have shown that low back pain, which is the most prevalent reason for sickness absence, may actually heal faster with light activity. Markussen *et al*. (2012) show based on Norwegian data that sick-listed employees who are issued graded (partial) absence certificates from their physicians actually experience shorter absence spells and higher employment rates two years after than those issued full certificates. A similar therapeutic effect of work may be present for some less serious mental disorders (Waddell and Burton, 2006).

When workers suffer a permanent injury or illness they are eligible for disability pension. Like most other Western industrialized nations, the Nordics also have relatively high rates of disability pensioning (seen in relation to objective measures of health), with about 5–10% of the 16/18–64/66 population receiving disability pension in 2000–2013 (as for sickness absence, the rate for women exceeds that for men) (Thorsen *et al*., 2015). The high disability rates are also remarkable in relation to the low unemployment rates of these countries. Together with the Netherlands, the Nordics rank top six in terms of disability shares out of 28 countries with comparable data (OECD, 2009b). Remarkably, there has been little change in this share over the period for most Nordics, the exceptions being rising disability in Iceland after 2009 and falling disability in Sweden after 2007. One study shows that for Denmark, take-up of disability seems to be on its own track and quite unrelated to a number of health indicators (Bingley *et al*., 2014). The compensation under disability in most cases is a flat-rate benefit, although in Sweden it is a function of previous earnings. The implication of a flat-rate benefit in disability and other pension programs in most Nordic countries is that the replacement rate is comparatively higher for the less-educated. Most Nordic countries have redefined eligibility for disability pension in recent years in terms of strict medical criteria and have expanded partial disability provisions, but have not altered the generosity of pension benefits (Andersen, 2011). Still, these reforms have most likely kept disability from rising even further, and have seemingly reversed the trend in Sweden even as the population continues to age.

Early retirement before the age of 65 in Nordic countries is prevalent in Denmark and Finland due to the existence of comprehensive

public early retirement schemes. In 1979, a voluntary early retirement program (*Efterløn*) available at age 60 to workers who had contributed to an unemployment insurance (UI) fund for a sufficient length of time was introduced in Denmark. The rationale given at the time for its introduction was the need to ameliorate the youth unemployment problem and to give workers in physically demanding jobs an exit route out of the labor market without having to satisfy formal medical criteria. One year after, in 1980, the labor force participation rate of men aged 60–64 dropped 20 percentage points and remained at this lower level for the next two decades (Bingley *et al*., 2012). Following the Retirement Reform in 2011, the program has been restricted substantially. In the future, the duration of benefits is shortened from 5 to 3 years, along with considerable implied deductions in pension assets, implying the de facto cancellation of the program in due course (Andersen, 2011).

In Finland, too, access to early retirement is no longer possible for cohorts born after 1943 (Finnish Pension Reform of 2005). While Denmark and Finland have historically had the highest rates of early retirement, Iceland has had the lowest rate among the Nordics until the financial crisis hit, after which the share has become more variable. Sweden and Norway fall somewhere in between. There are larger differences in retirement patterns than in sickness absence and disability patterns due to somewhat different structuring of the pension systems and differing relative importance of occupational pensions. Nonetheless, all the Nordic countries have now embraced automatic indexation of pensions to longevity, which ought to improve the future sustainability of the system (Andersen, 2011).

As funded pension plans (occupational and private) grow in importance in the Nordic countries, a question for the future is whether statutory retirement programs will become irrelevant as people begin to retire when they have built up sufficient pension funds, and what the resulting implications of such behavior will be for labor supply and public finances. A study from Denmark shows that individuals are increasingly beginning to combine work and retirement, i.e., partial retirement (Larsen and Pedersen, 2012). Earnings-related pensions systems are also undergoing reforms currently. The 2005 reform of the Finnish private-sector earnings-related pension system curbed the increases in the

contribution rate without hurting replacement rates and simulations show the reform will lead to a postponing of retirement (Lassila and Valkonen, 2006).

8.3 Is there evidence of moral hazard induced retirement?

Not all retirement options are freely available to a worker at any point in time. For example, disability pension is reserved for cases of permanent injury or illness. As already alluded to, this implies an incentive on the part of the worker to misrepresent the true health condition, i.e., a moral hazard problem (Rothschild and Stiglitz, 1976, Diamond and Mirrlees, 1978). This may not be a case of consciously trying to cheat the system, in the sense of insurance fraud involving a completely false claim by a healthy individual with purely financial motives, but could be a case of exaggerating an actual health problem, in order to increase the probability of obtaining the desired pension and leisure. Further, the moral hazard problem has to be viewed in the context of evolving social norms regarding how to use the system in place. Thus, the decision about when it is acceptable to exit may depend on the local participation rate, an idea pushed in the literature on reference points in the labor decision (Lindbeck *et al*., 1999). Similarly, in the case of sickness absence, Lindbeck *et al*. (2011) present evidence from Swedish full population panel data of geographical differences consistent with local variation in social norms, but unexplained by socioeconomic differences. A justification issue may arise, as well, with individuals overstating their health problem after receiving the benefit, as an explanation or excuse. Nevertheless, regardless the exact motives for the potential over-exploitation of the system, the phenomenon has caused authorities to implement increasingly rigorous eligibility tests, e.g., requiring specific medical conditions to be met for disability, to prevent free-riding.

While false or overstated reports of injury, illness, or other medical conditions for the desired retirement program may be classified as the results of ex post moral hazard (i.e., the claim is that the condition for admission to the program has already been met), another mechanism

known as ex ante moral hazard may be in play, as well. According to this, individuals may be less careful at work and engage in more risky activities, simply because of the existence of the retirement program insuring their continued income in the event of a loss of work capacity caused by the hazardous behavior. In this case, the strict medical criterion, e.g., for disability retirement, may truly be met at the time of the claim, i.e., there is no misreporting, but it is debatable to what extent admission to the program is deserved or socially desirable.

The usual solution to the ex ante moral hazard problem is partial insurance, e.g., a lower replacement ratio. This is no longer optimal when taking into account the simultaneous presence of ex ante and ex post moral hazard (Mookherjee and Png, 1989, Bond and Crocker, 1997). A main problem is costly state verification, i.e., the government is unable to catch misreports in a cheap manner. Indeed, the more difficult it is to detect the true health condition, the greater the incentive to misreport. The upshot is overinsurance for small losses and underinsurance for large, i.e., most people pay too much to keep the system in place, and those who really need it receive insufficient benefits. These issues have been studied, e.g., in the case of workers' compensation and workplace injuries in the U.S. (Butler *et al*., 1996) and Canada (Bolduc *et al*., 2002). Leth-Petersen and Rotger (2009) study whiplash claimants in Denmark. Diagnosing lasting disorders associated with whiplash (e.g., from a rear-end car collision) is difficult, and hence the incentive to misreport the condition should be high. The results show that claimants who are compensated fail to regain their previous earnings levels, thus suggesting little misreporting. On the other hand, about half the claims are turned down, and these claimants do in fact on average regain their pre-injury earnings levels. This indicates that misreporting and thus ex post moral hazard exists in this group, and that the claims system is capable of revealing this. Ex ante moral hazard may be present, too, in that those compensated used more health care in the year prior to the injury, consistent with poorer general health and perhaps risky behavior. The similar issues apply in the retirement case.

Different forms of insurance against health-related loss of work capacity are subject to different screening mechanisms, and different degrees of moral hazard problems. Sickness absence programs involve

follow-up health assessments, and disability programs include strict medical eligibility criteria. On the other hand, early retirement eligibility criteria typically do not involve health conditions. Furthermore, an additional exit route may present itself to the individual strongly determined to leave work in favor of paid leisure, namely, unemployment accompanied by the associated UI benefits. Thus, the issue arises whether unemployment may in some cases serve as a voluntary route to retirement, as well, and, if so, whether take-up in this case is induced by deteriorating health.

Some evidence on this is presented by Christensen and Kallestrup-Lamb (2012). This study is based on merged register data from Denmark on individual objective medical ICD diagnosis codes upon hospitalization and labor market and socioeconomic variables. Workers are followed from age 50 until labor market exit through either of a number of distinct routes. Duration models are fit to explain the time elapsed until exit, controlling for time-varying conditioning variables including gender, marital status, education, experience, occupation, financial variables such as income and wealth, and a host of diagnosis-based health shock indicators, including diseases of the circulatory, respiratory, musculoskeletal, nervous, digestive, and genitourinary systems, cancer, nutritional and metabolic diseases, mental disorders, and injuries. The five exit routes considered are disability (DI); early retirement (ER) immediately preceded by work; unemployment (UI) followed by ER; UI followed by other retirement states; and other schemes (e.g., civil service pension). Durations are censored at the old age pension (OAP) age, 67 for the cohort under study, aged 50 in 1985. For this age group, many UI spells are followed by one of the retirement states (DI, ER, OAP, or other), and hence this sort of terminal unemployment is considered a form of labor market exit.

During the study period, UI benefits could generally extend up to a maximum of 4 years. The ER age was 60, contingent on eligibility criteria including having paid into a UI fund for sufficiently long. Of particular interest for the moral hazard question is a rule that was in place during the period and that waived the 4 year maximum for those unemployed workers between ages 55 and 59 who would become eligible for early retirement by age 60. In effect, individuals wishing to exit the la-

bor market could do so already by age 51 by becoming voluntarily unemployed, getting the waiver from the 4-year limit to UI benefits at age 55 and thus continuing on benefits until transiting to (official) ER at age 60. This route was clearly illegal, as unemployed workers must remain available to the labor market, but it would appear attractive to individuals wishing to exit, hence posing a moral hazard problem. Again, this should be seen in the context of social norms, i.e., unemployment may be a social norm (Clark, 2003), in that the decision whether to pursue this combined UI-ER exit route may depend on the behavior of peers that the worker uses as reference points.

Whether individuals observed to follow the UI-ER route in question initially became unemployed by chance, or with a plan to remain unemployed until ER, is of course not directly observable. Instead, in a competing risk framework, with the five exit routes (DI, ER, UI-ER, UI-other, other) representing the competing risks, Christensen and Kallestrup-Lamb (2012) compare the routes in terms of the effects of explanatory variables, including health shocks. In this sense, the UI-ER route is similar to the ER route per se, whereas UI-other is not (actually, it is more similar to DI). This suggests that there are differences between UI spells, even these terminal spells leading into ER or other retirement, with some of the unemployment possibly being planned and voluntary (that leading to ER), and some rather being associated with the type of health shocks that could also lead to disability.

In terms of the detailed results, increased wealth reduces exits through both the DI and UI-other routes, but has only an insignificant effect (in the opposite direction) on UI-ER exits, and indeed induces ER. Greater household income strongly increases DI and UI-other, i.e., the income effect dominates the substitution effect, but not so for ER or UI-ER. Naturally, DI is induced by many of the serious health conditions, such as diseases of the circulatory, respiratory, musculoskeletal, and nervous systems, cancer, nutritional and metabolic diseases, mental disorders, and injuries, but not by diseases of the genitourinary systems, which include, e.g., kidney stone, renal failure, and other diseases of the urinary and genital organs. In contrast, the latter group of diseases strongly increases exits through the UI-ER route, perhaps suggesting that selection of this combined route may not purely reflect suspicious

behavior, but also indicate health conditions that in fact lead to a loss of work capacity, without being sufficiently recognized by disability criteria in place. Interestingly, UI-ER exits are furthermore induced by endocrine, nutritional and metabolic diseases, significantly and more strongly so than the other exit routes. These diagnoses include life style diseases such as diabetes, obesity, etc. The results are consistent with risky behavior (ex ante moral hazard) leading to life style diseases and prolonged voluntary unemployment followed by early retirement.

Many other studies investigate issues of risk, insurance, and health from different perspectives. Datta Gupta, Kleinjans and Larsen (2015) compare the effects of health shocks on the probability of not working across elderly workers in Denmark and the U.S., using Danish registers and the Health and Retirement Survey (HRS), respectively. With universal insurance and nationalized health care in Denmark, compared to the job lock effect of employer-based health insurance and the stronger income effect stemming from higher out-of-pocket medical expenditures in the U.S., it is expected that Danish workers are more likely than their U.S. counterparts to exit the labor force following health shocks. However, the results show little difference across the two countries in this respect. This is shown to be a result of differences in mortality, baseline health, and the way health care is provided. Datta Gupta and Larsen (2010b) offer a further comparison across register and survey data of health effects on planned labor supply, here with both data sources from Denmark. This allows a direct measurement of the bias arising due to the survey responses being self-reported and thus potentially involving misreports. This could reflect moral hazard, or a justification bias may exist, as the severity of a health conditions is overstated to provide an excuse for early labor market exit. The results indicate that men's self-reports on myalgia and back problems and women's on osteoarthritis yield biased estimates of the impact on planned retirement age. This is consistent with strategic misreporting, i.e., the existence of insurance covering the retired state makes it worthwhile to misreport.

Further evidence on the economic incentives for misreporting of own health is offered by Datta Gupta *et al*. (2010), demonstrating in a multi-country comparison that it is exactly in welfare state countries with significant social transfers that workers are most likely to act op-

portunistically and falsely self-report disability, whereas this phenomenon is less marked in Southern European countries. Sweden represents the only deviation from this pattern, possibly due to the strictness of vocational assessments in disability cases in Sweden compared to other SHARE (Survey of Health and Retirement in Europe) countries (Börsch-Supan, 2007). This indicates a case of a successful state-verification mechanism.

Although welfare state countries clearly offer incentives to report that conditions for retirement are met, it is difficult to associate the labor supply response to health shocks with specific welfare state institutions. Datta Gupta and Larsen (2007) use Danish register data to show that the increase in retirement rate following a health shock is unaffected by eligibility to early exit programs (eligibility dummies are insignificant), by the long duration of sickness benefits (results are unchanged when including receipt of sickness benefits in the definition of the retirement state), and by the promotion of corporate social responsibility (CSR) since 1994 and subsidized employment (Flexjob) since 1998 (no significant difference between pre- and post-promotion estimates). This complicates policy choices. Thus, although increasing longevity calls for policies to retain workers longer, the best method for identifying such policies remains elusive, based on these results. Nevertheless, that there is scope for longer work lives in the future is established by Bingley *et al*. (2015) using SHARE data for Denmark. The health capacity to work longer is estimated by comparing how much older individuals work now with how much those (younger individuals) with similar mortality rates worked in the past, and with how much younger individuals with similar self-reports of health work now. Any bias due to moral hazard and misreporting may thus cancel when constructing the estimate based on this comparison or matching procedure.

8.4 Does early retirement bring about cognitive decline?

As the population in Nordic countries continues to age, dealing with the effects of mental deterioration will add substantially to health care costs. The Nordic Council estimates there will be a doubling of the number of people living with dementia in the Nordic countries over the next 35 years (Nordic Council, 2010). In Denmark alone, the Alzheimer's Association estimates that annual costs on treatment and care (direct plus indirect) exceed 15 billion kroner. Longer life expectancy coupled with early retirement programs tends to add to years spent out of the labor market. On the other hand, the employment rate in especially the 60–64 age group has been rising in most OCED countries, including the Nordics, and the recent spate of reforms may intensify this trend.

Can raising the retirement age ward off cognitive decline? A wave of studies has looked at this question, but employing survey-based tests of cognitive functioning as their outcome (Coe & Zamarro, 2008, Rohwedder and Willis, 2008, Bonsang *et al*., 2012, Mazzona and Peracchi, 2012, and Bingley and Martinello, 2013). We bring novel evidence from the comprehensive Danish registers using as outcome a more reliable measure of mental decline than unincentivized tests of word recall, namely, diagnosis of Alzheimer's disease (AD)/dementia made in the hospital. In the Danish system, specialist visits are only possible by GP referral and a diagnosis of AD/dementia usually requires scanning and tests performed by a specialist. Medical diagnoses, however, may be unrelated to functioning on the labor market, especially if the disease progresses slowly. Furthermore, getting a diagnosis may vary systematically with education or work status. Those who are retired have lower time costs of investing in their health (getting diagnosed). At the same time, they also face less of a monetary incentive to invest in their own health, so effects may cancel out.

We use data on the population of older males[2] observed in the labor force in 1998 and followed annually up to 2007. Information on the di-

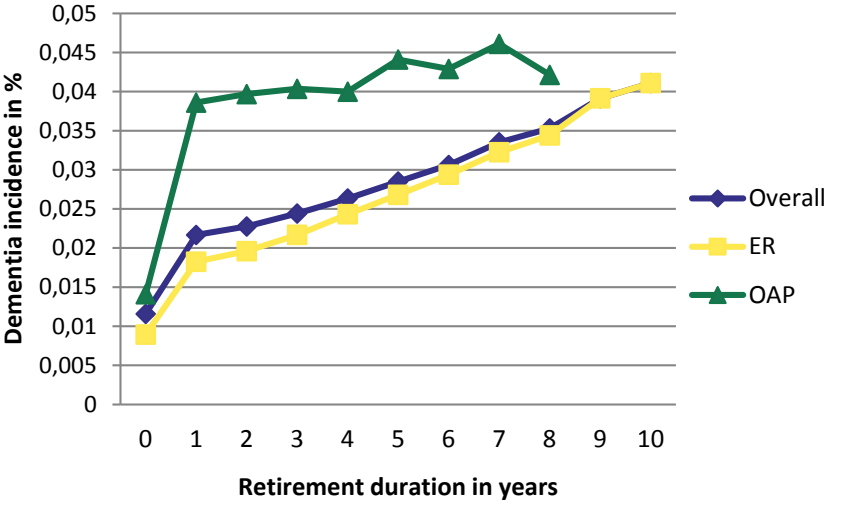
[2] For brevity we only show results for males. Results for females actually do not differ much.

agnosis of AD/dementia at the time of hospitalization is obtained from the National Patient Registry and linked to labor market registers.[3] We condition on absence of dementia/AD in 1998 and on being in the work force in 1998 (employed or unemployed). We control for age and education (completed basic schooling or not). We limit the sample to those individuals who in each year are observed to be either in the labor force or retired. That is, transitions to disability or other types of benefit schemes are excluded from the sample. The resulting sample consists of 3,449,037 person-year observations. Unlike much of the previous literature, we distinguish between early retirement (ER) and normal old-age pension (OAP). Basic descriptive statistics reveal that average sample age is 58, varying from 47 to 74. Dementia/AD incidence is 1.4%. This definition is purely based on hospital diagnoses and therefore does not include any diagnosis made in a doctor's office and not necessitating a hospital visit. Omitting the non-retired, retirement duration is on average 3.7 years for ER and 3 years for OAP. 57% of the sample has completed their basic schooling.

Figure 1 shows the incidence of dementia/AD according to retirement duration in the sample. The incidence obviously rises with retirement duration as people get older. ER retirees are a much larger subsample and therefore resemble the overall sample while OAP retirees have both higher incidence of dementia/AD and a sharper rise in this when going from work to retirement. By definition, OAP retirees are older when they retire, compared to ER retirees, and the diagram does not control for age. It is also important to be able to disentangle whether the abrupt rise among OAP pensioners when retiring indeed is the causal effect of going on retirement or whether it reflects reverse causation, i.e., people going on OAP when diagnosed with dementia/AD.

[3] In practice, the broader ICD category we access includes other mental and behavioral disorders, and nervous system disease, but many of these are rare in the population and do not exhibit increasing incidence with age.

Figure 1: Dementia incidence by retirement duration, males aged 47–74, 1998–2007



We estimate a fixed effects regression of the probability of obtaining a diagnosis of dementia/AD in the sample period as a function of retirement, a polynomial in age,[4] and education, including an individual time-constant effect, and an idiosyncratic error term. Comparison within individuals allows us to see whether the individual is more likely to obtain a diagnosis for dementia/AD after retirement. The fixed effects model removes all time-constant heterogeneity from the model: Indeed, education is most likely superfluous, as it rarely changes at older ages, although adult education is fairly prevalent in Denmark.

ER is available from age 60 in the sample period, although a financial incentive is in place to defer it to age 62, and OAP for most of the period from age 67, as in Section 3 (changed to 65 starting in 2004). Since the decision to retire is obviously endogenous, we instrument ER by being 60 or above, and OAP by being 65 or above, as clear spikes have been shown in the retirement hazard in Denmark over this time period at ages 60, 62, and 65 (Bingley *et al.*, 2014).

[4] A quadratic in age has been found to capture the age dependence of cognitive decline quite well, see Bonsang *et al.* (2012).

Table 1: Effect of retirement & retirement duration on dementia/ad diagnosis

Variable	FE_IV[a]	FE_IV[b]	FE[c]	FE[d]
Has begun early retirement	-0.0031*** (0.0006)			
Has begun normal OAP retirement		0.0037*** (0.0011)		
Log early retirement duration			-0.0043*** (0.0011)	
Log normal retirement duration				0.0268*** (0.0033)
Age	-0.0048*** (0.0002)	-0.0035*** (0.0003)	0.0051** (0.0020)	-0.0935*** (0.0112)
Age2/10	0.0007*** (0.0000)	0.0005*** (0.0000)	0.00001 (0.0000)	0.0067*** (0.0007)
Education above basic school	-0.0016 (0.0022)	-0.0020 (0.0021)	-0.0025 (0.0119)	0.0076*** (0.0025)
Constant	0.0689*** (0.0066)	0.0377*** (0.0078)	-0.8205** (0.3288)	3.2904*** (0.4047)
#Obs	3,363,437	2,874,615	574,422	90,159

Note: FE is fixed effects. FE IV is fixed effects with instrumental variables estimation.
[a] Sample includes all workers and all ER retirees, 1998–2007, 47 ≤ age < 74. Has begun ER retirement instrumented by Age ≥ 60.
[b] Sample includes all workers and all OAP retirees (not including ER retirees), 1998–2007, 47 ≤ age ≤ 74. Has begun OAP instrumented by Age ≥ 65. F tests show that both instruments are positive and very strong. F tests of individual effects show joint significance in all cases.
[c] Sample includes all ER retirees only with positive retirement duration, 1998–2007.
[d] Sample includes all OAP retirees only with positive retirement duration, 1998–2007.

Results of the regression analysis appear in Table 1 (standard errors in parentheses). In columns 1 and 2 we see that the rise in dementia incidence at retirement occurs mainly among OAP retirees. The column 1 sample includes all workers and all ER retirees. The column 2 sample includes all workers and OAP retirees only in each year (i.e., excluding ER retirees who would automatically transit to OAP at age 65). Thus, working individuals contribute to both samples. The size of the effect on

the cognitive decline of OAP retirees is 0.37 percentage points on a mean overall incidence of 3.3%, about 11% of baseline.

As argued in Rohwedder and Willis (2008), the effect of non-work on cognitive decline may appear only with a lag. Thus, we next identify the effect of retirement duration on dementia diagnosis given that one is retired, i.e., for the sample of retirees only. Retirement duration is measured in logs as in the previous literature (see Bonsang *et al*, 2012). Retirement duration is not instrumented for in this analysis, and endogeneity only controlled for via fixed effects. Arguably, the more severe selection is encountered for the decision of when to retire, and not for the duration of retirement, which is mainly driven by age and program structure.

Results appear in columns 3 and 4 of Table 1. Longer retirement duration is associated with cognitive decline, but again only evident in the sample of OAP retirees. The effect size is a 2.7 percentage point increase in dementia incidence for every log point increase in retirement duration. Since mean OAP duration is 3 years, for every 3.6 months increase in retirement duration, dementia incidence rises approximately 0.27 percentage points, equalling 8.6% of baseline, i.e., a strong effect.

To the extent that the policy variables which we use to instrument for retirement and the fixed effects specification together eliminate the endogeneity bias, the results show both an immediate increase in dementia incidence when going on normal (OAP) retirement, which is not just reverse causation, and a further increase as more time is spent on OAP.

For early retirees, the results actually show positive (protective) effects, both of going on ER, and of ER duration (although the effect size of duration is an order of magnitude smaller than for OAP). This result may be interpreted in a number of ways. First, ER retirees are younger when retiring, and so their cognitive reserves more intact. With time, ER retirees may well begin to show the same patterns of cognitive decline as normal retirees. It could also be that those diagnosed with dementia in the ER group transit to disability (DI) and hence disappear from the sample, leaving behind a healthier group. In a robustness test, individuals in the relevant age-interval who were on DI and were observed sometime during the period in the ER state were pooled with the sample in column 1, Table 3. This increased the sample size slightly, but

did not change the coefficient appreciably. It is possible that ER is not health related, whereas OAP is, meaning that people who work beyond ER age have stronger work preferences, and probably more health knowledge, and only stop working when they are diagnosed with mental decline. We also saw that the effect of retirement duration given that one was retired was stronger among OAP retirees, controlling for age and its square. Greater health knowledge might mean that OAP retirees visit the doctor more frequently, although some of this should be captured by the fixed effect. A further explanation can be that ER retirees are healthier when they retire, and are able to stave off cognitive decline by pursuing a more active and healthy lifestyle during retirement, whereas those in the OAP age group are not able to, because of poorer health status and comorbidities when they retire (we control for age, but not health). Relief from work-related stress and strain and more frequent exercise have been found to be the primary mechanisms through which retirement has a positive effect on health (Eibich, 2015). A final explanation for the findings could be that the nature of work is different for ER and OAP retirees – if the latter are giving up mentally demanding work, this could bring about faster cognitive decline for them during retirement. This is also an area that can be studied further.

8.5 Recent policy changes targeting disabled with remaining working capacity

Here, we consider recent policy initiatives undertaken in Sweden and Denmark to encourage individuals with some remaining working capacity to remain on the labor market.

8.5.1 Sickness absence and disability policy changes in Sweden since 2006

Sweden has traditionally had very high rates of sickness absence (it is the world leader in terms of long-term sickness absence). It is also one of the countries with the highest rates of inflow to disability from sickness absence. This has been due to relatively lenient eligibility criteria

and easy access to generous benefits (also temporary benefits), compared to the other Nordic countries. Since 2007, however, the trend in disability has been reversed (except among the youngest age group), and sickness absence has fallen (OECD, 2009, and Thorsen *et al*., 2015).

A series of reforms have been implemented by the Swedish government, establishing a "rehabilitation chain" with fixed time limits and regular reassessments. Among other things, the sickness absence period was reduced to one year, employers' copayment was increased, and permanent benefits replaced by temporary benefits coupled with work incentives for young and prime-aged workers. Effectively, disability pension would henceforth only be granted to those with a permanent reduction in working capacity. An important element of the reforms was to centralize the Social Insurance Agency (SIA), to increase its contact and cooperation with the Public Employment Service (PES), and to give county councils and regions financial incentives to improve the quality and efficiency of the sickness certification process (the so-called "Swedish sick leave billion", see Thorsen *et al*., 2015). In 2008, a rehabilitation guarantee targeted workers with mental disorders such as depression, anxiety, or stress, providing them with free cognitive behavioral therapies and employing multimodal rehabilitation efforts.

How much of the reduction in sickness absence and disability in recent times in Sweden can be attributed to the government's reforms? Hartman (2011) looks at the evolution of sickness absence and disability since the 1990s and points out that they began trending downward even before the current government came into office. Although evaluation studies of these reforms are still in the making, the review of a first batch of studies in Hartman (2011) indicates that the rehabilitation chain appears to have had a substantial impact, and that the transfer to the PES during the first year of sick leave appears to be functioning smoothly. However, in her opinion, the government moved too hastily to implement these reforms, and in the process, groups that were already in the system according to the old rules clearly lost out. The eligibility criteria for permanent disability may have become too stringent and thereby difficult for the oldest disabled to fulfill, and it is worrisome that the use of disability benefits among the young is increasing in Sweden, in parallel with grow-

ing academic requirements from school-leaving and vocational programs (Hägglund and Skogman Thoursie, 2010).

8.5.2 4.2. Partial sick leave and wage subsidy schemes in Denmark

Partial sick leave has been prevalent in many Nordic countries for decades – it offers sick-listed workers the possibility to return to work on a part-time basis. Finland was a relative latecomer, instituting partial sick leave in 2007 (Thorsen *et al*., 2015). The partial sick leave policy in Denmark has been evaluated by Høgelund, Holm and McIntosh (2010). Using both Danish register and survey data, this study examined the return to normal working hours of sick-listed workers covered by a national graded RTW program. The program was designed such that during work hours the workers receive the normal wage, and for the hours off work spent sick they receive sickness pay. The approach takes into account unobservable differences between 265 program participants and 669 non-participants, all of whom were sick-listed for at least eight weeks. The results show that graded RTW significantly increases the likelihood of returning to regular working hours. Similar evidence of the success of partial sick leave programs is found for other Nordic countries, too, by Andrén and Svensson (2012) for Sweden, Kausto *et al*. (2014) for Finland, and by Markussen *et al*. (2012) for the full Norwegian population.

While partial sick leave policies appear to increase the probability of RTW of the sick-listed, wage subsidy programs are potentially powerful for reducing unemployment of the disabled, by replacing passive social insurance payments by employer subsidies to hire disabled workers and improve their attachment to the labor force. The Danish Flexjob program is a unique subsidy program for the disabled that has been cited as a good example of a support scheme by the OECD, but only recently has it been reliably evaluated. Datta Gupta, Larsen and Thomsen (2015) use Danish register data and exploit a 2002 change in the reimbursement to government units (but not to municipal or regional units) within a difference-in-difference framework. A lowering of the subsidy from full to partial for governmental units only changed their hiring

practices to favor insiders over outsiders. This suggests that an increase in wage subsidies can reduce rehiring of previous employees who are disabled and increase the hiring of non-employed disabled.

8.6 Conclusion

The Nordic welfare states share in common an elaborate tax-financed system of income compensation in the event of a loss of working capacity. Benefits are generous and the coverage period long in an international context. The criteria for eligibility for such programs were also relatively lax in the past, but have recently been tightened up. Although this policy shift has been viewed as problematic by some, it may have been justified by the existence of moral hazard problems. Indeed, several studies reviewed here point to the existence of moral hazard in the Nordic context, which may manifest itself in an increased tendency to exaggerate health problems for the purpose of claiming disability (ex post moral hazard), and in risky behavior (ex ante moral hazard), e.g., leading to life style diseases and prolonged voluntary unemployment followed by early retirement. Clearly, misreports of health generate biases in estimates of the true effect of health on retirement. On the other hand, when health is more objectively measured, such as via hospital diagnoses, the effects on retirement are expected to be real.

Nordic countries are also experiencing a rise in longevity, and recent results show that there is scope for having people work longer in the future. We provide novel analyses based on comprehensive register data from Denmark, showing that retaining individuals on the labor market beyond age 65 has a causal effect on reducing their incidence of dementia/Alzheimer's disease. Interestingly, however, no such effect is present for early retirees. A number of possible reasons for this are considered, including that early retirees are younger when retiring, and their cognitive reserves more intact. With time, they may well begin to show the same patterns as normal retirees. Thus, beside coping with aging, there may be additional benefits to reforms retaining older workers on the labor market, in terms of reductions in cognitive decline.

One of the issues not touched upon in this paper is the interdependence of couples' retirement decisions. Any policy affecting retirement or disability may have spousal multiplier effects beyond the affected individual worker, and this may extend to concerns of moral hazard. For instance, Johnsen and Vaage (2015) found that the effect of the husband being eligible for ER is a 2.9 percentage points increase in the likelihood of the wife receiving disability pension. Similarly, spouses' retirement decisions, shared leisure, and lifestyle choices will affect both partners' physical and mental functioning.

Some of the most successful initiatives undertaken in recent times by Nordic governments to try to improve labor market attachment and return-to-work of the sick and disabled appear to be partial sick leave policies and wage subsidy programs, coupled with a tightening of disability criteria, shortening of the sickness absence period, frequent reassessments, employer co-payments, and close cooperation between disability councils and employment services. The evidence provided in this paper suggests that these changes have been necessary for curbing the outflow of workers with remaining stocks of working capacity from the labor market, thereby improving dependency ratios, labor market shortages, and public finances, while also impacting individual own health.

References

Andersen, J. G. (2011). *Ageing and welfare reform in the Nordic Countries, 1990–2010. Multipillar pensions and "activation" of social and labour market policies*. Aalborg University Working Paper.

Andrén, D. & Svensson, M. (2012). Part-time sick leave as a treatment method for individuals with musculoskeletal disorders. *Journal of Occupational Rehabilitation* 22, 418–426. http://dx.doi.org/10.1007/s10926-011-9348-7

Bach, H., Datta Gupta, N. & Høgelund, J. (2007). *Employment effects of educational measures for work-injured People*. IZA discussion paper 2657.

Bingley, P., Datta Gupta, N., Jørgensen, M. & Pedersen, P.J. (2014). Health, disability insurance and retirement in Denmark. In D. A. Wise (ed.), *Social Security Programs and Retirement Around the World: Disability Insurance Programs and Retirement.* forthcoming.

Bingley, P., Datta Gupta, N. & Pedersen, P.J. (2012). Disability programs, health, and retirement in Denmark since 1960. In D. A. Wise (ed.), *Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms*. University of Chicago Press, Chicago. http://dx.doi.org/10.7208/chicago/9780226921952.003.0006

Bingley, P., Datta Gupta, N. & Pedersen, P.J. (2015). *Health capacity to work at older ages in Denmark*. Aarhus University Working Paper.

Bingley, P. & Martinello, A. (2013). Mental retirement and schooling. *European Economic Review* 63, 292–298. http://dx.doi.org/10.1016/j.euroecorev.2013.01.004

Bolduc, D., Fortin, B., Labrecque, F. & Lanoie, P. (2002). Workers' compensation, moral hazard and the composition of workplace injuries. *Journal of Human Resources* 37, 623–652. http://dx.doi.org/10.2307/3069684

Bond, E. W. & Crocker, K.J. (1997). Hardball and the soft touch: The economics of optimal insurance contracts with costly state verification and endogenous monitoring costs. *Journal of Public Economics* 63, 239–264. http://dx.doi.org/10.1016/S0047-2727(96)01594-0

Bonsang, E., Adam, S. & Perelman, S. (2012). Does retirement affect cognitive functioning? *Journal of Health Economics* 31, 490–501. http://dx.doi.org/10.1016/j.jhealeco.2012.03.005

Börsch-Supan, A. (2007). *Work disability, health and incentive effects*. Sonder Forschungs Bereich 504, 07–23.

Butler, R. J., Durbin, D.L. & Helvacian, N.M. (1996). Increasing claims for soft tissue injuries in workers' compensation: Cost shifting and moral hazard. *Journal of Risk and Uncertainty* 13, 73–87. http://dx.doi.org/10.1007/BF00055339

Christensen, B. J. & Kallestrup-Lamb, M. (2012). The impact of health changes on labor supply: Evidence from merged data on individual objective medical diagnosis codes and early retirement behaviour. *Health Economics* 21, 56–100. http://dx.doi.org/10.1002/hec.2811

Clark, A. (2003). Unemployment as a social norm: Psychological evidence from panel data. *Journal of Labor Economics* 21, 323–351. http://dx.doi.org/10.1002/hec.2811

Coe, N. B. & Zamarro, G. (2011). Retirement effects on health in Europe, *Journal of Health Economics* 30, 77–86. https://ideas.repec.org/a/eee/jhecon/v30y2011i1p77-86.html, http://dx.doi.org/10.1016/j.jhealeco.2010.11.002

Datta Gupta, N., Kleinjans, K.J. & Larsen, M. (2015). The effect of a severe health shock on work behavior: Evidence from different health care regimes. *Social Science & Medicine* 136–137, 44–51. http://dx.doi.org/10.1016/j.socscimed.2015.05.008

Datta Gupta, N., Kristensen, N. & Pozzoli, D. (2010). External validation of the use of vignettes in cross-country health studies. *Economic Modelling* 27, 854–865. http://dx.doi.org/10.1016/j.econmod.2009.11.007

Datta Gupta, N. & Larsen, M. (2007). Health shocks and retirement: The role of welfare state institutions. *European Journal of Ageing* 4, 183–190. http://dx.doi.org/10.1007/s10433-007-0062-4

Datta Gupta, N. & Larsen, M. (2010a). *Evaluating labour market effects of wage subsidies for the disabled. The Danish Flexjob scheme*. SFI Working Paper 07:2010. http://www.sfi.dk/s%c3%b8geresultat-10668.aspx?Action=1&NewsId=2854&PID=32427

Datta Gupta, N. & M. Larsen (2010b). The impact of health on individual retirement plans: Self-reported versus diagnostic measures. *Health Economics* 19, 792–813. http://dx.doi.org/10.1002/hec.1523

Datta Gupta, N., Larsen, M. & Thomsen, L.S. (2015). Do wage subsidies for disabled workers reduce their non-employment? – Evidence from the Danish Flexjob Scheme. *IZA Journal of Labor Policy* 4:10, http://dx.doi.org/10.1186/s40173-015-0036-7.

Diamond, P. A., & Mirrlees, J. A. (1978). A model of social insurance with variable retirement. *Journal of Public Economics* 10, 295–336.

Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, 43, 1–12. http://dx.doi.org/10.1016/j.jhealeco.2015.05.001

Hartman, L. (2011). *Swedish policies for the disabled. What do we know about the effects?* Uppsala University Working Paper.

Hägglund, P., & Skogman Thoursie, P. (2010). *De senaste reformerna inom sjukförsäkringen: En diskussion om deras förväntade effekter* (The latest reforms within sickness insurance: A discussion of the expected effects). Studier i Finanspolitik 2010/5, Finanspolitiska Rådet.

Høgelund, J., Holm, A. & McIntosh, J. (2010). Does graded return-to-work improve sick-listed workers' chance of returning to regular working hours? *Journal of Health Economics* 29, 158–169. http://dx.doi.org/10.1016/j.jhealeco.2009.11.009

Johnsen, J.V. & Vaage, K. (2015). *Spouses' retirement and the take-up of disability pension*. University of Bergen Working Papers in Economics, No. 3/15.

Kausto, J., Viikari-Juntura, E., Virta, L.J., Gould, R., Koskinen, A. & Solovieva, S. (2014). *Effectiveness of new legislation on partial sickness benefit on work participation: A quasi-experiment in Finland*. BMJ Open 2014;4:e006685. http://dx.doi.org/10.1136/bmjopen-2014-006685

Larsen, M. & Pedersen, P. J. (2012). *Paid work after retirement: Recent trends in Denmark*. IZA Discussion Paper No. 6537.

Lassila, J & Valkonen T. (2006). The *Finnish pension reform of 2005, The Research Institute of the Finnish Economy (ETLA)*. Discussion Papers, No. 1000.

Leth-Petersen, S., & Rotger, G.P. (2009). Long-term labour market performance of whiplash claimants. *Journal of Health Economics 28*, 996–1011. http://dx.doi.org/10.1016/j.jhealeco.2009.06.013

Lindbeck, A., Nyberg, S., & Weibull, J. (1999). Social norms and economics incentives in the welfare state. *Quarterly Journal of Economics* 114, 1–35. http://dx.doi.org/10.1162/003355399555936

Lindbeck, A., Palme, M., & Persson, M. (2011). *Sickness absence and local benefit cultures*. IZA Discussion Paper No. 6245.

Markussen, S., Mykletun, A. & Roed, K. (2012). The case for presenteeism – evidence from Norway's sickness insurance program. *Journal of Public Economics* 96, 959–972. http://dx.doi.org/10.1016/j.jpubeco.2012.08.008

Mazzona, F., & Peracchi, F. (2012). Aging, cognitive abilities and retirement. *European Economic Review* 56, 691–710. http://dx.doi.org/10.1016/j.euroecorev.2012.03.004

Mookherjee, D., & Png, I. (1989). Optimal auditing, insurance and redistribution. *Quarterly Journal of Economics* 104, 205–228. http://dx.doi.org/10.2307/2937855

NOMESCO (2014). *Health Statistics in the Nordic Countries in 2014*. Nordic Medico-Statistical Committee (NOMESCO), 102:2014, Copenhagen.

Nordic Council (2010). *Crack the dementia code*. http://www.norden.org/en/news-and-events/news/crack-the-dementia-code.

OECD (2009a). *Sickness, disability and work: Breaking the barriers*. Sweden: Will the Recent Reforms Make It? OECD iLibrary. http://dx.doi.org/10.1787/9789264090606-en

OECD (2009b). *Sickness, disability and work – keeping on track in the economic downturn, background paper*. OECD.

Rohwedder, S. & Willis, R.J. (2010). *Mental retirement, Journal of Economic Perspectives 24*, 119–138. http://dx.doi.org/10.1257/jep.24.1.119

Rothschild, M., & Stiglitz, J. E. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics* 90, 629–649.

Saltman, R. B., Vrangbaek, K., Lehto, J. & Winblad, U. (2012). Consolidating national authority in Nordic health systems. *Eurohealth incorporating Euro Observer* 18, 21–24.

Thorsen, S.V., Friborg, C., Lundstrøm, B., Kausto, J., Örnelius, K., Sundell, T., Kalstø, Å. M., Thune, O., Gross, B.-O., Petersen, H. & Haram, Ö. (2015). *Sickness absence in the Nordic countries*. Nordic Social Statistical Committee (NOSOSCO), Copenhagen.

Waddell, G. & Burton, A. K. (2006). *Is work good for your health and well-being?* Report for Department for Work and Pensions, TSO (The Stationery Office), London.

9. The future of Welfare services:
How worried should we be about Wagner, Baumol and Ageing?[1]

Andreas Bergh, Research Institute of Industrial Economics (IFN), Box 55665, 102 15 Stockholm, Sweden

Abstract

Welfare services are an important part of the Nordic welfare states both financially and for welfare state redistribution. Baumol's cost disease, Wagner's law, and population ageing are often said to bring challenges for the future provision of welfare services. While none of the three poses an immediate threat against the financial sustainability of the welfare state, they have important implications for distribution and for the political support for the welfare state. The combination demographic change, a higher relative price of welfare services and increasing demand for welfare services may force politicians to make a difficult choice between increasing taxes, allowing people to top up publicly financed services with additional private financing, or risk eroding support for the welfare state.

[1] Financial support from the Swedish Research Council and Torsten Söderberg Foundation is gratefully acknowledged. The author also thanks Sebastian Jävervall for research assistance.

9.1 Introduction

In welfare state research, a distinction is often made between cash benefits and benefits in kind. The former consists of income transfers and social insurance schemes; the latter largely consists of a set of welfare services (such as schooling, health care and elder care) that are mainly or entirely tax financed and available for citizens at low or no monetary costs. The discussion about role of benefits in kind in the welfare state goes far back. Based on traditional welfare economics it can be argued that in-kind redistribution is inefficient, because the potential beneficiary could typically do better (and never worse) if he were given the cash equivalent of the in-kind subsidy. One might thus ask why not all redistribution is done using cash transfers. There are different types of answer to that question. First, the traditional welfare economic argument that it is sufficient to redistribute income and to rely a markets for allocative efficiency rests on assumptions such as markets being sufficiently competitive (Arrow, 1963) and that equilibria are unique (Foldes, 1967). A second more fundamental objection is that the traditional welfare economic approach implicitly assumes that tax payers care only about the utility of beneficiaries, which may not be true. As noted by Buchanan (1968) taxpayers may well aim to support only "specific spending patterns" (p. 189), in which case re-distribution in kind is not obviously inefficient.[2]

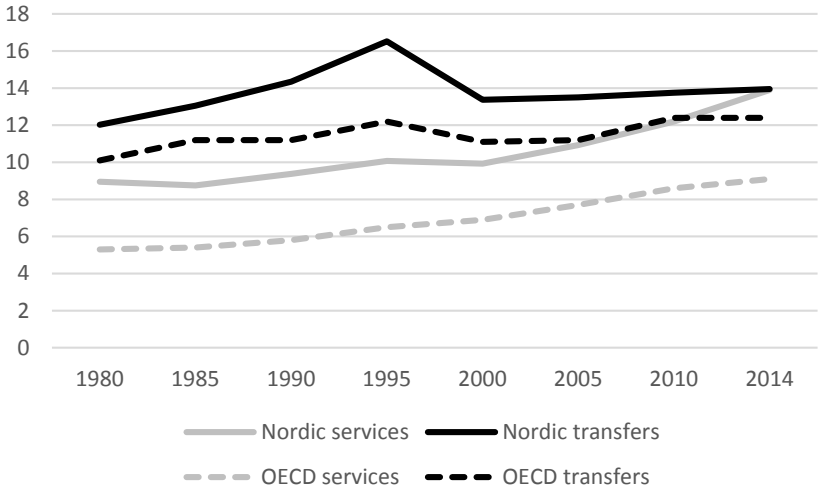
Regardless of how welfare services are motivated, they are an important part of the welfare state and they matter a lot for welfare state redistribution. Compared to other countries, welfare services are particularly extensive in the Nordic welfare states. In the literature on welfare state classifications (Titmuss, 1974, Esping-Andersen, 1990), welfare states are typically divided into three categories: the "marginal" (typical for Anglo-Saxon countries), the "industrial achievement" (typical for Central European countries) and the "institutional" (typical for Scandinavia). As noted by both Bambra (2005) and Jensen (2008), the research on welfare regimes has focused on cash transfers, paying less

[2] See Garfinkel (1970) for an analysis that accounts for both beneficiary and tax payer preferences.

attention to the role of welfare services. The contributions of Bambra and Jensen confirm however that the Nordic countries are different also when welfare services are accounted for.[3]

Compared to the OECD average, the Nordic countries actually differ more from the OECD average in the spending on welfare services than they differ in the spending on cash transfers. As shown in Figure 1, spending on welfare services in the Nordic countries are now financially as important as cash benefits, with both at 14% of GDP in 2014.

Figure 1: Services and transfers (% of GDP), average for Sweden, Denmark, Finland, Norway compared to OECD average 1980–2014



Source: OECD.

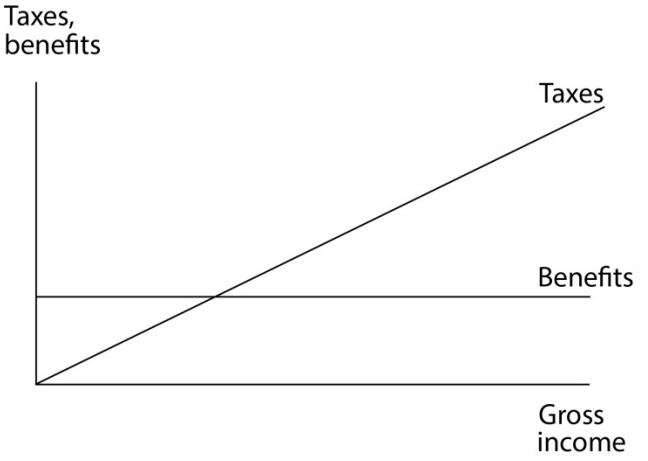
9.2 Welfare services and redistribution

Though sometimes neglected, the provision of welfare services has important consequences for welfare state redistribution. Rothstein (2001) provides a simple model of how universal programs aimed at the entire

[3] The one exception is health care expenditure, which according to Jennsen (2008) is characterized by uniform levels of expenditure across countries.

population have a redistributive effect even when they are financed using proportional taxes, illustrated in figure 2.

Figure 2: Redistribution in a stylized welfare state

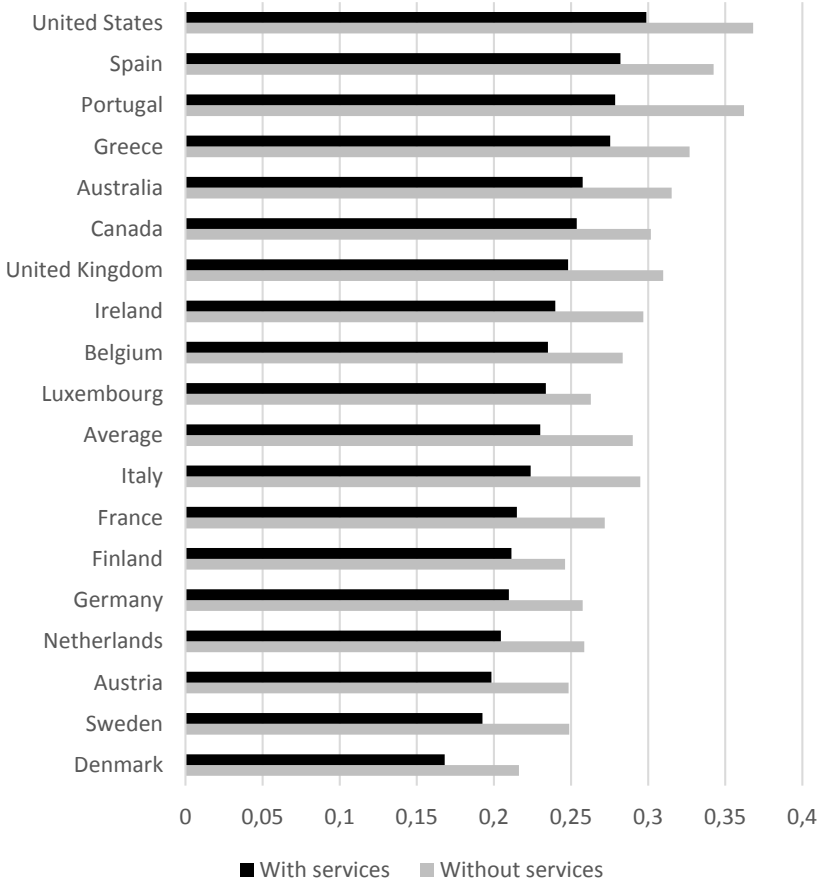


Source: Based on Rothstein 2001.

Because cash transfers and welfare state services both matter for welfare state redistribution, inequality measures based only on the distribution of household income will not fully capture all political efforts to increase equality. It is far from obvious, however, how the redistributive effect of welfare services should be accounted for.

Welfare services contribute to equality both directly and indirectly. Figure 3 illustrates an attempt by Verbist and Matsaganis (2012) to approximate the extra redistribution created by welfare services by simply adding the value of welfare services to household disposable income, and calculating Gini coefficients that include the value of welfare services. The OECD average Gini coefficient in 2008 decreases from 0.29 to 0.23 when the value of welfare services is included. The Nordic countries are similar to the average, with Sweden falling from 0.25 to 0.19, and Denmark falling from 0.22 to 0.17.

Figure 3: Gini coefficient 2008 for disposable household income including and excluding the value of tax financed services



Source: Verbist and Matsaganis (2012).

The calculations in Figure 3 are problematic in several ways. Most importantly, they are based on the assumption that receiving welfare services worth a certain amount is similar to receiving the same amount in cash. That is typically not the case. When the content of publicly provided services differs more from what households would buy if given cash instead, there is less extra direct redistribution associated with welfare

services. The estimates in Figure 3 should therefore be seen as an upper bound on the extra redistribution created by welfare services.

A second problem is that the degree to which the content of the welfare services provided matches what households would buy with the cash equivalent is likely to differ between countries because political systems differ. As a result, the cross country ranking shown in figure 3 is not entirely reliable.

Both problems are related to the discussion above on benefits in kind versus cash benefits. Should the value of welfare services and their impact on distribution be calculated using the preferences of the beneficiaries or the preferences of the tax payers, or possibly using some other set of social preferences? The welfare services provided by the welfare state may differ not only from what beneficiaries would buy with the cash equivalent, but also from what tax payers would want to provide.

A third problem with the data in Figure 3 is that they do not capture the indirect or dynamic channels by which welfare services can affect the income distribution. For example, if publicly provided schooling contribute to a more egalitarian distribution of human capital in the population, one should expect that incomes are also more equally distributed. The main reason why primary schooling contributes to a more equal income distribution is not that households receive schooling worth a certain amount of money yearly, but rather the way in which schooling affects the distribution of human capital in the population. A substantial and significant equalizing effect of primary education has been identified for Sweden by Meghir and Palme (2005) and a survey by Abdullah *et al*. (2015) demonstrates that a large literature has found a similar link between education and income equality.

To summarize: The provision of welfare services is financially important. As shown in Figure 1, the Nordic countries have during the recent two decades become more similar to the OECD countries when it comes to expenditure on cash transfers, but they have actually become more different in spending on welfare services. Welfare services have thus become an increasingly important characteristic of for the Nordic welfare states. Welfare services are also an important mechanism by which the welfare states affects the income distribution and promotes equality more generally.

Against this background, it is motivated to ask if there are threats or challenges for the future provision of welfare services. There is no shortage of analyses that point to potential problems for the welfare state connected to the provision of welfare services. In particular, problems related to Baumol's cost disease, Wagner's law and the population aging are often mentioned. These three are discussed in the following.

9.3 The Baumol effect

The Baumol effect, also known as Baumol's cost disease, can be traced back to Baumol (1967) where the explicit premise is that economic activities can be grouped into two types: "technologically progressive activities in which innovations, capital accumulation, and economies of large scale all make for a cumulative rise in output per man hour" and "activities which, by their very nature, permit only sporadic increases in productivity" (p. 415–6). More generally, activities differ in the relative importance of man hours as a production factor. For some activities, typically those that require face to face human interaction, productivity increases are rare and difficult to achieve. While technological progress has increased output per hour worked substantially in the manufacturing industry, the time needed to produce many services, for example the time it takes to help an old man eat a meal, has remained more or less constant. More generally, if wage increases tend to be uniform across the labour market, the relative price of services will increase.

The implications of Baumol's analysis is sometimes said to have dramatic consequences for the public sector, and in particular for the provision of welfare services. As recently noted by Andersen and Kreiner (2015), both the IMF (2012) and the European Commission (2013) refer to Baumol's cost disease as an important expenditure driver for the public sector. But under what circumstances must will higher productivity in some parts of the economy lead to cost increases or other problems for the public sector? First, two clarifications must be made.

First, the effect noted by Baumol (1967) applies to all services where manual labor time and human interaction are important inputs with no or very imperfect substitutes, regardless of whether these are

part of the public or the private sector. Baumol (2012) discusses health care and education but also the performing arts. The Baumol effect has also been used to explain the decline of the big bands and the growth of rock 'n' roll (DeBoer, 1985). While it may well be the case, especially in universal welfare states, that many of the activities where productivity increases are rare and difficult to achieve occur within the public sector, the Baumol effect itself is independent of the public-private dimension.

Second, it is far from clear how productivity should be measured in the public sector. Despite these measurement problems, there are many welfare services where both potential and actual productivity increases are possible, among other things because of labor saving technologies and improved management (see, for example OECD, 2006, Productivity Commission, 2005, Carter *et al*., 2011 and Arnek, *et al*. 2013).

When it applies, Baumol's disease implies that the tax rate is determined by the fraction of total labor employed in the public sector (Lindbeck, 2006, Andersen and Kreiner, 2015). It is a necessity that taxes must be increased when the fraction of the labor force employed in the public sector increases, but for a given fraction, the Baumol effect does not imply any upward pressure on expenditure and thus not on taxes. A higher relative price of services caused by increasing productivity in manufacturing is in itself is no expenditure driver. It does, however, increase the opportunity cost of services, and policymakers that wish to maintain or even increase public financing of welfare services may find it harder to motivate this spending pattern.

It bears emphasizing that the driving force for the Baumol effect is a fundamentally desirable development, namely increasing productivity in some parts of the economy, leading to rising real incomes. The challenges come from the fact that these productivity increases do not occur uniformly. For any given average rate of productivity growth in an economy, it is only to be expected that in some sectors it will be higher and in others it will be lower. If large parts of the economy experiences high productivity growth, average income will increase and demand patterns will change as a result of income effects. The result may well be

increasing demand for services where productivity has increased less.[4] If this is the case, it is actually be a sign of progress that low-productive sectors of the economy are expanding.

It is thus incorrect to describe Baumol's disease as an expenditure driver, and it does not suggest that the welfare state will run into problems due to rapidly increasing expenditure on public services. Andersen and Kreiner (2015) formally shows that the Baumol effect does not imply that the welfare state is unsustainable. In fact they show that, under certain assumptions, the Baumol effect is compatible with constant principles for the supply of services, a constant distribution of well-being and a balanced budget, and also leaves room for Pareto improvements.

While the challenges for the welfare state caused by the Baumol effect are less about financial sustainability, they are probably more about political sustainability and distribution. As noted above, welfare services play an important role in welfare state redistribution. When welfare services become relatively more expensive, it can be seen as increasing the relative price of redistribution. As possible response is to make publicly financed welfare services less redistributive. Policymakers who wish to maintaining public support for redistributive welfare services may thus find it increasingly difficult to do so.

9.4 The Wagner effect

Adolph Wagner (1893) proposed that there is a positive relationship between the level of economic development and the size of the public sector. This general proposition is compatible with several exact meanings. Henrekson (1993) discusses various interpretations and concludes that it is a reasonable interpretation is that Wagner claimed that public spending as a share of GDP is positively related to real GDP per capita.

As noted by Durevall and Henrekson (2011) Wagner's view that the public sector's share of GDP tends to grow as real per capita GDP increas-

[4] It should be noted that in many cases, productivity is notoriously hard to measure because products and services, as well as the quality with which they are associated, are hard to define and measure.

es was an accepted fact in public economics among influential observers such as Atkinson and Stiglitz (1980) and Easterly and Rebelo (1993), and also in mainstream textbooks (such as Hindriks and Myles (2006)). Until the early 1990s, Wagner's proposition had received strong empirical support (see, for example Musgrave 1969, Mann 1980 and Abizadeh and Gray, 1985) and is still today often referred to as a law.

Surveying the more recent literature, Durevall and Henrekson (2011) categorize 40 studies published after 1990 and show that about 35% fail to find evidence for Wagner's Law, 30% find support when controlling for other variables or focusing on specific types of expenditures, and 35% obtain unqualified support for the hypothesis. Their own study of the long run development within Sweden and the UK since 1800 suggest that Wagner's Law does not hold in the long run, but that data are consistent with Wagner's Law for the periods 1860–1913 and 1920–1975.

Can it be said that Wagner's law is now falsified? Strictly speaking, any observation where it clearly does not hold is arguably sufficient to change the label from law to regularity. Given how rare laws are in the social sciences, the positive correlation between the public sector's share of GDP and real GDP per capita remains an important regularity.

An important explanation for the Wagner regularity is that the long run income elasticities are well above 1 for many welfare services that in many countries are provided mainly by the state. An often cited source is Fogel (1999), who calculated long run income elasticities for various consumption categories in USA as shown in Table 1.

Table 1: Long-term income elasticities for USA (1875–1995)

Food	0.2	Health Care	1.6
Clothing	0.3	Education	1.6
Shelter	0.7	Leisure	1.4

Source: Fogel (1999).

According to Fogel (1999) the long run income elasticity for welfare services such as health care and education is well above unity.[5] Hirsch (1961) calculates an income elasticity at 1.1 for the US over the period 1900–1958 for public education expenditure only. High income elasticities for health care in other OECD countries are also documented by OECD (1985) for the period 1960–1983 period. More recent studies based on micro data (e.g. Parkin *et al*. (1987)) have found lower elasticities, but according to Getzen (2000) the diversity of findings reflects the fact that individual income elasticities are typically near zero, while national health expenditure elasticities are commonly greater than 1.0.

As a result, we expect expenditure on health care and education to expand as a share of total expenditure as countries grow richer. This is also the trend documented by OECD (2006) for both public and private expenditure in OECD over the period 1970–2005. The fact that countries that grow richer tend to spend a larger share of the GDP on these services does not necessarily mean that public expenditure on these areas must also increase their GDP share. While lots of historical data are in line with the Wagner regularity, the crucial question with regard to the future of the welfare state is if further income growth will imply further increases in public sector size. The answer ultimately depends on political decisions, discussed further in section 5.

9.5 Aging

Demographic change is an often mentioned challenge for the welfare state. The basic idea is that a relatively older population increases expenditure via both the pension system and increasing demand for welfare services such as health care and elder care. For example, Morrow and Roeger (1999) estimated that the ratio of the population above 65 to the population in labor active groups will increase from 24 to 49% within the EU15 countries between 2000 and 2040. Similar predictions

[5] Note also the high income elasticity reported by Fogel for leisure. It is natural that people want to expand non-material activities as incomes increase. That trend poses yet another challenge for the public sector because labour income is taxed whereas leisure is not.

are made in more recent analyses: According to the 2015 Ageing Report from the European Commission, the ratio is projected to increase from 27.8% today to 50.1% in 2060 for the EU as a whole.

Based on such population forecasts it is easy to predict a dismal future for ageing societies, in particular where the welfare state plays an important role in the income smoothing over the life cycle. Again, one must be careful not to overstate the problems.

While the demographic changes are often mentioned as a challenge for the welfare state, the demographic situation is actually less alarming in the Nordic countries compared to for example Greece, Poland and Portugal (European Commission, 2015), and a likely explanation is that the family and labor market policies of the welfare state have contributed to keeping fertility rates (Rovny, 2011).[6]

Italy, possibly as a result of child and parent friendly welfare state policies (see for example OECD, 2006). Most importantly, the long run problems associated with population ageing are smaller if people can be expected to work roughly a constant share of their lifespan. Still, it must be stressed that substantial challenges during the 2020s and 2030s remain even if the average retirement age increases. To see this, a simple calculation based on the official population forecasts in the Nordic countries is sufficient.

Figure 5a shows how the working age share of the Swedish population (15 to 64 years old) is expected to decrease from 63% today to 60% in the 2030s, and falling further during the 2050s to reach 58% around 2060. Over the same time period, the share aged 75 or older is predicted to increase from 8.5% today to 11% in the 2020s and to reach 13% in the 2040s. Based on numbers like these it seems safe to predict an increasing demand for elder care and health care that coincides with a shrinking tax base made up of those in working age. Such predictions, however, ignore the counter effect that the definition of "working age" is likely to change as life expectancy increases.

[6] Interestingly, Rovny also find that the presence of employment protection legislation — rules concerning hiring and firing – hinders the growth of fertility rates.

Looking back one decade in Sweden reveals that the employment rate among 65 to 74 year olds has been steadily increasing, from 9.9% in 2005 to 16.5% in 2014. This corresponds to a yearly increase of 0.73 percentage unit in the employment rate. Moreover, the increasing employment trend for 65 to 74 year olds was not visibly affected by the financial crisis, in contrast to the employment for the entire adult population. The trend towards increasing employment rates among 65 to 74 year olds in Sweden thus seems to be relatively robust. Adding employed 65 to 74 year olds to those who are counted as working age, changes the prediction for the future substantially if we are willing to assume that employment among 65 to 74 years olds continues to increase the way it has the most recent decade. As a more conservative scenario, assume instead that the future yearly increase is 70% of the yearly increase from 2005 to 2014 (which for Sweden means a yearly increase of 0.7*0.73 = 0.51 percentage units). The difference between the two scenarios is illustrated by the fact that employment among 65 to 74 year olds will have risen to 40% in 2060 in the conservative scenario, and to 50% in the optimistic scenario.

Figure 5a–d shows the result of doing these calculations using the official population forecast for Sweden, Denmark, Norway and Finland.[7] As can be seen in Figure 5, even the conservative scenario improves the picture substantially in all countries except Denmark (where employment for 65 to 74 years olds fell from 10% in 2005 to 9.1% in 2009, and then increased to 11.8 in 2014 (corresponding to a yearly increase of 0.2 percentage units over the latest decade).[8] On the other hand, even in the optimistic scenario demographic balance will still worsen in all countries at least until the 2030s.

Figure 5a–d. Population share in working age 16 to 64 years (dashed) and including an increasing share of 65–74 year olds.

[7] Note that only Sweden and Norway have forecasts that reach the year 2110, whereas Denmark's stops at 2050 and Finland's at 2060.

[8] In Finland the trend in employment for 65 to 74 years olds is similar to the trend in Sweden, with employment increasing from 5.1% in 2005 to 10.2& in 2014, corresponding to a yearly increase of 0.56 percentage units. Norway is also similar, with an increase in employment for 65 to 74 years during the same period from 12.8 to 19.3%, or 0.72 percentage units yearly.

Upper solid: optimistic scenario with future increase based on trend 2005 to 2014. Lower solid: conservative scenario based on trend 2005 to 2014 multiplied by 0.7). Source: Author's calculations based on official population forecasts.

Figure 5a: Sweden

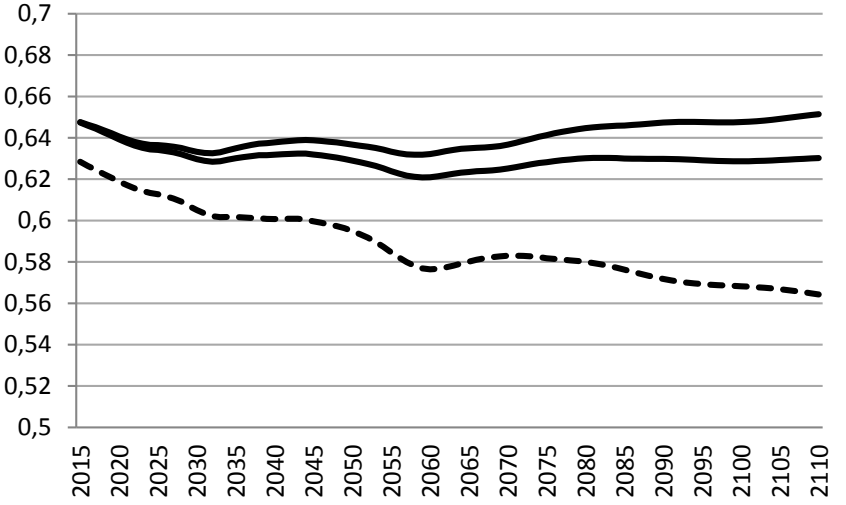


Figure 5b: Denmark

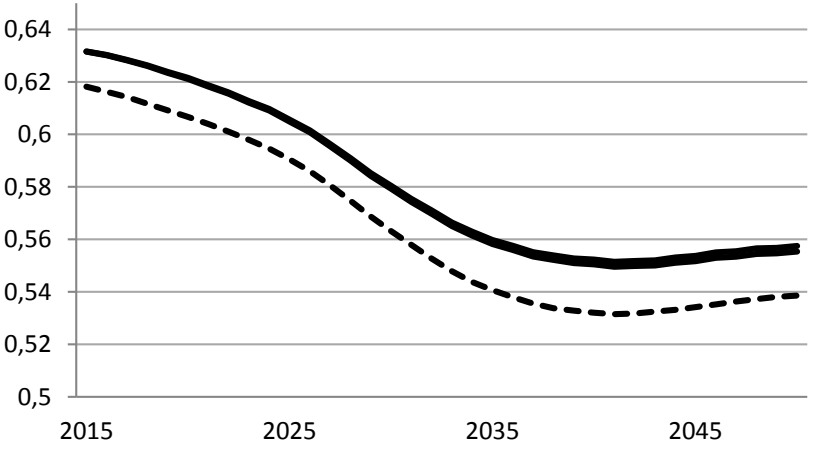


Figure 5c: Norway

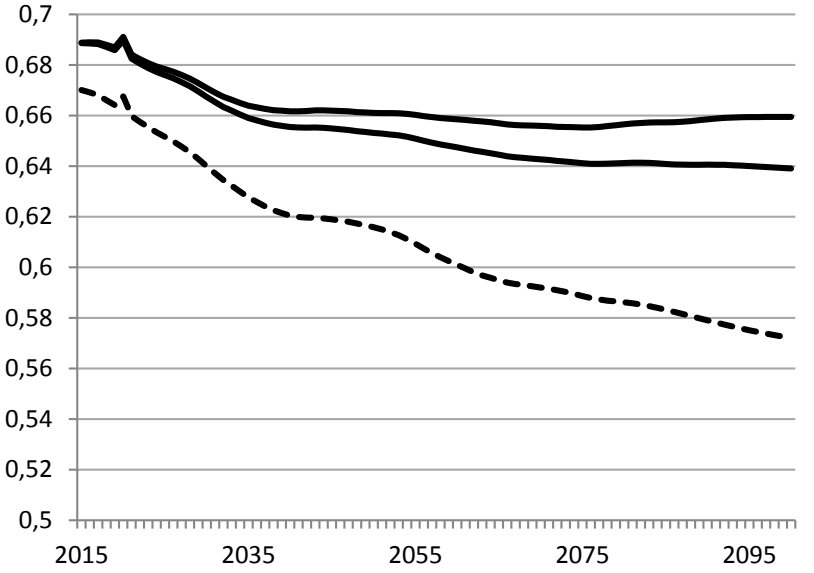
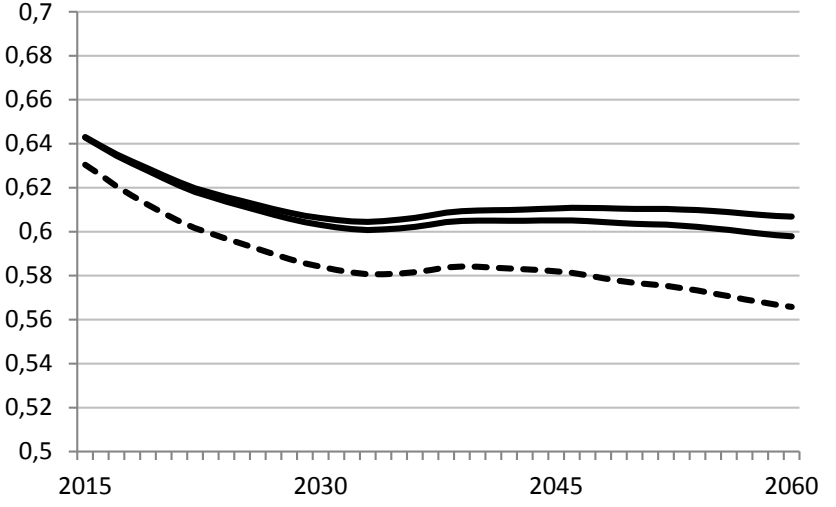
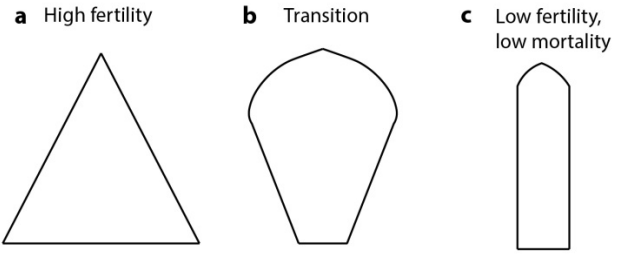


Figure 5d: Finland



As shown by Bengtsson and Scott (2010) the main driver behind the demographic changes is the transition from high fertility to low fertility, and only a small part is explained by increased longevity. A schematic explanation is given by Bergh (2010), reproduced in Figure 6.

Figure 6: Schematic population pyramids during before, during and after the transition from high to low fertility



Source: Bergh (2010).

When fertility and mortality is high, the demographic structure is akin to the left triangle in Figure 6. When fertility drops, the population pyramid is transformed to the well-known mushroom shape, with a relatively large share of old in the population. The mushroom shape is the reason why the demographic balance in Figure 5a to 5d will deteriorate for most countries during the next decades, but the situation should improve once the cylinder stage is reached.

Finally, it is motivated to as how increased longevity will affect health care expenditure. If we live longer, will we also need more health care? The answer depends on the extent to which increasing life expectancy adds healthy years to our lives, or years when we need expensive health care to stay alive. This is a well-researched topic, and there is agreement in the literature that health care costs increase by increased proximity to death, suggesting that longer lives should only have small effects on total expenditure. Christiansen *et al*. (2012) summarize the state of research and confirm that ageing as such can be expected to cause only a modest increase in health care expenditure per capita in the future. Their own empirical study, based on 15 EU countries, support this conclusion. They also give examples of cross-country studies

where demography loses significance when institutional characteristics are controlled for factors. OECD (2006) reaches a similar conclusion: rising health care expenditure is mainly explained by rising incomes, and only to some extent by population aging.

9.6 Taxes, topping up or paying twice – can the welfare state muddle through?

For the welfare state, the combination of the Baumol effect, the high income elasticity of welfare services and the aging population creates a challenge that must be acknowledged and discussed. As productivity increases, society grows richer. With higher income, people are likely to demand more welfare services. But some incomes grow faster than others. When increasing average incomes are combined with increasing income dispersion, and the income elasticity for welfare services is above unity, the difference between the level and quality of welfare services demanded by high income earners and low income earners will grow. Furthermore, welfare states services have traditionally played an important role in making universal redistributive towards low-income earners. As the Baumol effect means that they become relatively more expensive, they are increasingly demanded by high income earners. When more is demanded from services that are to be produced by the public sector, higher taxes are required. Raising taxes may be problematic for efficiency reasons, especially in countries where average taxes are already relatively high. Moreover, the need to raising taxes in order to improve welfare services in ways demanded by high income earners adds a potentially problematic political dimension.[9]

If tax increases are ruled out, two remaining options are the so-called topping up strategy and the paying twice strategy. In this context, topping up entails that those who desire to do so may add private financing on top of the publicly financed welfare service level in order to

[9] However, see Blomquist *et al*. (2010) for an interesting idea involving increased tax progressivity as a way to make high income earners to pay more for publicly provided services.

achieve the desired level, quantity or quality of the service. If topping up is not allowed, high income earners still have the option to pay twice, in the sense that after paying taxes to a welfare state where the provision levels are insufficient by their standards, they buy private insurance arrangements on the market for privately provided welfare services.

In the short run, the difference between the "topping up" and the "paying twice" strategies may seem small, but the long run political dynamic is fundamentally different. Allowing topping up can be increase the political support for public provision (Epple and Romano, 1996, Gouveia, 1997), while having high income earners paying twice is likely to erode their willingness to pay taxes to finance the public system. Empirical support for the latter mechanism has been provided by Hall and Preston (1998) who showed that people who opt out from publicly provided health care and pay for private health insurance support less spending on the public system.

To put it harshly, policy makers are facing an unpleasant choice between three alternatives: Increasing taxes, facilitating topping up and accepting increasing inequality of access to welfare services, or having high income earners paying twice and risk eroding welfare state support. In practice, the three strategies are not mutually exlusive. As discussed by Bergh (2008) there are signs that Swedish policy makers implement at least the incremental changes that are necessary to secure majority support for a high tax welfare state. In many cases, the changes are likely to imply that the vertical income distribution of the welfare state decreases, as many welfare services are changed according to the preferences of the middleclass. The tendency for welfare state programs to adjusted to the interests of the middle class for political reasons is well-documented (Goodin and Le Grand, 1987)

If the strategy continues, policy makers in the future may well opt for a mix of higher taxes, topping up and having some groups paying twice, resulting in the Nordic welfare states muddling through the challenges ahead. In any case it can be concluded that while the problems caused by the Baumol effect for financial sustainability have been somewhat exaggerated, the political problems related to vertical income redistribution may well be underestimated.

References

Abdullah, A., Doucouliagos, H., & Manning, E. (2015). Does education reduce income inequality? A meta-regression analysis. *Journal of Economic Surveys* 29, 301–316. http://dx.doi.org/10.1111/joes.12056

Abizadeh, S. & Gray, J. (1985). Wagner's Law: A Pooled Time-Series, Cross-Section Comparison. *National Tax Journal* 38:209–218.

Andersen, T. M., & Kreiner, C. T. (2015). Baumol's cost disease and the sustainability of the welfare state. *EPRU Working Paper Series 2015-02*.

Arnek, M., Bergman, M., Edquist, H., Hultkrantz, L., Lundberg, S. & Svensson, M. (2013). Den offentliga sektorn – en antologi om att mäta produktivitet och prestationer. *Rapport till expertgruppen för studier i offentlig ekonomi 2013:8*.

Arrow, K. J. (1963). Uncertainty and the Economics of Medical Care. *American Economic Review* 55, 942–943.

Atkinson, A. B. & Joseph E. Stiglitz. (1980). *Lectures on Public Economics*. New York: McGraw-Hill.

Baumol, W. J. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *American Economic Review* 57, 415–426.

Bengtsson, T. (2010). Po*pulation Ageing, a Threat to the Welfare State? – The Case of Sweden*. Springer. http://dx.doi.org/10.1007/978-3-642-12612-3

Bergh, A. (2010). Towards a New Swedish Model? In T. Bengtsson (ed.), *Population Ageing – A Threat to the Welfare State?* pp. 109–118. Springer. http://dx.doi.org/10.1007/978-3-642-12612-3_7

Buchanan, J. M. (1968). *What Kind of Redistribution Do We Want?* p. 185.

Carter, K, Dorgan, S. & Layton, D. (2011). Why hospital management matters. *Health International* 11.

Christiansen, T., Lauridsen, J., & Bech, M. (2012). Ageing populations: More care or just later care? *Nordic Economic Policy Review*, 23–53.

Durevall, D. & Henrekson, M. (2011). The futile quest for a grand explanation of long-run government expenditure. *Journal of Public Economics* 95:708–722. http://dx.doi.org/10.1016/j.jpubeco.2011.02.004

Easterly, W., & Rebelo, S. (1993). Fiscal policy and economic growth. *Journal of Monetary Economics* 32, 417–458. http://dx.doi.org/10.1016/0304-3932(93)90025-B

Epple, D. & Romano, R. E. (1996). Public Provision of Private Goods. *Journal of Political Economy* 104, 57–84.

European Commission (2013). Report on Public Finances in the EMU. *European Economy* 4-2013.

European Commission(2015). *The 2015 Ageing ReportEconomic and budgetary projections for the 28 EU Member States (2013–2060)*. European Economy 3/2015.

Fogel, R. W. (1999). Catching up with the economy. *American Economic Review* 89, 1–21. http://dx.doi.org/10.1257/aer.89.1.1

Foldes, L. (1967). Income Redistribution in Money and in Kind. *Economica* 34, 30–41. http://dx.doi.org/10.2307/2552511

Garfinkel, I. (1970). Is in-kind redistribution efficient? *Institute for Research on Poverty Discussion Paper 74-70*.

Getzen, T. E. (2000). Health care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures. *J Health Econ* 19, 259–70. http://dx.doi.org/10.1016/S0167-6296(99)00032-6

Goodin, R. E., & Grand, J. L., (1987). *Not Only the Poor: The Middle Classes and the Welfare State.* London: Allen and Unwin.

Gouveia, M. (1997). Majority rule and the public provision of a private good. *Public Choice* 93, 221–244. http://dx.doi.org/10.1023/A:1017929005280

Henrekson, M. (1993). Wagner's Law – A Spurious Relationship? *Public Finance* 48, 406–415.

Hindriks, J. & Myles, G. M. (2006). *Intermediate Public Economics.* Cambridge, MA.: MIT Press.

Hirsch, W. Z. (1961). Income Elasticity of Public Education. *International Economic Review* 2, 330–339. http://dx.doi.org/10.2307/2525434

Horngren, L. (2001). Pension reform: The Swedish case. *Journal of Pensions Management* 7, 131–138.

IMF. (2012). The Economics of Health Care Reform in Advanced and Emerging Economies.

Lindbeck, A. (2006). Sustainable social spending. *International Tax & Public Finance* 13, 303–324. http://dx.doi.org/10.1007/s10797-006-9175-5

Mann, A. J. (1980). Wagner's Law: An Econometric Test for Mexico, 1925–1976. *National Tax Journal* 33, 189–201.

Musgrave, R. A. (1969). *Fiscal Systems*. New Haven and London: Yale University Press.

OECD (1985). Measuring health care 1960–1983. Expenditure, costs and performance. *OECD Social Policy Studies No. 2*.

Parkin, D., McGuire, A., & Yule, B. (1987). Aggregate health care expenditures and national income. Is health care a luxury good? *J Health Econ* 6, 109–27. http://dx.doi.org/10.1016/0167-6296(87)90002-6

Productivity Commission. (2005). Impacts of Medical Technology in Australia (Progress Report). *Melbourne, April 2005*.

Rovny, A. E. (2011). Welfare state policy determinants of fertility level: A comparative analysis. *Journal of European Social Policy*, 21(4). 335–347. http://dx.doi.org/10.1177/0958928711412221

The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't (2012).. USA: Yale University Press.

The drivers of public expenditure on health and long-term care: An integrated approach. (2006). *OECD Economic Studies No. 43, 2006/2*.

Wagner, A. (1893). *Grundlegung der politischen Ökonomie (3rd edition)*. Leipzig.

Verbist, G. & Matsaganis, M. (2012). The Redistributive Capacity of Services in the EU. *Gini Discussion Paper 53*.

10. Ethnic fractionalization and the demand for redistribution – Potential implications for the Nordic model

Johanna Mollerstrom[1]

Summary

A distinctive feature of the Nordic model is that economic resources are redistributed between citizens to a high degree. Historically, the Nordic voters have expressed a higher demand for such redistribution than people in other parts of the world. This paper considers the factors that determine individual preferences for redistribution, giving special attention to heterogeneity in the form of ethnic fractionalization. Such heterogeneity is generally linked to reductions in demand for redistribution. The paper takes as its starting point the fact that the populations in all the Nordic countries are, to varying degrees, becoming increasingly heterogeneous. Potential mechanisms for why this may have a negative impact on demand for redistribution, and potential consequences for the Nordic model are discussed.

---

[1] Interdisciplinary Center for Economic Science, George Mason University, Fairfax, VA, USA. Research Institute for Industrial Economics, Stockholm, Sweden. Email: jmollers@gmu.edu. I am grateful to Henrik Jordahl, Magnus Henrekson, Bjørn-Atle Reme, Jesper Roine and the participants in the 2015 NEPR Conference for very useful comments and suggestions. Thank you also to Swan Lee for excellent research assistance.

10.1 Introduction

The Nordic model is a combination of a free market economy and an extensive welfare state, and a system that all the Nordic countries (Denmark, Finland, Norway, Iceland, and Sweden) exhibit. The model is intensely debated, admired by many (e.g. Logue, 1979; Popenoe, 1994; Sachs, 2006; Hilson, 2008; and Brandal *et al*, 2013), but often also heavily criticized (see Sanandaji, 2015 for an overview). This paper focuses on one particular aspect of the Nordic model: extensive economic redistribution between citizens.

Economic redistribution means that economic resources are reallocated between citizens in such a way that the distribution becomes more equal. Such redistribution can be done through the public sector or through charities (this distinction is discussed in e.g. Alesina and Glaeser, 2004), with many people having strong opinions about which entity is most well suited to conduct the redistribution. The extensive economic redistribution in the Nordic countries is almost exclusively conducted through the public sector.

Basically all countries in the world redistribute resources between citizens to some extent, and policies with redistributive components, such as social security, universal health insurance, and progressive taxation, have increased in importance over time (Alesina *et al*. 2004). The fact that within country inequality has increased in most Western democracies over the last decade(s) is hence primarily reflecting that economic redistribution has not kept up with the pace of the increase in the pre-redistribution inequality, rather than mirroring decreases in public redistribution (Wang *et al*., 2014).

Historically, the Nordic countries have redistributed resources between citizens to a much larger extent than what is the case in other Western democracies (Alesina and Glaeser, 2004). This is still the case, but there are tendencies of growing inequality also in the Nordic countries (Atkinson and Morelli, 2014; Morelli *et al*., 2014; and Wang *et al*., 2014). Not only the implemented redistribution but also the preferences of the population are different in the Nordic countries with Danes, Finns, Norwegians, Icelanders and Swedes being more politically sup-

portive of economic redistribution than people in other parts of Europe or in the US.

Preferences for redistribution are not, however, exogenously given but something that change over time. This paper is discussing how to understand the political support for redistribution in general. As many of the Nordic countries are currently experiencing increasing ethnic heterogeneity, in particular through immigration, the paper specifically focuses on how this may impact demand for redistribution.

The paper continues as follows: Section 1 discusses cross-country differences in demand for redistribution from the empirical perspective and goes on to give an overview of the vast literature on which factors that determine demand for redistribution. Section 2 focuses more narrowly on the research that relates demand for redistribution to heterogeneity and ethnic fractionalization. Section 3 looks specifically at the empirics of immigration in the Nordic countries, and discusses historical and current trends. Section 4 concludes.

10.2 The demand for redistribution

10.2.1 Cross-country differences

To measure exactly how much a country redistributes between its citizens is challenging as systems for taxation and welfare spending look very different in different countries. Direct cross-country comparisons are therefore hard to make. There are also other policies, such as labor market regulations (which also tend to differ a lot in structure between countries), that are designed to benefit lower income groups in particular. Despite these measurement difficulties there is vast agreement on the general patterns of redistribution in the Western world: European countries has much more of it than the US, and in Europe the Nordic countries top the redistribution league.

There are several ways to illustrate this. The gini coefficient is one of the most commonly used measures of how equal incomes are distributed between people in a country (but many other measures exist and using them instead yields very similar results). The gini coefficient

is a measure of statistical dispersion, with 0 indicating perfect equality and 1 indicating perfect inequality. Table shows gini coefficients for primary (pre-redistribution) and disposable (post-redistribution) income, for four Nordic countries (Denmark, Finland, Norway and Sweden), three other European countries (France, Germany, UK) and the United States. By definition, the difference between the gini coefficient for primary and disposable income is a measure of fiscal redistribution. It is important to note, however, that the shape of the welfare state will influence not only post- but also pre-redistribution inequality, so this difference is not the only relevant redistribution measure.

Some patterns are immediately obvious from Table 1. First, we see a trend of increasing inequality over time which can be observed in most countries. This is especially pronounced in the distribution of primary income.

Second, compared to the US, the Nordic countries have both a more equal distribution of primary income and a (much) more equal distribution of disposal income. This indicates that, compared to the US (and to some extent the UK), the higher equality in the Nordic countries stems both from structural factors that make pre-redistribution income more equal, and from the fact that the public sector is larger and more redistributive.

Third, if we instead compare the Nordic countries to the three other European countries in the table, France, Germany and the UK (but again, in some regards the UK has more in common with the US than with most European countries), we see a different pattern. The gini coefficient for primary income indicates more equality in the Nordic countries, but it is not immediately obvious that the public sector in other European countries is much smaller or less redistributive than in Denmark, Finland, Norway and Sweden.

Table 2 shows other indicators of the size of the public sector in the form of tax revenue, and public expenses, in percentage of GDP. To some extent this reveals information about the degree of redistribution (since a lot of redistribution is done through taxes and expenses) but there is of course, also significant public spending which is not redistributive (military spending is one, but not the only example). We see, again, the sharp divide between Europe and the US, with both taxes and

public expenses making up a significantly higher share of GDP on the European side of the Atlantic. In terms of taxation, the Nordic countries display a relatively larger public sector than the other European countries, but the difference is not anywhere as large as the difference between Europe and the US. For public expenses, the difference between the Nordic countries and the other European countries in the table is less pronounced.

Table 1a: Trends in income inequality and redistribution, 1985–2005 (Girni primary income)

	Girni primary income						
	mid-1970s	mid-1980s	around 1990	mid-1990s	around 2000	mid-2000s	2010
Denmark		0.373	0.396	0.417	0.416	0.416	0.429
Finland		0.387		0.479	0.478	0.483	0.479
Norway		0.351		0.404	0.426	0.447	0.423
Sweden	0.389	0.404	0.408	0.438	0.446	0.432	0.441
France				0.473	0.490	0.485	0.505
Germany		0.439	0.429	0.459	0.471	0.499	0.492
UK	0.378	0.469	0.490	0.507	0.512	0.503	0.523
United States	0.406	0.436	0.450	0.477	0.476	0.486	0.499

Source: Morelli *et al*. (2014). Data from the OECD inequality data base.

Table 1b: Trends in income inequality and redistribution, 1985–2005 (Gini disposable income)

	Gini disposable income						
	mid-1970s	mid-1980s	around 1990	mid-1990s	around 2000	mid-2000s	2010
Denmark		0.221	0.226	0.215	0.227	0.232	0.252
Finland		0.209		0.218	0.247	0.254	0.260
Norway		0.222		0.243	0.261	0.276	0.249
Sweden	0.212	0.198	0.209	0.211	0.243	0.234	0.269
France				0.277	0.287	0.288	0.303
Germany		0.251	0.256	0.266	0.264	0.285	0.286
UK	0.269	0.309	0.355	0.337	0.352	0.335	0.341
United States	0.316	0.340	0.349	0.361	0.357	0.380	0.380

Source: Morelli *et al*. (2014). Data from the OECD inequality data base.

Table 1c: Trends in income inequality and redistribution, 1985–2005 (Redistribution)

	Redistribution						
	mid-1970s	mid-1980s	around 1990	mid-1990s	around 2000	mid-2000s	2010
Denmark		0.152	0.170	0.202	0.189	0.184	0.177
Finland		0.178		0.261	0.231	0.229	0.219
Norway		0.129		0.161	0.165	0.171	0.174
Sweden	0.177	0.206	0.199	0.227	0.203	0.198	0.172
France				0.196	0.203	0.197	0.202
Germany		0.188	0.173	0.193	0.207	0.214	0.206
UK	0.109	0.160	0.135	0.170	0.160	0.168	0.182
United States	0.090	0.096	0.101	0.116	0.119	0.106	0.119

Source: Morelli *et al*. (2014). Data from the OECD inequality data base.

Table 2: Tax revenues and public expenses as % of GDP

	Tax revenue in % of GDP	Public expenses in % of GDP	Public expenses in % of GDP, excluding military
Denmark	33.4	41.3	39.9
Finland	19.9	39.1	37.7
Iceland	22.2	34.3	34.2
Norway	26.8	33.7	32.3
Sweden	20.7	30.9	29.8
France	21.4	47.0	44.8
Germany	11.5	28.3	27.0
UK	25.3	42.1	39.9
US	10.2	24.0	19.8

Source: The World Bank. Data from 2012.

Lastly, we consider expressed demand for redistribution on behalf of the population. The World Value Survey, a global research project which investigates people's values and beliefs and how these change over time, ask a question about to what extent the respondent would like incomes in society to be made more equal. Answers are given on a scale between 1 and 10, where 1 is defined as "Incomes should be made more equal" and 10 as "We need larger income differences as incentives for individual effort". This data is only available for a subset of the countries that we are interested in, and as the World Value Survey is conducted continuously, it can differ quite substantially when the latest da-

ta point for a country is dated. People in the US tend to give a higher answer to this question, on average, with the latest data point being 5.58. For Finland, Sweden, France, Germany and the UK, the corresponding figures are 4.97, 4.88, 5.12, 4.08, 5.36 (World Value Survey, 2015 – data from 2005–2014). Again we see that the Nordic countries are among those who demand most redistribution but that for example Germany is not lagging far behind.

Another way to look at cross country differences in demand for redistribution is to compare political party preferences. Given that the US has never had an influential Social Democratic party, there is a clear difference between the two sides of the Atlantic in this regard. Within Europe, the Nordic countries have historically had stronger Social Democratic parties than many other European countries. This can be seen as an indicator of a strong Nordic demand for redistribution, but it should be noted that party structures are impacted not only by political preferences but also by institutional factors, such as the degree of federalism and whether the country has proportional or majoritarian representation (Alesina and Glaeser, 2004).

10.2.2 Underlying factors that shape demand for redistribution

The theoretical start of the quest to understand how individual demand for redistribution is shaped started with seminal contributions by Romer (1975), Roberts (1977), and Meltzer and Richard (1981). They present simple models that suggest that a person's relative income or wealth is decisive for her demand for redistribution: since a relatively richer person benefits less, in monetary terms, from redistribution she should demand less of it. A relatively poor person, on the other hand, has more to gain from redistribution and should therefore demand more.

These early theoretical predictions about individuals who are relatively rich, compared to others in society, wanting less redistribution has found empirical support: People with a higher income generally want less redistribution (see, e.g., Pelzman, 1985; Alesina and Giuliano, 2010; Margalit, 2013; Durante *et al*., 2014; Powdthavee and Oswald, 2014; and Elinder *et al*., 2015). Other studies utilize the fact that people

often have misperceptions about their own position in the income distribution. Mollerstrom *et al*. (2015) conduct a survey experiment in Sweden and inform a treatment group about their true relative position, thereby exogenously (but truthfully) manipulating perceptions of relative income. They show that this causes changes in individual demand for redistribution and that the direction of these changes indicate that a substantial fraction of people do exhibit the theoretically proposed negative correlation between (perceived) relative income and demand for redistribution (see also Cruces *et al*., 2013, for evidence corroborating this conclusion).

Income and wealth are not the only individual heterogeneities that generate differences in demand for redistribution and understanding the role that relative income, or the perceptions of relative income, play for the demand for redistribution is made more difficult by the fact that other underlying variables may also cause a correlation between income and political preferences. Mollerstrom and Seim (2014) find, for example, that high-IQ individuals in Sweden favor less redistribution. This could reflect that high ability individuals, who more easily succeed economically, lean toward a more individualistic, right-wing world view. In general, there is ample evidence, however, that also other factors, that have less to do with pure monetary self-interest, matter for a person's demand for redistribution. For example, people are often found to care also about the consumption of others (see e.g. Fehr and Schmidt, 1999; and Bolton and Ockenfels, 2000) and such other regarding preferences tend to be positively correlated with the demand for redistribution (Sears and Funk, 1990; Fong, 2001; and Alesina and Giulino, 2010).

Individual beliefs about the income generating process have also been studied theoretically (Piketty, 1995; and Benabou and Tirole, 2006) with special emphasis put on beliefs about the extent to which individual economic success can be attributed to effort, rather than to luck, and how such beliefs can be self-fulfilling. Empirically, these beliefs have been found to be a stronger determinant of preferences for redistribution than income itself (Fong, 2001) with people who believe that individual economic success can primarily be attributed to individual effort wanting significantly less redistribution in society than those who perceive luck as being the most important underlying factor.

A third factor which has been studied theoretically in order to understand how individual demand for redistribution is shaped is risk aversion. The provision of insurance against future negative economic shocks have been suggested to be part of the attraction of redistribution. Harsanyi (1953) was one of the first to illustrate how the insurance motive links the demand for redistribution to risk aversion: in his model, individuals are asked to state their preferences for redistributive policy behind a veil of ignorance, before their position in the income distribution is determined by a lottery. In such a situation, demand for redistribution should be increasing in a person's degree of risk aversion. Benabou and Ok (2001) extends and enriches a version of the Meltzer and Richards (1981) model and show that when individuals care not only about their current position in the income or wealth distribution, but also about their future position, insurance motives, and hence risk aversion, play a role for demand for redistribution. Empirically, the proposed positive relation between individual risk aversion and demand for redistribution has been shown to hold (see, e.g., Rainer and Siedler, 2008; Gaertner *et al*., 2015).[2]

In addition to theoretically founded discussions between demand for redistribution and its correlates there is also evidence that other factors matter. There is, for example, consistent, cross-country evidence that women want more redistribution than men, that those who are unemployed want more redistribution than the employed, that married people prefer less redistribution than singles and that having more education is negatively correlated with demand for redistribution (Alesina and Guiliano, 2010). In the US context, blacks have been shown to want more redistribution than whites (Alesina and Guiliano, 2010).

[2] The question of risk exposure and the size of the welfare state has mostly been analyzed from the perspective of individuals, but Rodrik (1998) discusses this also from an institutional perspective.

10.3 Ethnic heterogeneity and preferences for redistribution

In their seminal work seeking to understand why European countries exhibit much more extensive redistribution than the United States, Alesina and Glaeser (2004, see also Alesina *et al*., 2001) argue that two factors are key to understanding the difference across the Atlantic: political institutions and ethnic heterogeneity.

Regarding political institutions, Alesina and Glaeser argue that the fact that almost all European countries (with the notable exceptions of the UK and France) have proportional representation systems have contributed strongly to the redistributive differences between the US and Europe. According to them, this system facilitated the growth of communist and social democratic parties that support increased redistribution and a large welfare state. The majoritarian American system, on the other hand, made it difficult for such parties (which at the beginning where very small) to get any representatives elected. The American federalism and decentralization, which is stronger than what all European countries except Switzerland display, may also have contributed. Alesina and Glaeser (2004) conclude that institutional differences can explain about half of the difference in social spending between Europe and the US.

The main explanatory factor for the remaining half of the difference in redistribution between Europe and the US is, according to Alesina and Glaeser (2004), differences in racial and ethnic fractionalization. The US has historically been a much more racially and ethnically fragmented society than European countries. Such fractionalization appears, in turn, to make people demand less redistribution, especially when poverty is concentrated to minority groups. Given the changes in the direction of more ethnic fractionalization that the Nordic countries (to a varying degree) are experiencing, it is important to understand what impact this may have on the Nordic model with its extensive redistribution.

10.3.1 Ingroup/outgroup bias

A person's social identity is often partly defined as the sense of self that she derives from perceived membership in a social group. The seminal work of Billig and Tajfel (1973, see also Tajfel 1982 and Tajfel *et al*., 1971) was the starting point for the empirical research on intergroup behavior in laboratory settings. They showed that people have a tendency to put both themselves and others into categories and that this categorization gives rise to favorable treatment of the people in the same social group as oneself (the ingroup) compared to those in other groups (the outgroup), i.e. people exhibit an ingroup bias. The way the groups are created has been shown to matter for the strength of the ingroup bias, but also completely random formation of groups give rise to the effect (Billig and Tajfel, 1973; and Locksley *et al*., 1980).

The first experiments investigating the ingroup/outgroup-effect were done with participants dividing a valuable asset (usually money or lottery tickets) between another member of the ingroup and a member of the outgroup, i.e. they could not allocate anything to themselves (see Tajfel and Turner, 1986; and MacDermott, 2009 for surveys). Experimental economists have also considered situations, both in laboratory and in field settings, where the individual herself has money at stake in the decision, for example in the form of a dictator game or a prisoners' dilemma. Bernard *et al*. (2006), Goette *et al*. (2006) and Chen and Li (2009) show that an ingroup bias arises in such situations as well, with people behaving more altruistically and cooperatively towards people in their ingroup (see also Mollerstrom, 2015).

10.3.2 Evidence on the importance of racial/ethnical heterogeneity for demand for redistribution

There is significant evidence that feelings of altruism and willingness to redistribute in general weaken across racial and ethnic lines. Especially when those with lower income consist disproportionately of people of racial or ethnical minorities, the majority prefers less redistribution (Alesina and Giuliano, 2010). We may not like it, but it is widely observed that individuals are more generous towards others who are

similar to themselves racially, ethnically and linguistically (see also Alesina *et al*., 1999; Alesina *et al*., 2004; Luttmer, 2001; and Fong and Luttmer, 2011).

The data in Table 3 come from Alesina and Glaeser (2004) and show evidence on racial, ethnic, linguistic and religious fractionalization across Europe and the US. Fractionalization is measured on a range between 0 and 1, and a lower value indicates more homogeneity. For all measures, except linguistic fractionalization, the US exhibits significantly more fractionalization than what the European countries included in the table do on average. The three Nordic countries included in the table, Denmark, Norway and Sweden, are even less fractionalized.[3]

Alesina and Glaeser (2004, see also Alesina *et al*., 2001) go on to show that the correlation between racial fractionalization and social spending as a share of GDP is statistically significant and strongly negative for a global sample of 52 countries. This indicates that the more fractionalized a country is, the smaller is the share of GDP that is allocated to social welfare spending. The results is equally strong and statistically significant if only high-income countries are included. Racial fractionalization is the best predictor of social welfare spending, but the results for other types of fractionalization are similar. Alesina and Glaeser (2004) also present similar results regarding the relation between fractionalization and social spending from the 52 US states.[4]

The evidence presented above is to a large extent correlational (with the exception of the laboratory studies which show a causal relation between heterogeneity and demand for redistribution). There are, however, also field studies that attempt to study the causal relation between ethnic fractionalization and preferences for redistribution. One influential study, conducted in Sweden, is Dahlberg *et al*. (2012). They use a Sweden-wide program for placing refugees in municipalities to

[3] As discussed below, these patterns are changing. More recent data for the Nordic countries are considered in Table 5.

[4] Lindqvist and Östling (2010) have in a related paper showed that there is a negative correlation between size of government and political polarization. Their paper does not study ethnic fractionalization directly, but the results are of interest to this context as there is a positive relation between ethnic fractionalization and political polarization.

generate the exogenous variation needed to study the relationship between increased immigration and demand for redistribution.

Table 3: Fractionalization indices

	Racial fractionalization	Ethnic fractionalization	linguistic fractionalization	Religious fractionalization
Austria	0.03	0.11	0.15	0.41
Belgium	0.05	0.56	0.54	0.21
Denmark	0.02	0.08	0.10	0.23
France	0.10	0.10	0.12	0.40
Germany	0.06	0.17	0.16	0.66
Italy	0.02	0.11	0.11	0.30
Netherlands	0.11	0.11	0.51	0.72
Norway	0.06	0.06	0.07	0.20
Portugal	0.05	0.05	0.02	0.14
Spain	0.03	0.42	0.41	0.45
Sweden	0.05	0.06	0.20	0.23
Switzerland	0.05	0.53	0.54	0.61
United Kingdom	0.10	0.12	0.05	0.69
Average:	0.06	0.19	0.23	0.40
United States	0.49	0.49	0.25	0.82

Source: Alesina and Glaeser (2004). Data from 1990–2000.

The starting point of Dahlberg *et al*. (2012) is the fact that immigration of workers and refugees to European countries in general (and some of the Nordic countries in particular) have increased substantially over time. They ask how this affects native citizens' views on redistribution and the size of the welfare state. The main contribution of their paper is that they, in contrast to the (predominantly negative) correlational evidence that several other studies have documented between immigration and demand for redistribution (in addition to Alesina *et al*., 2001 and Alesina and Glaeser, 2004, see also Shayo, 2009; Stichnoth and Van der Straeten, 2013; Harmon, 2014),[5] claim to be able to study the causal relationship.

During the years 1985 to 1994, a refugee placement program was in place in Sweden to achieve a more even distribution of refugees over

[5] Note that even though the majority of studies find a negative correlation between immigration and the demand for redistribution, there are some paper which find no correlation (e.g. Gerdes and Wadensjö, 2008; Brady and Finnigan, 2013) or a conditional negative one (e.g. that the negative correlation only holds when immigrants are overrepresented among those with lower income, e.g. Finseraas, 2012).

the country. 277 out of the 286 Swedish municipalities participated, neither the refugees nor the municipalities could affect the placement and the total arrival of refugees in the relevant years exceed 200,000. Given these circumstances the identifying assumption of Dahlberg *et al*. (2012), that the placement of refugees was exogenous with respect to the inhabitants of the municipalities' preferences for redistribution, seem justified.

The authors match the municipal-level data on refugee placements, immigration shares, and other municipal covariates, with individual level survey information. The survey data come from the Swedish National Election Studies Program which is a rotating panel. The measure of individual preferences for redistribution was extracted from a survey question regarding if a person would be "in favor of decreasing the level of social benefits."

Using the refugee placement program as an instrument for the share of immigrants living a particular municipality, Dahlberg *et al*. (2012) show that there is a negative and statistically significant relation between increasing the share of immigrants in a municipality and how much social benefits people in that municipality prefer to see. Moreover, the effect is especially pronounced among white-collar, high-income earners, meaning that the respondents who contribute more extensively to the redistribution scheme are those whose support for redistribution is reduced as the group of potential recipients become more ethnically diverse (Dahlberg *et al*., 2012, p. 69).[6]

10.3.3 Potential underlying mechanisms

Having concluded that most research indicates that there is a negative correlation between the increased ethnic fractionalization caused by immigration and support for redistributive policies, and that this relation may well be causal, the question arises what the underlying mecha-

[6] It should be noted that the results of Dahlberg *et al*. (2012) have also invoked criticism. See e.g. Nekby and Pettersson-Lidbom (forthcoming), who claim that the measure of demand for redistribution used in Dahlberg *et al*. (2012) is suboptimal, that the results are not generalizable to the full Swedish population, and that the measurement of the refugee placement program is flawed.

nisms are. Existing research suggests a number of potential mechanism, but very few studies have attempted to disentangle them. Here we will discuss three potential mechanism.

First, it may be the case that altruism is lower when directed towards people of different race or ethnicity as compared to when it is directed to people who are more similar to the person offering the help. This is suggested and confirmed in, for example, Fong and Luttmer (2011), where it is showed, in a setting with charitable giving, that people perceive those who are more similar to themselves as more worthy of help.

A second hypothesis is that it is not altruism, but rather trust and reciprocity that is impacted by ethnic fractionalization. Trust and reciprocity are both positively related to willingness to redistribution. The former because you need to trust that people are not taking unfair advantage of the redistributive system in order to demand extensive redistribution, and the latter because you have to be convinced that if you were the one in need of help, others would provide you with it, just as you would if you are in the position to help. In general, differences between people, both ethnic, racial and other, seem to generate distrust (see e.g. Glaeser *et al*., 2000; Zak and Knack, 2001; Alesina and La Ferrara, 2002; Knack and Zak, 2002, Dienesen and Sønderskov, 2015. The work of Alesina and Zhuravskaya (2011) is also related although it has a slightly different focus). Gustavsson and Jordahl (2008) consider this in a Nordic setting (Sweden) and find a similar negative association between proportion of people born in a foreign country on the one hand and trust and reciprocity on the other hand.

Third, it is also possible that larger interpersonal differences, for example in the form of ethnic or racial difference, impacts how people view themselves and the likelihood that they may be on the receiving end of the redistributive system in the future. As concluded by Benabou and Ok (2001), people generally do not care only about their current position in the income distribution when their demand for redistribution is formed – their beliefs about their future position also matters. Finseraas (2012) find evidence that it is especially the people with a higher income who demand less redistribution when fractionalization is high. He presents evidence that suggests that this does not happen because of a decrease in altruism or in trust. Instead, people with a higher

income are less concerned with downward income mobility, and believe that the likelihood that this will happen to them, is lower when ethnic fractionalization is higher.

These hypotheses are not necessarily exclusive and all three mechanisms may be at play when the negative relation between immigration and demand for redistribution arises. For those interested in affecting this relation, the policy implications of the various mechanisms are however very different. Hence, research is most likely going to continue to try to disentangle these and other underlying mechanisms in order to gain a better understanding of exactly why we observe a negative relation between immigration and ethnic fractionalization, and demand for redistribution.

10.4 The Nordic countries: moving towards more heterogeneity

Up until World War II, Denmark, Finland, Norway and Sweden were all very ethnically homogenous in comparison with most other European nations. Since then, the paths have been diverging. The differentiation began during the war, when Sweden took in a substantial number of refugees. After World War II, immigration was to a large extent intra-Nordic, but at the beginning of the 1960s demand for labor increased in Sweden, Norway and Denmark which led to substantial immigration from other parts of Europe. Finland also delivered a substantial number of emigrants to Denmark, Norway and Finland (Yousfi, 2010; Kouvo and Lockmer, 2012).

Thereafter, immigration to Sweden, Norway and Denmark has consisted of refugee migration and family reunification. After the fall of the Soviet Union, and after joining the European Union, migration flows to Finland have increased rapidly with migrants predominantly coming from the former Soviet area. (Kouvo and Lockmer, 2012.)

Regarding regulations, the Nordic countries also exhibit considerable differences. Sweden has been, by far the most permissive whereas Finland and Norway, and especially Denmark, are more regulated. Taking more types of policies into account, including the allowance of dual

citizen-ships and the funding of ethnic group organizations or activities, Kymlicka and Banting (2006) classify 21 OECD countries according to their multiculturalism. Sweden is the only Nordic country having modest multiculturalism policies, whereas Denmark, Finland and Norway are categorized as having weak multiculturalism policies.

The Nordic countries are also different in the sense that the anti-immigration political movement grow strong earlier in Denmark and Norway as compared to Finland and Sweden (Kouvo and Lockmer, 2012). Currently, there is substantial anti-immigration sentiment in all four countries, however, and this is currently gaining more strength. Despite recent tightenings, Sweden still has the most generous policies on immigration. Considering recent trends, all Nordic countries are experiencing a rise in the inflow of immigrants in general and refugees in particular and the trend for the proportion of the population that is foreign, or foreign-born, is positive in most of the countries. These trends are illustrated in Table 4 below.

Table 4: Recent trends in immigration for the Nordic countries

	2005	2010	2012	2013
A. Immigration, per 1,000 citizens				
Denmark	3.7	6.0	6.3	7.5
Finland	2.4	3.4	4.3	4.4
Norway	6.8	13.3	13.9	13.8
Sweden	5.7	8.4	8.7	9.9
B. Inflow of asylum seekers, per 1,000 citizens				
Denmark	0.4	0.9	1.1	1.4
Finland	0.7	0.7	0.5	0.6
Norway	1.2	2.1	1.9	2.4
Sweden	1.9	3.4	4.6	5.6
C. Foreign-born, percentage of population				
Denmark	6.5	7.7	8.2	8.6
Finland	3.4	4.6	5.3	5.6
Norway	8.2	11.6	13.2	14.5
Sweden	12.5	14.8	15.5	16.0
D. Foreign, percentage of population				
Denmark	5.0	6.2	6.7	7.2
Finland	2.2	3.1	3.6	3.8
Norway	4.8	7.6	8.9	10.0
Sweden	5.3	6.8	7.0	7.2

Source: OCED (2015).

As panel A of Table 4 shows, Norway has the highest immigration among the four Nordic countries. However, this consists to a large extent of temporary workers, often from the other Nordic countries. Panels A and B together paints the picture of significantly more asylum seekers in Sweden than in any of the other Nordic countries. All countries except Finland are experiencing a rising trend, however.

As for the stock of immigrants, the percentage of foreign-born (panel C) is highest in Sweden, followed by Norway. The trend is positive in all four countries. The fact that Sweden is more generous regarding naturalization as compared the other three countries is evident from the data in panel D in relation to panel C: fewer people remain foreign in Sweden as it is easier to eventually become a citizen. Again, the trend is that of a rising share of foreign population in all the four countries.

Taken together, these data support the view of Sweden as the most immigration-friendly Nordic country, in particular when it comes to immigrants with a different ethnic background (which is more common among refugees than among other types of immigrants in general). All four Nordic countries are experiencing an increase in immigration and the population is becoming more diverse. This ethnic fractionalization is most and least manifest in Sweden and Finland, respectively.

10.5 Concluding remarks

What the Nordic model, with its extensive redistribution of economic resources between citizens, will look like in the future depends on the preferences of the voters in the Nordic countries. Historically, they have demanded more redistribution than voters in almost all other parts of the world. Preferences for redistribution are not fixed however but change over time.

The academic quest for understanding how individual demand for redistribution is shaped has provided ample insights. We know that more inequality in itself generally leads to a higher demand for redistribution, that women want more redistribution than men, that people with higher income, more education and higher cognitive ability tend to

demand less redistribution and that a person who is very risk-averse is generally going to want more redistribution than a person who is more willing to take on risks.

We also know that the heterogeneity of a country's population matter for how much redistributions its citizens want to see. More heterogeneous groups have been shown to be less generous and cooperative in laboratory settings – also when the heterogeneity is artificially created. This holds empirically also when considering political preferences within countries. More racial, ethnic and literacy fractionalization, for example due to historical factors and to immigration, is generally found to be negatively associated with demand for redistribution, and there are reasons to believe that this relation is causal.

The Nordic countries are, to varying degrees, experiencing a rise in ethnic heterogeneity. This may lead the Nordic model to change in the direction of less redistribution, simply because citizens are likely to demand less of it. The underlying mechanism is not clear, but it is probable that feelings of altruism, trust and reciprocity are involved. It is also likely that especially the part of the native population with a higher income regard it as more unlikely that they will be on the receiving end of the redistributive system and hence demand less economic redistribution.

A striking example (which in itself is, of course, only illustrative and does not provide a proof in any way) is the fact that at the same time as Denmark has employed restrictive immigration policies it has overtaken Sweden as the Nordic country with the highest taxes. Sweden, on the other hand, which has a very generous immigration policy, has witnessed a more pronounced trend towards lower taxes and a less extensive welfare state.

Needless to say, there are also other trends that are impacting demand for redistribution in the Nordic countries. For example, income inequality in general is increasing which in itself may lead to higher demand for redistribution and thus a mitigation of the effect discussed above. Also, as immigrants become eligible to vote, it is unclear in which direction this will impact demand for redistribution. What the net effect will be is hence impossible to say. Given the rise in anti-immigration sentiment in the Nordic countries, we can however conclude that it is a cur-

rently an extremely salient topic; the likelihood that it has a substantial impact on preferences for redistribution is therefore likely to be high.

We may not like that more heterogeneity reduces the willingness to redistribute – but this is what is observed empirically. Some would like to see a situation where substantial heterogeneity is combined with extensive redistribution, and there are of course many individuals who hold those preferences. But research indicates that the majority of people behave differently. And hence the question that Alesina and Glaeser (2004, p. 181) is more topical than ever: "We shall see whether the generous [Nordic] welfare state can really survive in a heterogeneous society." If current trends continue, it is, in light of the research presented here, unlikely that the exceptionally high levels of redistribution that have historically been demanded in the Nordic countries will persist.

To what extent these changes in demand for redistribution are perceived as problematic depend, of course, on individual views on the optimal size of the welfare state. For a person who is generally positive to immigration and at the same time wishes to see a smaller welfare state, the negative relation between ethnic heterogeneity and demand for redistribution does not necessarily pose a problem.

On the other hand, for a person who is likewise positive to immigration but wants the welfare state to remain large, there is a challenge. Importantly and as discussed above, research tells us that the negative correlation between immigration and redistributive preferences is especially pronounced when ethnic minorities are overrepresented among those who are dependent on the welfare state. This implies that the integration of immigrants is crucial as it, as is making it possible for them to be successful in the labor market. Doing this successfully may weaken the negative link between ethnic heterogeneity and demand for redistribution, and could hence make it easier to successfully combine generous immigration policies with a continued high demand for an extensive welfare state.[7]

[7] Unfortunately the experiences of immigrants on the Nordic labor markets is not particularly positive, something which is further discussed by Bratsberg and Røed (2015) in this volume. They present facts regarding labor market participation and welfare dependency for immigrants to the Nordic countries, and discuss policy changes that could make the welfare state more robust to handling extensive immigration.

References

Alesina, A., & Giuliano, P. (2010). Preferences for redistribution. In Jess Benhabib, Matthew O. Jackson and Alberto Bisin (eds.), *Handbook of Social Economics*, 1A:93–131.

Alesina, A., & La Ferrara, E. (2002). Who trusts others? *Journal of Public Economics*, 85(2): 207–234. http://dx.doi.org/10.1016/S0047-2727(01)00084-6

Alesina, A., Baqir, R, & Hoxby, C. (2004). Political Jurisdictions in Heterogeneous Communities. *Journal of Political Economy*, 112(2): 348–396. http://dx.doi.org/10.1086/381474

Alesina, A., Baqir, R., & Easterly, W. (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics*, 114(4): 1243–84. http://dx.doi.org/10.1162/003355399556269

Alesina, A., E. Glaeser, & B. Sacerdote (2001). Why Doesn't the United States Have a European-Style Welfare State? *Brookings Papers Econ. Activity*, no. 2: 187–254. http://dx.doi.org/10.1353/eca.2001.0014

Alesina, A., Tella, R. D., & MacCulloch, R. (2004). Inequality and happiness: Are Europeans and Americans different? *Journal of Public Economics*, 88(9–10): 2009–2042. http://dx.doi.org/10.1016/j.jpubeco.2003.07.006

Alesina. A., & Glaeser, E. (2004). *Fighting Poverty in the US and Europe – A World of Difference*. Oxford University Press, Oxford, UK. http://dx.doi.org/10.1093/0199267669.001.0001

Alesina, A., & Zhurayskaya, E. (2011). Segregation and the Quality of Government in a Cross Section of Countries. *American Economic Review*, 101(5): 1872–1911. http://dx.doi.org/10.1257/aer.101.5.1872
Atkinson, A.B., & Morelli, S. (2014). *The Chartbook of Economic Inequality*. www.chartbookofeconomicinequality.com. Accessed on 11 September 2015. http://dx.doi.org/10.2139/ssrn.2422269

Benabou, R., & Ok, E.A. (2001). Social Mobility and the demand for redistribution. *Quarterly Journal of Economics*, 116(2): 447–487. http://dx.doi.org/10.1162/00335530151144078

Benabou, R., & Tirole, J. (2006). Belief in a just world and redistributive politics. *Quarterly Journal of Economics*, 121(2): 699–746. http://dx.doi.org/10.1162/qjec.2006.121.2.699

Bernard, H., Fehr, E., & Fischbacher, U. (2006). Group Affiliation and Altruistic Norm Enforcement. *American Economic Review*, 96(2): 217–221. http://dx.doi.org/10.1257/000282806777212594

Billig, M., & Tajfel, H. (1973). Social Categorization and Similarity in Intergroup Behavior. *Journal of Social Psychology*, 3(1): 27–52.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1): 166–193. http://dx.doi.org/10.1257/aer.90.1.166

Brady, D., & Finnigan, R. (2014). Does Immigration Undermine Public Support for Social Policy? *American Sociological Review*, 79(1): 17–42. http://dx.doi.org/10.1177/0003122413513022

Brandal, N., Bratberg, Ø., & Thorsen, D.E. (2013). *The Nordic Model of Social Democracy*. Palgrave MacMillan, New York, NY, USA. http://dx.doi.org/10.1057/9781137013279

Bratsberg, B. & Røed, K. (2015). The Nordic Welfare Model in an Open European Labor Market. *Nordic Economic Policy Review*: this volume.

Chen, Y., & Xin Li, S. (2009). Group Identity and Social Preferences. *American Economic Review*, 99(1): 431–457. http://dx.doi.org/10.1257/aer.99.1.431

Cruces, G., Perez-Truglia, R., & Tetaz, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98: 100–112. http://dx.doi.org/10.1016/j.jpubeco.2012.10.009

Dahlberg, M., Edmark, K., & Lundqvist, H. (2012). Ethnic Diversity and Preferences for Redistribution. *Journal of Political Economy*, 120(1): 41–76. http://dx.doi.org/10.1086/665800

Dinesen, P. T., & Sønderskov, K. M. (2015). Ethnic Diversity and Social Trust: Evidence from the Micro-Context. *American Sociological Review*, 80(3): 550–573. http://dx.doi.org/10.1177/0003122415577989

Durante, R., Putterman, L., & van der Weele, J. (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association*, 12(4): 1059–1086. http://dx.doi.org/10.1111/jeea.12082

Elinder, M., Jordahl, H., & Poutvaara, P. (2015). Promises, policies and pocketbook voting. *European Economic Review*, 75(April): 177–194. http://dx.doi.org/10.1016/j.euroecorev.2015.01.010

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3): 817–868. http://dx.doi.org/10.1162/003355399556151

Finseraas, H. (2012). Poverty, ethnic minorities among the poor, and preferences for redistribution in European regions. *Journal of European Social Policy*, 22(2): 164–180. http://dx.doi.org/10.1177/0958928711433655

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, 82(2): 225–246. http://dx.doi.org/10.1016/S0047-2727(00)00141-9

Fong, C., & Luttmer, E. (2011). Do Race and Fairness Matter in Generosity? Evidence from a Nationally Representative Charity Experiment. *Journal of Public Economics*, 95(5–6): 327–394. http://dx.doi.org/10.1016/j.jpubeco.2010.07.010

Gaertner, M., Mollerstrom, J., & Seim, D. (2015) *Risk and redistribution*. Mimeo, Stockholm University.

Gerdes, C., & Wadensjö, E (2008). *The Impact of Immigration on Election Outcomes in Danish Municipalities*. IZA Working Paper 3586.

Glaeser, E. Laibson, D., Sheinkman, J., Soutter, C. (2000). Measuring Trust. *Quarterly Journal of Economics*, 115(3): 811–846. http://dx.doi.org/10.1162/003355300554926

Goette, L., Huffman, D., & Meier, S. (2006). The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups. *American Economic Review*, 96(2): 212–216. http://dx.doi.org/10.1257/000282806777211658

Gustavsson, M., & Jordahl, H. (2008). Inequality and trust in Sweden: Some inequalities are more harmful than others. *Journal of Public Economics*, 92: 348–365. http://dx.doi.org/10.1016/j.jpubeco.2007.06.010

Harmon, N. (2015). *Immigration, Ethnic Diversity and Political Outcomes: Evidence from Denmark*. Mimeo.

Harsanyi, J.C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61: 434–435. http://dx.doi.org/10.1086/257416

Hilson, M. (2008). *The Nordic Model*. Reaktion Books, London, UK.

Knack, S., & Zak, P.J. (2002). Building trust: public policy, interpersonal trust, and economic development. *Supreme Court Economic Review*, 10: 91–107. http://dx.doi.org/10.1086/scer.10.1147139

Kouvo, A., & Lockmer, C. (2012). Imagine all the Neighbours: Perceived Neighbourhood Ethnicity, Interethnic Friendship Ties and Perceived Ethnic Threat in Four Nordic Countries. *Urban Studies*, 50(16): 3305–3322. http://dx.doi.org/10.1177/0042098013484538

Kymlicka, W., & Banting, K. (2006). Immigration, multiculturalism, and the welfare state. *Ethics & International Affairs*, 20(3): 281–304. http://dx.doi.org/10.1111/j.1747-7093.2006.00027.x

Lindqvist, E., & Östling, R. (2010). Political Polarization and the Size of Government. *American Political Science Review*, 104(3): 543–565. http://dx.doi.org/10.1017/S0003055410000262

Locksley, A., Ortiz, V., & Hepburn, C. (1980). Social Categorization and Discriminatory Behavior: Extinguishing the Minimal Intergroup Discrimination Effect. *Journal of Personality and Social Psychology*, 39(5): 773–783. http://dx.doi.org/10.1037/0022-3514.39.5.773

Logue, J. (1979). The Welfare State: Victim of its Success. Daedalus, 108(4): 69–78.

Luttmer, E. (2001). Group Loyalty and the Taste for Redistribution. *Journal of Political Economy*, 109: 500–528. http://dx.doi.org/10.1086/321019

MacDermott, R. (2009). Psychological Approaches to Identity: Definitions, Measurement and Experimentation. In Abdelal, R., Y. Herrera, A. Johnson, & R. McDermott (eds.), *Measuring Identity: A Guide for Social Scientists,*. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511810909.013

Margalit, Y. (2013). Explaining social policy preferences: Evidence from the great recession. *American Political Science Review*, 107(01): 80–103. http://dx.doi.org/10.1017/S0003055412000603

Meltzer, A. H., & Richard, S. F. (1981) A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927. http://dx.doi.org/10.1086/261013

Mollerstrom, J. (2015). *Favoritism reduces cooperation*. Mimeo, George Mason University.

Mollerstrom, J., & Seim, D. (2014). Cognitive ability and the demand for redistribution. *PloS one*, 9(10): e109955. http://dx.doi.org/10.1371/journal.pone.0109955

Morelli, S., Smeeding, T., & Thompson, J. (2014). *Post-1970 Trends in Within-Country Inequality and Poverty: Rich and Middle Income Countries*. IRP Discussion Paper, No. 1419-14.

Mounir, K., Mollerstrom, J., & Seim, D. (2014). *Richer (and Holier) than Thou? The Effect of Relative Income Improvements on Demand for Redistribution*. IFN Working Paper No. 1042.

Nekby, L. & Pettersson-Lidbom, P. (forthcoming). Revisiting the Relationship between Ethnic Diversity and Preferences for Redistribution. *Scandinavian Journal of Economics*.

OCED (2015). *International Migration Outlook 2015*. OECD Publishing, Paris, France.

Pelzman, S. (1985). An economic interpretation of the history of congressional voting in the twentieth century. *American Economic Review*, 75(4): 656–678.

Piketty, T. (1995). Social mobility and redistributive politics. The *Quarterly Journal of Economics*, 110(3): 551–584. http://dx.doi.org/10.2307/2946692

Popenoe, D. (1994). *Scandinavian Welfare. Society*, 31(6): 78–81. http://dx.doi.org/10.1007/BF02693281

Powdthavee, N., & Oswald, A. J. (2014). *Does money make people right-wing and inegalitarian? A longitudinal study of lottery winners*. Warwick University Economics Working Paper 1039.

Rainer, H., & Siedler, T. (2008). Subjective income and employment expectations and preferences for redistribution. *Economics Letters*, 99: 449–453. http://dx.doi.org/10.1016/j.econlet.2007.09.011

Roberts, K. (1977). Voting over Income Tax Schedules. *Journal of Public Economics*, 8(3): 329–340. http://dx.doi.org/10.1016/0047-2727(77)90005-6

Rodrik, D. (1998). Why do More Open Economies have Bigger Governments? *Journal of Political Economy*, 106(5): 997–1032. http://dx.doi.org/10.1086/250038

Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics*, 4(2): 163–185. http://dx.doi.org/10.1016/0047-2727(75)90016-X

Sachs, J.D. (2006). *The Social Welfare State, Beyond Ideology*. Scientific American, 16 November.

Sanandaji, N. (2015). *Scandinavian Unexceptionalism*. The Institute of Economic Affairs, London, UK.

Sears, D. O., & Funk, C. L. (1990). The limited effect of economic self-interest on the political attitudes of the mass public. *Journal of Behavioral Economics*, 19(3): 247–271. http://dx.doi.org/10.1016/0090-5720(90)90030-B

Shayo, M. (2009). A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution. *American Political Science Review*, 103: 147–174. http://dx.doi.org/10.1017/S0003055409090194

Stichnoth, H., & Van der Straeten, K. (2013). Ethnic Diversity, Public Spending, and Individual Support for the Welfare State: A Review of the Empirical Literature. *Journal of Economic Surveys*, 27(2): 364–389. http://dx.doi.org/10.1111/j.1467-6419.2011.00711.x

Tajfel H., Billig M., Bundy R., & Flament C. (1971). Social categorization and intergroup behaviour. *Eur J Soc Psychol* 1:149–178. http://dx.doi.org/10.1002/ejsp.2420010202

Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, 33: 1–39. http://dx.doi.org/10.1146/annurev.ps.33.020182.000245

Tajfel, H., & Turner, J. (1986). The Social Identity Theory of Intergroup Behavior. In (eds. Stephen Worchel, and William Austin) *The Psychology of Intergroup Relations*, 7–24. Chicago: Nelson-Hall. http://dx.doi.org/10.1146/annurev.ps.33.020182.000245

Wang, C., Caminada, K., & Goudswaard, K. (2014). Income Redistribution in 20 Countries over Time. *International Journal of Social Welfare*, 23: 262–275. http://dx.doi.org/10.1111/ijsw.12061

World Value Survey (2015). http://www.worldvaluesurvey.org. Accessed on 15 September 2015.

Yousfi, S. (2010). The development of immigration. In: R. Andersson, H. Dhalmann, E. Holmqvist, T. M. Kauppinen *et al*. (eds.) *Immigration, Housing and Segregation in the Nordic Welfare States*, pp. 271–274. University of Helsinki Press, Helsinki, Finland.

Zak, P.J., & Knack, S. (2001). Trust and growth. *Economic Journal*, 111(470): 291–321. http://dx.doi.org/10.1111/1468-0297.00609

11. The Social Upper Class under Social Democracy

Kalle Moene, Department of Economics, University of Oslo[1]

In *Capital in the 21st Century* Thomas Piketty explores the history of the distribution of income and wealth. Among other things he demonstrates how top income shares have soured recently. Although top income shares are somewhat lower in the Nordic countries, the recent developments follow a similar pattern as elsewhere (Roine and Waldenström, 2015). So also in countries characterized as social democratic the rich obtain a higher share of a growing pie.

These trends in the distribution of income and wealth may have political consequences. Piketty rightly insists "that economic and political changes are inextricably intertwined and must be studied together", and those who have a lot of wealth, he continues, "never fail to defend their interests" politically (p. 577). The concentration of income and wealth in the hands of the rich, may therefore give them more political clout as well.

Yet without discussion Piketty seems to assert that the distribution of political power is linked directly to the distribution of top income shares. I agree that the prosperous rich are the political powerful in most countries, but I am less confident that the share of income to the richest x percent is a good measure of the political power of the upper class. In this essay I therefore emphasize not only how rich the rich are

[1] I'm grateful for discussion with Rolf Aaberge, Halvor Mehlum, Debraj Ray and Gaute Torsvik. I have benefited from useful comments by Torben Andersen and efficient assistance by Kristian Harald Myklatun. This paper is part of the research activities at ESOP, a research center funded as a center of excellence by the Norwegian Research Council.

relative to the rest of society, but also how many rich people there are relative to the size of the population. Accounting for the number of rich – for instance the share of people with wealth above y – gives us another picture of the wealth concentration in the Nordic countries than what is normally associated with social democracy. It gives us a picture of a larger upper class, at least in Sweden and Norway, compared to other countries.

The upper class is the class composed of the wealthiest members of society who also wield the greatest political power, Wikipedia rightly points out. The social democratic upper class is not a class of the wealthiest social democrats, but the class of the richest people created by social democracy. Many of its members are critical of welfare spending and small income differentials even though these social democratic institutions may have laid the foundation for their affluence and for the size of the upper class.

11.1 Prosperity, power and poverty

All too many of us have a great respect for people with money. The social habit is not new. Writing in 1759, Adam Smith was worried about "the corruption of our moral sentiments, which is occasioned by this disposition to admire the rich and the great, and to despise or neglect persons of poor and mean condition" (*Moral Sentiments*, p. 84). He also pointed to social and political consequences. In my reading, the respect for money includes the respect for the wealthy, and therefore also for the number of wealthy people.

The number of wealthy people is even more important in democratic societies than in the autocratic ones that Adam Smith considered. The division into groups and classes is still important, and so is how much resources each of them has. Yet a person who owns everything, has less power than a group that owns everything.

The wealthy Norwegian family Løvenskiold is a case in point. The family owns the huge forest and recreational area, Nordmarka outside Oslo, a local public good. Obviously the family could make a lot of money by developing the area commercially for housing, shopping

centers and businesses. It is equally clear, however, that if the family tried, it would most likely never get the permission to develop the area. Democratic institutions can easily control one family. But if the same area had thousands of wealthy owners, they would have a much easier time getting permission to develop the area. Numbers count under democratic rule.

The influence of the rich is well illustrated in a study by Martin Gilens and Benjamin Page (2014). They look at the connection between public opinion polls and political decisions in the US, based on 1779 cases or propositions in the period 1981 to 2002. The study shows that whether ordinary people support or oppose a proposition has little influence for whether the proposition is adopted. If the richest tenth of the population support the proposition, however, it has a higher probability of being adopted.

How many wealthy people that support the proposition plays a significant role. A proposition supported by one in five wealthy people, has only 18% chance of being adopted, while a proposition supported by four out of five wealthy people, has a 45% chance of being adopted. These correlations are worth noting even though we don't quite know whether the propositions were adopted because the rich supported them, or not. Interpreted with the same caveat, the paper also shows that the rich have an effective veto right over politics. If they are against a proposition, the chances that the proposition is adopted are very small.

All in all the study reminds us that "votes count, but resources decide", as Stein Rokkan said in 1966. The influence of the rich seems to increase when they are many, and when there is a large gap between the rich and the rest (see for example Bartels, 2008, Gilens, 2005 and 2009, Peters and Ensik, 2014).

The interests of a high number of wealthy people count both because they are wealthy, and because they are many. This is important for how to measure their potential power. It is not only their share of the national income that matters. Measuring richness, I suggest, is similar to measuring poverty. In the low end of the income distribution we could also look at income shares, for instance the share of the national income that goes to the z poorest percent of society. Such a measure of poverty, however, tells us nothing about how many people

who live in poverty – and thus how long the hunger march can become. To get an approximation of the magnitude of the problem, it is better to look at the so-called head count measure of poverty. This tells us the number of people with a lower income than a certain threshold, the poverty line, and who therefore are not able to achieve the socially decided minimum level of material well being – the lowest level required to appear in public without shame, as Adam Smith said.

In the same way, we should also look at the number of extremely wealthy people, the number of people with a wealth beyond a certain threshold which we could call the richness line – perhaps the highest level tolerated to appear in public without shame. To measure wealth as the relative number of wealthy people in the population – the head count measure of richness – should be as natural as measuring the relative number of poor people. In practice, of course, it is not obvious where the richness line should be set, just as it is not obvious where the poverty line should be. There is no objective distinguishing line, neither for extreme poverty nor for extreme richness.

11.1.1 Dollar billionaires

We can use the Forbes list of billionaires as an illustration. It uses an implicit richness line of wealth equal to USD 1 billion, which is the threshold wealth to be included on the list. By using the Forbes list of the world's billionaires, we can examine which countries have the highest number of billionaires, relative to the population. Is US ahead of the Nordic countries? The answer is not obvious. We know that the wealthiest persons in the US are wealthier than the wealthiest persons in the Nordic countries, For example, the 0.1% richest in the US get 11% of the national income, while the 0.1% richest in Norway get 2.5% of the national income. So according to these numbers social democracy can be viewed as effective in abolishing extreme wealth.

Wealth concentration represented by the head count measure of richness, however, provides an entirely different picture. While the US has 1.7 billionaires per million inhabitants, the corresponding numbers for the Nordic countries are:

- Denmark 0.9.
- Finland 0.9.
- Norway 2.0.
- Sweden 2.4.
- Iceland 3.1.

In comparison the head count measure of billionaires in other European countries is in the range of 0.5 to 1.3 per million inhabitants. For instance, Spain has 0.4, Italy 0.6, France, 0.7, UK 0.8 and Germany 1.2. So, compared to the rest of Europe there are more billionaires relative to the population in the Nordic countries.

As seen, Norway, Sweden and Iceland have even more billionaires relative to the population than United States. Measured by the head count measure of richness, the social democratic model in Norway and Sweden creates more wealthy people than the American model. Studying the Forbes list, we find that only countries that are pure tax havens, or would like to become tax havens – Iceland, Singapore, Switzerland Cyprus, Hong Kong, St. Kitts and Nevis and Monaco – have a higher share of billionaires than Norway and Sweden.

In sum, all this indicates that the social democratic upper class can be larger, relatively speaking, than the upper class in most countries, and in particular it can be larger than the American upper class. But Forbes operates with a very high richness line. The question is whether we find the same pattern when we use a lower richness line of wealth, and when we account for differences in national income per capita.

Being in the start of a research project on these issues, my plan is to explore several data sets to compare countries and the development over time. At this stage I can offer preliminary illustrations. In addition to the observations above based on the Forbes list, I can add more

detailed observations based on another data source. I focus on the comparison between Norway and the US.

11.1.2 An example: the super rich in Norway and in the US

In this illustration I use numbers from *The Wealth Report 2014*, published by the international consultancy Knight Frank. Statistics on wealth are often incomplete, and usually build on rough approximations. Like Forbes list, the Wealth Report 2014 is no exception. It is also difficult to check the quality of the numbers, but the report has a good reputation.[2]

The Wealth Report provides numbers on the wealth distribution for a many countries, including the US and Norway, concentrating on wealthy individuals. In the discussion below I use two richness lines: the wealth levels of USD 30 million and USD 100 million. Those with higher wealth than 30 million are called rich, those with wealth higher than 100 million are called super rich. The two richness lines are not corrected for purchasing power. A purchasing power adjustment would only show how much the wealth is worth when it is exclusively used to purchase goods and services in the country of its owner.

Per million inhabitants in 2013 we have:

- 124.6 rich and 36.5 super rich individuals in the US.
- 472.8 rich and 63.6 super rich individuals in Norway.

In other words, Norway has almost four times as many rich individuals per capita compared to the US, and almost two times as many super rich.

When we measure richness as the relative number of rich people in the population, it does not look like social democracy limits the size of the upper class. On the contrary, a social democratic system seems to increase the recruitment to this class. Relative to the population, social

[2] I have started checking the numbers, comparing them to information from other sources. There is a clear spread in the reported numbers (surveys) that should make us less confident. Yet, the basic pattern that I focus on in this essay seems to prevail in the data I've checked so far.

democracy creates more rich and super rich than the American model does. But are these differences only a reflection of the fact that Norway is richer per capita compared to the US? And, should we not therefore correct for this fact?

It is reasonable to assume that richer countries have more rich and super rich members than not quite so rich countries. However, we do not correct the head count measures of poverty because some countries have a higher level of income than others. It is far from obvious that we should do this when measuring and evaluating the head count meausre of richness.

Yet, it is still interesting to know whether the US would have the same relative number of rich and super rich if the US had the same income per capita as Norway. Does the fact that Norway is richer than the US explain why Norway has more multimillionaires than the US?

11.1.3 If the US was as rich as Norway

Norway has an income per capita of USD 65,461, while the US has USD 53,042, such that Norway is 23.5% richer than the US. We will now try to estimate how many rich and super rich individuals would have been lifted above the richness line if the US was 23.5% richer than what the country is today. Since the wealth distribution has a thin tail for the very wealthy, the relative increase in the number of multimillionaires, is usually disproportionately large compared to the increase in the income per capita. A ten percent increase in income would give a larger than ten percent increase in the number of multimillionaires.

By how much does the share of rich and super rich individuals increase as income goes up? The answer depends on three things:

Firstly, we have to know how large the income growth will be for the Americans that are almost rich, and almost super rich, as the per capita income in the US increases by 23.5%. I emphasize three alternatives where the income of the rich and the super rich increase by a percentage which is a) equal to the growth of national income, b) twice as high as the national income growth, c) triple the size of the national income growth.

Secondly, we have to know how much higher wealth this increase in income gives us. As an approximation I assume that the wealth is proportional to income, such that wealth increases at the same rate as income.

Thirdly, we have to know how fat the tail of the wealth distribution is, as this determines how many almost rich and almost super rich that are lifted above the richness line. Here we can utilize that the Pareto-distribution can give a good approximation of the upper part of the wealth distribution. The Pareto-distribution is determined by two parameters, one that controls how fat the tail is, so we can find out how many almost rich and almost super rich that pass the richness line when their incomes increase. In the simple calibration I perform I use the shares of the population with wealth of at least USD 30 million, and with a wealth of at least USD 100 million (see the appendix for details).

Following this recipe, the results should be compared to the case of Norway with 472.8 rich, and 63.6 super rich per million inhabitants. I get the following hypothetical upper class in the US:

- If the income growth of all wealthy individuals was equal to the growth of national income, the US would have: 154.5 rich, and 45.2 super rich per million inhabitants.
- If the income growth of all wealthy individuals was twice the growth in national income, the US would have: 184.5 rich, and 54 super rich per million inhabitants.
- If the income growth of all wealthy individuals was three times the growth of national income, the US would have: 215.3 rich, and 63.1 super rich per million inhabitants.

Even if the income of the rich grows at three times the rate of the rest, the US would not have a higher share of rich or super rich than Norway. So if we choose US as the bench-mark for the link between income and wealth, we would have to conclude that Norway has a higher share of rich and super rich individuals than the income per capita predicts. In comparison with the US, Norwegian social democracy creates a higher share of people

with wealth around NOK 700–800 million, and in particular a higher share of people with wealths around NOK 200–300 million.

I have also tried other methods, including calibrating the parameters of the Pareto-distribution to the inequality of the entire wealth distribution expressed by the Gini coefficient – and I obtain similar results (see the appendix). The Gini of the wealth distribution in the US is slightly above 0.8. The large wealth inequality that such a high Gini coefficient indicates, means that there is a low number of rich and super rich individuals that would be lifted over the richness line when their incomes increase. Norway has a lower Gini coefficient around 0.6 for the wealth distribution. This indicates that Norway has a higher density of rich and super rich individuals who would be lifted upwards when their incomes grow.

The higher shares of rich and super rich in Norway does not mean that Norway has more concentration of wealth than the US. Both in the US and in Norway the population share with wealth over 30 million USD possesses a considerable share of the entire wealth in the nation. Norway has more wealth equality among the rich than the US, such that Norway has a higher share of rich and super rich individuals in the population than the US. In other words, the US has fewer super rich than Norway, but they own a greater share of the national wealth.

To be clear, if one person owned the entire wealth of the country, the Gini coefficient would be equal to 1. The head count measure of richness would then be approximately equal to zero, and increasing the income would not raise the share of rich individuals. Of course, the concentration of wealth is not that high in the US. But the example illustrates that the US can have a so high concentration of wealth that income growth only lifts rather few people above the richness line.

11.2 Why so many wealthy?

Is the wealth of the social democratic upper class old or new, inherited or recently accumulated? A check of the family background of the members of the richest groups in Norway (done together with the newspaper Dagens Næringsliv, indicates that the upper class to a large

extent, but not entirely, is produced under social democracy. Going through the background of the 100 richest persons according to the Norwegian tax register in 2014, reveals that less than one third of them had inherited their wealth, while more than two thirds of them had built up their wealth under social democratic rule.

If this is true more generally, it remains to be explained how social democracy produces so many wealthy persons. In my view it follows naturally from the social democratic development strategy of combining social protection with capitalist dynamics. Contrary to what most people believe, the strategy does not represent a compromise between socialist and capitalist ideals that leads to neither. Rather the two ideals are complementary in the sense that we get more out of each by combining them.[3] This complementarity can also give more multimillionaires and a larger upper class once we incorporate how wage equality can raise the sum of asset values to capital, and how social mobility can expand the numbers of capital owners.

Over time, new technologies and new organizations replace the old and outdated. The goal of the Scandinavian union movement has for long been to speed up this process of creative destruction, as Joseph Schumpeter called it, under a policy of maintaining full employment. The strategy could also be seen as an attempt to raise long term real wages by wage restraint in the short run.

Unions have followed a wage policy that raises the profitability of investments and modernization by holding back the highest wages in the name of solidarity. The goal has also been to make the export sector more competitive. Both goals were reached by taking the wage determination out of market competition and placing it in a system of collective decision making. The result was a highly compressed export-lead wage structure, eliminating employment rents in high productivity jobs and raising the pay in low productivity jobs. Over time the new wage structure has also altered the composition of high and low productivity jobs.

---

[3] Together with co-authors I have for long argued this. See for instance Moene and Wallerstein (1993) and Barth, Wallerstein and Moene (2003), and Barth, Moene and Willumsen (2014).

The isolated effect of wage restraint in high-skilled jobs is to raise profits and thus to induce more investments and in turn higher demand for all types of labor. The unions could therefore raise the lowest wages without increasing unemployment. Aiming for small wage differentials and full employment lead to more equal wages. It also lead to higher profits, increasing the asset values of capital owners. The social democratic growth strategy therefore raised the total wealth of the upper class.

The high number of wealthy people is in part explained by this expansion of wealth and in part by the access to free higher education for all and the corresponding social mobility that enabled more people to take advantage of the possibilities to become rich. Individuals without inherited wealth simply has a higher probability of succeeding economically in the Nordic countries than in the US. The Nordic countries have more equal opportunities. For instance, the chances that children of manufacturing workers in the 1960–70s end up in a higher class of income than their parents are considerably higher than in the US. The probability that a son with a father in the bottom fifth of the income distribution actually ends up in the top fifth of the income distribution, is 8% in the US, and more than 12% in Norway (Jänti *et al*., 2006).

Some might argue that the mobility in the Nordic countries is higher than in other countries because income equality is more pronounced. Thus, the differences between high and low income is relatively small, and the distance between classes is low, which enables more people to move upwards socially and financially. This is not a counter argument, however. It is part of the explanation for what we observe – that the upper class is largest where the differences among workers are smallest.

11.3 Policy implications or implied policy?

Political actors may be worried or enthusiastic about the size of the upper class. Both groups may nevertheless consider to adjust wealth taxes, inheritance taxes and other re-distributive means to achieve their political goals – if they can. Policy implications, however, are never as straight forward as most economists assert. Economists easily

exaggerate the independence of policy – and pay too little attention to implied policy by endogenous forces in society.

In our case one relevant issue is what it really means to ask for the policy implications if the distribution of wealth to some extent determines policies – which again may affect the ability of the rich to accumulate wealth. When money gives power, a higher gap between the wealthy and the rest may imply more unequal power and influence over policy. The reinforcing development can be strong: the political influence of the rich may affect the economy; the economic changes may make it easier to become rich.

If this is right, more inequality might lure us into a trap where those who should be the object of regulation, becomes those who in fact regulate – and where the new multi-millionaires influence the economic policies in their own favor. Wealth may bring power, and power may bring more wealth to the wealthy – such that the rich grow both in number and in wealth.

To illustrate the possible consequences of having a big upper class, let us return to the paper by Gilens and Pages and take their numbers literally: The fraction of wealthy people supporting or opposing a political proposition, is imperative to whether it passes or not. The numbers suggest that a quadrupling of the support by the richest 10% of Americans, increases the chances of the proposition passing, by a factor of 2.5. If we use this factor on the numbers we have calculated for the relative size of the upper classes in the US and Norway, we find that the rich in Norway potentially has 2.4 times the influence of the rich in the US.

Of course this approximation is too simpleminded. There are also opposing forces as can be seen by the fact that almost equally rich countries have very different income distributions – one hundredth of the population gets almost 25% of the income in the US, but less than 8% in Norway (Aaberge and Atkinson, 2011). One reason for the difference may be that labor unions represent a countervailing power to the political and economic influence that the rich otherwise would have.

The logic of this countervailing power is simple: When workers are poorly organized, it is much easier to become exploited by a big capitalist

than by many smaller ones. When workers are well organized, in contrast, it is easier for them to tame a few big capitalists than many small.

Some central union leaders, at least in Norway and Sweden, may think of the capitalists as their saving machines, and the large retained profits as the savings of the union members. For this reason, they have not found it beneficial to squeeze the capitalists too hard. As mentioned they may have thought that some wage restraint of the best paid workers would benefit the great majority of union members via the investments that the retained profits generated.

In this context capitalist consumption is a cost. Few and big savings machines may thus also be viewed as more cost efficient than many and small savings machines. Union leaders might also for that reason have preferred to have few and large capitalists and thus more concentrated ownership.

The comparison between the US and Norway shows that countries do not always get the upper class that fits the best. Norway, with its strong union associations has a large upper class with more dispersed wealth. The US, in contrast, has a more concentrated and top-heavy upper class, but almost no unions in the private sector. In the US, where it is easier to buy influence through gifts to political campaigns, political advertising and otherwise, the concentrated wealths are a threat to democracy. The inequality in Norway could also become a similar threat, especially when unions are in decline.

The richest capitalist in Sweden has long been the Wallenberg family. At times, the family has been seen as a social democratic capitalist light, acting as the saving machine for the Swedish union members. The most famous capitalist in the US, in contrast, John D. Rockefeller, was perceived more as a threat to democracy than as a saving machine of the country. Even when he started to donate parts of his fortune, it could be seen as an unreasonable concentration of power over what should be financed.

Perhaps the clearest policy implications for those who are concerned with the low level inequality of the Nordic model, is to emphasis the importance of a well organized labor market – both on the workers' side and on the employers' side. A strong union movement is decisive to reduce the power that a large upper class otherwise would

imply. A strong employers' association is important to voice long term producer interests as opposed to short term distributional share-owner interests. A strong employers' association is also important as a threat against sheltered unions that otherwise might prefer to operate alone and not coordinate with the unions in traded goods industries.

The irony is, perhaps, that the structure of organized interests that can prevent the bad political consequences of a large upper class also may lay the foundation for its expansion. Similarly, while the goal of social democracy as a political movement used to be to work for socialist egalitarian ideals, the simple empirics in this essay indicate that the social democratic combination of worker security and capitalist dynamics may generate not only high growth and small wage differentials, but also a large upper class. Having a larger upper class than the US teach us that it cannot be particularly difficult to become rich under social democracy in Norway and Sweden. A large upper class may also be a sign of financial success.

Yet, it may undermine democratic ideals. In practice, of course, democracy always entails a combination of one person one vote and one dollar one vote. The weights on each depend on the social organization and on the wealth distribution. When many individuals become wealthy, the political system may put less weight on the average vote of citizens. In this way social democracy in Norway and Sweden can be a victim of its own success.

References

Aaberge, R. & Anthony B. Atkinson (2010). Top Incomes in Norway. In A.B. Atkinson and T. Piketty (eds.), *Top Incomes A Global Perspective,* Oxford: Oxford University Press, 448–482.

Bartels L.M (2008). *Unequal Democracy: the political Economy of the New Gilded Age*. Princeton: Princeton University Press.

Barth E., Moene, K., & Wallerstein, M. (2003). *Likhet under press: utfordringer for den skandinaviske fordelingsmodellen*. Oslo: Gyldendal Akademisk.

Forbes list. The World's Billionaires. www.forbes.com/lists/(Accessed 15 November 2015).

Gilens, M. (2005). Inequality and Democratic Responsiveness. *Public Opinion Quarterly*, 69(5):778–896. http://dx.doi.org/10.1093/poq/nfi058

Gilens, M. (2009). Preference Gaps and Inequality in Representation. *PS: Political Science and Politics*, 42(2): 335–341. http://dx.doi.org/10.1017/s1049096509090441

Gilens, M. & Benjamin, I. P. (2014). Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens. *Perspectives on Politics*, 12(03), 564–581. http://dx.doi.org/10.1017/S1537592714001595

Jänti, M., Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Österbacka, E., Björklund, A., & Eriksson, T. (2006). *American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States*. IZA Discussion Paper no. 1938.

Moene, K. & Wallerstein, M. (1993). *What is Wrong with Social Democracy*. In Pranab Bardhan and John E. Roemer (eds.), *Market Socialism. The present debate*. Cambridge, Cambridge University Press.

Barth, E., Moene, K., & Willumsen, F. (2014). The Scandinavian Model – an interpretation. *Journal of Public Economics*. http://dx.doi.org/10.1016/j.jpubeco.2014.04.001

Ensik S.J. & Peters, Y. (2014). *Differential Responsiveness in Europe. The Effects of Preference Difference and Electoral Participation*. West European politics.

Piketty, T. (2014). *Capital in the Twenty-First Century.* Cambridge, MA: Harvard University Press. Statistical series and technical appendix. http://piketty.pse.ens.fr/ capital21c.

Roine, J. & Waldenström, D. (2015). Long-Run Trends in the Distribution of Income and Wealth. In A.B. Atkinson and F. Bourguignon, F. (eds.), *Handbook of Income Distribution,* vol. 2A. Amsterdam: North-Holland. http://dx.doi.org/10.1016/b978-0-444-59428-0.00008-4

Smith, A. (1759). *The Theory of Moral Sentiments* Library of Economics and Liberty. 26 November 2015. http://www.econlib.org/library/Smith/smMS.html

The Wealth Report 2014. (2014). London: Knight Frank.

Appendix

To see the impact of a higher national income (in the US) on the share of rich and super rich people we ustilize that the Pareto-distribution is a good approximation to the wealth distribution.

We first use how the Pareto-distribution fits around the two richness lines of 30 and 100 millions USD, but not necessary for the entire wealth distribution. Let the share of people with wealth higher than x millions USD (in the US) be $r(x)$ given by

$$r(x) = Ax^{-\alpha} \tag{1}$$

where $\alpha > 0$ and $A > 0$. With the numbers we have for the US we can then calculate

$$r(30) = \frac{39.378}{316100} = 0{,}000125 \quad \text{and} r(100) = \frac{11.544}{316100} = 0{,}0000365 \tag{2}$$

Using (1) and (2) we have two equations to "determine" A og α. This calibration yields

$$A = 0{,}004 \quad \text{and} \quad \alpha = 1.02 \tag{3}$$

We also assume that wealth is proportional to income y implying that a λ-doubling of the national income raises the income of the rich and super rich by the factor λ.

Clearly if the income increases by a factor 1.235, people with wealth between $x/1.235$ and x would be lifted up to a wealth higher than x.

If the income growth of all wealthy individuals was equal to the growth of national income of 23.5%, $r(30/1.235) = 0.0001544$ rich, and $r(100/1{,}235) = 0{,}0000452$ super rich.

If the income growth of all wealthy individuals was twice the growth in national income, i.e. 47%, $r(30/1.47)=0.0001845$ rich, and $r(100/1.47)=0.000054$ super rich.

If the income growth of all wealthy individuals was three times the growth of national income, i.e. 71%, $r(30/1.71)=0.0002153$ rich, and $r(100/1.71)=0.0000631$ super rich.

Alterantively we could calibrate α in the Pareto-distribution such that it fits to the rough estimates of the Gini coefficient for welath equal to 0.8 in the US, assuming that the Pareto-distribution fits fits for the entire wealth distribution.

The Gini for the wealth distribution can be written

$$G=\frac{1}{2\alpha-1} \tag{4}$$

If $G\approx 0.8$ we have $\alpha\approx 1.1$, (not far from our calibrated value of 1.02.

Taking expectations

$$Ex=\frac{\alpha}{\alpha-1}A^{\frac{1}{\alpha}} \tag{5}$$

Utilizing that the average wealth is proportional to average income, i.e. $Ex=kEy$, we obtain

$$r(x,Ex)=\left[\frac{(\alpha-1)}{\alpha}\frac{kEy}{x}\right]^{\alpha} \tag{6}$$

We see that a λ-doubling of Ey yields

$$r(x,\lambda Ex)=\lambda^{\alpha}r(x,Ex) \tag{7}$$

The calculation shows that if the US increased its income by 23.5%: the US would have had 163.8 rich and 47.8 super rich per million inhabitants.

Calibrating the Norwegian wealth distribution in the same manner, using NBER's estimate of $G = 0.633$ for Norway, we get $\alpha = 1.29$. an increase in the mean income in Norway with 23.5% would raise the share of rich and super rich by a factor $(1.235)1.29 = 1.31$. this is far from enough to explain that Norway has a much higher share of rich and super rich than the US.

By calibrating $Ax^{-\alpha}$ locally so that it fits for $r(30)$ and $r(100)$ for Norway, we get $\alpha = 1.67$ og $A = 0.037$. The high value of α here shows that Norway has a much thicker tail in the wealth distribution implying that a rise in income relative to the US leads to more rich and super rich in Norway compared to the US.

TemaNord 2016:503

norden

Nordic Council of Ministers

Ved Stranden 18
DK-1061 Copenhagen K
www.norden.org

Nordic Economic Policy Review

The Nordic Economic Policy Review is published by the Nordic Council of Ministers. This year's issue is part of the Danish presidency programme for the Nordic Council of Ministers in 2015. The review addresses policy issues in a way that is useful for informed non-specialists as well as for professional economists. All articles are commissioned from leading professional economists and are subject to peer review prior to publication.

The review appears once a year. It is published electronically on the website of the Nordic Council of Ministers: www.norden.org/en. On that website, you can also order paper copies of the Review (enter the name of the Review in the search field, and you will find all the information you need).

TemaNord 2016:503
ISBN 978-92-893-4450-0 (PRINT)
ISBN 978-92-893-4452-4 (PDF)
ISBN 978-92-893-4451-7 (EPUB)
ISSN 0908-6692



NORDIC ECONOMIC POLICY REVIEW NUMBER 2 / 2015