

UNIVERSITY OF COPENHAGEN



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

Søgaard, Jakob Egholt

Publication date:
2014

Citation for published version (APA):
Søgaard, J. E. (2014). *Labor Supply and Optimization Frictions: Evidence from the Danish student labor market*. Economic Policy Research Unit. Department of Economics, University of Copenhagen. EPRU Working Paper Series

Labor Supply and Optimization Frictions: Evidence from the Danish student labor market^{*}

Jakob Egholt Sogaard[†]
University of Copenhagen

Draft

July 2015

Abstract

In this paper, I investigate the nature of optimization frictions by studying the labor market of Danish students. This particular labor market is an interesting case study as it features a range of special institutional settings that affects students' incentive to earn income and comparing outcomes across these settings effectively allow you to distinguish between different types of frictions. I find that the considered labor market is significantly affected by optimization frictions that mask the bunching at kink points normally associated with a positive labor supply elasticity under standard theory. More concretely I find the dominant optimization frictions to be individuals' inattention about their earnings process, while real adjustment costs and gradual learning appear to be of less importance.

JEL: H21, H24, J22

Keywords: Optimization frictions; labor supply; bunching; inattention; student labor market;

^{*} I thank Henrik Kleven, Claus Thustrup Kreiner, Thomas H. Jørgensen, Anders Munk-Nielsen, Peer Skov and participants at the CAM and EPRU seminars at the University of Copenhagen for valuable comments and suggestion. I am grateful to the Kraka Foundation for providing access to Statistic Denmark's data. The finale version of this paper is based on data available via the servers at the Danish Ministry of Finance.

[†] Email: jes@econ.ku.dk. Homepage: www.sites.google.com/site/jesoegaard.

1 Introduction

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 frictions limit the researcher to identify only bounds on the elasticities. Bounds that in many cases are so wide that they are likely to dwarf many of the econometric issues involved in the identification.

In this paper, 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 reflects 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.¹

The labor market for Danish students represents an interesting case study on optimization frictions for several reasons: 1) students face a sharp kink in their budget set created by the phase out of student benefits, 2) a reform in 2009 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.

Having all of these institutional settings within a single labor market covered by rich register data allows you to distinguish between three 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 gradual learning as the

¹ In other contexts such as e.g. consumption, Chetty et al. (2009) show that salience of taxes can affect demand.

main types of frictions, as these types of frictions would imply 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 thus implicit their target 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 target. This prevents 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 of how to quantify the behavioral responses. This is not a trivial task as the lack of a 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 difference-and-difference estimations.²

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 for which I obtain lower bound estimates around 0.05-0.06

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 3 be related to the actual institutional settings facing Danish students. More concretely, I consider the following three 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 first are standard institutional settings considered in the public finance literature, whereas the 3rd needs some additional explanation.

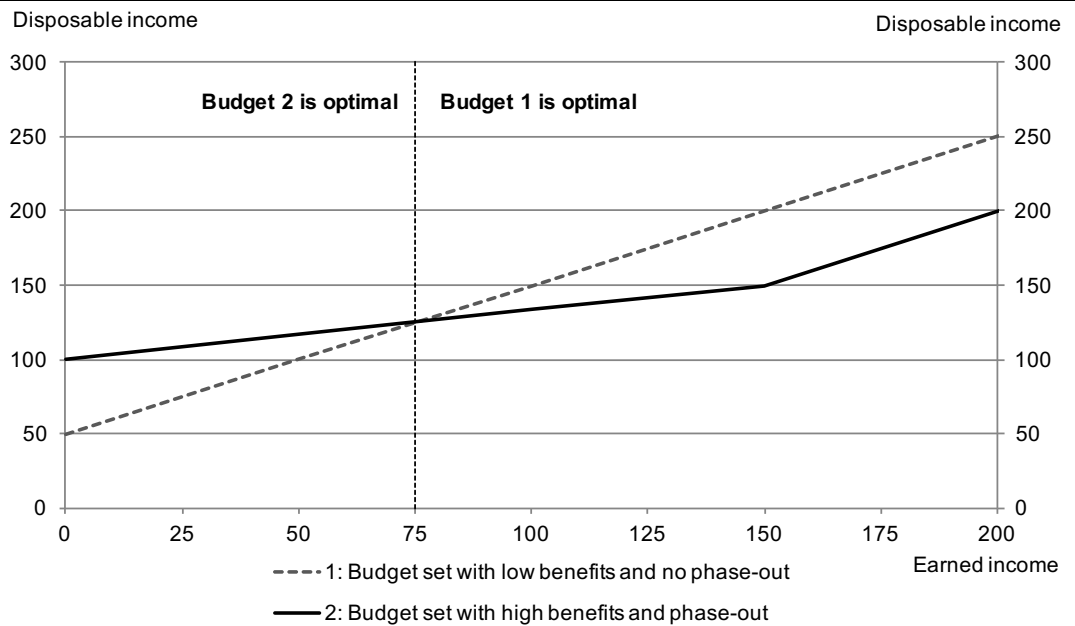
The basic point is that the take up of benefits might only be optimal for individuals in certain earnings intervals. Consider e.g. a stylized benefit system consisting of a lump sum grant that is phased out with earnings according to some schedule. With differences in the level of the

² This method is e.g. used by Feldstein (1995) and Gruber and Saez (2002). See Kleven and Schultz (2014) for an application on Danish data.

lump sum grant and the phase out rate, the optimality of taking up a given benefit may depend on the individual's earned income.

Such a situation is shown in figure 1, which contains two budget sets. Budget set 1 with a low level of benefits, but no phase out, and budget set 2 with a higher level of benefits that is phased out with earnings at a rate of 67 percent. At this phase-out rate net benefits reach 0 at an earned income of 150 and the budget set 2 thus kinks in this point. Comparing these two budgets, we see that the disposable income with budget set 2 is higher than budget set 1 for earnings below 75, while the opposite is the case with earnings above 75.

Figure 1
Illustration of the potential sub-optimality of taking up of higher benefits.



Notes: The figure shows two stylized budget set. One budget set (budget set 1) with a low lump sum grant and a no phase out rate and one (budget set 2) with a higher lump sum grant but a phase out rate of 67 percent. At this phase-out rate net benefits reach 0 at an income of 150. The disposable under the two budget sets are the same at earnings of 75.

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

1. Real adjustment costs 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 distri-

bution at this point. Furthermore, the excess mass would be 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 above 75 in figure 1 should take up the higher benefits, but stay on the low 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 utility gain of finding a better job match does 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 earnings 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 optimally choose 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, e.g. 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 pay 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 learning 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 where 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 a 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 presumption 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 1 and as the table shows each type of optimization frictions lead 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 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
<i>Benchmark:</i>			
No frictions	Clear bunching	Immediate change	Optimal take up
<i>Optimization frictions:</i>			
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

3 Institutional settings: students' incentive to earn income

In this section I present the key features of the Danish student benefit system and relate it to the stylized institutional settings discussed in section 2.³

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 the 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.⁴ If they earn more than this baseline income limit (on a yearly 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.⁵

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.⁶

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.

³ A more detailed description can be found in appendix A.

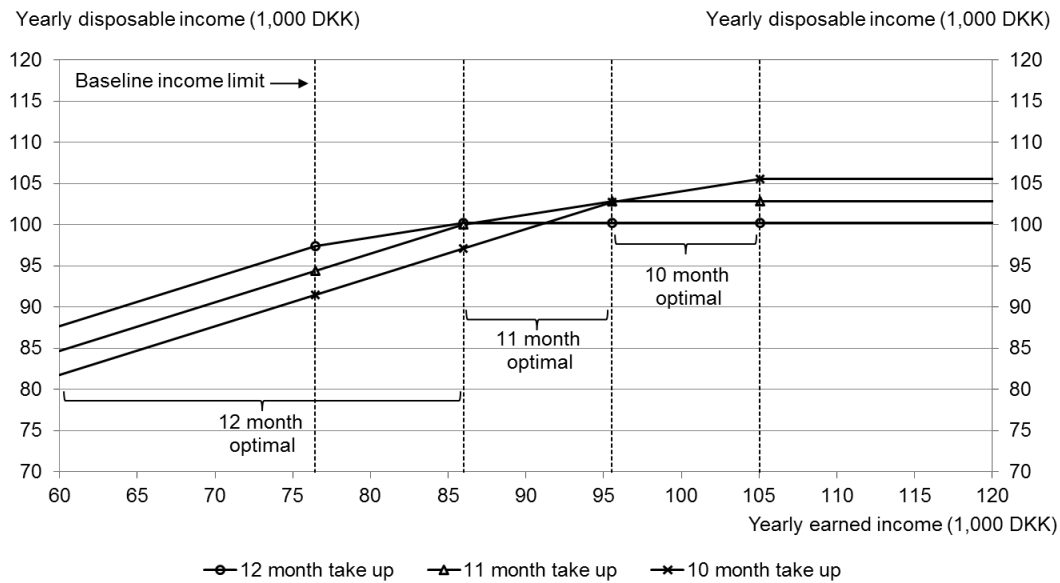
⁴ Income counted against the income limit is called "own income" and includes labor income, transfers other than student benefits and capital income with the exceptions of certain types of stock income. All relevant variables are drawn from detailed register data organized by Statistics Denmark (DST) covering the entire Danish population. A more detailed description of these registers and the variables used can be found in appendix B.

⁵ 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.

⁶ 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 paper and are discussed in section 4 and 6 below.

This problem corresponds to the problem of optimal take up of benefits described in section 2, where 12 months of benefits are optimal for students earning up to 86,000 DKK annually. 11 months are optimal for students with income between 86,000 and 95,500 DKK. For students earning extra 9,500 DKK 10 months is optimal etc., as illustrated in figure 2.

Figure 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. 6 DKK ≈ 1 USD.

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

However, the switch different 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 has remained largely unchanged through the period 2004-2011, which is considered in this analysis, except from a reform in 2009 that increased the baseline income limit by 25 percent for students enrolled in tertiary educations, while leaving it un-

changed for lower levels of education, cf. table 2. At the same time the phase out rate for tertiary students was also increased from 52 to 62 percent and thus causing an increase in the effective marginal tax rate from 72 to 78 percent.

Table 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 educations 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).

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.

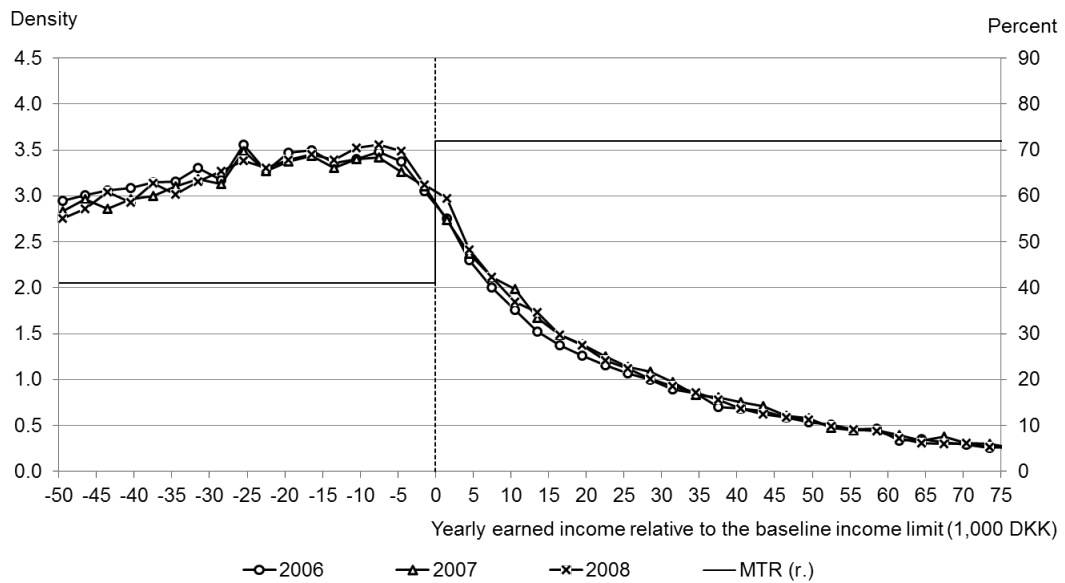
4 Graphical evidence on labor supply responses and optimization frictions

In section 3, I linked the specific features of the Danish student benefit system to the stylized institutional settings listed in section 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.

4.1 Evidence from bunching at the kink point

Figure 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.

Figure 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 $MTR = 1 - (1-t) \cdot (1-q)$, 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

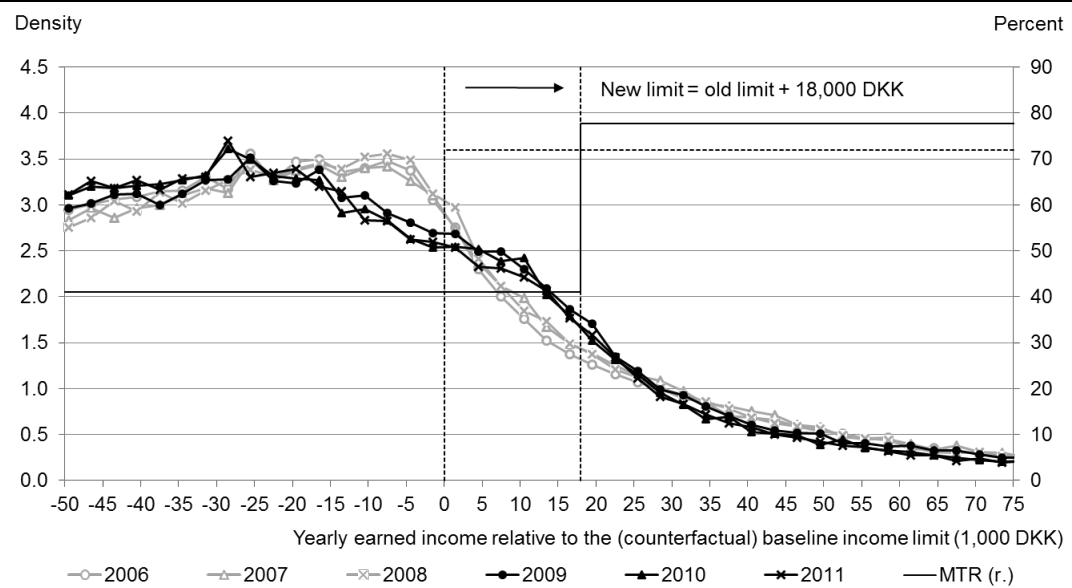
The figure shows that the earnings distribution was very stable during the 3 years prior to the reform and with no clear sign of excess mass around the kink point. In a frictionless world this would imply that the labor supply elasticity was negligible, but from the cross sectional evidence alone – which 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.

4.2 Evidence from the 2009 reform

When comparing the pre-reform earnings distribution with the distributions after the 2009 reform, we see in figure 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.⁷

⁷ 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.

Figure 4
The earnings distribution for tertiary students before and after the 2009 reform



Notes: See notes to figure 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.

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.

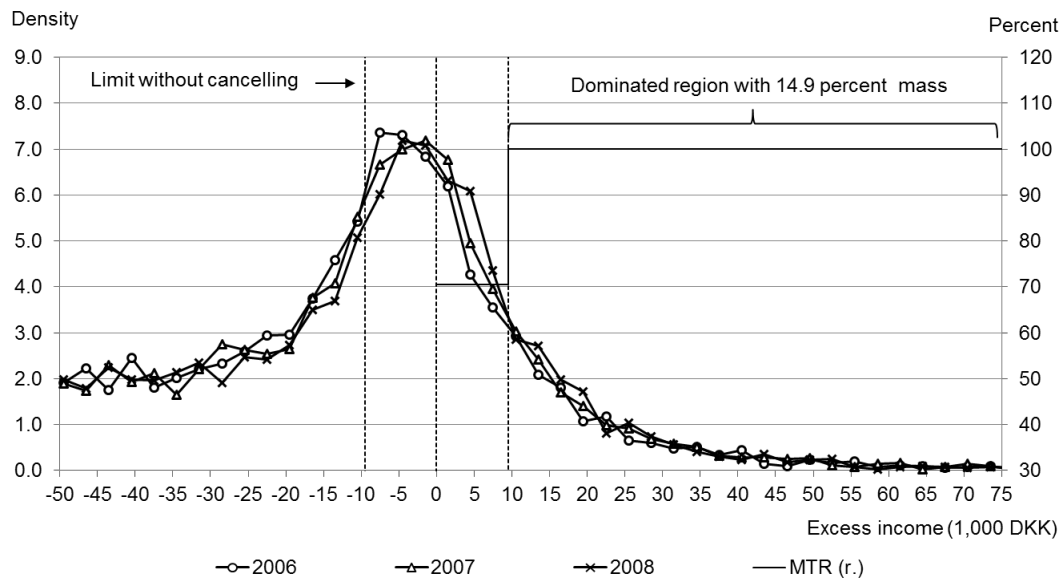
It may finally be noted that the “excess mass” revealed by the shift in the distribution is centered below the kink point. I return to this finding in section 6 and discuss it in greater details in appendix E.

4.3 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 5 this is done for students who have cancelled exactly 1 month and thus taken up 11 months of benefits.

By cancelling 1 month of benefits these students increase 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 5
The earnings distribution for tertiary students who cancel 1 month of 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 $MTR = 1 - (1-t) \cdot (1-q)$, 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 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.⁸

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 (≈ 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

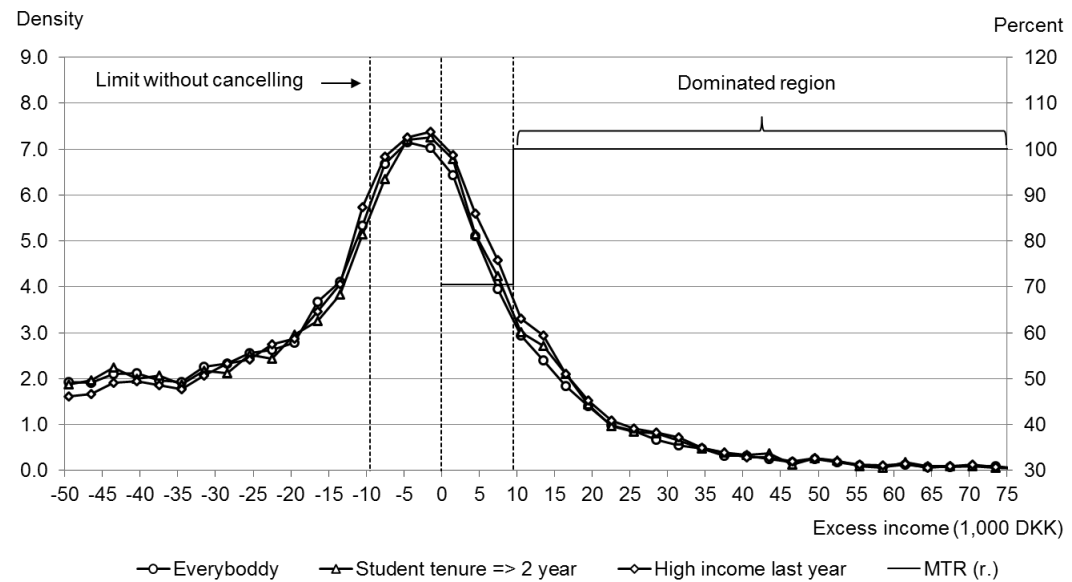
⁸ When interpreting the distribution in figure 5 as a result of optimization frictions it is important to eliminate measurement errors from the data, as these will otherwise result in an upward bias of the amount of frictions. An assessment of the amount of measurement errors and the results robustness to these are presented in appendix B and C.

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 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 6 the distribution from figure 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.

Figure 6
The earnings distribution for tertiary students who cancel 1 month of student benefits split on student history



Notes: See notes to figure 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 are fundamentally no differences between the distributions, and this 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 students 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 5), and the shape of the earnings distribution outside this range is therefore informative about the size of earnings errors that the individuals 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.⁹

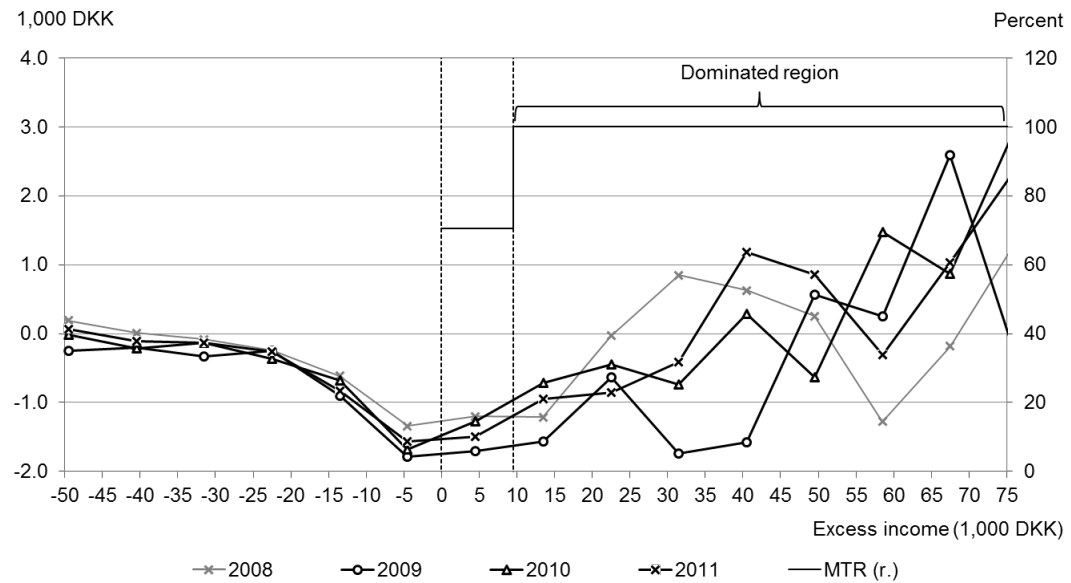
5 The nature of inattention

The graphical evidence in section 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 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 proxy 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 7.

⁹ For the exact calculation and description of the method see appendix D.

Figure 7
Average end of year income surprise over the income distribution, 2008-11



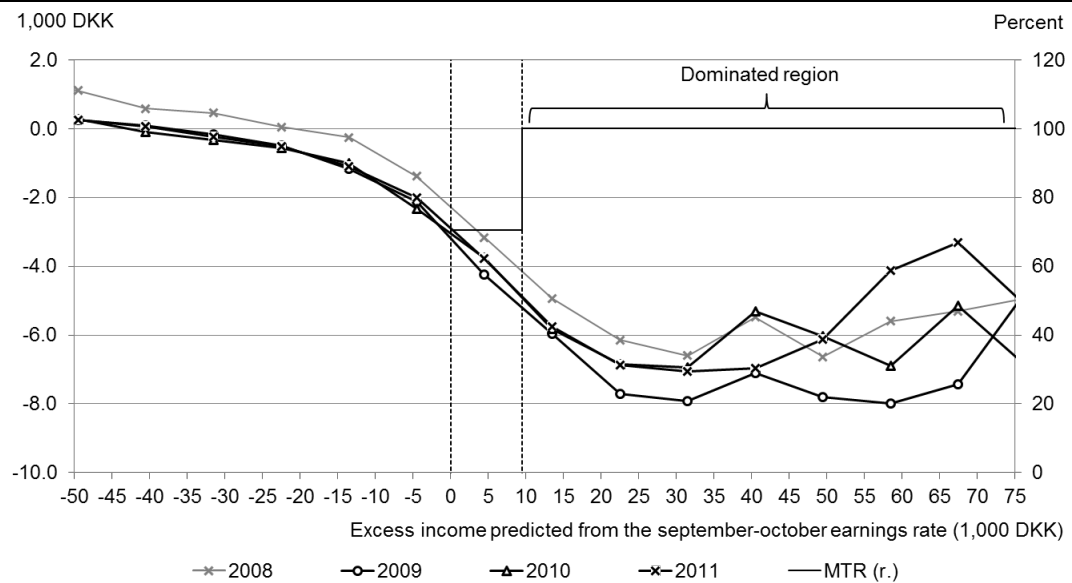
Notes: The figure only includes individuals who cancelled either 0 or 1 month if student benefits. The individual end of year income surprise is calculated as the difference between the sum of November and December pay and the sum of the September and October pay. Only labor income is included in this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

Sources: Own calculations based on DST

From this figure, we see a clear 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 8.

Figure 8
Average end of year income surprise over the predicted income distribution, 2008-11



Notes: See notes to figure 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 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.¹⁰

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 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 their earnings.

¹⁰ Indeed, most of the shift in the distribution after the 2009 reform observed in figure 4 can be attributed to the drop in the November-December earnings first occurring at higher earnings levels after the reform.

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 reduced earnings incentives created by the phasing out of student benefits and lowered their labor supply. However, perhaps more realistically the inattention that individuals exhibit in this labor market is not fully rational.

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 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 uncover 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.¹¹ 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 4 it is clear that the shift in the distribution happens over a much wider 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.

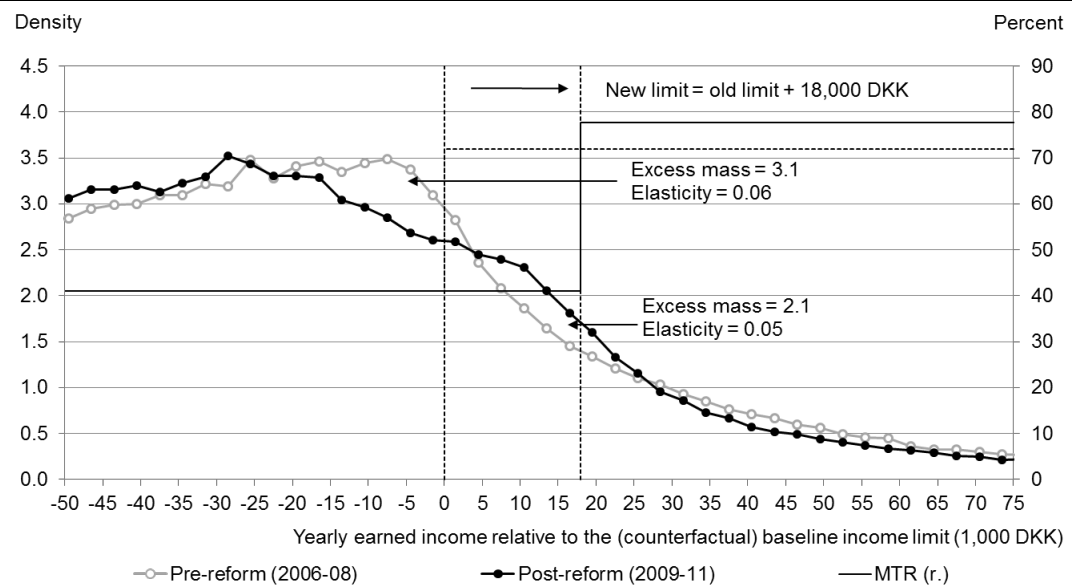
To uncover a labor supply elasticity I instead employ a method that resemble the method use by Chetty et al. (2013) and utilize the shift in the distribution created by the 2009 reform to uncover the (local) counterfactual distribution and hence the excess mass created by

¹¹ 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.

the pre- and post-reform kink.¹² Finally, I turn this excess mass into a labor supply elasticity using the Saez (2010) bunching formula.¹³

Figure 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 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.

Figure 9
Identifying excess mass using the 2009 reform



Notes: See notes to figure 4. For the calculations of the elasticities see table 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:

¹² 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.

¹³ 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 E, which yields almost the same elasticity estimates as here.

$$dz = \frac{B}{f(z)} \tag{1}$$

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:

$$\varepsilon = \frac{dz}{d(1-t)} \frac{1-t}{z} = \frac{B}{f(z)} \frac{1-t}{z} \tag{2}$$

yields an elasticity of 0.06 for the pre-reform kink point and 0.05 for the post-reform kink point, cf. table 3.

Table 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	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. 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 many students probably have a large degree of flexibility in increasing their earnings if desired.¹⁴ 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 distribution 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 optimization frictions this will not necessarily hold.

Examining figure 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.

¹⁴ Working in the other direction is the fact that students might use a student job to gain valuable job experience, 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.

This implies that the phase out rate used so far – and hence the size of the kink point – is an upper bound of the actual phase out rate further implying that the estimated elasticity is a lower bound. Assuming e.g. that 20 percent of the students in a given year is indifferent between receiving benefits within the year or saving them for later imply that the average kink size will be 20 percent lower than the one used above. Scaling down $d \log(1 - t)$ 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 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 capped. 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.

7 Conclusion

In this paper, 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 earnings process during the year as being 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 larger 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 frictions to mask the bunching expected at kinks points 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 elasticity 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 a more important role, as this would cause the labor supply responses to be more gradual following a reform – a gradual responses 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.
- le Maire, Daniel 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 Economics*, 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 Journal: 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 A: The Danish student benefit system

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 A1. Benefits for students aged 18-19 in lower educations (ISCED2011 level 3-4) furthermore depend on their parents' income.

Table 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 A2.

Table 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 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.

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.

A3 Income control

When students receive student benefits they are subject to an income test. The test is automatically done after the end of the income year by the student benefit administration, 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 student 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 paper 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 $12 \cdot 6,370 = 76,440$ 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 amount¹⁵) 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.

¹⁵ After the 2009 reform the low amount for the lower education is used for the tertiary educations.

Appendix 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 drawn 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.

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 where 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 themselves a child in a household (individuals not living with their parents).

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

$$\text{Income limit} = \text{Amount}_{\text{Low}} \cdot \text{No}_R + \text{Amount}_{\text{Medium}} \cdot \text{No}_E + \text{Amount}_{\text{Child}} \cdot \text{No}_C \quad (\text{B1})$$

Where No_R is the number of months, where the individual receives student benefits. No_E is the number of months where the individual is eligible for student benefits without receiving them (student benefits have been cancelled), and No_C is the number of children below 18 years. The amounts are the corresponding contributions to the income limit described in appendix A.

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 B1.

Table 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 funds¹⁶ 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 B1, but these variables do not precisely correspond to the 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 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.

¹⁶ But not direct dividend payments and capital gain taxes under the stock income scheme.

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 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 differences 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 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.

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.¹⁷

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 A, but this type of error of less importance for the analysis of labor supply responses in the paper. 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 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 B2

¹⁷ This is in contrast to "regular" regression analysis, where measurement error in the outcome variable only will lead to higher standard errors.

Assessing the accuracy of the own income definition

Hit rate (percent)	2007	2007	2008	2009	2010	2011
Aggregate hit 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 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).

Source: Own calculations based on DST.

Among the tertiary students the hit rate is naturally higher for the student below the income 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 C by replicating the key graphs in the paper 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.

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 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.

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 B3. The core sample consists of students, who are fully eligible for student benefits and employed. They numbers around 85,000 per year.

Table B3**The size of the Danish student population**

Labor Supply and Optimization Frictions

	2006	2007	2008	2009	2010	2011
1,000 persons						
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	

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

Sources: Own calculations based on DST

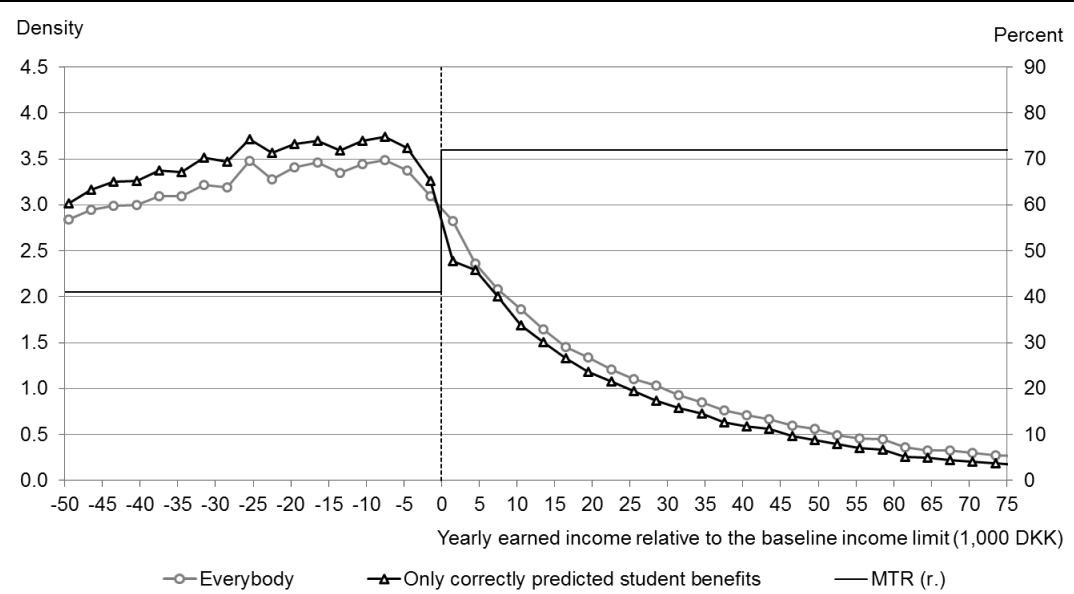
Appendix C: Robustness check wrt. measurement error

As shown in appendix 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 paper, 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 paper I therefore repeat the key figures in the paper (figure 3-5) using only the part of the sample, where I can actually predict their final student benefits.

Figure C1 corresponds to figure 3 in the paper 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 C1
The income distribution for tertiary students, 2006-2008



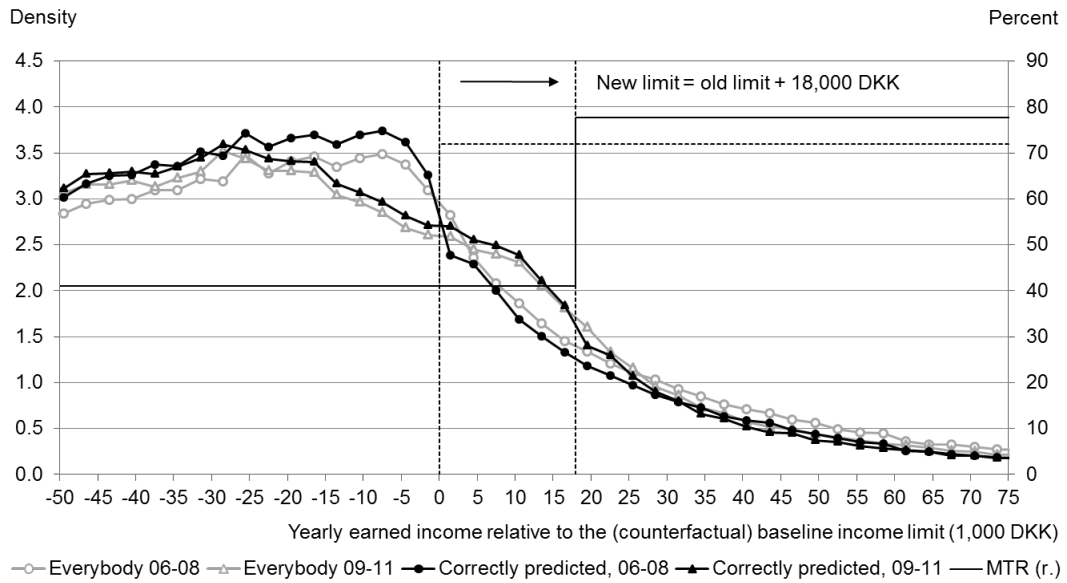
Notes: See notes to figure 3. The line for “Everybody” corresponds to average over the years in figure 3. “Only correctly predicted student benefits” only includes individuals with predicted student benefits with +/- 10 DKK of the actual student benefits received.

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 C2 and figure C3.

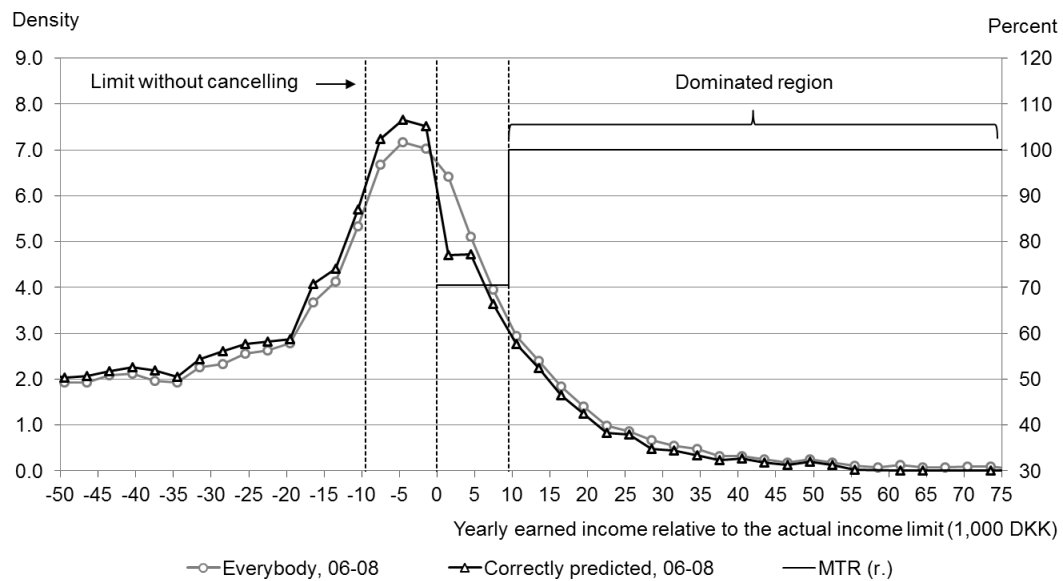
Figure C2

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



Notes: See notes to figure 4 and figure C1.
Sources: Own calculations based on DST

Figure C3
The income distribution for tertiary students who cancel 1 month of student benefits



Notes: See notes to figure 5 and figure C1. For the correctly predicted sample the mass in the dominated region is 12 percent.
Sources: Own calculations based on DST

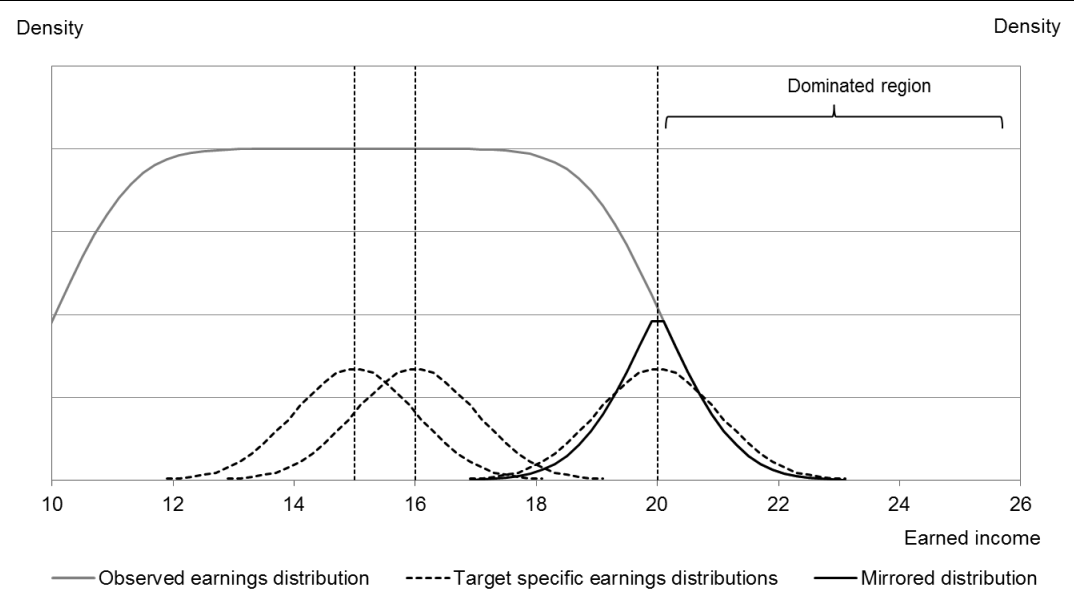
Appendix D: Deriving the costs of inattention

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 4.3 in the paper. 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 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 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 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.

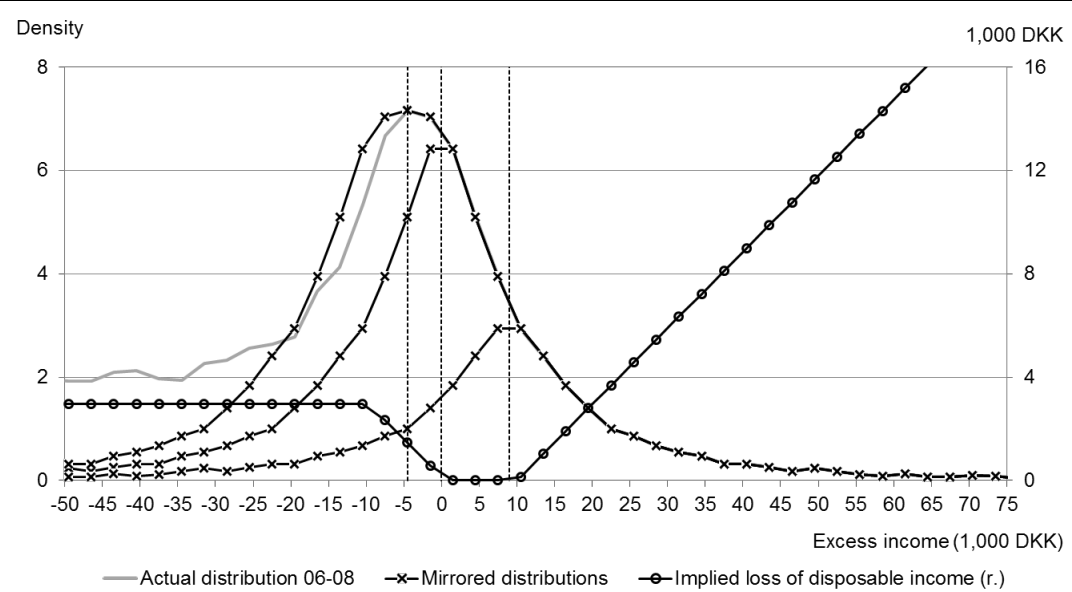
In contract the presence of dominated regions enables you to put bounds on the earning levels that individuals target. In figure D1 this is illustrated with a dominated region from earnings 20 and above, and as a consequence all observed earnings above 20 must be due to earnings errors among individuals 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 distribu-

tion in figure 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 5 in the paper) 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 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 5 in the paper. 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.

Source: Statistics Denmark and own calculations

Still, in the figure 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 D1.¹⁸

Table 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

¹⁸ 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.

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 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 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 E: GMM estimation of the labor supply of students

As a supplement to the non-parametric estimates of the labor supply elasticity in the paper 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.¹⁹

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 described in the paper 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 9. In this way the approach falls into the frame of GMM (Generalized method of moments) estimation.

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):

$$u_i = c_i - \frac{\mu n_i}{1 + \mu} \left(\frac{\hat{z}_i}{n_i} \right)^{\frac{1+\mu}{\mu}} \quad (\text{E1})$$

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:

$$z_i = \hat{z}_i + \varepsilon_i \quad (\text{E2})$$

where ε_i is an iid. error term.

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

$$c_i = (1 - t)[SB - q(T_i - L) \cdot 1(T_i > L) + z_i \cdot 1(z_i \leq T_i) + T_i \cdot 1(z_i > T_i)] \quad (\text{E3})$$

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 that ε_i is normal $N(0, \sigma)$ distributed. In this setting maximizing expected utility only depends on income through expected consumption, which given (E3) can be written as:

¹⁹ 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.

$$E(c_i) = (1 - t) \cdot \left[SB - q(T_i - L) \cdot 1(T_i > L) + \left(\hat{z}_i - \sigma \frac{f(\theta_i)}{F(\theta_i)} \right) \cdot F(\theta_i) + T_i \cdot (1 - F(\theta_i)) \right] \quad (\text{E4})$$

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:

$$\frac{\partial E(u_i)}{\partial T_i} = 0 \Leftrightarrow \frac{\partial E(c_i)}{\partial T_i} = 0 \Leftrightarrow 1 - F\left(\frac{T_i - \hat{z}_i}{\sigma}\right) = q, \text{ for } T_i > L \quad (\text{E5})$$

$$\frac{\partial E(u_i)}{\partial \hat{z}_i} = 0 \Leftrightarrow \frac{\partial E(c_i)}{\partial c_i} = \left(\frac{\hat{z}_i}{n_i}\right)^{\frac{1}{\mu}} \Leftrightarrow \hat{z}_i = \left((1 - t) F\left(\frac{T_i - \hat{z}_i}{\sigma}\right) \right)^{\mu} n_i \quad (\text{E6})$$

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.

E2 Simulation

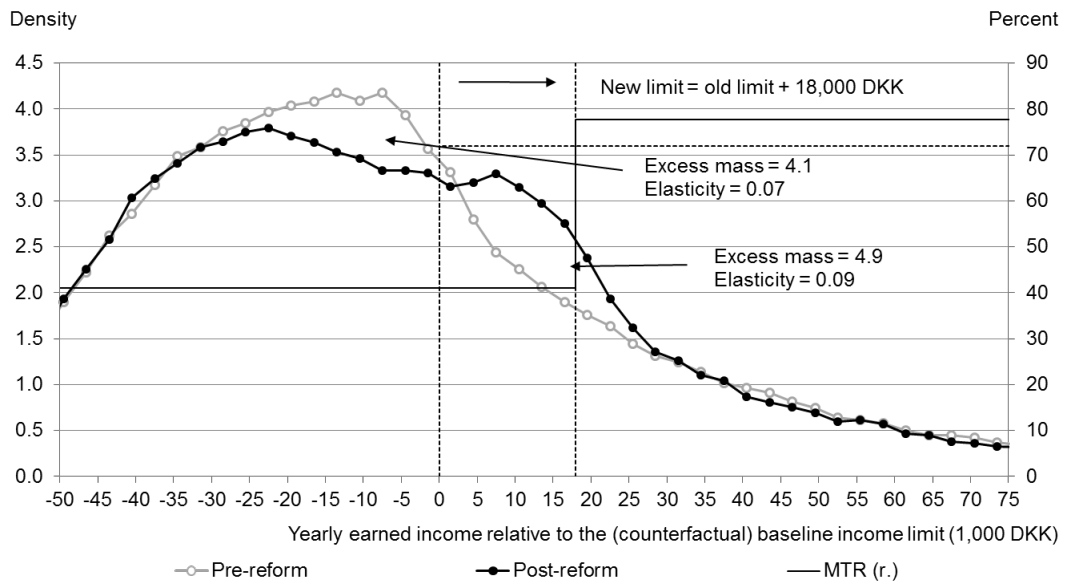
Before moving into the actual estimation, I present the performance of the model based on a simulation with fixed parameter values. The simulation is done by solving the model for a large number of individuals with different drawn of the distribution of potential earnings and with different realizations 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 E1, which shows the same shift in mass from below the pre-reform kink point to a range above as in figure 4 in paper. The figure also reports elasticity estimated using the same non-parametric method as in section 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 6.

Figure E1

Simulated earnings distribution before and after the 2009 reform.



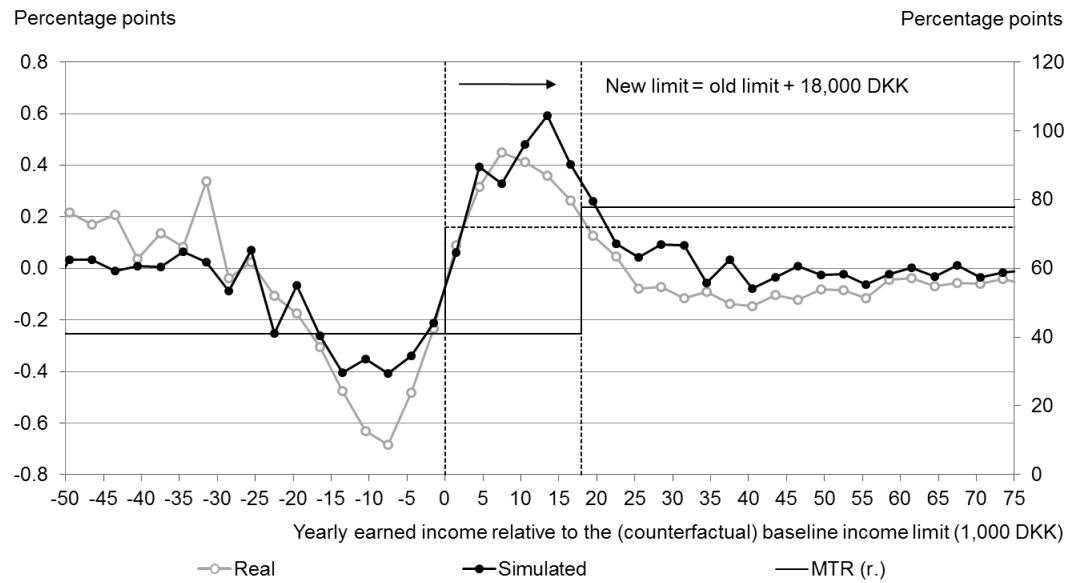
Notes: Simulated distribution of realized earnings based on a draw of 100,000 individuals with log earnings normally distributed with mean 4.3 and standard error 0.5. $\mu = 0.1$ and $\sigma = 7$.
Sources: Own calculations

E3 GMM estimation

From the simulation above I can calculate a change in the frequency in each bin and map this to the actual changes seen in figure 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.²⁰ The estimated labor supply elasticity is in other word more or less the same as the non-parametric estimate in the paper, 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 E2.

Figure E2
Simulated and actual change in the earnings distribution following the 2009 reform

²⁰ 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.



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

E4 The position of the excess mass

As mentioned in the paper 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 expect. However as already seen in figure E1 this is a consistent feature of the model, where individuals can cancel benefits in order to avoid the 100 percent marginal tax rate.

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.²¹

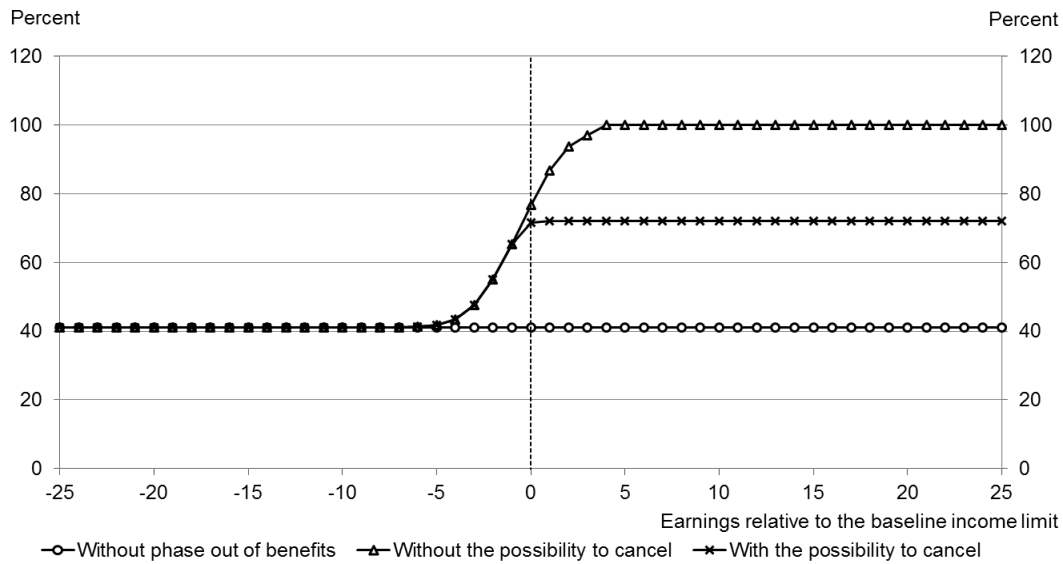
Without the possibility to cancel benefits the kink point faced by students is effectively a jump from 41 to 100 percent marginal tax rate, and so with earnings uncertainty the effective marginal tax rate increases smoothly between these 2 rates symmetrically around the kink point, cf. figure E3.

With the possibility to cancel benefits students can effectively move up the kink point by phasing out benefits, and from equation E5 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 effec-

²¹ 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.

tive 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 capped (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.

Figure E3
Effective marginal tax rates with and without the possibility to cancel benefits

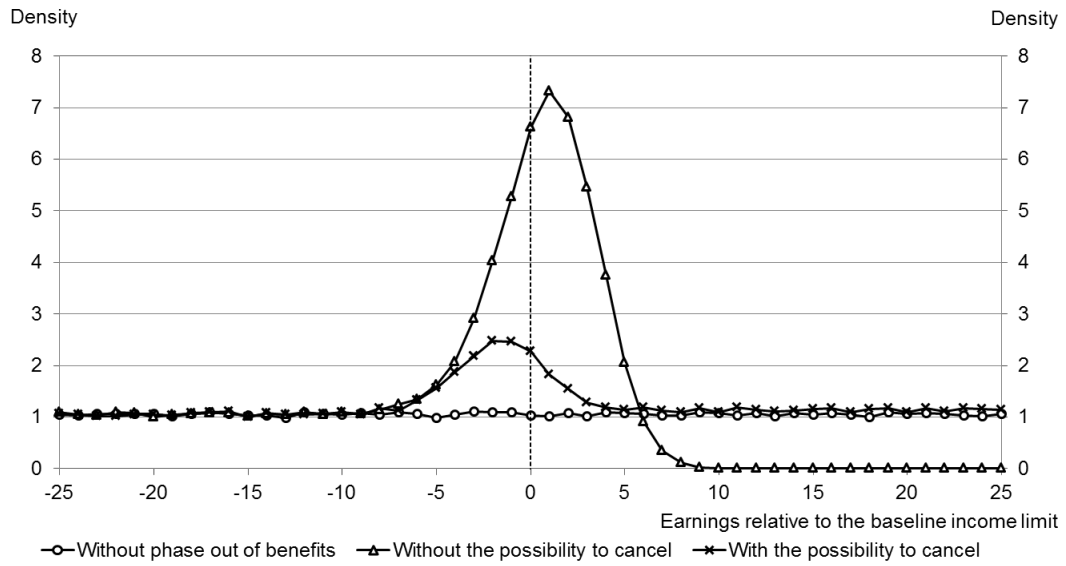


Notes: The effective marginal tax rate is calculated as $1 - (1 - t)F\left(\frac{T_i - \hat{z}_i}{\sigma}\right)$, where T_i is set so $1 - F\left(\frac{T_i - \hat{z}_i}{\sigma}\right) = q$, for $T_i > L$. (equations E5 and E6 above). In the case without phase out of benefits $q = 0 \Rightarrow T_i \approx \infty \Rightarrow F\left(\frac{T_i - \hat{z}_i}{\sigma}\right) = 1$. In the case without the possibility to cancel benefits $q = 1 \Rightarrow T_i = L \Rightarrow F\left(\frac{T_i - \hat{z}_i}{\sigma}\right) = F\left(\frac{L - \hat{z}_i}{\sigma}\right)$. 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 shown in figure 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 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 out 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.