# CEBI WORKING PAPER SERIES

## LABOR SUPPLY AND OPTIMIZATION FRICTIONS: EVIDENCE FROM THE DANISH STUDENT LABOR MARKET

Jakob Egholt Søgaard

# Labor Supply and Optimization Frictions:

## Evidence from the Danish Student Labor Market[*]

Jakob Egholt Søgaard[†]

University of Copenhagen and CEBI

January 2019

## Abstract

Using Danish administrative data, I investigate the magnitude and nature of optimization frictions in the labor market of Danish students. Danish students face a unique institutional setting that makes it possible to distinguish between different types of frictions and estimate their effect on individual utility. I find that frictions significantly affect observed labor market outcomes. In particular, the empirical evidence points to inattention as the dominant type of friction. In contrast, my findings appear inconsistent with real adjustment costs, price misperception and gradual learning. Overall, optimization frictions reduce the utility of individuals by approximately 2-3 percent of disposable income.

---

# 1   Introduction

One of the key insights that have emerged from the last two decades of research in public and labor economics is just how important optimization frictions are in shaping observed labor market outcomes. Optimization frictions are the likely cause for the lack of observed bunching at kink points (Saez, 2010 and Chetty et al., 2011) as well as the range of different labor supply elasticities estimated using tax reforms (Chetty, 2012).

Following these insights, one strain of literature has sought to estimate structural labor supply elasticities, explicitly taking into account the biases created by optimization frictions, but often remaining agnostic about the underlying types of frictions (Kleven and Waseem, 2013). However, even though the structural elasticities are still key in assessing revenue effects of tax reforms and designing optimal policy, an increasing number of scholars point to optimization frictions as not just "noise" preventing us from observing the structural economy, but as having first order effects on the design of optimal policy (Farhi and Gabaix, 2018, Goldin and Reck, 2018 and Reck, 2016). This nascent literature essentially makes the point that structural elasticities alone are not necessarily sufficient statistics when designing policy.

Yet, despite the growing recognition that both the magnitude and nature of optimization frictions are important for our understanding of the labor market, concrete evidence of optimization frictions is still limited. This reflects that identification of optimization frictions is demanding in terms of both high-quality data and special institutional settings. High-quality data are needed to separate optimization errors by individuals from measurement errors in the data and special institutional settings are needed to separate optimization errors from preference heterogeneity.

I contribute to this limited literature by presenting graphical evidence of both the magnitude and the nature of optimization frictions. I do so by studying the labor supply of Danish students, which represents a promising case study for several reasons. First, students face a sharp kink in their budget set created by the phase out of student benefits. The kink is large with a change in the net-of-tax rate of at least 75 log points. This is more than twice the size of the largest kinks in both the US and Danish tax systems. Second, a reform in 2009 significantly increased the earnings level at which students reach the kink point. Third, students need to adjust their benefits take-up before actually receiving the benefits to remain on a lower phase-out rate. If they fail to do so sufficiently, they risk hitting an ex post phase-out rate of 100 percent. This feature creates several dominated regions in their budget set. Finally, the labor market is covered by rich administrative data on earnings and benefits take-up at the monthly frequency.

By examining the individuals' earnings and take-up behavior across each of the three institutional features, I am able to distinguish between four of the main types of optimization frictions discussed in the literature:

- Real adjustment costs, such as effort associated with job search (Attanasio, 2000).
- Price misperception, such as mistaking the average tax rate for the relevant marginal tax rate (Liebman and Zeckhauser, 2003 and Ito, 2014).
- Gradual learning, e.g. related to learning about new or changed institutional features following a reform (Evans and Honkapohja, 2001).
- Inattention, e.g. related to changing circumstances that would otherwise warrant re-optimization of behavior (Sims, 2003 and Gabaix, 2017).

Overall, my main findings point to the presence of significant optimization frictions. First, following the 2009 reform I find an immediate and non-trivial shift in the earnings distribution compared to stable distributions both before and after the reform. The shift is consistent with a labor supply elasticity of 0.1. Second, despite this clear evidence of a positive labor supply elasticity, I find no sign of sharp bunching at the kink point, created by the phase-out of student benefits, which is predicted by standard theory (Saez, 2010). Third, I find that a significant share of individuals fails to take up benefits optimally given their final end-of-year earnings.

While pointing to the presence of significant optimization frictions, these findings are inconsistent with real adjustment costs or gradual learning as the main types of frictions, as these types of frictions would have created a more gradual transition to a new earnings distribution following the 2009 reform. The observed shift in the earnings distribution is also inconsistent with price misperception, as the responses to the 2009 reform appear relatively close to the kink point. This suggests that the individuals do in fact respond to changes in their effective marginal tax rate and not the average tax rate as often found with price misperception.

Instead, the findings are consistent with a model, where individuals (rationally) choose their desired labor supply and, thus, implicitly their desired yearly earnings, but where their actual earnings over the year may be affected by unexpected shocks to, for example, the wage rate, work commitments etc. As a consequence, final end-of-year earnings may deviate from the desired level, if individuals do not pay sufficient attention to their earnings process and re-optimize their labor supply and benefits take-up throughout the year. This prevents the formation of sharp bunching in the earnings distribution, and pushes individuals into dominated regions, even if individuals quickly change their desired labor supply in response to changes in the institutional setting.

Exploiting monthly data on earnings and benefits take-up, I further investigate when and how individuals re-optimize over the course of the year. I find that a significant share of

individuals passively maintains the default of full benefits take-up, despite that passivity can entail losses of more than 10 percent of disposable income. Looking over the entire year, 20 percent of individuals remain passive, and even towards the end of the year, where there is less uncertainty of final end-of-year earnings, 40 percent remain passive. Overall, individuals appear to accept average losses of approximately 2-3 percent of disposable income from optimization errors.

Combining the estimated labor supply elasticity and accepted utility losses provides a natural explanation for my finding of inattention as the dominate optimization friction. To see this, note that the cost of ignoring a tax change increases quadratically in the size of the tax change and linearly in the labor supply elasticity. This implies many tax reforms and kinks are too small to prompt behavioral responses if individuals are willing to accept even small utility losses from optimization errors. However, because the kink created by the phase-out of student benefits is very large, the utility loss of ignoring the kink or the 2009 reform is much more significant and above 3 percent for labor supply elasticities as low as 0.05-0.1.

The kink in my setting is, in other words, sufficiently large to overcome any real adjustment costs or cost of learning and prompts individuals to respond immediately to the higher kink point following the 2009 reform. However, this does not imply that individuals perfectly overcome all optimization frictions. Earning shocks over the year will still push them away from their desired earnings, but the expected utility losses from these deviations are not high enough for individuals to pay closer attention.

My paper contributes primarily to two literatures. First, and foremost, to the literature on optimization frictions in the labor market as well as in other settings. A key paper in this literature is Chetty (2012), who argues that the widely different estimates of the labor supply elasticity from a range of studies can be reconciled with a single structural elasticity if individuals are willing to accept utility losses from optimization frictions as low as 0.5-1.0 percent of disposable income. My paper provides the first non-parametric estimate of such accepted utility losses.[1]

My paper is not the first to document optimization frictions non-parametrically but differs in key ways from the previous literature. First, Kleven and Waseem (2013) quantify optimization frictions as the share of individuals in dominated regions. In their setting, this measure is essentially a sufficient statistic along with the observed bunching at notches to uncover a structural labor supply elasticity. However, the share in the dominated region is most likely a function of the notch size (a bigger notch increases the utility loss of being in the dominated

---

[1] The structural literature on labor supply often incorporates parameters or model specifications that can be interpreted as optimization frictions, see e.g. Keane, Todd, and Wolpin (2011).

region) and is less likely to be a structural parameter. Other papers often focus on one particular type of optimization friction such as adjustment costs (Gelber, Jones, and Sacks, 2013, 2017) or discrete earnings responses (Matikka and Kosonen, 2017) without explicitly considering other types of frictions. In this respect, my paper is more along the lines of Hoopes, Reck, and Slemrod (2015), who examine how tax payers collect information and contrast it with predictions from different behavioral models. They also reach the conclusion that inattention plays a key role in shaping taxpayer behavior. I complement their method by showing the same in a labor market setting using a natural experiment and the presence of dominated regions.

Second, I contribute to the literature on bunching estimation. The Saez (2010) bunching estimation has been applied in a number of settings, but most often the standard estimator only uncovers small behavioral responses (see e.g. Chetty et al., 2011 and Bastani and Selin, 2014). In the case, where the standard estimator does uncover large behavioral responses, they are often driven by reporting responses (e.g. tax avoidance) rather than real responses (see e.g. le Maire and Schjerning, 2013). My paper is the first to document clear but fuzzy excess mass around a kink point using reform variation. This result provides a firsthand example of the amount of frictions required to smooth out the excess mass at kinks predicted by standard labor supply models. In this way, the paper complements the work by Chetty, Friedman and Saez (2013), who use differences in knowledge about the Earned Income Tax Credit in the US to uncover (fuzzy) earning responses across the whole income distribution.

The rest of the paper is organized as follows. Section 2 formulates the expected labor market outcomes at different stylized institutional features and under different types of optimization frictions. Section 3 describes the institutional background surrounding Danish students and the data. Section 4 presents the key graphical evidence of optimization frictions and compares the observed outcomes to the predictions outlined in section 2. Section 5 examines when and how individuals re-optimize over the year and section 6 quantifies the costs of optimization frictions to individuals. Section 7 concludes.

## 2   Optimization frictions and labor market outcomes

Before moving into the empirical analysis, I start by outlining a number of predictions about how different types of optimization frictions affect observed labor market outcomes around the key institutional features surrounding Danish students. I will lay out the institutional setting in more details in section 3, but overall the institutional setting can be summarized by three key features:

1. Phase-out of benefits with earnings above a given threshold.
2. A reform that significantly increased the earnings threshold.

3.   Voluntary take-up of benefits, where more take-up risks reducing disposable income.

Of these, the two first are standard institutional features in the public finance literature. The first creates a kink point in the budget set due to a jump in the effective marginal tax rate, while the second is a reform that changes effective marginal tax rates in a part of the distribution. In contrast, the third feature is unique and specific to the setting facing Danish students. Essentially, the Danish student benefit system allows students to select their initial take-up of benefits. Taking up less benefits reduces the base amount received but increases the earnings threshold at which point benefits are phased out. As a consequence, the system creates an optimal level of benefits take-up that vary with earnings, where taking up either more or less benefits compared to this level reduces disposable income.

For each of the three institutional features, it is possible to draw a number of predictions about the labor market outcomes you would expect to find given the presence of different types of optimization frictions.

To see the possible outcomes, I start by considering the labor market outcomes in a benchmark situation without optimization frictions. In this case, individuals will bunch exactly at the kink point created by the jump in the marginal tax rate, and thereby create a sharp excess mass in the earnings distribution at the kink (Saez 2010). Following a tax reform that changes tax rates in part of the income distribution, we should observe an immediate change in earnings for the individuals, who are directly affected by the changed incentives. Finally, we should expect individuals to take up benefits only if it increases their disposable income.

Against the frictionless benchmark, I consider the effects of four types of frictions commonly discussed in the economics literature.

First, I consider the effect of real adjustment costs (see e.g. Attanasio, 2000). Real adjustment costs imply that it is costly for individuals to change their earnings, e.g. because it involves spending time and effort finding a new job. In this case, individuals are willing to accept jobs located in an interval around, rather than at, the optimal point, as the expected utility gain of finding a better job does not outweigh the search costs (Chetty et al., 2011). Consequently, only a fraction of the individuals, who in a frictionless world would bunch at the kink point, does so, causing the excess mass to distribute across an interval around the kink point (fuzzy bunching).

Following a tax reform, the presence of real adjustment costs implies that not all individuals will find it optimal to change their earnings immediately. Instead, they might choose to keep their current job if, for example, they expect to change job for other reasons in the near future. Consequently, we should expect to see only a gradual change in the earnings distribution.

Finally, real costs of adjusting labor supply should not necessarily lead individuals to take up benefits sub-optimally. As long as the administrative system is simple and the economic costs of taking up benefits are trivial, we should 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.

Second, I consider price misperception (see e.g. Liebman and Zeckhauser, 2003). There are in principle a number of ways that individuals can misperceive price or tax schedules, but in the context of non-linear price and tax schedules evidence points to ironing as the most common type of misperception (Ito, 2014 and Rees-Jones and Taubinsky, 2016). With ironing, individuals fail to take the complicity of the tax schedule into account, and instead use their average tax rate as the relevant "marginal" tax rate. As a result, the ironed-out tax schedule contains no kink and we should not expect any bunching around kink points in the actual tax schedule. On the other hand, we should expect both optimal take up as this reduces the average tax rate and immediate changes in the earnings distribution following a tax reform. However, ironing should imply a different set of responses compared to the frictionless benchmark, as the individuals with the largest changes in their average tax should respond more. In the frictionless model, the largest responses should be found among individuals with the largest change in their marginal tax rate.

Third, I consider gradual learning (see e.g. Evans and Honkapohja, 2001). Gradual learning implies that individuals do not have perfect information about the institutional setting. This includes knowledge of the precise position of kink points and the design of the benefit system. Consequently, 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 setting.[2] Likewise, gradual learning implies that knowledge of a reform expands gradually after its implementation, and we should expect to see only gradual changes in the earnings distribution.

Fourth, I consider the effect of inattention (see e.g. Sims, 2003). 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 re-optimization of individual behavior, might not be noted by individuals and leave them with ex post sub-optimal behavior. Formulated in this way there is a significant overlap between gradual learning and inattention, as inattention to changes in the institutional setting exactly corresponds to the gradual learning described above. Conse-

---

[2] I refer to sub-optimal behavior or optimization errors whenever behavior deviates from the optimal behavior in the benchmark case with no optimization friction. This does not necessarily imply that the individuals are irrational. The optimization errors due to real adjustment cost are, for example, perfectly rational.

quently, I will make the following distinction between gradual learning and inattention. Gradual learning refers to learning about the institutional setting, which we would normally think of 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 such as individual wages, working requirements etc. In a world, where these individual factors are partly random, an individual will never learn the true values of these by accumulated experience, but only 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 end-of-year earnings may deviate from this level due to changes in individual economic factors, which the individuals fail to realize and thus do not re-optimize. Consequently, we should expect only fuzzy bunching around a kink point in the budget set and some individuals to take up benefits even though it turns out to be sub-optimal ex post. However, despite the inattention, we should expect to see an immediate change in the earnings distribution following a tax reform, as individuals adjust their desired earnings in response to the new incentives.

Finally, it should be noted that, while Sims (2003) discusses the implications of rational inattention, inattention might as well be irrational, just as inattention may be related to the effects of the individual's own actions.[3] For example, as most countries base taxation on the cumulative earnings over the year, knowing the effect of extra earnings in one month requires individuals to keep track of (and predict) earnings in all months. These types of inattention also lead to the predictions mentioned above.

All of the above predictions are summarized in Table 1. From the table, it is clear that each type of friction leads to a unique set of predictions across the different institutional features. Hence, combining the observed outcomes across these features will, in principle, make it possible to distinguish between the different types of frictions.

---

[3] The notion of rational inattention relies on the assumption that the costs of paying closer attention to changes in the economic circumstances outweigh the expected benefits of smaller optimization errors.

**Table 1**
**Predictions: What to expect under different types of optimization frictions?**

|  | Bunching at the kink point | Effect of a tax reform | Take up of benefits |
|---|---|---|---|
| *Benchmark:* | | | |
| No frictions | Sharp bunching | Immediate change | Optimal take up |
| *Optimization frictions:* | | | |
| Real adjustment costs | Fuzzy bunching | Gradual change | Optimal take up |
| Price misperception | No bunching | Immediate change[a] | Optimal take up |
| Gradual learning | Fuzzy bunching[b] | Gradual change | Sub-optimal take up[b] |
| Inattention | Fuzzy bunching | Immediate change | Sub-optimal take up |

a) The distribution of the responses will differ compared to the frictionless model, as individuals would respond to the change in their average tax rate instead of the marginal.
b) These effects will be more pronounced among individuals, who are relatively new to the system.

## 3 Institutional background and data

### 3.1 The Danish student benefit system

In this section, I present the key features of the Danish student benefit system and relate it to the stylized institutional features discussed in section 2.

On top of tuition free access to public education, Danish students enrolled in education above primary school (ISCED2011 level 3 and above) are eligible to state financed student benefits from age 18. Benefit rates vary depending on the type of education and civil status, but the basic rate for the typical students enrolled in tertiary educations (ISCED2011 level 5 and above) is DKK 6,100 per month corresponding to approximately USD 950 (2018-levels). This rate applies to vast majority of the population used in this analysis.

In addition to receiving benefits, students (prior to 2009) could earn income of up to DKK 7,700 per month without having their benefits reduced.[4] If they earn more than this baseline income limit (on a yearly basis), their benefits are phased out. For the first DKK 11,600 above the limit the phase-out rate is 50 percent, while further earnings entail a phase-out rate of 100 percent.[5] If students want to earn more, they can increase their individual limit by reducing the amount of benefits taken up by one or more months. For each month not taken

---

[4] Income counted against the income limit includes labor income (net of labor market contribution), transfers other than student benefits and capital income with the exceptions of certain types of stock income. All monetary amounts in this paper have been scaled to 2018-levels using the official (wage) indexation of the student benefits. DKK 6.4 ≈ USD 1 given 2018 exchange rates.

[5] Finally, if the amount of student benefits that a student has to pay back exceeds DKK 9,200, the entire payback is increased by 7 percent. This creates a notch in the budget set, which is ignored in Figure 1 and Figure 2.

up the student increases the income limit by DKK 11,600, which translates into a phase out rate of 6,100/11,600 = 52 percent.

This system creates a set of budget sets that students can choose from as illustrated in Figure 1. Taking up more benefits increases disposable income for students with low earnings, but holding constant the amount of benefits taken up, students will eventually hit the 50 and 100 percent phase-out rate if their earnings are too high. As a result, students maximize their disposable income by taking up 12 months of benefits for earnings below DKK 104,600, by taking up 11 months for earnings between DKK 104,600 and DKK 116,200 etc. The choice of budget set corresponds to the problem of optimal take-up of benefits described in section 2.

**Figure 1**
**Effective budget sets depending on benefits take up before the 2009 reform**
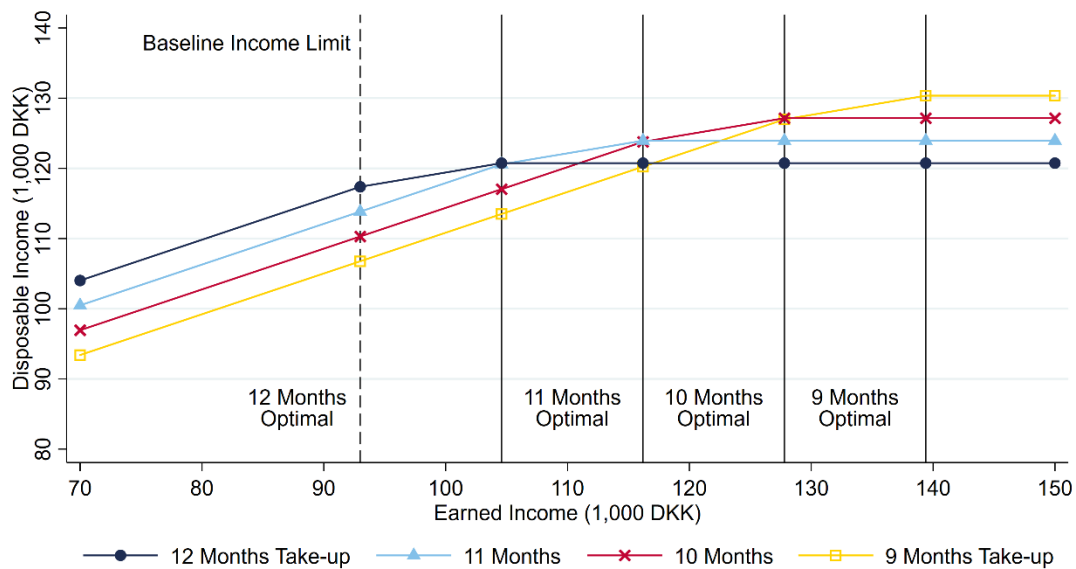


Notes:  The figure shows the disposable income for students depending on their earned income and the amount of benefits they take up. The baseline income limit is calculated as 12 x the monthly basic amount of DKK 7,700. Yearly disposable income is calculated in three steps: First, gross income is calculated as DKK 6,100 x the number of months of benefits taken up plus earned income up to the income limit, which increases by DKK 11,600 for each month not taken up. For the first DKK 11,600 in earned income above this income limit, 50 percent is deducted from the student benefits, while further excess income is deducted at a rate of 100 percent. Second, the tax liability is calculated as 42 percent of gross income above a personal allowance of DKK 49,900. Finally, disposable income is calculated as the difference between gross income and the tax liability. DKK 6.4 ≈ USD 1.

By default, eligible students are assigned to full benefits at the beginning of each year, but students can easily adjust their take-up through a simple webpage, where individual months of benefits can be cancelled/taken up with a few clicks (see Figure A.I). The real economic costs of adjusting benefits are therefore trivial. However, students face the constraint that

benefits for a given month cannot be changed after the 15th of the previous month. Compared to this, students typically receive their pay check at the end of the month or with an additional month's lag. Consequently, the system creates a difference between the point in time, where students have to decide on their benefit take-up, and the point in time, where students have precise information about their realized earnings. For example, students have to decide on their benefits paid out in December prior to November 15, and at this point in time students have in general only seen their pay checks up to October or September.

This time gap implies that students have to pay close attention to their earnings during the year and to some extent predict what they will earn a couple of months into the future to take up benefits optimally.

Together with the normal income tax system, which – for incomes in the range considered here – imposes a marginal tax rate of 42 percent, the phase-out of benefits causes the effective marginal tax rate to jump from 42 percent to $1–(1–0.52)(1–0.42) = 72$ percent when earnings exceed DKK 93,000 annually.[6,7] However, the effective marginal tax rate jumps even more, if students fail to take up the right amount of benefits. In the case of full take-up of benefits (12 months), students would (prior to 2009) hit the 100 percent phase-out rate at earnings exceeding DKK 104,600 as illustrated in Figure 2. In this case, the student could have increased his disposable income by taking up fewer months of benefits.

---

[6] Excluding VAT of 25 percent.

[7] There is a caveat to the calculation of the effective marginal tax rate: 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 taking up fewer months of benefits, the student can save them for later use. Consequently, some students might not see less than full take-up as the full loss assumed here. This probability does not significantly affect the conclusions drawn in this paper, but simply implies a smaller effective kink. Benefits that are phased-out, when earnings exceed the income limit, are not available for later use.

**Figure 2**
**Effective marginal tax rates before and after the 2009 reform**
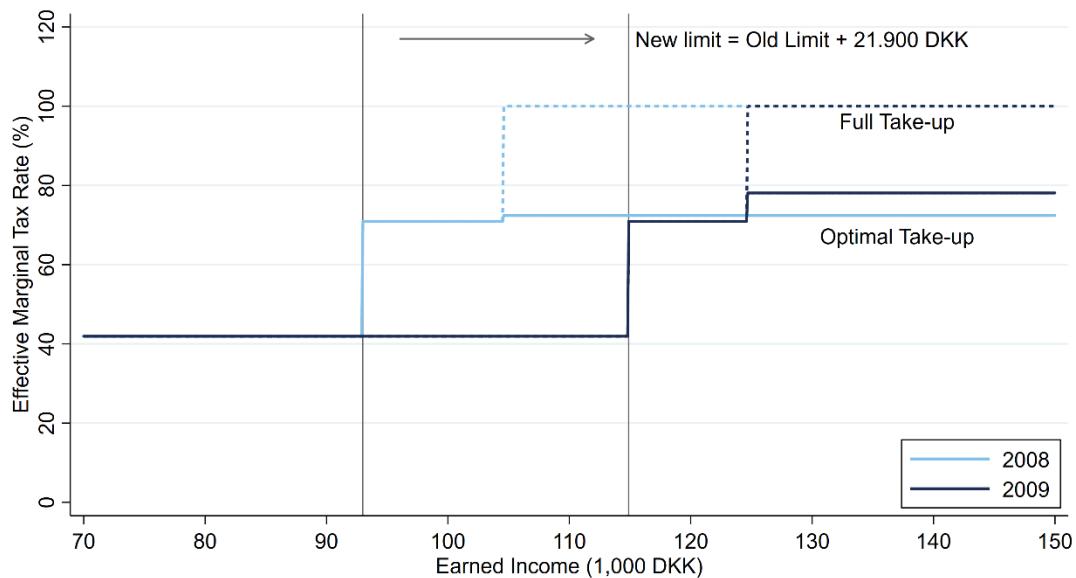


Notes: The figure shows effective marginal tax rates implied by the budget sets illustrated in Figure 1. The solid lines (optimal take-up) illustrate effective marginal tax rates, when students take up the amount of benefits that maximize their disposable income. This corresponds to the upper envelope of the budget sets in Figure 1. The dashed lines (full take-up) illustrate effective marginal tax rates, when students take up 12 months of benefits regardless of their earned income. DKK 6.4 ≈ USD 1.

The student benefit system has remained largely unchanged throughout the period 2004-2011, which I considered in this analysis, with the exception of a reform in 2009. The reform increased the baseline income limit by 24 percent (on top of the regular wage indexation) corresponding to DKK 21.900 as illustrated in Figure 2. At the same time, the phase-out rate increased from 52 to 62 percent and thus causing an increase in the effective marginal tax rate from 72 to 78 percent.

The change in the effective marginal tax rate created by the phase-out of student benefits constitutes a large kink in the incentives faced by students. The jump from 42 to 72 percent corresponds to a change in the net-of-tax rate of 75 log points. This is more than twice as large as the kink created by the current Danish top tax, which increases the marginal tax rate from 42 to 56 percent, and larger than the current kinks in the US tax schedule including the kink created by the phase out of the EITC. The 2009 reform also created changes in the log net-of-tax rate that are twice the size of the changes created by US Tax Reform Act of 1986.[8]

---

[8] The US Tax Reform Act of 1986 lowered the top marginal tax rate from 50 to 28 percent (Feldstein, 1995) corresponding to an increase in the net-of-tax rate of log(1–0.28)–log(1–0.50) = 36 log points.

## 3.2 Data

I use administrative data for the full population in Denmark. The Danish data combine several administrative registers (linked at the individual level via personal identification numbers) and contain detailed information on the entire education history of each individual, earnings, benefits, as well as demographic information. The data allow me to link individuals over time, family members, students to schools and workers to firms.

The frequency of the data varies. Demographic, education and income information from the tax returns are on a yearly basis, while information on benefits take-up is on a monthly basis. From 2008, I also have access to monthly records of earnings and hours worked for all wage earners in Denmark. Almost all income data are third party reported (Kleven et al., 2011).

I focus on individuals between 18-35, enrolled in tertiary education, fully eligible for student benefits and active in the labor market. These restrictions leave me with a core sample of around 75,000 observations per year as shown in Table 2. The core sample only represents a small proportion of the initial total population. This reflects that I consider an 18-year age span, while individuals are typically only enrolled in tertiary education for 3-5 years. At age 35 around half the population have completed a tertiary education.

**Table 2**
**Data restrictions and average sample size per year, 2006-2011**

|  | No. of individuals | Percent |
|---|---|---|
| Total population, age 18-35 | 1,195,500 | 100.0 |
| + Enrolled in tertiary education | 171,400 | 14.3 |
| + Fully eligible for student benefits | 94,700 | 7.9 |
| + Positive labor earnings = Core sample | 75,500 | 6.3 |

Notes: Positive labor earnings is defined as having earned income above DKK 8,000 per year.

Table 3 presents a list of descriptive statistics for the core sample focusing on 2008, where the monthly income data are first available. The average student is just below 25 years old and half the sample has an age between 23 and 26 years old. 70 percent of the sample are university student, while the remaining 30 percent are enrolled in lower levels of tertiary educations such as schoolteacher and nurses' training. Average earnings (excl. student benefits) are around DKK 64,000 per year corresponding to approx. USD 10,000. The majority of these are labor earnings while the rest comes from capital income and transfers excluding student benefits. The students' own earnings together with the student benefits paid by the state constitute the main sources of income for most students. Parental transfers play only a minor role for most students.

Around half the sample work in non-study related jobs such as supermarkets, restaurants and childcare.[9] A final feature of the student labor market is the relative flexibility of the majority of the jobs that students occupy. One way to see this is to compare the within-individual standard deviation of monthly earnings (normalized to average monthly earnings) to that of the rest of the population. For students the average normalized standard deviation is 0.7 compared to 0.3 for the general population.[10]

**Table 3**
**Descriptive statistics for the core sample, 2008**

|  | P25 | Average | P75 |
|---|---|---|---|
| Age | 23.0 | 24.6 | 26.0 |
| Enrolled in university education (%) | 0.0 | 68.9 | 100.0 |
| No. of past years in tertiary education[a] | 0.0 | 1.7 | 3.0 |
| Benefits take-up (months) | 12.0 | 11.0 | 12.0 |
| Total earnings excluding student benefits | 34,200 | 64,600 | 77,600 |
| - Labor earnings | 31,500 | 61,200 | 75,400 |
| - Non-labor earnings[b] | 0 | 3,400 | 1,600 |
| Avg. hours worked per week | 4.5 | 8.8 | 11.2 |
| Std. of monthly labor earnings[c] | 2,100 | 4,100 | 5,100 |
| Normalized std. of monthly labor earnings[d] | 0.5 | 0.8 | 1.0 |
| Employed in a study related job (%) [e] | 0.0 | 68.6 | 100.0 |
| Living with parents (%) | 0.0 | 6.2 | 0.0 |

Notes:  a) The number of years in tertiary education over the last four years. b) Capital income and transfers other than student benefits. c) Within individual standard deviation of monthly earnings. d) Within individual standard deviation of monthly earnings divided by the individual's average monthly earnings. e) Study related jobs are defined as all jobs not in wholesale (NACE 46), retail (NACE 47), postal service (NACE 53), hotels and restaurants (NACE 55, 56), cleaning (NACE 81), elderly, social and childcare (NACE 87, 88) and sports and amusement facilities (NACE 93). DKK 6.4 ≈ USD 1.

In Figure A.II in the appendix, I present additional descriptive statistics focusing on the relative high earning individuals, who are affected by the (risk of) phase-out of benefits. These individuals have a relative constant average hourly wage at around DKK 175 across the income distribution. Instead, the differences in total earnings are mainly driven by differences in the number of hours worked. Individuals with yearly earnings below the baseline income limit work on average 10 hours per week, while individuals with earnings above work 15-20 hours per week.

---

[9] Defining study related jobs are difficult. A definition based on where final graduates primarily find employment would e.g. not capture research and teaching assistants as only few graduates continue in academia. The definition employed here is based on manually picking the mayor occupations that typically are considered non-study related.

[10] Non-students aged 30-35 with labor earnings above DKK 8,000 similar to the definition used for students.

Individual marginal tax and phase-out rates are not directly observed in data.[11] I therefore develop a model that simulate individual phase-out rates, distance to the income limits and payback of student benefits for each individual based on the available administrative data (a model similar to the NBER TAXSIM model for the United States). The model accurately predicts individual student benefits within DKK +/- 2 for 98 percent of the core sample as described in appendix C.

# 4  Graphical evidence on labor supply responses and optimization frictions

In this section, I examine the labor market outcomes around each of the institutional features described in section 3 and compare it to the predictions in section 2.

## 4.1  Evidence from bunching at the kink point

Figure 3 shows the earnings distributions for students enrolled in tertiary education before and after the 2009 reform centered on the baseline income limit. Only students fully eligible for student benefits the entire year are included in the figure. However, I do not condition on the actual take-up of benefits (i.e. students are allowed to take up less than 12 months of benefits). Under the assumption that students always chose the optimal take-up, their effective marginal tax rate jumps from 42 to 72 percent at the baseline income limit as described in section 3.

In the 3 years prior to as well as the 3 years after the 2009 reform the earnings distributions were stable with no sign of sharp 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 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 sharp excess mass. Naturally, we cannot distinguish between different types of optimization frictions either.

---

[11] Individual phase-out rates are based on calculations of individual income limits and total earned income counted toward the income limit. These are performed by the Student Benefits Administration, but not transmitted to Statistics Denmark. The marginal tax rates are straight forwards as all individuals in the earnings interval relevant for this analysis face a marginal tax rate of 42 percent above a personal allowance of DKK 49,900.

**Figure 3**
**The earnings distribution before and after the 2009 reform**
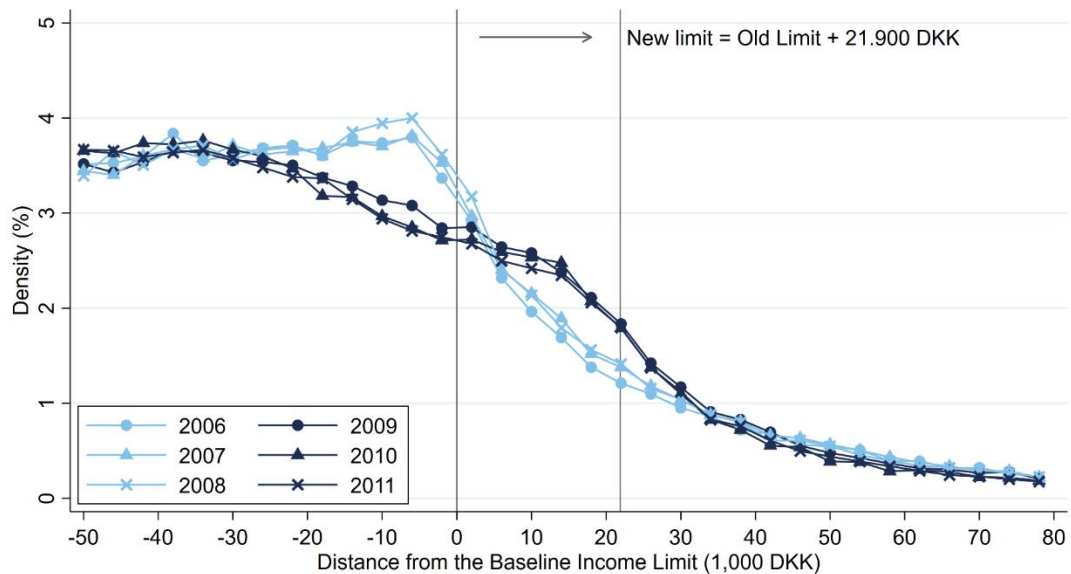


Notes:  The figure shows the distributions of earned income for students enrolled in tertiary education, fully eligible for student benefits and with yearly earnings above DKK 8,000. The distribution is not conditional on the actual take-up of benefits. Under the assumption that students always choose the optimal take-up, their effective marginal tax rate jumps from 42 to 72 percent at the initial baseline income limit and from 42 to 72/78 percent at the new income limit as described in section 2. The shift in the distribution is highly significant as illustrated in appendix Figure B.I. DKK 6.4 ≈ USD 1.

## 4.2  Evidence from the 2009 reform

Comparing the earnings distributions before and after the 2009 reform in Figure 3 reveals a clear shift in the distribution with mass moving from below the initial kink point to a range above.[12] Given the stability of the distribution in the years prior to the reform, this shift constitutes compelling graphical evidence for a positive labor supply elasticity and suggests that the lack of bunching at the kink points is due to optimization frictions.[13] Furthermore, the fact that the shift in the distribution appears to happen instantaneously from 2008 to

---

[12] When comparing the distributions before and after the 2009 reform, it is important that they are conditional on the actual take-up of benefits, as the 2009 reform changes the incentives to take up benefits for any given earnings level. Conditioning on the actual take-up would therefore "mechanically" create a shift in the earnings distribution. To see this point, consider individuals who take up full benefits (12 months). As illustrated in Figure 1, full take-up is optimal for earnings up to DKK 104,600 before the 2009 reform, but after the 2009 reform the range increases by DKK 21,900. After the 2009 reform, we should expect to see full take-up for a new group of individuals with higher earnings, and conditional on full take-up (or any other amount) we would see a shift in the distribution even in a situation with perfectly inelastic labor supply.

[13] 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 distributions for individuals in the same age range, but who were unaffected by the 2009 reform, did not change in a similar way (see Figure A.III in the appendix).

2009 speaks against both real adjustment costs and gradual learning as the dominant frictions as both would predict a more gradual response to the 2009 reform.

In appendix B, I use the 2009 reform to estimate an underlying labor supply elasticity. In a first approach, I apply the Saez (2010) bunching estimator. The key element in this approach is the estimation of the excess mass at the kink points, measured as the difference between the observed distribution and the counterfactual distribution absent at the kinks. In a standard cross-sectional setting, the counterfactual distribution would be estimated by fitting a high order polynomial to the observed distribution leaving out a range with observed bunching (see e.g. Kleven, 2016). This approach is obviously not feasible in the current setting, where there is no sharp bunching in each of the cross sections. Instead, I use the shift in the distributions to estimate the counterfactual distribution and arrive at an elasticity around 0.1.[14]

A labor supply elasticity of 0.1 is well within the range estimated by Kleven and Schultz (2014) for the full Danish population. However, there are reasons to expect a lower elasticity for students than ordinary workers as students may work to gain experience that will translate into higher earnings in the future. Such intertemporal consideration reduces individuals' responsiveness to contemporary incentives (Best and Kleven, 2013).

The Saez (2010) bunching estimator relies theoretically on a set of individuals responding to the kink in the tax schedule by reducing their earnings to exactly match the kink point. However, such precise earnings responses are unrealistic once we move to a world with optimization frictions.[15] Therefore, I supplement the bunching estimates using a structural model described in appendix B. This approach delivers a labor supply elasticity similar to the bunching estimator.

The structural model also explains why the excess masses revealed by the 2009 reform are centered below the kink points. In a standard setting, optimization frictions would smooth out kinks in the budget set symmetrically around the actual kink point, which would translate into a (fuzzy) excess mass centered on the actual kink point. The same is true in the current setting when holding benefit take-up fixed. For given benefits take-up students face a jump in the effective marginal tax rate from 42 to 100 percent as described in section 3 and optimization frictions would smooth symmetrically around the kink point. However, as students have the possibility to take-up less benefits ex ante at a phase-out rate of 52 percent, they

---

[14] This resembles the estimation strategy in Chetty, Friedman and Saez (2013) except that the source of the variation in this study comes from the time series variation created by a reform, while Chetty, Friedman and Saez (2013) use cross-sectional variation in the knowledge about the tax schedule.

[15] 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.

will optimally choose to do so once the expected effective marginal tax rate equals the ex ante rate of 72 percent. This caps the smoothed effective marginal tax at 72 percent and shifts the excess mass below the actual kink point as illustrated in Figure B.V in the appendix.

## 4.3 Evidence from the take-up of benefits

Turning to the final institutional feature, I show in Figure 4 the distribution of earnings relative to the individual income limits. The difference gives individual excess income, which (for positive amounts) translates into payback of benefits. I show the distribution for individuals, who have taken up less than full benefits. These individuals have actively changed their take-up of benefits from the default of full benefits take-up. The advantage of looking at this group is that their choice of take-up implies a relatively narrow region of optimal earnings as illustrated in Figure 1 and observing earnings falling outside this region is a strong indicator of optimization frictions.[16]

As illustrated in Figure 4, only 16 percent of individuals have earnings in the optimal region (excess income between DKK 0 and 11,600) before the 2009 reform. This is the less than the share of individuals, who end up with earnings above the optimal region. Having excess income above DKK 11,600 places the individual in a dominated region with an effective marginal tax rate of 100 percent. These individuals could with relatively little effort have taken up less benefits, thereby reducing the effective marginal tax rate to 72 percent. Therefore, ending up with excess income above DKK 11,600 is sub-optimal, even for individuals, who supply labor perfectly inelastically, for example because they work to gain experience and thereby higher future income.[17]

---

[16] Looking at individuals who take up full benefits does not change the conclusions drawn in this section, although the share of individuals in the dominated region is smaller due to the fact that a significant share of individuals has earnings well below the income limit. For these individuals it is not possible to separate optimization friction from heterogeneity in preferences/productivity.

[17] Using excess income to identify optimization frictions relies on the assumption that the calculated earnings and income limits are accurate. In presence of significant measurement errors, the amount of optimization frictions risk being biased upwards. To investigate this, I replicate, in Figure A.V, the distributions in Figure 4 using imputed excess income based on the observed payback of student benefits. The distributions using this measure are close to identical to the distributions in Figure 4.

**Figure 4**
**The distribution of excess income conditional on less than full take-up**
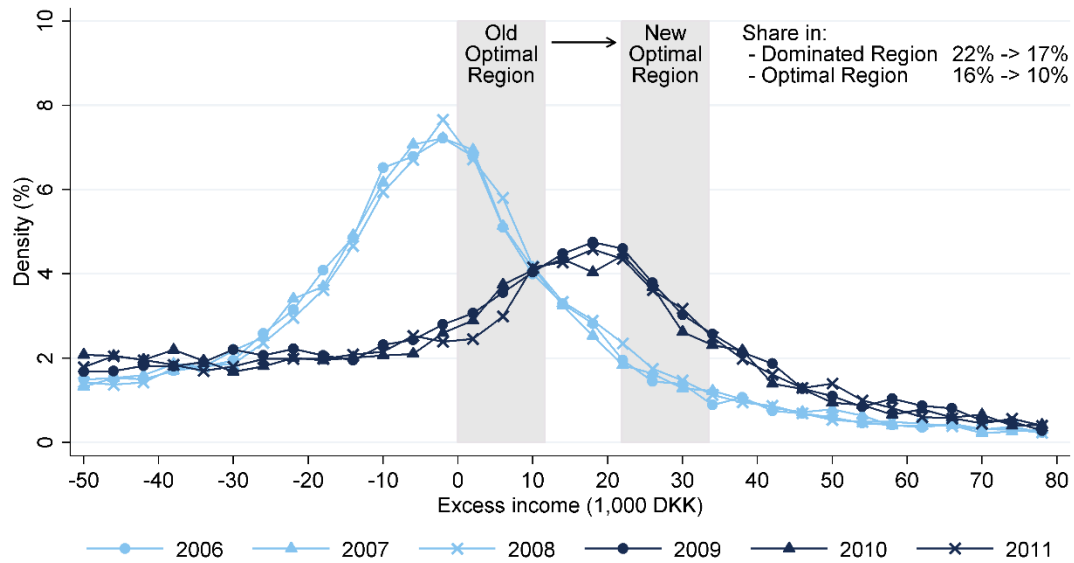


Notes:   The figure shows the distributions of earned income relative to the individual income limit for stu-
dents enrolled in tertiary education, fully eligible for student benefits and with yearly earnings
above DKK 8,000. I only include individuals who take up less than full benefits as these individu-
als have a well-defined region of optimal earnings (excess income between DKK 0 and 11,600).
Earnings above this region face a 100 percent effective marginal tax rate, which could have been
avoided by taking up less benefits ex ante. The reported share in the dominated region only counts
individuals above the optimal region.
DKK 6.4 ≈ USD 1.

Considering the lower part of the distribution, we also see a significant share (60 percent) of
individuals, who earned less than the individual income limit. These individuals could have
taken up more benefits without facing a higher effective marginal tax rate and thereby in-
crease their disposable income. However, there might be intertemporal considerations that
partly rationalize this behavior. Student benefits are typically limited to 6 years (compared to
a nominal study time of 5 years) and individuals might choose to save benefits for later use.
In a lifetime perspective, such behavior only represents a loss of disposable income if the
individual fails to use the benefits later.[18] In contrast to the upper part of the distribution, it
is not unambiguously a dominated region.

Figure 4 includes all individuals who take up between 1 and 11 months of benefits. Consid-
ering each level of take-up separately, appendix Figure A.IV reveals that the fewer benefits
individuals take up, the less able they are to hit the optimal region. This is likely due to the

---

[18] Depending on the interest rate and the rate of wage indexation of benefits, there might also be a (small) gain
or loss in the present value of the benefits.

fact that individuals with higher earnings have to adjust their benefits earlier to take up the optimal amount, which is more demanding in terms of paying attention throughout the year.

The observed sub-optimal take-up of benefits is consistent with both gradual learning and inattention, however with the distinct difference that under gradual learning, sub-optimal take-up should be more pronounced among individuals who have less experience with the system. I investigate this in two way. First, in appendix Figure A.VI, I show the distributions of excess income for individuals depending on their experience with the benefits system. The distributions are close to identical for all levels of experience and, in particular, more experience with the system does not reduce the share of individuals, who end up with earnings in the dominated region.

**Table 4**
**Estimated effects of knowledge and experience on sub-optimal benefit take-up, 2008**

| Dependent variable: | Dummy for earnings in the dominated region | | Amount of earnings in the dominated region | |
|---|---|---|---|---|
| *Years in tertiary education* | | | | |
| - 1 year | -0.048*** | (0.010) | -1,416*** | (327) |
| - 2 years | -0.055*** | (0.010) | -1,720*** | (324) |
| - 3 years | -0.095*** | (0.010) | -3,329*** | (359) |
| - 4 years | -0.126*** | (0.010) | -4,068*** | (393) |
| | | | | |
| *Old siblings in tertiary education* | | | | |
| - Currently | -0.002 | (0.010) | -33 | (315) |
| - Previously (completed) | -0.019** | (0.008) | -286 | (266) |
| | | | | |
| *Previously sub-optimal take-up* | | | | |
| - in 2007 (t-1) | 0.082*** | (0.008) | 1,897*** | (311) |
| - in 2006 (t-2) | 0.062*** | (0.011) | 1,658*** | (435) |
| | | | | |
| *Other factors* | | | | |
| - Studying economics | -0.031*** | (0.008) | -548* | (287) |
| - Female | -0.009 | (0.006) | -483** | (225) |
| - Log no. of students in the firm | -0.023*** | (0.004) | -555*** | (139) |
| - Log no. of employees in the firm | 0.020*** | (0.003) | 489*** | (109) |
| | | | | |
| $R^2$ | 0.336 | | 0.866 | |
| No. of observation | 16,743 | | 16,743 | |
| Sample average | 0.41 | | 12,888 | |

Notes: The estimations are run on the sample of individuals, who in 2008 have earnings above the baseline income limit and consequently are at risk of ending up in the dominated region (above the optimal region). I control non-parametrically for income in bins of DKK 4,000. Hence, the variation in earnings in the dominated region is only created by individuals' take-up of benefits and not their earnings potential. Robust standard errors in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$
DKK 6.4 ≈ USD 1.

Second, in Table 4, I estimate the effect on the propensity of having earnings in the dominated region of different sources of knowledge and experience. In the regressions, I control non-parametrically for earnings (dummies for each DKK 4,000 earnings bin) as the risk of ending in the dominated region is present only for relatively high earnings. Hence, the variation in the propensity to end up in the dominated region is only driven by individuals' take-up of benefits and not their earnings potential. The results point to only weak and incomplete learning over time. More years in the educational system is associated with a lower propensity to end up in the dominated region. However, even four years of experience only reduce the propensity by 13 percentage points compared to a sample average of 41 percent.

Interestingly, Table 4 also reveals that individuals, who previously ended up in the dominated region, are more likely to do so again. Learning would predict the opposite. Instead, this result points to the presence of different types of individuals, where some pay less attention than others. Consequently, ending up in the dominated region becomes a signal of low attention, which dominates any learning effects associated with it.

## 5    When and how do individuals re-optimize?

The observed labor market outcomes around each of the three institutional features considered in the previous section consistently point to the presence of significant optimization frictions. In particular, the graphical evidence points to inattention as the dominate optimization friction, while frictions such as real adjustment costs, price misperception and gradual learning appear to be playing a much smaller role.

In this section, I investigate when and how individuals re-optimize as well as the heterogeneity of responses. In Figure 5, I start by showing the share of individuals, who actively adjust their benefit take-up across the earnings distribution. As aforementioned, all individuals are automatically assigned to full benefits at the beginning of the year and for individuals with sufficiently low earnings it is optimal to stay with the default. Once individuals earn DKK 11,600 above the baseline income limit, it becomes optimal to adjust take-up.

The observed behavior in Figure 5 broadly follows these incentives with the majority of individuals staying with the default for low earnings, while increasingly adjusting their take-up for higher earnings. Yet, the alignment between behavior and incentives are far from perfect and even among individuals with earnings significantly above the optimal region 20 percent stay with the default. This result is in line with the strong effect of defaults found in the pension literature (see e.g. Chetty et al., 2014 and Madrian and Shea, 2001) with the difference that inactivity in my setting is associated with a first order utility loss in the form of lost disposable income.
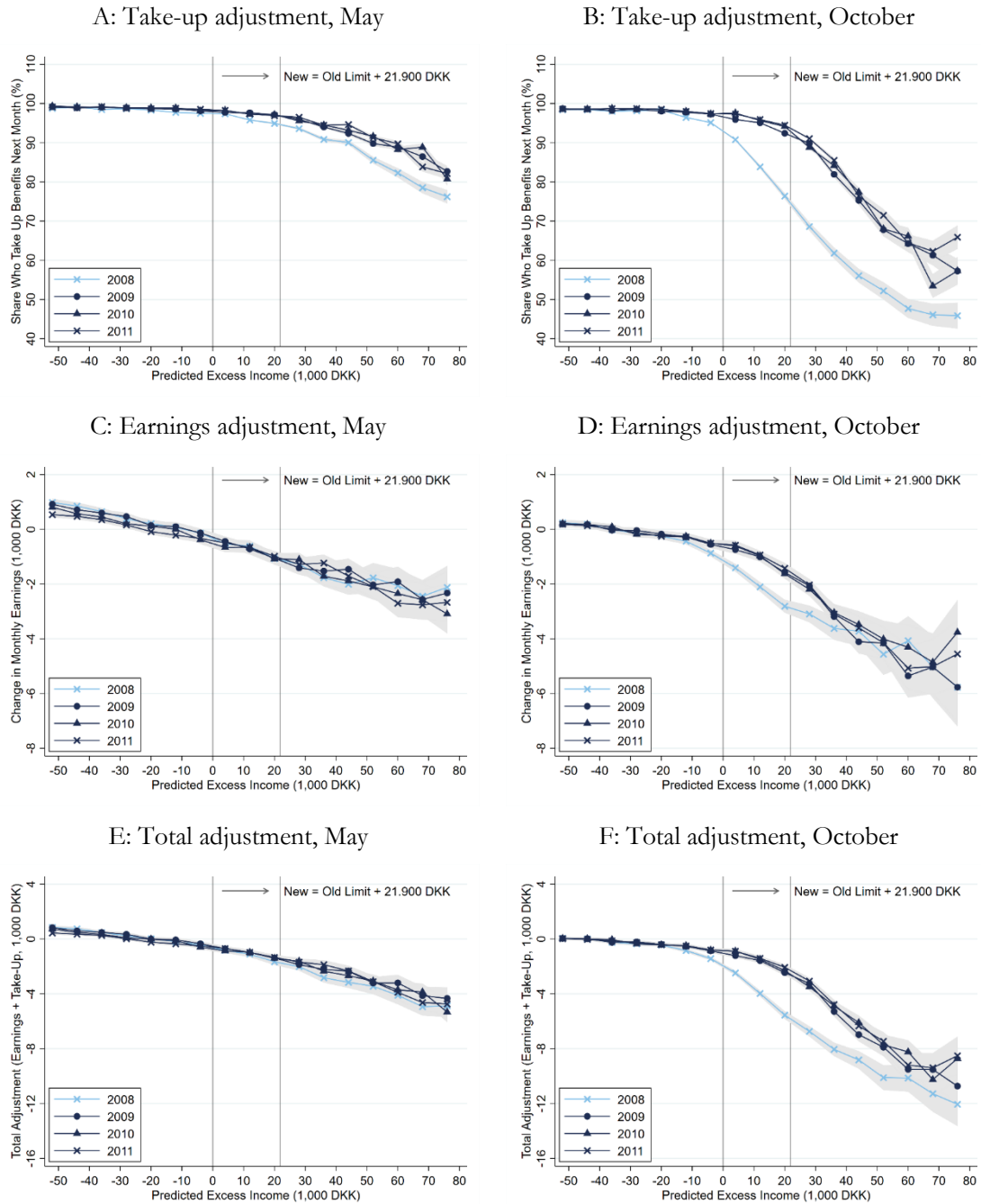
**Figure 5**
**Share who adjust their benefits take-up, 2006-2011**
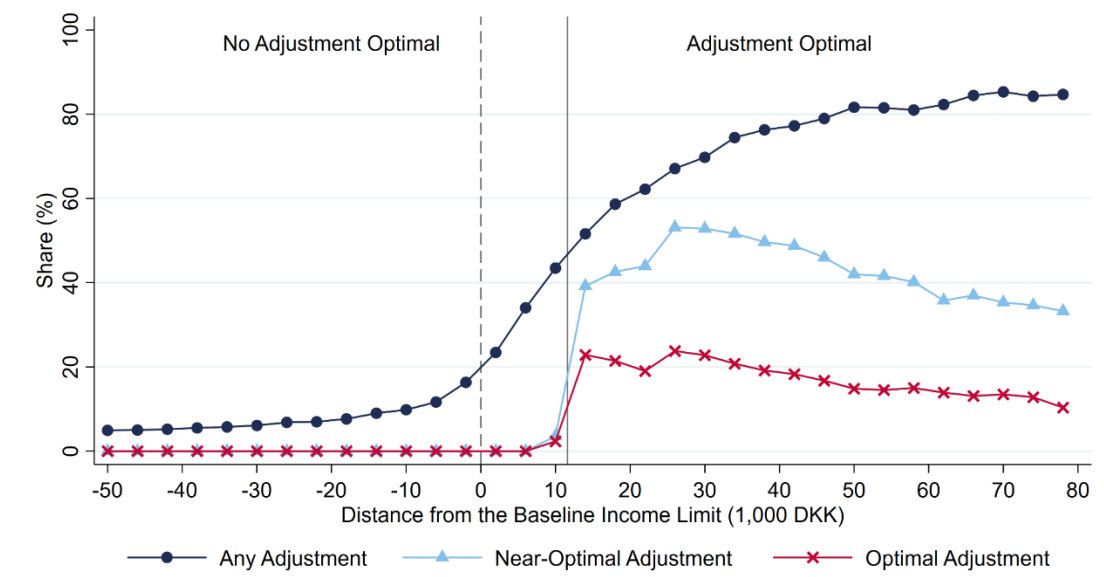


Notes:  The figure shows the share of individuals in each income bin, who adjust their benefits take-up. "Any Adjustment" counts adjustment whether it is optimal or not. "Optimal Adjustment" counts only adjustment that brings the individual into the optimal regions (as illustrated in Figure 1). Below the solid line taking up full benefits is optimal. All individuals are automatically assigned to full benefits at the beginning of the year and this passive optimal behavior is not counted as an adjustment. "Near-Optimal Adjustment" allows for +/- 1 month take-up relative to the optimal. As the figure covers the entire period 2006-2011 it includes the effect of the 2009 reform that shifted the point at which it became optimal to adjust benefits. The effect if the 2009 reform is illustrated in appendix Figure A.VII.
DKK 6.4 ≈ USD 1.

In Figure 5, I also show the share of individuals who adjust their take-up optimally (the take-up that brings them into the optimal region illustrated in Figure 1) and near-optimally (+/- 1 month take-up relative to the optimal). The definition of optimal and near-optimal adjustment is conditional on actually adjusting take-up and are, therefore, by definition zero for excess income below DKK 11,600. From Figure 5, it is clear that a significant share of the individuals who adjust their benefits fail to do so optimally, in particular for individuals with relatively high earnings where less than half of individuals adjust to a near-optimal level. One likely reason for this result is that individuals can only adjust their take-up prior to actually receiving the benefits. Individuals with relative high earnings who optimally have to adjust their take-up by a number of months will therefore have to pay close attention to their earnings process over the year and start adjusting benefits early.

**Figure 6**
**Take-up and earnings adjustments over the year**

A: Take-up adjustment, May

B: Take-up adjustment, October

C: Earnings adjustment, May

D: Earnings adjustment, October

E: Total adjustment, May

F: Total adjustment, October



Notes: The figure shows how individuals' behavior the following month depends on their predicted excess income. For every considered month (m), I construct predicted excess income as the sum of realized earnings up to and including month m plus predicted rest-of-year earnings based on the average earnings in month m and m–1. From this, I subtract the individual income based on observed take-up of benefits up to month m and assuming full take-up the rest of the year. Panel A and B show the share of individuals who take up benefits in month m+1. Panel C and D show the changes in monthly earnings from the average of month m and m–1 to m+1. Panel E and F add together the effect of take-up and earnings adjustments to form the total change in excess income. DKK 6.4 ≈ USD 1.

In Figure 6, I explore how individuals respond to an expected excess income over the year. Using the monthly income data available from 2008 and monthly information on benefits take-up, I can for every month compute an estimate of expected excess income based on individual earnings and take-up behavior so far and the current monthly earnings. Next, I compare observed behavior the following month for different levels of excess income.

In Panel A and B, I show the share who take up benefits in June and November depending on their expected excess income in May and October. The figure reveals that individuals do respond to an expected excess income by taking up less benefits already at the beginning of the year, but the size of the responses increases significantly towards the end of the year. Yet, even close to the end of year where there is less uncertainty on final end-of-year earnings, a significant share of individuals fails to adjust their take-up despite having earnings well above the baseline income limit.[19]

Panel C and D show the change in monthly earnings from the average of the current and previous month to the following month. In almost all months, there is a negative correlation between expected excess income and the change in earnings the following month. However, this pattern might just as well reflect mean reversion and not behavioral responses to an expected excess income. Having relatively high earnings in one month makes it more likely to experience a fall in earnings the next month. To separate these effects, I exploit the shift in the baseline income limit following the 2009 reform, which reveals a clear shift in behavior towards the end of the year (but not earlier). Interestingly, these behavioral responses towards the end of the year appear to be the main driver of the excess mass around the kink points. Looking at the earnings patterns for the first part of the year, there is no difference between before and after the 2009 reform as illustrated in appendix Figure A.VIII.

Panel E and F add together the effect of take-up and earnings adjustments to the total effect on excess income. These panels reveal only limited responses throughout the first half of the year. Towards the end of the year, the observed responses become much stronger, but in Panel F, the average total adjustments are still capped at round DKK 10,000 even for expected excess incomes of more than DKK 50,000. These adjustments are in other words too little, too late in terms of moving individuals into the optimal region.

## 6   Quantifying the costs of optimization frictions

The presence of optimization frictions will typically entail loss of utility for individuals as it pushes them away from their otherwise optimal choice. However, the risk of such utility losses does not necessarily imply that individuals behave irrationally. On the contrary, it

---

[19] The figures for all months are shown in appendix Figure A.IX and Figure A.X. Responses remain very limited until August and builds gradually from September to December.

might be perfectly rational for an individual to accept an expected loss of utility if it is less than the cost of overcoming the frictions. This would be the case in a model with real adjustment cost (Attanasio, 2000) or disutility of paying attention (Sims, 2003).

Chetty (2012) shows that the size of the expected utility losses that individuals are willing to accept plays a key role in determining how individuals respond to incentives. Due to the envelope theorem, behavioral changes to marginal price changes have no first order effect on individual utility and ignoring even sizable price changes will often only entail minor utility losses. This implies that even small frictions limit researchers to identify only bounds on underlying elasticities from observed behavioral responses to price changes. Bounds that in many cases are so wide that they are likely to dwarf many of the econometric issues involved in the estimation. Applied to the labor market, Chetty (2012) shows that widely different estimates of the labor supply elasticity from a range of studies can be reconciled with a single structural elasticity if individuals are willing to accept utility losses from optimization frictions as low as 0.5-1.0 percent of disposable income.

In this section, I use the unique institutional setting surrounding Danish students to provide the first non-parametric estimate of the size of accepted utility losses from optimization frictions. There are two feasible approaches to estimate the size of accepted losses in the current setting. One based on quantifying errors in labor supply and one on quantifying errors in take-up of benefits.

Using the first approach, Chetty et al. (2009) show that the utility loss from choosing labor supply (z) different from the optimal labor supply (z*) can be approximated by

$$u(z) - u(z^*) \approx \frac{1}{2}\frac{1}{\varepsilon}c^* \left(\frac{z - z^*}{z^*}\right)^2,$$ (1)

where c* is disposable income given z.[20] In most cases, this approach is infeasible because we only observe z and not z*. However, in the current setting there is a narrow region of optimal earnings for all individuals who take up less than full benefits, as illustrated in Figure 4. Taking take-up as given, these individuals reveal z* within a narrow interval.

The second approach starts by taking earnings as given and quantifies the utility losses from sub-optimal take-up. For a given level of earnings, taking up too many or too few benefits risks pushing the individual into dominated regions. This creates a utility loss that is linear in excess income as every additional dollar is taxed at 100 percent instead of the 72 percent had

---

[20] Based on a second order Taylor approximation of a standard quasi-linear utility function: $u(x,z) = x - 1/(1+1/\varepsilon)z^{1+1/\varepsilon}$ subject to the budget constraint $c = z - T(z)$. The second order derivative evaluated at z* is $-T''(z^*) - 1/\varepsilon z^{*1/\varepsilon - 1}$. Assuming a locally constant marginal tax rate and plugging in z* gives $-1/\varepsilon(1-T'(z^*))/z^*$. Finally, assuming that the average tax rate equals the marginal, $1-T'(z^*)$ can be replaced by $c^*/z^*$, which gives equation (1). As the approximation is based on quasi-linear utility, the utility loss $u(z)-u(z^*)$ has a direct monetary interpretation.

the individual taken up benefits optimally. This loss is therefore first order compared to the second order loss given by equation (1), and it is independent of the labor supply elasticity and functional form assumptions.

In Figure 7, I compare the two approaches and show the utility loss for an individual, who takes up 10 months of benefits, at different levels of excess income. Assuming that all deviations from the optimal region are labor supply errors, equation (1) implies relatively small utility losses for excess incomes of DKK +/-5,000, but further away the utility losses become significantly larger. For example, an excess income of DKK 20,000 implies a utility loss of around DKK 20,000 corresponding to 15 percent of (optimal) disposable income. Assuming instead that all deviations from the optimal region reflect errors in take-up of benefits, the utility losses are given by the increase in disposable income from moving to optimal take-up. For excess income below zero, disposable income could be increased by taking up more benefits and for excess income above the optimal region, disposable income could be increased by taking up fewer benefits.

**Figure 7**
**Comparing utility losses from labor supply and take-up errors**
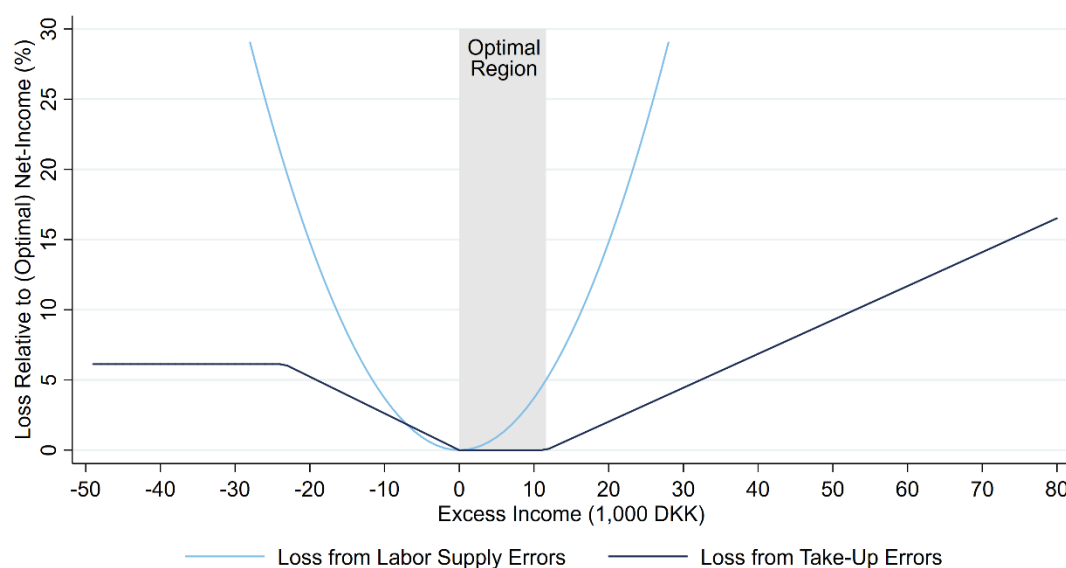


Notes: The figure shows the utility loss for different excess incomes under two different assumptions. First, assuming that all deviations in excess income (relative to excess income = 0) are labor supply errors, the utility loss is given by equation (1). The elasticity in equation (1) is set to 0.1 corresponding to the estimated elasticity in appendix B. Second, assuming instead that earnings always reflect the individuals' optimal labor supply, and that all deviations reflect take-up errors, the utility loss is given by the increase in disposable income by taking up benefits optimally. The increase is illustrated for an individual, who has taken up 10 months of benefits. DKK 6.4 ≈ USD 1.

From Figure 7 it is clear that the first approach based on labor supply errors will generate significantly larger utility losses than the approach based on take-up errors. In Table 5, I

therefore show the average utility losses using the second approach based on errors in benefits take-up.

**Table 5**
**Estimated utility losses from take-up errors**

| Initial take-up of benefits | | Excess income relative to the optimal region | | | |
|---|---|---|---|---|---|
| | | Below | In | Above | Total |
| 12 months (full benefits) | Average loss (DKK) | . | 0 | 6,500 | 300 |
| | Average loss (percent) | . | 0.0 | 5.2 | 0.3 |
| | Share of population | . | 94.9 | 5.1 | 100.0 |
| | | | | | |
| 11 months | Average loss (DKK) | 3,300 | 0 | 5,600 | 3,000 |
| | Average loss (percent) | 2.6 | 0.0 | 4.4 | 2.4 |
| | Share of population | 63.8 | 20.1 | 16.1 | 100.0 |
| | | | | | |
| 10 months | Average loss (DKK) | 5,600 | 0 | 6,000 | 4,700 |
| | Average loss (percent) | 4.2 | 0.0 | 4.6 | 3.5 |
| | Share of population | 62.5 | 17.7 | 19.8 | 100.0 |
| | | | | | |
| 9 months | Average loss (DKK) | 7,500 | 0 | 6,800 | 6,000 |
| | Average loss (percent) | 5.3 | 0.0 | 4.9 | 4.3 |
| | Share of population | 59.0 | 17.6 | 23.4 | 100.0 |
| | | | | | |
| 9-11 months | Average loss (DKK) | 5,000 | 0 | 6,000 | 4,200 |
| | Average loss (percent) | 3.7 | 0.0 | 4.6 | 3.2 |
| | Share of population | 62.3 | 18.7 | 18.9 | 100.0 |

Notes: The table shows the losses in disposable income from sub-optimal take-up of benefits in levels and in percent measured relative to the level of disposable income had the individual taken up benefits optimally. The average utility losses are computed for three parts of the population. Individuals with excess income below zero, where disposable income is increased by taking up more benefits, individuals in the optimal region and individuals above the optimal region, where disposable income is increased by taking up less benefits. Utility losses in the optimal region are by definition 0. For individuals who take-up full benefits, the optimal region covers all earnings from the baseline income limit to zero earnings (see Figure 1), and no one is below the optimal region in this case. DKK 6.4 ≈ USD 1.

Splitting the population by their excess income, Table 5 reveals that individuals who end up with excess income above the optimal region accrue significant utility losses in the order of DKK 6,000 (≈ USD 900) or 4-5 percent of disposable income. For individuals with excess income below the optimal region the average loss is higher than for individual who take up fewer benefits. This reflect that benefits are capped at full take-up and for individuals who take up close to full benefits, the maximum gain of taking up more is therefore lower. Still, average losses are sizable and increase from 2.6 percent of disposable income for individuals who take up 11 months of benefits to 5.2 percent for individuals who take up 9 months of benefits. For individuals who take up between 9 and 11 months of benefits (including those in the optimal region) the average utility loss is DKK 4,200 corresponding to 3.2 percent of disposable income.

As discussed in section 4.3, some of the individuals with excess income below the optimal region might choose to take up fewer benefits to save them for later use. In this case, we should put less weight on the individuals with excess income below the optimal region, when calculating the average utility loss for all students. In the extreme case, where I completely exclude individuals below the optimal region, the average utility loss drops from 3.2 to 2.3 of disposable income.

There are, however, strong indications that a significant share of individuals below the optimal region are there by mistake. First, those with earnings significantly below the optimal region are not predominately individuals, who are likely to exhaust their benefits. In contrast, the distribution of individuals split by the amount of benefits taken up in previous years is almost constant for all levels of excess income as illustrated in appendix Figure A.XI. Second, the individuals below the optimal region are much more likely to have adjusted their benefits take-up in the first half of the year as illustrated in appendix Figure A.XII. This combined with the fact that most individuals adjust in the first half of the year as a response to a predicted excess income above the optimal region (see Figure 6) indicates that individuals end up below the optimal region as a result of negative income shocks towards the end of the year.

Before moving forward, it is worth comparing the approach to quantifying optimization applied here to the earlier approaches in the literature. Kleven and Waseem (2013) quantify optimization frictions as the share of individuals in dominated regions, and they derive formulas that relate this measure (together with the observed excess mass at notch points) to the underlying structural labor supply elasticity. The share of individuals in dominated regions is in other words a sufficient statistic for optimization frictions in their setting. However, compared to the measure of frictions considered by Chetty (2012) and estimated here, it is less likely to be a deep parameter that can be applied across settings. Instead, the share of individuals in a dominated region likely depends on the size of the notch. A small notch would only imply a small utility loss from locating in the dominated region, and we should expect a larger share of individuals located in the region compared to a large notch. A given level of accepted utility losses would predict such differences in behavior.

Gelber, Jones, and Sacks (2013, 2017) quantify optimization frictions from persistent bunching at kink points even after the kink point has been reduced or removed, which they translate into an adjustment cost. This approach uses the same type of variation in the data as I exploited in section 4.2 to uncover labor supply responses. However, while they rely on observing (sharp) bunching in both cross-sections along with parametric assumptions to estimate both the labor supply elasticity and cost of frictions, I rely on the labor supply adjust-

ment to happen fast enough to uncover pre- and post-reform (fuzzy) excess mass and estimate a labor supply elasticity.[21] I estimate the cost of optimization errors non-parametrically using dominated regions.

An estimated range of accepted utility losses between 2-3 percent of disposable income is well above the level considered in the analysis by Chetty (2012) and is sufficient to explain the lack of sharp behavioral responses to even significant kinks or changes in the tax schedule. To see this point, I show, in Figure 8, the utility loss of ignoring a change in the marginal tax rate and instead keeping labor supply fixed at the initial optimal level. As in equation (1), the utility loss of ignoring a change in the marginal tax rate is second order and therefore close to zero for small changes. However, contrary to the apparent intuition in equation (1), the utility loss of ignoring a change in the marginal tax rate is not decreasing in the elasticity. It is true that for any given deviation from the optimal labor supply the utility loss is decreasing in the elasticity, but at the same time, a larger elasticity implies a larger change in the optimal labor supply, which enters squared in equation (1). Both the Danish top tax kink and the kink created by the phase out of the EITC in the US constitute a significant change in the marginal after tax rate of approximately 28 log points, but even for an elasticity as high as 0.5 the utility loss from ignoring the kink is below 3 percent of disposable income.[22]

---

[21] The estimation strategy used to uncover the excess mass in appendix B and section 4.2 is essentially a before/after analysis with the lower parts of the income distribution acting as a placebo group. Hence, it relies on the observed shift to be caused only by the reform, which is more credible if the shift happens quickly following the reform.

[22] In 2018, the Danish top tax kink increased the marginal tax rate from 42.1 to 55.9 percent implying a $\Delta\log(1-m) = 0.283$. The phase out of the US EITC increases the effective marginal tax rate from 15 to 36 percent implying $\Delta\log(1-m) = 0.284$.

**Figure 8**
**Utility loss of ignoring a tax change for different labor supply elasticities**
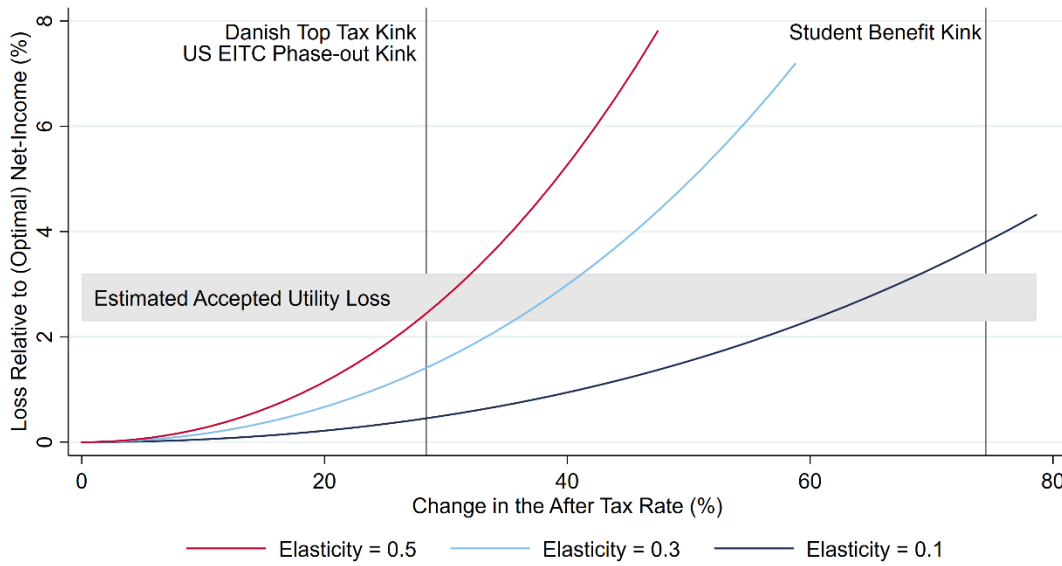


Notes: The utility loss of ignoring a tax change corresponds to the utility loss in equation (1) with $z–z^*$ given as the optimal behavioral response to the tax change: $\Delta\log(z) = \varepsilon\Delta\log(1–m)$. However, here I have computed the utility losses using an exact functional form of the utility function. I start from a standard utility function: $u(x,z) = x–1/(1+1/\varepsilon)z^{(1+1/\varepsilon)}$ and budget constraint: $x = z–T(z)$, where x is consumption, z is labor supply and $T(z)$ is the tax function. I compute the optimal labor supply $(z_0,z_1)$ given both an initial marginal tax rate $m_0 = T_0'(z_0)$ and a new marginal tax rate $m_1 = T_1'(z_1)$. Each of these situations is associated with a utility level $(u_0,u_1)$. The utility loss of ignoring a tax change from $m_0$ to $m_1$ is computed as min $E = x–(z_0–T_1(z_0))$ sub. $u(x,z_0) = u_1$. With quasi-linear utility this is the same as the utility loss from choosing $z_0$ instead of $z_1$ when facing the new $T_1'(z)$. The utility loss is measured relative to the optimal net-of-tax income $z_1–T_1(z_1)$ and as a function of the log change in the net-of-tax rate: $\log(1–m_0) – \log(1–m_1)$. The grey area illustrates the accepted utility losses estimated in Table 5 with and without the individuals below the optimal region.

Figure 8 also provides an explanation as to why inattention appears to be the dominate friction in the student labor market. The kink created by the phase-out of student benefits is large (more than twice the size of Danish top tax kink and the US EITC kink) and even for the estimated elasticity of 0.1, the utility loss of ignoring the kink is above the estimated range of accepted losses. One interpretation of the results is that the kink is sufficiently large to overcome any real adjustment costs or cost of learning and prompts individuals to respond immediately to the 2009 reform. However, this does not imply that individuals perfectly overcome all optimization frictions. Earning shocks over the year will still push them away from their desired earnings, but the expected utility losses from these deviations are not high enough for individuals to pay closer attention.

## 7   Conclusion

Despite growing recognition of the importance of optimization frictions both in understanding labor market outcomes and designing optimal policy, concrete evidence of optimization frictions is still limited. The limited evidence essentially reflects that identification of optimization frictions is demanding in terms of high-quality data and in particular in terms of special institutional settings. In this paper, I exploit the unique institutional setting facing Danish students to study both the magnitude and nature of optimization frictions in a real-life setting. I have presented two sets of results.

First, following a reform in 2009 that significantly increased the earnings threshold above which student benefits are phased out, I find an immediate and non-trivial shift in the earnings distribution compared to stable distributions both before and after the reform. The shift is consistent with a labor supply elasticity of 0.1, which is well within the interval found by Kleven and Schultz (2014) for the full Danish population. However, despite this clear evidence of a positive labor supply elasticity, I find no sharp bunching at the kink point created by the phase-out of student benefits predicted by standard theory (Saez, 2010).

Second, I find that a significant share of individuals fails to take up benefits optimally given their final end-of-year earnings. As a unique feature of the Danish student benefit system, individuals can actively adjust their take up of benefits and the stakes of this adjustment are large. For individuals with relative high earnings, passively maintaining the default of full benefits take-up can entail losses of more than 10 percent of disposable income. Yet, among this group 20 percent remain passive throughout the entire year, and  even towards the end of the year, where there is less uncertainty of final end-of-year earnings, 40 percent remain passive. Overall, individuals appear to accept average losses of approximately 2-3 percent of disposable income from optimization errors.

I take these findings as strong evidence for the presence of optimization frictions. However, at the same time the findings are inconsistent with many of the main types of frictions commonly discussed in the economic literature. For example, both real adjustment costs and gradual learning would have predicted a much more gradual response to the 2009 reform.

Instead, the findings are consistent with a model with inattention, where individuals (rationally) choose their desired labor supply, but where their actual earnings over the year may be affected by unexpected shocks. As a consequence, final end-of-year earnings may deviate from the desired level, if individuals do not pay sufficient attention and re-optimize behavior.

Of course, the relative strength of the different types of frictions is not necessarily directly transferable to other labor markets. On the contrary, my finding of inattention as the dominate friction is well explained by the institutional setting in combination with the estimated

labor supply elasticity of 0.1 and the utility losses due to optimization frictions that individuals appear willing to accept. Simply put, the kink in my setting is more than twice the size of the largest kinks in the current US and Danish tax schedules and hence sufficiently large to overcome any real adjustment costs or cost of learning and prompts individuals to respond immediately to the higher kink point following the 2009 reform.

# References

Attanasio, O. P. 2000. "Consumer Durables and Inertial Behaviour: Estimation and Aggregation of (S, s) Rules for Automobile Purchases." *The Review of Economic Studies* 67 (4): 667–96. https://doi.org/10.1111/1467-937X.00149.

Bastani, Spencer, and Håkan Selin. 2014. "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule." *Journal of Public Economics* 109 (January): 36–49. https://doi.org/10.1016/J.JPUBECO.2013.09.010.

Best, Michael Carlos, and Henrik Jacobsen Kleven. 2013. "Optimal Income Taxation with Career Effects of Work Effort." Working Paper.

Chetty, Raj. 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica.* The Econometric Society. https://doi.org/10.2307/41493842.

Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark." *The Quarterly Journal of Economics* 129 (3): 1141–1219. https://doi.org/10.1093/qje/qju013.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2009. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." 15617. NBER Working Paper Series. Cambridge, MA. https://doi.org/10.3386/w15617.

———. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics* 126 (2): 749–804. https://doi.org/10.1093/qje/qjr013.

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* 103 (7): 2683–2721. https://doi.org/10.1257/aer.103.7.2683.

Evans, George W., and Seppo Honkapohja. 2001. *Learning and Expectations in Macroeconomics.* Princeton University Press. https://www.jstor.org/stable/j.ctt7s6t9.

Farhi, Emmanuel, and Xavier Gabaix. 2018. "Optimal Taxation with Behavioral Agents." Working Paper.

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* 103 (3): 551–72. https://doi.org/10.1086/261994.

Gabaix, Xavier. 2017. "Behavioral Inattention." 24096. NBER Working Paper Series. https://doi.org/10.3386/w24096.

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2013. "Earnings Adjustment Frictions:

Evidence from the Social Security Earnings Test." 19491. NBER Working Paper Series. https://doi.org/10.3386/w19491.

———. 2017. "Estimating Earnings Adjustment Frictions: Method and Evidence from the Social Security Earnings Test." Working Paper.

Goldin, Jacob, and Daniel Reck. 2018. "Optimal Defaults with Normative Ambiguity." Working Paper. https://www.danreck.com/s/defaults.pdf.

Gruber, Jonathan, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84 (1): 1–32. https://doi.org/10.1016/S0047-2727(01)00085-8.

Hoopes, Jeffrey L., Daniel Reck, and Joel Slemrod. 2015. "Taxpayer Search for Information: Implications for Rational Attention." *American Economic Journal: Economic Policy* 7 (3): 177–208. https://doi.org/10.1257/pol.20140050.

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review* 104 (2): 537–63. https://doi.org/10.1257/aer.104.2.537.

Keane, Michael, Petra E. Todd, and Kenneth I. Wolpin. 2011. "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, 4th ed. Elsevier. https://www.sciencedirect.com/science/article/pii/S0169721811004102.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8 (1): 435–64. https://doi.org/10.1146/annurev-economics-080315-015234.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez. 2011. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." *Econometrica* 79 (3): 651–92. https://doi.org/10.3982/ECTA9113.

Kleven, Henrik Jacobsen, and Esben Anton Schultz. 2014. "Estimating Taxable Income Responses Using Danish Tax Reforms." *American Economic Journal: Economic Policy* 6 (4): 271–301. https://doi.org/10.1257/pol.6.4.271.

Kleven, Henrik Jacobsen, and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *The Quarterly Journal of Economics* 128 (2): 669–723. https://doi.org/10.1093/qje/qjt004.

Liebman, Jeffrey B., and Richard J. Zeckhauser. 2003. "Schmeduling." Working Paper. http://www.people.hbs.edu/bhall/NOMTalks/papers/schmeduling_apr62003.pdf.

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 1149–87. https://doi.org/10.1162/003355301753265543.

Maire, Daniel le, and Bertel Schjerning. 2013. "Tax Bunching, Income Shifting and Self-Employment." *Journal of Public Economics* 107: 1–18. https://www.sciencedirect.com/science/article/pii/S0047272713001618.

Matikka, Tuomas, and Tuomas Kosonen. 2017. "Discrete Earnings and Optimization Errors: Evidence from Student's Responses to Local Tax Incentives." Working Paper.

Reck, Daniel. 2016. "Taxes and Mistakes: What's in a Sufficient Statistic?" Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2268617.

Rees-Jones, Alex, and Dmitry Taubinsky. 2016. "Measuring 'Schmeduling.'" 22884. NBER Working Paper Series. dge, MA. https://doi.org/10.3386/w22884.

Saez, Emmanuel. 1999. "Do Taxpayers Bunch at Kink Points?" 7366. NBER Working Paper Series. https://doi.org/10.1257/pol.2.3.180.

———. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212. https://doi.org/10.1257/pol.2.3.180.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–90. https://doi.org/10.1016/S0304-3932(03)00029-1.

# Appendix

## Intended for online publication

## Appendix A: Supplementary figures and tables

**Figure A.I**
**Homepage through which students adjust their benefits take-up**



Notes: Screenshot from www.su.dk

**Figure A.II**
**Descriptive statistics for students close to the baseline income limit, 2008**

A: Hourly wage rate

B: Hours worked per week

C: Take-up of benefits

D: Share with a study related job

E: Std. of monthly earnings

F: Normalized std. of monthly earnings
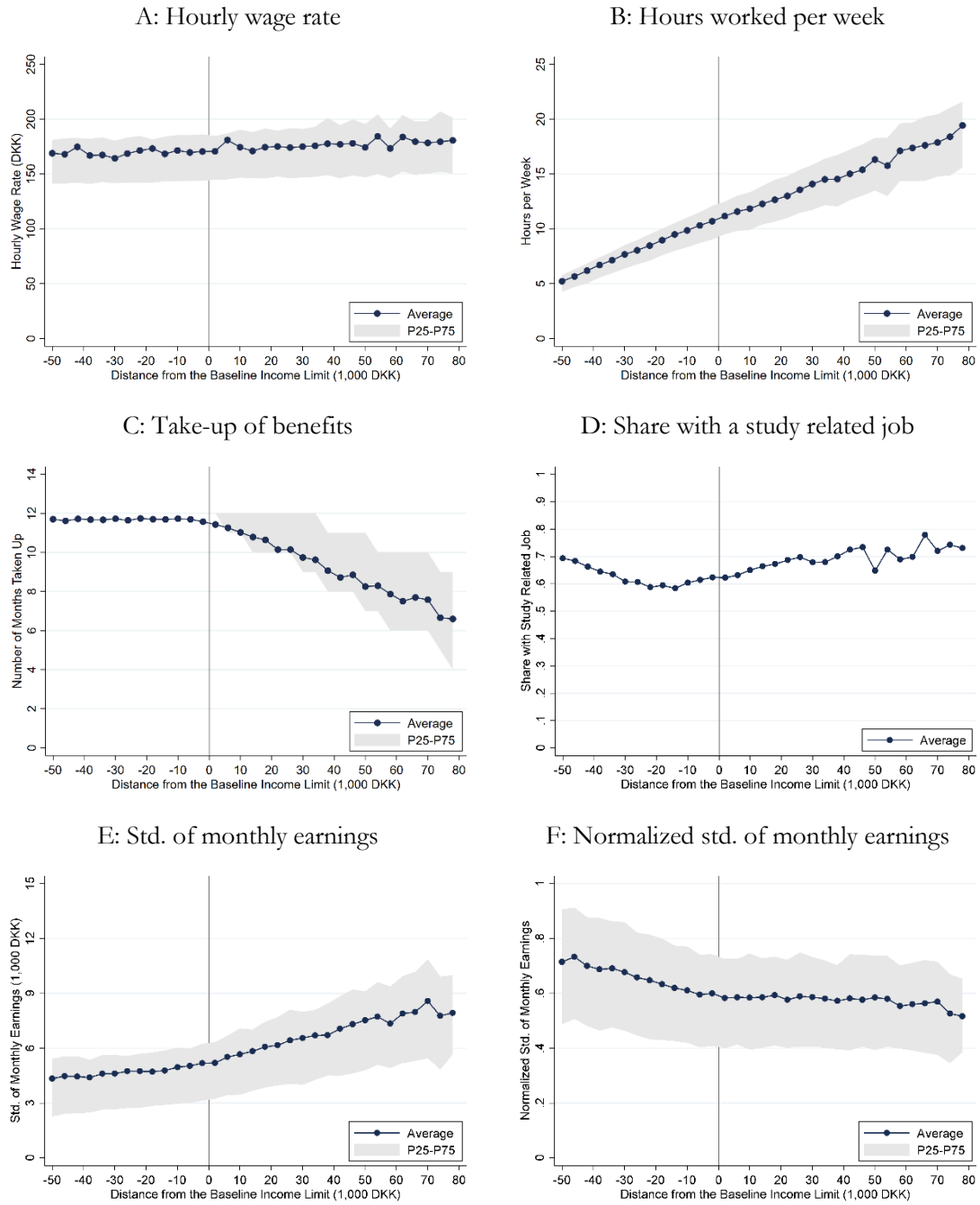
Notes: The figure shows key descriptive statistics for students with earnings relatively close to the baseline income limit (where the phase-out of benefits starts). The dotted lines show averages within each DKK 4,000 bin, while the shaded area shows the range between the 25th and 75th percentile of the distribution.
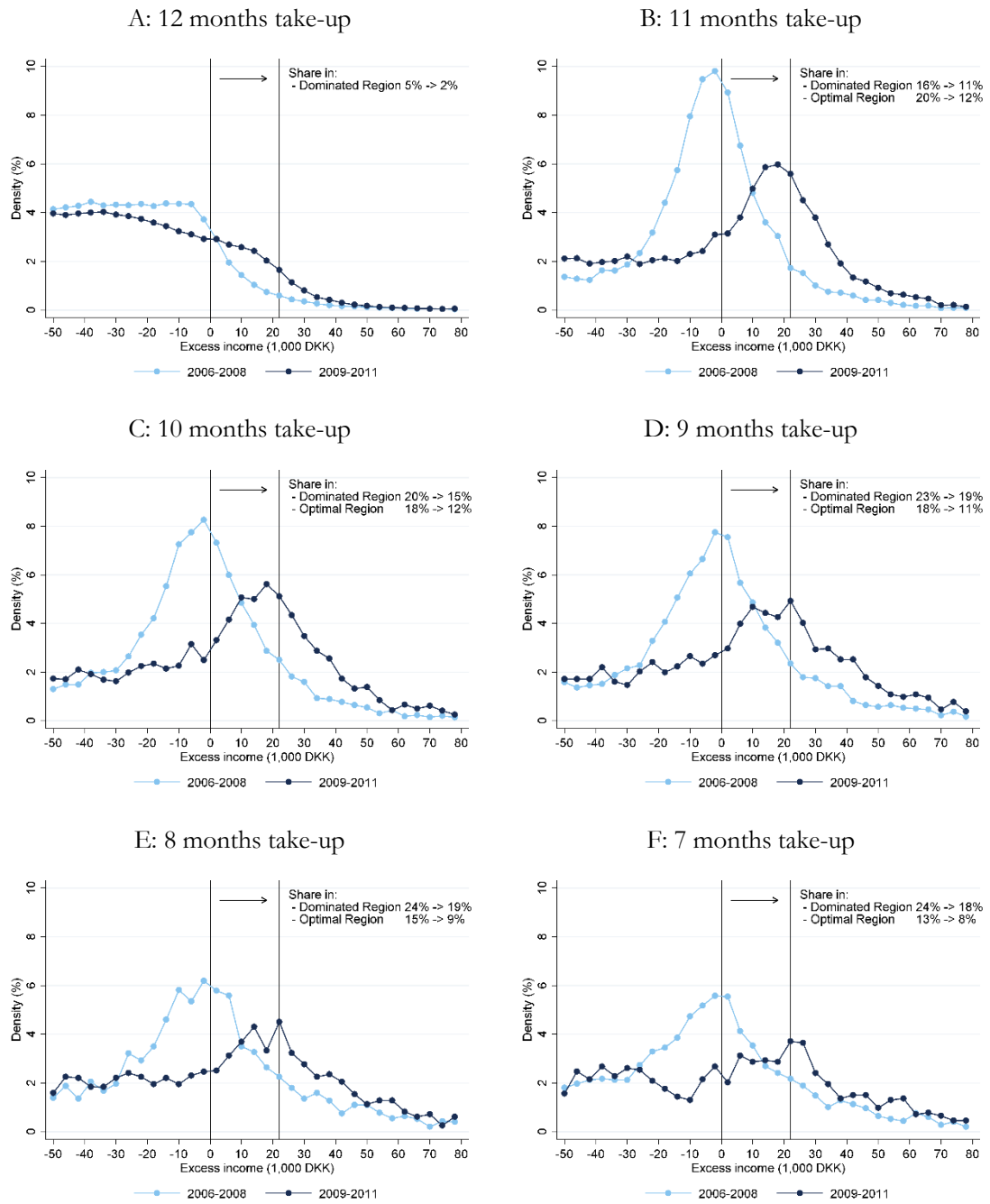DKK 6.4 ≈ USD 1.

**Figure A.III**
**The earnings distribution for individuals not affected by the 2009 reform, 2006-2011**

A: Individuals age 18-35 not in enrolled in education
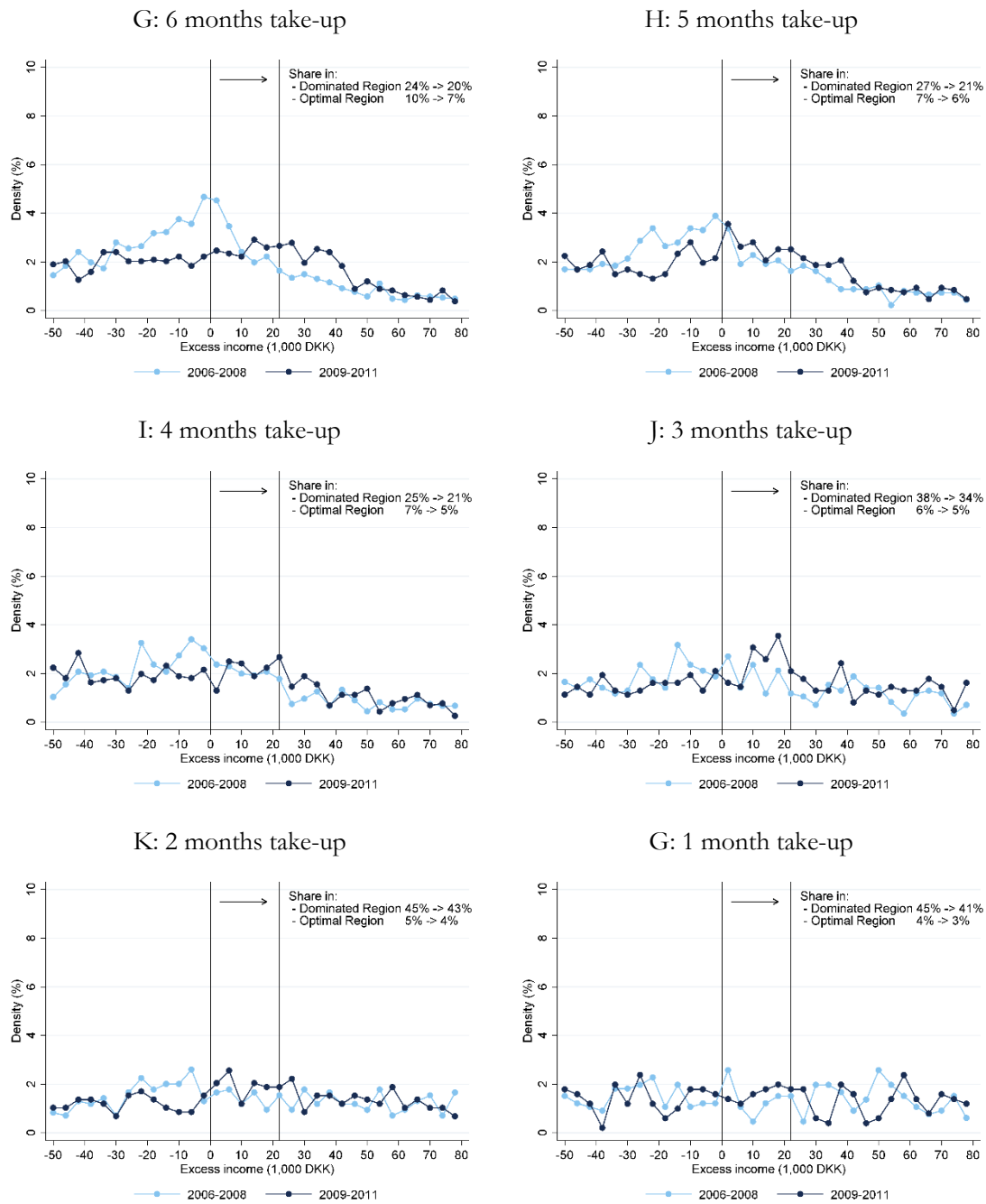


B: Students in lower educations
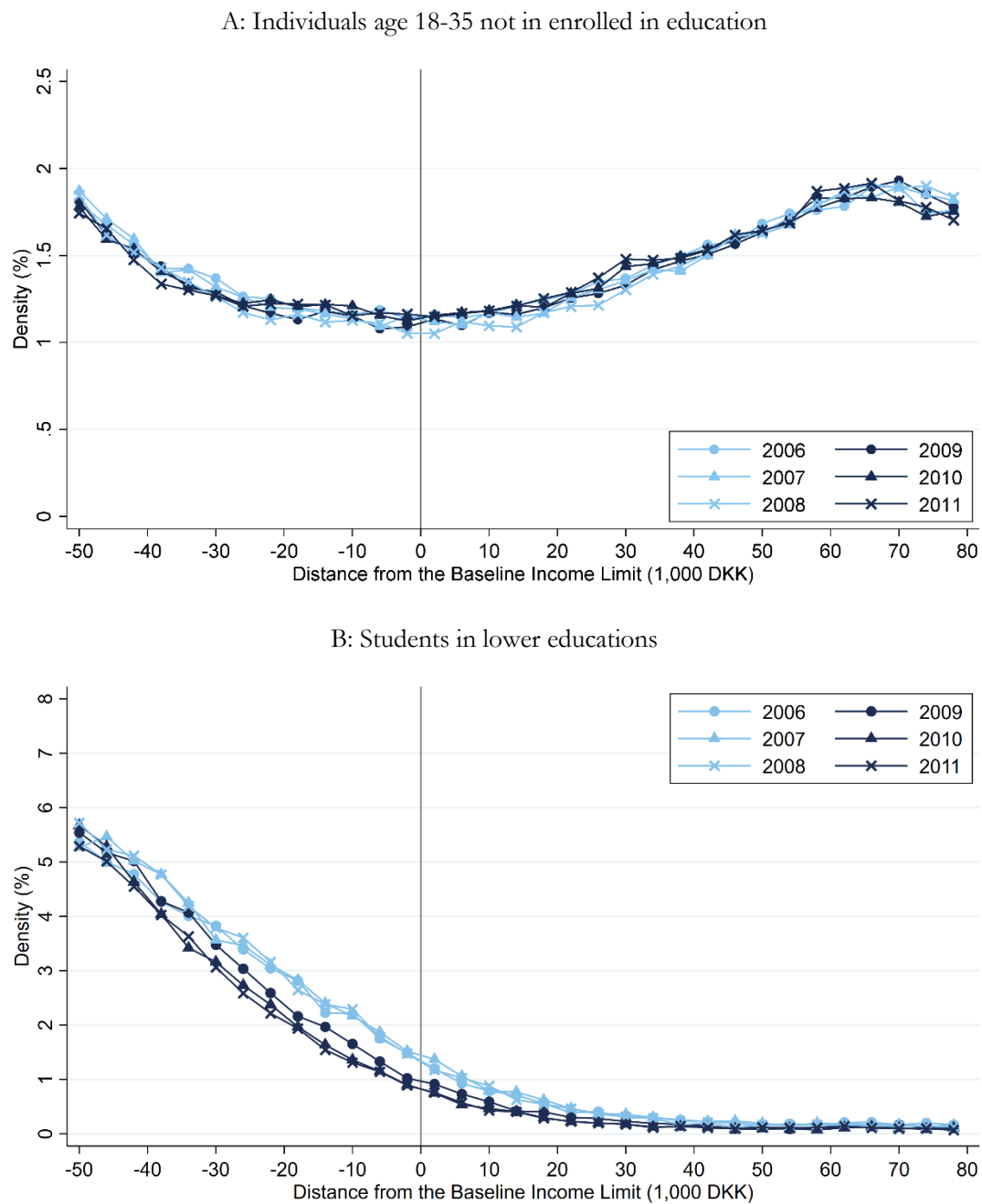


Notes: The figure shows the distributions of earned income for two groups not affected by the 2009 re-form. In Panel A, individuals age 18-35 not enrolled in education with their income centered around a (placebo) baseline income limit of DKK 93,000. In Panel B, students enrolled in lower education (high schools and vocational training), fully eligible for student benefits and with yearly earnings above DKK 8,000. The distribution is not conditional on the actual take-up of benefits. DKK 6.4 ≈ USD 1.

**Figure A.IV**
**The distribution of excess income conditional on benefits take-up, 2006-2011**

A: 12 months take-up



B: 11 months take-up



C: 10 months take-up



D: 9 months take-up



E: 8 months take-up



F: 7 months take-up

## G: 6 months take-up



Share in:
- Dominated Region 24% -> 20%
- Optimal Region    10% -> 7%

## H: 5 months take-up



Share in:
- Dominated Region 27% -> 21%
- Optimal Region    7% -> 6%

## I: 4 months take-up



Share in:
- Dominated Region 25% -> 21%
- Optimal Region    7% -> 5%

## J: 3 months take-up



Share in:
- Dominated Region 38% -> 34%
- Optimal Region    6% -> 5%

## K: 2 months take-up



Share in:
- Dominated Region 45% -> 43%
- Optimal Region    5% -> 4%

## G: 1 month take-up



Share in:
- Dominated Region 45% -> 41%
- Optimal Region    4% -> 3%

Notes: The figure shows the distributions of earned income relative to the individual income limit for students enrolled in tertiary education, fully eligible for student benefits and with yearly earnings above DKK 8,000. Each panel shows the distribution conditional on the amount of benefits taken-up. For less than full take-up (Panel B-G) there is a narrow region of optimal earnings, while the optimal region for students with full benefits (Panel A) is given by all earning levels below an excess income of DKK 11,600. The solid vertical lines indicate the shift in the baseline income limit following the 2009 reform.
DKK 6.4 ≈ USD 1.

**Figure A.V**
**The distribution of imputed excess income conditional on less than full take-up**
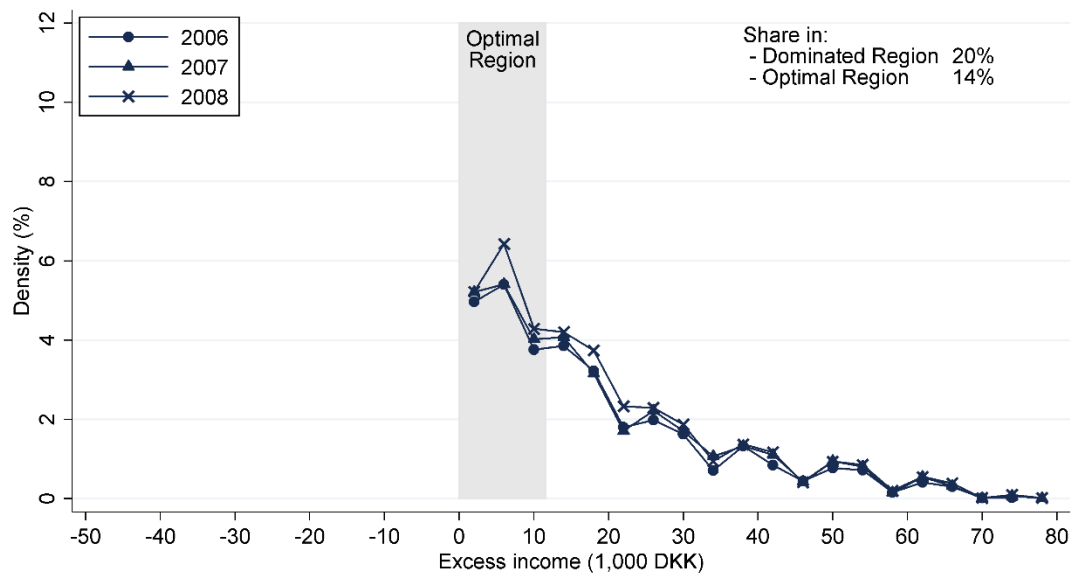


Notes: The figure shows the distributions of imputed excess income for students enrolled in tertiary education, fully eligible for student benefits and with yearly earnings above DKK 8,000. I only include individuals who take up less than full benefits as these individuals have a well-defined region of optimal earnings (excess income between DKK 0 and 11,600). Excess income is imputed based on the observed difference between the initial payout of benefits from the Student Benefit Administration and the final end-of-year benefits observed in the tax record. This measure is independent of earnings and income limits used in Figure 4, but the imputation is only feasible for earnings above the income limit. The reported share in the dominated region only counts individuals above the optimal region.
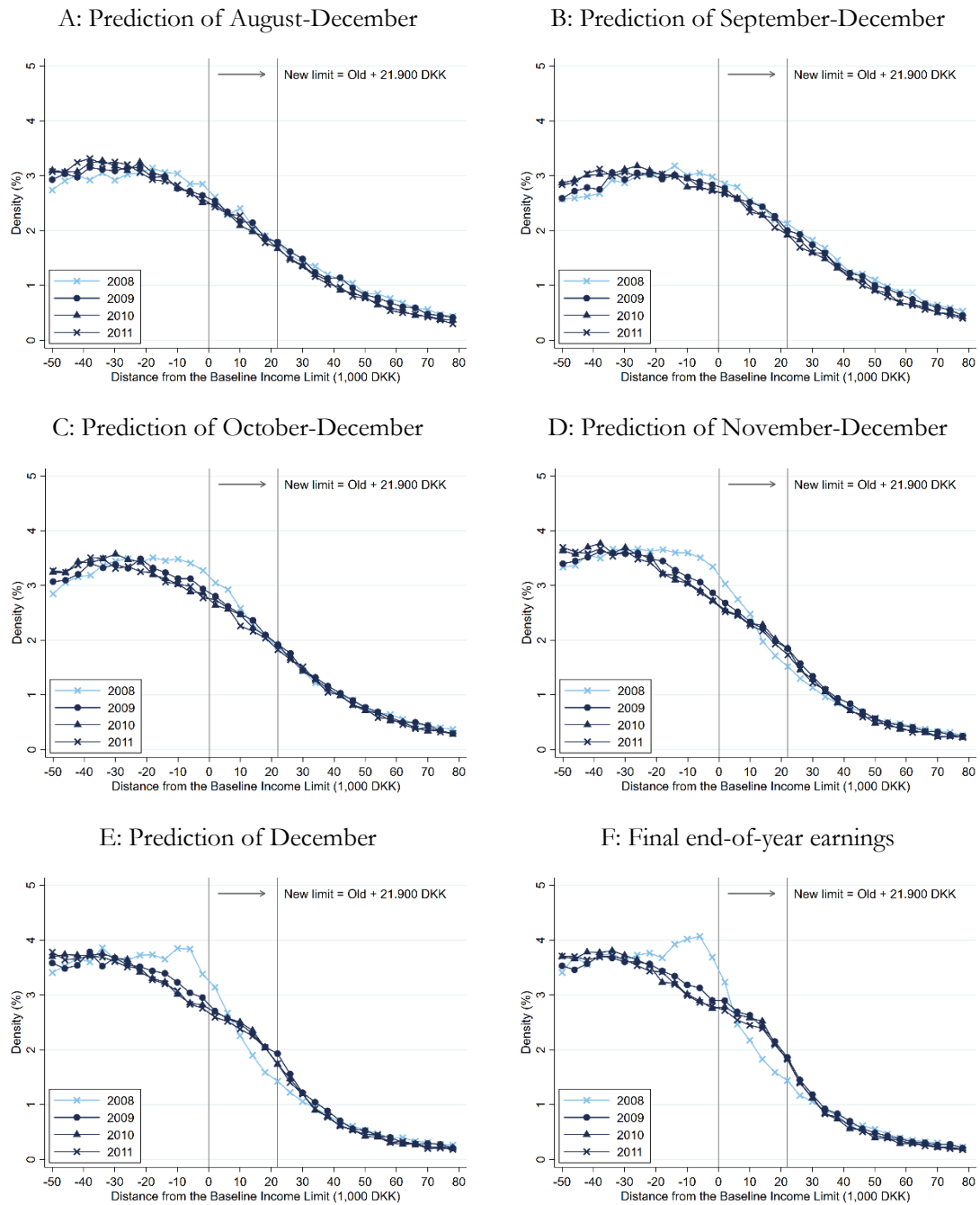
DKK 6.4 ≈ USD 1.

**Figure A.VI**
**The distribution of excess income for different levels of student experience, 2008**
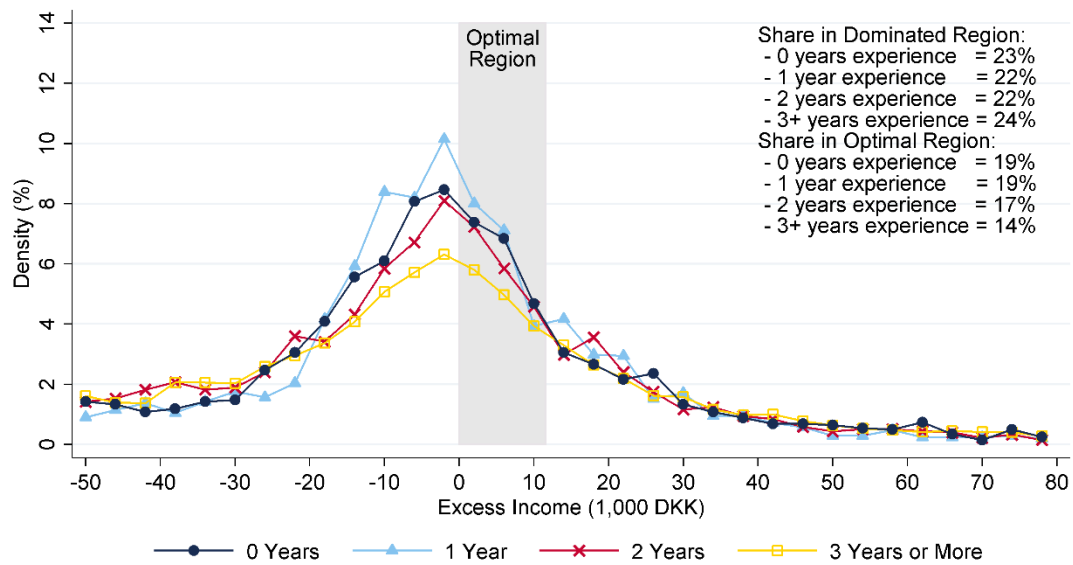


Notes: The figure shows the distributions of earned income relative to the individual income limit for students enrolled in tertiary education, fully eligible for student benefits and with yearly earnings above DKK 8,000. I only include individuals who take up less than full benefits as these individuals have a well-defined region of optimal earnings (excess income between DKK 0 and 11,600). Student experience is measured as the number of years in tertiary education between 2004 and 2007. The reported share in the dominated region only counts individuals above the optimal region.
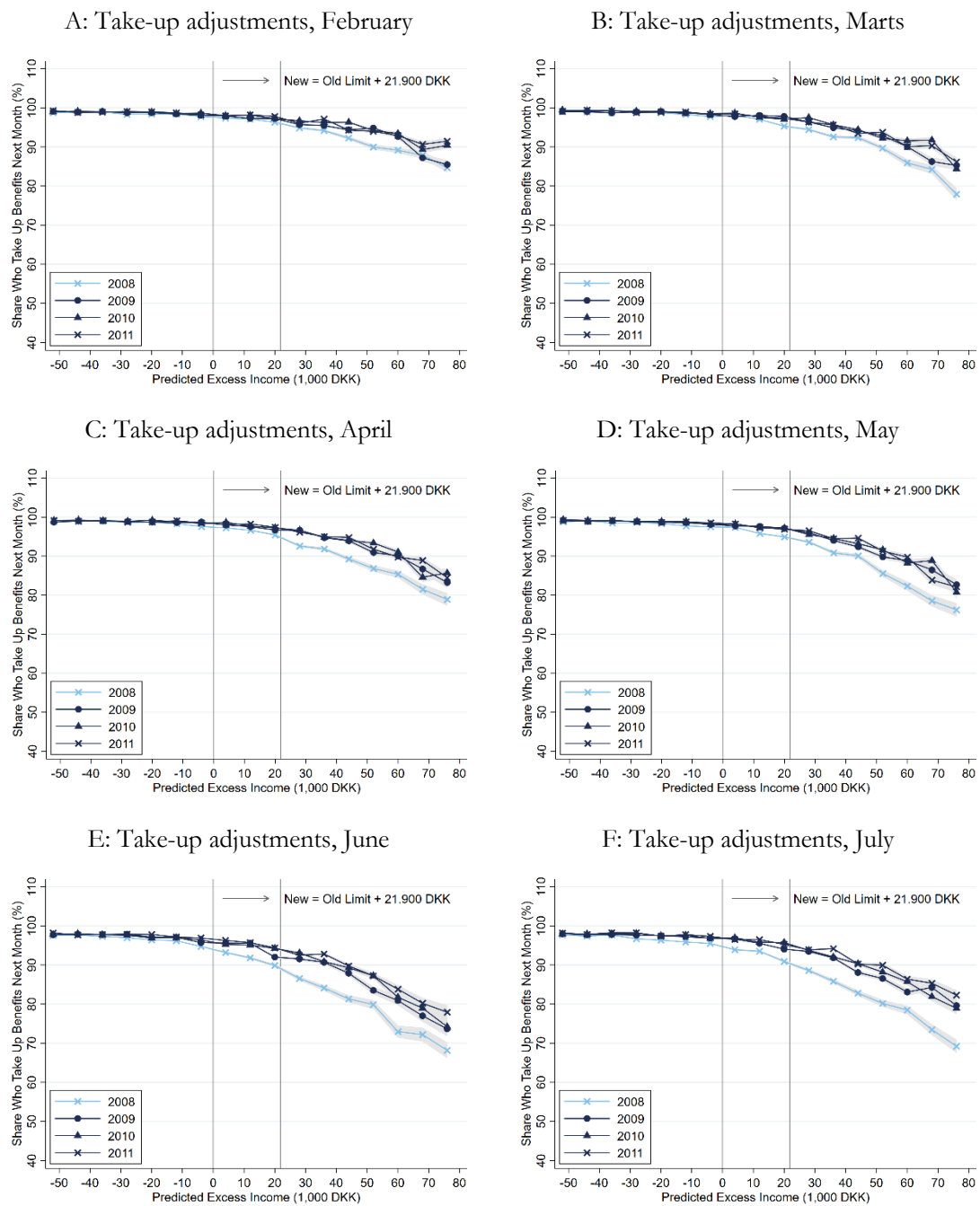DKK 6.4 ≈ USD 1.

**Figure A.VII**
**Share who adjust their benefits take-up before and after the 2009 reform**
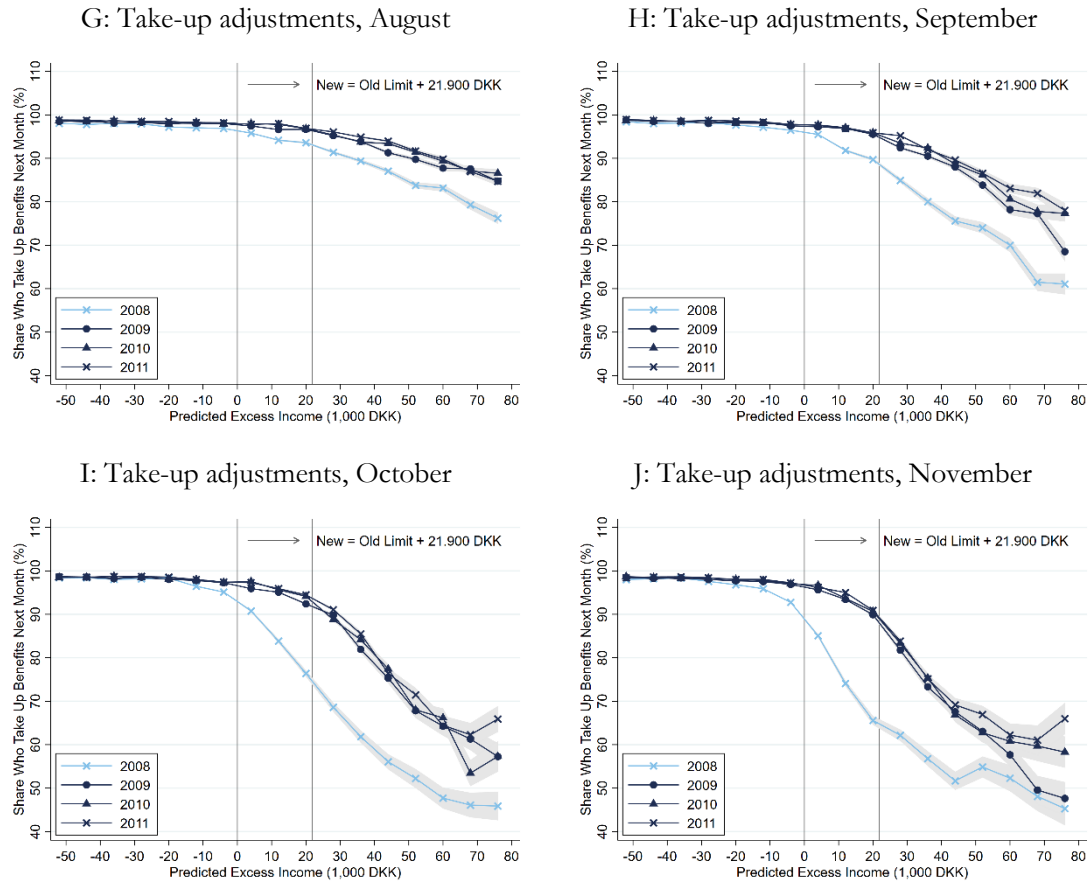
A: Any Adjustment



B: Optimal Adjustment
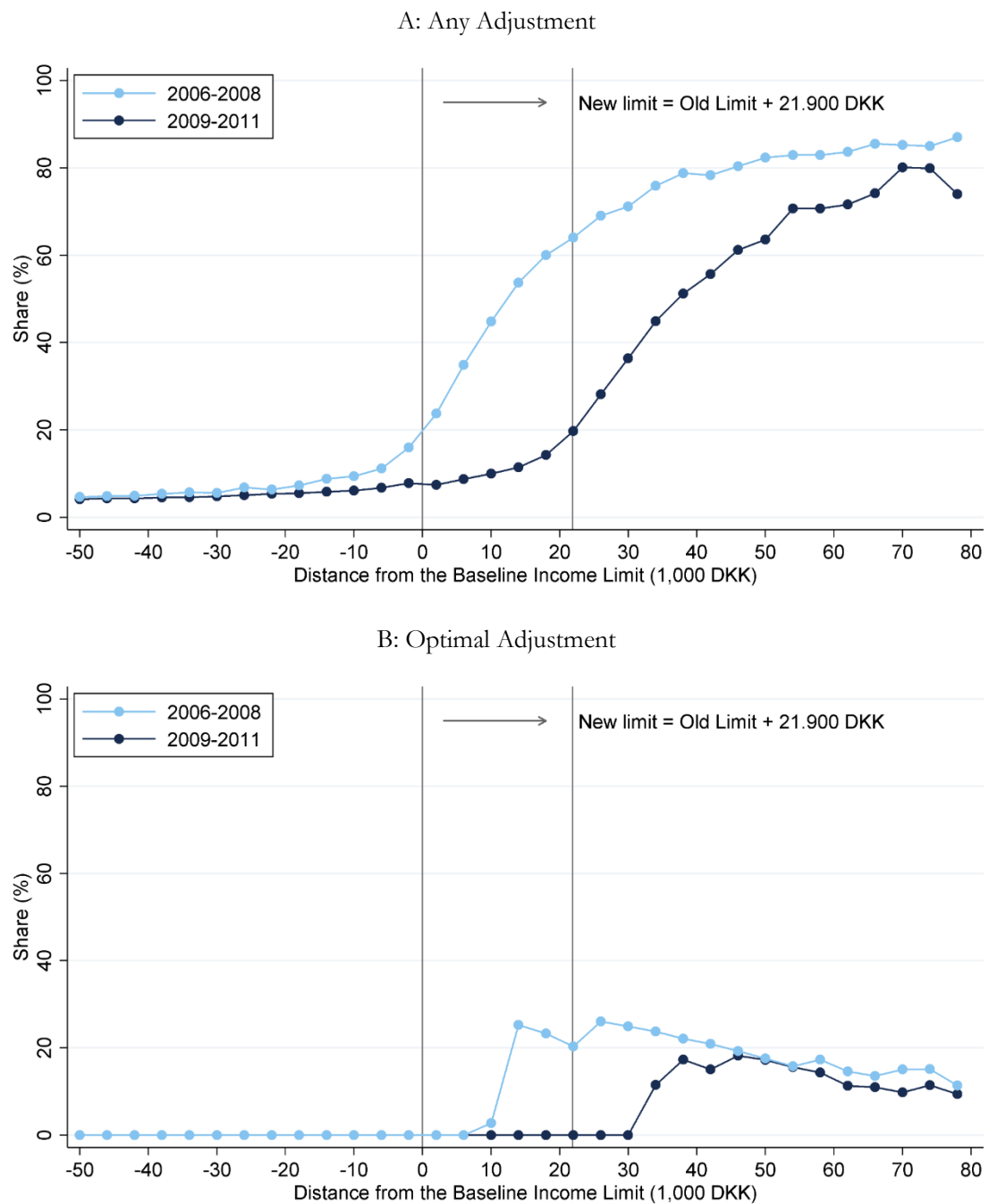


Notes: The figure shows the share of individuals in each DKK 4,000 income bin, who adjust their benefits take-up. "Any Adjustment" counts adjustment whether it is optimal or not. "Optimal Adjustment" only counts adjustment that brings the individual into the optimal regions (as illustrated in Figure 1 in the main paper). All individuals are automatically assigned to full benefits at the beginning of the year and this passive optimal behavior is not counted as an adjustment. The 2009 reform shifted the level at which it becomes optimal to adjust take-up to the right as illustrated in Panel B. DKK 6.4 ≈ USD 1.

**Figure A.VIII**
**Distributions of predicted earnings before and after the 2009 reform**

A: Prediction of August-December



B: Prediction of September-December



C: Prediction of October-December



D: Prediction of November-December



E: Prediction of December



F: Final end-of-year earnings



Notes: The figure shows the distribution of predicted earnings (Panel A-E) and final end-of-year earnings (Panel F) before and after the 2009 reform. For each month m, I construct predicted earnings as the sum of realized earnings up to and including month m plus predicted rest-of-year earnings based on the average earnings in month m and m–1.

DKK 6.4 ≈ USD 1.

**Figure A.IX**
**Take-up adjustments over the year**

A: Take-up adjustments, February



B: Take-up adjustments, Marts



C: Take-up adjustments, April



D: Take-up adjustments, May



E: Take-up adjustments, June



F: Take-up adjustments, July

## G: Take-up adjustments, August



## H: Take-up adjustments, September



## I: Take-up adjustments, October



## J: Take-up adjustments, November



Notes: The figure shows individuals' behavior the following month depending on their predicted excess income. For every considered month (m), I constructed predicted excess income as the sum of realized earnings up to and including month m plus predicted rest-of-year earnings based on the average earnings in month m and m–1. From this, I subtract the individual income based on observed take-up of benefits up to month m and assuming full take-up the rest of the year. Each panel show the share of individuals who take up benefits in month m+1 within each bin (DKK 8,000 bins).
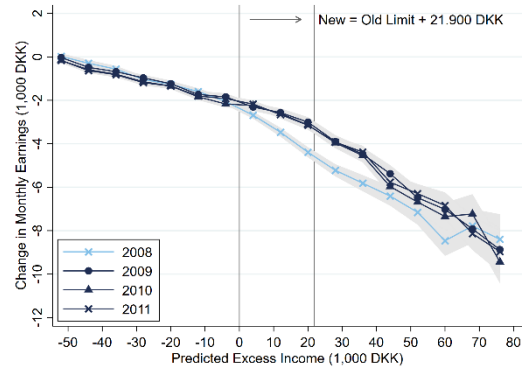DKK 6.4 ≈ USD 1.
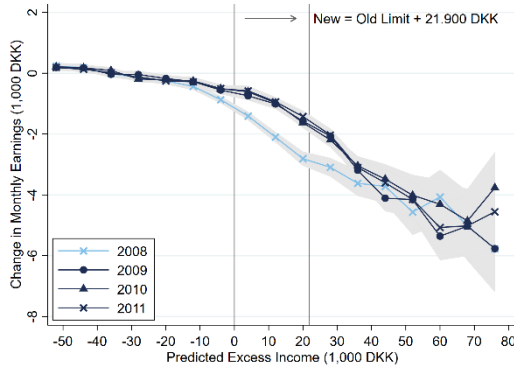
**Figure A.X**
**Earnings adjustments over the year**

A: Earnings adjustments, February



B: Earnings adjustments, Marts



C: Earnings adjustments, April



D: Earnings adjustments, May



E: Earnings adjustments, June



F: Earnings adjustments, July
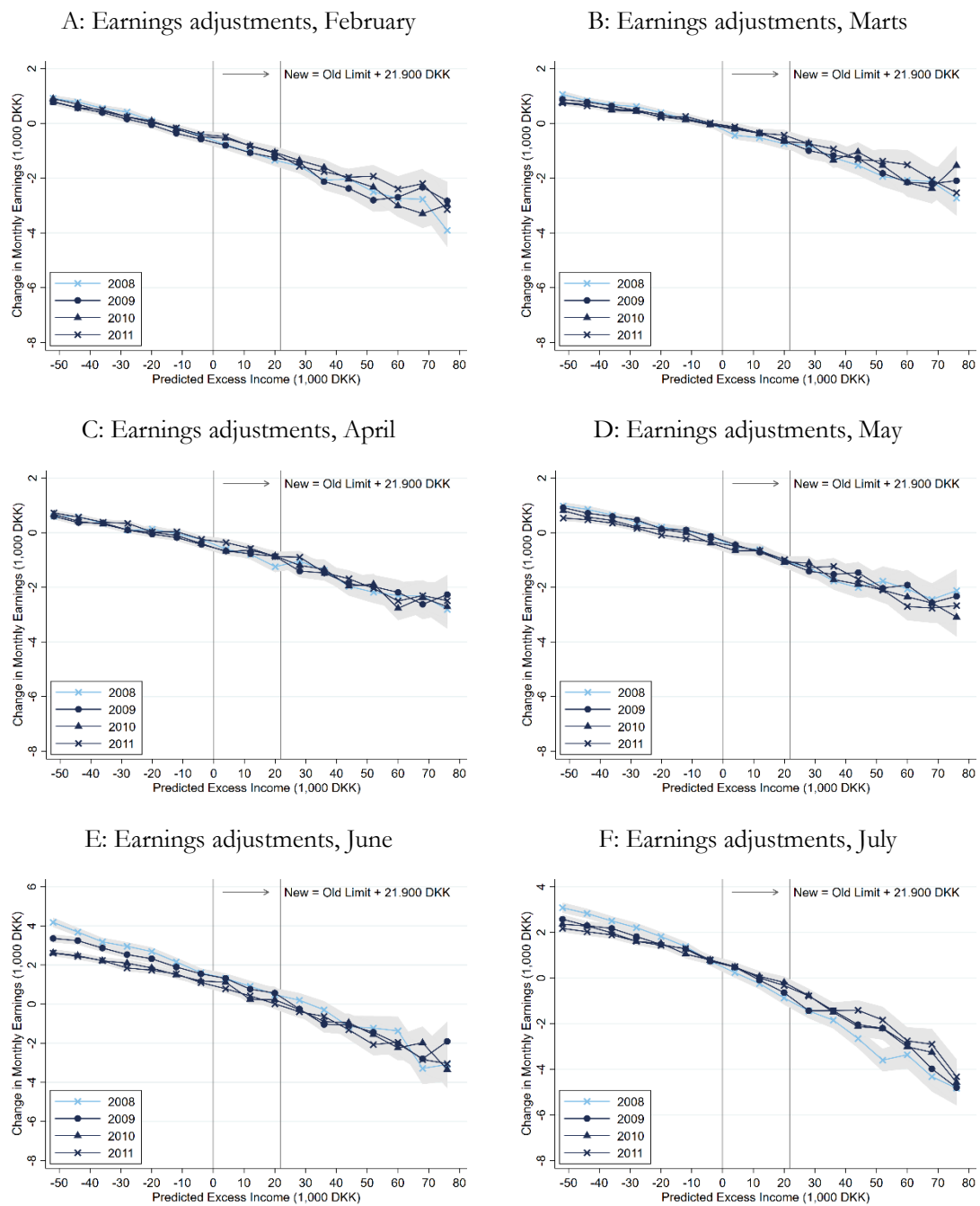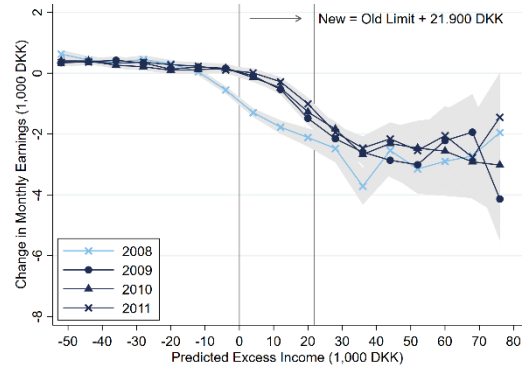
### G: Earnings adjustments, August
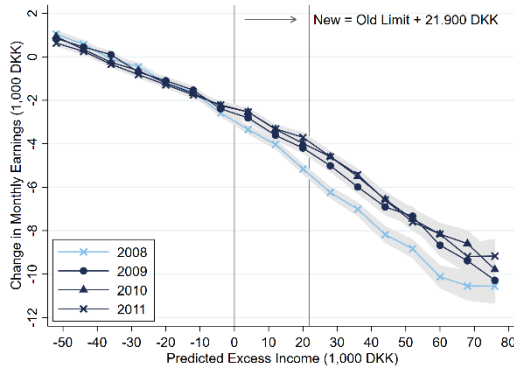


### H: Earnings adjustments, September



### I: Earnings adjustments, October



### J: Earnings adjustments, November



Notes: The figure shows individuals' behavior the following month depending on their predicted excess income. For every considered month (m), I construct predicted excess income as the sum of realized earnings up to and including month m plus predicted rest-of-year earnings based on the average earnings in month m and m−1. From this, I subtract the individual income based on observed take-up of benefits up to month m and assuming full take-up the rest of the year. Each panel shows the changes in monthly earnings from the average of month m and m−1 to m+1 within each bin (DKK 8,000 bins).
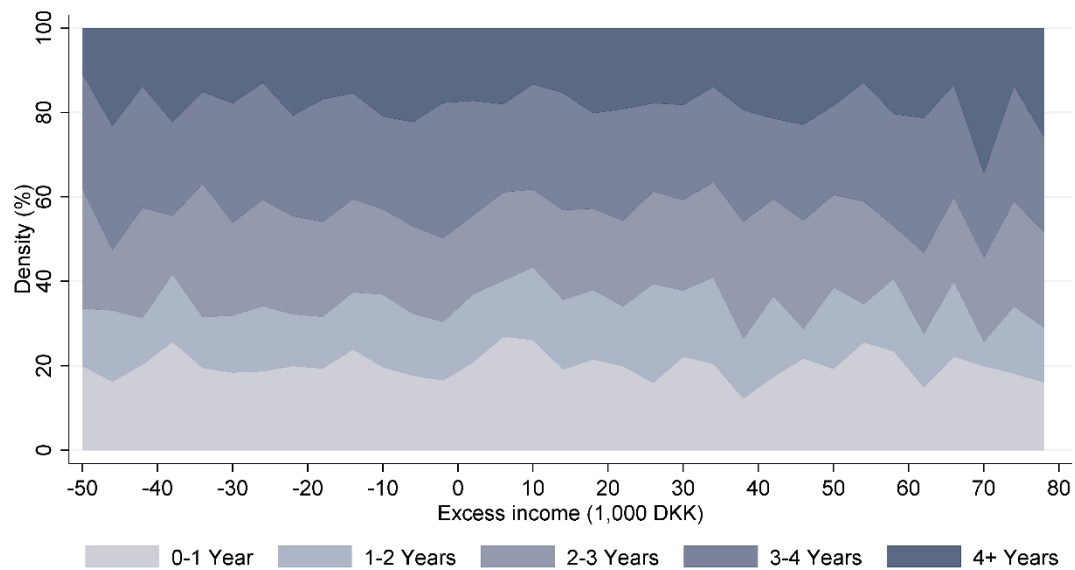
DKK 6.4 ≈ USD 1.

**Figure A.XI**
**The distribution of excess income split by previous benefit take-up, 2011**
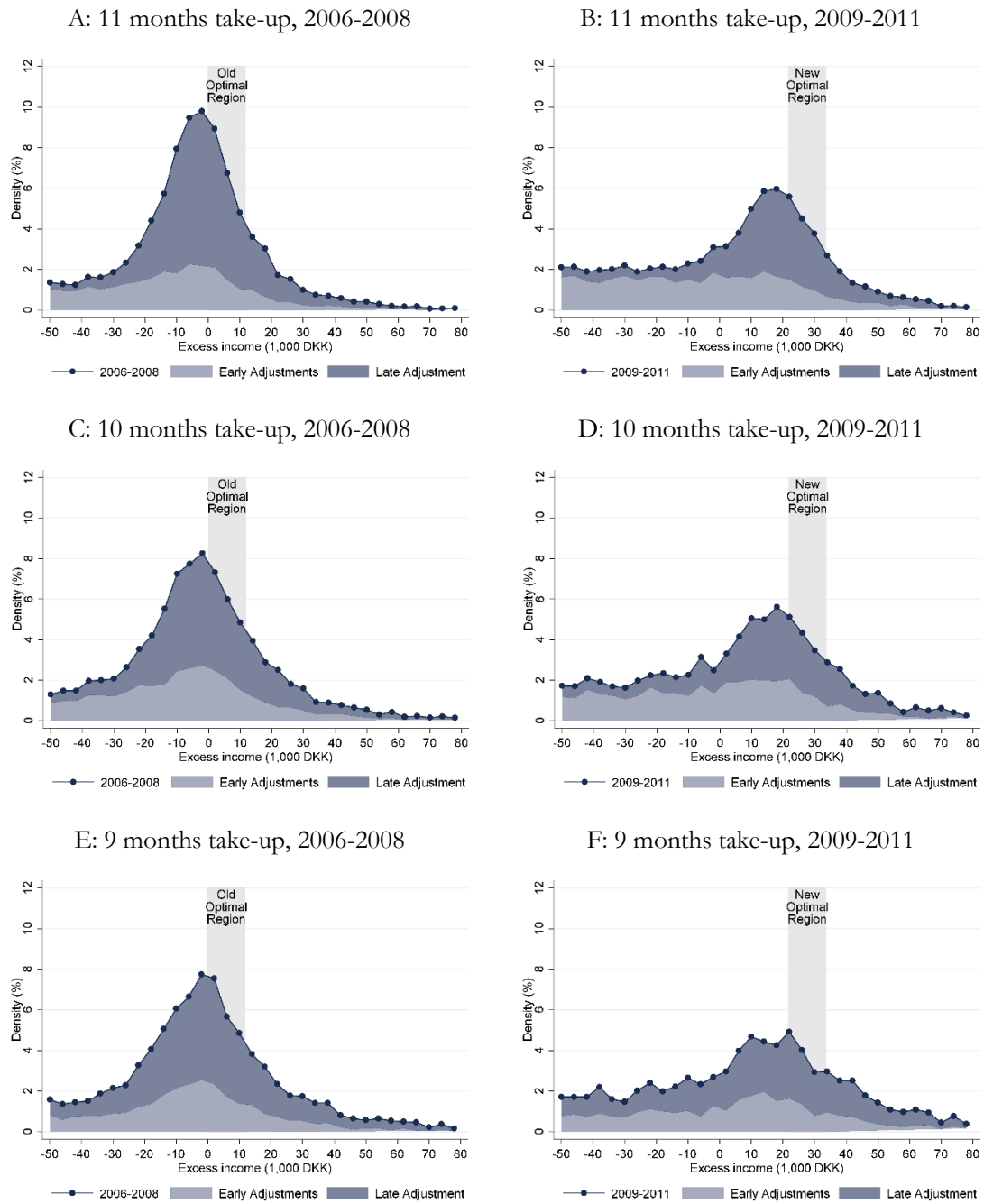
A: Actual distribution
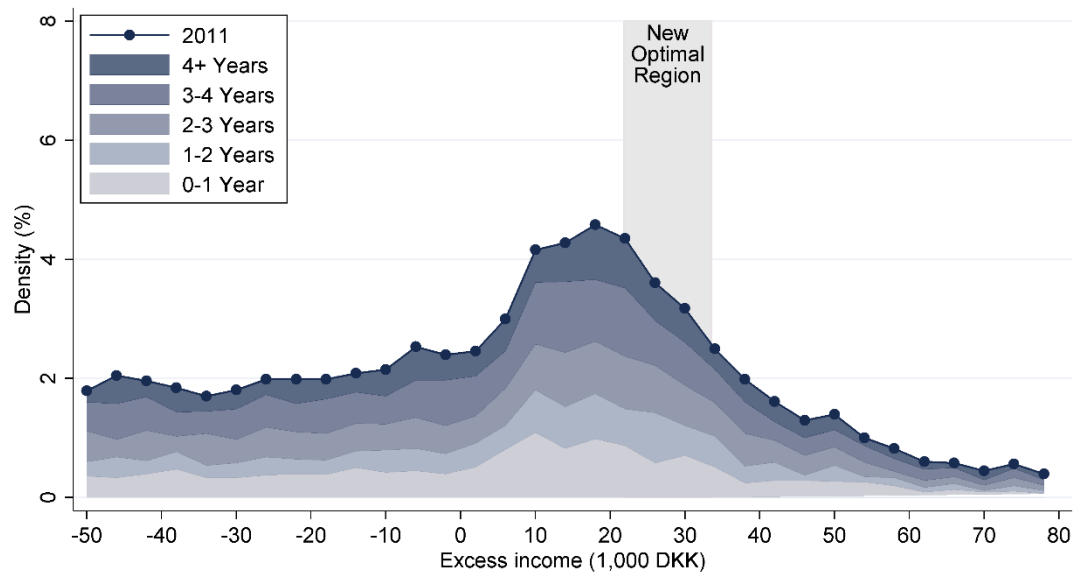


A: Scaled to 100 percent



Notes: The figure shows the distribution of excess income in 2011, similar to Figure 4 in the main paper, split by the amount of benefits taken up over the past 5 years (2006-2010). In panel A, the split adds up to the total density within each DKK 4,000, while panel B shows the split adding up to 100 percent.
DKK 6.4 ≈ USD 1.

**Figure A.XII**
**The distribution of excess income split by time of adjustment and total take-up**

A: 11 months take-up, 2006-2008

B: 11 months take-up, 2009-2011

C: 10 months take-up, 2006-2008

D: 10 months take-up, 2009-2011

E: 9 months take-up, 2006-2008

F: 9 months take-up, 2009-2011



Notes: The figure shows the distribution of excess income in 2011, similar to Figure 4 in the main paper, split by time of adjustment and total take-up. Early adjustments are defined as adjustments taking place from January to August, while late adjustments take place from September to December. DKK 6.4 ≈ USD 1.

## Appendix B: Estimating labor supply elasticities

### Appendix B1: Bunching estimation

In this appendix, I apply the Saez (2010) bunching method to estimate the labor supply elasticity of Danish students. The key element in this approach is the estimation of the excess mass at the kink points, measured as the difference between the observed distribution and the counterfactual distribution absent of kinks. In a standard cross-sectional setting counterfactual distribution would be estimated by fitting a high order polynomial to the observed distribution leaving out a range with observed (sharp) bunching (see e.g. Kleven, 2016). This approach is obviously not feasible in the current setting, where there is no sharp bunching in each of the cross sections.

Instead, I use the shift in the earnings distribution following the 2009 reform to identify the counterfactual distribution. This approach has the advantage of being able to identify excess mass over a much wider range (fuzzy bunching) and hence identify a labor supply elasticity even in the presence of optimization frictions.

More concretely, I construct the counterfactual earnings distribution as the predicted values from an estimation of a 6th order polynomial on both the pre-reform and post-reform earnings distributions (illustrated in Figure 3). In this way, I generally have two density points for each earnings interval (one pre-reform and one post-reform). However, I leave out two regions with visible fuzzy bunching: DKK -28,000 to 6,000 in the pre-reform distribution and DKK -8,000 to 32,000 in the post-reform distribution. This implies that the counterfactual distribution is identified solely from the post-reform distribution for earnings between DKK -28,000 and -8,000 and solely from the pre-reform distribution for earnings between DKK 6,000 and 32,000. The resulting counterfactual distribution is illustrated in Figure B.I and reveals an excess mass of 4.5 percentage points at the pre-reform kink and 3.2 percentage points at the post-reform kink.

**Figure B.I**
**Using the 2009 reform to uncover excess mass and the labor supply elasticity**
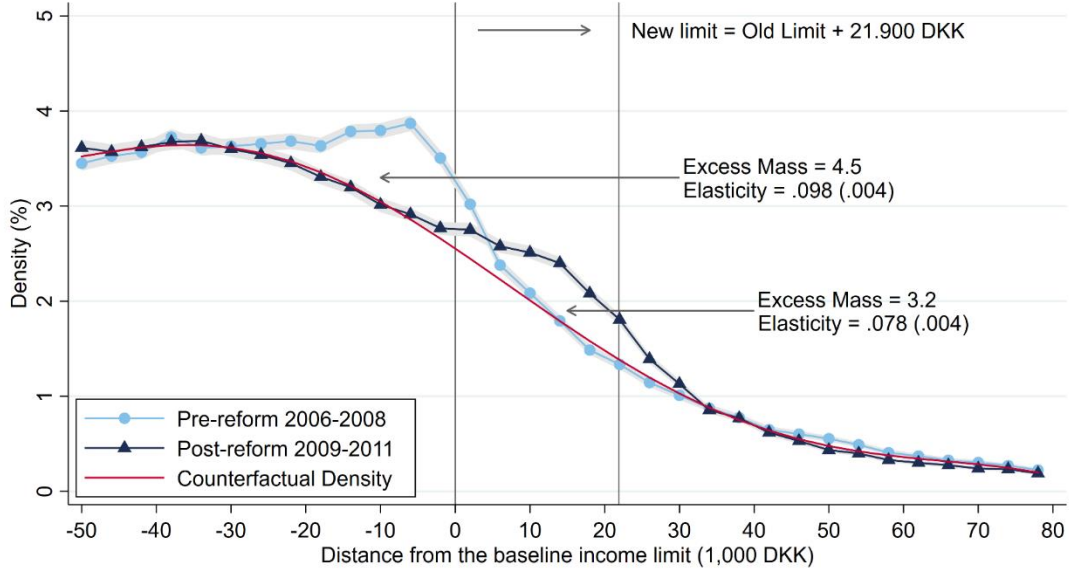


Notes: The figure shows the distributions of earned income for students enrolled in tertiary education, fully eligible for student benefits and with yearly earnings above DKK 8,000. The distribution is not conditional on the actual take-up of benefits. Under the assumption that students always chose the optimal take-up, their effective marginal tax rate jumps from 42 to 72 percent at the initial baseline income limit and from 42 to 72/78 percent at the new income limit as described in section 2 in the main paper. The counterfactual density is estimated as a 6th order polynomial using both the pre- and post-reform distribution except for two regions with visible fuzzy bunching (DKK -28,000 to 6,000 in the pre-reform distribution and DKK -8,000 to 32,000 in the post-reform distribution). The standard errors in brackets and the grey shaded 95 percent confidence intervals are bootstrapped (1,000 replications).
DKK 6.4 ≈ USD 1.

In translating an observed excess mass into a labor supply elasticity, Saez (2010) starts by considering the so-called "marginal bunch'er" defined in the following way. Starting from a standard utility maximization problem

$$\max u(c_i, z_i) = c_i - \frac{n_i}{1 + \frac{1}{\varepsilon}} \left(\frac{z_i}{n_i}\right)^{1+\frac{1}{\varepsilon}}, \tag{B1}$$

subject to the budget constraint

$$c_i = z_i - T(z_i), \tag{B2}$$

where $z_i$ is labor supply, $c_i$ is consumption, $n_i$ is potential earnings and $T(z_i)$ is the tax function, gives the following optimal labor supply

$$z_i^* = n_i \left(1 - T'(z_i)\right)^{\varepsilon}. \tag{B3}$$

Considering a two-bracket tax system with a low marginal tax of $t_1$ below a kink point $K$ and high marginal tax of $t_2 > t_1$ above, the marginal bunch'er is defined as the individual with potential earnings $(n)$ that exactly fulfils

$$z_2^* = n(1 - t_2)^\varepsilon = K \tag{B4}$$

$$z_1^* = n(1 - t_1)^\varepsilon \geq K. \tag{B5}$$

In other words, the marginal bunch'er is the individual who has an indifference curve tangent to the higher marginal tax rate exactly at the kink point. Everybody with counterfactual earnings (without the kink) between $K$ and $z_1^*$ will bunch at the kink point, leaving the following excess mass $(B)$ on this point:

$$B = \int_K^{z_1^*} h(x)dx \approx h(K)(z_1^* - K) = h(K)dz, \tag{B6}$$

where $h(x)$ is the counterfactual earnings distribution. The last part of the equation is an approximation using the counterfactual density at $K$ as the average density between $K$ and $z_1^*$. With this assumption, the labor supply response $(dz)$ of the marginal bunch'er is proportional to the observed excess mass, and the labor supply elasticity can be derived as

$$\frac{z_1^*}{z_2^*} = \frac{n(1-t_1)^\varepsilon}{n(1-t_2)^\varepsilon} \Leftrightarrow \ln\left(1 + \frac{dz}{K}\right) = \varepsilon \ln\left(\frac{1-t_1}{1-t_2}\right) \Leftrightarrow \varepsilon = \frac{\ln\left(1 + \frac{B/h(k)}{K}\right)}{\ln\left(\frac{1-t_1}{1-t_2}\right)} \tag{B7}$$

Based on equation (B7), I arrive at an elasticity of 0.098 at the pre-reform kink and 0.078 at the post-reform kink.

You might expect individuals in study relevant jobs to work partly with the aim of obtaining experience relevant for their future career, in which case they would respond less to a change in contemporary incentives (Best and Kleven, 2013). To investigate this, I split the sample by whether they are employed in a study relevant job. However, as Figure B.II shows, there is no systematic higher elasticity for the individuals in student relevant jobs. The estimated labor supply elasticity for individuals in study relevant jobs is higher at the pre-reform kink point and lower at the post-reform kink point compared to individuals in non-study relevant jobs.

**Figure B.II**
**Labor supply responses depending on the type of job**

A: Study relevant jobs



B: Non-study relevant jobs
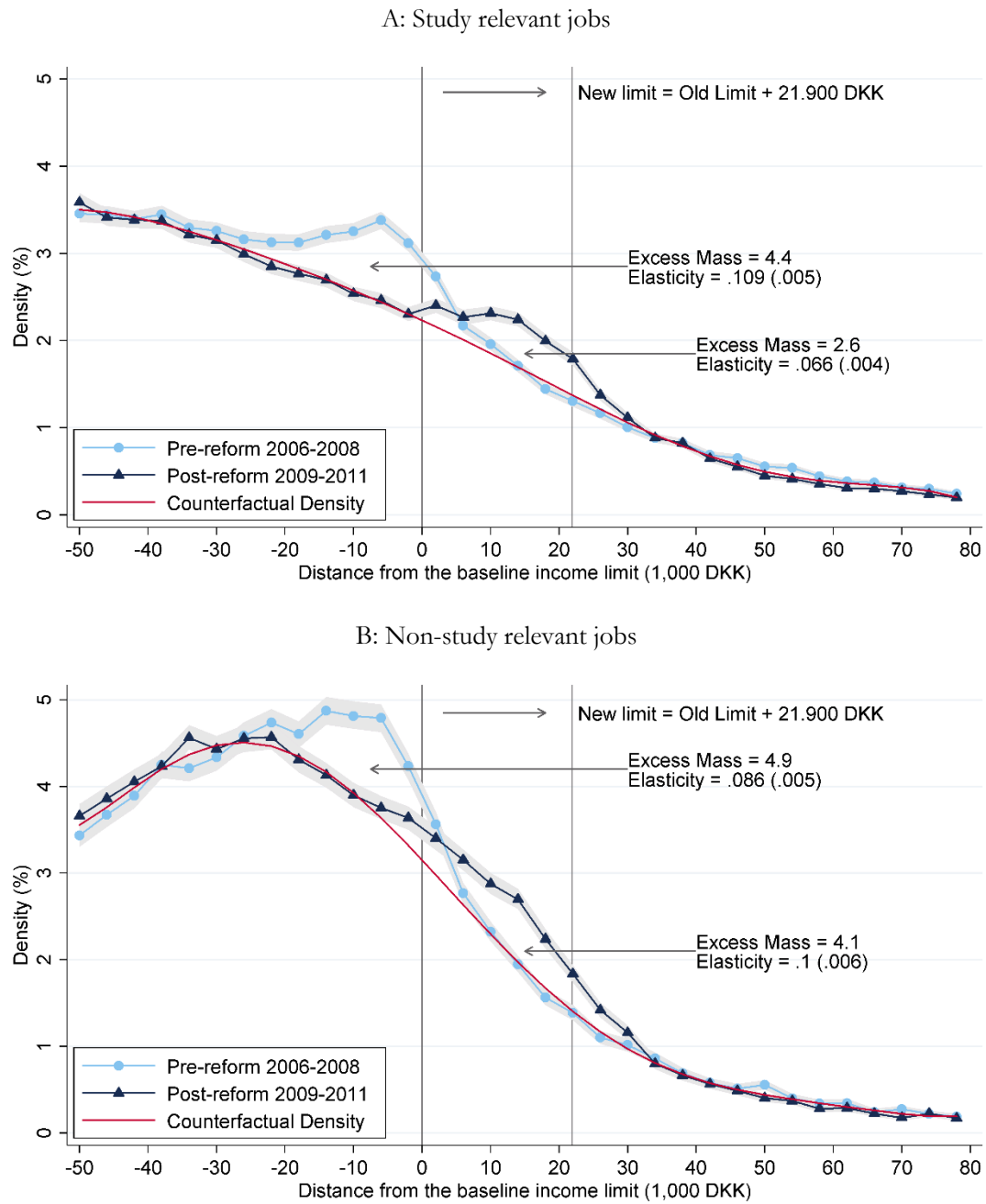


Notes: The figure replicates the bunching estimation in Figure B.I on two sub-samples of individuals split by their job type. Study related jobs are defined as all jobs not in wholesale (NACE 46), retail (NACE 47), postal service (NACE 53), hotels and restaurants (NACE 55, 56), cleaning (NACE 81), elderly, social and childcare (NACE 87, 88), sports and amusement facilities (NACE 93). The standard errors in brackets and the grey shaded 95 percent confidence intervals are bootstrapped (1,000 replications).
DKK 6.4 ≈ USD 1.

It is worth discussing one aspect of the use of the bunching estimator in this setting. Ideally, we want to measure the excess mass relative to the true counterfactual distribution given a constant marginal tax rate of $t_1$. Using the post-reform distribution as a counterfactual for the pre-reform comes close to this ideal setting (except that the fuzzy bunching at the post-reform kink point may affect the distribution at the pre-reform kink as indicated in Figure B.I). Using the pre-reform distribution as the counterfactual for the post-reform is on the other hand theoretically more problematic, as the pre-reform distribution above the kink point is affected by the higher marginal tax rate $(t_2)$.

The use of the pre-reform distribution creates two forms of bias. First, the higher marginal tax rate increases the density. To see this note that equation (B3) implies that

$$F(z(n)) = G(n) \Rightarrow F(z) = G\left(\frac{z}{(1-t)^\varepsilon}\right), \tag{B8}$$

where $F(z)$ is the (observed) cumulative distribution of labor earnings $(z)$ as a function of potential earnings $(n)$ and where $G(n)$ is the underlying cumulative distribution of potential earnings. From equation (B8) we can derive the density function of labor earnings as

$$f(z) = \frac{g(n)}{(1-t)^\varepsilon}, \tag{B9}$$

which, all else equal, is increasing in $t$.

Second, the higher marginal tax rate shifts the earnings distribution to the left, which with downwards density function as here, implies a lower density compared to the counterfactual. The net result of these two biases is ambiguous. If the estimated counterfactual is too high, I underestimate the elasticity and vice versa. For that reason, we should probably put more weight on the elasticity estimated at the pre-reform kink point. Note, however, that the structural estimation in appendix B2 appears to be consistent with the estimated elasticities from the bunching method at both kink points.

Finally, it could be noted that the effect of the higher marginal tax rate above the kink point also creates potential bias even in the standard setting with sharp bunching at a single kink point. In the standard setting, the counterfactual density function is estimated using a smooth polynomial. Thus, the counterfactual density at the kink point will be an average of the density below and above the kink point, where the density above is affected by the higher marginal tax rate as described above. Consequently, the bunching estimator in the standard setting is, to some extent, biased in the same way as my estimation using the post-reform kink point. However, the bias is in most cases likely to be small, particularly in cases with downwards sloping density functions.

## Appendix B2: GMM estimation

As a supplement to the non-parametric estimates of the labor supply elasticity in appendix B1, I present here a structural approach that jointly identifies the labor supply elasticity and a measure of optimization frictions (defined below).

The basic idea behind this estimation strategy is as follows: I build a standard labor supply model augmented with a simple form of optimization frictions and simulate earnings distributions given the institutional setting facing Danish students both before and after the 2009 reform. For each simulation, I compute the sum of squared errors between the observed changes in the earnings distribution following the 2009 reform and the simulated change. Finally, I minimize sum of squared errors by changing the key parameters in the model.

*The model*
I start with a simple quasi-linear utility function

$$u_i = c_i - \frac{\mu n_i}{1 + \mu}\left(\frac{\hat{z}_i}{n_i}\right)^{\frac{1+\mu}{\mu}}, \tag{B8}$$

where $c_i$ is private consumption and $\hat{z}_i$ is the earnings level that the individuals target. $\mu$ and $n_i$ are parameters of the utility function that can be interpreted as the labor supply elasticity and potential earnings. Final end-of-year earnings ($z_i$) are stochastic and given by

$$z_i = \hat{z}_i + \zeta_i, \tag{B9}$$

where $\zeta_i$ is an iid. error term.[23]

The student benefits system creates the budget constraint

$$c_i = (1 - t)(x_i + \min(z_i, L_i)) + tA, \tag{B10}$$

where $x_i$ is the student benefits taken up by the individual, $L_i$ is the individual income limit, $t$ is the tax rate and $A$ is a personal allowance in the tax system. The key feature of the equation is that students will see their benefits phased out at a 100 percent rate for earnings above their income limit.[24]

The individual income limit is given by

$$x_{max} - x_i = q(L_i - \bar{L}) \Leftrightarrow L_i = \tfrac{1}{q}(x_{max} - x_i) + \bar{L} \quad \text{for} \quad x_i < x_{max} \tag{B11}$$

---

[23] I do not model inattention endogenously, but simply assume that individuals cannot observe/re-optimize their earnings during the year. In this way, the estimated variation in final end-of-year earnings should be interpreted as the underlying earnings variance net of re-optimization during the year.
[24] To keep the model simple, I ignore that student benefits are phased out at 50 percent for the first DKK 11,600 above the income limit.

That is, if the student increases his income limit above the baseline limit $(\bar{L})$, his student benefits are phased out at rate $q$.

Assuming that $\zeta_i \sim N(0, \sigma^2)$ with the density function $g(\zeta)$, I can combine equation (B9)–(B11) and derive expected consumption given $x_i$ and $\hat{z}_i$

$$c_i = (1-t)\left(x_i + \hat{z}_i + \min\left(\zeta_i, \frac{1}{q}(x_{max} - x_i) + \bar{L} - \hat{z}_i\right)\right) + tA$$

$$\Rightarrow E(c_i) = (1-t)\left(x_i + \hat{z}_i + \int_{-\infty}^{\infty} \min\left(\zeta_i, \frac{1}{q}(x_{max} - x_i) + \bar{L} - \hat{z}_i\right)g(\zeta)\,d\zeta\right) + tA$$

$$\Leftrightarrow E(c_i) = (1-t)\left(x_i + \hat{z}_i - \sigma(f(\theta) - (1 - F(\theta))\theta)\right) + tA, \qquad (B12)$$

where $\theta = \frac{\frac{1}{q}(x_{max} - x_i) + \bar{L} - \hat{z}_i}{\sigma} = \frac{L_i - \hat{z}_i}{\sigma}$ and $f(\theta)$ is the standard normal density function.

Optimal behavior implies the following two first order conditions for $x_i$ and $\hat{z}_i$.

$$\frac{\partial E(u_i)}{\partial x_i} = 0 \Leftrightarrow \frac{\partial E(c_i)}{\partial x_i} = 0 \Leftrightarrow 1 - F\left(\frac{L_i - \hat{z}_i}{\sigma}\right) = q, \text{ for } x_i < x_{max} \qquad (B13)$$

$$\frac{\partial E(u_i)}{\partial \hat{z}_i} = 0 \Leftrightarrow \frac{\partial E(c_i)}{\partial \hat{z}_i} = \left(\frac{\hat{z}_i}{n_i}\right)^{\frac{1}{\mu}} \Leftrightarrow \hat{z}_i = \left((1-t)F\left(\frac{L_i - \hat{z}_i}{\sigma}\right)\right)^{\mu} n_i \qquad (B14)$$

Both conditions have a straightforward economic interpretation. First, when an individual takes up less than full benefits, she effectively increases her income limit and thereby reduces the risk of having her marginal earnings taxed at 100 percent. The cost of raising the income limit by one unit equals the phase-out rate $q$ and it is therefore optimal to take up less benefits until the probability of hitting the 100 percent tax rate equals $q$.

Second, given the choice of take-up the individual chooses a target earnings (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 100 percent tax rate. The strength of the responses to the effective marginal tax rate depends on labor supply elasticity $(\mu)$.

*GMM estimation*

The model above has three unknown parameters: the labor supply elasticity $(\mu)$, the standard deviation of earnings shocks $(\sigma)$ and the (distribution of) potential earnings $(n_i)$. I estimate the two first parameters by minimizing the sum of squared errors between the actual and simulated change in the earnings distribution to the 2009 reform, while drawing potential earnings from the estimated counterfactual distribution in Figure B.I (corrected for the general tax rate of 42 percent).

More concretely, I preform the estimation as a grid search over $(\mu, \sigma)$ with the following steps. For a given value of $\mu$ and $\sigma$, I simulate the model for a large number of individuals

given both the pre-reform setting ($\bar{L} = 93.0, q = 0.525, t = 0.419, x_{max} = 73.1$) and the post-reform setting ($\bar{L} = 114.9, q = 0.623, t = 0.419, x_{max} = 73.1$). For each simulation, I generate an earnings distribution and compute the change in density at each earnings interval (DKK 4,000 bins). Next, I compute the sum of squared errors of the differences between the simulated changes in densities and the observed change in the densities following the 2009 reform.

In Figure B.III, I show the sum of squared errors of the differences between the simulated changes in densities and the observed changes in the densities following the 2009 reform (the minimization object) for different values of $\mu$ and $\sigma$. The figure shows a positive relationship between the standard deviation of earnings shocks ($\sigma$) and the labor supply elasticity that minimizes the sum of squared errors. In other words, because I only observe fuzzy bunching in the considered labor, a higher labor supply elasticity requires more optimization frictions to spread out the excess mass. The global minimum is at $\mu = 0.0925$ and $\sigma = 7.25$.

In Figure B.IV, I show the resulting simulated earnings distributions compared with the observed distributions before and after the 2009 reform (averages over 2006-2008 and 2009-2011). Panel A compares the changes in the distributions following the 2009 reform. These changes are the moments that have been used in the estimation of the model. Panel B shows pre- and post-reform distributions directly, which besides $\mu$ and $\sigma$ also depend on the distribution of potential earnings taken from the counterfactual estimation in appendix B1.

In general, Figure B.IV indicates a good match between the actual and simulated data. The model slightly underestimates the drop in the density at the pre-reform kink point and overestimates the increase at the post-reform kink point. This is consistent with the results in appendix B1, where the bunching estimator gave an estimate of 0.1 at the pre-reform kink point and 0.08 at the post-reform kink point.

**Figure B.III**
**Minimization object as a function of the elasticity and sigma**
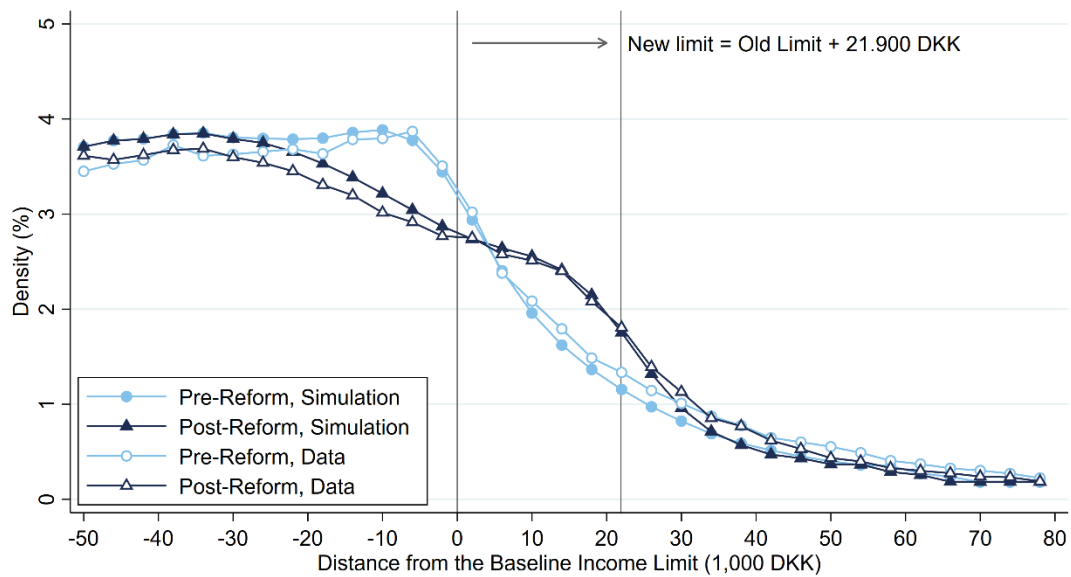
A: Coarse grid



B: Fine grid
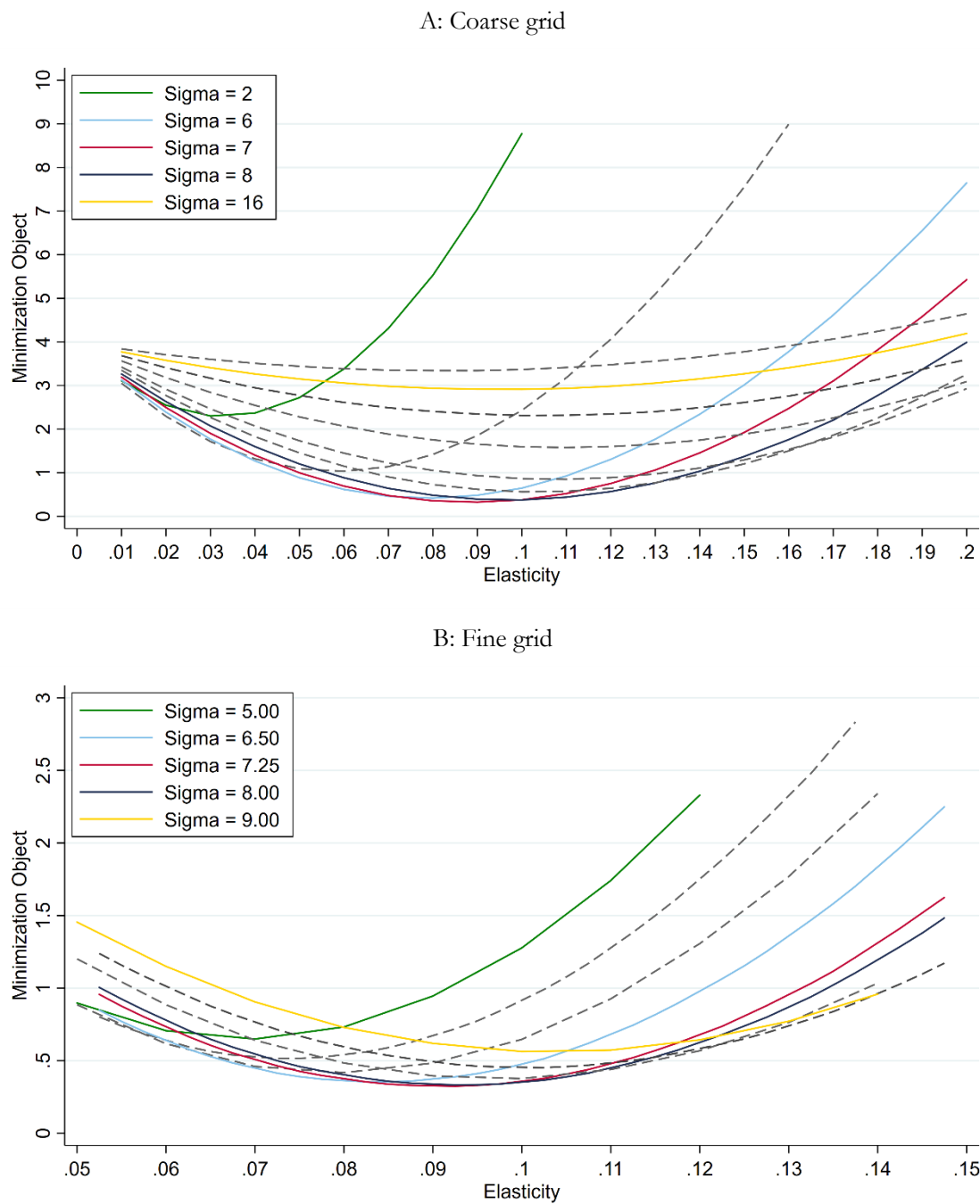


Notes: The figure shows the sum of squared differences between the change in the simulated and observed distributions as functions of the labor supply elasticity ($\mu$) and standard deviation of earnings shocks ($\sigma$).

**Figure B.IV**
**Simulated and observed earnings distributions before and after the 2009 reform**

A: Simulated and actual changes in the earnings distribution
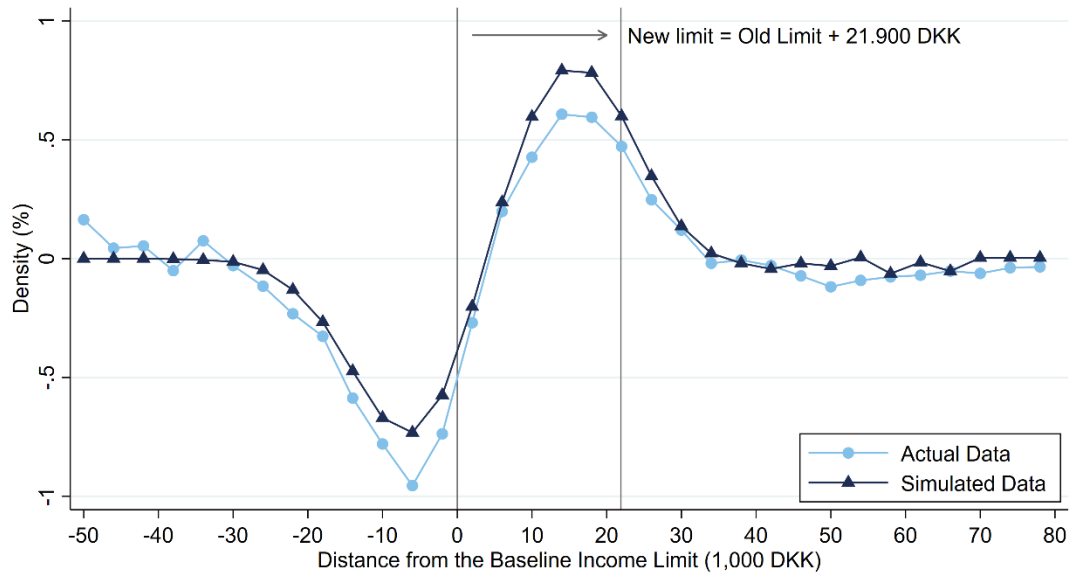


B: Simulated and actual earnings distributions



Notes: The figure compares the observed changes in the distribution of earnings for students enrolled in tertiary education, fully eligible for student benefits and with yearly earnings above DKK 8,000 with the simulated distributions given $\mu = 0.09$ and $\sigma = 7$. These parameters minimize the sum of squared differences between the change in the simulated and observed distributions illustrated in Panel A.

*The position of the excess mass*

As mentioned in the main paper, it might at first glance appear strange that the excess mass uncovered by the shift in the earnings distribution following the 2009 reform is located significantly below the kink points. However, as already seen in Figure B.IV, this is fully consistent with the model set up in this appendix. In the model, the non-centered excess mass is created by the (ex ante) possibility of adjusting benefits. Without this possibility, students would face a jump in the marginal tax rate from 42 to 100 percent and with earnings uncertainty, the expected marginal tax rate would increase smoothly and symmetrically around the kink point from the low tax to the high tax rate, as illustrated in Figure B.V. Such a symmetrically smoothed tax rate would also create an excess mass centered at the kink point.

**Figure B.V**
**Expected marginal tax rate with and without voluntary take-up of benefits**
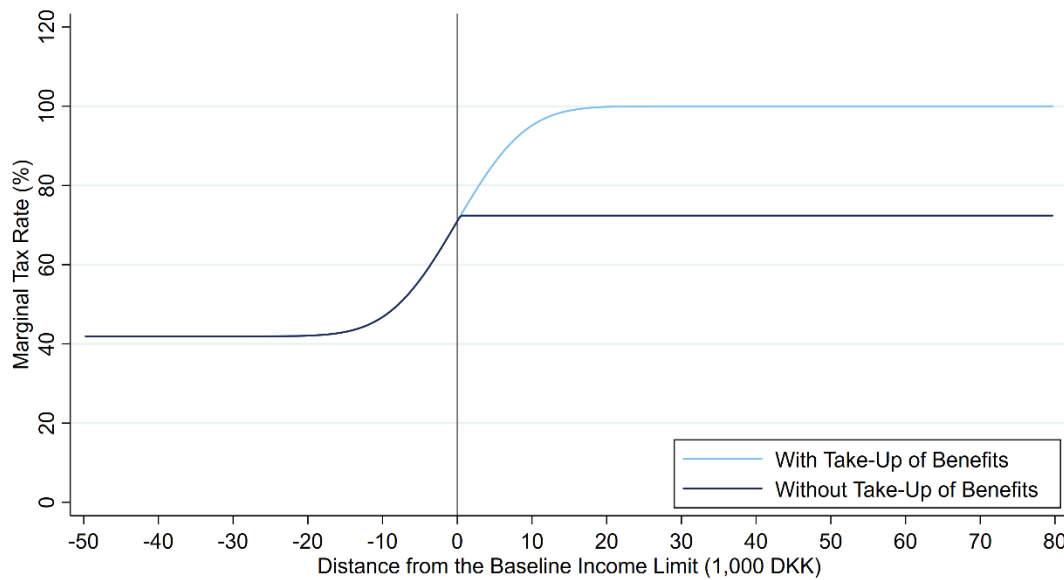


Notes: The figure shows the expected marginal tax rate given by $1 - (1 - t)F\left(\frac{L_i - \hat{z}_i}{\sigma}\right)$, where $L_i = \frac{1}{q}(x_{max} - x_i) + \bar{L}$ is the individual income limit as a function of the target earnings $\hat{z}_i$. In the case without take-up $x_i$ is fixed at $x_{max}$, while $x_i$ is set optimally in accordance with equation (B13) in the case with take-up of benefits.

When individuals have the possibility to adjust their benefits take-up, they will choose to do so as soon as the expected marginal tax rate given their current take-up exceeds the ex ante phase-out rate. This behavior effectively caps the expected marginal tax rate at the ex ante phase-out rate as illustrated in Figure B.V and the expected marginal tax rate is no longer symmetric around the kink point. Consequently, the excess mass is also no longer symmetric around – but shifted below – the kink point.

## Appendix C: Accuracy of the predicted student benefits

As mentioned in the main paper, individual marginal tax and phase-out rates are not directly observed in data. I therefore develop a model that simulates individual phase-out rates, distance to the income limits and payback of student benefits for each individual based on the available administrative data (a model similar to the NBER TAXSIM model for the United States).

The precision of these simulations is important when analyzing optimization frictions as any errors in the calculated incentives risk being interpreted as individuals behaving sub-optimally. Contrary to studies of individual responses to marginal tax rates (such as Gruber and Saez, 2002), where errors in the simulated tax rates create measurement bias towards zero, they would here create a positive bias in the estimated magnitude of frictions.

To check the quality of the simulations, I perform the following validating exercise. For each individual, I compute total earned income based on information from the tax records and the individual income limit based on information on benefits take-up from the Student Benefit Administration. The difference between total earned income and the individual income limit gives individual excess income, which (for positive amount) translates into a predicted payback of benefits. Subtracting the predicted paybacks from the total initial payout of benefits observed in data from the Student Benefit Administration yields predicted end-of-year benefits, which should match the benefits reported in the tax records.

Table C.I shows the distribution of the numerical differences between predicted and final end-of-year benefits. For 98 percent of the sample, I can accurately predict final end-of-year benefits within DKK +/-2, which essentially reflects rounding errors.

**Table C.I**
**Accuracy of the predicted student benefits (all numbers in percent)**

| Error (numerical) | Core sample |
|---|---|
| Less than 2 DKK | 97.9 |
| Between 2 and 100 DKK | 0.3 |
| Larger than 100 DKK | 1.8 |
| Total | 100.0 |

Notes: The table shows the distribution of differences between predicted benefits, based on information on initial payouts and total earned income, and final end-of-year benefits.
DKK 6.4 ≈ USD 1.

The remaining errors can come from two sources. Errors either in total earned income and the individual income limit or in the initial payouts and final end-of-year benefits. The sources are likely independent from each other and I validate the critical parts of my analysis using both my predicted excess income and the excess income implied by the observed differences between the initial payouts and final end-of-year benefits.