Assorted Essays in Economics

Inequality, Labor Supply and Gender Roles

PhD dissertation by

Jakob Egholt Søgaard

January 2015

Presented at the Department of Economics

Faculty of Social Sciences, University of Copenhagen

Acknowledgment

There are so many people that I need to thank and without whom this PhD thesis would not have been completed that I hardly know where to start.

Firstly I would like to thank the Ministry of Finance for financing this PhD and letting me keep my connection to the Macro Economic Policy Center thoughout the PhD process. I have really enjoyed the involvement in actual policy making that unavoidably follows from having an office in the heart of the Danish central administration.

Secondly I would like to thank my supervisor professor Claus Thurstrup Kreiner, and my co-author on my first paper professor Anthony Atkinson (Oxford University) for taking a novice PhD student under their wings and introducing me to the world of science.

The same thanks go to my co-authors on my third paper professor Henrik Kleven (LSE) and professor Camille Landais (LSE). One of the challenges of being a reseacher in Denmark is to keep track of reseach frontier, but working with people like Henrik and Camille this is never a problem, and the work on our joint project over the last year has been super inspiring. Henrik furthermore provided me with the possibility to visit LSE in the auturm 2012, which set me on a steep learning curve early in my PhD - something that I have been grateful for ever since.

I would also like to thank Louise, Peer, Thomas, Anders, Mette, Anna and all the other PhD students and faculty members at the Department of Economics. The atmospheres created by you at the department is simply incredible and I always feel at home there - something that I do not take as given, considering that I am partly an outsider who show up very sporadically.

Finally I would lile to thank my family and inparticular Astrid and Julie for bearing with me in this - as well as in all other of life's aspects.

English summary

The chapters collected in this PhD dissertation consist of 3 self-contained papers within the fields of public finance and labor economics. Even though this means that it is not required to read the entire dissertation to understand the individual chapters, readers who do so, will find that there are some similarities between the chapters. Firstly, all chapters utilize data from Denmark – something that is unintentional, but perhaps not unexpected given the richness and quality of the Danish data and in particular the Danish register data. Secondly, the argumentation in all chapters relies to a large extent on visualization of data – a style that this summary also will follow by summarizing each chapter in one key graph.
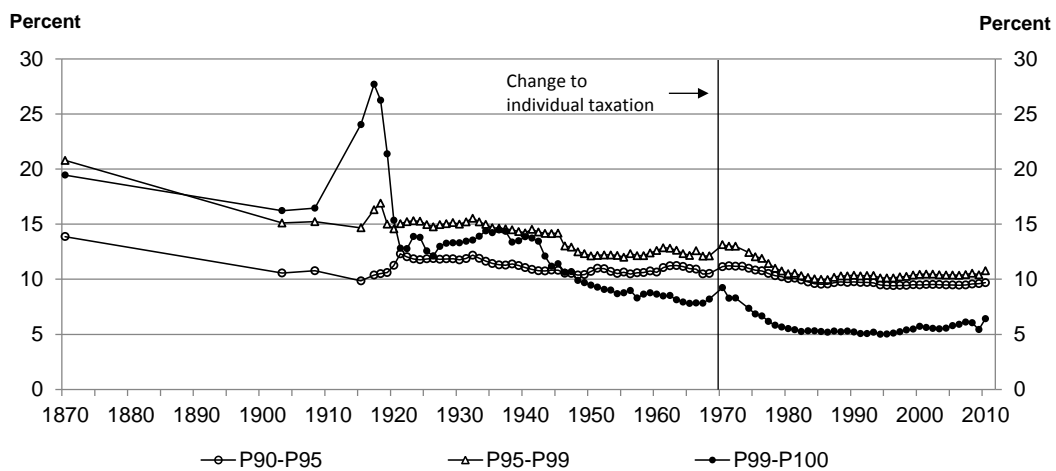
Chapter 1 – Long run income inequality

The first chapter written jointly with Anthony Atkinson focuses on the development in income inequality in Denmark over the past 100-140 years and should be seen as a contribution to the literature sparked by Thomas Piketty and other some 15 years ago and which so far has culminated – at least in the public – with his book "Capital in the 21st Century". What makes this literature special is that a large number of researchers around the world have worked to create a database of comparable inequality estimates for as many countries and as far back in time as possible. This type of coordinated effort is not that normal in economic research.

The key graph in this chapter shows the development in the income shares of different income groups in Denmark over the last century. The graphs show that the income share of the top percentile (P99-P100) drops from around 15-20 percent at the beginning of the 20th century to 6-7 percent today. This drop implies that the 1 percent of the population with the highest income has gone from an income that on average was 15-20 times higher than the mean income in the population to a factor ranging between 6 and 7 today. In other words, the graph shows that income inequality has dropped significantly over the last century.

Key Graph 1

The development in income shares in Denmark 1870-2010: Top 10 percentiles

Percent   **Percent**

Notes: See chapter 1

It may appear strange to measure income inequality by only considering the income share of the top income earners and not the incomes of everybody else. The choice is also primarily driven by data limitations. The reason is that the sources behind these inequality series are tabulations of the income distributions, which became available in many countries when progressive income taxation was introduced in the beginning of the 20th century. However, in the beginning it was typically only people in the upper part of the income distribution, who initially paid income taxes and thus were included in the statistics. This makes it infeasible to calculate broader inequality measures such as the modern day Gini coefficient.

This is however not the case for Denmark, where the data cover a large part of the income distribution almost from the beginning of the 20th century, and one of the contributions, we make in the chapter, is to examine who "gained" from the drop in the top income share and whether the development in the top income share is a good proxy for the development in the overall inequality level.

We find that the top income share – at least for Denmark – indeed is a good proxy for the development in the overall inequality level, and that in particular individuals below the 70th percentile in the income distribution benefited from the decline in the top income share, while the "upper middle class" (individuals between the 70th and the 95th percentile) did not. In contrast, the income share of the latter has been remarkably stable throughout the last 100 years, cf. key graph 1.

Chapter 2 – Labor supply and optimization frictions

The second chapter focuses on the effect of taxation – or incentives broadly speaking – on labor supply. Empirically, this literature has undergone relative large changes during the last 10-15 years due to the increasing availability of large amounts of high quality register data. Something, which has highlighted a number of weaknesses in the way labor supply responses was estimated previously.

Simply put: 30 years ago researchers would write up a complicated model with optimization agents and estimate the model parameters so that the model would fit observed behavior in data sets containing only a few 100 observation of self-reported data. Today researchers typically employ register data from the tax authorities covering the entire population and this gives the possibility to test some of the prediction of the labor supply models, which was impossible to test previously.

One of these predictions is that people should tend to cluster (or bunch as it is called) at points in the tax schedule, where the marginal tax rate jumps up, e.g. at the point where you start to pay top tax in Denmark. In standard labor supply models such bunching follows unavoidably, when individuals react to taxes by reducing their labor supply, but as Emmanuel Saez show in his 2010 paper, there is hardly any bunching to be found in actual income distributions.[1] Something that would imply that individuals did not react to taxes, which again would have stark consequences for the way we think about tax policy.

Following the findings of Emmanuel Saez, there has been a research agenda lead by in particular Raj Chetty to reconcile the missing bunching (among other things) with individuals actually responding to taxes, and the main explanation given in this literature is that individuals' actions are subject to optimizations frictions – i.e. that they for various reasons do not choose the actions that maximize their utility according to the standard economic models. In his 2011 and 2012 papers, Raj Chetty shows that small optimizations frictions in labor supply decisions – such as e.g. that it takes time and effort to search for a new job – are enough to explain both the missing bunching and other paradoxes in the labor supply literature.

The contribution of my research to this literature is to shed light on the precise nature of these optimization frictions, and to do so I consider the Danish student labor market as a case study. So far concrete evidence on frictions has been relatively limited in the economics literature on labor supply, which reflects that identification of optimization frictions typically requires both high quality data and special institutional settings –

---

[1] The paper already came out in 1999 as a working paper.

high quality data in order not to confound optimization errors by individuals with measurement error in the data and special institutional settings that allow separation of rational behavior from optimization errors.

Both of these requirements are met with the Danish student labor market, and the key graph in this chapter shows the earnings distribution of Danish students 3 years prior and 3 years after the reform of the student benefit system in 2009. The central feature of the Danish student benefit system is that students receive benefits, which are phased out with earnings above a given threshold, and this phase out creates a large jump in the effective marginal tax rate at this point. Yet, consistent with the findings of Emmanuel Saez, there is no visible bunching in the earnings distribution: neither before the reform (where the jump happens at 0 DKK on the x-axis in the graph) or after the reform (where the jump happens at 18,000 DKK). So, taken to the standard labor supply models this would imply that students do react to taxes.

Key Graph 2

The income distribution for tertiary students before and after the 2009 reform
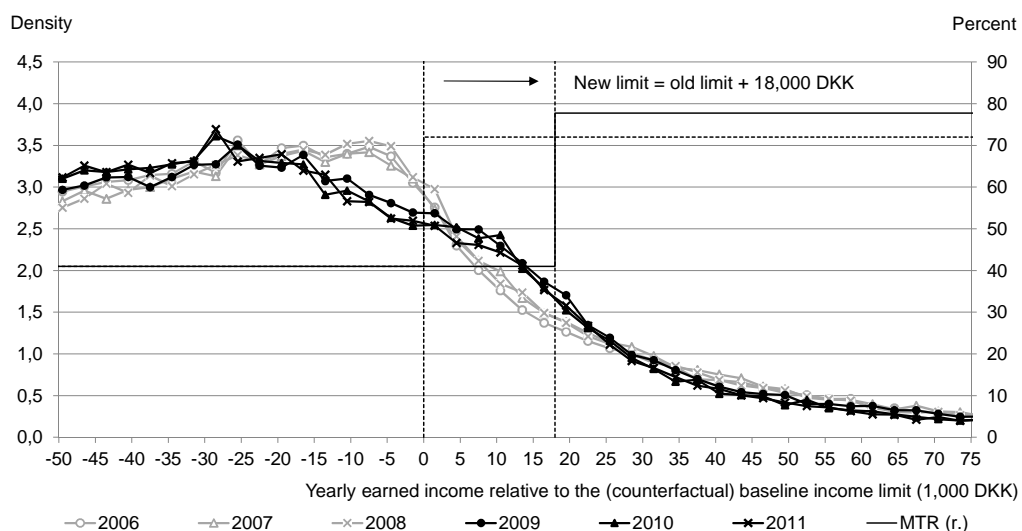
Notes: See chapter 2.

However, despite the lack of bunching, the graph also shows a non-trivial shift in the earnings distribution with individuals moving from a range below the initial point, where the marginal tax rate jumps to a range above. Considering how stable the earnings distribution is both before and after the reform this finding is to me a strong indication of a labor response to taxation and that something in people's behavior is masking the bunching that you should expect to find.

Next in the chapter I turn to investigate what this "something" is and here I consider 3 of the main competing explanations in the literature, namely: 1) real adjustment costs (it takes time and effort to find a new job), 2) gradual learning (it take time to learn exactly how the tax and benefit system works) and 3) inattention (unexpected things happen during the year, which people do not pay attention to and thus fails to counteract with reoptimization).

While all of these optimization frictions are likely to play a role in all labor markets I find inattention to be the dominating friction and that the effects are large. A finding that is already hinted by the key graph, as the reform effect appears instantaneously, while both real adjustment costs and gradual learning predict a more gradual response.

Chapter 3 – Gender inequality

In the third chapter written jointly with Henrik Kleven and Camille Landais I return to the issue of inequality, but this time with the focus on earnings differences between men and women. It turns out that even though there has been considerable narrowing of the earnings differences between men and women in many countries over the past 30 years there is still a sizable gap remaining. In Denmark in 2011 the median earnings of full-time employed women is 15 percent lower than the median for men and this finding is surprisingly similar across countries.

The aim of the chapter is to analyze the how the arrival of children affects the earnings of men and women respectively and our main finding is shown in the key graph below. This graph is constructed using all child births by first time parents in Denmark from 1985 to 2001, where data on both parents are available in the registers in all 5 years prior to the birth of their first child and 10 years after. This leaves us with a sample of around 350,000 births or 11,200,000 individual-year observations.

Key Graph 3

Male and female earnings around the time of the birth of their first child
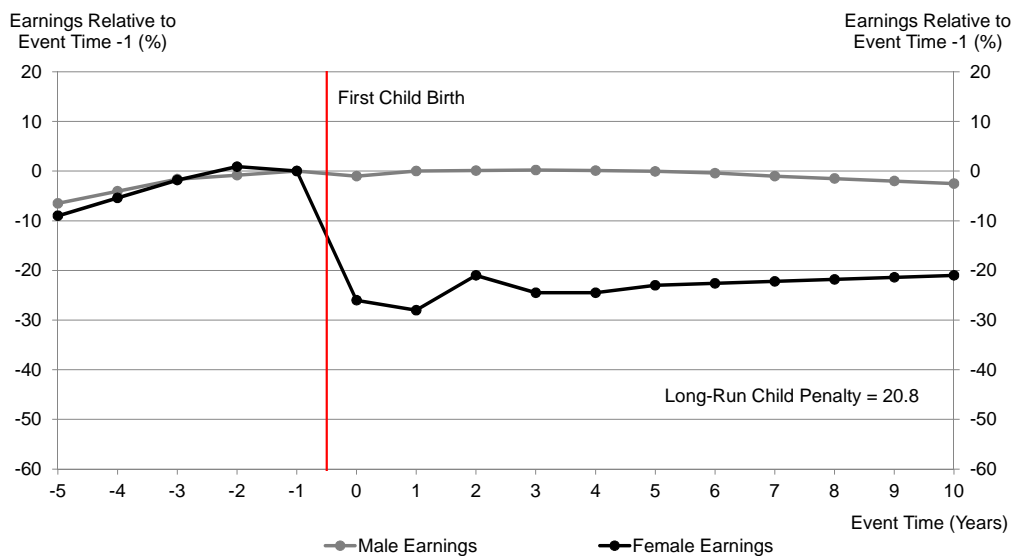
Notes: See chapter 3.

In the figure we have plotted the development in male and female earnings around the birth of their first child, and this reveals a distinct drop in female earnings of 25-30 percent just after the birth of the first child. This is perhaps not surprising given that women still take the majority of the parental leave, but what is surprising is that there is very little convergence back in female earnings. In contrast, after 10 years female earnings are still 20 percent below those of men, whose earnings are largely unaffected by the arrival of a child.

Interpreting this figure, it seems to be the case that women are changing career paths after the arrival of children, which puts them on a lower earning path than their male partners. Something that is also confirmed when we look at other measures of career progression such as e.g. the probability of becoming a top manager in a firm. Instead we find that women increasingly select into "family friendly" firms after the arrival of a child, which on average is likely to result in lower earnings.

Taking these estimated "child penalties" back to the gender earnings gap in the whole population, we find that most of the remaining gender gap can be attributed to the effect children have on the women's earnings. More concretely, we estimate that 80 percent of the remaining gap in earnings between men and women in 2011 can be attributed to the differential effect of child on male and female earnings. In 1980 the corresponding number is 30 percent.

It is of course an open question why men and women react so differently to the arrival of children. One explanation is that women have a comparative advantage in taking care of children, so that the couples rationally choose to have the women focusing on childcare. Another explanation is that the observed behavior is a result of a norm in society that women ought to take care of the children. While we do not come up with a definitive answer to this question, we do find that there is a significant correlation between the relative working history of the mother's parents and the child penalty that she experiences, i.e. in families, where the grandmother worked very little compared to the grandfather, the mother experiences a larger earnings penalty than mothers in families, where the grandmother worked relatively more compared to the grandfather. This could indicate that the women tend to inherent their own mothers' preferences for professional work relative to child caring.

Dansk resume

Kapitlerne samlet i denne ph.d.-afhandling består af 3 selvstændige forskningsartikler inden for områderne offentlige økonomi og arbejdsmarkedsøkonomi. Selvom det betyder, at det ikke er nødvendigt at læse hele afhandlingen for at forstå de enkelte kapitler, vil læsere, der læser det hele, finde, at der er visse ligheder mellem kapitlerne. For det første anvender alle kapitlerne data fra Danmark - noget der er utilsigtet, men måske ikke uventet i betragtning af den mængde og kvalitet der kendetegner de danske data og især de danske registerdata. Og for det andet bygger argumentationen i alle kapitlerne i vid udstrækning på visualisering af data - en stil, som dette resumé også vil følge ved sammenfatte hvert kapitel i én hovedfigur.

Kapitel 1 – Langsigtet indkomstulighed

Det første kapitel, som er skrevet sammen med Anthony Atkinson, fokuserer på udviklingen i indkomstuligheden i Danmark gennem de sidste 100-140 år. Kapitlet skal ses som et bidrag til den litteratur, der blev startet af Thomas Piketty m.fl. for ca. 15 år siden, og som indtil videre har kulmineret - i hvert fald i det offentlige - med hans bog "Kapital i det 21. århundrede". Denne litteratur er speciel, da et stort antal forskere i hele verden har arbejdet på at oprette en database med sammenlignelige skøn for uligheden for så mange lande og så langt tilbage i tiden som muligt. Denne type af koordineret arbejde er ikke så udbredt i økonomisk forskning.

Hovedfiguren i dette kapitel viser udviklingen i forskellige indkomstgrupper andel af den samlede indkomst i Danmark i det sidste århundrede. Heraf fremgår det, at indkomstandelen for den øverste percentil (P99-P100) falder fra omkring 15-20 procent i begyndelsen af det 20. århundrede til 6-7 procent i dag. Dette fald betyder, at den rigeste procent af befolkningen er gået fra en indkomst, der i gennemsnit var 15-20 gange højere end gennemsnitsindkomst for hele befolkningen til 6-7 gange i dag. Figuren viser med andre ord, hvordan indkomstulighed er faldet betydeligt i løbet af det seneste århundrede.

Hovedfigur 1

Udviklingen i indkomstandele i Danmark 1870-2010: Top 10 percentiler
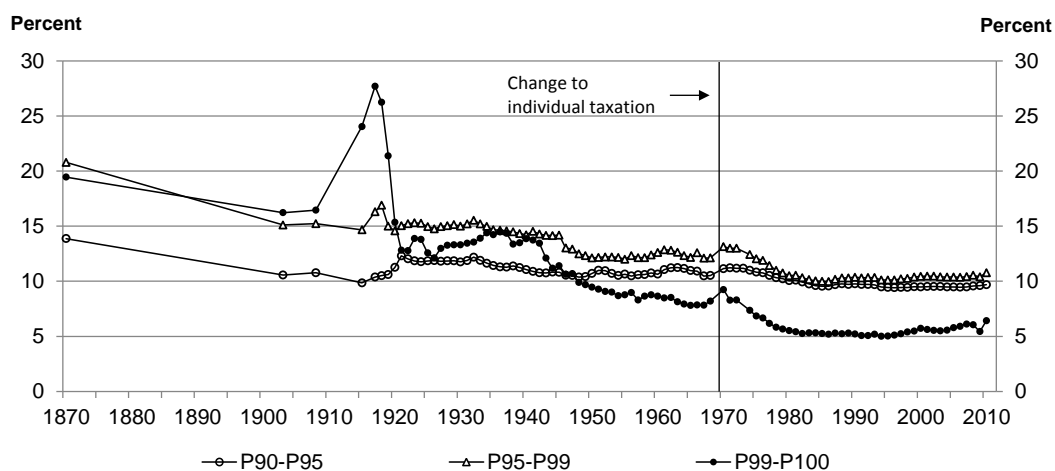
Anm: Se kapitel 1.

Det kan forekomme mærkeligt at måle indkomstuligheden ved kun at betragte de rigestes indkomstandel og ikke indkomstfordelingen blandt alle andre. Valget er da også primært drevet af begrænsninger i det data, der er tilrådelig. Årsagen er, at kilden til disse ulighedsserier er skattemyndighedernes opgørelser af indkomstfordelingen, som blev tilgængelige i mange lande, da progressiv indkomstbeskatning blev introduceret i begyndelsen af det 20. århundrede. Det var dog typisk kun personer i den øvre del af indkomstfordelingen, som oprindeligt betalte indkomstskat og dermed indgik i statistikken. Derved er det umuligt at lave præcise skøn for bredere ulighedsmål som fx Gini-koefficienten.

Dette er dog ikke tilfældet for Danmark, hvor data dækker en stor del af indkomstfordelingen næsten fra starten af det 20. århundrede og et af de bidrag, som vi laver i artiklen, er at undersøge, hvem der har oplevet stigende indkomstandele, når indkomstandelen blandt de rigeste er faldet, og om udviklingen i den øverste indkomstindkomstandel er en god proxy for udviklingen i det samlede ulighedsniveau.

Vi finder, at indkomstandelen blandt den rigeste procent – i det mindste for Danmark – faktisk er en god proxy for udviklingen i det generelle ulighedsniveau, og at især personer under den 70. percentil i indkomstfordelingen har nydt af den faldende top indkomstandel, mens den "øvre middelklasse" (personer mellem den 70. og den 95. percentil) ikke har. Deres indkomstandel har derimod været overraskende stabil i de sidste 100 år, jf. hovedfigur 1.

11

Kapitel 2 – Arbejdsudbud og optimeringsfejl

Det andet kapitel fokuserer på effekten af beskatning - eller incitamenter mere generelt - på udbuddet af arbejdskraft. Empirisk har denne litteratur gennemgået relativt store ændringer i løbet af sidste 10-15 år som følge af den øgede tilgængelighed af store mængder af registerdata af høj kvalitet, hvilket har fremhævet en række svagheder i den måde arbejdsudbudsreaktioner blev estimeret tidligere.

Kort fortalt: for 30 år siden ville forskere opstille en kompliceret model med optimering agenter og estimere modelparametrene, så modellen matcher den observerede adfærd i datasæt indeholdende kun nogle få 100 observationer af selvrapporteret data. Til sammenligning anvender forskere i dag typisk registerdata fra skattemyndighederne, der dækker hele befolkningen, og det giver mulighed for at afprøve nogle af de forudsigelser fra arbejdsudbudsmodellerne, som ikke var muligt at teste tidligere.

En af disse forudsigelser er, at folk burde være tilbøjelige til at klumpe samme (eller bunche som det kaldes) ved punkter i skattesystemet, hvor marginalskatten springer op, fx på det indkomstniveau, hvor man begynder at betale topskat i Danmark. I standard arbejdsudbudsmodeller følger sådan sammenklumpning uundgåeligt, når folk reagerer på skatter ved at reducere deres arbejdsudbud, men som Emmanuelle Saez viser i hans artikel fra 2010, er der stort set ingen sammenklumpning at finde i de faktiske indkomstfordelinger.[2] Noget der ville indebære, at folk ikke reagerede på skatter, og som igen ville have store konsekvenser for den måde, vi laver skattepolitik.

I kølvandet af Emmanuel Saez' resultater har der været en forskningsdagsorden anført af især Raj Chetty om at forene den manglende sammenklumpning (med mere) med, at folk faktisk reagerer på skatter, og hovedforklaringen givet i denne litteratur er, at den enkeltes handlinger er underlagt optimeringsfejl - dvs. at folk af forskellige årsager ikke vælge de handlinger, der maksimerer deres nytte i henhold til de almindelige økonomiske modeller. I sine 2011 og 2012 artikler viser Raj Chetty, at små optimeringsfejl i arbejdsudbudsbeslutningerne - som fx at det tager tid og energi at søge efter et nyt job - er nok til at forklare både den manglende sammenklumpning og andre paradokser i arbejdsudbudslitteratur.

Bidraget fra min forskning til denne litteratur er at kaste lys over, hvad disse optimeringsfejl præcis dækker over, og dette gør jeg ved at anvende de danske studerendes arbejdsudbud som case. Hidtil har konkret evidens på optimeringsfejl været forholdsvis begrænsede i arbejdsudbudslitteraturen, hvilket afspejler, at identifikationen af optimeringsfejl typisk både kræver højkvalitetsdata og særlige institutionelle rammer - data

---

[2]Artiklen udkom allerede som et arbejdspapir i 1999.

af høj kvalitet for at folks optimeringsfejl ikke sammenblandes med målefejl i data og særlige institutionelle rammer, der tillader adskillelse af rationel adfærd fra optimeringsfejl.

Begge disse krav er opfyldt for de danske studerendes arbejdsudbud, og hovedfiguren i dette kapitel viser indkomstfordelingen blandt danske studerende 3 år før og 3 år efter reformen af SU systemet i 2009. Det centrale element i det danske SU system er, at de studerende modtager en udbetaling fra staten, som udfases med erhvervsindkomst over et bestemt niveau, og denne udfasning skaber et stort spring i de studerendes effektive marginalskat. Men i overensstemmelse med Emmanuel Saez' resultater, er der ingen synlig sammenklumpning i indkomstfordelingen: hverken før reformen (hvor springet sker ved 0 kr. på x-aksen i figuren) eller efter reformen (hvor springet sker ved 18.000 DKK). Fortolket inden for rammerne af standard arbejdsudbudsmodellerne skulle dette betyde, at de studerende ikke reagerer på højere skatter.

Hovedfigur 2
Indkomstfordelingen for studerende før og efter 2009 reformen
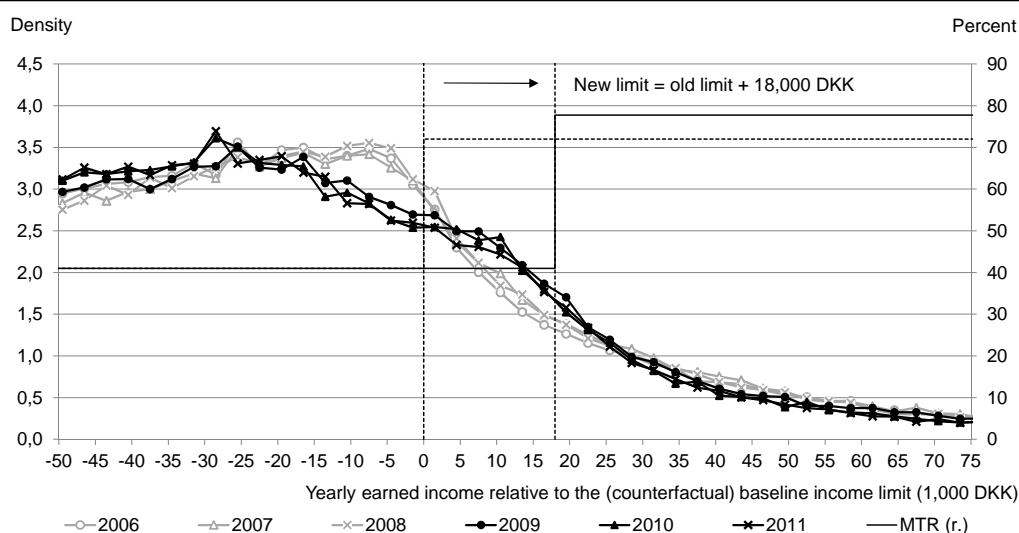
Anm: Se kapitel 2.

På trods den manglende sammenklumpning viser figuren dog også et ikke-trivielt skift i indkomstfordelingen med personer, der flytter fra et interval under det initiale punkt, hvor marginalskatten springer til et interval ovenfor. I betragtning af hvor stabil indkomstfordelingen er både før og efter reformen er dette en stræk indikation på, at de studerende reagerer på de incitamenter de står over og at der er noget i folks adfærd slører den sammenklumpning, som man skulle forvente at se.

Derefter undersøger jeg, hvad dette "noget" er og fokuserer i den forbindelse på 3 alternative forklaringer givet i litteraturen, nemlig: 1) reale justeringsomkostninger (det tager tid og kræfter at finde et nyt job), 2) gradvis læring (det tager tid at lære, hvordan skattesystemet mv. præcis fungerer) og 3) uopmærksomhed (uventede ting sker i løbet af året, som folk ikke er opmærksomme på og dermed ikke indregner i deres reoptimering).

Mens alle disse typer af optimeringsfejl sandsynligvis spiller en rolle på alle arbejdsmarkeder, finder jeg, at uopmærksomhed er den dominerende årsag til optimeringsfejl blandt studerende og at virkningerne heraf er store. En konklusion, der allerede er antydet af hovedfiguren, idet reformeffekten indtræder med det samme, mens både reale tilpasningsomkostning og gradvis læring tilsiger en mere gradvis reformeffekt.

Kapitel 3 – Lønforskelle mellem mænd og kvinder

I det tredje kapitel, som er skrevet sammen med Henrik Kleven og Camille Landais, vender jeg tilbage til spørgsmålet om ulighed, men denne gang med fokus på indkomstforskelle mellem mænd og kvinder. Det viser sig, at selvom der har været en betydelig indsnævring af indkomstforskellene mellem mænd og kvinder i mange lande i løbet af de seneste 30 år, er der stadig betydelige forskelle tilbage. I Danmark i 2011 er median indkomsten for fuldtidsbeskæftigede kvinder 15 procent lavere end medianen for mænd, og denne forskel er overraskende ens på tværs af landene.

Formålet med kapitlet er at analysere den rolle, som børn spiller i forhold til indkomstforskellene mellem mænd og kvinder, og vores vigtigste resultat fremgår af hovedfiguren nedenfor. Denne figur er konstrueret ud fra alle børnefødsler af førstegangsforældre i Danmark fra 1985 til 2001, hvor der er registerdata for begge forældre for alle 5 år forud for fødslen af deres første barn og 10 år efter. Det efterlader os med en population på omkring 350.000 fødsler eller 11.200.000 individ-år observationer.

Lønudviklingen for mænd og kvinder omkring tidpunktet for deres første barn
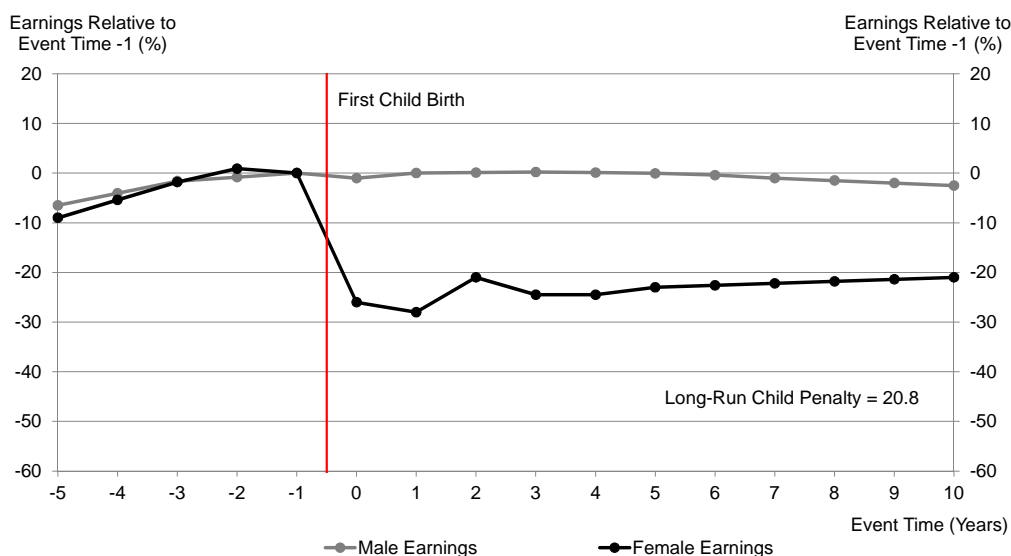
Anm: Se kapitel 3.

I figuren har vi plottet udviklingen i mænds og kvinders indkomst omkring fødslen af deres første barn, og det afslører en tydelig nedgang i kvindernes indkomst på 25-30 procent lige efter fødslen af det første barn. Dette er måske ikke overraskende, da kvinder stadig tager størstedelen af forældreorloven, men det er derimod overraskende, at der er meget lidt konvergens tilbage til før-niveauet i kvindernes indkomst. Tværtimod ligger deres indkomst stadig 20 procent under mændenes 10 år efter. I mændenes tilfælde er indkomsten stort set upåvirket af ankomsten af et barn.

Tolkes figuren, ser det ud til, at kvinder ændrer karrierespor efter ankomsten af børn, og at det giver dem en lavere indkomst over tid sammenlignet med deres mandlige partnere. Noget der også bekræftes, når vi undersøger andre mål for karriereudviklingen, såsom fx sandsynligheden for at blive en topleder i en virksomhed. I stedet finder vi, at kvinder i stigende grad vælger ind i "familievenlige" virksomheder efter ankomsten af et barn, hvilket i gennemsnit kan forventes at resultere i lavere indkomst.

Tager vi disse skønnede "child penalty" tilbage til de overordnede indkomstforskelle mellem mænd og kvinder i hele befolkningen, finder vi, at det meste af de resterende indkomstforskelle kan tilskrives den effekt, som børn har på kvindernes indkomst. Mere konkret vurderer vi, at 80 procent af de resterende indkomstforskelle mellem mænd og kvinder i 2011 kan tilskrives den forskel, der er i effekten af at få børn på mænds og kvinders indkomst. I 1980 er det tilsvarende tal 30 procent.

Der er selvfølgelig et åbent spørgsmål, hvorfor mænd og kvinder reagerer så forskelligt på det at få af børn. En forklaring kunne være, at kvinder har en komparativ fordel i at tage sig af børn, således at parrene rationelt vælger, at kvinderne fokuserer børnepasningen. En anden forklaring kunne være, at den observerede adfærd er et resultat af en norm i samfundet om, at kvinder bør tage sig af børnene. Selvom vi ikke kommer med et endeligt svar på det spørgsmål, finder vi, at der er en signifikant sammenhæng mellem den relative arbejdsmarkedstilknytning mellem moderens forældre og den "child penalty", hun oplever. Konkret finde vi at i familier, hvor mormoren arbejdede meget lidt i forhold til morfaren, oplever mødre en større "child penalty" end mødre i familier, hvor mormoren arbejdede relativt meget i forhold til morfaren. Dette kunne indikere, at kvinder har tendens til arve deres egne mødres præferencer for professionelt arbejde relativt til børnepasning.

Contents

1 The Long Run History of Income Inequality in Denmark

Joint with A. B. Atkinson (Oxford University)

1.1 Introduction[1]

The long run history of income inequality in Denmark is of considerable interest. Denmark is often portrayed as a country that has successfully combined economic performance with social safety. Certainly, in today's terms, Denmark scores well in league tables of income inequality. In the OECD 2011 report, Divided we stand, Denmark has one of the lowest Gini coefficients, and in the World Top Incomes Database (WTID), the share of the top 1 per cent is among the lowest recorded. This leads naturally to the question whether this has always been so. Or has Denmark in the past brought about a significant reduction in inequality? If so, when did it take place and how was it achieved?

With the focus on long run income inequality the present chapter contributes to the recent litera-ture on top income shares that has emerged since the studies by Piketty (2001,2003), Piketty and Saez (2003) and Atkinson (2005). These studies have been influential in highlighting the dramatic changes in income inequality across many different countries since the beginning of the 20th century. Studies of this type have already been conducted on Sweden (Roine and Waldenström, 2010), Norway (Aaberge and Atkinson, 2010) and Finland (Jäntti et al., 2010) and the present chapter thus completes the set of the Scandinavian countries.

At the same time the focus on top income shares stems from the fact that historical sources on the income distribution in most countries primarily covered the top of the

distribution, hence making the estimation of broader measures of inequality such as the Gini coefficient infeasible. In contrast, the data sources for Denmark extend well down into the income distribution from almost the beginning of the 20th century and we can – quite unusually – construct a series covering nearly a hundred years. This allows us to address two further questions. How well do top income shares work as proxies for the development in the overall income distribution? What other parts of the income distribution mirror the changes at the top?

The study on Denmark is aided by the fact that the income tax data provide a rich historical source, which can be coupled with micro data covering the entire universe of taxpayers in Denmark available from 1980. There has long been research based on these tax records. The 1928 textbook, Den økonomiske fordeling, by Zeuthen contained analyses of the distribution in the 1920s. Bjerke (1957, 1965) examined the period 1939-64, while later studies included Egmose (1985) covering 1939-80, Pedersen and Smith (2000) covering 1981-96 and the series for top income shares constructed from micro data for 1980 to 2005 by Kleven and Schultz (2014). Finally a long run perspective is taken by Sørensen (1989, 1993), whose study covers 1903-1986. This chapter benefits from these earlier investigations, and seeks to join up the findings for the different sub-periods, while paying strict attention to the comparability of the data over time.

In this respect, the concept of "assessed income" used in many of these earlier Danish studies represents an obstacle (in effect, it deducts taxes paid in the previous year), and compared to these earlier studies a major contribution of the present chapter is to make estimates of the distribution of taxable income – a concept similar to that employed in other countries. In doing so, we obtain internationally comparable estimates of income inequality in Denmark from 1870 to 2010 – 140 years spanning two world wars, and the Great Depression as well as the recent financial crisis. At the outset we should emphasise that, as normal when considering time series from more than a century, data have limitations. The income tax data used here arise from an administrative process and reflect both the tax legislation and the reactions of taxpayers in the forms of avoidance and evasion. They should be seen as providing an imperfect source of evidence about a long period for which no alternative source (such as household surveys) exists.

With this in mind our study shows firstly that there have been epochs when Denmark has seen significant reductions in income inequality: (possibly) in the last 30 years of the 19th century spanning the start of the industrialization in Denmark, and definitely over the Second World War and in the 1970s. As a consequence the decline is not simply a secular downward trend. Instead there have been periods in between where the top income shares have remained stable and, most notably, spiked during the First World

War. This time path resembles to a large extent of that of Sweden, which adds to the picture of strong co-movement in income inequality across countries until the 1970s and increasing divergence since then.

Secondly, these patterns are present both when we consider the top income shares and when we examine the Gini coefficient, indicating that the top income shares indeed are a good proxy for the underlying development in income inequality. Likewise our study shows that the reduction in the top income shares over the past century in particular benefitted the part of the population with incomes below the 70th percentile, while in contrast the income shares of the population between the 70th and 95th percentile have remained remarkably stable. This has implications for the possible explanations behind the development.

1.2 Methodology and data

In using income tax data we follow a line of research that began in the United Kingdom in the 19th century (e.g. Baxter, 1868), was taken up in the United States when the present income tax was established, developed further by Kuznets (1953), and which has recently been revived in a series of studies following Piketty (2001, 2003). These later studies combine administrative income tax data with external sources for total income and total population (often referred to as control totals) that allow the tax data to be placed in context (see Atkinson, 2007a). This implies that our estimate of e.g. the income share of the top 1 per cent refers to the richest per cent of the total population and their share of total income.

The great advantage of income tax data is that they provide a long run time series – in the present case more than a century, but at the same time, it represents a challenge.[2] Firstly, because the early income taxation schemes only levied taxes on the part of the population with relatively high incomes, parts of the income distribution were not included in the data and control totals for both income and the population therefore have to be established on a consistent basis. Secondly, tax systems are over time subject to changes that affect the consistency of the statistics. However in tackling these challenges, the construction of the time series for Denmark benefits from a number of advantages.

Firstly there was a stable tax code over a long time span, with the tax code from the establish-ment of a permanent national income tax system in 1903 remaining the foun-

---

[2]More detailed information about the data and methods is provided in the working paper version of the chapter, Atkinson and Søgaard (2013).

dation of the income tax system until the end of the 1960s.[3] And secondly the statistics that Statistics Denmark (DS) have collected since the beginning of the income taxation are very detailed and cover a large part of the population compared to that from other countries. This reduces the extent to which we have to rely on external income totals and the uncertainty associated with the calculation of the inequality measures.

At the same time, there are features of the Danish income tax system that differ from those in other countries and which have to be carefully treated. In what follows, we describe these features and the principal respects in which the tax system has changed over the period.

Definition of income

The income concepts applied in determining tax liability, and the income concepts reported in the statistics resemble – with one major exception – what is used in other studies of long run income inequality. The foundation of the income tax system was a comprehensive income concept, in which – in principle – all income streams were added along with deductions of all costs associated with "acquiring, securing and maintaining" the income. This meant that imputed rents, positive and negative interest payments, and income in kind all were included, while gifts, inheritances, lottery winnings etc. were exempt.

The income concept includes public transfers such as unemployment benefits, sickness benefits and public pensions as they were all taxable.[4] However before 1994 some other transfers (cash-benefits and supplement provisions for pensioners) were exempt. We deal with this data break by assuming that the inclusion of these transfers (including the increase in benefit rates needed to compensate for the new tax liability) only affects the income total (not the top income brackets) and add back an estimate of the size of the exemptions to years before 1994.[5]

Until 1970, tax liability in year T was based on income accrued in year T-1. In 1970 this changed to a current year basis, which meant that 1969 was a tax free year, and no estimates are given for that year. In all cases the year refers to the year of receipt not the

---

[3]Since then, a number of tax reforms have been passed, notably the change from family to individual taxation in 1970 that is discussed below.

[4]Even though most transfers were included in taxable income, individuals who had only these as income would historically not pay taxes due to high personal allowances.

[5]We do this by first regressing the growth rates of total taxable income on total labour and capital income (excluding 1994) and scale up the 1993 income total, so that the growth rate from 1993 to 1994 equals the predicted value (3.8 per cent) instead of the actual (7.9 per cent). The implied increase in the 1993 income total of 4.0 per cent is indexed to the development in the income transfers relative to GDP and applied to all years before 1993, so that e.g. total income in 1950 is increased by 1.0 percent.

year of assessment.

Treatment of capital income

There have been two significant exceptions from a comprehensive definition of income. The first is that capital gains were only included if they were accrued on intent, i.e. if they were accrued in relation to a taxpayer's livelihood or speculation. In practice it was often difficult to determine whether an asset was bought with the intent of speculation, and so from 1922 it became legal practice to presume speculation if an asset was bought and sold within 2 years. It therefore seems reasonable to assume that capital gains for ordinary tax payers only rarely entered our income definition.

After 1960, the treatment and placement of capital gains in the tax system were changed a number of times, but the changes in general kept capital gains not related to a tax payer's livelihood outside the income concept used in the statistics and as a consequence we interpret the income series as excluding capital gains. Inspection of the Danish National Financial Accounts from 1995 indicates year-to-year variation, but no trend in the amount of capital gains accruing to the household sector.

The second exception is that, from 1991, dividends were taxed under a separate scheme and not included in taxable income. However this only had a minor impact on the income distribution as we show in section 1.3.

"Assessed income"

Where the Danish data deviate from those in most other countries – the major exception re-ferred to earlier – is that, until 1966, the tax was levied on so-called "assessed income", which was given by taxable income as defined above minus all paid personal taxes, cf. table 1.1. This procedure did not involve circularity, since the tax paid in year T was based on income and deductions in year T-1. The deductions included all personal taxes paid to state, municipalities and the church, so that in effect the tax base in year T was equal to the net-of-tax income in year T-1 (see Bjerke, 1957, p. 99).

Table 1.1

Overview of the income concepts available in the tax statistics

Period	Income concept	Definition
1903-66	Assessed income	= Taxable income – paid personal taxes[a]
1967-	Taxable income	= Gross income – deductions (interest payments, etc.)
1976-	Gross income	= Wage income[b] + interest received + net business income + transfers + dividends[c] + imputed rents[d]
1980-	Micro data allowing a variety of definitions in addition to the above.	

Notes: a) Relating to assessed income in the previous year. b) Including compensation in the form of stock options etc. c) Until 1990. d) Until 1999.

Sources: The first three rows correspond to Sørensen (1989, p. 63).

As such the assessed income definition found in the Danish statistics is close to the post-tax post-transfer income measure urged by Burkhauser and Larrimore (2014), and we show below the results obtained on this basis. However, the deductibility of paid personal taxes up to 1966 stands in the way of making a comparison with the estimates of top income shares in other countries, where pre-tax incomes are employed. Correcting tabulated data for this deductibility is not trivial, as the individual size of the deduction depends on the assessed income the year before. Presumably this is why earlier studies have simply used the "raw" data. In this study we have treated it by adding back an estimate of paid personal taxes for each reported income bracket in the statistics, where the estimate are based on the bracket mean income and information on the tax schedule in each year. We thereby create an income concept closer to taxable income used after 1966. To our knowledge, this is the first time such an adjustment has been made.

The calculation is only approximate for a number of reasons. It is based on income in year T ra-ther than year T-1. It excludes municipal taxes, church taxes and certain other taxes such as the state wealth taxation; this exclusion affects the income share to the extent that they were not levied proportionally.[6] Using the interval average as the tax base is likely to underestimate the average tax payment in a progressive tax system, and hence to understate the degree of inequality. However, operating in the opposite direction is the fact that the groups at the top of the distribution in year T can be ex-

---

[6]Municipalities had especially in the beginning of the 20th century a large degree of autonomy in specifying their own tax system, resulting in progressive taxation in some and probably effectively regressive taxation in others. It is therefore difficult to say anything about the aggregate level of progressivity in municipal taxation.

pected to have had a relatively higher income growth between T–1 and T compared to other income groups in the same year, thus giving them a relative higher income compared to the year before in which the actual tax payments were calculated. Not being able to control for income mobility therefore presumably overestimates the size of the taxes paid at the top of the distribution and underestimates them at the bottom, giving a more unequal income distribution. Leaving out the wealth tax counters this to the extent that income and wealth are correlated.

The resulting totals for taxable income are shown in figure 1.1 below together with the totals for assessed income and for reported income (excluding the income of those not included in the income tax statistics). In each case, the totals are expressed as a percentage of Gross Factor Income taken from the national accounts. The switch from assessed income to taxable income between 1966 and 1967 creates a large jump in the total reported income, but, despite the limitations of our estimates of the paid personal taxes, the implemented correction is more or less able to remove the jump in the total income, which indicates that the net effect of our correction is broadly correct.[7] The remaining gap is attributed to the deduction of municipal, wealth and other minor taxes, which – taken together – are assumed to be proportional and thus to not affect the calculated inequality measures.

The totals relate to recorded household income and fall short of Gross Factor Income (100 per cent in figure 1.1) for two major reasons.[8] The first reason is that factor income accrues to other sectors of the economy (such as the corporate sector) and is not fully passed on to households (in the form, for example, of company dividends or interest payments). The second reason is that the valuation of income (for example, imputed rent) may differ in the national accounts. From figure 1.1, it may be seen that gross income was broadly stable as a percentage of Gross Factor Income from 1920 to the end of the 1960s. It increased in the 1970s and then fell by some 10 percentage points. A similar fall has been observed in other countries such as the United Kingdom (Atkinson, 2007b).

---

[7]Unfortunately there are no years in which DS reports the income distribution both with and without the deductibility of paid personal taxes. However, also considering our series for the top income shares our correction appears to be correct, as we return to below.

[8]Working in the opposite direction is the inclusion in household income of transfers and payments of interest by the gov-ernment.

Figure 1.1

Income totals as share of gross factor income
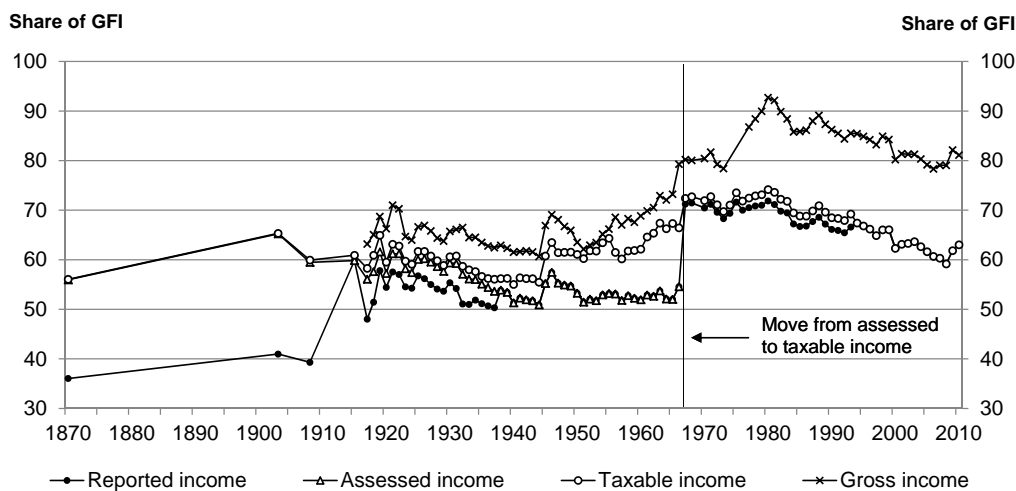
Notes: The income concepts refer to the following: a) Reported income: The income total of the legal tax base from the DS tabulated income statistics and the micro data from 1980. b) Assessed income: Reported income plus the DS estimates of the income of those who are not included in the income tax statistics. From 1917-1937 the excluded income has been estimated by Sørensen (1989). b) Taxable income: Before 1966, assessed income plus our own estimates of the deductions for ordinary state income taxes (with effect from 1908). Before 1994 the taxable income series have further been adjusted for the grossing up of income transfers in 1994 as described in the text. c) Gross income: Before 1970 gross income is given by the contemporary estimates from DS including the income of individuals not included in the income tax statistics and adding back all deductions. After 1970 it is given by the legal gross income, which is collected automatically by the tax authorities. From 1980 the totals are taken from the micro data; the years 1980-82 overlap with tabulated totals. All income totals are expressed as a percentage of Gross Factor Income as defined in the national accounts given by Hansen (1974) (1870-1936) and DS (1937-2010).

Sources: Statistics Denmark (DS), Bjerke and Ussing (1957), Hansen (1974), Sørensen (1989) and own calculations.

With the above adjustments we obtain a comparable series on taxable income covering the whole century, which we take as our main series. They differ from estimates using gross income in that items such as interest payments, contributions to unions and unemployment insurance, travel to work costs, etc. are deducted. As may be seen from figure 1.1, there is a widening gap in recent decades between total gross income and total taxable income. However comparing our main series with estimates based

on gross income for the years 1977-2010, where we have overlapping data, we find a very parallel development, as discussed below. This gives us some confidence that the development measured by our series using taxable income is also historically a good proxy for the underlying development in gross income.

Finally it should be kept in mind that the income used here is reported for tax purposes, which implies that tax evasion is a potential problem. The presence of tax evasion of course gives rise to some caution in terms of interpreting the observed income distribution as the real distribution of economic resources, but it only constitutes a problem for our measures of inequality and of the top income shares, if the evasion is disproportional to reported income. As discussed in Atkinson and Søgaard (2013), there is some evidence that evasion is indeed relatively proportional to reported income.

Definition of population

Until 1969 the tax unit was the family with the incomes of husbands and wives being added to-gether, and the required control total for the population for this period is therefore the total number of individuals aged 15+ minus the number of married women. Both of these numbers are available from the population censuses. In 1970 this changed to an individual based tax system so that relevant control total became individuals aged 15+.[9] From this point on we use the actual number of tax units as the population total, which corresponds closely to the population of age 15+. We do not attempt to bridge this break in the series, but we discuss the implications of using different population totals and the change between them in section IV.

Data sources

The sources for the period up to 1979 are tabulated data.[10] There are a number of gaps, but the series is particularly rich for the first part of the twentieth century: e.g. between 1903 and 1939 there are 26 observations. The Danish data are less strong for the 19th century, having only the one observation for 1870.[11] In what follows, we make use of the data for 1870, but it should be borne in mind that the long gap – a third of a century

---

[9]Some individuals below 15 years also filed a tax return if they earned a sufficiently high income.

[10]As the tabulated income statistics in general do not correspond to the income percentiles of interest, it is necessary to derive an estimate of the income distribution within each interval in order to get the desired percentile cut-off. We use the split histogram as the interpolation method (see Atkinson, 2005). No extrapolation is made to obtain income shares in an open upper interval. The tabulated data do not include any intervals with negative incomes. Assessed income is assumed to be non-negative.
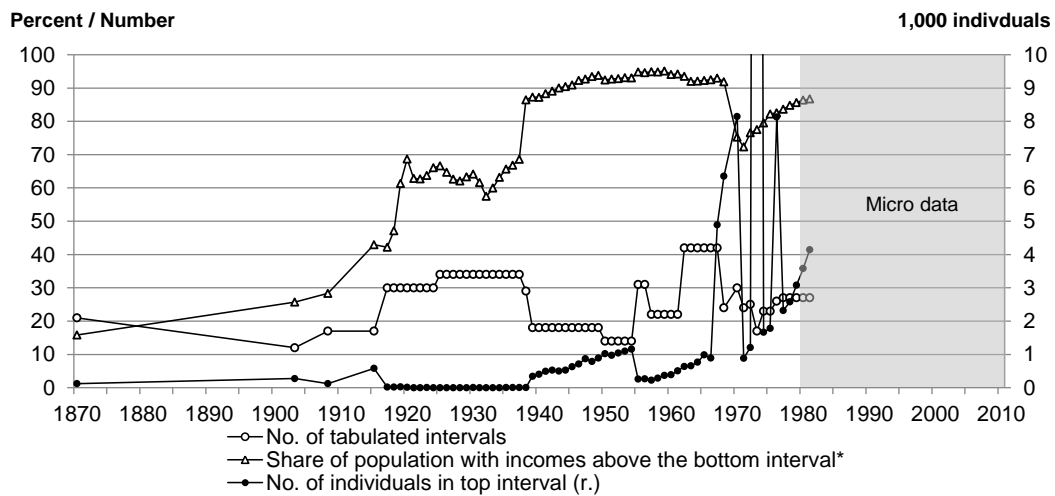
[11]As a reference the corresponding number of observations between 1903 and 1939 is much smaller for Sweden (10) and Norway (6), although for the 19th century Norway has 10 observations.

– means that the figures may be less comparable even though the principles in the tax code were the same.

A further strength of the tabulated data for Denmark is that the data from an early point in time covered a substantial part of the population, which is in contrast to many other countries, where the historical income tax data typically are limited to the very top of the income distribution. Prior to 1938, DS collected income assessments for families with an assessed income above 800 DKK, which meant that 26 per cent of the population was included in 1903. As is shown in figure 1.2 this number had increased to around 66 per cent by 1917 and remained around this level until 1937. From 1938 the statistics are assumed by DS to cover all potential tax payers and the overall coverage is therefore in principle 100 per cent from this point. Of these, a substantial proportion (86 per cent in 1938) had incomes above the bottom interval. The wide coverage together with the relatively high number of tabulated intervals (shown in figure 1.2) and close to micro data at the top of the income distribution (the numbers of taxpayers in the top interval are shown in figure 1.2) imply that from this point we can derive tighter bounds on both top income shares as well as broader inequality measures and that we have to rely less on external controls for total income.

Figure 1.2

Assessing the quality of the data

Percent / Number   **1,000 indivduals**

—o— No. of tabulated intervals
—△— Share of population with incomes above the bottom interval*
—●— No. of individuals in top interval (r.)

Notes: * For the years before 1938 the "residual" interval, with those who had an income below the
threshold needed to be included in the income tax statistics, is treated as the bottom interval and
counted in the number of intervals. For the years 1921-1938 the number of individuals in the top
interval is between 0-3. This is what we mean by saying that the data are "close to micro data". In
1973 the number of individuals in the top interval was 75,049.

Sources: Statistics Denmark and Sørensen (1989).

However despite of the wide coverage of the income statistics we still need estimates
of the in-come of the families not included in the tabulated data for the period before
1938 and for this we rely on the contemporary estimates made by DS until 1915 and
the estimates made by Sørensen (1989) for the years 1917-37. These estimates typically
imply that an amount of around 400 DKK is added per excluded family, i.e. half of
the income threshold required to be included in the statistics. In figure 1.1, the result-
ing estimated income of the families with an income below 800 DKK is given by the
difference between reported and assessed income. It may be seen that the addition is
substantial in the years before 1915.

From 1980 and onwards the income records of the entire universe of Danish tax pay-
ers are available as micro data. From this point we can therefore calculate income
inequality using a variety of measures and income definitions, as well as at both an
individual and a family level.

1.3 The development over time of the Danish income distribution

We now turn to the long run development in income inequality starting in figure 1.3 and 1.4 with the income share of the top income percentiles (P90-P100) as well as the income shares of the percentiles further down the income distribution. For the top income percentiles the quality of the Danish data implies that we can calculate the income shares (figure 1.3) throughout the entire period from 1870 to 2010, while this is generally not the case before 1915 for the income shares of the lower percentiles (figure 1.4).

Figure 1.3

The development in income shares in Denmark 1870-2010: Top 10 percentiles
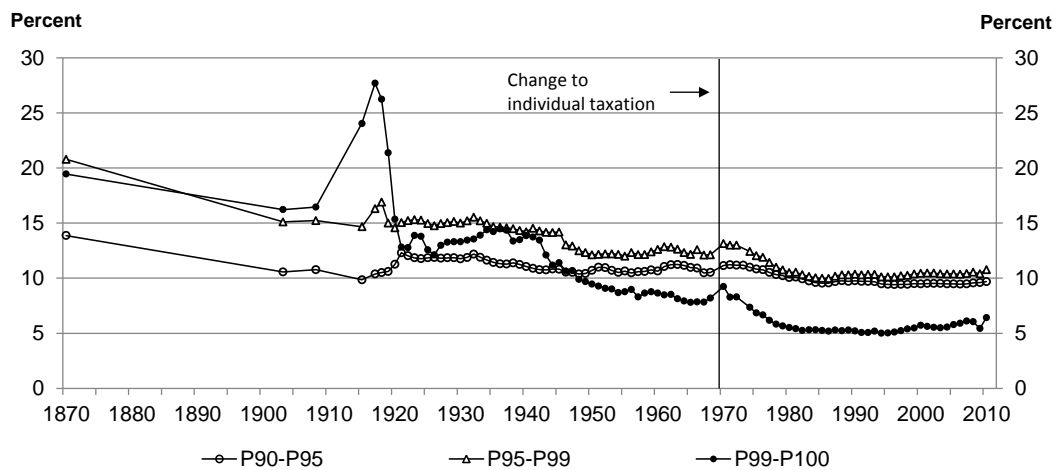
Notes: The income shares have been calculated using the definition of taxable income (gross income minus deductions), which we have adjusted for the grossing up of transfers in 1994 (with effect in the years prior to 1994) and the deductibility of paid personal taxes (with effect in the years 1908-1966). The vertical line in 1970 indicates the change from family to individual taxation.

Sources: Own calculations.

Figure 1.4

The development in income shares in Denmark 1870-2010: Bottom 90 percentiles
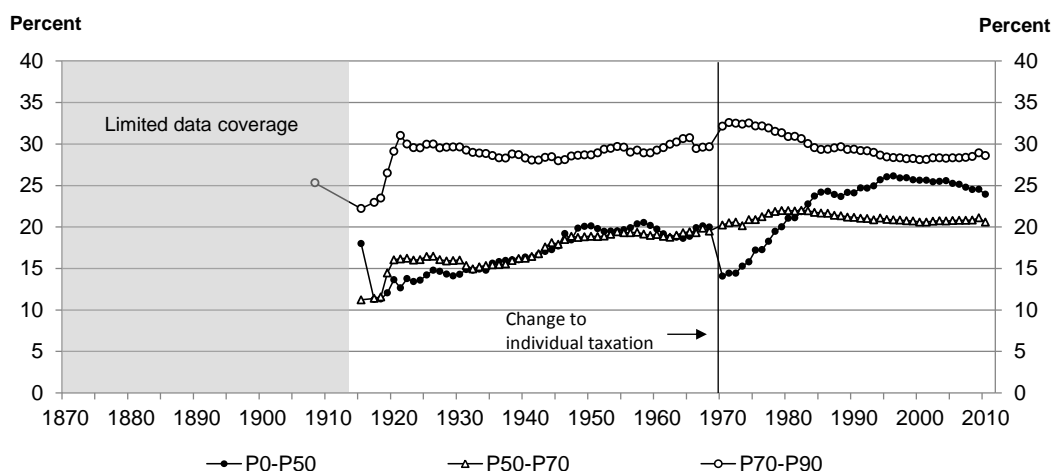
Notes: See figure 1.3.

Sources: Own calculations.

Looking at the top income shares, we see a substantial decline between 1870 and 1903 of 12 percentage points distributed more or less equally among the 3 subgroups (the P90-P95, P95-P99 and the P99-P100). The facts that the 1870 figures were the result of a one-off tax and that we have no evidence about the intervening years, mean that the fall must be interpreted with caution. However the indication that income was much more unequally distributed before the 20th century is supported by Soltow (1979). He uses data from another one-off tax in 1789 to analyse the distribution of both income and wealth and arrive at a Gini coefficient at around 90 per cent. Applying the methodology used here to his numbers gives a lower bound on the top 1 per cent income share of around 30 per cent compared to our estimate of 19.4 per cent in 1870.

The indications of high inequality in 1789 and 1870 are interesting because they pre-date the 1890s that most historians set as the start of the industrialization in Denmark. The data thus speak against a standard Kuznets (1955) type of explanation for the development in inequality, where inequality follows an inverse-U shape: first increasing as only a few workers initially move to high productivity/wage sectors in the beginning of the industrialization and then decreasing as more and more do so.

From 1903 – where we have a nearly continuous series – we first see a dramatic rise and fall in top shares during the First World War, and then a decline that took place particularly in the 1940s and in the 1970s. Taken together, this means that the share of the top 1 per cent has fallen from around 16 per cent at the beginning of the twentieth

century to around 6 per cent at the end of the century. Although the top income shares in the recent years have remained relatively low by historical standards, there has been a tendency to rising inequality at the very top. In 2010 the share of the top 1 per cent was 6.4 per cent, which was its highest level over the past 30 years, although only 1 percentage point higher than in 1980.

From figure 1.3 it is clear that most of the changes in the income shares of the top income groups have happened in the top 1 per cent, while the P90-P95 income share has been remarkably stable at around 10 per cent since the beginning of the 20th century. This pattern has been found in many other countries (see e.g. Piketty (2003) for France, Piketty and Saez (2003) for the USA and Roine and Waldenström (2010) for Sweden), and naturally leads to the question of which income groups were affected by the changes in the top income shares, i.e. as the income share of the top 1 per cent fell from 16.2 in 1903 to 5.5 in 1980, where did this income mass go?

The answer to this question can be seen in figure 1.4, where we show the income shares of the bottom half of the population (P0-P50), as well as for the P50-P70 and the P70-P90. From this figure it is clear that the stability of the income shares of the P90-P95 can also be found for the P70-P90. In contrast the income shares of the P0-P50 and the P50-P70 grew almost in line from 11-12 per cent in 1917 to 19-20 per cent in 1968. At least for the case of Denmark the decrease in the top income shares since the beginning of the 20th therefore seems to have benefitted the income groups from the 70th percentile and below, but not the income groups between the 70th and the 95th percentile.

Top income shares and broader measures of inequality

The above description of the development in income shares of different groups throughout the income distribution raises the question as to how far changes in the top income shares can serve as an indicator of the change in overall inequality, as measured for example by the Gini coefficient.

In considering measures of overall inequality, we have to take account of the incomplete cover-age of the income tax data. This may be summarized in terms of three variables: (a) the proportion of the population, F, for whom we have effective income data (those above the income threshold for inclusion in the tabulation, shown in figure 1.2 earlier), (b) the share, Ω, of income attributable to this group, and (c) the starting value of income, y, for those covered, expressed as a fraction of the mean income. In 1903 the excluded group were essentially those with less than mean income, which means that any measure of overall inequality is likely to be surrounded by considerable un-

certainty. We can however calculate bounds on this uncertainty.

A lower bound for the Gini coefficient is given by assuming that all individuals in the excluded group receive the same income, $(1-\Omega)/(1-F)$, expressed relative to the mean, and an upper bound is obtained by assuming that they were divided between two groups: one receiving zero and the other receiving the maximum, y.[12] The difference between the two assumptions provides a measure of the maximum possible margin of error. For 1903, this is quite large – around 14.1 percentage points – but from 1915 the difference is generally below 2 percentage points. Using the same technique to calculate the upper and lower bounds for the contribution to the Gini coefficient from each tabulated interval, one can calculate the overall bounds on the Gini coefficient. In 1903 the difference between the two bounds is 14.2 percentage points, indicating that the bulk of the uncertainty comes from the excluded bottom group.

Figure 1.5 shows the Gini coefficients (upper and lower bounds) for the entire period covered. The bounds are virtually indistinguishable from 1915 onwards. In considering these figures, it is important to bear in mind the change in 1970 from a family to an individual basis and the inclusion of certain transfers in 1994 (marked by the vertical lines in figure 1.5), but allowing for the breaks, it is clear that the Gini coefficient has fallen substantially in Denmark over the past century. It also appears that, as with the top shares, the fall has been episodic rather than a constant downward trend.

---

[12]The bounds can be visualized using the Lorenz curve. If P denotes the first point for the included population, then the lower bound is found by joining the origin to P. The upper bound is obtained from the Lorenz curve that follows the horizontal axis to F*, and then the line joining F* to P, where F* is chosen so that there is the same mean income for the excluded population.

Figure 1.5

The Gini coefficient and the top 1 per cent income share
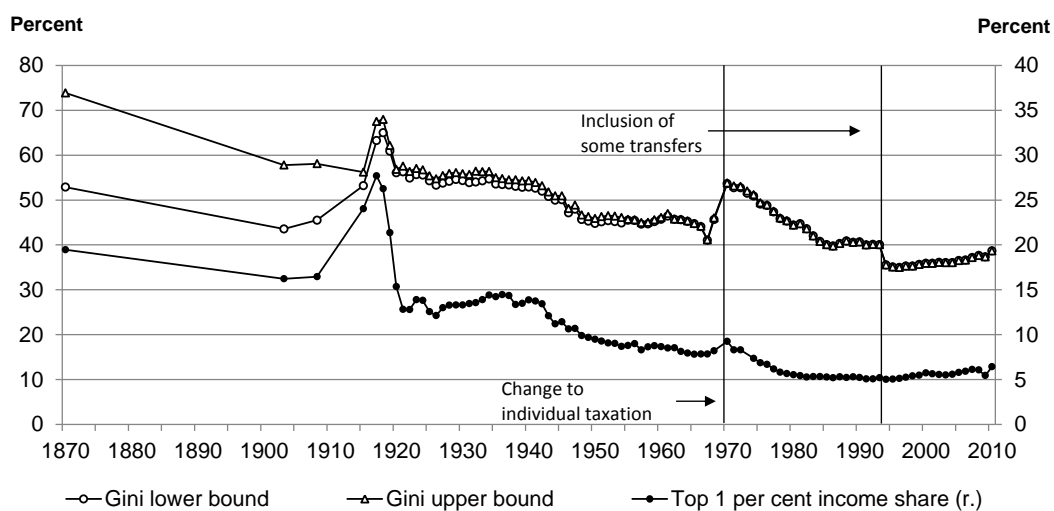
Notes: The first vertical line indicates the change from family to individual taxation in 1970, the second the inclusion of certain transfers in 1994. The top income series have been adjusted for the later data break by assuming that the grossing up only affected the income total as described above. Something similar is not possible for the Gini co-efficient.

Sources: Own calculations.

From figure 1.5 it is evident that there qualitatively is a strong co-movement between the Gini coefficient and the top 1 per cent income share. The next question is whether this is also the case quantitatively in the sense that there is a stable relationship between changes in the top 1 per cent income share and the Gini coefficient. If this is the case, the magnitude of changes in the top income share is also informative about the magnitude of changes in the Gini coefficient.

In order to examine this question formally we run a basic co-integration analysis on the two variables over the time period 1915 to 2010, where we have relatively precise estimates of the Gini coefficient. From this analysis we learn that there is a stable co-integration relationship in most of the sample period with the exception of the years around the First World War (1915-1920) and the period after the change to individual taxation, which was dominated by entry of women into the labour market (1970-1983). Including dummies for these years in the co-integration relationship along with a dummy for the data break in 1994 yields the long run relationship presented in table 1.2.

Table 1.2

Estimated reduced form long run relationship between the Gini coefficient and the top 1 per cent income share

	Gini lower bound		Gini upper bound	
	Estimate	**Standard error**	**Estimate**	**Standard error**
Gini	-1.000		-1.000	
Top income share	1.877	0.176	2.096	0.168
Constant	27.470	1.717	26.157	1.638
No. of observations	94			
Rank of long-run matrix	1			

Notes: The table shows the reduced-form beta estimates from a co-integrated VAR of rank 1 with the Gini coefficient and the top 1 per cent income share as endogenous variables – both of which are included with two lags. The model further includes unrestricted year dummies for the years 1917-1920, 1969-1983 and 1994. Missing observations in 1916, 1969 and 1973 are imputed by linear interpolation.

Sources: Own calculations.

The interpretation of this table is that the equation:

$$Gini_t = \beta_0 + \beta_1 TopIncomeShare_t \qquad (1.1)$$

constitutes a stable long run relationship in the sense that deviations from this relationship tends to be closed over time, and on the basis of the estimated coefficients, a 1 percentage point fall in the share of the top 1 per cent is associated with a fall in the Gini coefficient of 1.877-2.096 percentage points. This relationship is as mentioned stable throughout most of the sample period, except under extreme increases in the top 1 per cent income share as experienced during the First World War (year dummies for 1917-1920) or under the major change in the labour market structure represented by the entry of women into the labour market (year dummies for 1969-1983).

It is further interesting to note that the regression coefficients exceed the purely arithmetic con-tribution of changes in the top share, which can be approximated by 1-G* ≤ 1, where G* is the Gini coefficient for the bottom 99 per cent of the population (Alvaredo, 2011). This means that there must be co-movement between the top income share and inequality among the rest of the population. However this comes as no surprise, as we saw above that the decrease in the top 1 per cent income share primarily benefitted the income groups below the 70th percentile and therefore worked to reduce G*. If the contribution of the changes in the top share should have equalled the purely arithmetic contribution, the changes should have been distributed to the income groups

proportionally to their initial income.

Implications of different definitions of income and population unit

In section 1.2 we identified three different definitions of income – gross income, taxable income and assessed income – and distinguished between family and individual income. With the micro data from 1980 we can examine the robustness of the findings described above to the differences in defini-tion, allowing us to address some of the issues raised by Burkhauser and Larrimore (2014). The first is that the series relates to taxable income, defined as gross income minus the deductions allowed in the tax code (such as interest payments). As we have seen, the overall importance of deductions has in-creased in recent years as the difference between total gross income and total taxable income has widened. Figure 1.6 shows the share of the top 1 per cent calculated on a gross income basis (where we also use tabulated data from 1977-1979). The gross share is initially higher, but the gap narrows over time and by 2000 there is little difference. As a result, the rate of decline in the top share is rather more marked for gross income.

In figure 1.6 we also show the effects of the omission of dividend income from 1991 onwards. It is not possible to add back this income source to our taxable income concept as dividends and realised capital gains are mixed together in the tax records after 1990, but removing dividends from taxable income before 1990 yields a drop of only around 0.1 percentage points in the income share of the top 1 per cent. This reflects the fact the level of dividends recorded by the tax authorities were close to zero in the beginning of the period, and as a result the removal of dividends from taxable income in 1991 did not create a visible break in the time series.[13]

---

[13]Including both dividends and capital gains in our income definition gives an income share of the top 1 per cent of 8.2 per cent in 2010. However we do not make this adjustment to our main series, as the total income of this type reported to DS increases from almost nothing in the beginning of the 1980s to around 2 per cent of taxable income from 2000 and onwards. This implies that, while the level of inequality using income including dividends and capital gains presumably is accurately measured at the end of the period, the increase over the period is likely to be exaggerated.

Figure 1.6

The top 1 per cent income share under different definition: Gross versus taxable income and the effect of family-basis
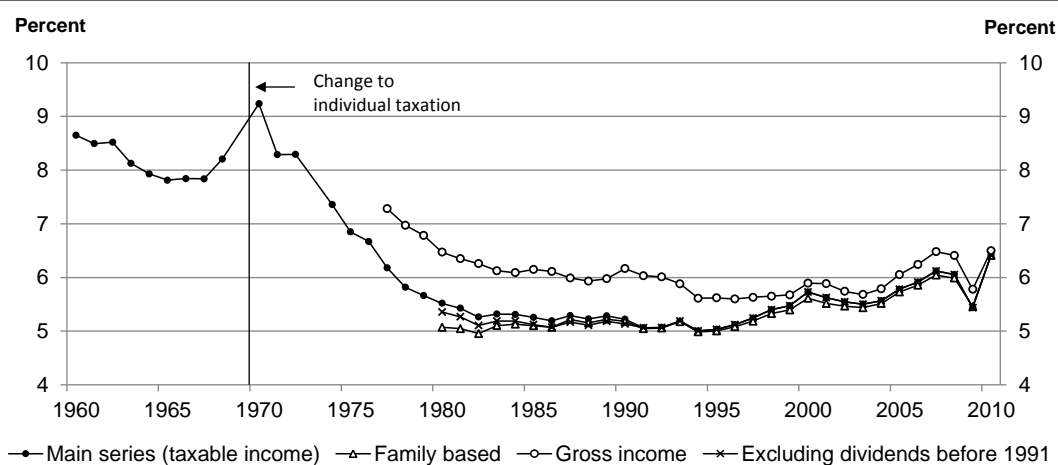
Notes: The income share on taxable income have been calculated using the definition of taxable income (gross income minus deductions), which we have adjusted for the grossing up of transfers in 1994 (with effect in the years prior to 1994) and the deductibility of paid personal taxes (with effect in the years 1908-1966). The vertical line in 1970 indicates the change from family to individual taxation.

Sources: Own calculations.

In figure 1.7, we compare the series for taxable income with that for assessed income, where per-sonal taxes have been deducted, allowing us to make a long-run comparison of these two measures.[14] First, it may be seen that the correction we have made effectively removes the jumps in the series between 1966 and 1967, which indicate that our correction method works well.

Second, from figure 1.7 it is visible that the effect of taxation on the measured income shares was marginal before 1935 and builds up thereafter, where already from 1940 the effect on the top 1 per cent is stable at around 2 percentage points, while the effect on the top 10 per cent continues to increase to 4-5 percentage points. This is in line with evidence from other studies, see e.g. Egmose (1985, p. 53). In the figure we have also included estimates of the top income shares using assessed income constructed from

---

[14]The same is indicated when we use the micro data for 1980-2010 to replicate our correction. More concretely we take the taxable income of each individual in the micro data in a given year and subtract the personal paid taxes calculated from last year's income and add the taxes calculated from the current year's income. Use this income definition in the calculation of the income shares, we obtain series that follows the same overall development as in the original series.

the micro data. From these estimates we see that the difference between the income shares using the two income definitions are approximately the same in 1980 as in 1966. Since then the difference has narrowed so that the series for the top 10 per cent income share was 2.5 percentage points lower in 2010 using assessed income while the series for the top 1 per cent was 1.1 percentage points lower. This implies that the overall level of progressivity in the tax system has been reduced over the last 30 years.

Figure 1.7

The top income shares using taxable and assessed income
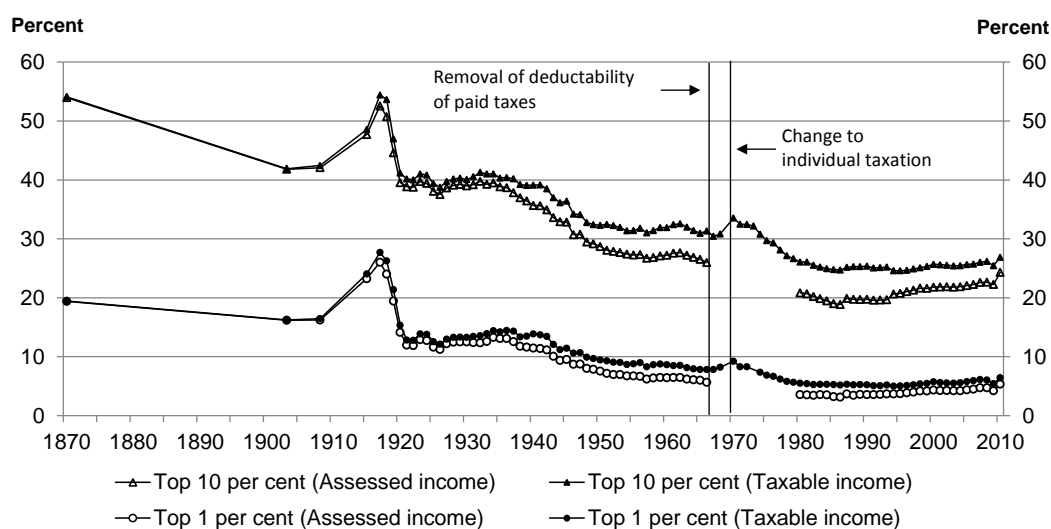
Notes: The pre-1967 taxable income is given by assessed income plus the estimated deduction for ordinary state income taxes (with effect from 1908). The vertical lines indicate the two data breaks of the removal of the deductibility of paid personal taxes in 1967 and the change from family to individual taxation in 1970. Both series have been adjusted for the grossing up of transfers. Assessed income from 1980 and onwards has been constructed from the micro data as taxable income minus paid personal taxes.

Sources: Own calculations.

Finally, the difference between the individual and the family as the population unit may be seen from figure 1.6. The change from family to individual taxation in 1970 resulted in a jump in the top income share (explored further below), but interestingly the results on a family basis from 1980 are very close to those on an individual basis when considering the top 1 per cent income share. This implies that the decline in the top income share between 1970 and 1980 is not just due to the changeover to individual

taxation, but would also have been present without it.[15]

The Danish development of top income shares in international perspective

We began the chapter by citing the good standing of Denmark in contemporary inequality league tables and asking how this had been reached. Before moving into these explanations, it is therefore interesting to compare the long-term development of top income shares in Denmark with that of other countries, both Scandinavian neighbours (Norway, Sweden and Finland) with relatively similar economic/fiscal systems and two countries (France and the US) where the differences are greater.[16] In making such a comparison, it is important to bear in mind the differences across countries in fiscal systems which compound those that arise from changes over time within countries, and of course modern day comparisons of inequality could be done using the official measures reported routinely by e.g. the OECD.

Having this in mind, figure 1.8 nonetheless shows a striking similarity between Denmark, Sweden, Norway, France and USA up to around 1970. In 1965, for example, the top 1 per cent share was, in round numbers, 8 per cent in the US as well as Denmark, 6 per cent in Norway and Sweden and 10 per cent in France. Finland is somewhat different with inequality increasing in the first two decades after the Second World War.

After 1970 there was a divergence. While the top 1 per cent income shares in USA and France remained constant between 1970 and 1980, Denmark alongside Sweden saw its top income share fall additionally and this has only been moderately reversed in the succeeding years. As a consequence Denmark in 2010 had a top income share of taxable income of 6.4 per cent, which is low by historical standards and was only 1 percentage point higher than in 1980. In contrast, the 1 per cent income share in the United States rose from 8.2 to 17.4 per cent over the same period.[17]

---

[15]The same conclusion emerges if we consider the top 0.1 per cent, but not for the top 5 per cent or top 10 per cent shares or broader measures such as the Gini coefficient.

[16]The choice of France and USA reflect that previous studies have identified similar trends in income inequality among Continental European and Anglo Saxon countries respectively with France and USA being good representatives for the two groups (see e.g. Alvaredo et. al., 2013).

[17]As is discussed by Piketty and Saez (2003), part, but only part, of the increase is associated with the Tax Reform Act of 1986.

Figure 1.8

Top 1 per cent income share for Denmark 1870-2010 in international perspective
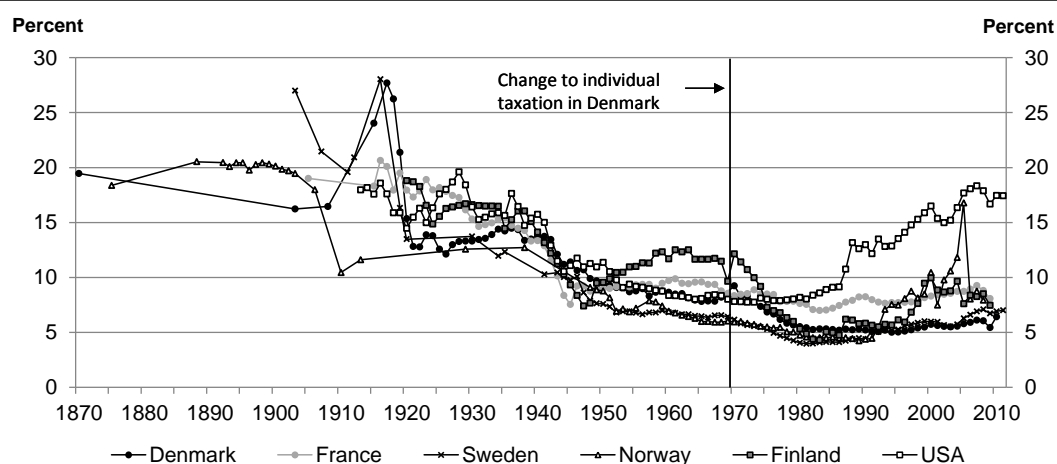
Notes: The income share for Denmark has been calculated using the definition of taxable income (gross income minus deductions), which we have adjusted for the grossing up of transfers in 1994 (with effect in the years prior to 1994) and the deductibility of paid personal taxes (with effect in the years 1908-1966). The series for other countries apart from Finland relate to gross income. The Finish series change income definition from taxable to gross income in 1990-92 and the two series have been linked by scaling the pre-1990 series to the post series. The vertical line in 1970 indicates the Danish change from family to individual taxation.

Sources: Own calculations and the World Top Incomes Database.

To find a share of more than 17 per cent in Denmark, one has to go back to the First World War and this highlights a second interesting feature: the sharp increase in income inequality during the First World War, which is present for both Denmark and Sweden, where the top 1 per cent income share reached a staggering 27-28 per cent in 1916-17. However, while the increase in one country might be dismissed as a statistical anomaly, the fact that the sharp rise is found independently in both countries suggests that something dramatic indeed happened during the First World War. Note also that the increase during the First World War is not a consequence of a collapse of the income total: the income total in Denmark increased on average 10 per cent annually (in nominal terms) from 1908 to 1918 compared to an average increase of 6 per cent from 1903 to 1960.

1.4 Factors behind the evolution of top income shares in Denmark

The evidence presented in the previous section show that the level of income inequality in Denmark has changed dramatically over the past 140 years. Top income shares appear to have fallen between 1870 and the beginning of the twentieth century. There was a sharp rise and then fall in inequality associated with the First World War. Inequality then fell over the rest of the century in an episodic manner, not as a continuing trend, with marked falls in the Second World War and after 1970. In this section, we discuss some of the forces lying behind the observed evolution of income inequality.

World Wars

The sharp rise in inequality during the First World War, found also for top shares in Sweden, cannot, as we argued in the previous section, be dismissed as a statistical anomaly. The contrast be-tween the dramatic increase in measured inequality during the First World War and the decrease (in both Denmark and Sweden) during the Second World War is therefore interesting. It is true that the two situations were different in that Denmark managed to stay neutral during the First World War and was occupied during the Second World War (the situations were less different in the case of Sweden). But during the occupation Denmark was able to maintain its own government with a high level of autonomy over internal affairs until 1943, and economically both episodes meant a large increase in aggregate demand in particular for agricultural products, while imports such as fuel and coal were in short supply.

The different development in inequality may instead lie in the fact that during the First World War the Danish government largely expected the war to be over quickly and was thus slow to adopt measures such as rationing and price/rent control. Furthermore, the unions and employer organiza-tions had in 1911 settled on a 5 year collective agreement, which more or less dictated the nominal wage growth until 1916; this resulted in a large drop in real wages as documented by Lindberg (1921). In contrast, the potential economic consequences of the Second World War were much better foreseen by the Danish government, which therefore was quicker to implement rationing, price control etc. Also the unions reacted faster and demanded quarterly automatic wage adjustment to inflation in the collective agreement signed in March 1941. The two wars therefore point to the potential distributional consequences of increases in aggregate demand under sticky wages and prices, respectively. We recognise that the circumstances of wartime mean that money incomes may be a less reliable guide to living

standards than in peacetime, but believe that there are grounds for exploring further the wartime experiences, in Denmark and in other countries.

The effect of moving to an individual basis for taxation

A major change in the tax system was the move to individual taxation in 1970. As may be seen from figure 1.9, this occurred at a time when the share of the bottom 99 per cent was rising. It has also been suggested that this coincides with an increased proportion of Danish women entering the labour market.[18] This indicates that we should explore more fully the change from a family to an individual basis.

Figure 1.9

The income share of the bottom 99 per cent and the female employment rate
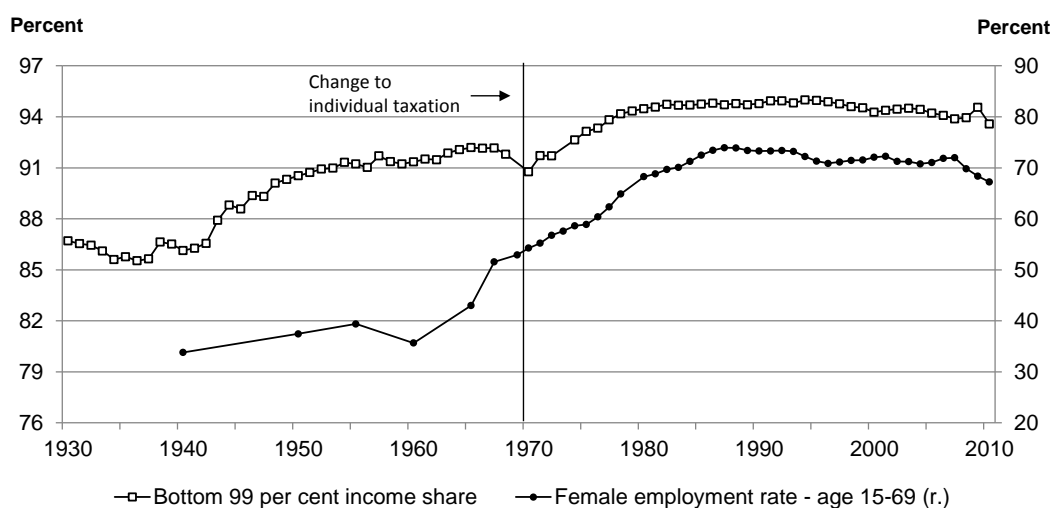
Notes: The vertical line in 1970 indicates the change from family to individual taxation. The participation rate is defined as total participation divided by the number of women between the ages of 15-69. For a description of the income definition, see figure 1.3.

Sources: Own calculations and Statistics Denmark.

As explained in Atkinson (2007a, p. 27), a move from a family to an individual unit could raise or lower top income shares, depending on the joint distribution of the incomes of husbands and wives. At one extreme, assuming that all individuals in the top income groups are either unmarried or married to someone with zero income, the

---

[18]Sørensen (1989, 1993) notes that the decline in inequality after 1970 is mainly is driven by a decline in inequality among secondary taxpayers (many of whom were outside the labour market at the beginning of the period), while inequality among primary earners is stable.

change only affects the top income shares through a change in the total population, in which case we can remove the break in the series by simply changing the population total to all individuals age 15. Doing this for the year 1968 yields an increase in the income share of 8.4 percentage points for the top 10 per cent and 2.0 percentage points for the top 1 per cent, cf. table 1.3, which can be compared to the actual increases from 1968 to 1970 of 2.7 and 1.0 percentage points respectively.

Table 1.3

The effect of the change from family to individual taxation

	Individual		Family	
Per cent	**Top 10 per cent**	**Top 1 per cent**	**Top 10 per cent**	**Top 1 per cent**
1967	38.7*	9.8*	30.4	7.8
1968	39.2*	10.2*	30.8	8.2
1970	33.5	9.2		
1971	32.5	8.3		

Notes: Families refer to age 15+ excluding married women. Individual refers to all individuals age 15+. 1969 was a "tax free year" and is thus excluded from the table. * Calculated based on the assumption that all taxpayers in the top income shares were either unmarried or married to someone with zero income.

Sources: Own calculations based on the series for taxable income.

The fact that this calculation overestimates the effect of the change in tax units indicates that the families in the top income groups before 1970 had a non-negligible income from secondary earners. The change to individual taxation thereby resulted in a division of the family income between two individuals thus tending to reduce the top shares. With no information from the tabulated data on the income distribution within each family, there is no easy fix to join the series across the change in tax units. One cannot simply buckle the two series together by scaling, as the timing in the change from family to individual taxation can have a big impact on not just the recorded level, but also the development in income inequality.

The role of taxation

In the recent studies of the long run development of top incomes one of the key elements has been the effect of taxes, see Piketty et. al., 2014 and Alvaredo et. al., 2013. These authors conclude that the rising marginal tax rates were one of the reasons why top income shares did not recover after the Second World War, as high marginal tax rates impaired the incentive (or capacity) to accumulate capital at the top. Similarly, the

global divergence in top income shares since the 1970s might be explained by differences in the development of marginal tax rates.

The rise in tax progressivity in Denmark is shown in figure 1.10, which depicts the marginal tax rate at the income cut-off for the top 1 per cent together with the income share of the top 1 per cent. The marginal tax rate is shown both as the statutory rate (τ), and the effective rate, where the latter takes into account that the tax prior to 1967 was levied on "assessed income" where paid personal taxes in the previous year had been deducted. At the time policy makers typically calculated the effective tax rates as a so-called "equilibrium tax rate", which was based on the fact that under the assumption of a constant income and tax schedule, the effective marginal tax rate converges to $\tau/(1+\tau)$.[19] A statutory marginal tax rate of 50 per cent thus corresponded to an equilibrium marginal tax rate of 33.3 per cent. This implied that the statutory marginal tax rates could be higher than 100 per cent, which they indeed were in the 1950s and 1960s.

---

[19]The formula can be derived by considering a one-year increase in a tax payer's income. The first year this increases his tax liability by t. The next year the extra tax paid decreases his tax liability by t2, which the third year increases his tax liability by t3 and so forth. Disregarding discounting this process can be shown to converge to t/(1+t).

Figure 1.10

The income share of the top 1 per cent and the marginal tax rate at the top 1 percent cut-off
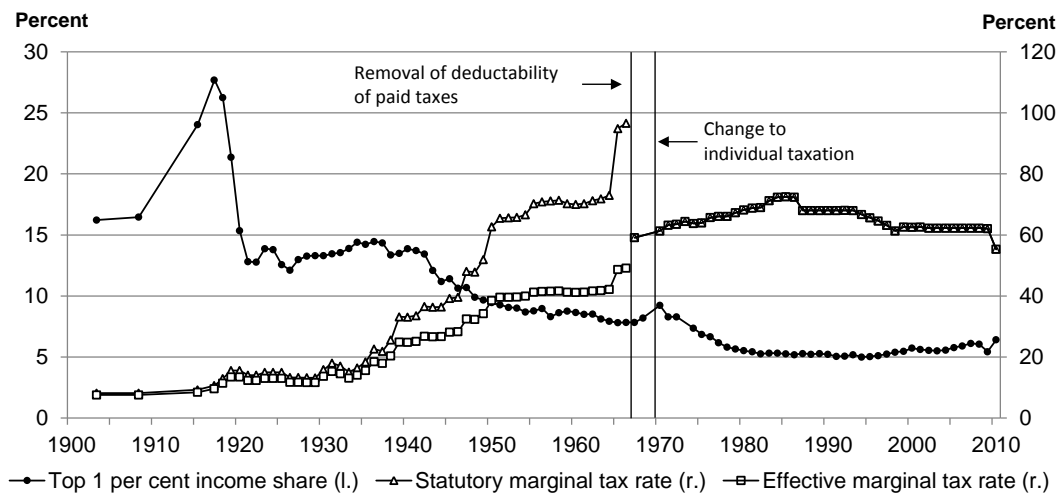
Notes: The marginal tax rate is measured at the income cut-off for the top 1 per cent and before 1970 it takes into account the ordinary state taxation, the common municipal fund law and average municipal taxation. The municipal average taxation (as per cent of the assessed income to the state) can be found in the statistical yearbook back to 1927. These rates have been adjusted by a factor 1.25 to take into account of the fact that the municipalities gave larger deductions before calculating the taxable income (in 1967 and 1968 the municipal income tax base was only 80 per cent of that of the state). Before 1927 the average municipal tax rate has been assumed constant at a level of 6.6 per cent. From 1970 the marginal tax rates are published by the Danish Ministry of Taxation. The effective marginal tax takes into account the effect of the deductibility of paid personal taxes under the assumption of a constant income level. From 1967 the two tax rates are identical, since taxes paid could no longer be deducted. Until 1987 the marginal tax rate applies to almost all income courses. After this point capital income is taxed at a lower rate. The vertical lines indicate the two data breaks of the removal of the deductibility of paid personal taxes in 1967 and the change from family to individual taxation in 1970.

Sources: Johansen (2007), Philip (1965), the Ministry of Taxation and own calculations.

Our calculation of effective tax rates uses the same formula and, as may be seen from figure 1.10, before 1967 the deductibility of paid personal taxes created a marked difference between statutory and effective tax rates. It should further be noted that the marginal tax rate at the top 1 per cent cut-off was not necessarily the top marginal tax rate: a general feature of the Danish tax system at the beginning of the 20th century was

the large number of tax brackets going high up in the income distribution. Figure 1.10 shows that the development of the effective marginal tax rates at the top of the income distribution have followed an inverted U since the beginning of the 20th century, which inversely mirrors the top 1 per cent income share.

Looking at the effective tax rates, we see that the first rise in the marginal tax rate (at this income level) came during the First World War. It then levelled off until the mid-1930s, after which it increased quite substantially until the beginning of the 1950s. The increase continued until the mid-1980s, where the marginal tax rate peaked at the same time as the top 1 per cent income share reached a historical low point. Since then marginal tax rates have decreased, while the top 1 per cent income share has increased slightly.

At a first glance the Danish case therefore seem to support the conclusions from the earlier studies that the development in (pre-tax) income inequality to some extent is driven by changes in tax rates. However, just as in these earlier studies, we have to question whether this effect can be interpreted as a standard labour supply effect. The reason is that tax rates did not only increase at the top but also for the lower income groups, and interpreting the development in income shares as effects of changes in labour supply therefore require relatively large differences in labour supply elasticities across income groups.

To see this formally note that we can write the relative income shares (S) of two groups (i, j) as:

$$\frac{S_i}{S_j} = \frac{w_i N_i}{wN} / \frac{w_j N_j}{wN} = \frac{w_i N_i}{w_j N_j},\tag{1.2}$$

where w_i is the average income in group i and N_i is the number of individuals in the group, while the same variables without the subscripts are the average income and number of individuals in the total population. Now, decomposing changes in S into effects of taxation given by the group specific mar-ginal tax rate (τ) and other factors (z) and log differentiating with respect to time (t) yields the following:

$$\frac{\partial}{\partial t} \ln \left(\frac{S_i}{S_j} \right) = \varepsilon_i \frac{\partial \ln (1 - \tau_i)}{\partial t} - \varepsilon_j \frac{\partial \ln (1 - \tau_j)}{\partial t} + \frac{\partial \ln (w_i)}{\partial z_i} \frac{\partial z_t}{\partial t} - \frac{\partial \ln (w_j)}{\partial z_j} \frac{\partial z_j}{\partial t},\tag{1.3}$$

where we have used that $\partial ln(N_i)/\partial t = \partial ln(N_j)/\partial t$ by definition, when the considered income groups constitute a fixed faction of the total population and the definition of the labour supply elasticity (more correctly the elasticity of taxable income) $\varepsilon = \partial ln(w_i)/\partial ln(1 - \tau_i)$. In this setting, interpreting the changes in income shares over time as solely driven by changes in taxation corresponds to assuming that the two

last terms in equation (1.3) cancel out. The relationship between the two labour supply elasticities is thus given by:

$$\frac{\partial}{\partial t} \ln (S_i) - \frac{\partial}{\partial t} \ln (S_j) = \varepsilon_i \frac{\partial \ln (1 - \tau_i)}{\partial t} - \varepsilon_j \frac{\partial \ln (1 - \tau_j)}{\partial t} \tag{1.4}$$

Applying this formula to the Danish series of top income shares and assuming an elasticity of 0.2 for the P90-P95 income group (corresponding to the estimate obtained for Denmark by Kleven and Schultz (2014) for the period 1980-2005) implies an elasticity of 0.64 for the P95-P99 group and an elasticity of 1.49 for the top 1 per cent of the income distribution, when considering the period 1903 to 1965, cf. table 1.4. Interpreting the changes in top income shares as solely an effect of changes in taxation thus requires that the underlying elasticities increase relatively quickly with income and to levels that seem to high compared with the consensus in the literature (see Saez et. al., 2013).

Table 1.4

Changes in income shares and marginal tax rates and implied elasticities

	Income share			Marginal net-of-tax rate[a)]		
	P90-P95	**P95-P99**	**P99-P100**	**P90-P95**	**P95-P99**	**P99-P100**
1903	10.57	15.10	16.21	0.93	0.93	0.92
1965	10.94	12.13	7.79	0.54	0.53	0.51
Log change	0.03	-0.22	-0.73	-0.53	-0.56	-0.59
Implied elasticity[b)]		0.64	1.49			

Notes: a) Measured at the bottom income threshold for the group. b) Assuming an elasticity of 0.2 for the P90-P95 group. If we instead assume an elasticity of 0.1 the elasticities for the P95-P99 and P99-P100 groups become 0.54 and 1.39 respectively.

Sources: Own calculations.

1.5 Summary

This chapter has been concerned with the long run development of income inequality in Denmark. To this end, we have constructed estimates of the top income shares in Denmark dating back 140 years. Compared to earlier studies of income inequality in Denmark these series are unique in that we pay strict attention to the comparability of the estimates over time and bring the methodology in line with that used in the recent studies of other countries. Using the constructed time series, it is possible to give some answers to the questions posed at the outset. There have been periods in the past when Denmark has seen significant reductions in inequality: (possibly) in the last 30

years of the 19th century spanning the start of the industrialization in Denmark, and definitely over the Second World War, and in the 1970s. Income inequality has thus declined during several distinct phases, and even though there has been an increase in inequality since the 1980s, inequality has by historical standards remained low. This time path follow relatively close the time path in Sweden, and the Danish case thus adds to the picture of global co-movement in inequality until the 1970s and thereafter divergences.

In contrast to the studies of other countries, the Danish tax records cover a relatively large pro-portion of the population almost from the beginning of our sample period, and this enables us to an-swer two important questions. Firstly we are able to assess to what extent changes in the top income shares proxy the underlying development in inequality and secondly we can address the question as to which groups of the income distribution benefitted from the decline in the top income shares.

The answer to the first question is that there is a relatively stable relationship between changes in the income share of the top 1 per cent and changes in overall income inequality as measured by the Gini coefficient, implying that top 1 per cent income share indeed is a good proxy for underlying de-velopment in inequality, when the data quality restrict the calculation of such measures. Furthermore the development in the top 1 per cent income share seems to be relatively robust to the use of different income and population definitions, except during radical changes of the in structure of the economy, such as women entering the labour market.

The answer to the second question is that the decline in the top 1 per cent income shares pri-marily benefitted the part of the income distribution below the 70th percentile, while the income shares of the income groups between the 70th and 95th percentile have been remarkably stable over the past century. The fact that primarily the top 1 per cent share has fallen, while the income share of e.g. the group between the 90th and 95th income percentile has remained constant, seems to support the conclusions drawn by Piketty et. al., (2014) and Alvaredo et. al., (2013) that the long run trends do not merely reflect standard labour supply responses to increases in tax rates.

References

Aaberge, R. and Atkinson, A. B. (2010), Top Incomes in Norway, in Atkinson, A. B. and Piketty, T. (eds.), Top Incomes: A global perspective, Oxford University Press.

Alvaredo, F. (2011), A Note on the Relationship Between Top Income Shares and the Gini Coefficient, Economics Letters 110, 274-277.

Alvaredo, F., Atkinson, A. B., Piketty, T. and Saez, E. (2013), The Top 1 Percent in International and Historical Perspective, Journal of Economic Perspectives 27, 3-20.

Atkinson, A. B. (2005), Top Incomes in the UK Over the 20th Century, Journal of the Royal Statistical Society 168, 325-343.

Atkinson, A. B. (2007a), Measuring top incomes: Methodological issues, in Atkinson, in A. B. and Piketty, T. (eds.), Top Incomes over the Twentieth Century - A Contrast Between European and English-Speaking Countries, Oxford University Press.

Atkinson, A. B. (2007b), The distribution of top incomes in the United Kingdom 1908-2000, in A. B. and Piketty, T. (eds.), Top Incomes over the Twentieth Century - A Contrast Between European and English-Speaking Countries, Oxford University Press.

Atkinson, A .B. and Søgaard, J. E. (2013), The long-run history of income inequality in Denmark: Top incomes from 1870 to 2010, EPRU working paper series 2013-01.

Atkinson, A. B., and Piketty, T. (2007), Top Incomes over the Twentieth Century - A Contrast Between European and English-Speaking Countries, Oxford University Press.

Atkinson, A. B., and Piketty, T. (2010), Top Incomes: A global perspective, Oxford University Press.

Atkinson, A. B., Piketty, T., and Saez, E. (2011), Top Incomes in the Long Run of History, Journal of Economic Literature 49, 3-71.

Baxter, R.D. (1868), National income, The United Kingdom, Macmillan.

Bjerke,   K. (1957), Changes in Danish Income Distribution 1939-52, Review of Income and Wealth 6, 98-154.

Bjerke,   K. (1965), Forskydninger i den Personlige Indkomstfordeling 1939 til 1964 (Shifts in Personal Income Distribution 1939-1964), Socialt Tidsskrift 41, 220-230.

Bjerke,   K., and Ussing, N. (1957), Studier over Danmarks National Produkt 1870-1950 (Studies of the Danish Domestic Product 1870-1950), Danmarks Statistik.

Burkhauser, R. and Larrimore, J (2014), Journal of Economic Perspectives 28, Correspondence sections, 243–248.

Egmose,   S. (1985), Udviklingen i den Personlige Indkomstfordelingen (Trends in the Personal Income Distribution), in Socialforskningsinstituttet, Uligheden, Politikerne og Befolkningen 139.

Jäntti,   M., Riihelä, M., Sullström, R. and Tuomala, M. (2010), Trends in Top Income Shares in Finland; in Atkinson, A. B. and Piketty, T. (eds.), Top Incomes: A global perspective, Oxford University Press.

Johansen,   H. C. (2007), Indkomstskatter og Offentlig Vækst 1903-2005, Dansk Skattehistorie VI (Income taxes and public growth 1903-2005, Danish Tax History VI). Told- og Skattehistorisk Selskab.

Kuznets,   S. (1953), Shares of Upper Income Groups in Income and Savings, National Bureau of Economic Research, New York.

Kuznets,   S. (1955), Economic Growth and Economic Inequality, American Economic Review XLV, 1-28.

Kleven,   H. and Schultz, E. (2014), Estimating Taxable Income Responses Using Danish Tax Reforms. American Economic Journal: Economic Policy 6, 271-301.

Lindberg,   J. K. (1921), Indtægtsforskydningen i Danmark 1908-20 (Income shifts in Denmark 1908-20), Nationaløkonomisk Tidsskrift 3.

Pedersen,   P. J., and Smith, N. (2000), Trends in Danish Income Distribution, Labour 14, 523–546.

Philip,   K. (1965), Skattepolitik (Tax policy), Gyldendal, 2 edition.

Piketty,   T. (2001), Les hauts revenus en France au 20ème siècle (High earners in France in the 20th century), Grasset, Paris.

Piketty,   T. (2003), Income Inequality in France 1901–1998, Journal of Political Economy 111, 1004-1042.

Piketty,   T., and Saez, E. (2003), Income Inequality in the United States, 1913-1988, Quarterly Journal of Economics 118, 1-39.

Piketty   T., Saez, E. and Stantcheva, S. (2014), Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities, American Economic Journal: Economic Policy 6: 230-271

Roine,   J. and Waldenström, D. (2010), Top Incomes in Sweden over the Twentieth Century; in Atkinson, A. B. and Piketty, T. (eds.), Top Incomes: A global perspective, Oxford University Press.

Soltow,   Lee (1979), Wealth distribution in Denmark in 1789, Scandinavian Economic History Review 27, 121-138

Sørensen,   R. S. (1989), Udviklingen i den Danske Personlige Indkomstfordeling i 1900-tallet, Master's thesis, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Bygning 26, 1353 Copenhagen K.

Sørensen,   R. S. (1993), Changes in the Personal Income Distribution 1870–1986, in Persson, K. G. (eds.), The Economic Development of Denmark and Norway since 1870. Aldershot, U.K.

World   Top Incomes Database, http://www.parisschoolofeconomics.eu/en/expertise-dissemination/wtid-world-top-incomes-database/

2 Labor Supply and Optimization Frictions: Evidence from the Danish student labor market

2.1 Introduction[1]

Labor supply elasticities – or more generally earning elasticities – are key parameters in many areas of economics, e.g. optimal income taxation (Saez et al., 2012). However empirical identification of these parameters remains a challenge – especially in the likely presence of optimization frictions, where Chetty (2012) shows that even small optimization frictions limits the researcher to identify only bounds on the elasticities. Bounds that in many case are so wide that it is likely to dwarf many of the econometric issues involved in the identification.

In this chapter I shed light on the presence and underlying nature of these frictions by studying the labor supply of Danish students. So far concrete evidence on frictions has been relatively limited in the economics literature on labor supply, which reflect that identification of optimization frictions typically requires both high quality data and special institutional settings – high quality data in order not to confound optimization errors by individuals with measurement error in the data and special institutional settings that allow separation of rational behavior from optimization errors. Kleven and Waseem (2013) is one of the few papers that fulfils both of these requirements, which enables them to estimate both a structural labor supply elasticity and the level of optimization frictions in a Pakistani setting, while remaining agnostic about the underlying

nature of frictions.[2]

The labor market for Danish students represents an interesting case study for learning about optimization frictions for several reasons: 1) students face a sharp kink in their budget set created by the phase out of student benefits, 2) in 2009 a reform significantly increased the earnings level at which students reach the kink point and 3) students face a special institutional setting, where they effectively can choose between different budget sets. The use of the Danish student labor market as a case study further benefits from the fact that the labor market is covered by rich register.

The strength of having all of these institutional settings within a well-defined labor market is that it allows you to distinguish between 3 of the main types of optimization frictions discussed in the literature – namely real adjustment costs (Attanasio, 2000), gradual learning (Mankiw and Reis, 2002 and Evans and Honkapohja, 2001) and (rational) inattention (Sims, 2003) – by examining the outcomes around each setting.

My main findings are the following: First, following the 2009 reform I find an immediate and non-trivial shift in the students' earnings distribution compared to a very stable distribution both before and after the reform. Second, despite this clear evidence of a positive labor supply elasticity I find no sign of bunching at the kink point created by the phase out of student benefits. Finally, I find that a significant share of students fail to choose the budget set that is optimal given their final level of earnings.

Taken together, these findings point to the presence of significant optimization frictions that mask the bunching at the kink point predicted by a standard labor supply model (Saez, 2010). However, the findings do not point to real adjustment costs or learning as the main types of frictions, as these types of frictions would lead to a more gradual transition to a new earnings distribution following the 2009 reform.

Instead the findings are consistent with a model, where individuals (rationally) choose their desired labor supply and earnings, but where final earnings may deviate from this level due to unexpected shock to e.g. the wage rate. If individuals fail to realize such shocks and reoptimize behavior, their final earnings will deviate from their desired level of earnings, which prevent the formation of clear bunching in the earnings distribution, even if individuals quickly change their desired earnings in response to change in the institutional settings. Put differently the findings suggest that the dominant optimization frictions are individuals' inattention about their earnings process during the year.

After presenting graphical evidence on the above findings I proceed with a discussion

---

[2]In other contexts such as e.g. consumption, Chetty et al. (2009) show that salience of taxes can affect demand.

of how to quantity the behavioral responses. This is not a trivial task as the lack of clearly visible excess mass in the cross sectional setting makes it impossible to employ the standard bunching method developed by Saez (2010), and because the differences between individual desired and final earnings effectively mix the treatment and control groups as they are normally defined in a difference-and-difference estimation.[3]

Instead I propose a method that resembles the method used by Chetty et al. (2013) and use the shift in the distribution following the 2009 reform to uncover the (local) counterfactual distribution at the kink point. Having the counterfactual distribution I use the bunching method to translate the observed responses into elasticities from which I obtain lower bound estimates around 0.05-0.06.

2.2 Optimization frictions and labor market outcomes

Before moving into the empirical analysis I start by drawing a number of hypotheses about how different types of optimization frictions affect observed labor market outcomes around different stylized institutional settings. These will in section 2.3 be related to the actual institutional settings facing Danish students. More concretely I consider the following 3 stylized settings:

1. A kink point in the budget set created by a jump in the marginal tax rate.

2. A tax reform that changes tax rates in some parts of the income distribution.

3. Voluntary take up of benefits.

Of these, the 2 firsts are standard institutional settings considered in the public finance literature, whereas the 3rd needs additional explanation:

The basic point is that the take up of benefits might only be optimal for individuals in certain earnings intervals. Consider e.g. a stylized benefit system consisting of a lump sum grant that is phased out with earnings according to some schedule. In most real life benefit systems taking up benefits would from an economic perspective always be optimal, as the phase out stops once (net) benefits reach 0. The budget set created by taking up benefits would thus always (weakly) dominate the budget set without benefits. However if the phase out is "prolonged" beyond the breakeven point, taking up benefits is not always optimal as illustrated in figure 2.1.

---

[3]This method is used by among others Feldstein (1995) and Gruber and Saez (2002). See Kleven and Schultz (2013) for an application on Danish data.

Figure 2.1

Illustration of the potential sub-optimality of taking up of benefits
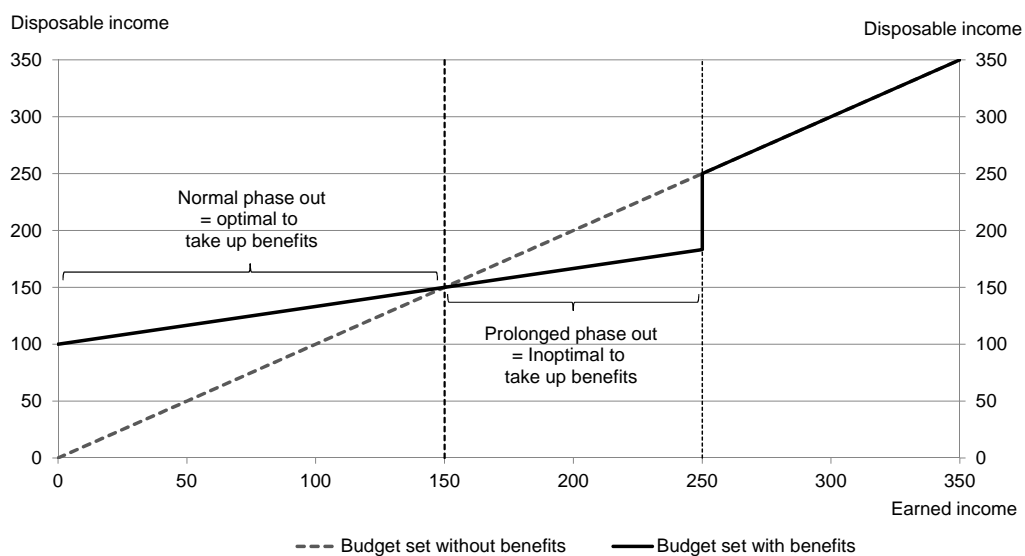
Notes: The budget set without benefits (or taxation) equal the 45 degree line, while the budget set with benefits that gives a lump sum grant of 100, which is phased out with earnings at a rate of 75 percent. At this phase out rate net benefits reach 0 at earned income of 150. If the phase out is "prolonged" beyond this point, it create a range of earnings where it is sub-optimal to take up benefits.

The figure shows a stylized budget set without benefits (or taxation) equal to the 45 degree line and a budget set under a benefit system that gives a lump sum grant of 100 that is phased out with earnings at a rate of 75 percent. At this phase out rate net benefits reach 0 at an income of 150 and take up of benefits is therefore optimal with income below this point. In contrast, it is inoptimal to take up benefits with income above this point if the phase out is prolonged.

From these 3 stylized policy settings it is possible to draw a number of hypotheses about what outcomes we should expect to find under the presence of different types of optimization frictions. More concretely I consider the effect of 3 broad groups of optimization frictions – namely:

1. Real adjustment cost on the labor market.

2. Gradual learning about the institutional settings.

3. (Rational) inattention.

However before doing so I start by considering the labor market outcomes in a world without optimization frictions. In this case individuals would bunch at the kink point created by the jump in the marginal tax rate and thereby create clear excess mass in the earnings distribution at this point, with the excess mass being proportional to the labor supply elasticity (Saez, 2010). Following a tax reform that changes tax rates in some part of the income distribution, we should find an immediate change in earnings for the individuals who are directly affected by the change in incentives and finally, we should expect individuals to only take up benefits if it increases their disposable income – i.e. no one with earnings in the "prolonged" range shown in figure 2.1 should take up benefits.

Against this benchmark I start by considering the effect of real adjustment costs in the labor market (see e.g. Attanasio, 2000). Real adjustment costs imply that it is costly for individuals to change their earnings, e.g. because it requires finding a new job, which might take time and effort. In this case, individuals are willing to accept jobs located in an earnings interval around their optimal point, as the expected benefits of renewed search do not outweigh the search costs (Chetty et al., 2011). As a consequence only a fraction of the individuals, who in a frictionless world would bunch at the kink point, do so in this case causing the excess mass to be spread over an interval around the kink point (fuzzy bunching).

When it comes to the effect of a tax reform, the presence of real adjustment costs imply that not all individuals will find it optimal to change their income immediately. Instead they might choose to keep their current job if they e.g. expect that they in the near future have to change job for other reasons. As a consequence we should expect to see only a gradual change in the earnings distribution.

Finally, real adjustment costs in the labor market should not necessarily have anything to do with individuals being able to decide whether or not to take up benefits. As long as the administrative system is fairly simple, the economic costs of taking up benefits are trivial, and we should therefore expect individuals to take up benefits optimally given their current job choice, even if this choice deviates from what they would have chosen in a frictionless world.

The second general class of optimization frictions that I consider is gradual learning (see e.g. Mankiw and Reis, 2002 and Evans and Honkapohja, 2001). Gradual learning implies that individuals do not have perfect information about the institutional setting, when they are new to the system or when the system is changed. This would e.g. include knowledge of the precise position of the kink point and the design of the benefit system, and as consequence we should expect only fuzzy bunching around the actual

kink point and sub-optimal take up of benefits – especially among individuals with less experience with the institutional settings.

Likewise, gradual learning implies that the knowledge of a reform would expand gradually after its implementation and we would therefore expect to see a gradual change in the earnings distribution.

Finally I consider the effect of (rational) inattention (see Sims, 2003). Rational inattention builds on the idea that economic circumstances might change over time, but that it is costly for individuals to keep close attention to these changes. Changing circumstances, which in a frictionless world would have warranted reoptimization of individual behavior, therefore might not be noted by individuals leaving them with ex post sub-optimal behavior.

Formulated in this way there is a potential big overlap between gradual learning and inattention, as e.g. inattention about changes in the institutional setting will be exactly the same as the gradual learning described above. I will therefore make the following distinction between gradual leaning and inattention: Gradual learning refers to learning about institutional settings that we normally would think as constant in the long run (changes in institutional settings such as tax rates only happen as a result of reforms). In contrast, inattention refers to inattention about individual economic factors that may vary even in the long run – factors such as individual wages, working requirements etc. In a world were these individual factors are partly random, individuals will never learn the true values of these by accumulated experience, but can only know them by paying close attention to their evolution.

Applied to the labor market, inattention implies that individuals will aim at a desired level of labor supply and earnings, but that their final earnings will be distributed around this level due to random shocks to individual economic factors, which the individuals fail to realize and thus offset by reoptimization. As consequence we should expect only fuzzy bunching around a kink point in the budget set. Likewise we should expect to see some individuals take up benefits even though it ex post turns out to be a sub-optimal choice. However, despite of the inattention about the evolution of individual economic factors, we should expect to see an immediate change in the earnings distribution following a tax reform, as individuals adjust their desired earnings to the new incentives.

Finally it should be noted that the notion of inattention as being rational rely on the pre-sumption that the costs of paying closer attention to changes in the economic circumstances outweigh the expected benefits of smaller optimization errors. However more generally inattention might also be irrational just as the inattention might also be

related to the effects of the individuals' own actions – e.g. in the labor market, where individuals' labor supply and earnings may vary from month to month, while taxation is based on the cumulative earnings over the year. In this case, knowing the effect of extra earnings in one month requires the individuals to keep track of (and predict) earnings in all months.

The predictions from the different hypotheses described above are summarized in table 2.1 and as the table shows each type of optimization frictions leads to a unique set of predictions across the different institutional settings. Combining the observed outcomes across these settings therefore in principle allow you to distinguish between different types of frictions.

Table 2.1

Hypotheses: What to expect under different types of optimization frictions?

	Bunching at the kink point	**Effect of a tax reform**	**Take up of benefits**
Benchmark:			
No frictions	Clear bunching	Immediate change	Optimal take up
Optimization frictions:			
Real adjustment cost	Fuzzy bunching	Gradual change	Optimal take up
Gradual learning	Fuzzy bunching	Gradual change	Sub-optimal take up
(Rational) inattention	Fuzzy bunching	Immediate change	Sub-optimal take up

2.3 Institutional settings: students' incentive to earn income

In this section I present the key features of the Danish student benefit system and relate them to the stylized institutional settings discussed in section 2.2.[4]

Danish students enrolled in education above primary school (ISCED2011 level 3 and above) are eligible to state financed student benefits from the age of 18. Benefit rates vary depending on the type of education and civil status, but in 2008 the basic rate for typical students enrolled in tertiary educations (ISCED2011 level 5 and above) was 5,000 DKK per month (1 USD $\approx$ 6 DKK).

In addition to receiving these benefits, students are allowed to earn income of up to 6,400 DKK per month.[5] If they earn more than this baseline income limit (on a yearly

---

[4]A more detailed description can be found in appendix 2.A.

[5]Income counted against the income limit is called "own income" and includes labor income, transfers other than student benefits and capital income with the exceptions of certain types of stock income. All relevant variables are drawn from detailed register data organized by Statistics Denmark (DST) covering the entire Danish population. A more detailed description of these registers and the variables used can be found in appendix 2.B.

basis) the excess is deducted from the amount of benefits they are eligible for thus creating a kink in their budget set. Of the first 9,500 DKK 50 percent is deducted, while further excess earnings is deducted 100 percent.[6]

If students want to earn more they can increase the limit by cancelling one or more months of benefits. By cancelling one month of benefits a student increases the income limit by 9,500 DKK, which translates into a phase out rate of 5,000/9,500 = 52 percent. Administratively, it is fairly easy for students to cancel benefits, as it is done through a simple webpage, where students can click benefits in individual months on and off.

Taken together with the normal income tax system, which – for incomes in the range considered here – imposes a marginal tax rate of 41 percent (excl. VAT) the phase out of benefits causes the effective marginal tax rate jumps from 41 to 72 percent when students' earnings exceed 76,400 DKK annually.[7]

However, the effective marginal tax rate might jump even more if students fail to cancel the right amount of benefits and thereby end up hitting the phase out rate of 100 percent. If e.g. a student earns more than 9,500 DKK above the baseline limit and does not cancel student benefit he faces as marginal tax rate of 100 percent. In this case it would be optimal to cancel one month of benefits in order to lower the marginal tax rate to 72 percent.

This problem corresponds to the problem of optimal take up of benefits described in section 2.2, where 12 months of benefits are optimal for students earning up to 86,000 DKK annually. 11 months are optimal for students with earnings between 86,000 and 95,500 DKK. For students earning extra 9,500 DKK 10 months is optimal etc., as illustrated in figure 2.2.

---

[6]Finally, if the amount of student benefits that a student has to pay back exceeds 7,600 DKK (2008 level), the entire payback is increased by 7 percent. This notch implies that the marginal tax rate for excess earnings above this amount exceeds 100 percent. This is not shown in figure 2.2.

[7]There is a caveat to the calculation of the effective marginal tax rate, when students cancel student benefits. For most university students student benefits are limited to a period of 6 years (compared to a standard study time of 5 years) and by cancelling a number of months of benefits, the student can save them for later use. Some student might therefore not see the cancelling of benefits as the full loss assumed here. The probability of this does not significantly affect the conclusions drawn in the chapter and are discussed in section 2.4 and 2.6 below.

Figure 2.2

Effective budget sets for students depending on benefits take up, 2008
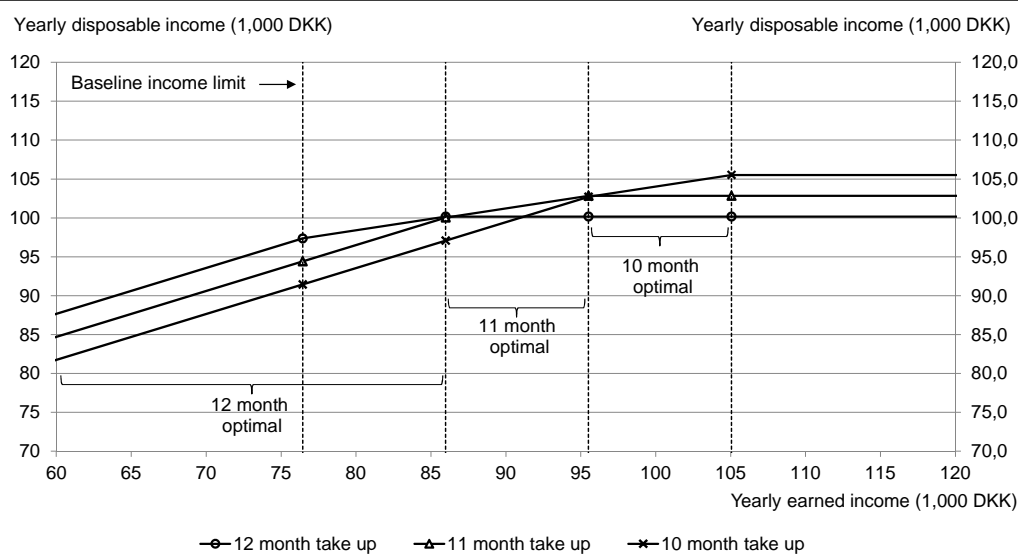
Notes: The baseline income limit is calculated as 12 x the monthly basic amount of 6,400 DKK. Yearly disposable income is calculated as first gross income consisting of 5,000 DKK x the number of months of benefits taken up plus earned income up to the income limit, which increases by 9,500 DKK for each month not taken up. Above this income limit the first 9,500 DKK in earned income is deducted in student benefits at 50 percent, while further excess is deducted 100 percent. Finally gross income is turned into disposable income based on a personal allowance of 41,000 DKK and a marginal tax rate in the normal income system of 41 percent. 1 USD $\approx$ 6 DKK.

Sources: Own calculations based on www.su.dk.

However the switch between budget sets by cancelling benefits is complicated by the fact that students have to do so actively prior to actually receiving the benefits. Cancelling benefits for a given month has to be done prior to the 15th the month before, while students typically receive their wage check at the end of the month or with an additional month's lag. E.g. cancelling benefits in December has to be done prior to November 15th, where students in general only have seen their wage checks up to October or September.

This time difference between, when students have to cancel benefits and when they have the actual information about the monthly (or yearly) income implies that students have to pay close attention to their income process during the year and to some degree predict what they will earn a couple of months into the future in order to cancel the right amount.

The student benefit system have remained largely unchanged through the period 2004-2011, which is considered in this analysis, except from a reform in 2009 that increase the baseline income limit by 25 percent for students enrolled in tertiary educations, while leaving it unchanged for lower levels of education, cf. table 2.2. At the same time the phase out rate for tertiary students was also increased from 52 to 62 percent thus causing an increase in the effective marginal tax rate from 72 to 78 percent.

Table 2.2

Development in the yearly baseline income limit

1,000 DKK	2006	2007	2008	2009	2010	2011
Students in tertiary education	72.1	74.1	76.4	97.7	101.6	103.5
Students in lower education	72.1	74.1	76.4	79.0	82.2	83.8

Notes: The baseline income limit refers to the income limit for students, who receive full benefits (12 months). Tertiary education includes university education and educations such as nurses and school teachers (ISCED2011 level 5 and above). Lower educations includes high school (gymnasium) and vocational educations (ISCED2011 level 3-4).

Sources: www.su.dk.

In what follows all numbers related to income variables have been translated to 2008 values using the indexation implied by the baseline income limit for students in lower educations.

2.4 Graphic evidence of labor supply responses and optimization frictions

In section 2.3 I linked the specific features of the Danish student benefit system to the stylized institutional settings listed in section 2.2. In this section I examine the labor market outcomes around each of the institutional settings and compare it with the hypotheses drawn in section 2.2.

Evidence from bunching at the kink point

Figure 2.3 shows the earnings distribution for students enrolled in tertiary educations before the 2009 reform centered on the baseline income limit. Only students, who are fully eligible for student benefits the entire year is included in this figure, however inclusion is not conditional on actually receiving student benefits (i.e. students are allowed to cancel benefits). Under the assumption that students cancel the right amount

of benefits, their effective marginal tax rate jump from 41 to 72 percent at the baseline income limit as described in section 2.2.

Figure 2.3

The earnings distribution for tertiary students, 2006-2008
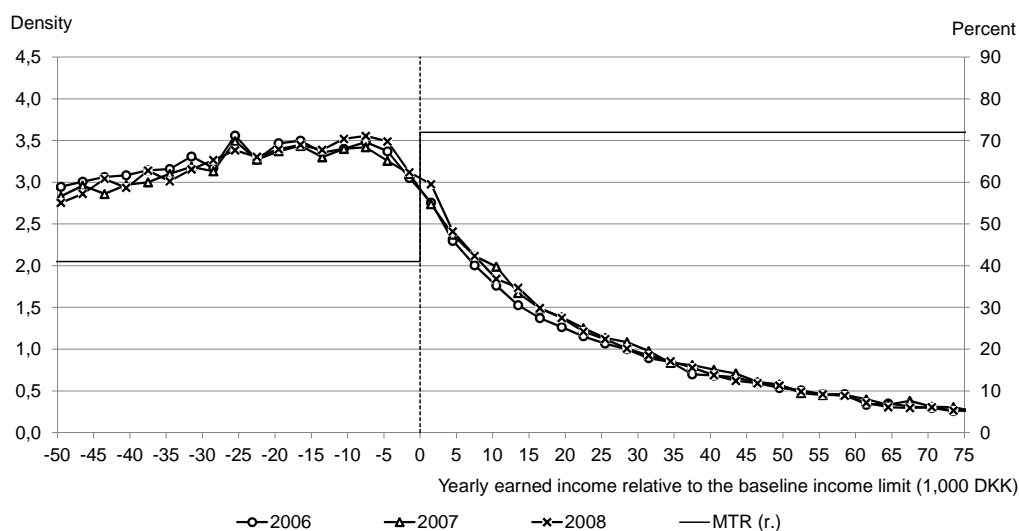
Notes: Students have to be fully eligible for student benefits (but necessarily receive student benefits) and have yearly earnings above 6,500 DKK to be included in the distributions. The marginal tax rate (MTR) is calculated under the assumption that students always cancel the optimal amount of student benefits. In that case MTR = 1 – (1-t)•(1-q), where t = 0.41 and q = 0 below the baseline income limit and q = 0.52 above. The baseline income limit was 76,400 DKK in 2008. Bin size = 3,000 DKK.

Sources: Own calculations based on DST

The figure shows that the earnings distribution was very stable during the 3 years prior to the reform and with no clear excess mass around the kink point. In a frictionless world this would imply that the labor supply elasticity was negligible, but from the cross sectional evidence alone – which most bunching studies rely on – we are not able to determine whether this outcome is truly driven by a zero labor supply elasticity or whether optimization frictions prevent the formation of a clear excess mass at the kink point. Naturally, we cannot distinguish between different types of optimization frictions either.

Evidence from the 2009 reform

When comparing the pre-reform earnings distribution with the distributions after the 2009 reform, we see in figure 2.4 a clear shift in the distribution with mass moving from below the initial kink point to a range above. Given the fact that the distribution was very stable in the years prior to the reform this shift constitutes compelling graphical evidence for a positive labor supply elasticity, suggesting that the lack of bunching at the kink points is due to optimization frictions.[8]

Figure 2.4

The earnings distribution for tertiary students before and after the 2009 reform
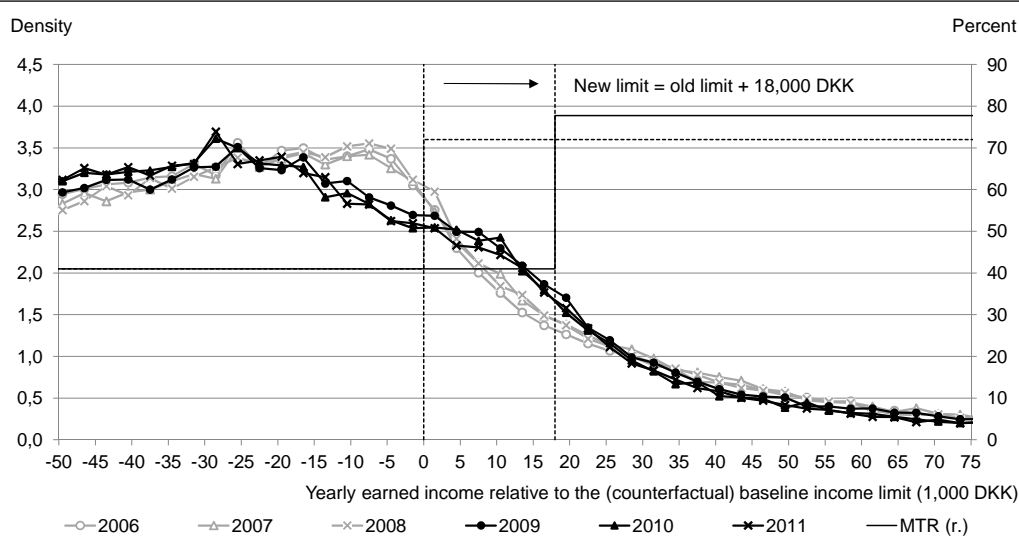
Notes: See notes to figure 2.3. For the years 2009-11 income is measured relative to the baseline income limit without the 2009 reform. This corresponds to the baseline income limit for students in lower education listed in table 2.2.

Sources: Own calculations based on DST

Furthermore, the fact that shift in the distribution appears to happen instantaneously from 2008 to 2009 speaks against both real adjustment cost and gradual learning as the dominant frictions. Taken together, the two first pieces of empirical evidence thus points to inattention as the dominant optimization frictions in this labor market.

It may finally be noted that the "excess mass" revealed by the shift in the distribution is centered below the kink point. I return to this finding in section 2.6 and discuss it in

---

[8]The interpretation of the shift in the earnings distribution as an indication of a positive labor supply response to the 2009 reform is also supported by the fact that the earnings distribution for students in lower educations, who was unaffected by the 2009 reform, remained stable.

greater details in appendix 2.E.

Evidence from the cancelling of student benefits

Turning to the cancelling of student benefits I consider the earnings distribution for students conditional on the amount of student benefits they cancel. In figure 2.5 this is done for students who have cancelled exactly 1 month and thus taken up 11 month of benefits.

By cancelling 1 month of benefits these students increased their income limit to 86,000 DKK (before the 2009 reform) and we should not expect to find students with earnings 9,500 DKK above this amount (where they reach the 100 percent marginal tax rate). If they wished to earn more they should have cancelled an extra month of student benefits in order to increase the income limit and lower their effective marginal tax from 100 to 72 percent.

Figure 2.5

The earnings distribution for tertiary students who cancel 1 month of student benefits
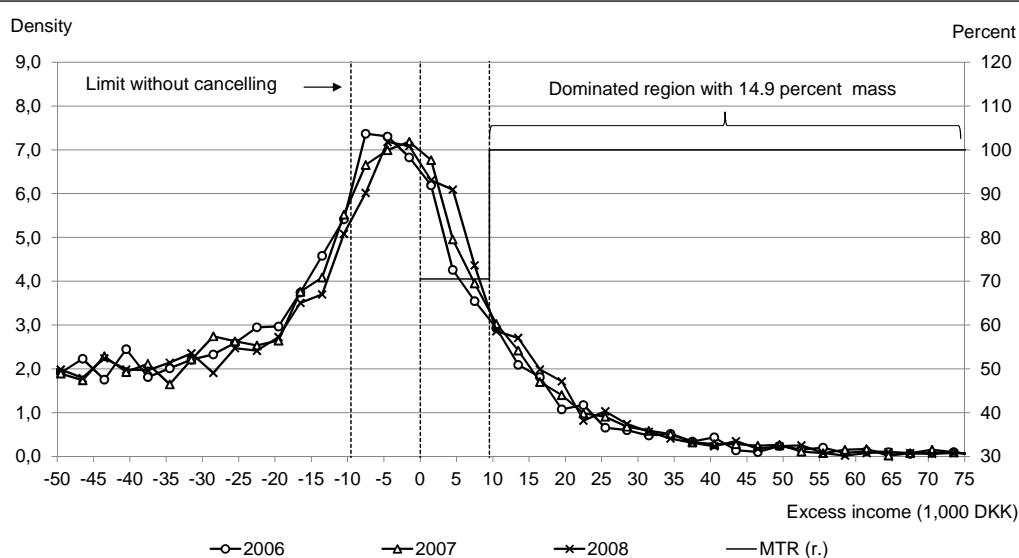
Notes: Excess income is defined as the yearly earning income relative to the actual income limit that the individual is facing. The marginal tax rate (MTR) is calculated using the formula MTR = 1 – (1-t)•(1-q), where t = 0.41 and q = 0.50 for the first 9,500 DKK above and q = 100 above this level. Bin size = 3,000 DKK.

Sources: Own calculations based on DST

From figure 2.5, however, we see that, even though the earnings distribution for this group of students is more or less centered on the actual income limit that they faced after cancelling 1 month of benefits, a significant proportion of students deviate from this earnings level.[9]

Considering e.g. the upper part of the distribution, 14.9 percent of the students, who have cancelled exactly 1 month of benefits, earned more than 9,500 DKK above their actual income limit and thus hit the effective marginal tax rate of 100 percent. As a consequence these students could with relatively little effort have cancelled another month of benefits and thereby increased their disposable income. For the 6.3 percent, who had an excess income of more than 20,000 DKK, the increase in disposable income would have been at least 3,000 DKK ($\approx$ 500 USD).

Considering the lower part of the distribution we also see a significant proportion (70 percent) of students, who earned less than the actual income limit. In principle these students cancelled benefits without the need to do so and therefore received fewer benefits than they could have, however there might be intertemporal considerations that rationalize this behavior. As student benefits are limited to typically 6 years, student might find it optimal to save benefits for later use by cancelling some months even in years, where their earnings are below the income limit. In contrast to the upper part of the distribution, it is therefore less straight forward to take this as firm evidence of sub-optimal cancelling.

While the sub-optimal cancelling of benefits – as argued above – speaks against real adjustment costs as the dominant type of optimization frictions, it might be consistent with both gradual learning and inattention, cf. table 2.1. However a key difference between these two explanations is that under gradual learning we should expect to find sub-optimal cancelling of benefits primarily among new students.

In order to investigate this, I show in figure 2.6 the distribution from figure 2.5 split into 2 sub-samples of students, who have either be a student for 2 or more years or had a high income the year before – with the idea being that these two sub-samples should have better information about the structure of the student benefit system.

---

[9]When interpreting the distribution in figure 2.5 as a result of optimization frictions it is important to eliminate measurement errors from the data, as these will otherwise result in an upward bias of the amount of frictions. An assessment of the amount of measurement errors and the results robustness to these are presented in appendix 2.B and C.

Figure 2.6

The income distribution for tertiary students who cancel 1 month of student benefits split on student history
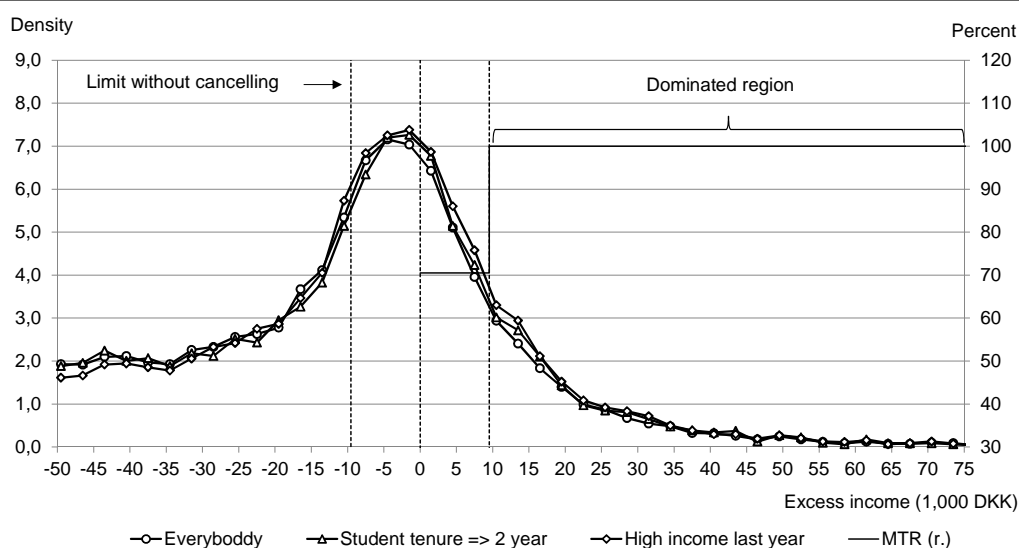
Notes: See notes to figure 2.5. Student tenure is measured from the start of the student's first tertiary education. High income last year is defined as having an income no less than 20,000 DKK below the baseline income limit.

Sources: Own calculations based on DST

As this figure shows, there is fundamentally no differences between the distributions, and the evidence does therefore not support that the sub-optimal cancelling is caused by gradual learning.

Above the level of optimization frictions is quantified by the share of students in the dominated region. However this metric is problematic as it depends crucially on the part of the sample that is included in the calculation. Considering e.g. the students, who do not cancel benefits, only 5.0 percent end up in the dominated region (compared to 14.9 percent above), but this is of course due to the inclusion of a large number of students, who are well below and not targeting the income limit.

Interpreting the frictions as earnings uncertainty and inattention a more natural way to quantify the level of frictions is to ask how much variance in their final earnings (relative to their desired earnings) individuals are will to accept and what the expected loss of disposable income from this variance amounts to.

One way to quantify this is to exploit that the dominated region bounds the range in which individuals rationally can set their desired earnings. For the students who

cancel exactly 1 month of benefits this range is limited to earnings between 86,000 and 95,500 DKK (excess income of 0-9,500 DKK in figure 2.5), and the shape of the earnings distribution outside this range is therefore informative about the size of earnings errors that the students make. Combining this information with the increase in disposable income that students could have gained by cancelling more or less student benefits, the costs of inattention for the students near the income limits can be estimated to 2-3,000 DKK.[10]

2.5 The nature of inattention

The graphical evidence in section 2.4 points to inattention about the earnings process during the year as the dominant optimization frictions in the labor market for Danish students. However, because of the time lag of 1-2 months between, when students have to decide whether or not to cancel benefits and when they have precise information about their current accumulated earnings, the sub-optimal cancelling we observe in figure 2.5 might simply reflect income surprises in the end of the year. In this case we should expect to find a positive correlation between positive individual income surprises and the amount of income exceeding their income limit.

In order to investigate this I use monthly income register data available from 2008 and define a proxy for an end of the year income surprise as the difference between the sum of November and December pay and the sum of the September and October pay. Plotting this measure against the individual excess income gives the picture presented in figure 2.7.

---

[10]For the exact calculations and description of the method see appendix 2.D.

Figure 2.7

Average end of year income surprise over the income distribution, 2008-11
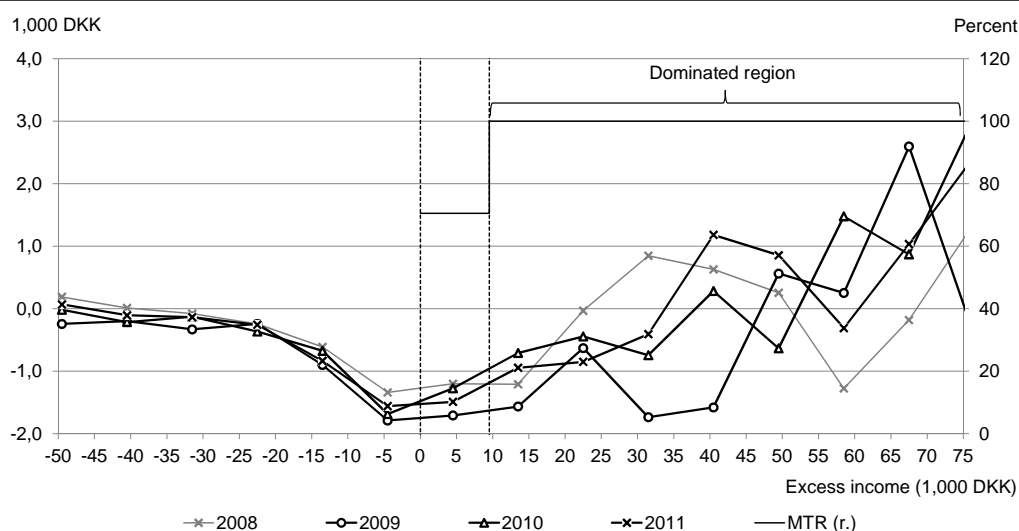
Notes: The figure only includes individuals who cancelled either 0 or 1 month if student benefits. The individual end of year income surprise is calculated as as the difference between the sum of November and December pay and the sum of the September and October pay. Only labor income is included in this data and months without employment are treated as 0 income. Bin size = 9,000 DKK.

Sources: Own calculations based on DST

From this figure it is clear that there is a tendency to find larger end of year income surprises among the individuals who end up with larger excess income. However the magnitude of the effect is not enough to explain the level of sub-optimal cancelling. Going e.g. from an excess income of 10,000 DKK to 50,000 DKK the average income surprise only increases by around 2,000 DKK, which therefore only explains 5 percent of the excess.

The figure, however, reveals another interesting feature from the monthly income data. It seems to be the case that students reduce their earnings when they approach the income limit. This behavior is more clearly visible when plotting the average end of year income surprise against the level of earnings that the students would have had without the income surprise – i.e. the yearly level of earnings if the November and December pay had equaled the earnings in September and October, cf. figure 2.8.

Figure 2.8

Average end of year income surprise over the predicted income distribution, 2008-11
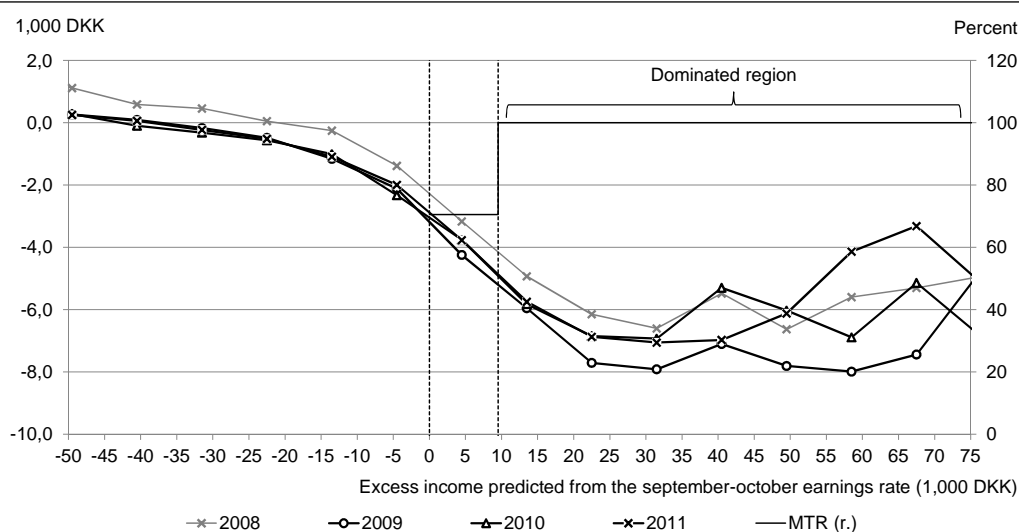
Notes: See notes to figure 2.7. Predicted excess income is the excess income that the individual would have had without the end of year income surprise – i.e. the actual earned income minus the difference between the sum of the November and December pay and the sum of the September and October pay. Bin size = 9,000 DKK.

Sources: Own calculations based on DST

From figure 2.8 we see a consistent drop in the average end of year income surprise of magnitude of 6-8,000 DKK for individuals, who at their September-October earnings rate were in risk of exceeding their income limit by the end of the year.

This drop could of course just be due to mean reversion following a positive income shock in September-October, but note that the drop is the same in the pre-reform year 2008 as in the post-reform years despite that the baseline income limit has been increased by 25 percent. That the drop occurs over the same range of excess income therefore reflects that the behavior has moved up in the earnings distribution.[11]

This type of behavior is not straight forward to reconcile with standard rationale inattention. Under risk neutrality standard rational inattention would suggest that individuals choose a job, which in expectation would give them their desired level of earnings. In the labor market considered here it appears that individuals take a job, which in

---

[11]Indeed, most of the shift in the distribution after the 2009 reform observed in figure 2.4 can be attributed to the drop in the November-December earnings first occurring at higher earnings levels after the reform.

expectation gives them a level of earnings above their desired level. Something that they first realize in the end of the year and instead of cancelling an extra month of student benefits – which would be a relative easy way to avoid the 100 percent effective marginal tax rate – they seek to reduce their labor supply and thus earnings.

One way to rationalize it is to assume that individuals are relatively risk adverse and thus take a job that with a high probability will give them their desired level of earnings, but once this level has been achieved they react to the lower earnings incentives created by the phase out of student benefits and reduce their labor supply. However, perhaps more realistically the inattention that individuals exhibit in this labor market is not fully rational.

2.6 Estimation of the labor supply response

After having shown in the sections above the likely presence of significant optimization frictions in the Danish student labor market, I proceed in this section with a discussion of how this is likely to affect the way labor supply elasticities are normally estimated.

Considering the labor supply responses observed in section 2.4 it clear that the two "standard" methods for estimating labor supply responses in public finance - the Saez (2010) bunching method and the Feldstein (1995) difference-in-difference (DiD) method - may fail to undercover the true elasticity.

When applying the bunching method researchers typically calculate the excess mass by fitting a high order polynomial to the distribution around the kink point excluding a range, where there is "visible bunching". However, in the student labor market considered here there is no visible bunching and a credible counterfactual distribution using this method in the purely cross sectional setting would therefore in practice follow the actual distribution yielding a zero excess mass and elasticity.

Likewise, when applying the DiD method, the labor supply elasticity is estimated by comparing individuals who are treated by (tax) reforms to different extent, where treatment statuses typically are assigned based on pre-reform earnings.[12] In the case considered here, this would imply that students with earnings between the pre-reform and the post-reform kink point would be assigned a lower marginal tax rate and the students above the-post reform kink point a slightly higher marginal tax rate. However, from figure 2.4 it is clear that the shift in the distribution happens over a much wider

---

[12]In practices the estimation procedure is more advanced using the treatment status based on pre-reform earnings as an instrument and controlling for underlying income dynamics such as mean reversion. See Weber (2014) for a recent discussion of the DiD method.

range than is directly affected by the changes in effective marginal tax rates and as a consequence the assigned treatment and control groups would consist of a mix of the true treatment and control groups.

To undercover a labor supply elasticity I therefore instead employ a method that resemble the method use by Chetty et al. (2013) and utilize the shift in the distribution created by the 2009 reform to undercover the (local) counterfactual distribution and hence the excess mass created by the pre- and post-reform kink.[13] Finally, I turn this excess mass into a labor supply elasticity using the Saez (2010) bunching formula.[14]

Figure 2.9 shows the average income distribution over the 3 pre- and post-reform years considered in this analysis, and illustrates the shift in the distribution after the reform also seen in figure 2.4. From this figure we can identify two areas with excess mass: Taking the post-reform distribution as a (local) counterfactual we find an excess mass 3.1 percentage points at the pre-reform kink point. Likewise, taking the pre-reform distribution as a counterfactual we find an excess mass of 2.1 percentage points at the post-reform kink point.

---

[13]The method resembles the method used by Chetty et. al. (2013) except that the source of the variation in the distribution here does not come from differences in knowledge about the tax schedule in a cross sectional setting, but from the time series variation created by a reform.

[14]One caveat has to be mentioned in connection with the translation of the excess mass into a labor supply elasticity. The formula derived by Saez (2010) rely theoretically on the marginal indifference individual, who bunch at the kink point, to change his earnings the same amount found when comparing two linear tax systems. In the presence of earnings uncertainty, where individuals not necessarily hit their desired income, this will no longer be the case and it is therefore not trivial that the formula is valid in this setting. Saez (1999) performs simulations of the income distribution and assess the amount of bunching under various model setups, incl. income uncertainty, but he does not evaluate the performance of the bunching estimate in these simulations. As a robustness check I therefore preform a more structure estimation of the labor supply elasticity in appendix 2.E, and yields almost the same elasticities as presented here.

Figure 2.9

Identifying excess mass using the 2009 reform
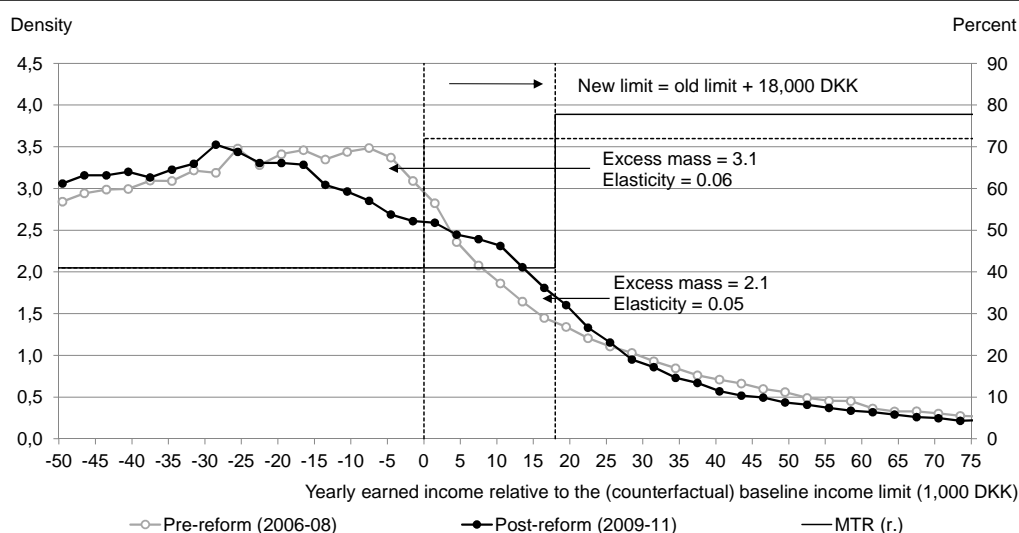
Notes: See notes to figure 2.4. For the calculations of the elasticities see table 2.3.

Sources: Own calculations based on DST

Using the Saez (2010) bunching formula, the change in earnings in responses to a tax change (dz) can be expressed as:

$$dz = \frac{B}{f(z)},\qquad(2.1)$$

where B is the excess mass and $f(z)$ is the counterfactual density at the kink point (z), and inserting this into the formula for the elasticity (ε) as:

$$\varepsilon = \frac{dz}{d(1-t)}\frac{1-t}{z} = \frac{B}{f(z)}\frac{dz}{d(1-t)}\qquad(2.2)$$

yields an elasticity of 0.06 for the pre-reform kink point and 0.05 for the post-reform kink point, cf. table 2.3.

Table 2.3

Calculating the labor supply elasticity for the tertiary students

	Pre-reform kink point	Post-reform kink point
Excess mass (B)	3.11	2.06
Counterfactual density (f(z))	0.87	0.47
Kink point (z)	76,4	97,7
dz/z = B / f(z) / z	0.05	0.05
d(1-t)/ (1-t)	0.74	0.98
Elasticity	0.06	0.05

Notes: The counterfactual density is estimated as the average density in the two bins around the
relevant kink point divided by the bin size (3,000 DKK).

Sources: Own calculations based on DST

This elasticity estimate is perhaps surprisingly small compared to the consensus in
the literature of around 0.25 according to Saez et. al. (2012) and considering that the
many students probably have a large degree of flexibility in increasing their earnings if
desired.[15] However there are a couple of reasons why the estimated elasticity is a lower
bound.

First of all taking the post-reform distribution as the (local) counterfactual for the
pre-reform distribution (and vice-versa) rely on the assumption the post-reform distri-
bution at the pre-reform kink point is unaffected by the post-reform kink point. This
would be true in a frictionless world, but with the fuzzy bunching created by optimiza-
tion frictions this will not necessarily hold.

Examining figure 2.9 it indeed seems to be the case that the excess mass around
the post-reform kink point start to build up already at the pre-reform kink point and
thereby biasing both the pre-reform and the post-reform excess mass downwards.

Secondly, as student benefits are limited to typically 6 years, some students might not
see it as a full loss to cancel benefits as assumed above. If students expect to use the
saved benefits later the real loss is only in terms of the difference in present value.

This implies that the phase out rate used so far - and hence the size of the kink -
is an upper bound of the actual phase out rate and thereby further implying that the
estimated elasticity is a lower bound. Assuming e.g. that 20 percent of the students in a
given year are indifferent between receiving benefits within the year or saving them for

---

[15]Working in the other direction is the fact that students might use a student job to gain valuable job expe-
rience, in which case the low intratemporal elasticity reflect future career concerns. However, dividing
student job into non-relevant jobs (retail, waitering and postal service) and relevant jobs (everything
else) does not give different elasticity estimates, which indicate that the future career concerns are not
the prime reason for the low estimates.

later imply that the average kink size will be 20 percent lower than the one used above. Scaling down $d(1-t)/(1-t)$ by this amount, increases the elasticities to 0.08 and 0.06, respectively.

Finally I return to the fact that the excess mass revealed by the shift in the earnings distribution following the 2009 reform is centered below and not on the kink point, as you would expect in a normal tax system under earnings uncertainty. However, as I show in appendix 2.E, this is fully consistent with the model under the institutional settings considered here. The reason is that, while earnings uncertainty in a normal tax system "smoothes" the jump in the marginal tax rate symmetrically around the kink point, this not the case, when students have the possibility to cancelling benefits. Without this possibility the jump in the marginal tax would be from 41 to 100 percent, and the smoothed effective margin tax rate faced by students follow the symmetric profile of this kink until the effective rate equals the phase out rate, where after it is caped. In this way the smoothed profile of the effective marginal tax is longer symmetric around the kink point, which causes the excess mass to be centered below the kink point.

2.7 Conclusion

In this chapter I have investigated the nature and impact of labor market optimization frictions among Danish students. This labor market represents an interesting case study as it features a number of special institutional settings, which allow you to distinguish between different types of optimizations frictions.

Examining labor market outcomes across these institutional settings I find clear evidence of a positive labor supply response following a reform in 2009 that substantially increased the earnings level at which phase out of student benefits begins. Yet, despite of this clear evidence of a positive labor supply elasticity, I find no visible bunching at the kink point created by the phase out in contrast to what standard theory suggest (Saez, 2010).

I take this as evidence of significant optimization frictions that mask the labor market outcomes suggested by standard theory – a finding that might be surprising given that student labor markets in general are associated with a lot of job turnover and part time workers and thus expected to have a high degree of flexibility. However this is not at odd, as a closer examination of the observed outcomes speaks against real adjustment costs or gradual learning about the institutional settings as the dominant optimization frictions. In particular because the positive labor supply responses after the 2009 reform materialize immediately. Instead, the evidence is consistent with inattention about the

individual earnings process during the year as the dominant frictions among the individuals in the considered labor market.

Of course, the relative strength of the different types of frictions might not be directly transferable to other labor markets and in particular you would probably expect real adjustment to play a bigger role in the regular labor market, where workers in general tend to be more specialized full time employees. However, the finding that inattention in itself can create large enough optimization errors to mask the bunching expected at kinks points in the tax schedule is interesting even for the broader labor market.

Following the investigation of the relative importance of the different optimization frictions I discuss the implications for identifying the underlying labor supply elasticity and propose a method that utilizes the shift in the earnings distribution created by the 2009 reform to uncover the local counterfactual distribution around the kink points created by the phase out of student benefits. Having this counterfactual distribution I use the Saez (2010) bunching formula and estimate a labor supply elasticities in with a lower bound in the range of 0.05-0.06.

This method is in many ways a compelling method for estimating labor supply elasticities, but at the same it time puts high requirements on the data being used. Indeed, as the presence of optimization frictions causes a mixing of treatment and control groups in the way they are typically assigned in the commonly used Feldstein (1995) difference-in-difference method, you are forced to rely more heavily on the time series variation and this is only credible if the earnings distribution is stable in the non-reform year. This is potential a problem in labor markets where real adjustments or gradual learning play more dominant roles, as this would cause the labor supply responses to be more gradual following a reform – a gradual response that often will be difficult for the researcher to credibly attribute to the reform.

References

Attanasio, O. P. (2000). "Consumer Durables and Inertial Behaviour: Estimation and Aggregation of (S,s) Rules for Automobile Purchases," Review of Economic Studies, vol. 67, 667-696.

Chetty, Raj, John N. Friedman, Tore Olsen and Luigi Pistaferri (2011)."Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." The Quarterly Journal of Economics, vol. 126(2): 749-804.

Chetty, Raj (2012). "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." Econometrica, 80(3): 969-1018.

Chetty, Raj, John N. Friedman and Emmanuel Saez (2013). "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." American Economic Review, vol. 103(7): 2683-2721.

Evans, George W. and Seppo Honkapohja (2001). "Learning and Expectations in Macroeconomics." Princeton University Press.

Gelber, Alexander M., Damon Jones and Daniel W. Sacks (2013). "Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test." NBER Working Papers 19491.

Gruber, Johathan and Emmanuel Saez (2002). "The Elasticity of Taxable Income: Evidence and Implications." Journal of Public Economics, vol. 84(1), 1-32.

Feldstein, Martin (1995). "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." Journal of Political Economy, vol. 103(3), 551-72.

Kleven, Henrik J. and Esben A. Schultz (2014). "Estimating Taxable Income Responses Using Danish Tax Reforms." American Economic Journal: Economic Policy, 271-301.

Kleven, Henrik J., Martin Knudsen, Claus Thustrup Kreiner, Soren Pedersen, and Emmanuel Saez (2011). "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." Econometrica, vol. 79(3), 651-692.

Kleven,   Henrik J. and Mazhar Waseem (2013). "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." The Quarterly Journal of Economics, vol. 128(2): 669-723.

Maire,   Daniel le and Bertel Schjerning (2013). "Tax Bunching, Income Shifting and Self-employment." Journal of Public Economics, vol. 107, 1-18.

Mankiw,   N. Gregory and Ricardo Reis (2002). "Sticky Information Versus Sticky Prices: A Proposal to Replace The New Keynesian Phillips Curve," The Quarterly Journal of Eco-nomics, vol. 117(4), 1295-1328.

Saez,   Emmanuel (1999)."Do Taxpayers Bunch at Kink Points?" NBER working paper 7366.

Saez,   Emmanuel (2010)."Do Taxpayers Bunch at Kink Points?" American Economic Jour-nal: Economic Policy, vol. 2(3), 180-212.

Saez,   Emmanuel, Joel Slemrod and Seth H. Giertz (2012). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." Journal of Economic Literature, vol. 50(1): 3-50.

Sims,   Christopher A. (2003). "Implications of Rational Inattention," Journal of Monetary Economics, vol. 50(3), 665-690.

Weber,   Caroline E. (2014)." Toward Obtaining a Consistent Estimate of the Elasticity of Taxable Income Using Difference-in-Differences" Journal of Public Economics, vol. 117, 90-103.

Appendix 2.A: The Danish student benefit system

2.A1 Student benefit rates

Danish students enrolled in most educations above the primary school (ISCED2011 level 3 and above) are eligible to state financed student benefits from the age of 18. Benefit rates vary depending on the type of education and civil status with the main rates (2008 level) listed in table 2.A1. Benefits for students aged 18-19 in lower educations (ISCED2011 level 3-4) furthermore depend on their parents' income.

Table 2.A1

Overview over basic student benefit rates, 2008

Monthly rate (DKK)	Baseline rate	Reduced with parents' income[1]	Minimum rate
Lower education and aged 18-19			
Living with parents	2,489	8.76 / 1,000 DKK	1,108
Not living with parents[2]	5,007	4.45 / 1,000 DKK	3,211
Tertiary education or lower education and aged 20+			
Living with parents	2,489	0 / 1,000 DKK	2,489
Not living with parents	5,007	0 / 1,000 DKK	5,007

Notes: Tertiary education include university education and educations such as nurses and school teachers (ISCED2011 level 5 and above). Lower educations include high school (gymnasium) and vocational educations (ISCED2011 level 3-4). 1) Benefits to student in lower education below age 20 depend on the parents' income in the way that the baseline rate is reduced by the listed amount for parent income exceeding 273.644 DKK until the minimum rate is reached. An extra allowance for the parents' income of 29.046 DKK is given for each sibling under the age of 18. 2) Students in lower education below age 20 have to apply for the higher benefits even if they are not living with their parents.

Sources: www.su.dk

On top of these basic rates it is possible to obtain a number of supplement payments summarized in table 2.A2.

Table 2.A2

Overview over supplement student benefit rates, 2008

	DKK per month
Supplement for single parents	5,007
Supplement if both parents are on student benefits	2,000
Disability supplement on tertiary educations	7,120
Supplement for tuition fees (maximum)	1,954

Notes: The 2008 special rates were no longer available online. The rates listed here are therefore based on the 2009 rates indexed back using the increase in the basic rates.

Sources: www.su.dk

The criteria for the different rates can be updated on a monthly basis and individual rates may therefore change during the year. This is likely to be a source of error in the prediction of final student benefits described in appendix 2.B given that the demographic information in the registers only is available on a yearly basis.

On top of these rates students have under some circumstances the possibility to "double clip", which means that the students receive a double benefit rate for that month. Prior to the 2009 reform this was possible in 3 situations:

1. During the last 12 month of the education if the student have cancelled student benefits in previous months.

2. The last month before paid internship (where it is not possible to get student benefits).

3. In connection with childbirth or adoption.

After the 2009 reform only the two last situations still apply.

In most educations student benefits are limited to the standard study time, except on university educations where student benefits are limited to 72 "clip" = 6 years, which is 1 year extra compared with the standard study time of most university educations.

2.A2 Student loans

While receiving student benefits students also have the possibility to take up a state administrated subsidized loan that payout 2,562 DKK per month (2008 level). The loan cannot be received if the student cancelled student benefits and student loans might therefore give an additional incentive not to do so. The loans are paid back after the student leaves the educational system according to a fixed schedule.

2.A3 Income control

When students receive student benefits they are subject to an income test. The test is au-tomatically done after the end of the income year by the student benefit adminis-tration, who draw the relevant information from the tax authorities income register of which most is 3rd party reported (see Kleven et al, 2011 for details). Based on this information the student benefit administration calculate a so-called "own income" (in Danish: egenindkomst), which consist of all income components except from the stu-dent benefits themselves, child benefits, employer administrated pension contributions and income taxed under the stock income tax scheme (dividends and capital gains).

The own income is compared to an individual income limit, which is generated as the sum of monthly amounts depending on the student's actions:

- In months where the student is eligible and receives student benefits a "low amount" of 6.370 DKK is added to the income limit.

- In months where the student is eligible, but does not receive benefits (the student has cancelled benefits) a "medium amount" of 15.908 DKK is added.

- In months where the student is ineligible for student benefits a "high amount" of 30.619 DKK is added.

On top of these amounts the income limit for parents is further increased by a yearly amount of 23.008 DKK per child below 18.

As the analysis in the chapter only focuses on students who are fully eligible for student benefits the entire year the key variation in the individual income limits comes from the students' cancelling of benefits, which moves them from the low to the medium amount and thereby increase their income limit by 15,908 − 6,370 = 9,538 DKK per month cancelled relative to a baseline amount of 12 • 6,370 = 76,440 DKK per year.

If a students' own income exceeds his/her income limit the excess has to be paid back to the student benefit administration according to the following formula: of the first 9,538 DKK (= Medium − Low amount[16]) 50 percent has to be paid back, while further excess income is paid back 100 percent. Finally if the amount that is to be paid back exceeds 7,569 DKK (= basic student benefit rate for student not living with their parents + student loan payout) the payback is further increase by 7 percent. In the register the payback – except the 7 percent increase – is treated as a reduction in the received student benefits.

---

[16]After the 2009 reform the low amount for the lower education is used for the tertiary educations.

Appendix 2.B: Calculating own income and determining eligibility

The data are constructed by drawing from a number of register data sets organized by Statistics Denmark (DST). In particular income data from the tax return (INDH), education information (UDDA) and weekly information about recipient status for public transfers (DREAM) along with standard demographic information (BEF). Finally, the individual monthly earnings are draw from the E-income register (BFL) available from 2008 and onwards. All of these registers contain the entire Danish population and can be linked using a unique identification number.

2.B1 Student benefits and income limit

Eligibility for and payout of student benefits are determined from the DREAM data set, where the first challenge is to aggregate the weekly information into monthly information (the interval at which student benefits are paid out). This is done by first allocating weeks to months based on the position of Wednesday and then counting the number of weeks where student benefits have been paid out (code 651) and the number of weeks where an individual has been eligible for student benefits without receiving them (code 652).

In a month with 4 weeks, 3 or more weeks with pay outs are coded as a month were the individual has received student benefits. Similarly 3 or more weeks with eligibility for student benefits without receiving them is coded as an eligible month (the individual has cancelled student benefits). In months with 5 weeks the number of weeks has to be 4 or more.

These numbers are coupled with the educational and demographic registers to determine the benefit rate the each individual is eligible for and the income limit that the individual faces. The key variables here are the level of the current ongoing education (UDD) and the civil status (FM_mark), which can be used to determine whether individuals are not living with their parents (code 6).

Finally, the number of children, which affects both the income limit and the benefit rate is calculated from the number of children below 18 in the household (variable PLADS, code 3) for the individuals who are not them self a child in a household (individuals not living with their parents).

With the above variables the individual income limit is calculated as:

$$IncomeLimit_i = Amount_{Low} \cdot No_{Ri} + Amount_{Medium} \cdot No_{Ei} + Amount_{Child} \cdot No_{Ci}, \quad (2.3)$$

where No_{Ri} is the number of months, where the individual receives student benefits. No_{Ei} is the number of months where the individual is eligible for student benefits without receiving them (student benefits have been cancelled), and No_{Ci} is the number of children below 18 years. The amounts are the corresponding contributions to the income limit described in appendix 2.A.

2.B2 Own income

When it comes determining "own income", the income registers unfortunately do not contain the own income variable constructed by the student benefit administration and this variable therefore has to be constructed. A challenge in this respect is that the registers only contain pre-aggregated income variables and not the full set of information available on the tax return and it is therefore not possible simply to apply the code used by the student benefit administration. Instead the own income variable is constructed by adding together labor income, capital income (earned interests) and transfers other than student benefits (excluding child related transfers) defined from the variables listed in table 2.B1.

Table 2.B1

Variables used in the constructed of own income

Variables that is always included	
Labor income excl. labor market contribution	LOENMV – SLUTBID
Capital income (earned interest)	RENTEINDK
Other transfers	QMIDYD
Additional variables that sometimes is included	
Business income	NETOVSKYD
Capital income from investment funds	PEROEVRIGFORMUE – AKTIEINDK – SKATFRIYD
Other types of income (scholarships etc.)	RESUINK_GL

Notes: A more detail description of the variables (in Danish) can be found at www.dst.dk/times.

These 3 income components, however, do not fully cover the income that is included in the student benefit administrations definition of own income. In particular, business income among self-employed students, capital income from investment funds[17] and other types of income such and certain types of scholarships are included in the student benefit administrations definition but not in the three main components included here.

The additional income components can in principle be found in the register data from the variables listed in table 2.B1, but these variables do not precisely correspond to the

---

[17]But not direct dividend payments and capital gain taxes under the stock income scheme.

variables that the student benefit administration uses – either because they are calculated net of certain deductions (NETOVSKUD) or because they include additional income components. A general inclusion of these variables therefore adds as much error to the own income variable as leaving them out. Instead I apply the following strategy for determining the individual own income.

First I calculate each individual's own income based on the 3 main income components and the individual income limit based on the number of months of student benefits and the level of his current study and the number of children. The difference between the own income and the income limit identify the excess income that is to be deducted according to the formula described in appendix 2.A in the benefits that the student benefit administrations initially have paid out.

Second I identify the actual deduction based on the difference between the student benefits that initially have been paid out and the final level of student benefits registered in the tax returns (variable: STIP). For the individuals with positive deductions I can uniquely identify the excess income that would correspond to the observed deduction.

Finally, if difference between the excess income calculated in step 1 and the excess income calculated in step 2 exactly (+/- 2 DKK) corresponds to a combination of the 3 additional income components listed in table 2.B1, I add these income components to the own income variable for that individual.

Of course this procedure is potential problematic as it only add to the precision of the variable for the individuals who exceeds the income limit and because the procedure risk adding wrong income components that simply by chance matches the difference between the excess income calculated in step 1 and step 2, while the error might come from errors in the applied benefit rates.

However, given that the additional income components have to exactly match the differ-ences in own income it seems safe to assume that risk of addition wrong components is minimal and given that the amount of frictions in section 2.4 is identified from the individuals exceeding the income limit, I choose to do this adjustment to the own income definition. Over the 6 years 2006-2011 the adjustment is applied to 32,000 individuals or 5 percent of the student sample in tertiary educations.

2.B3 Assessing the accuracy of the own income variable and income limit

With the above construction of the own income variable it is important to assess the accuracy of the variables – especially because measurement error in the outcome variable

will create an upward bias in the estimations of optimization frictions.[18]

In order to do this, I calculate each individual's predicted student benefits based on the number of month the individual have received student benefits during the year, their income limit and their own income. If the predicted student benefits lies within +/- 10 DKK of the actual student benefits I define it as a "hit".

There is however two problems with this way of assessing the accuracy of the own income and income limit. First of all a hit also depend on an accurate modelling of the student benefit rates – a potentially large source of error given the number of rates described in appendix 2.A, but this type of error of less importance for the analysis of labor supply responses in the chapter. Secondly – and more problematic – (small) errors in the own income variable only affected the predicted student benefits, if the own income excess the income limit. The assessment of the accuracy of the own income variable is therefore only precise above the income limit.

Table 2.B2 summarizes the proportions of hits (the hit rate) for different parts of the sample. At an aggregate level the procedure accurately predicts the student benefits for 2/3 of the sample with better hit rate for the tertiary students (80 percent hit rate) than for the students in lower educations (40 percent hit rate), which is probably due to larger variety in the benefits rates for students in lower educations.

Table 2.B2

Assessing the accuracy of the own income definition

Hit rate (percent)	2007	2007	2008	2009	2010	2011
Aggregate his rate	66.4	66.0	65.1	65.6	67.8	67.9
- Lower educations	40.7	42.2	41.1	35.1	42.1	42.2
- Tertiary education	81.6	80.4	80.1	83.3	84.7	84.4
Among the tertiary students						
- below the income limit	85.1	84.2	83.4	85.1	86.1	85.6
- above the income limit	58.6	58.1	61.3	59.9	65.1	65.1

Notes: A "hit" of the own income definition is definitions as a predicted student benefits within +/- 10 DKK of the actual final student benefits received. 99.9 percent of the hits are with +/- 1 DKK, which can be attributed to rounding errors. Tertiary education include university edu-cation and educations such as nurses and school teachers (ISCED2011 level 5 and above). Lower educations include high school (gymnasium) and vocational educations (ISCED2011 level 3-4).

Sources: Own calculations based on DST.

Among the tertiary students the hit rate is naturally higher for the student below the

---

[18]This is in contrast to "regular" regression analysis, where measurement error in the outcome variable only will lead to higher standard errors.

in-come limit, where the marginal errors in the own income does not affected the predicted student benefits. Some of these errors can be attributed to errors in the applied student benefit rate due to e.g. student moving from their parents during the year, child birth and "double clipping" prior to 2008, however trying to control for these types of errors does not significantly improve the hit rate – especially for the individuals above the income limit.

As a consequence of this potential measurement error in either the own income and/or the income limit I conduct a robustness test in appendix 2.C by replicating the key graphs in the chapter only with the part of the sample, where I can accurately predict the final student benefits. As the appendix shows this sample restriction does not affect the conclusions significantly.

2.B4 The monthly income data (E-income)

The monthly income data is collected from the E-income statistics from 2008, which is collected by the Danish tax authorities. It is mandatory for all firms to report their wage payments to this register.

From this statistics I draw the variable AJO_SMALT_LOENBEGREB, which corresponds to the labor income variable used in table 2.B1 gross of labor market contribution. As the labor market contribution is 8 percent the variable is made net by multiplying by 0.92. With this correction the yearly income in the E-income statistics almost exactly matches the labor income in the yearly income register. Put into numbers, a regression of labor income on yearly E-income yields a parameter estimate of 0.997 with a R^2 of 0.989.

2.B5 Sample size

With the data drawn from the registers I get the breakdown of the size of the Danish student population shown in table 2.B3. The core sample consists of students, who are fully eligible for student benefits and employed. They numbers around 85,000 per year.

Table 2.B3

The size of the Danish student population

1,000 persons	2006	2007	2008	2009	2010	2011
Everybody aged 18-30	813.5	820.3	833.4	841.1	855.4	869.2
In education	327.5	333.1	334.4	338.6	354.4	379.3
Of these:						
- Lower education	183.6	188.5	190.4	194.2	202.1	214.2
- Tertiary education	143.9	144.6	144.1	144.4	152.3	165.1
Among the tertiary students						
- Fully eligible	97.6	97.8	97.7	98.2	105.8	115.9
-- Employed[1] (core sample)	85.0	86.2	86.1	84.4	88.8	95.6
--- Also the year after	52.3	52.9	53.7	53.0	55.4	

Notes: 1) Employed is defined as having a positive labor income.

Sources: Own calculations based on DST

## Appendix 2.C: Robustness check wrt. measurement error

As shown in appendix 2.B it is not possible to precisely predict the student benefits received for the entire sample of students. In the case these errors are a result of errors in the coding of the benefit rates it will not affect the analyses conducted in the chapter, however if the errors stems from errors in the coding of the individual income limits or individual own income it poses a threat, as these measurement errors will make some individuals behavior appear sub-optimal.

As a robustness check to the analyses in the chapter I therefore repeat the key figures in the chapter (figure 2.3-2.5) using only the part of the sample, where I can actually predict their final student benefits.

Figure 2.C1 corresponds to figure 2.3 in the chapter and shows the same general patterns as the original figure, except from a slightly steeper drop in the density at the kink point. However this steeper drop is partly mechanical, as the predicted student benefits only depend on marginal changes in the own income and the individual income limit above the baseline income limit. Small errors in these components will therefore only lead to an exclusion from the sample above this limit and thereby create the steeper drop.

Figure 2.C1

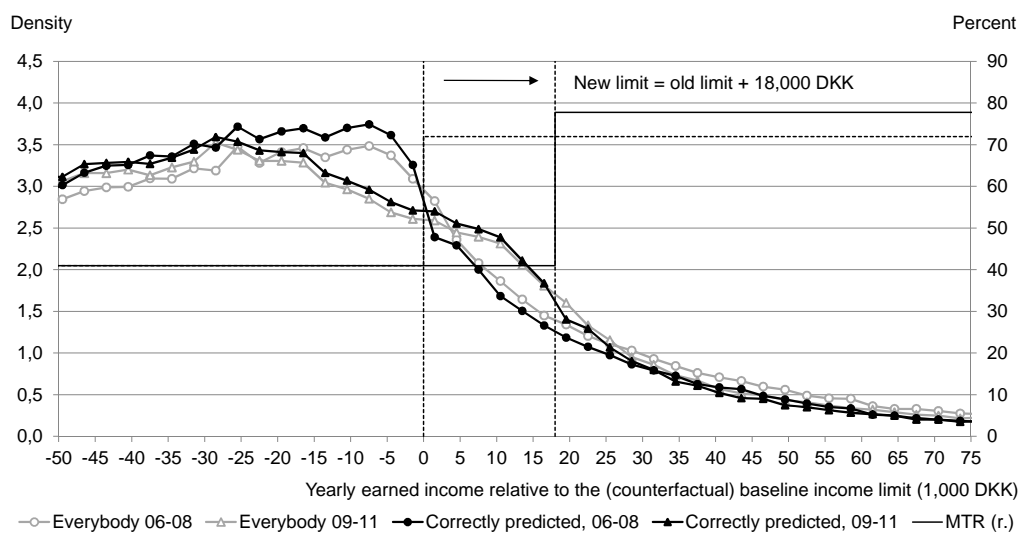The income distribution for tertiary students, 2006-2008
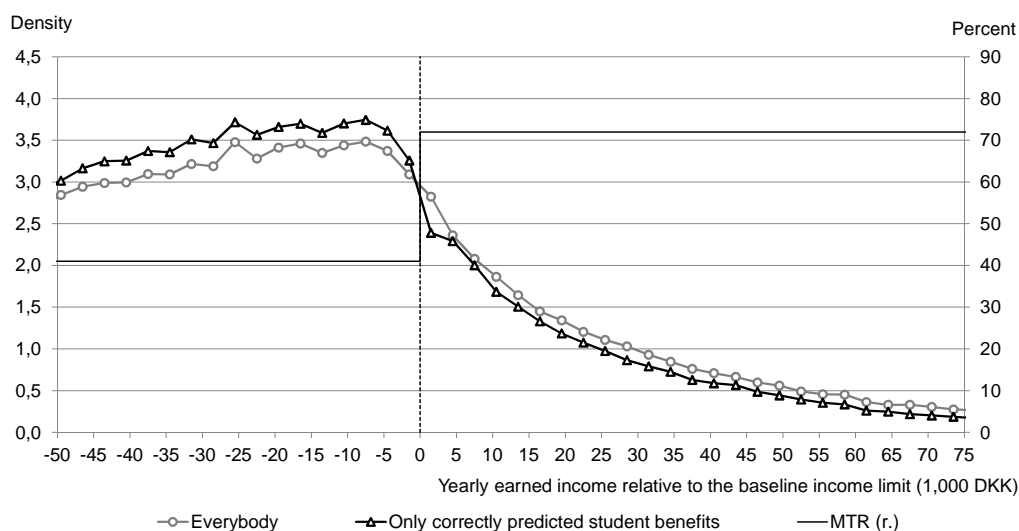
Notes: See notes to figure 2.3. The line for "Everybody" corresponds to average over the years in figure 2.3. "Only correctly predicted student benefits" only includes individuals with predicted student benefits with +/- 10 DKK of the actual student benefits received.

Sources: Own calculations based on DST

Similar the exclusion of the individuals, where I cannot accurately predict student benefits, does not significantly affect the conclusions drawn from the other key figures, cf. figure 2.C2 and figure 2.C3.

Figure 2.C2

The income distribution for tertiary students before and after the 2009 reform
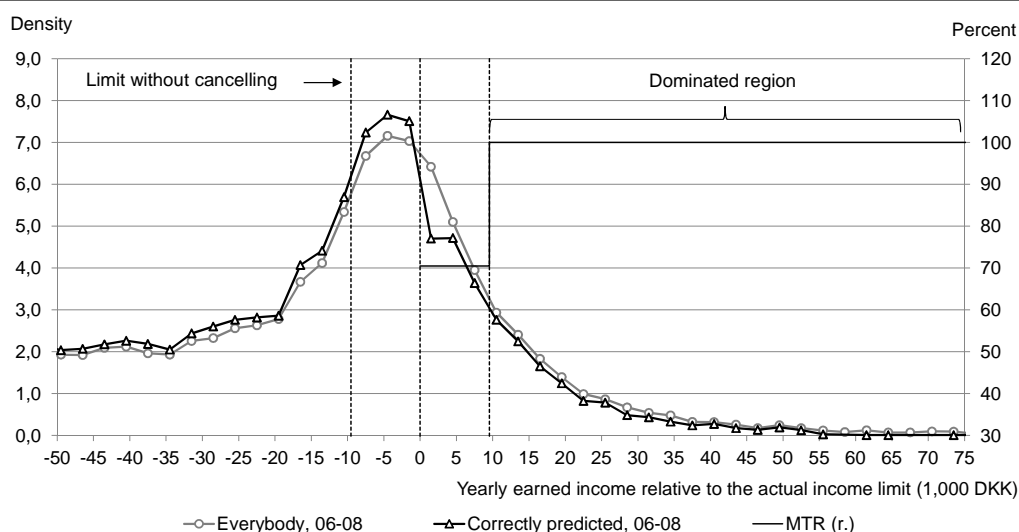

Notes: See notes to figure 2.4 and figure 2.C1.

Sources: Own calculations based on DST

Figure 2.C3

The income distribution for tertiary students who cancel 1 month of student benefits

Notes: See notes to figure 2.5 and figure 2.C1. For the correctly predicted sample the mass in the domi-nated region is 12 percent.

Sources: Own calculations based on DST

Appendix 2.D: Deriving the costs of inattention

This appendix describes how I use the shape of the mass found in dominated regions to quantify the level of optimizations frictions, as mentioned in section 2.4 in the chapter. More precisely this appendix provides estimates of the variance of the earnings errors (caused by inattention) and the expected cost for individuals associated with these errors.

In a normal setting it is not possible estimate the variance of earnings errors as it is not possible to split an observed individual earnings level into the earnings that the individual targeted and a earnings error. This is illustrated in figure 2.D1, which shows a simulated earnings distributions, where individuals target earnings are uniformly distributed from 10 to 20, while realized earnings is given by this target plus and a normally distributed error. Considering e.g. individuals in this setting with an observed earnings level of 16, these individuals include both individuals, who targeted this earnings level, as well as individuals who targeted other earnings levels but ended up for deviating from this target.

Figure 2.D1

Illustration of the identification of earnings errors
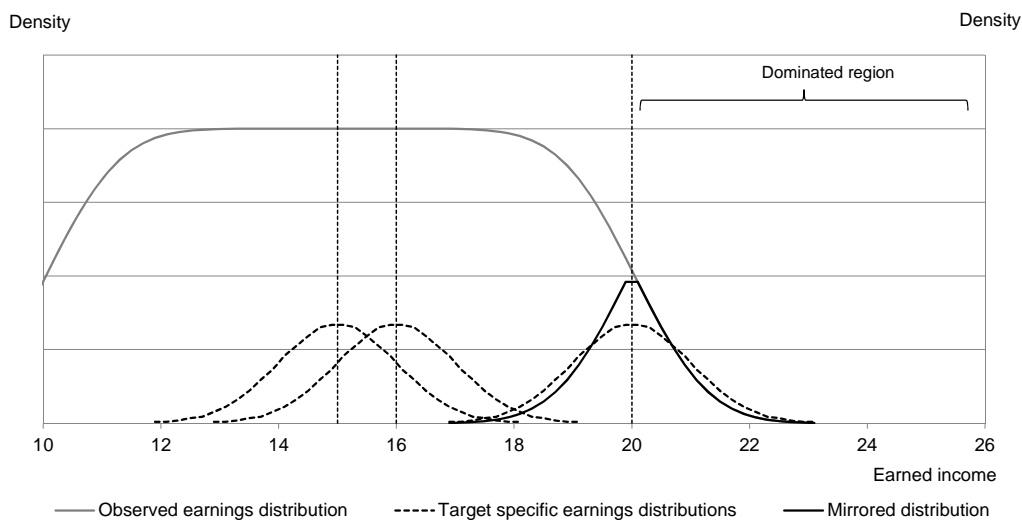
Notes: The observed earnings distribution shows a simulated earnings distributions, where individuals target earnings are uniformly distributed from 10 to 20, while realized earnings is given by this target plus and a normally distributed error. The mirrored distribution shows the mirror of the observed distribution around the mirror point 20. The target specific earnings distributions shows the distribution of earnings errors for a given earnings target. The mass of these distributions have been scaled to equal the mass under the mirrored distribution.

In contract the presence of dominated regions enables you to put bounds on the earning levels that individuals target. In figure 2.D1 this is illustrated with a dominated region from earnings 20 and above, and as a consequence all observed earnings above 20 must be due to earnings errors among individuals with earnings targets below 20. A lower bound on each individuals earnings error is therefore their observed earnings minus 20. This is a lower bound as some individuals might have target earnings below this level.

Further assuming symmetry of the errors distribution, you can mirror the observed earnings distribution in the dominated region to get an estimate of total error distribution and from there calculate measures such as e.g. a standard error. Doing this for the mirrored distribution in figure 2.D1 yields a standard error of 2/3 compared to an actual standard error of 1, which precisely indicate the lower bound nature of the method in this setting.

Turning to the actual earnings distribution for the students who cancelled exactly 1 month of benefits (shown in figure 2.5 in the chapter) I benefit from the fact that

the range of earnings in which it is optimal to cancel this amount of student benefits is relative narrow and – as a consequence – the room for error when assign a target earnings level to individual is reduced.

Figure 2.D2

Calculation of the costs of inattention for tertiary students who cancel 1 month of student benefits

---
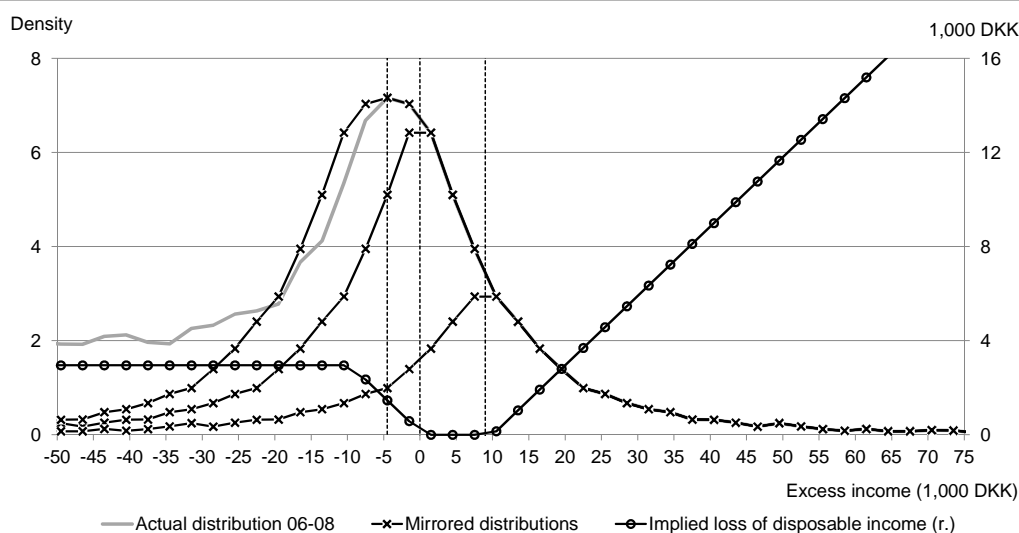
Notes: The actual distribution is the average density for the years 2006-08 also shown in figure 2.5 in the chapter. The mirrored distributions shows the actual distribution mirrored around 3 different mirror point (-4,500, 0 and 9,000 respectively). The implied loss of disposable income shows the maximum increase in disposable income that a student with a given excess income could have obtained by cancelling more or less student benefits.

Sources: Statistics Denmark and own calculations

Still, in the figure 2.D2 I consider 3 different mirrorings of the earnings distribution: 1) the actual start of the dominated region (excess income = 9,000 DKK), 2) 0 excess income and 3) the mode of the distribution (excess income = -4,500 DKK). These mirrored distributions yield a standard error of 20-25,000 DKK, cf. table 2.D1.[19]

---

[19]These standard errors are relatively large, which reflect that the distributions have relatively fat tails. If I instead calculate the cut-offs levels for the 95% confidence intervals these the absolute distances to the mirror point becomes 28,000 for the 0 mirror point and 10,000 for the 9,000 mirror point.

Table 2.D1

Quantifying the costs of inattention, 2006-08

Benefits cancelled:	0 months			1 month		
Mirror point:	>-4.500	> 0	> 9,000	>-4.500	> 0	> 9,000
Mass[1]	14.5	11.1	5.0	37.4	30.3	14.9
Standard error[2]	19.1	20.5	25.0	21.6	22.4	26.5
Expected costs[3]	0.7	1.1	2.4	2.3	2.2	3.0

Notes: Excess income = earned income – income limit. 1) Share of the sample with earnings above the mirror point (percent). 2) Standard error calculated from the mirrored distribution (1,000 DKK). 3) The expected cost is calculated by computing the increase in disposable income from optimal cancelling of student benefits for each level of excess income and integrating over these amounts using the mirrored distributions (1,000 DKK).

Sources: Own calculations based on DST

Figure 2.D2 also shows the maximum increase in disposable income that a student with a given excess income could have obtained by cancelling more or less student benefits. At negative excess incomes this increase comes from the fact the a student could have obtained the same income limit without having cancelled student benefits, while the increase at positive excess incomes comes from the fact that students could have avoided the 100 percent marginal tax rate by cancelling additional months of student benefits. Integrating over this loss function with the densities from the mirrored distribution gives an expected cost of the earnings errors – which can be interpreted as a result of inattention – of around 2-3,000 DKK.

Replicating the same calculations for the students who do not cancel student benefits yields standard error estimates of the same size as for the students who cancel 1 month, while the estimated expected costs are lower for the lower mirror points, cf. table 2.D1. The lower expected costs reflect that there are no costs associated with negative earning errors for the students who do not cancel benefits, as their already receive the maximum amount.

Appendix 2.E: GMM estimation of the labor supply of students

As a supplement to the non-parametric estimates of the labor supply elasticity in the chapter I present in this appendix a more structural approach that jointly identifies the labor supply elasticity and the amount of variance in their final earnings relative to

their desired earnings students are willing to accept.[20]

The idea behind the structural approach is to formula a model of labor supply under earnings uncertainty and the Danish student benefit system, simulate the effect of a reform similar to the 2009 reform descripted in the chapter and estimate the two parameters by minimizing the squared difference between the simulated changes in the earnings distribution and the observed change in the distribution shown in figure 2.9. In this way the approach falls into the frame of GMM (Generalized method of moments) estimation.

2.E1 The model

Following the norm in most recent empirical papers in public finance I start with a simple quasi-linear utility function (see e.g. Saez et. al., 2012):

$$u_i = c_i - \frac{\mu n_i}{1+\mu} \left(\frac{\hat{z}_i}{n_i} \right)^{\frac{1+\mu}{\mu}} \tag{2.4}$$

where c is private consumption and $\hat{z}$ is the income level that the individuals target. μ and n is parameters of the utility function that can be interpreted as the labor supply elasticity and potential earnings, respectively. Final earnings (z) is stochastic and given by:

$$z_i = \hat{z}_i + \varepsilon_i \tag{2.5}$$

where ε_i is an iid. error term.

The budget constraint that the students are facing can be written as follows:

$$c_i = (1-t)\left[SB - q(T_i - L) \cdot 1(T_i > L) + z_i \cdot 1(T_i \geq z) + T_i \cdot 1(T_i < z)\right] \tag{2.6}$$

This equation states that if students raise their announced income target (T) above the baseline income limit (L) the baseline student benefits (SB) is phase out at a rate q. Next, given the announced income target the students are allowed to keep any income below this target, while any excess income is taxed at 100 percent. The announced income target thus effectively constitutes an income ceiling for the student. Finally, both student benefits and earned income is subject to the ordinary tax system, which here is summarized by the (marginal) tax rate t.

In order to simplify the optimization I assume that the students are risk neutral and

---

[20]I do not model inattention endogenously, but simply assume that individuals cannot observe/reoptimize their earnings during the year. In this way the estimated end-of-year earnings variation should be interpreted as the underlying earnings variance net of reoptimization during the year.

that ε_i is normal $N(0, \sigma)$ distributed. In this setting maximizing expected utility only depends on income through expected consumption, which given equation 2.6 can be written as:

$$E(c_i) = (1-t) \cdot$$

$$\left[SB - q(T_i - L) \cdot 1(T_i > L) + \left(\hat{z}_i - \frac{\sigma f(\theta_i)}{F(\theta_i)} \right) \cdot F(\theta_i) + T_i \cdot (1 - F(\theta_i)) \right] \quad (2.7)$$

where $\theta_i = \frac{T_i - \hat{z}_i}{\sigma}$. Optimal behavior implies the follows two first order conditions for T and $\hat{z}$ respectively:

$$\frac{\partial E(u_i)}{\partial T_i} = 0 \Leftrightarrow \frac{\partial E(c_i)}{\partial T_i} = 0 \Leftrightarrow 1 - F\left(\frac{T_i - \hat{z}_i}{\sigma} \right) = q, \, for \, T_i > L \quad (2.8)$$

$$\frac{\partial E(u_i)}{\partial \hat{z}_i} = 0 \Leftrightarrow \frac{\partial E(c_i)}{\partial T_i} = \left(\frac{\hat{z}_i}{n_i} \right)^{\frac{1}{\mu}} \Leftrightarrow \hat{z}_i = \left((1-t)F\left(\frac{T_i - \hat{z}_i}{\sigma} \right) \right)^{\mu} n_i \quad (2.9)$$

Both conditions have a straightforward economics interpretation. When it comes to raising the announced income target students have to balance the decrease in the probability that their marginal income will hit the income ceiling with the phase out of student grant. Because I have assumed risk neutrality this probability has to exactly equal the phase out rate. Second, given the announced income target the students choose a target income (labor supply) as a function of not only the standard tax rate (t) but also the implicit tax rate created by the risk of hitting the income ceiling. The strength of the responses to the effective marginal tax rate depend on labor supply elasticity (μ). Finally, note that the students in the absence of taxes and phase out of student benefits in this model will target an earnings of n_i, which therefore can be interpreted as potential (expected) earnings.
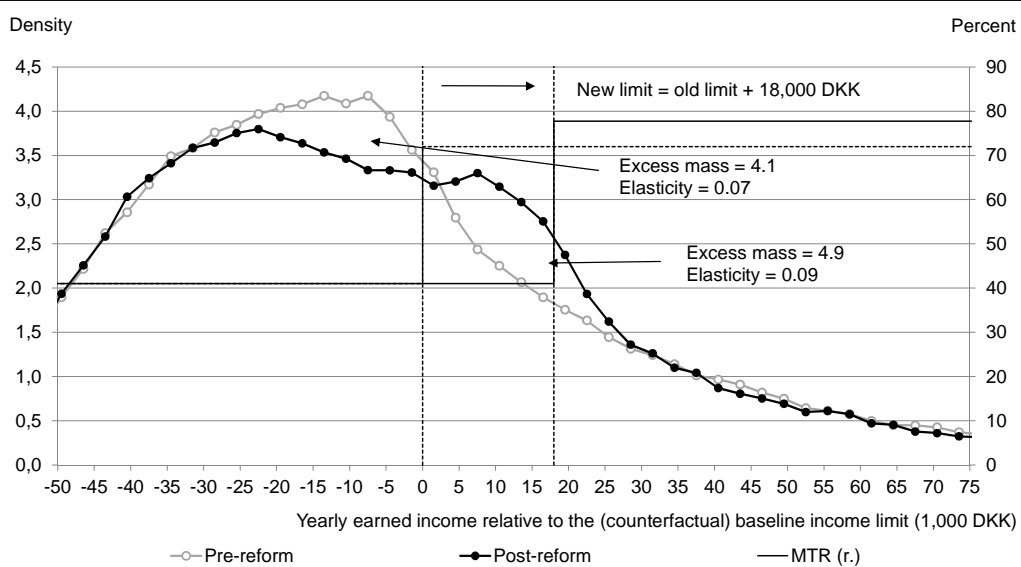
2.E2 Simulation

Before moving into the actual estimation, I present the performance of the model based on a simulation with fixed parameter values. The simulation is done by solving the model for a large number of individuals with different drawn of the distribution of potential earnings and with different realizations of the stochastic component of income (ε). More concretely I draw log potential earnings (measured in 1,000 DKK) form a normal distribution with mean 4.3 and standard error 0.5 and set the labor supply elasticity (μ) to 0.1 and the standard error of the stochastic component of earnings (σ) to 7.

In this setting I implement both the pre-reform policy setting ($L = 76.4, q = 0.525$) and the post-reform setting ($L = 94.4, q = 0.623$). The tax rate (t) is in both cases set to

0.41.

The resulting earnings distributions are show in figure 2.E1, which shows the same shift in mass from below the pre-reform kink point to a range above as in figure 2.4 in chapter. The figure also reports elasticity estimated using the same non-parametric method as in section 2.6. The method is able to recover the true elasticity with a small downwards bias, which stems from the fact that the post-reform distribution that is used as the local counterfactual distribution at the pre-reform kink point, is affected by the post-reform kink due to the optimization frictions as also discussed in the section 2.6.

Figure 2.E1

Simulated earnings distribution before and after the 2009 reform

Notes: Simulated distribution of realized earnings based on a draw of 100,000 individuals with log earnings normally distributed with mean 4.3 and standard error 0.5. $\mu = 0.1$ and $\sigma = 7$.

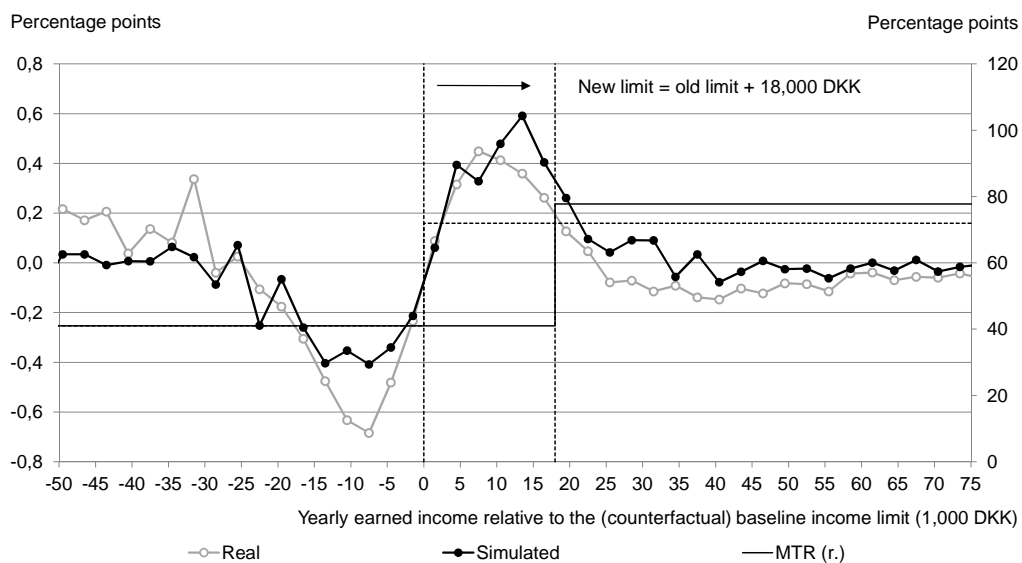Sources: Own calculations

2.E3 GMM estimation

From the simulation above I can calculate a change in the frequency in each bin and map this to the actually changes seen in figure 2.4 and from there, choose the parameter values of μ and σ that minimizes the sum of squared errors between the actual and simulated data. This procedure yields an estimate of the labor supply elasticity of 0.06

and standard error of ε of 6,000 DKK.[21] The estimated labor supply elasticity is in other word more or less the same as the non-parametric estimate in the chapter, while the standard error of individuals' final earnings is significant smaller. Given these parameter estimates I obtain a simulated change in the earnings distributions compared to the actual change as shown in figure 2.E2.

Figure 2.E2

Simulated and actual change in the earnings distribution following the 2009 reform

Notes: Simulated distribution of realized earnings based on a draw of 100,000 individuals with log earnings normally distributed with mean 4.3 and standard error 0.5. $\mu = 0.05$ and $\sigma = 6.5$.

Sources: Own calculations

2.E4 The position of the excess mass

As mentioned in the chapter it might appear strange that the excess mass uncovered by the shift in the earnings distribution following the 2009 reform appeared significantly below the kink point and not centered on the kink point as you would except. However as already seen in figure 2.E1 this is a consistent feature of the model, where individuals can cancel benefits in order to avoid the 100 percent marginal tax rate.

---

[21]The estimation is done as a grid search going from $\mu=0.01$ to $\mu=0.20$ in steps of 0.01 and from $\sigma = 1$ to $\sigma = 20$ in steps of 0.5. If the objective function is defined as the change in the distribution relative to the pre-reform distribution I obtain $\mu =0.09$ and $\sigma = 8,000$ DKK.

The reason behind this non-centered excess mass in the case with a possibility to cancel benefits comes from the effect that this possibility has on the effective marginal tax rate. In the standard setting without earnings uncertainty this is simply equal to the statutory marginal tax rate and a kink in the tax schedule thus creates a discrete jump in the marginal tax rate. Adding earnings uncertainty to this setting smoothes the jump, so that the effective marginal tax rate increases "symmetrically" from the low tax rate to the high tax rate around the kink point.[22]
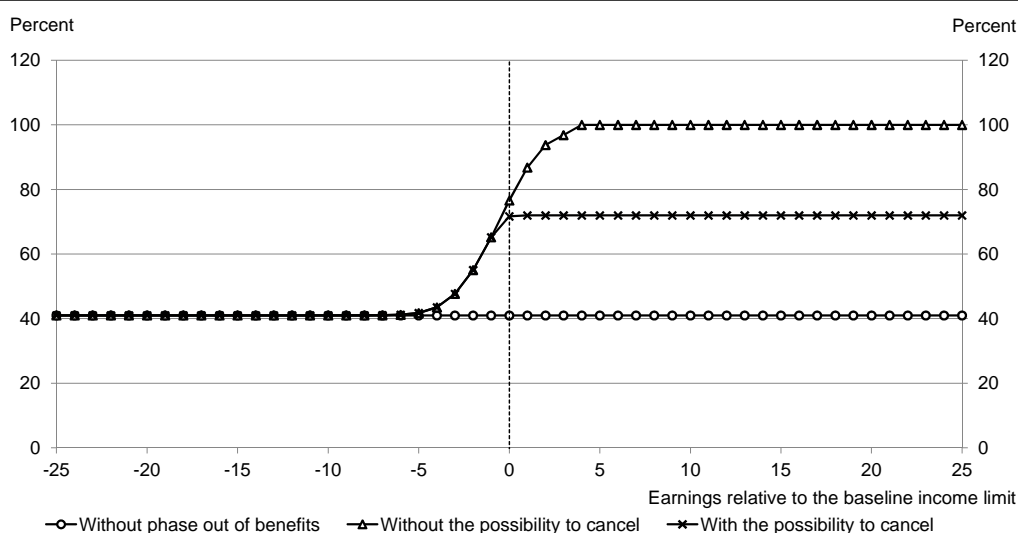
Without the possibility to cancel benefits the kink point faced by students is effectively a jump from 41 to 100 percent marginal tax rate, and so with earnings uncertainty the effective marginal tax rate increases smoothly between these 2 rates symmetrically around the kink point, cf. figure 2.E3.

With the possibility to cancel benefits students can effectively move up the kink point by phasing out benefits, and from equation 2.8 we see that they will do this until the probability of hitting the 100 percent tax rate is equal to the phase out rate. As a consequence the effective marginal tax rate profile will follow the profile without the possibility to cancel benefits until it equals the phase out rate, where after it becomes caped (in the present case at 72 percent), cf. figure E3. As a result the smoothed increase in the effective marginal tax will no longer be symmetric around the kink point.

---

[22]The symmetry comes from the symmetry of the distribution of earnings errors. If this distribution is not symmetric the change in the marginal tax rate will neither be symmetric.

Figure 2.E3

Effective marginal tax rates with and without the possibility to cancel benefits

Notes: The effective marginal tax rate is calculated as $1 - (1-t)F\left(\frac{T_i - z_i}{\sigma}\right)$, where T_i is set so $1 - F\left(\frac{T_i - z_i}{\sigma}\right) = q$, for $T_i > L$. (equations 2.8 and 2.9 above). In the case without phase out of benefits $q = 0 \Rightarrow T_i \approx \infty \Rightarrow F\left(\frac{T_i - z_i}{\sigma}\right) = 0$. In the case without the possibility to cancel benefits $q = 1 \Rightarrow T_i \approx L \Rightarrow F\left(\frac{T_i - z_i}{\sigma}\right) = F\left(\frac{L - z_i}{\sigma}\right)$. Simulations are done with $\sigma = 2$.

Sources: Own calculations

Translating the profiles of the effective marginal tax rates into earnings distributions I again simulate the model, where I in order to simplify matters assume a uniform distribution of potential earnings. The resulting distributions are shows in figure 2.E4. In absence of phase out of benefits the earnings distribution simply follows the distribution of potential earnings, while phase out without the possibility to cancel creates a large excess mass more or less centered on the kink point. Compared to this outcome it is clear from the figure that the possibility to cancel benefits shifts the excess mass below the kink point.

Figure 2.E4

Simulated earnings distribution with and without the possibility to cancel benefits
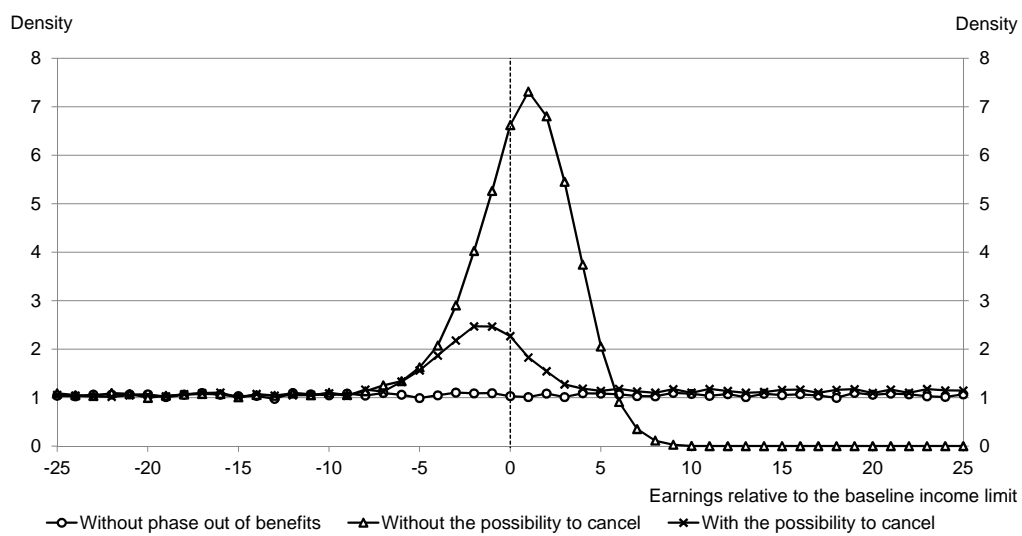
Notes: Simulated distribution of realized earnings based on a draw of 100,000 individuals with potential earnings uniformly distributed from 50 to 150 with the baseline income limit = 100. $\mu = 0.1$ and $\sigma = 5$.

Sources: Own calculations

It should be noted that the earnings distribution without phase of benefits does not equal the distribution of potential earnings as the presence of the linear tax reduces earnings and hence increases the density compared the density of potential earnings (except at the very top of the earnings distribution, where the density drops to 0). This is also the reason why the excess mass in the setting without the possibility to cancel is not exactly centered on the kink point, as the increased marginal tax rate to the right of the kink point even without the excess mass increases the density just above the kink point. In the extreme case here where the marginal tax rate jumps to 100 percent, this creates the perception that the excess mass is centered to the right of the kink point.

3 Parenthood and the Gender Gap: Evidence from Denmark

Joint with H. Kleven (LSE) and Camille Landais (LSE)

3.1 Introduction

Despite considerable gender convergence over the last century, substantial gender inequality persists in all countries and the process of convergence has slowed down. The early literature on gender inequality in the labor market focused on the role of education and discrimination (Altonji and Blank, 1999), but the disappearance of gender differences in education and the implementation of anti-discrimination policies suggest that the explanation for the remaining gender gap lies elsewhere. Based on administrative data for the full population in Denmark since 1980, we provide a simple explanation for the persistence of the gender gap: the effects of parenthood on the careers of women relative to men are large and have not fallen over time. Hence, most of the remaining gender gap can be attributed to children. Our findings are surprising given that Scandinavian countries have been leaders in the implementation of legislation and policies that are supposed to allow women to balance career and family.

To provide context, Figure 3.1A shows the evolution of the gender gap in earnings for full-time workers in different countries. It is striking that the cross-country differences in gender inequality have largely disappeared over time. For example, while gender inequality in Denmark was dramatically lower than in the United States around 1980, today the gender pay gap is between 15-20% in both countries and appears to have plateaued at that level. That is, gender convergence happened earlier in Scandinavia than elsewhere, but the process also slowed down earlier in Scandinavia allowing other countries to catch up. So even though these countries feature different public policies and labor markets, they are no longer very different in terms of overall gender inequality.

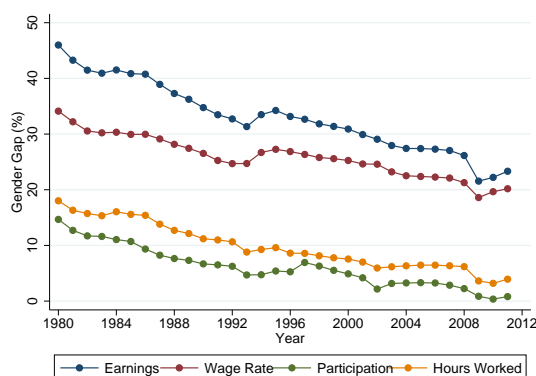Figure 3.1: Gender Gaps Across Countries 1980-2011

A: Convergence of the Gender Pay Gap Across Countries

Median Earnings for Full-Time Workers
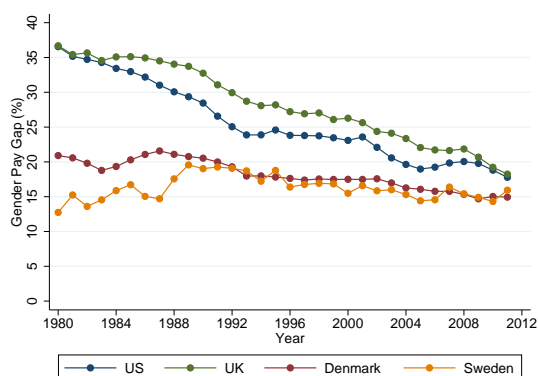
B: Evolution of Gender Gaps in Denmark

Means for All Workers

Notes: The time series in Panel A are drawn from OECD.org, except for Denmark where we use our own calculation of median earnings for full-time workers aged 16-64 (as defined by the ATP hours measure described in section 3.2). Our calculation for Denmark uses the same underlying data as the official OECD series, but is more consistent with the sample definitions used for the other countries. In Panel B the gaps in earnings and participation are calculated among the entire population aged 16-64 regardless of employment status, while the gaps in hours worked and the wage rate is calculated conditional on participation (ATP hours > 0).

To estimate the effect of parenthood on the careers of women relative to men, we adopt a quasi-experimental approach based event studies around the birth of the first child. For a range of labor market outcomes, we find large and sharp effects of children: women and men evolve in parallel until the birth of their first child, diverge sharply immediately after child birth, and do not converge again. The long-run child penalty in female earnings equals 21% over the period 1980-2011. This should be interpreted as a total penalty that includes the costs of children born after the first one, and we show that the penalty is increasing in the number of children. The earnings penalty can come from three margins—labor force participation, hours of work, and the wage rate—and we find sharp effects on all three margins that are roughly equal in size. Our ability to precisely estimate hours and wage rate effects relies on a unique administrative and third-party reported measure of working hours that is available for the full population.

Based on the event study approach, we find effects on several other margins that can shed light on some of the underlying mechanisms. Just after the birth of the first child, women start falling behind men in terms of their occupational rank (ordered by earnings level) and their probability of being manager. Furthermore, women switch

jobs to firms that are more "family friendly" as measured by the share of women with young children in the firm's workforce, or by an indicator for being in the public sector which is known to provide very flexible conditions for working women (see Nielsen et al., 2004). Since family friendly firms are associated with lower earnings and wage rates, this response explains part of the gender gaps described above.

We proceed to decompose the gender gap in the full population (with and without children) into what can be attributed to children and the residual. We show that the fraction of the total earnings gap that can be explained by children has risen from 30% in 1980 to 80% in 2011. This dramatic change reflects a combination of two effects: (i) the child-related earnings gap has increased from about 14% to 18%, and (ii) the total earnings gap has fallen from about 45% to 22%.[1] To understand the first effect, note that although the female child penalty in percentage terms has fallen slightly over time, the penalty operates on a larger base due to the general increase in the earnings of women relative to men. Our findings have implications for future work on gender inequality, which should focus on understanding what drives gender roles and gender outcomes in relation to parenting. This is consistent with the views expressed by Goldin (2014) on what the "last chapter" of gender convergence must contain, but the persistence of child penalties in a country with generous family policies suggests that the last chapter may not be written any time soon.

Further insight can be obtained by analyzing the heterogeneity of child penalties across families. This analysis shows that relative skill within families—as measured by relative wage rates in the years prior to child birth—do not affect child penalties in the direction one might expect. Both the earnings penalty and the wage rate penalty are increasing in the skill of the mother relative to the father, conditional on a rich set of covariates. Even in families where the woman is the primary earner before having children she takes the major hit when children arrives. These findings are interesting in relation to the evidence on the disappearance of the gender gap in education (Goldin et al., 2006 and Goldin and Katz, 2008). While the closing of the education gap has a direct and positive effect on gender equality in earnings (consistent with the narrowing of the non-child gender gap that we document), the potential gain will not be fully realized if child penalties are borne to a larger degree by highly skilled women. The large child penalties for high-skill women that we estimate are consistent with evidence for the US by Wilde et al. (2010) and Bertrand et al. (2010).

---

[1]These gender gaps are larger than those reported in Figure 3.1A discussed above. This is because the cross-country evidence in Figure 3.1A is based on *median* earnings for *full-time* workers, whereas we are now considering *mean* earnings for *all* workers as shown in Figure 3.1B.

The size and persistence of female child penalties, along with their heterogeneity across skill, are difficult to reconcile with comparative advantage alone. A recent literature discusses the importance of social norms and gender identity for explaining the different labor market outcomes of men and women, although causal testing of these ideas is difficult (Bertrand, 2011 and Bertrand et al, 2010). We explore the potential role of such effects by showing that the female child penalty is strongly related to the labor supply history of her parents, conditional on the socio-economic characteristics of the family. For example, in traditional families where the mother works very little compared to the father, their daughter pays a much larger child penalty when she eventually becomes a mother herself. We estimate the intergenerational transmission of child penalties by exploiting that our administrative measure of hours worked goes back to 1964, allowing us to relate the estimated child penalties between 1980-2011 to the within-family work history one generation before. Our findings are consistent with an influence of nurture in the formation of women's preferences over family and career. This analysis is related to the work by Fernandez et al. (2004), but focusing on the intergenerational transmission of child penalties (as opposed to labor supply levels) between parents and their daughters (as opposed to daughters-in-law).

Our paper speaks to the large literature on gender inequality in the labor market (surveyed by Altonji and Blank 1999 and Bertrand, 2011), and it is closely related to a body of work emphasizing the importance of parenthood for gender differences (e.g. Waldfogel, 1998; Paull, 2008; Bertrand et al., 2010; Wilde et al., 2010; Adda et al., 2011 and Goldin, 2014).). We push this literature based on an event study methodology that take advantage of the quality and comprehensiveness of the Danish administrative data. Although we find that the dynamic effects of children on the gender gap are very large, one could argue that the event study approach represents a *lower* bound due to a potential effect of children that it misses. If women select family-friendly career paths (offering flexible hours and generous maternity leave, but lower wages) based on their *planned* fertility prior to child birth, then the pre-child gender gap partly reflects a child penalty. This idea is consistent with our finding that women are working in relatively family-friendly firms and sectors prior to child birth, and that this by itself reduces the child penalty following child birth (see also Nielsen et al., 2004). From this perspective our estimates may be viewed as conservative.

Our paper is also related to the literature on family labor supply and fertility. This literature has tried to estimate the causal effect of children on female labor supply using potentially exogenous variation in family size coming from twin births, sibling sex mix, miscarriage and infertility (see e.g., Browning, 1992; Bronars and Grogger, 1994; Angrist

and Evans, 1998; Hotz et al., 2005; Aguero and Marks, 2008). The empirical approach in our paper is different from this body of work, and the objective is different too. Our identification is based on sharp breaks in career trajectories that occur *just after* having children for women, but not for men. These sharp dynamic patterns are unlikely to be driven by omitted variables or selection on unobservables as these factors should be smooth around the precise moment of child birth.

While we are thus estimating a causal effect of children on labor market outcomes, it is important to keep in mind that having children is a choice and this affects the interpretation. In particular, by estimating the relationship between children and career choices, our results are most naturally interpreted as measuring complementarity in preferences. We show that this complementarity is very strong for women but not for men, and that these gendered preferences can account for most of the remaining gender inequality. The key question is why preferences are so strongly gendered. Is it biology or is there a role for environmental influences in the formation of preferences? We start probing into these questions by considering patterns of heterogeneity and intergenerational transmission, but future work should go further in the investigation of the underlying mechanisms as this will ultimately determine the welfare and policy implications of the patterns we uncover here.

The paper is organized as follows. Section 3.2 describes the context and data, section 3.3 lays out the event study methodology, section 3.4 presents the empirical results, and section 3.6 concludes.

3.2 Context and Data

Context

Scandinavian countries have been praised for offering better opportunities for women to balance career and family than most other countries. This view is based on the presence of generous family policies—job-protected parental leave and public provision of child care—and a perception that gender norms are comparatively egalitarian in Scandinavia. Consistent with this view, Denmark has one of the highest female labor force participation rates in the world, currently around 80% as opposed to around 70% in the United States, and it has almost no remaining gender gap in participation rates. However, Figure 3.1 shows that this is far from the full story. The cross-country comparisons in Panel A were discussed above and shows that Denmark is no longer a strong outlier in terms of the gender gap in earnings for full-time workers. Panel B focuses on Denmark alone and shows gender gaps in different labor market outcomes for all workers.

We see that the gender gap in participation has gradually disappeared over the last three decades and that the gender gap in hours worked has fallen substantially, but that large gaps persist in earnings and wage rates (defined as earnings/hours worked among those who are working). The earnings gap is now around 22% and is created mostly by differences in wage rates and to a smaller degree by differences in hours worked.[2]

Figure 3.2 probes the idea that gender norms are more egalitarian in Scandinavia than elsewhere. The evidence in the figure is based on questions from the International Social Survey Program (ISSP) regarding the attitudes that people have towards market work by women with and without children. Specifically the survey asks participants whether they think women should work outside the home full-time, part-time or not at all when they have no children (Panel A), have children under school age (Panel B), have children in school (Panel C), and have children who have left home (Panel D). Two striking insights emerge from the figure: one is that gender attitudes are still quite traditional—essentially that women should work full-time before having children and after the children have left home, while they should work only part-time or not at all when they have children living at home—and the other is that different countries are very similar in holding this view. The only noticeable cross-country difference is that the Scandinavian populations are somewhat more open to the idea that women with young children work part time (rather than staying at home entirely) compared to the US and UK populations, but overall the similiarities in gender attitudes stand out much more than the differences. The figure is based on samples that include both men and women, but interestingly there is very little difference in these gender attitudes between men and women. Overall and in contrast to common wisdom, the evidence presented in Figures 3.1 and 3.2 raises doubts about the degree to which Scandinavian countries are positive outliers in terms of gender equality in the labor market.

---

[2]As we describe below, the way we measure hours worked means that we understate somewhat the gender hours gap and by implication overstate the gender wage rate gap. That is, the decomposition in Figure 3.1B of the earnings gap into the underlying margins is tilted somewhat from hours worked to wage rates.
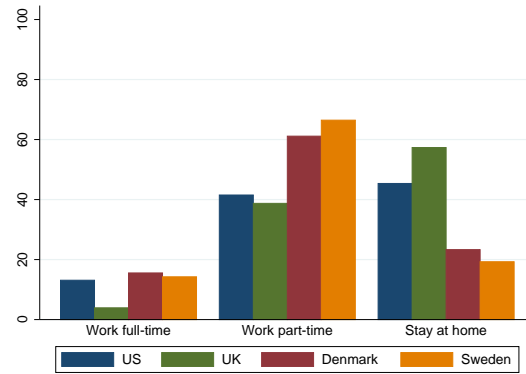
Figure 3.2: Gender Norms Across Countries

A: Women Without Children

Do you think that women should work outside the home full-time,part-time or not at all when they are married but with no children?
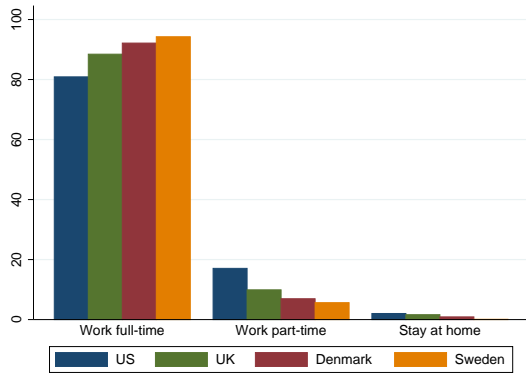
B: Women With Children Under School Age

Do you think that women should work outside the home full-time, part-time or not at all when there is a child under school age?

C: Women With Children In School

Do you think that women should work outside the home full-time, part-time or not at all when the youngest child is still in school?
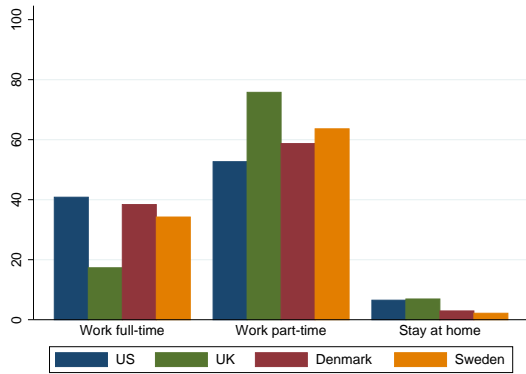
D: Women With Children Who Have Left Home

Do you think that women should work outside the home full-time, part-time or not at all when the child has left the home?
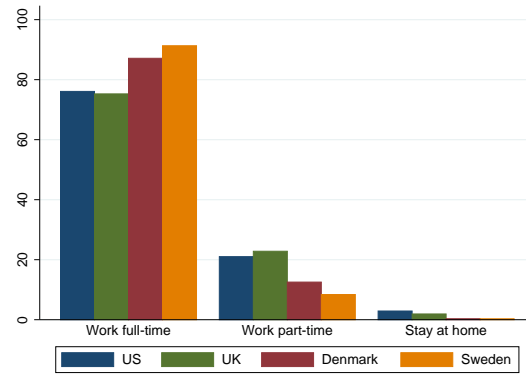
Notes: The figure is based on data from the International Social Survey Program (ISSP) in 2002. Each panel shows shares (in percent) choosing each of the 3 listed categories.

The policy environment in Denmark is one which combines large tax-transfer distortions (which may affect the gender gap due to differential labor supply elasticities between men and women) and generous family policies intended to support female labor

supply. As shown by Kleven (2014), the effective tax rate on labor earnings is exceptionally large in Denmark, but so are the implicit subsidies to labor supply through publicly provided child care and public spending on other goods that are complementary to working (transportation, elder care, education, etc.). Over the period we consider, public child care is universally provided at a heavily subsidized price from around 6-12 months after birth. Until the child reaches the age where public child care becomes available, there is job-protected and paid maternity and parental leave. Up until 2001, parents were offered 14 weeks of maternity leave followed by 10 weeks of parental leave to be shared between the mother and father. Since 2002 this has been extended to 18 weeks of maternity leave and 32 weeks of parental leave. Hence, throughout the period we consider, parents were covered first by paid leave and then by public child care, with no gap between the two.

Data

The analysis is based on administrative data for the full population in Denmark between 1980-2011. For the study of intergenerational transmission we exploit additional administrative data going back to 1964. The Danish data combines several different administrative registers (linked at the individual level via personal identification numbers) and therefore contains rich information on children, earnings, labor supply, occupation, firm, education, and many other outcomes. Furthermore, the data allows us to link family members, generations, and workers to firms.

The Danish population is currently 5.5 million people and there were around 2 million child births during the period 1980-2011. For our main event study analysis we focus on *first* child births where both parents are observed every year between 5 years before having a child and 10 years after. We are thus focusing on first child births between 1985-2001 where both parents are alive and reside in Denmark throughout a 15-year window around the birth. We condition on both parents being at least 20 years of age when having their first child (teenage births constitute only 2.3% of all births during the 1985-2001 period). We do not impose restrictions on the relationship status of the parents: we include all individuals who have a child together in a given year and follow them through the 15-year window whether or not they are married, cohabiting, separated, divorced, or have not formed a couple yet in any given year. This leaves us with a core estimation sample of around 350,000 births or 11,200,000 individual-year observations.

We estimate child penalties in earnings, labor force participation, hours worked, and wage rates (earnings/hours worked for those who are working). Our ability to esti-

mate child penalties in hours worked and wage rates using sharp event studies relies on a unique administrative and third-party reported measure of hours worked that is available for the full population. This measure comes from a mandated pension scheme introduced in 1964—*Arbejdsmarkedets Tillægspension* (ATP)—that require all employers to contribute on behalf of their employees based on individual hours worked.[3] The pension contribution is a function of hours worked in discrete steps, namely four bins of weekly hours (0-8, 9-17, 18-26, 27-) for someone paid weekly or four bins of monthly hours (0-38, 39-77, 78-116, 117-) for someone paid monthly, with the latter being much more common. Hence the annual pension contribution for someone paid monthly depends on $\sum_{i=1}^{12} h_i$ where monthly hours h_i has 4 steps, which gives an annual hours measure in 37 steps (= 4 × 12 - 12 + 1). Our measure of the wage rate is defined as earnings divided by this ATP hours measure.

Because the ATP hours measure is capped, it does not capture marginal hours adjustments for those working every month of the year in the highest hours bin. For a given child penalty in earnings, this will make us underestimate the penalty in hours worked and correspondingly overestimate the penalty in wage rates. The hours measure does capture larger labor supply adjustments such as switches to different levels of part-time work and work interruptions within the year, which are important adjustments for women with children. The key advantage of our measure is that it is precisely measured for the full population over a very long time period, unlike labor market surveys that have considerable measurement error and small samples. Moreover, we are able to validate our results for the discrete hours measure against a continuous hours measure reported by all firms on behalf of their employees, but only since 1997.

3.3 Event Study Approach

The idea of the event study approach is to estimate female child penalties based on (sharp) changes that occur just after child birth for mothers relative to fathers. For each parent in the data we denote by $t = 0$ the year in which the individual has his/her first child and index all years relative to that year. As described above, we consider a balanced panel of parents who we observe every year between 5 years before having their first child and 10 years after, and so event time t runs from -5 to +10. We then study the evolution of different outcomes (earnings, labor supply, wage rates, etc.) as a function of event time.

Specifically, denoting by Y_{ist} the outcome of interest for individual i in year s at event

---

[3]The scheme also covers self-employed individuals who contribute on their own behalf.

time t, we run the following regression separately for men and women

$$Y_{ist} = \sum_{t \neq -1} \alpha_t \cdot EVENT_{it} + \sum_j \beta_j \cdot AGE_{is}^j + \sum_s \gamma_s \cdot YEAR_s + \nu_{ist}, \qquad (3.1)$$

where $EVENT_{it}$ is an event time dummy, AGE_{is}^j is an age dummy for being j years old, and $YEAR_s$ is a year dummy. We omit the event time dummy at $t = -1$, implying that the event coefficients α_t measure the impact of children relative to the year just before the first child birth. If we did not include the age and year dummies in the specification, the estimated event coefficients $\hat{\alpha}_t$ would correspond simply to the mean value of the outcome at event time t, measured relative to the year before birth. By including a full set of age dummies we control non-parametrically for non-child related career progression, and by including year dummies we control non-parametrically for non-child related time changes such as wage inflation and business cycles. In other words, the age and year dummies remove any underlying life-cycle and time trends in the outcomes we consider.

We specify equation (3.1) in levels rather than in logs to be able to keep the zeros in the data (due to non-participation). We convert the estimated level effects into percentages by calculating $\hat{\alpha}_t / E\left[\hat{Y}_{ist} \mid t\right]$ where $\hat{Y}_{ist}$ is the predicted outcome when omitting the contribution of the event dummies, i.e. $\hat{Y}_{ist} \equiv \sum_j \hat{\beta}_j \cdot AGE_{is}^j + \sum_s \hat{\gamma}_s \cdot YEAR_s$. This captures the year-t effect of having a child as percentage of the counterfactual outcome absent the child. We estimate this separately for men and women and denote the gender-specific effects by $P_t^k \equiv \hat{\alpha}_t^k / E\left[\hat{Y}_{ist}^k \mid t\right]$ where $k = m, w$. We then define the *long-run child penalty* on women as the average effect of children over a 10-year horizon for women relative to men, i.e.

$$\Delta P \equiv P^m - P^w \quad where \quad P^k \equiv E\left[P_t^k \mid 0 < t \leq 10\right]. \qquad (3.2)$$

Hence, the child penalty ΔP is the percentage by which women are falling behind men due to children over a 10-year period following the first child birth. The choice of a 10-year window is based on the empirical analysis below, which shows that the effect is roughly at a steady state by that time.

To gain insight into the potential determinants of child penalties, we present a detailed study of heterogeneity using the rich observational data. Here we consider penalties at the family level and regress these on a range of non-parametric controls. The

long-run child penalty on the female in family *i* is defined as

$$\Delta p_i \equiv p_i^m - p_i^w \quad where \quad p_i^k \equiv \frac{E\left[Y_{it}^k \mid 0 < t \le 10\right] - E\left[Y_{it}^k \mid -5 \le t < 0\right]}{E\left[Y_{it}^k \mid -5 \le t < 0\right]}, \quad (3.3)$$

that is, the percentage change in a given outcome between the 5-year period before birth and the 10-year period after birth for the man relative to the woman within a family. The family-level penalty in equation (3.3) is conceptually similar to the aggregate-level penalty in equation (3.2), but in general it does not aggregate to the same number: the average over family-level penalties in percentages ($E[\Delta p_i]$) is not the same as the percentage penalty in the average levels (ΔP) due to a potential correlation between penalties and the outcome level. Furthermore, when focusing on family-level penalties, we have to drop families in which one or both of the parents have zeros (non-participation) in all five years preceding the arrival of the child. The key advantage of defining family-level penalties is that it allows us to study heterogeneity in a given dimension while controlling for the (correlated) variation in other relevant determinants of penalties.

We regress the penalty Δp_i on a range of variables that capture the socio-demographics, work environment, and relative skill of the two parents in the family. The richest specification we consider is the following

$$\begin{aligned}
\Delta p_i \quad = \quad & \sum_c \beta_{1c} \cdot COHORT_i^c + \sum_k \sum_j \beta_{2jk} \cdot AGE_i^{jk} + \sum_j \beta_{3j} \cdot KIDS_i^j + \sum_j \beta_{4j} \cdot INCOME_i^j \\
& + \sum_j \beta_{5j} \cdot EDUCATION_i^j + \sum_j \beta_{6j} \cdot SKILL_i^j + \sum_k \sum_j \beta_{7j} \cdot EXPERIENCE_i^{jk} \quad (3.4) \\
& + \beta_8 \cdot PUBLIC_i + \sum_j \beta_{9j} \cdot FRIENDLY_i^j + \sum_k \sum_j \beta_{10j} \cdot PROFESSION_i^{jk} + \mu_{ic}.
\end{aligned}$$

The explanatory variables in (3.4) are dummies defined as follows: $COHORT_i^c$ is a dummy for the first child being born in year c, AGE_i^{jk} is a dummy for parent k being j years old when having the first child, $KIDS_i^j$ is a dummy for having j children in total (1, 2, 3, 4+), $INCOME_i^j$ is a dummy for being in the jth decile of the household income distribution just before having the first child, $EDUCATION_i^j$ is a dummy for being in the jth quartile of the distribution of relative years of education between parents (based on completed education before having a child), $SKILL_i^j$ is a dummy for being in the jth decile of the distribution of relative wage rates between parents (based on the five years prior to having the first child), $EXPERIENCE_i^{jk}$ is a dummy for parent k having j years of experience between completing education and the arrival of the first child (bottom

coded at zero if education is completed after child birth), $PUBLIC_i$ is a dummy for the woman working in the public sector at the time of having her first child, $FRIENDLY_i^j$ is a dummy for the woman working in a firm belonging to the jth quartile of the distribution of family friendliness when she has her first child, and $PROFESSION_i^{jk}$ is a dummy for parent k being in profession j (based on 22 categories education fields). Family friendliness $FRIENDLY_i^j$ is based on the share of women with young children in the firm's workforce (or by the presence of women with young children in the firm's management).

3.4 Empirical Results I: Child Penalties and the Gender Gap

Estimating Child Penalties

We start by documenting a set of stark changes in the labor market outcomes of women relative to men around the event of having the first child. Figure 3.3 plots P_t^m and P_t^w as defined in section 3.3: these are outcomes for men and women as a function of event time t, measured relative to the year just before the birth of the first child ($t = -1$) and controlling non-parametrically for age and time trends. The figure includes 95% confidence bands around the event coefficients, although these are not always clearly visible due to the enormous amount of precision in the administrative data. Panel A starts by considering total earnings before taxes and transfers. We see that, once life-cycle and time trends are taken out, the earnings of men and women evolve in a strikingly parallel way until they become parents. But at the precise moment that the first child arrives, the earnings paths of men and women begin to diverge: women experience an immediate drop in gross earnings of almost 30%, while men experience no significant variation in their earnings. Importantly, in the years following the initial drop, the earnings of women never converge back to their original level. Ten years after the birth of a first child, female earnings have plateaued around 20% below its level just before child birth, whereas male earnings are essentially unaffected by children. As shown in the figure, these estimates imply a long-run child penalty in the earnings of women relative to men (ΔP defined in equation (3.2)) equal to 20.8%.

While we take an event study approach using the birth of the first child, the evidence presented in Figure 3.3 is based on the full population of women with children, irrespective of the total number of children a woman ends up having. This implies that the dynamic patterns we document include the effects of children born after the first one, and the estimated long-run child penalty is naturally interpreted as the aggregate penalty of all children. In appendix Figure 3.9 we replicate the earnings event study

from Figure 3.3A for different numbers of children (1, 2, 3, 4+). The impact of children is sharp for all four family sizes and the magnitude of the long-run child penalty varies with the number of children in the expected direction. The child penalty increases by roughly 10 percentage points per child. Note that the event study graph for families with two children in Panel B of Figure 3.9 looks very similar to the graph for all families in Figure 3.3, which is natural given that the average completed fertility per woman in Denmark is close to two conditional on having children.

Figure 3.3: Child Penalties on Women
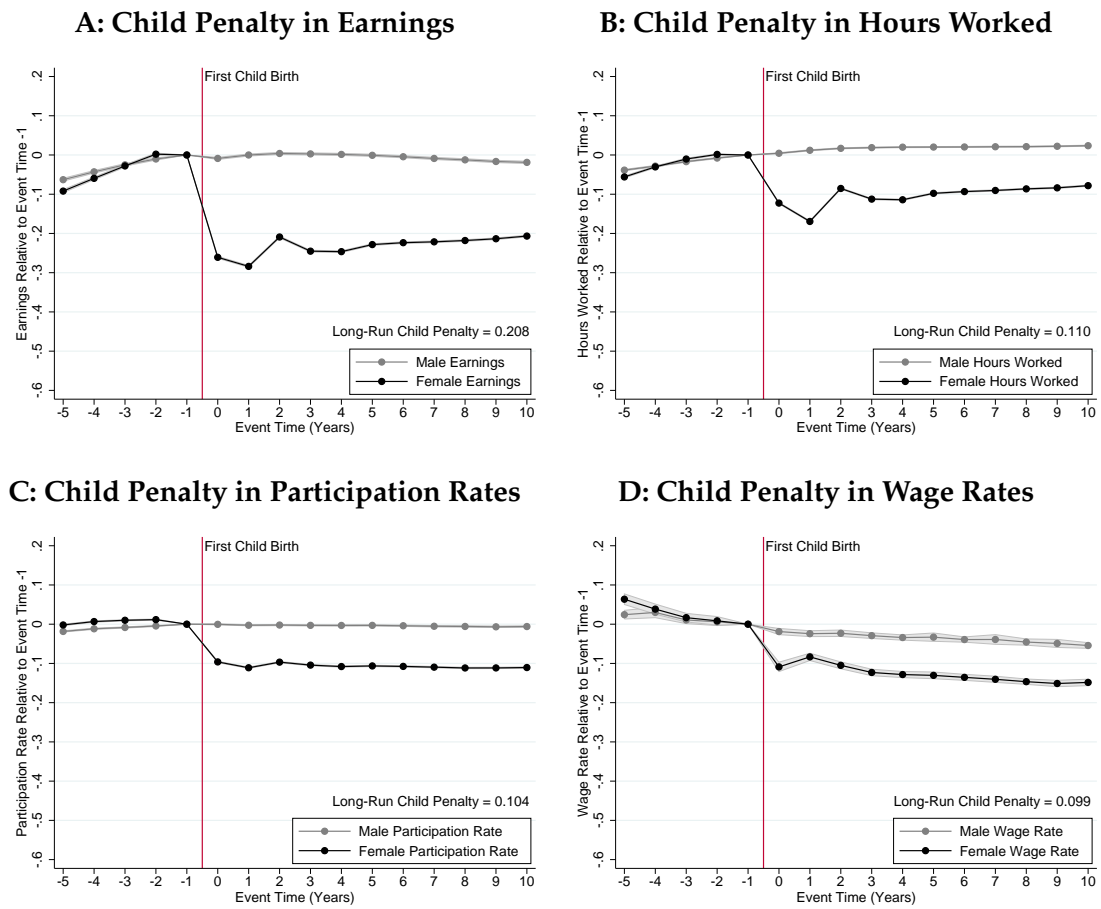
Notes: The panels show estimated coefficients on the event time dummies in equation (3.1) as a fraction of the predicted outcome when omitting the contribution from the event dummies (i.e., $P_t^k = \hat{\alpha}_t^k / E[\hat{Y}_{ist}^k | t]$ defined in section 3.3). The coefficients are estimated on a balanced sample of parents, who have their first child between 1985-2001 and who we observe in the data during the entire period between 5 before and 10 years after child birth. The effects on earnings and participation are estimated unconditional on employment status, while the effects on hours worked and the wage rate are estimated conditional on participation (ATP hours > 0). The shaded 95 % confidence intervals are based on robust standard errors.

In general, earnings responses can be driven by three margins: labor force participation, hours worked conditional on participation, and the wage rate. In panels B-D of Figure 3.3 we analyze how parenthood affects each of these three margins separately using the same methodology as above. Panel B plots the evolution of hours worked for men and women relative to the year before the first child birth. Hours worked follow the same qualitative pattern as earnings, with a sharp and persistent drop after child birth for women relative to men. Three years after birth, hours worked by women are 10% lower than before birth, while hours worked by men are almost unchanged. Ten years after birth, there is no sign of convergence; a persistent 10% gender gap in hours worked has been created due to children. Panel C displays the evolution of the labor force participation rates of men and women. Again, the labor force participation trends of men and women are perfectly similar until the birth of a first child, at which point they sharply diverge with a 10% relative drop for women that fully persists over time. Finally, Panel D shows that wage rates feature a similar dynamic pattern: men and women are on very similar trends prior to the birth of the first child, diverge immediately after birth, creating a 10% gap between men and women that does not fade over time.

These results show that the female child penalty in earnings is in part a direct consequence of intensive and extensive labor supply adjustments made by the family after the birth of the first child. At the same time, the wage rate patterns suggest that there is more going on than these quantitative labor supply adjustments. A possibility is that women, once they become mothers, make active career decisons in other more qualitative dimensions (choice of occupations, tasks, firms) that have immediate and persistent effects on their wage rates. We provide direct evidence on such margins of adjustment in the next section. Interestingly, the estimated long-run penalties at the intensive, extensive, and wage rate margins are roughly similar in magnitude, suggesting that these margins are equally important for understanding the long-run effect of children on the earnings paths of women relative to men.[4]

In the event graphs presented so far, the drop in earnings and labor supply in event year 0 is not much larger than the drop in subsequent event years. While this may seem surprising, note that the use of calendar-year measures of earnings and labor supply tend to create attenuation bias in the first-year dip compared to a continuous time representation: as women give birth some time during year 0, some of the earnings and

---

[4]The child penalties in panels B-D of Figure 3.3 are unconditional penalties: when estimating the effect of parenthood on one particular margin, we are not controlling for the other two margins in the regression. This explains why the long-run penalties on the three margins do not sum up to the overall earnings penalty.

hours in calendar-year 0 were realized prior to child birth. To investigate this point, we reproduce Figure 3.3 on a sample restricted to child births in January for which the definition of calendar years and event years coincide. The results are shown in appendix Figure 3.10 from which the following insights emerge. First, when focusing on January births alone we do see a pronounced dip in event year 0 as one would expect. This dip reflects the extra time taken out of the labor market immediately following delivery. Second, focusing on January births also reveal a slight drop in labor market outcomes in event year -1, which can be explained by sick leave and parental leave (for which women are eligible during the last four weeks of pregnancy) taken just before giving birth. Third and most important, the long-run child penalties over a 10-year horizon are very similar for January births and all births, which implies that the calendar-year presentation in Figure 3.3 does not introduce any bias in the long run.

We have presented estimates on the career cost of children using child births between 1985-2001 and an event study horizon that includes 10 post-birth years. It is of course interesting to study how these career patterns evolve in the very long run, which is feasible to explore with our data. In appendix Figure 3.11 we consider child births between 1985-1991 and a 20-year post-birth horizon. The figure is otherwise similar to Figure 3.3 and shows results for earnings, hours worked, participation, and wage rates. The long-run child penalty estimates shown in each panel is based on an average over event years 10-20 (ΔP in equation (3.2) for $10 < t \leq 20$). The figure shows how strikingly persistent the effects of children are. In fact, the earnings penalty 20 years after child birth is almost the same as the penalty 10 years after. The only qualitative difference that emerges from considering the very long run is that hours worked do eventually begin to converge, while at the same time wage rates keep diverging. The combination of the narrowing hours gap and the widening wage rate gap produces a constant earnings gap.

It is useful to take a step back in order to discuss identification and how to interpret the event study estimates we have presented. Consider first the relationship between our approach and the vast literature on family labor supply and fertility (e.g., Browning, 1992). This literature has discussed the difficulties of interpreting the observed negative correlation between children and female labor supply, noting that causal inference is difficult due to omitted variables and reverse causation. A series of papers try to estimate the causal effect of children on female labor supply using potentially exogenous instruments for family size such as twin births, sibling sex mix, miscarriage, and infertility (e.g., Bronars and Grogger, 1994; Angrist and Evans, 1998; Hotz et al., 2005; Aguero and Marks, 2008). While this literature has been constrained by data

limitations—having to rely on cross-sectional variation in survey data—we leverage full-population administrative panel data in order to pursue an event study strategy that exploits sharp breaks in career trajectories occurring *just after* having children for women relative to men. The sharp dynamic patterns that we document are unlikely to be driven by omitted variables (such as unobserved heterogeneity in career preferences) as these should be smooth around the moment of child birth, nor are they driven by reverse causality as the labor market changes occur after child birth. Broadly speaking, our identification is related to the fundamental insights of Sims (1972) and the concept of Granger causality: we exploit the fact that the arrival of children is sharply related to future career trajectories, but not to past career trajectories. Examples of papers that come close to our event study strategy include Paull (2008) on the impact of children on hours worked in the UK and Wilde et al. (2010) on the impact of children on female wages in the US, although they do not push the analysis of anatomy, mechanisms, and secular composition changes as we do in this paper.

While the preceding arguments suggest that we are uncovering a causal relationship between children and labor market outcomes, it is important to keep in mind that having children is a choice and this affects the interpretation. Three points are worth noting. First, we are estimating the effect of children on the sample of individuals who have selected parenthood as opposed to the effect of an exogenous change in children on the full population. As in the IV approaches discussed above, what we obtain is a treatment effect on the treated. Second, since we are estimating the relationship between choice variables—having children and various labor market choices—the results are most naturally interpreted as measuring complementarity in the utility function. The decision to have children and a less ambitious career is strongly complementary for women, but not for men. The deeper question is why this complementarity is so strongly gendered, a question to which we return in section 3.5. Third, because parenthood is a *planned* choice, some of the labor supply and career decisions that are complementary to parenthood could be made prior to the birth of the first child. Although the striking similarity of pre-child trends for men and women suggests that such anticipatory responses are relatively limited, we cannot rule out that some women make child-related career choices far in advance, before our event study window starts. For example, this would be the case if a woman decides to never enter the labor force in anticipation of becoming a mother. For such a woman, the estimated child penalty using our event study methodology would be zero, although the true child penalty is positive and already incorporated in the pre-child gender gap in participation rates. Hence the large child penalties that we find are, if anything, *lower bounds* on the total career effects of

children.

Anatomy of Child Penalties

We have seen that motherhood is associated with large and persistent penalties in earnings driven in roughly equal proportions by participation, hours of work, and wage rates. These empirical patterns, and especially the large effects of children on wage rates, beg the question of what are the underlying mechanisms that drive the effects. This section focuses on this question, leveraging the rich administrative data to explore if child birth changes women's careers in qualitative dimensions that affect their productivity in the labor market. The results are presented in Figure 3.4, which is based on the same event study methodology used above.

Figure 3.4: Anatomy of Child Penalties

A: Occupational Rank

Levels 1-5 from Unskilled Labor to Manager

B: Probability of Being Manager

Manager Dummy

C: Probability of Public Sector Job

Public Sector Dummy

D: Family Friendliness of Firm

Share of Women with Young Children in the Firm

Notes: The panels show the estimated coefficients on the event time dummies in equation (3.1) as a fraction of the predicted outcome when omitting the contribution from the event dummies (similar to Figure 3). The effect on occupational rank is estimated conditional on not being self-employed or an assisting spouse. The effect on the probability of being employed in the public sector is estimated conditional on having a recorded sector variable. The effect of the family friendliness of the firm (i.e., share of women with young children—defined as children below age 15—in the firm) is estimated conditional on being in a firm with more than 10 employees. Moreover, the share of women with young children in a given woman's firm is calculated excluding the woman's own child (jack-knifed). The long-run child penalty/effect is calculated as the average effect for women relative to men over the period from event time 5 to 10. The long-run effect on the family friendliness of the firm takes into account the differential pre-trend between men and women.

Panels A considers occupational rank in five levels: unskilled labor, skilled labor,

white-collar work (low level), white-collar work (high level), and top manager. This ordering of occupations is consistent with an ordering based on average earnings or average wage rates in each occupation. This panel shows that men and women are on identical trends in terms of their occupational rank prior to becoming parents (controlling non-parametrically for age effects), but that women start falling behind men soon after parenthood. Note that the occupation graphs begin to diverge in event year 1 rather than in event year 0 as for earnings and labor supply. This makes sense given that women are giving birth during year 0, so that this year consists partly of a pre-birth period and partly of a period covered by job-protected parental leave. Hence, women do not have a strong incentive to change occupation within year 0, but can wait until year 1 when they come back to work. Panel B also explores occupational rank, but focuses specifically on the probability of being top manager. It is in general harder to uncover effects of children on this margin, because relatively few individuals have risen to the managerial level prior to becoming parents. Nevertheless, the graph suggests that parenthood has a negative effect on women's prospects of becoming managers. While the male and female trends are not perfectly similar prior to child birth, they do begin to diverge at a much faster pace following child birth.

The bottom panels turn to the choice of work environment and in particular its family friendliness. We first consider the link between parenthood and the decision to work in the Danish public sector, which has a long tradition of focusing on working conditions rather than on wages. This includes flexible working hours, leave days for those with sick children, and a favorable view on long parental leaves (see Nielsen et al., 2004 for a detailed description). It is therefore natural to expect that mothers would be induced to move into the public sector, a hypothesis that is clearly confirmed in Panel C. Women and men are on very similar pre-child trends in terms of their probabilities of working in the public sector, but begin to diverge strongly soon after having a child. Ten years after child birth, women have a 10pp higher probability of public sector employment than men, relative to the year before child birth. As with occupation, the divergence mainly starts in year 1 rather than in year 0, i.e., when women return to work after their parental leave.

Finally, Panel D considers a proxy for the family friendliness of a work environment that also encompasses heterogeneity across firms in the private sector. Here we take advantage of our employer-employee matched data by defining a firm's "family friendliness" as the share of women with children below 15 years of age in the firm's workforce (excluding the considered woman's own first child when it arrive). Having a larger share of female employees with young children may reflect that the firm offers more

family-friendly policies, or that the firm is more family friendly in the broader sense of employing people that women with children see as like-minded. Since our measure of firm family friendliness is negatively related to wage rates, if women move into more family-friendly firms following parenthood, this helps explain the wage rate penalties documented above. The outcome shown in Panel D is the percentile rank in the distribution of firm family friendliness for men and women, respectively, relative to the year before child birth. Although men and women are not on identical trends prior to birth (the female trend is steeper), there is a very clear break in the relative trends around the moment of having a child. Women begin to move into family-friendly firms at a much higher pace in the years following child birth, whereas the male trend is completely unaffected by child birth. The female trend is increasing somewhat already in event year -1, consistent with an anticipation effect of motherhood. Taking the differential pre-trend into account we estimate a long-run effect of parenthood on the percentile rank in the distribution of firm family friendliness for women relative to men equal to 4.36.

Overall, the results in Figures 3.3-3.4 show that women's career trajectories change sharply due to motherhood, creating substantial gender inequality in a range of quantitative and qualitative dimensions. The results demonstrate the difficulties that women continue to face in trying to balance career and family, and are broadly consistent with the arguments by Goldin (2014) on the "last chapter" of gender convergence. As discussed above, our large effects are, if anything, lower bounds as they do not include the potential anticipatory responses to planned parenthood. For example, while we find sharp effects on women's decision to work in the public sector or in a family friendly firm just after child birth, it is entirely conceivable that some women have made decisions to be in such sectors and firms far in advance in anticipation of eventual motherhood. Consistent with this, women are more likely than men to work in the public sector or in a more family friendly firm already prior to birth. Such lifetime effects are difficult to estimate without making strong parametric assumptions.

Decomposing Gender Inequality Over Time

In this section we study the long-run evolution in the composition of gender inequality into what is driven by children and what is driven by other factors (such as human capital or discrimination). For the reasons just discussed, our decomposition into child-related gender inequality and residual gender inequality will, if anything, be biased towards the latter, because there may be lifetime effects of anticipated parenthood that are not captured by our event study methodology. As we shall see, this potential bias

makes our findings all the more striking.

Our decomposition analysis is implemented as follows. The first step is to estimate cohort-specific child penalties, which we do by extending the baseline event study specification (3.1) in the following way

$$Y_{ist} = \sum_s \sum_{t \neq -1} \alpha_{st} \cdot EVENT_{it} \cdot YEAR_s + \sum_j \beta_j \cdot AGE_{is}^j + \sum_s \gamma_s \cdot YEAR_s + \nu_{ist}, \quad (3.5)$$

with the only innovation being that we interact the event time dummies by year dummies in order to estimate year-specific event coefficients α_{st}. Note that estimating event coefficients by year s and event time t amounts to estimating event coefficients by birth cohort $c = s - t$. As in the baseline specification we consider an event time window running from -5 to +10, but we expand from the previously balanced panel of individuals who have their first child between 1985-2001 to a semi-balanced panel of individuals who have their first child at any point during the data period 1980-2011. The sample is *semi*-balanced in the sense that early cohorts are not observed all the way back to event time -5 (for example, birth cohort 1981 is not observed before event time -1) and that late cohorts are not observed all the way up to event time +10 (for example, birth cohort 2009 is not observed after event time +2), but within each cohort we require both parents to be present in the maximum number of years possible within our data period. Expanding the sample in this way have no major impact on any of our conclusions, but it is helpful for separately identifying event×year coefficients and year coefficients by creating more independent variation in event time and calendar time towards the end points of the data period.

The earnings penalties for birth cohorts 1985-2001 obtained from specification (3.5) are shown in Figure 3.12. We show short-run earnings penalties in Panel A (an average over event time 0-4) and long-run earnings penalties in Panel B (an average over event time 5-10). We see that there is some cyclicality in the short-run child penalty faced by women, but no statistically significant long-run trend. On the other hand, the long-run child penalty features no cyclicality, but a statistically significant negative time trend.

The second step of the analysis requires us to take a stand on the child penalties faced by women who have their first child outside our event study period, but are in the labor market at some point during 1980-2001. For example, women who have their first child in 1978 will be at event time +7 in 1985, and we have to assign a child penalty associated with this event time and year for our historical decomposition analysis. The results presented in Figure 3.12 give guidance on how to do this. The child penalties for event time 5-10 are on an almost perfectly linear trend between 1985-2001, and so we

extrapolate linearly to obtain penalties for these event times outside our event study period. On the other hand, the child penalties for event time 0-4 are not trending between 1985-2001, and so we simply assume that they were constant at their 1985 level prior to that year and that they were constant at their 2001 level following that year.

The third step of the analysis requires us to take a stand on the child penalties faced by women after the end of our event time window, i.e. from event time +11 onwards. We already analyzed longer-run penalties in Figure 3.11, which showed clearly that earnings penalties are extremely stable from event year +10 onwards. Hence we assume uncontroversially that each women is at a steady state from 10 years after birth.

The fourth and final step is to decompose the gender gap using the estimates and assumptions described above. Building on the notation from section 3.3, the percentage child effect in event year t and calendar year s is denoted by P_{st}^k for $k = m, w$, and so the female child penalty associated with this event and calendar time is given by $\Delta P_{st} \equiv P_{st}^m - P_{st}^w$. Given the previous three steps, we have an estimate of ΔP_{st} for any event time and any year during 1980-2011, which we can use to decompose the gender gap. If the actual earnings of a woman with children are Y_{ist}, then we construct the earnings she would have had absent children as $\hat{Y}_{ist} \equiv Y_{ist} / (1 - \Delta P_{st})$. We do not adjust the earnings of men (as the adjustment for women is already based on an estimate relative to men), nor do we adjust the earnings of women before they become mothers or women who never have children. Using the adjusted earnings $\hat{Y}_{ist}$, we construct a new gender gap— this is the residual gap not related to children. The difference between the residual gender gap and the actual gender gap is the child-related gender gap.

The results of our decomposition exercise are shown in Figure 3.5. We see that the fraction of gender inequality in earnings that can be attributed to children has increased dramatically over time, from about 30% in 1980 to about 80% in 2011. This secular change reflects a combination of two effects: (i) the child-related gender gap in earnings has increased from about 14% to 18%, and (ii) the total gender gap in earnings has fallen from about 45% to 22%. To understand the first effect, note that although the *percentage* child penalty on women has fallen slightly over time (as shown in Figure 3.12), the penalty operates on a larger base due to the general increase in the earnings of women relative to men coming from the second effect. That is, at a time where non-child gender inequality is falling (for example, due to changes in education or discrimination) while child penalties are roughly constant or falling by less, there will be a tendency for child-related gender inequality to go up.

Figure 3.5: Decomposing the Gender Gap Over Time
Child-Related Earnings Gap vs Non-Child Earnings Gap
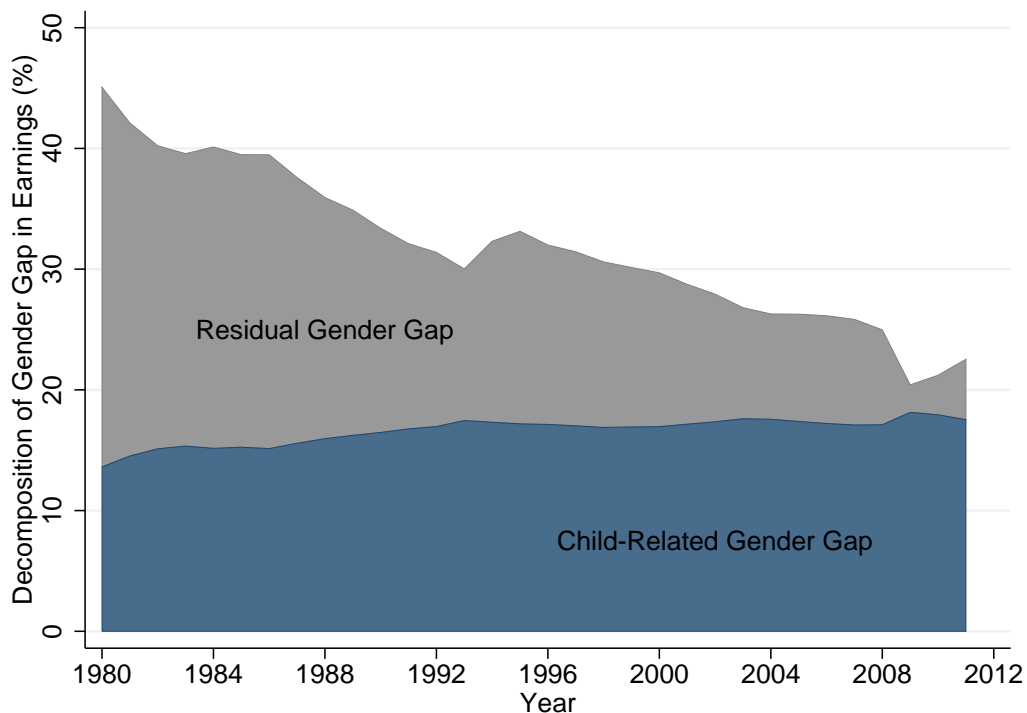
Notes: The gender gap in earnings is calculated using the entire population aged 16-64 regardless of employment status, and so the total gap (residual + child-related) corresponds to the earnings gap shown in Figure 3.1B. The decomposition of the total gap is based on the methodology developed in section 3.4. The residual gap is based on female earnings adjusted for effect of children using the estimated child penalties for each calendar and event year, namely earnings $\hat{Y}_{ist} = Y_{ist} / (1 - \Delta P_{st})$ where Y_{ist} is actual earnings and $\Delta P_{st} = P_{st}^m - P_{st}^w$ is the child penalty in calendar year s and event time t. The cohort-specific penalties ΔP_{st} are estimated based on equation (3.5) on a semi-balanced panel, with linear extrapolation to cover cohorts outside our data period as described in section 3.4.

Our findings imply that, to a first approximation, the remaining gender inequality is all about children. This has important implications for future work on gender inequality, which should focus on understanding what drives gender outcomes in relation to parenthood. There is likely to be some biological element to these gender differences (innate differences in childrearing abilities or preferences), but the more interesting question for economists is whether the strongly gendered parental outcomes are also driven by environmental influences, labor markets, and policy. These are the aspects that the gender inequality literature has already been focusing on (see Bertrand, 2011), but the findings in Figure 3.5 highlight that those topics must be studied specifically in the context of having children in order to shed light on the remaining gender

inequality.

3.5 Empirical Results II: Heterogeneity and Mechanisms

Heterogeneity in Child Penalties

We have documented the existence of large child penalties on women's careers in a number of dimensions, and we have shown that these penalties can account for almost all of the remaining gender inequality in earnings. The key question is why the labor market effects of parenthood are so strongly gendered after decades of legislation and policies that are supposed to foster gender equality (and in a country often seen as a leader on this front). Are these differences mostly due to biology or is there an important role for environmental influences? A first step in exploring this question is to document the degree and patterns of heterogeneity. The presence of large heterogeneity by itself speaks against the idea that child-related inequality is *only* about biology and lend support to the role of environmental influences. Moreover, the specific patterns of heterogeneity provide suggestive evidence of mechanisms, although our descriptive analysis of heterogeneity does not identify causal effects as such.

One way of analyzing heterogeneity would be to consider each dimension of interest separately. The problem with such a strategy is that many of the interesting variables are highly correlated, making it hard to draw meaningful conclusions when considering them one at a time. As an example take the one dimension of heterogeneity we have already considered, namely heterogeneity in the child penalty by the number of children shown in Figure 3.9. We saw there that the child penalty in earnings is strongly increasing in the number of children, with a roughly 10pp increase in the earnings penalty per child. Yet, having more children is correlated with many other parental variables that could influence the size of child penalties such as the age at first birth, education, occupation, income level, and so on. To properly analyze the patterns of heterogeneity in child penalties and shed light on potential mechanisms, we leverage the granularity and comprehensiveness of the Danish administrative data to document how the child penalty correlates with different dimensions of interest—for example, the number of children, the relative skills of the parents, and the work environment of the mother at the time of birth—holding constant all other potential determinants of child penalties in the data.[5]

---

[5]Of course, the family characteristics we consider are not randomly allocated, but may be partly driven by selection on unobserved preferences over careers and children. While this serves as a warning against causal interpretation, it is worth noting that the sharp event studies presented above suggest

The analysis is executed by regressing family-level penalties Δp_i defined in equation (3.3) on a rich set of non-parametric controls as specified in equation (3.4). Although conceptually similar to the aggregate-level penalties estimated in the previous section, the family-level penalties differ slightly in their definition. Instead of computing the penalty by comparing post-birth outcomes to the outcome at event time -1, the family-level penalties compare post-birth outcomes to the average outcome between event time -5 and -1. This minimizes the noise in family-level penalties introduced by potential mean-reversion in labor market outcomes at the individual level. Note also that if the mother or the father have zero earnings in all of the five years prior to child birth, then the family-level penalty cannot be computed. Finally, our measure of the family-level penalty is unbounded, which may create problems with extreme outliers. We address this concern in two ways. First, in our baseline specification we simply exclude penalties over 400% and under -400%. Second, we also consider results from robust regressions using a Huber M-estimator, which imposes weights on observations so as to reduce the influence of outliers.[6]
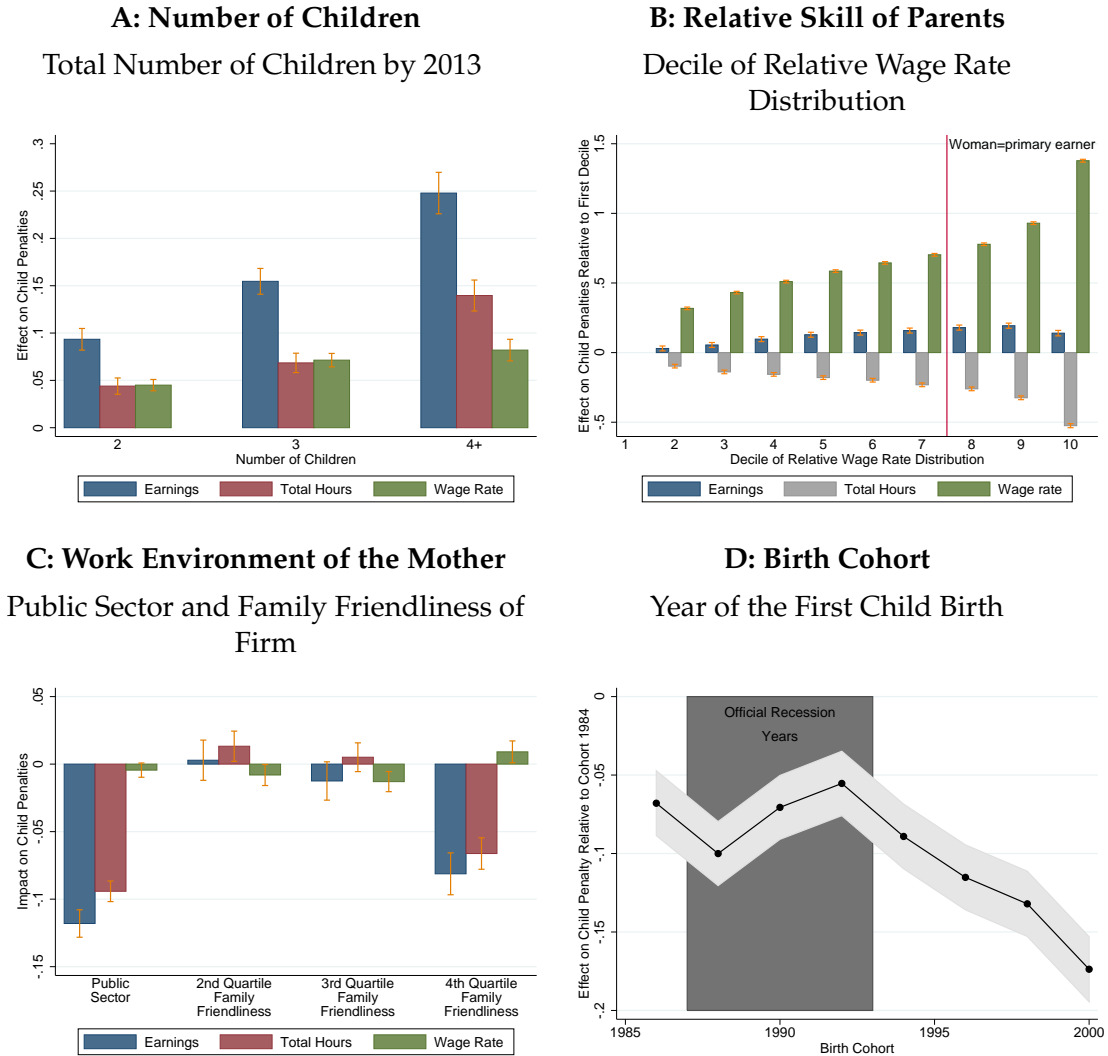
The full set of results on heterogeneity in earnings penalties, hours penalties, and wage rate penalties are presented in appendix tables 3.1-3.3, while a subset of these results are presented in Figure 3.6. The four panels in the figure display coefficient estimates from the specification in column (2) of the appendix tables, which corresponds to the specification shown in equation (3.4).

---

that the full effect of children come in part "as a surprise" to parents. Anticipation effects appear to be relatively limited, while immediately after child birth we see sharp changes in a multitude of dimensions. If the full implications of children come as a surprise, the dimensions that correlate with child penalties may well come as a surprise too. This would mitigate the concern that the heterogeneity we document is mainly driven by selection.

[6]Quantile regressions would have been a natural alternative, but they proved too computationally demanding with so many indicator variables and such a large sample.

Figure 3.6: Heterogeneity in Child Penalties

A: Number of Children

Total Number of Children by 2013

B: Relative Skill of Parents

Decile of Relative Wage Rate Distribution

C: Work Environment of the Mother

Public Sector and Family Friendliness of Firm

D: Birth Cohort

Year of the First Child Birth

Notes: These panels show coefficient estimates (with robust standard errors) based on specification (3.4) and correspond to the estimates reported in column (2) of the appendix tables 3.1-3.3. The dependent variable in this regression is the family-level penalty defined in equation (3.3). The total number of children is measured as of 2013, which leaves at least a 12-year window to have extra children as we consider first child births between 1985-2001. The relative wage rate of the parents in a family is based on an average over event time -5 to -1. The work environment of the woman (sector and firm) is measured at event time -1.

In panel A of Figure 3.6, we plot the coefficient estimates of the effect on child penalties of the total number of children that the woman has. The reference is having one child only. Consistent with the evidence in Figure 3.9, earnings penalties increase steadily with the number of children and in fact the size of the effect is roughly unaffected (slightly smaller) by the rich set of covariates. Interestingly, the earnings effect

is driven by both hours penalties and wage rate penalties, and in roughly equal proportions. These results confirm that larger families go hand-in-hand with lower career trajectories for women relative to men.

In panel B, we investigate how child penalties correlate with the relative wage rates that the mother and father had prior to child birth. Specifically, we compute for each family the average wage rate of the woman and the man w_i^w, w_i^m between event years -5 and -1, and rank families by deciles of the distribution of relative wage rates w_i^w/w_i^m.[7] Because we are including a rich set of controls for relative years of education, profession, and experience, the relative wage rate is meant to capture relative earnings abilities within the family. The reference category in the regression is the first decile of the relative wage rate distribution, i.e. families in which the earnings ability of the woman relative to the man is the smallest. The results show that the earnings penalty is strongly *increasing* in the earnings ability of the mother relative to the father. The 30% of women who are "primary earners" (according to their wage rate) prior to giving birth face an earnings penalty up to 10pp larger than women who are secondary earners. The larger earnings penalties on high-skill women are driven by much larger wage rate penalties, while there is an offsetting effect from smaller hours penalties. The appendix tables 3.1-3.3 show that these qualitative results are very robust to different sample definitions and specifications.[8]

These descriptive findings show that women at the top of the distribution face the hardest trade-offs between career and family, broadly consistent with evidence from the US by Wilde et al. (2010) and Bertrand et al. (2010). The fact that the *hours* penalty is declining in the relative skill of the woman is consistent with the presence of a comparative advantage channel, but this effect is swamped by other factors that create larger wage rate and earnings penalties. The evidence in section 3.4 on the effect of motherhood on occupational rank, sector, and firm provides insight into how high-skill women who face increased demands at home may reduce the intensity of their careers in a way that creates large earnings penalties.

Panel C shows that child penalties correlate strongly with the work environment of the woman at the time of having her first child (at event time -1). Working in the public sector is associated with a 10pp smaller penalty in earnings, driven mostly by a lower penalty in total hours. Furthermore, working in a relatively family-friendly

---

[7]By averaging individual wage rates over five years, we mitigate the potential concern about short-term mean reversion in wage rates at the individual level.

[8]In tables 3.1-3.3, we also investigate how child penalties correlate with relative years of education (prebirth) within the family. Conditional on all the other controls, relative years education within the family have only a modest and marginally significant effect on child penalties.

firm—proxied as above by the fraction of women with young children in the firm—is also associated with substantially smaller earnings and hours penalties. Note here that selection on unobservables is most likely to go against the effects we find. In particular, women with relatively strong preferences for family over career (those who would face larger child penalties other things equal) are more likely to select into family-friendly work environments ex ante, which by itself would increase observed child penalties in such environments.

Finally, panel D shows the evolution of the child penalty in earnings across birth cohorts, relative to the 1985 birth cohort and controlling for our rich set of covariates. While the conditional earnings penalty is overall declining with birth cohort, it also exhibits significant cyclicality. The recession years following the 1987 financial crisis were associated with larger child penalties than during boom periods. This evidence of cyclicality is consistent with the raw evidence on cohort-specific penalties presented in Figure 3.12.

Intergenerational Transmission of Child Penalties

The size and persistence of female child penalties, along with their heterogeneity across skill, are difficult to reconcile with comparative advantage alone. This suggests that the effects are partly driven by preferences over what men and women should do when becoming parents. Indeed, we have shown in Figure 3.2 that views on the appropriate gender roles in families with children are very conservative in all countries. The vast majority of both men and women hold the view that women should not work full-time as long as there are children living at home. This raises the question of where these gendered preferences are coming from? Are they biologically determined or is there a role for environmenal influences? In this section we present a set of findings that speak in favor of environmentally determined gender preferences.

A recent literature discusses the importance of social norms and gender identity in explaining the different labor market outcomes of men and women, although causal testing of these ideas has proved difficult (see Bertrand, 2011). We explore the role of such influences by showing that the female child penalty is strongly related to the work history of her parents, but not to the work history of her partner's parents. Our findings are consistent with the idea that a woman's preferences over family and career are shaped during her childhood. Our analysis is related to Fernandez et al. (2004), but they consider a transmission mechanism between the woman and the parents of the man (where we find no effect). Our approach also differ in that we consider the inter-generational transmission of child penalties—i.e., labor supply changes of women rela-

tive to men around child birth—rather than the intergenerational transmission of labor supply levels. Working with such labor supply changes directly takes care of some of the key omitted variable concerns encountered when working with labor supply levels.

Our analysis leverages the availability of the administrative ATP measure of hours since 1964. Specifically, for each family i we observe the cumulative sum of all recorded ATP hours between 1964-1979 of the mother's mother h_i^{mm} and of the mother's father h_i^{mf}. To capture the relative labor supply of the maternal grandparents, we rank families in deciles of the distribution of the difference $h_i^{mf} - h_i^{mm}$. We do the same for the relative labor supply of the paternal grandparents $h_i^{ff} - h_i^{fm}$.

Panel A and B of Figure 3.7 start by plotting the average family-level child penalties in earnings and hours by deciles of the relative labor supply of the maternal grandparent's (relative to first decile). They reveal a very strong and significant correlation between the child penalty in earnings and hours on the mother and the relative labor supply of the maternal grandparents. In families in the top decile of the relative labor supply of maternal grandparents, i.e. where the grandmother worked very little compared to the grandfather, the mother pays an earnings penalty that is 10pp larger and an hours penalty that is 7pp larger than in families from the bottom decile.

Figure 3.7: Intergenerational Transmission of Child Penalties (Maternal)

A: Earnings Penalties (Unconditional)

Effect of Maternal Grandparents' Work History
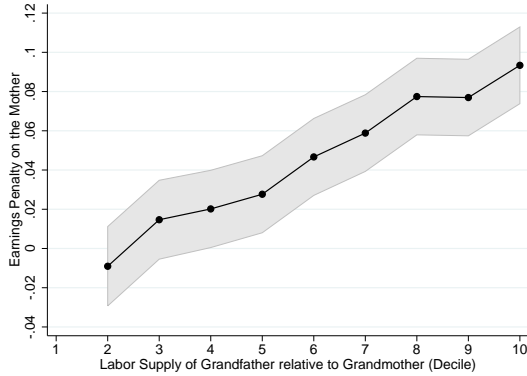
B: Hours Penalties (Unconditional)

Effect of Maternal Grandparents' Work History
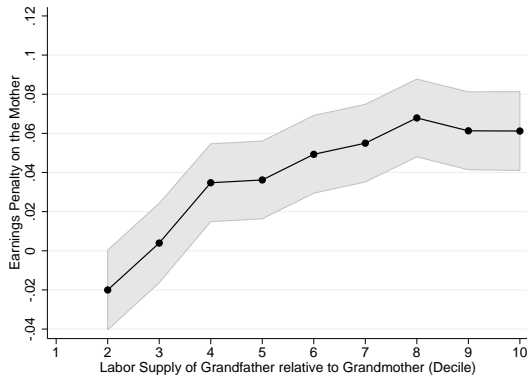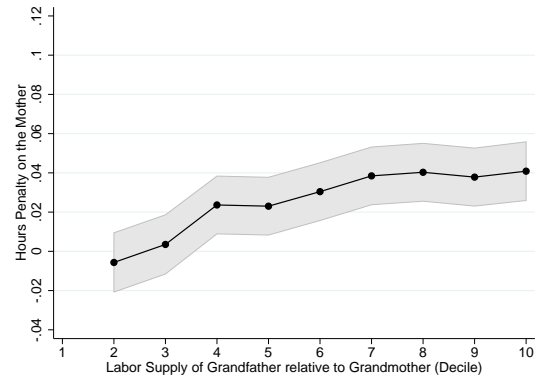
C: Earnings Penalties (Conditional)

Effect of Maternal Grandparents' Work History

D: Hours Penalties (Conditional)

Effect of Maternal Grandparents' Work History

Notes: The panels show the correlation between family-level child penalties (as defined in equation (3.3)) and the work history of the maternal grandparents, specifically total hours worked by the maternal grandfather relative to the maternal grandmother. Relative hours worked by the grandparents is based on the cumulative ATP contributions over the period 1964-79. Panels A-B show correlations without any controls, while panels C-D show correlations conditional on a rich set of socio-economic characteristics of the (maternal) grandparents, including their birth cohort, education level/fields, and within-generation wealth rank.

There are two potential concerns in the interpretation of these (unconditional) intergenerational correlations. First, rather than reflecting a transmission of gendered preferences, they could be driven by other transmissible characteristics of the maternal grandparents that correlate with child penalties in every generation. The prime candi-

dates are education and income/wealth levels. Second, while our cumulative measure of hours for grandparents between 1964-1979 has the advantage of minimizing noise compared to snapshots of cross-sectional data, part of the variation in this measure might be driven by grandparents in different cohorts being observed at different points in the lifecycle.

To investigate the robustness of our findings to these potential concerns, we adopt a methodology similar to section 3.5 by exploring the effect of the relative labor supply of grandparents, conditional on a rich set of socio-demographic controls for the maternal grandparents. We include in the regression a complete set of dummies for the birth cohort of the grandmother and the grandfather, which will control for observing grandparents at different points in the lifecycle. We also include detailed controls for the education of both the grandfather and the grandmother, with twenty-two dummies for each grandparent that capture education level and field. We finally control for the wealth level of the grandparents. We use the average net wealth of the grandfather over the years 1980-90 and control for deciles of the within-generation wealth rank of the grandfather.[9] The results, displayed in panels C and D of Figure 3.7, confirm that these intergenerational correlations are very robust to controlling for the characteristics of grandparents.

Interestingly, when we replicate the same exercice using paternal grandparents in appendix Figure 3.13, we find no significant relationship between grandparents' work history and female child penalties. This differential pattern is interesting for two reasons. First, it is consistent with child penalties being driven by female gender identity formed during her childhood, as opposed to child penalties being driven by male gender identity formed during his childhood. This makes sense when considering the event study evidence presented above. The effects in those event studies are entirely driven by sharp changes in the behavior of women, with men being essentially unaffected by parenthood, and so it makes sense that the underlying reason for such behaviors should be sought in the woman's childhood environment, not the man's. Second, the differential pattern of intergenerational correlations between maternal and paternal grandparents rules out the treat from omitted variables that are present in both sets of grandparents, i.e. family background variables (not captured by the detailed education and wealth controls) that both parents have and which affect child penalties.

Overall, these findings are pointing to an influence of nurture in the formation of gender attitudes regarding children and career. They suggest the existence of heteroge-

---

[9]This comprehensive wealth information on the universe of Danish taxpayers was collected for the purpose of the wealth tax.

neous preferences or norms across families regarding gender roles in the labor market and in the home, preferences that are transmitted in the family across generations, and differentially between daughters and sons. Such intergenerational transmission mechanisms may play an important role in the persistence of child penalties on women over time.

Parental Leave Policy: Part of the Solution or Part of the Problem?

As described in section 3.2 Denmark has been at the forefront of implementing family policies that offer job-protected and paid maternity and parental leave. An important question is whether such policies are helpful for reducing gender inequality or if they are counterproductive. While an evaluation of the long-run effect of such policies on gender inequality is outside the scope of this paper, in this section we provide evidence that the take-up of gender-neutral family leave policies is extremely unequal across gender. Our findings here are consistent with the evidence above that preferences over family vs career are strongly gendered, and this raises potential concerns with respect to the effect of family leave policies on the gender gap.

To study this question we consider a 2002 reform, which extended parental leave from 26 weeks to 52 weeks. Apart from the introduction of 4 weeks of pregnancy leave, the rest of the expansion took the form of parental leave that could be freely allocated between the two parents. Hence the reform did not directly change the relative prices of leave taken for women relative to men. However, if preferences over family vs career are strongly gendered as we have shown above, then a gender-neutral provision of parental leave may effectively subsidize leave taken by woman and this could potentially excaserbate gender inequality.

Our findings are presented in Figure 3.8. Panel A shows the amount of parental leave taken by men and women around the 2002 reform, relative to 2001. We see a very sharp increase in parental leave by women of about 15 weeks on average and only a tiny increase in parental leave by men. The difference-in-differences estimate of the effect on parental leave for women relative to men is 14.4 weeks

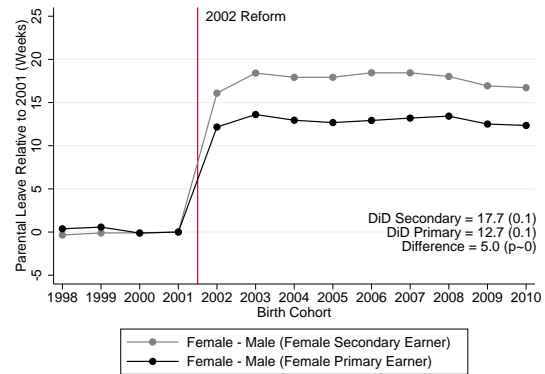Figure 3.8: Effects of Job-Protected Parental Leave

A: Effect of Parental Leave Extension

Parental Leave of Males and Females
around Reform
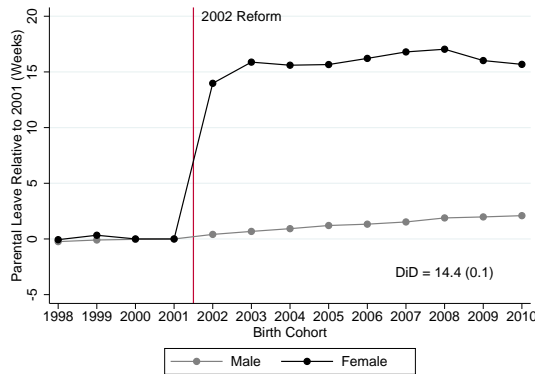
B: Effect of Parental Leave Extension by Skill

Parental Leave of Females − Males by
Relative Wage Rates

Notes: Panel A shows weeks on parental leave (relative to the level in 2001) for males and females, respectively, around the 2002 reform, which extended the total length of job-protected and paid parental leave from 24 to 50 weeks. Panel B considers the difference in parental leave between males and females (still relative to 2001) and splits the sample into two groups depending on whether the female has a lower or higher wage rate than the male in the year prior to child birth (labelled as the female being the secondary and primary earner, respectively, in the panel).

Panel B considers heterogeneity in treatment effects by skill. This panel plots the difference in parental leave between men and women (still relative to 2001) by relative skill within the family prior to child birth. Specifically, we split the sample into two groups depending on whether the woman has a lower wage rate than her partner prior to giving birth (in which case we label her the "secondary earner") or a higher wage rate than her partner prior to giving birth (in which case we label her the "primary earner"). Two key findings emerge from the figure. First, the increase in parental leave by women relative to men is smaller when she is the primary earner prior to birth, consistent with a comparative advantage channel being in operation. Second, while the differential treatment effect across relative skill levels is clear and statistically significant, it is economically very small. Even in families where the woman is the higher-skill person, she takes an extra 12.7 weeks of parental leave compared to the man when it is offered to both of them. Note that the additional leave comes on top of the 26 weeks already offered prior to the reform, and so the effects are not easily explained by comparative advantage in infant care. In other words, the effect of comparative advantage in the market place is there, but it is swamped by other factors that make women the prime caregivers to children.

While these findings do not identify the long-run effect of family policies on gender inequality (or welfare), they do raise some questions regarding the desirability of gender-neutral family leave policies in a world where preferences are extremely gendered (possibly because of environmental preferences as analyzed above). Future work will investigate further the link between these policy effects and the long-run child penalties and gender gaps analyzed above.

3.6 Conclusion

Despite considerable gender convergence over time, substantial gender inequality persists in all countries. Using full-population administrative data from Denmark and a quasi-experimental event study approach, we show that most of the remaining gender inequality can be attributed to the dynamic effects of childrearing. We have presented three main sets of results.

First, we have identified large child penalties on the careers of women relative to men in a broad range of dimensions. The female child penalty in earnings is around 20% even 20 years after the birth of the first child. Underlying this effect, we find sharp effects of children on labor force participation, hours of work, wage rates, occupation, sector, and firm choices. Together, these findings provide a quite complete picture of the behavioral margins that adjust in response to parenthood and how strongly gendered these margins are.

Second, by estimating child penalties in earnings in each year after birth and for each birth cohort, we have decomposed gender inequality into what can be attributed to children and and what is driven by other factors (such as human capital or discrimination). We have shown that the fraction of total gender inequality that can be explained by children has increased dramatically over time, from 30% in 1980 to 80% in 2011. In other words, to a first approximation, the remaining gender inequality is all about children. This has important implications for future work on gender inequality, which should focus on understanding what drives gender outcomes in relation to parenthood.

Third, we have provided evidence in favor of environmental influences in the formation of gendered preferences over family vs career. In particular, we have shown that the female child penalty is strongly related to the work history of her parents: for example, women who grow up in traditional families with a male breadwinner and a female homemaker end up paying much larger career penalties when they become mothers themselves. At the same time, the female child penalty is unrelated to the work history of her partner's parents. Overall, these findings are consistent with the

notion that child penalties are influenced by female gender identity formed during her childhood, as opposed to child penalties being driven by male gender identity formed during his childhood. We have argued that the differential pattern between maternal and paternal grandparents is consistent with the fact that the career effects of children are entirely driven by sharp changes in the behavior of women, not men, and so it is natural that the underlying reason for such behaviors should be sought in women's childhood environment.

While these patterns of intergenerational transmission of child penalties are suggestive of environmental influences on the strongly gendered preferences over family vs career, conclusive causal testing is obviously challenging due to the fact that family background is not randomly allocated. Future work should continue to study the underlying mechanisms as this will ultimately determine the welfare and policy implications of the patterns we have uncovered here.

References

Adda,   Jerome, Dustmann, Christian and Stevens, Katrien. (2011). "The Career Costs of Children". Working Paper, November 2011.

Aguero,   Jorge M., and Marks, Mindy S. (2008). "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks." American Economic Review, Papers and Proceedings, 98, 500–504.

Altonji,   Joseph G., and Blank, Rebecca M. (1999). "Race and Gender in the Labor Market." Chap. 48, pages 3143–3259 of: Ashenfelter, O., & Card, D. (eds), Handbook of Labor Economics, vol.3. North Holland: Elsevier Science Publishers.

Angrist,   Joshua D., and Evans, William N. (1998). "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." American Economic Review, 88, 450–477.

Bertrand, Marianne. (2011). "New Perspectives on Gender." Chap. 17, pages 1543–1590 of: Ashenfelter, O., & Card, D. (eds), Handbook of Labor Economics, vol.4b. North Holland: Elsevier Science Publishers.

Bertrand, Marianne, Goldin, Claudia, and Katz, Lawrence F. (2010). "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." American Economic Journal: Applied Economics, 2, 228–255.

Bertrand, Marianne, Kamenica, Emir, and Pan, Jessica. (2013). "Gender Identity and Relative Income Within Households." Working Paper, October 2013.

Bronars,   Stephen G., and Grogger, Jeff. (1994). "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." American Economic Review, 84, 1141–1156.

Browning, Martin. (1992). "Children and Household Economic Behavior." Journal of Economic Literature, 30, 1434–1475.

Fernandez, Raquel, Fogli, Alessandra, and Olivetti, Claudia. (2004). "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." Quarterly Journal of Economics, 119, 1249–1299.

Goldin,   Claudia. (2014). "A Grand Gender Convergence: Its Last Chapter." American Economic Review, 104, 1091–1119.

Goldin,   Claudia, & Katz, Lawrence F. (2008). The Race Between Education and Technology. Harvard University Press.

Goldin,   Claudia, Katz, Lawrence F., and Kuziemko, Ilyana. (2006). "The Homecoming of American College Women: The Reversal of the College Gender Gap." Journal of Economic Perspectives, 20, 133–156.

Hotz,   V. Joseph, McElroy, Susan Williams, and Sanders, Seth G. (2005). "Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment." Journal of Human Resources, 40, 683–715.

Kleven,   Henrik. (2014). "How Can Scandinavians Tax So Much?" Journal of Economic Perspectives, 28, 77–98.

Nielsen,   Helena Skyt, Simonsen, Marianne, and Verner, Mette. (2004). "Does the Gap in Family-Friendly Policies Drive the Family Gap?" Scandinavian Journal of Economics, 106, 721–744.

Paull,   Gillian. (2008). Children and Women's Hours of Work. Economic Journal, 118, F8–F27.

Sims,   Christopher A. (1972). "Money, Income, and Causality." American Economic Review, 62, 540– 552.

Waldfogel, Jane. (1998). "Understanding the Family Gap in Pay forWomen with Children." Journal of Economic Perspectives, 12, 137–156.

Wilde,   Elizabeth Ty, Batchelder, Lily, and Ellwood, David T. (2010). "The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Differing Skill Levels." NBER Working Paper No. 16582, December 2010.
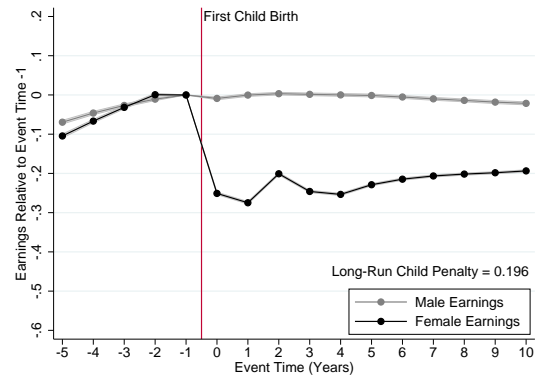
Appendix 3.A: Additional figures

Figure 3.9: The Child Penalty in Earnings by Number of Children
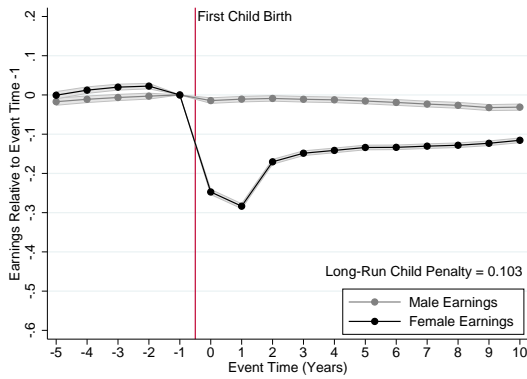
A: One-Child Mothers　　　　　　　　**B: Two-Child Mothers**

C: Three-Child Mothers　　　　　　　**D: Four-Child Mothers**
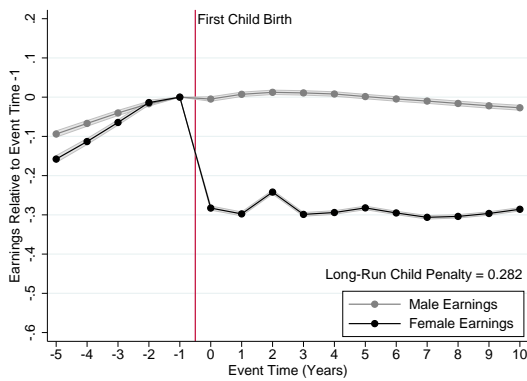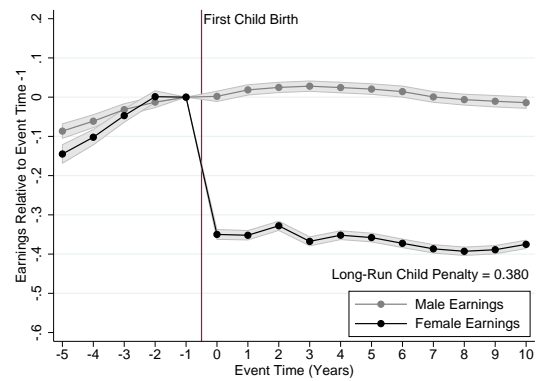
Notes: The panels show estimated coefficients on the event time dummies in equation (3.1) as a fraction of the predicted outcome when omitting the contribution from the event dummies (i.e., $P_t^k = \hat{\alpha}_t^k / E[\hat{Y}_{ist}^k | t]$ defined in section 3.3). The focus is on the earnings penalty as in Figure 3.3A, splitting the sample by the woman's total number of children as of 2013 (1, 2, 3, or 4 children). The long-run child penalty is calculated as the average effect for women relative to men over the period from event time 5 to 10.

Figure 3.10: Female Child Penalties for January Births

A: Child Penalty in Earnings

January Births

B: Child Penalty in Hours Worked

January Births

C: Child Penalty in Participation Rates

January Births

D: Child Penalty in Wage Rates

January Births
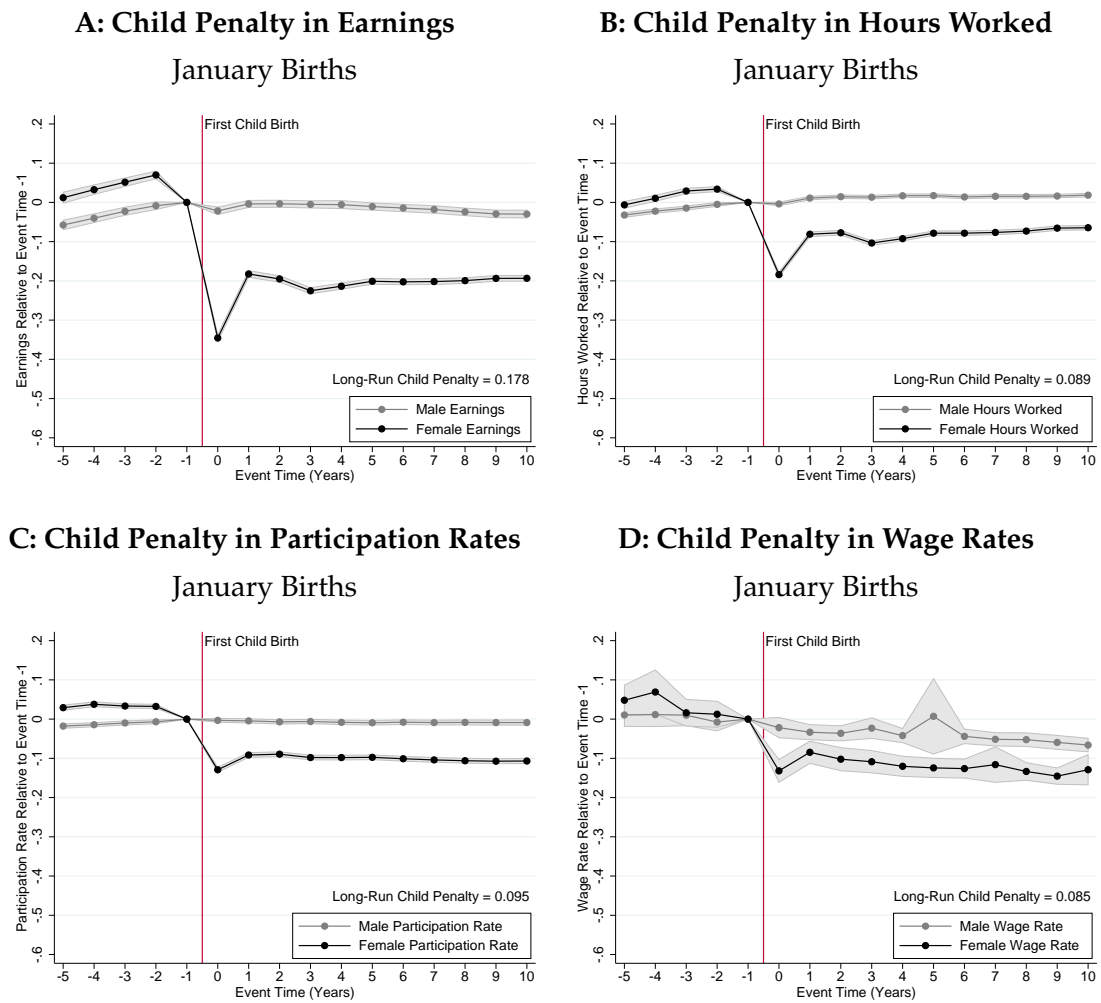
Notes: This figure is constructed exactly as Figure 3.3 except that the estimation is run only for those who have their first child in January.

Figure 3.11: Female Child Penalties Over a 20-Year Horizon

A: Child Penalty in Earnings

20 Years After Child Birth

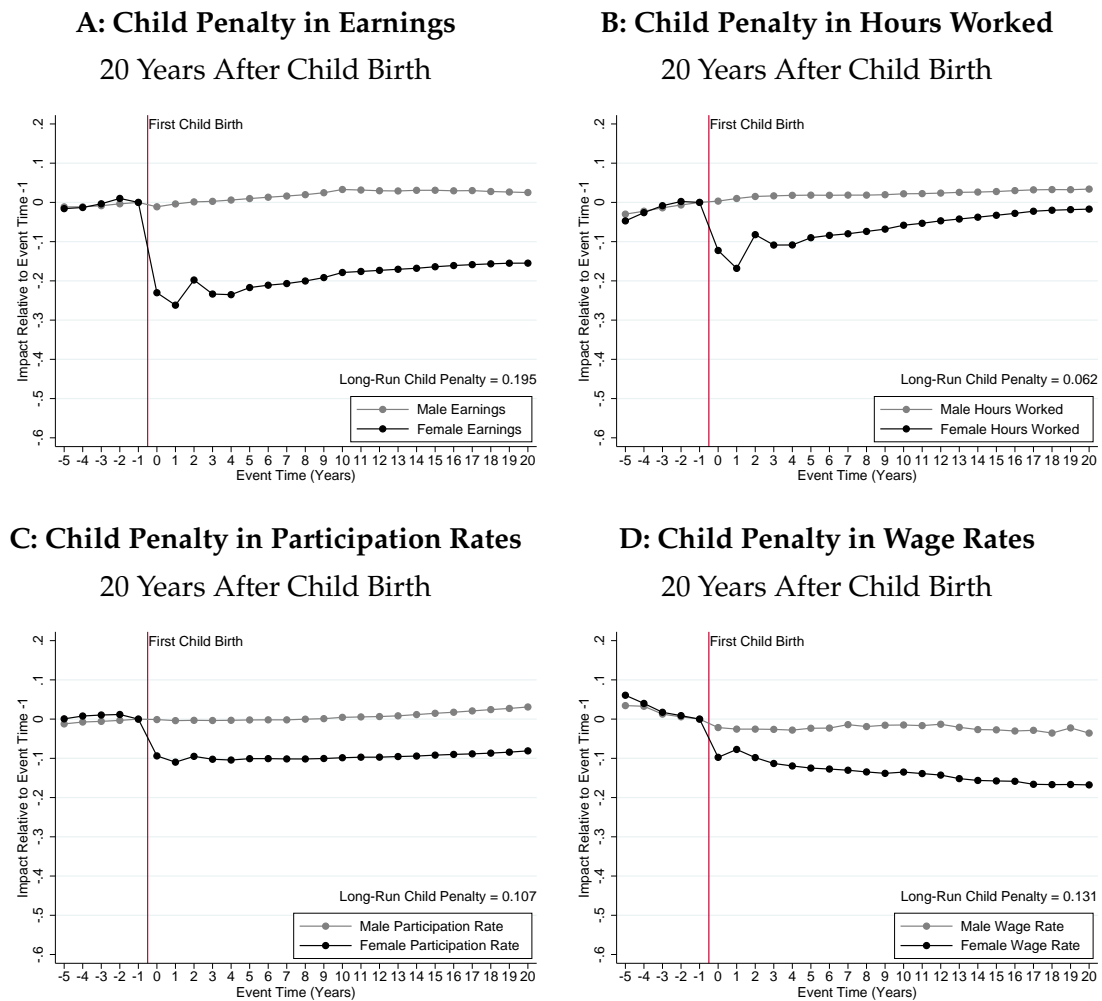

B: Child Penalty in Hours Worked

20 Years After Child Birth

C: Child Penalty in Participation Rates

20 Years After Child Birth

D: Child Penalty in Wage Rates

20 Years After Child Birth

Notes: The panels are based on the regression specification in (3.5) and a semi-balanced panel running from event time -5 to 20. Only parents who have their first child between 1985-1991 are observed in all 26 event years; parents who have their first child after 1991 are kept in the sample for the maximum possible number of event years. Given the long event time window and the interaction with year dummies, the specification includes more than 800 dummies and in order to reduce the computational requirements we aggregate the micro data into year-age-event time cells and run the regressions on this data weighted by the number of observations in each cell. The panels show average estimated coefficients across birth cohorts 1985-1991 (the cohorts observed over the entire 26-year event time window) as a fraction of the average predicted outcome when omitting the contribution from the event dummies, similar to the estimates reported in Figure 3.3.

Figure 3.12: Earnings Penalties by Birth Cohort and Extrapolation

A: Average Earnings Penalty From Event Time 0-4

No Significant Time Trend
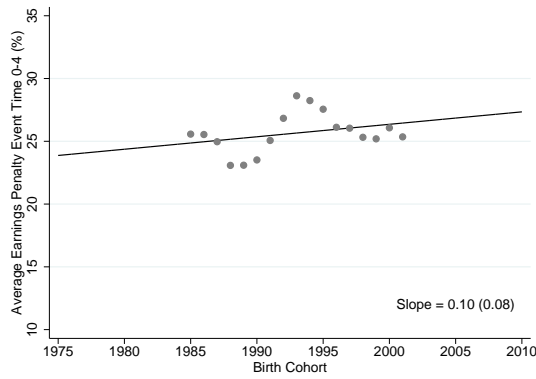
B: Average Earnings Penalty From Event Time 5-10

Significant Negative Time Trend
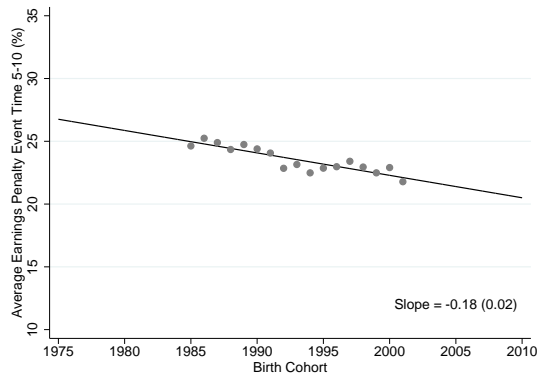
Notes: The panels show cohort-specific child penalties estimated based on (3.5). Panel A shows the average penalties over event time 0-4, while Panel B shows the average penalties over event time 5-10. Each panel also shows a linear OLS fit and the estimated slope coefficient (with standard errors in parentheses). There is no statistically significant trend in the short-run penalty, but there is a statistically significant downward trend in the long-run penalty.

Figure 3.13: Intergenerational Transmission of Child Penalties (Paternal)

A: Earnings Penalties (Unconditional)

Effect of Paternal Grandparents' Work
History

B: Hours Penalties (Unconditional)

Effect of Paternal Grandparents' Work
History

C: Earnings Penalties (Conditional)

Effect of Paternal Grandparents' Work
History
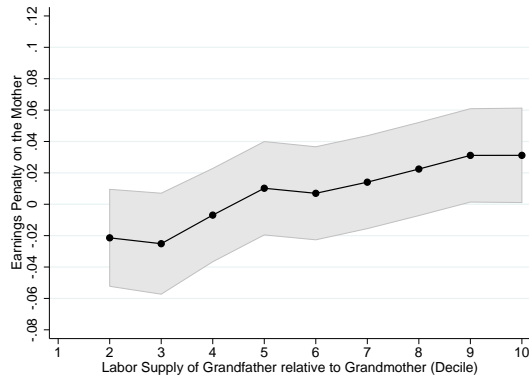
D: Hours Penalties (Conditional)

Effect of Paternal Grandparents' Work
History
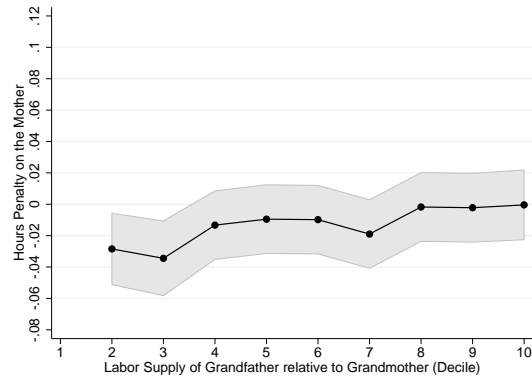
Notes: The panels show the correlation between family-level child penalties (as defined in equation (3.3)) and the work history of the paternal grandparents, specifically total hours worked by the paternal grandfather relative to the paternal grandmother. Relative hours worked by the grandparents is based on the cumulative ATP contributions over the period 1964-79. Panels A-B show correlations without any controls, while panels C-D show correlations conditional on a rich set of socio-economic characteristics of the (paternal) grandparents, including their birth cohort, education level/fields, and within-generation wealth rank.

Table 3.1: Heterogeneity in Family-Level Earnings Penalties

	(1) OLS	(2) OLS	(3) OLS Conditional on graduation before birth	(4) OLS Conditional on working full time before birth	(5) Robust regression
Number of children					
Two children	0.0898*** (0.00572)	0.0946*** (0.00557)	0.103*** (0.00567)	0.104*** (0.00560)	0.0970*** (0.00530)
Three children	0.155*** (0.00681)	0.170*** (0.00664)	0.196*** (0.00699)	0.209*** (0.00695)	0.181*** (0.00632)
Four children and +	0.259*** (0.0108)	0.269*** (0.0105)	0.315*** (0.0120)	0.343*** (0.0125)	0.293*** (0.0100)
Work environment: Family friendliness and sector					
Quartile 2	-0.0000305 (0.00738)	-0.000772 (0.00719)	-0.0184* (0.00750)	-0.0302*** (0.00749)	-0.0105 (0.00684)
Quartile 3	-0.0134 (0.00702)	-0.0143* (0.00684)	-0.0413*** (0.00716)	-0.0527*** (0.00717)	-0.0293*** (0.00651)
Quartile 4	-0.0787*** (0.00768)	-0.0624*** (0.00748)	-0.0600*** (0.00792)	-0.0660*** (0.00797)	-0.0705*** (0.00712)
Public sector	-0.112*** (0.00508)	-0.0649*** (0.00500)	-0.0345*** (0.00530)	-0.0354*** (0.00528)	-0.0701*** (0.00476)
Relative years of education					
Quartile 2	-0.00963 (0.00871)	0.00961 (0.00850)	-0.00164 (0.00904)	-0.00813 (0.00888)	0.0106 (0.00809)
Quartile 3	-0.0314*** (0.00910)	-0.00560 (0.00889)	-0.00762 (0.00935)	-0.00930 (0.00920)	0.00461 (0.00846)
Quartile 4	-0.0815*** (0.0146)	-0.0375** (0.0142)	-0.0396** (0.0149)	-0.0425** (0.0146)	-0.0242 (0.0135)
Deciles of relative wage rate distribution					
Decile 2	0.0299*** (0.00868)	0.00849 (0.00845)	0.0409*** (0.00904)	0.0439*** (0.00905)	0.0225** (0.00804)
Decile 3	0.0770*** (0.00867)	0.0306*** (0.00846)	0.0842*** (0.00909)	0.0848*** (0.00909)	0.0453*** (0.00805)
Decile 4	0.124*** (0.00870)	0.0528*** (0.00850)	0.115*** (0.00913)	0.116*** (0.00913)	0.0668*** (0.00809)
Decile 5	0.145*** (0.00875)	0.0525*** (0.00857)	0.128*** (0.00917)	0.125*** (0.00919)	0.0702*** (0.00816)
Decile 6	0.176*** (0.00880)	0.0681*** (0.00864)	0.153*** (0.00920)	0.153*** (0.00923)	0.0836*** (0.00822)
Decile 7	0.197*** (0.00886)	0.0781*** (0.00871)	0.165*** (0.00926)	0.165*** (0.00931)	0.0965*** (0.00829)
Decile 8	0.194*** (0.00989)	0.0686*** (0.00971)	0.178*** (0.00931)	0.176*** (0.00939)	0.0845*** (0.00924)
Decile 9	0.234*** (0.00914)	0.0912*** (0.00900)	0.214*** (0.00941)	0.224*** (0.00958)	0.113*** (0.00857)
Decile 10	0.175*** (0.00976)	0.0326*** (0.00959)	0.230*** (0.00986)	0.291*** (0.0105)	0.0569*** (0.00913)
Additional controls (full set of dummies)					
Age	×	×	×	×	×
Cohort	×	×	×	×	×
Occupation	×	×	×	×	×
Deciles of household earnings	×	×	×	×	×
Experience		×	×	×	×
N	257469	257469	185872	159595	257469

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.2: Heterogeneity in Family-Level Hours Penalties

	(1) OLS	(2) OLS	(3) OLS Conditional on graduation before birth	(4) OLS Conditional on working full time before birth	(5) Robust regression
Number of children					
Two children	0.0379***	0.0389***	0.0419***	0.0435***	0.0383***
	(0.00422)	(0.00413)	(0.00408)	(0.00378)	(0.00309)
Three children	0.0627***	0.0706***	0.0851***	0.0963***	0.0780***
	(0.00500)	(0.00490)	(0.00503)	(0.00468)	(0.00366)
Four children and +	0.137***	0.141***	0.161***	0.184***	0.154***
	(0.00788)	(0.00770)	(0.00859)	(0.00837)	(0.00576)
Work environment: Family friendliness and sector					
Quartile 2	0.0117*	0.0106*	-0.00579	-0.0217***	-0.00628
	(0.00541)	(0.00529)	(0.00539)	(0.00504)	(0.00396)
Quartile 3	0.00319	0.00191	-0.0240***	-0.0366***	-0.0204***
	(0.00515)	(0.00504)	(0.00515)	(0.00483)	(0.00377)
Quartile 4	-0.0668***	-0.0569***	-0.0503***	-0.0544***	-0.0583***
	(0.00564)	(0.00551)	(0.00569)	(0.00537)	(0.00412)
Public Sector	-0.0907***	-0.0573***	-0.0324***	-0.0303***	-0.0510***
	(0.00372)	(0.00368)	(0.00381)	(0.00355)	(0.00275)
Relative years of education					
Quartile 2	-0.0000864	0.0108	0.0109	0.00391	0.0121**
	(0.00635)	(0.00622)	(0.00647)	(0.00594)	(0.00465)
Quartile 3	-0.0175**	-0.00218	0.00955	0.00402	0.00979*
	(0.00665)	(0.00652)	(0.00669)	(0.00617)	(0.00488)
Quartile 4	-0.0339**	-0.00689	0.00183	-0.00379	0.00302
	(0.0106)	(0.0104)	(0.0107)	(0.00975)	(0.00778)
Deciles of relative wage rate distribution					
Decile 2	-0.0986***	-0.114***	-0.0586***	-0.0535***	-0.0771***
	(0.00636)	(0.00622)	(0.00651)	(0.00608)	(0.00465)
Decile 3	-0.131***	-0.163***	-0.0641***	-0.0602***	-0.120***
	(0.00636)	(0.00623)	(0.00654)	(0.00611)	(0.00466)
Decile 4	-0.149***	-0.198***	-0.0668***	-0.0624***	-0.144***
	(0.00638)	(0.00626)	(0.00657)	(0.00614)	(0.00468)
Decile 5	-0.178***	-0.242***	-0.0746***	-0.0698***	-0.176***
	(0.00642)	(0.00631)	(0.00659)	(0.00617)	(0.00472)
Decile 6	-0.195***	-0.268***	-0.0794***	-0.0716***	-0.197***
	(0.00645)	(0.00635)	(0.00663)	(0.00621)	(0.00476)
Decile 7	-0.223***	-0.304***	-0.0873***	-0.0773***	-0.221***
	(0.00649)	(0.00640)	(0.00667)	(0.00626)	(0.00479)
Decile 8	-0.255***	-0.340***	-0.0939***	-0.0832***	-0.248***
	(0.00733)	(0.00722)	(0.00670)	(0.00631)	(0.00540)
Decile 9	-0.308***	-0.405***	-0.103***	-0.0851***	-0.298***
	(0.00667)	(0.00659)	(0.00677)	(0.00643)	(0.00493)
Decile 10	-0.505***	-0.603***	-0.179***	-0.124***	-0.457***
	(0.00710)	(0.00701)	(0.00706)	(0.00701)	(0.00524)
Additional controls (full set of dummies)					
Age	×	×	×	×	×
Cohort	×	×	×	×	×
Occupation	×	×	×	×	×
Deciles of household earnings	×	×	×	×	×
Experience		×	×	×	×
N	267825	267825	190129	162649	267825

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.3: Heterogeneity in Family-Level Wage Rate Penalties

	(1) OLS	(2) OLS	(3) OLS Conditional on graduation before birth	(4) OLS Conditional on working full time before birth	(5) Robust regression
			Number of children		
Two children	0.0513*** (0.00308)	0.0523*** (0.00308)	0.0518*** (0.00358)	0.0506*** (0.00371)	0.0482*** (0.00256)
Three children	0.0812*** (0.00364)	0.0820*** (0.00364)	0.0840*** (0.00442)	0.0825*** (0.00460)	0.0763*** (0.00302)
Four children and +	0.0939*** (0.00574)	0.0948*** (0.00573)	0.101*** (0.00766)	0.103*** (0.00830)	0.0897*** (0.00477)
			Work environment: Family friendliness and sector		
Quartile 2	-0.00379 (0.00395)	-0.00408 (0.00394)	-0.00632 (0.00476)	-0.00146 (0.00496)	-0.00865** (0.00328)
Quartile 3	-0.00943* (0.00377)	-0.00974** (0.00376)	-0.0110* (0.00454)	-0.00869 (0.00476)	-0.0137*** (0.00312)
Quartile 4	0.0143*** (0.00411)	0.0175*** (0.00410)	0.000989 (0.00502)	-0.00134 (0.00528)	0.00957** (0.00341)
Public Sector	-0.00278 (0.00270)	0.00150 (0.00272)	0.00810* (0.00335)	0.00813* (0.00350)	-0.00307 (0.00226)
			Relative years of education		
Quartile 2	-0.00985* (0.00459)	-0.00585 (0.00459)	-0.0210*** (0.00565)	-0.0186** (0.00583)	-0.00323 (0.00382)
Quartile 3	-0.0192*** (0.00482)	-0.0148** (0.00482)	-0.0222*** (0.00585)	-0.0214*** (0.00605)	-0.0105** (0.00401)
Quartile 4	-0.0191* (0.00767)	-0.0153* (0.00766)	-0.0318*** (0.00932)	-0.0316*** (0.00956)	-0.0114 (0.00637)
			Deciles of relative wage rate distribution		
Decile 2	0.326*** (0.00459)	0.326*** (0.00458)	0.212*** (0.00571)	0.166*** (0.00599)	0.314*** (0.00381)
Decile 3	0.450*** (0.00459)	0.448*** (0.00459)	0.285*** (0.00573)	0.230*** (0.00602)	0.434*** (0.00381)
Decile 4	0.541*** (0.00460)	0.537*** (0.00461)	0.327*** (0.00575)	0.268*** (0.00604)	0.521*** (0.00383)
Decile 5	0.615*** (0.00462)	0.609*** (0.00463)	0.359*** (0.00577)	0.295*** (0.00608)	0.593*** (0.00385)
Decile 6	0.681*** (0.00465)	0.674*** (0.00466)	0.393*** (0.00580)	0.325*** (0.00611)	0.656*** (0.00388)
Decile 7	0.755*** (0.00467)	0.747*** (0.00469)	0.418*** (0.00583)	0.348*** (0.00615)	0.724*** (0.00390)
Decile 8	0.829*** (0.00528)	0.820*** (0.00530)	0.448*** (0.00586)	0.371*** (0.00620)	0.799*** (0.00440)
Decile 9	0.977*** (0.00476)	0.965*** (0.00479)	0.513*** (0.00592)	0.424*** (0.00632)	0.939*** (0.00398)
Decile 10	1.439*** (0.00496)	1.425*** (0.00498)	0.762*** (0.00617)	0.618*** (0.00690)	1.386*** (0.00414)
			Additional controls (full set of dummies)		
Age	×	×	×	×	×
Cohort	×	×	×	×	×
Occupation	×	×	×	×	×
Deciles of household earnings	×	×	×	×	×
Experience		×	×	×	×
N	273878	273878	187153	161089	273878

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$