Discrete Labor Supply: Empirical Evidence and Implications

Tuomas Kosonen and Tuomas Matikka*

May 30, 2022

Abstract

We provide evidence of discrete labor supply and study the broader implications of the labor supply model. We utilize a novel quasi-experimental setting where a reform shifted an income notch to a higher income level. We find transparent evidence of earnings responses from a wide range in the distribution, which is consistent with discrete but not continuous labor supply. We then illustrate that conventional empirical welfare-loss estimation methods can produce biased results when labor supply is discrete because they are typically based on a continuous labor supply model.

Keywords: discrete labor supply; welfare loss estimation; tax elasticity

JEL Classification Codes: H21, H24, J22

Acknowledgments: We would like to thank Spencer Bastani, Richard Blundell, Raj Chetty, Alexander Gelber, Jarkko Harju, Ohto Kanninen, Patrick Kline, Wojciech Kopczuk, Tomi Kyyrä, Attila Lindner, Tuomas Pekkarinen, Jukka Pirttilä, Jim Poterba, Daniel Reck, Emmanuel Saez, Matti Sarvimäki, Joel Slemrod, Marko Terviö, Mazhar Waseem, Danny Yagan and Gabriel Zucman as well as participants in numerous conferences and seminars for their helpful comments and discussions. Funding from the Academy of Finland (No. 299373, No. 303689, No. 322074) is gratefully acknowledged.

*Kosonen: VATT Institute for Economic Research (Helsinki, Finland), Academy of Finland and CESifo, tuomas.kosonen@vatt.fi. Matikka: VATT, Academy of Finland and CESifo, tuomas.matikka@vatt.fi.

1 Introduction

Analyzing labor supply is one of the most important topics in labor economics, public finance and macroeconomics. Textbook models often assume that labor supply can be adjusted flexibly and continuously in response to shocks or policy, but researchers have debated for quite some time between continuous and discrete labor supply (see e.g. Rosen 1976, Altonji and Paxon 1988, Dickens and Lundberg 1993, van Soest 1995, Blundell *et al.* 2008, Kreiner *et al.* 2015). Despite this discussion, there has been limited analysis on how the choice of the labor supply model affects empirical analysis and welfare loss estimates. The literature analyzing earnings responses to taxes mostly assumes continuous labor supply underlying the analysis, even when acknowledging various types of optimization frictions (Feldstein 1999, Saez *et al.* 2012, Chetty 2012, Kleven and Waseem 2013). However, this assumption seems contrary to many situations occurring in real-life labor markets. For example, switching to another job or getting a promotion or raise can lead to a discrete change in annual earnings even conditional on participating in the labor market. These examples appear to be particularly relevant for regular wage earners who constitute the bulk of taxpayers.

This paper provides novel empirical evidence supporting the discrete labor supply model and studies the implications of labor supply models on empirical welfare loss analysis. In our empirical analysis, we utilize a reform that shifted the location of an income notch and examine the changes in the whole income distribution caused by the reform. Our main finding is that earnings increase from a wide range in the income distribution below the original location of the notch, which is only consistent with discrete labor supply. We also provide divided sample and descriptive evidence on changes in annual earnings supporting the discrete labor supply model. We then discuss the important and previously under-explored consequences of discrete labor supply for empirical welfare loss analysis.

We make a number of contributions to the literature. We contribute to the long-standing debate between structural and quasi-experimental literature over the labor supply model mentioned above.[1] One reason for the ongoing debate could be the lack of other than descriptive evidence in favor of either of the labor supply models. We provide novel and clear quasi-experimental evidence in favor of a discrete labor supply model.[2]

We show that changing labor supply model from continuous to discrete has larger implications for the empirical welfare loss analysis than, for example, optimization frictions assumed in continuous labor supply context. First, the reduced-form methods aiming to estimate the elasticity of taxable income (ETI), such as differences-in-differences (Gruber and Saez 2002, Weber 2014) and bunching (Saez 2010, Kleven and Waseem 2013), can produce downward-biased estimates when labor supply is discrete instead of continuous. This is because these estimators are implicitly based on the continuous labor supply model, which implies that the effects of tax reforms remain local. In the empirical applications a counterfactual is often derived using those in the same distribution as a control group. However, with widespread effects in the distribution some individuals in the control group can also be affected by the reform, creating a bias in the empirical estimate.

Second, discrete labor supply can affect the sufficient statistic framework used to empirically assess the welfare losses of taxes. Empirical applications often rely on the average ETI estimate as a sufficient statistic for welfare analysis (see e.g. Chetty 2009a, Kleven 2020). Our analysis illustrates that the extent to which labor supply is discrete affects the income range in the distribution where income changes occur. This implies that the average ETI alone does not necessarily capture the full income response needed to evaluate the welfare loss, and the sufficient statistics approximation needs to be augmented with the number of people responding to a tax reform when labor supply is discrete. This notion favors distributional methods to capture the total earnings impact of tax

[1] Some of the quasi-experimental literature does allow for optimization frictions that hinder responses to taxes, including job search costs, optimization errors and inattention to taxes (see e.g Chetty *et al.* 2011, Kleven and Waseem 2013, Søgaard 2019 and Gelber *et al.* 2020). Note that some optimization frictions, such as fixed costs for job search analyzed in Chetty *et al.* (2011), could lead to discrete changes in earnings, but this literature has not studied in detail how such an assumption could lead to different labor supply models or the implications for empirical welfare analysis.

[2] However, in our analysis the more precise mechanism behind discrete labor supply remains open, such as a sparse menu of job offers or fixed job switching costs.

rate changes[3], and in this paper we propose a new method to capture these effects using a moving income threshold design.

Turning back to our empirical analysis, we utilize a combination of an income notch, i.e. a jump in the average tax rate creating strong local tax incentives, and a reform that shifted out the location of the notch. The institutional setting involves a monthly study subsidy for Finnish higher education students. A university student loses eligibility for one month of the subsidy (approximately 500 euros) if her annual earnings exceed a predetermined gross income threshold (9200 euros before 2008), causing a sharp drop in disposable income. Due to this feature, this kind of income threshold is called a notch. In 2008, the location of the notch was increased by 30%, allowing students to earn more income before they lose the subsidy they are eligible for. As Finnish university students typically participate in flexible part-time labor markets during their studies, the notch and the reform create a relevant change in incentives affecting the labor supply and earnings choices for a majority of students. Moreover, students face similar labor market institutions and regulations as other workers, such as employment contract requirements and minimum wage and working hours legislation.

We examine the impact of the reform on the whole income distribution rather than just within a narrow window around the notch. This allows us to observe whether the notch affected earnings from a wider income range, especially below the notch, which is revealing of different labor supply mechanisms. To capture a causal impact that accounts for changes in different parts of the distribution, we develop a new quasi-experimental method that resembles those proposed in recent literature (see e.g. Athey and Imbens 2006, Firpo *et al.* 2009, Firpo and Pinto 2016, Cengiz *et al.* 2019). We calculate the change in the relative density of the students' income distribution at the time of the reform and contrast it to changes in the distribution of a control group consisting of non-student young wage earners who are similar to students in their labor market attachment and earnings. Identification requires that the earnings distributions of treatment and

[3]For example, the changes-in-changes methods proposed by Athey and Imbens (2006) and recently applied, for example, by Kottelenberg and Lehrer (2017), methods to capture program effects in income distribution by e.g. Firpo and Pinto (2016), or changes by bins of the wage distribution as proposed by Cengiz *et al.* (2019).

control groups evolve similarly over time in the absence of treatment. We evaluate this assumption and find that the distributions of both groups remained practically constant over years both before and after the reform, which supports our identifying assumption.

As our main empirical result we find that the reform caused distinctive income changes among students from a broad income range, especially from below the original location of the notch. Thus, the overall shape of the earnings distribution changed due to the reform, evident in that the density at lower incomes reduced and at higher incomes increased. In contrast, we find no discernible changes in the income distribution of the control group, reducing the worry that the changes in the treatment group would be caused by e.g. local labor market shocks. Furthermore, our panel data evidence shows that large upward individual-level jumps in earnings were much more common for students the year following the reform compared to the period before it. These results suggest that the cross-sectional changes arise because of individual-level discrete responses and not, for example, because of participation of more productive students in the labor market after the reform.

We contrast our empirical results with predictions from discrete and continuous labor supply models and show that between these two models our results are only consistent with a discrete model. In a stylized discrete labor supply model, labor supply consists of discrete jumps where an individual considers whether or not to switch from one location to another. Utility maximizing individual for whom a location even well above the original location of the notch becomes more attractive after the reform will respond by switching to this new location.[4] With heterogeneity in discreteness, only a fraction of individuals respond and to a varying extent, leading to an overall smooth shift in the earnings distribution. This is exactly what we observe in our empirical analysis.

Instead, in a continuous model individuals well below the original notch are unaffected by the reform. In this model individuals only consider changes that affect their utility or budget constraint at the currently chosen optimal earnings level. A reform that takes place at a significantly higher income level affects neither of these elements. Note that also

[4]Note that this argument holds independently of the more precise mechanism that causes labor supply to be discrete, such as a sparse menu of available jobs or fixed job search costs.

in the continuous model those that were bunching right below the notch can relocate to a higher income level, but not those located more significantly below the notch. We present simulation results further illustrating that we need a discrete component in the model to even qualitatively match the distribution-wide changes we observe. Also, we show that augmenting the continuous model with optimization frictions that merely attenuate responses to taxes from prior literature cannot produce similar results as our empirical findings.

We find further empirical support for discrete labor supply by examining the responses of two specific subgroups of students: those who work in plausibly more discrete labor markets (public sector, or research, manufacturing and construction in the private sector) and those working in less discrete labor markets (restaurants, bars and cafes, hotels, cleaning and security services). The latter group faces arguably less discrete labor markets because they typically have more flexible working hours and are more likely subject to hourly rather than monthly wages compared to the first group. We find a significantly larger shift in the earnings distribution for students working in the more discrete labor markets, providing additional suggestive evidence in favor of a discrete model. Moreover, this finding suggests that discreteness stems from the functioning of the labor markets, a feature that is not specifically related to students. When looking at the labor force as a whole, we observe that a significant share of regular wage earners in Finland are employed in industries that we classified as more discrete above (61% in our data), suggesting that discrete labor supply can induce a relevant constraint for a large share of the labor force. We also find supporting descriptive evidence for the discrete model from the fact that earnings in general change quite often by large amounts from year to year among the whole Finnish wage earner population.

We turn to implications of discrete vs. continuous models on the methods aiming to estimate ETI and the of welfare loss of taxes. We use a simulation model calibrated to our empirical data that allows us to precisely track down the impact of various labor supply assumptions on the responses to different types of tax reforms. To illustrate how the labor supply model affects different ETI estimators, we simulate a tax reform that reduces the marginal tax rate for higher incomes, resembling many actual tax reforms. Our baseline estimator follows a differences-in-differences method that many previous studies have used to estimate ETI (see Gruber and Saez 2002, Saez *et al.* 2012, and Neisser 2021). This approach uses a control group consisting of individuals at lower income levels than to which the tax rate cut is applied. We also estimate a mobility elasticity which is consistent with the discrete labor supply framework (see Saez 2002 and Kreiner *et al.* 2015) for the whole distribution using a pure control group unaffected by the tax reform.

Our results show that the baseline estimator produces a downward biased ETI estimate when labor supply is discrete. This is because in the discrete model the whole distribution potentially responds to the reform, including the control group in this application. The standard ETI estimate in our baseline simulation model is 0.07 while the mobility estimate is 0.16, indicating a substantial bias. Similar bias can arise in the increasingly popular bunching method (Saez 2010, Kleven and Waseem 2013) that has already been criticized for various reasons (see e.g. Blomquist *et al.* 2019). Using the bunching method to estimate ETI essentially involves assuming the continuous model. The bunching method utilizes the estimated excess mass just around a notch or kink as evidence of behavioral responses to taxes, thus ignoring any effects further away in the distribution. Also, the counterfactual density that is typically estimated from the surrounding distribution is not valid if that part of the distribution is also affected by the income threshold. To demonstrate the extent of this bias, we estimate a mobility elasticity of 0.18 for students, which is approximately 2.5 times larger than a simple bunching ETI estimate of 0.07.

Finally, discrete labor supply can affect the empirical approximation of aggregate welfare losses created by taxes. The disutility from taxes captured by the average ETI governs to what extent individuals respond to taxes, and approximates the welfare loss in the standard continuous sufficient statistic framework (see e.g. Chetty 2009a, Kleven 2020). However, discreteness in labor supply expands the number of individuals affected by local tax changes depending on the extent of discreteness. Therefore, it is not necessarily sufficient to approximate welfare losses by solely estimating the average ETI using local tax reforms, as we need to also understand to what extent the overall distribution responds. We illustrate this point using the simulation model and show that an unbiased average elasticity can be much lower when labor supply is discrete instead of continuous, but when accounting for all income responses in the distribution the aggregate changes in income can be almost equal between the models.

Another illustration of the aggregate welfare loss is from comparing our results with those of Søgaard (2019). He estimates the income responses among Danish higher education students to a reform that shifted a kink in the study subsidy system. The reform caused a visibly more modest response in the income distribution than what we find, quite likely reflecting the fact that a kink is much weaker incentive than a notch, which we analyze. If one would evaluate the welfare losses caused by the two reforms solely by comparing elasticities, the calculation would miss the fact that aggregate welfare loss is much larger in the Finnish case, which is evident when accounting for the much larger change throughout the income distribution in the Finnish case.

This paper proceeds as follows: Section 2 presents the relevant institutions and empirical methods. Section 3 presents the empirical results. In Section 4 we discuss the theoretical mechanisms, and Section 5 presents our simulation results and discusses the broader implications of discrete labor supply. Section 6 concludes.

2 Institutions, data and empirical methods

2.1 Study subsidy for university students

In Finland, all students who are enrolled in a university or polytechnic can apply for a monthly study subsidy, administered nationally by the Social Insurance Institution of Finland (hereafter SII). The subsidy is intended to enhance equal opportunities in acquiring higher education, and to provide income support for students who often have low disposable incomes. In Finland, university education is publicly provided and there are no tuition fees. A large proportion of individuals receive higher education in Finland (approximately 40% of individuals aged 25-34 have a degree), and the study subsidy program is widely used among students.

The maximum amount of the subsidy was 461 euros per month in 2007. The default number of subsidy months per year is 9, which is provided if a student does not actively apply for a different number of subsidy months, and which a large proportion of students also receive. The eligibility for the study subsidy depends on personal annual gross income (labor income + capital income), and an academic criterion of completing a certain predefined number of credit points per academic year. Parental income or wealth do not affect eligibility nor the amount of the benefit for students not living with their parents.[5]

The discontinuity in labor supply incentives is created by an income threshold. If a student has annual gross income higher than a predetermined threshold, one month's subsidy is reclaimed by the SII. This results in an increase in the effective average tax rate, or an increase in the implied marginal tax rate of over 100%, in a region just above the threshold, creating a *notch* in the budget set of students. With 9 subsidy months the income threshold was 9260 euros in 2007. An additional month of the subsidy was reclaimed for an additional 1010 euros of income above the threshold. This implies that the schedule ultimately comprises multiple notches in an income range above the first notch. Students can deviate from the default of 9 months and alter the number of subsidy months by application, or by returning already granted subsidies by the end of March in the next calendar year. Having more study subsidy months reduces the income threshold, and vice versa.[6]

The study subsidy program was reformed in 2008. In the reform the income threshold was increased by approximately 30%. For a typical student with 9 study subsidy months, the annual income threshold increased from 9260 to 12,070 euros. In addition,

[5] The full study subsidy includes a study grant and a housing benefit. The standard study grant was 259 euros/month and the maximum housing benefit 202 euros/month in the academic year 2006/2007. Housing benefits are granted only for rental apartments, and the housing allowance is reduced if spousal gross income exceeded 15,200 euros per year (in 2007). In addition to the study subsidy, students can apply for repayable student loans secured by central government.

[6] In 2007, the formula for the annual income threshold was the following: 505 euros per study subsidy month and 1515 euros per month without the study subsidy, plus a fixed amount of 170 euros.

the monthly study subsidy was increased slightly from 461 to 500 euros per month. Other details of the system were not changed, including the academic criterion related to the required number of academic credits.[7]

Figure 1 illustrates the study subsidy schedule before and after 2008 for a student who collects the default 9 subsidy months. The figure shows that students face large local incentives not to exceed the first income threshold because of the initial threshold and multiple similar thresholds after that. Once an income threshold is exceeded, losing one month of the subsidy causes a significant drop in disposable income, and thus a dominated region right above the threshold. The figure underlines the distinctive change in incentives caused by the increase in the location of the income thresholds in 2008, highlighting that the reform encouraged students to increase their earnings above the old income threshold. Finally, Table A1 in Appendix A shows the income thresholds in numbers before and after 2008, and presents the relative loss in disposable income incurred if the income threshold was exceeded.

2.2 Data and descriptive statistics

Although the majority of students have access to the study subsidy and repayable student loans, most university students in Finland also work part-time during their studies within and between semesters. Therefore, the means-testing of the study subsidy creates a real constraint affecting the labor supply choices of a majority of students. In our analysis, we use panel data on all working-age individuals (15–70 years) living in Finland in 1999–2013, provided by Statistics Finland. These data include a rich set of register-based information, such as tax and social benefit registers and information on the study subsidy program. With these data, we can analyze responses to the incentives created by the program and learn how various individual characteristics affect labor supply responses.

Table 1 shows the descriptive statistics for a pooled sample of students in 1999–2013. In our analysis, we drop first-year students and students who graduate within a given

[7] As with the old regime, an additional month of the subsidy is reclaimed after an additional 1310 euros of gross income above the threshold. After 2008, the formula for the threshold was the following: 660 euros per study subsidy month and 1970 euros per month when no study subsidies are collected, plus a fixed amount of 220 euros.

year in order to avoid the effects of potential income shocks before enrollment and after graduation. However, dropping first-year students and graduates does not affect the main results in a meaningful way. Mean annual labor income among our sample of students is 8446 euros. We observe that on average 80% of students earned at least 500 euros of labor income in a year. In addition, only 55% of students received labor earnings from only one employer, indicating that students tend to work in different types of jobs during a year. These observations indicate that many students work in part-time or temporary jobs during their studies and breaks between semesters in order to increase their disposable income and/or to gain work experience while studying. In terms of sectors, 19% of students work in manufacturing (including construction), 16% in hospitality services such as hotels and restaurants, 39% in administrative and support services or in the public sector, and 25% in other sectors including those whose sector cannot be identified in the data. In terms of study fields, 17% of students in our sample study arts and humanities, 19% business and social sciences, 34% technology or health and social services and 29% in other fields including those whose field of study cannot be identified.

In our baseline analysis, we focus on students who received 9 months of study subsidy before and after 2008. For this group, the income threshold increased from 9260 to 12,020 euros. The advantage we gain by fixing the number of subsidy months is that we can isolate the effect of the change in the location of the threshold on the earnings distribution. This restriction is, however, not very selective as a large proportion of students receive 9 months of the study subsidy, partly because it is the default choice and partly because it presumably creates a good balance between subsidies and labor earnings for many students. The share of students receiving the default subsidy months is 42.4%. Furthermore, students who receive the default subsidy are similar to the overall student population, except that they are on average slightly younger (22.4) and have less labor income (5633 euros) than all students. Nevertheless, we test the robustness of our results by including students who deviate from the default subsidy choice.

2.3 Estimation method

The 2008 reform that shifted the location of the notch creates a unique empirical set-up to study earnings responses to a distinctive and salient change in tax incentives. We are particularly interested in investigating whether local tax incentives, such as notches or kinks, affect income distributions in a wide income range rather than just close to the threshold. Thus, we examine the changes in the whole income distribution at the time of the reform which enables us to test whether labor supply is discrete or continuous, as discussed in detail in Section 4.

Recent literature often uses a bunching method to estimate responses to a local tax discontinuity by relating an excess mass in the distribution just below the notch to an estimated counterfactual (see Kleven 2016 for a survey). Standard bunching method is presented in detail in Appendix B, and graph (a) in Figure 2 illustrates the method graphically. However, we do not apply the bunching method in our main analysis for two reasons: it produces downward-biased earnings response estimates if the notch affects the earnings distribution more broadly than just around the local discontinuity, and the surrounding density outside the bunching region cannot be used to estimate a credible counterfactual when that part of the distribution is also affected by the notch.

Instead of using the bunching method, we develop a new method in the spirit of differences-in-differences (DiD) and changes-in-changes (CiC) methods to estimate the distribution-wide responses to a reform that shifts the location of an income notch. Our method is similar to that used in Cengiz *et al.* (2019) who estimate the localized effects of minimum wages by income bins of the wage distribution, but our focus is to show more explicitly what happens to the shape of the overall distribution. We estimate a counterfactual change in the distribution utilizing the income distribution of a control group, similarly as in the DiD method. We estimate to what extent the whole earnings distribution is affected by the reform relative to this counterfactual. We measure the distributions relative to the total number of students in order to account for potential changes in the total number of students across years.[8]

[8]Note that in the bunching method, using relative distributions and frequency distributions produces

More formally, we first measure the change in the students distribution as follows:

$$\hat{b}(z) = \frac{\sum_{i=z_L}^{z_H} \left[(c_j^B/N^B) - (c_j^A/N^A)\right]}{\sum_{i=z_L}^{z_H}(c_j^A/N^A)/N_j} \tag{1}$$

where c_j is the count of individuals in an income bin j, and z_j denotes the income level in bin j. $\sum_{i=z_L}^{z_H}(c_j^B/N^B)$ is the share of students within a fixed income range $[z_L, z_H]$ relative to the total number of students in the distribution (N^B) before the 2008 reform, and $\sum_{i=z_L}^{z_H}(c_j^A/N^A)$ after the reform in the same income range. N_j denotes the number of bins within $[z_L, z_H]$. In the estimation, we set the lower limit z_L to zero and measure changes in the whole distribution below the old income threshold by setting z_H equal to the old income threshold (9260 euros). Graph (b) in Figure 2 graphically illustrates the estimation approach.

To complete our method, we utilize a control group to take into account potential changes in the earnings distribution for reasons other than the change in the study subsidy system, such as changes in the economic environment affecting the labor markets where students work. In our empirical application we use young part-time workers who are not higher education students as the control group. Those in the control group are not eligible for the study subsidy and thus not subject to the income threshold or the reform, but are of the same age as the students and work essentially in the same labor markets and in similar jobs. As a result, the control and treatment groups are similar in their labor market characteristics such as labor earnings, as described in Table 2.[9]

More formally, our method to estimate the change in the distribution caused by the reform calculates the change in the density in the treatment group between the two time periods as presented in equation (1), and subtracts from this the change in the control group over the same period:

identical estimates.

[9]The control group is selected to roughly match students' job and age characteristics. Students typically work in part-time jobs and they tend to be young. Thus the control group comprises individuals who we observe to have less than 12 working months per year, and who are 19–24 years old. This age interval matches the 25–75 percentile points of the students' age distribution. Our results are not sensitive to small changes in the composition of the control group.

$$\hat{b}_d(z) = \left[\frac{\sum_{i=z_L}^{z_H} \left[(c_j^B/N^B) - (c_j^A/N^A) \right]}{\sum_{i=z_L}^{z_H} (c_j^A/N^A)/N_j} \right]^S - \left[\frac{\sum_{i=z_L}^{z_H} \left[(c_j^B/N^B) - (c_j^A/N^A) \right]}{\sum_{i=z_L}^{z_H} (c_j^A/N^A)/N_j} \right]^P \quad (2)$$

where superscript S denotes students (treatment group) and P non-student part-time workers (control group). This estimate summarizes the broader change in the earnings distribution of students caused by the reform while taking into account any other potential changes in the labor market environment of part-time workers.[10]

Our identification assumption is that the changes in the earnings distribution of the control group reflect the changes in the treatment group in the absence of the reform. More precisely, we need to assume that the relative distributions would evolve similarly over time in the treatment and control groups in the absence of the reform. This resembles the parallel trends assumption familiar from the DiD approach. Note that as we estimate the changes in densities separately for the treatment and control groups, our approach is not as sensitive to linear functional form assumptions as the regression version of the DiD. Also, our identification assumptions are similar to those in the CiC approach (see e.g. Athey and Imbens 2006), but we estimate the overall change in the density due to the reform, in comparison to identifying changes at each quantile as in the CiC approach. Moreover, our approach is relatively straightforward to apply. We empirically evaluate the assumption that the distributions evolve similarly over time in both groups below by examining the development of the distributions before and after the 2008 reform.

3 Main results

We begin by presenting the earnings distributions of the treatment and control groups over time around the reform in 2008. Figure 3 shows the labor earnings distributions

[10]Following the bunching literature, the standard errors for $\hat{b}_d(z)$ are calculated using a residual-based bootstrap procedure (Kleven and Waseem, 2013). First, we fit a flexible polynomial function to both the pre- and post-reform relative earnings distributions of students and other young part-time workers. We then generate a large number of new estimates for the distributions by randomly re-sampling the residuals from these regressions (with replacement). The standard error is defined as the standard deviation of $\hat{b}_d(z)$ based on the bootstrapped distributions.

of the two groups before and after the reform for those with positive earnings within an income range of 0–18,000 euros in 2006–2007 and 2008–2009, denoting the pooled pre- and post-reform periods, respectively.[11]

Remarkably, the figure shows that the earnings distribution of students has a clearly different shape after the reform than before it; the earnings of students have increased over a wide income range. Especially intriguing is that the shifting of the earnings distribution to the right occurs from far below the old location of the income threshold. Contrary to students, the earnings distribution of the control group remained practically constant between 2006–2007 and 2008–2009. The earnings in the control group originate from the same labor markets as those in the treatment group. This suggests that the shifting of the earnings distribution cannot be explained by other contemporary changes in the local labor markets affecting the earnings development of both the control and treatment groups. We further discuss this below.

To quantify the changes in the distribution due to the reform, we estimate equation (2) that produces as its outcome the change in the relative density of students subtracted by the change in the density of the control group. The estimation is performed within an income range of 0–9260 euros, thus including the whole distribution below the old income threshold. The estimate is large (9.809 with a standard error of *1.01*), suggesting that the magnitude of the change in the overall earnings distribution is economically and statistically significant. This estimate is approximately three times larger than the standard bunching estimate, 2.931 (*0.875*), estimated following the methods of Kleven and Waseem (2013) within an income range just below the threshold (8100–9260 euros) before the reform.[12] In order to further characterize the general magnitude of the response, we estimate that earnings increased on average by 550 euros when accounting for the overall changes in the shape of the earnings distribution, which corresponds roughly to a 10% average annual increase in the labor earnings of students.

In order to ensure that the above estimates represent causal responses to the reform,

[11]The figure includes only labor earnings and not all income to which the income threshold applies because receiving capital income is very rare among students and other young part-time workers.

[12]Bunching results are discussed in more detail in Appendix B.

we need to assume that the distributions of the treatment and control groups evolve similarly over time in the absence of the reform, as discussed in Section 2.3. As a first check to this end, Figure 4 plots students' and other young part-time workers' earnings distributions over a longer time period of four years before and after the reform of 2008. The figure shows that the change in the earnings distribution of students occurred exactly at the time of the relocation of the income threshold, indicating that any gradual shifting of the distribution does not explain the observed pattern. Furthermore, the distribution of the control group remained almost unchanged throughout this period, therefore strongly supporting our identification assumptions. In more detail, both the treatment and control group distributions exhibit very similar minor changes at the bottom from 2004–2005 to 2006–2007, further strengthening the case that the distributions have developed very similarly over time in the absence of the reform.

Our second robustness check concerns potential changes in the characteristics of students over time. Figure A1 in Appendix A shows the distributions in 2006–2007 and 2008–2009 when we use bin-level inverse probability weighting to re-weight the distribution in the latter period to match the pre-reform characteristics of students in terms of age, field of study and industry of the firm they work for. Re-weighting does not change the outcomes in a significant manner, indicating that potential changes in the characteristics of the student population are not likely to explain the results.

Furthermore, one might be concerned that students can also respond to the reform by changing the number of study subsidy months they choose. First, Figure A2 in Appendix A shows the earnings distributions in 2006–2007 and 2008–2009 when including students with other than the default 9 study subsidy months. The figure illustrates that the broader changes in the distribution are clearly visible when including this group of students, implying that labor supply responses rather than the choice of study subsidy months are driving our main results. Second, we find no significant changes in the distribution of subsidy months over time or associated with the reform, and 9 months is the most typical choice in all of the years around the reform, as illustrated in Figure A3 in Appendix A. This indicates that current students responded to the reform by changing their earnings, but not, on average, by claiming more or less subsidies per year.

In addition, a closer examination of the students' earnings distribution in Figure 3 implies that lack of salience is not a significant factor in explaining the results. Instead, at least a fraction of students seem to be aware of the exact location of the income thresholds and are able to adjust their labor earnings in response to them, as the distribution exhibits clearly visible bunching just below the thresholds both before and after 2008. Also, the bunching response disappeared below the old threshold immediately after the reform.[13]

Next, we present panel data evidence on the nature of the responses by estimating how students starting from different parts of the base-year income distribution changed their earnings. These results highlight that many students who were located below the threshold before 2008 responded to the reform with a large increase in their earnings.

First, in graph (a) of Figure 5 we analyze average individual-level changes in earnings. Overall, the figure shows that average changes in individual income are very similar in the years before the reform (from 2005 to 2006 and 2006 to 2007), and that there is a visible pattern of mean reversion (on average, starting from a low income level leads to higher income in the next year, and vice versa). The figure shows that labor income increased significantly from 2007 to 2008 compared to the years before the reform for students originally below the threshold. This pattern is observable even for students with base-year earnings around 3000–6000 euros, which is well below the old threshold. However, we find no significant difference between the years for income bins above the new threshold, suggesting that the rapid increase in earnings below the old threshold stems from the change in the location of the income threshold.

Second, graph (b) of Figure 5 presents the likelihood of increasing individual earnings by 50% or more relative to base-year income. We observe that large increases in earnings were significantly more likely for students below the old threshold when the threshold was increased compared to previous years. The prevalence of increases larger than 50%

[13]Additional examination of excess bunching before and after the reform reveals, as further illustrated in Figure B2 in Appendix B, that bunching is slightly larger before the reform than after it in relative terms. One intuitive explanation for this finding is that the local incentives not to exceed the notch are somewhat smaller after 2008, since the relative significance of losing one month's subsidy in terms of disposable income is now smaller than before 2008 when the threshold was at a lower income level.

doubled from 5% to 10% in the income bins below the threshold at the time of the reform but remained constant between the years before 2008. In contrast, there are no significant differences in the likelihood of large earnings increases between the pre- and post-reform years in the bins above the new income threshold.

Third, graph (c) of Figure 5 shows that the likelihood of locating above the old income threshold in the next year increased significantly at the time of the reform, compared to the years prior to 2008. The fact that the likelihood of being located above the threshold increased even when starting from the income bins far below the old threshold further illustrates that a notable share of students responded to the reform with a large increase in their earnings when their budget constraint was relaxed at a higher earnings level.[14]

In summary, we find clear evidence that the 2008 reform that shifted the location of the income threshold for students induced clear earnings responses for students throughout their earnings distribution, and especially for those who were previously located well below the old income threshold. This indicates that the relaxed budget constraint induced large jumps in earnings for many students who were not directly targeted by the reform based on their pre-reform earnings. We explore the potential mechanisms explaining this result below.

4 Conceptual framework and implied mechanisms

The aim of this section is to discuss which labor supply models are and are not compatible with our empirical results presented above. The main feature we want to explain is the shifting of the income distribution from a wide income range below the notch following the change in the location of the notch. We highlight that a discrete labor supply model can explain these results while any version of the continuous models cannot, even when augmented with adjustment costs or optimization errors. We then discuss how discrete labor supply affects empirical welfare loss approximation and empirically estimate the

[14]Furthermore, these panel data results indicate that the observed changes in the cross-sectional earnings distributions of students discussed above stem from intensive-margin responses. To further support this, we find that the share of students not working at all (earning less than 500 euros per year) did not change significantly at the time of the reform. Therefore, potential extensive-margin responses do not explain the change in the shape of the observed earnings distributions.

welfare loss caused by the notch.

4.1 Background

Our starting point is that most of the recent empirical literature on labor supply and earnings responses to taxes assumes a standard continuous labor or earnings supply model either explicitly or implicitly (see e.g. Feldstein 1999 and Saez *et al.* 2012). This applies in particular to the recent literature studying the elasticity of taxable income (ETI), which is in turn a key component in sufficient statistic welfare loss analysis (Chetty 2009a).

The literature has discussed about the validity of the assumptions regarding the labor supply model. Recent argument favoring a continuous model in the ETI literature, following Feldstein (1999), is that because taxable income includes not only regular labor income possibly restricted by fixed working hours but also various amounts of effort, tax avoidance and tax evasion, the sum of all these is continuous even if one of the components is not. However, it is unclear whether the canonical continuous and frictionless model is a good assumption empirically, that is, whether typical taxpayers can vary their effort or engage in tax planning in such an extent that they could smoothly adjust their taxable income in response to taxes (see Kleven *et al.* 2011). Many institutional features could lead to non-smooth income adjustments. For example, the bulk of the labor force typically consists of wage earners, for whom the existence of employment contracts with either fixed terms or notice periods as well as restrictions on working hours and frictions created by wage setting institutions could prevent or hinder smooth income adjustments.

A recently popular extension and challenge to the view of smooth adjustments comes from the optimization frictions literature, which maintains the assumption of continuous labor supply but augments the model with various adjustment costs or optimization errors (see e.g. Chetty *et al.* 2011, Chetty 2012, Chetty *et al.* 2013, Kleven and Waseem 2013, Søgaard 2019, Gelber *et al.* 2020). The motivation for adjustment frictions considered in the literature include, for example, job switching costs, salience of tax rules or unawareness of tax incentives. This literature typically argues that these frictions lead to either mitigated responsiveness or delayed responses to tax changes. Note that fixed costs in job switching or job search analyzed in this literature could also lead to discrete labor supply (see e.g. Chetty *et al.* 2011 and Gelber *et al.* 2020), but this literature has not analyzed how discreteness would affect empirical reduced-form estimation and welfare analysis beyond mitigated or delayed responses to tax changes.

A more fundamental question for the canonical framework is whether labor supply is discrete instead of continuous. The debate over discrete or continuous labor supply has continued for a long time in the economics literature. Structural labor supply literature often assumes that working hours are discrete, stemming from the observation that working hours often tend to cluster at certain focal points in the distribution, such as typical full-time and part-time working hours (see e.g. Dickens and Lundberg 1993, van Soest 1995 , Beffy *et al.* 2018, and Löffler *et al.* 2018 for a recent review). In a similar vein, Saez (1999) argues in one of his extensions to the standard model that workers with specific education, training and occupational skills rarely face a large set of available jobs that match their skills and preferences, thus limiting their potential to flexibly adjust their earnings by changing their job. Relatedly, Kreiner *et al.* (2015) analyze a structural search model with discrete career choices, motivated with similar arguments as Saez (1999). As supporting reduced-form evidence, Blundell *et al.* (2008) find that the responses along the intensive-margin of labor supply for single mothers to changes in various in-work benefit programs in the UK are governed by discrete working hours responses between jobs. Finally, Saez (2002) proposes an optimal income tax model with both participation margin and discrete choices conditional on participation.

Thus, theoretical and structural models have well established how to analyze welfare losses under discrete labor supply, but reduced-form evidence supporting discrete labor supply and the implications of the model for reduced-form estimation methods and sufficient statistics welfare analysis are not that well established. The debate over the labor supply model could have been ongoing for so long partly because typical empirical evidence from tax reforms could often fit both models. We contribute to this discussion by utilizing our empirical setting including a notch that switched location in a reform, which allows us to disentangle between continuous and discrete models underlying the observed responses. Next, we explain in detail why we think our empirical results are consistent with discrete labor supply model, and study the broader implications of the model for empirical reduced-form analysis below in Section 5.

4.2 Labor supply models

We start the exposition with a canonical continuous model and then extend it with adjustment frictions, and then discuss a stylized discrete labor supply model. The canonical model includes a standard utility function $u(c, z)$, where c denotes consumption (net earnings) and z gross earnings, with properties $u_c{>}0$ and $u_z{<}0$. The budget set is $c = (1-\tau)z+R$, where $(1-\tau)$ is the net-of-tax rate and R is virtual income. We abstract from income effects following the earlier literature (Saez *et al.* 2012).[15]

For simplicity and to illustrate transparently how certain extensions modify the model, we parameterize the utility function to a quasi-linear form often used in the earlier literature:

$$u(c, z) = c - \frac{w^i}{1+\frac{1}{e}}\left(\frac{z}{w^i}\right)^{1+\frac{1}{e}}, \tag{3}$$

where w^i is an ability (productivity) parameter of individual i over which individuals are heterogeneous such that there is some underlying distribution of abilities. Thus, the utility maximization with respect to z gives the optimal choice for an individual, $z^* = w^i\,(1-\tau)^e$, where e is the earnings elasticity parameter with respect to τ. Thus, the earnings choices are governed by innate productivity w^i, marginal tax rate τ, and e parameter.

We want to understand how this model would explain the observed changes in the income distribution following an upward shift in the location of an income threshold, such as in the empirical reform we study. If individual's optimal choice is strictly below the original location of the notch, they will not respond at all to this kind of a reform. This is because none of the parameters determining the individual's originally chosen location

[15] Including income effects would not change the main results in qualitative terms but would complicate the formulas.

has changed, including τ. Therefore, this model cannot explain the responses in income levels reaching far below the original location of the notch we observed empirically. Panel (a) of Figure A4 in Appendix A illustrates this graphically.[16]

Following Kleven and Waseem (2013), optimization frictions that simply mitigate or hamper responses to taxes can be included in the parameterized continuous model by adding to the utility function a heterogeneous friction parameter $a \in (0, 1)$. If a is close to one on average, frictions are high and average individual responses to taxes would be heavily restricted, and if a is zero, responses to taxes would follow the standard model above. We assume that a enters the parametrized utility function as follows:

$$u(c, z) = c - \frac{w^i}{1 + \frac{1}{e(1-a)}} \left(\frac{z}{w^i}\right)^{1+\frac{1}{e(1-a)}}. \tag{4}$$

From the above equation it becomes clear that these types of optimization frictions merely reduce the responsiveness to taxes, but they do not alter individual responses in a more fundamental manner. In particular, this modification does not alter the above consideration of whether the modified model explains the empirically observed changes in the distribution following the reform. It continues to apply that individuals located below the original location of the notch would not consider responding to the reform because the parameters defining their earnings location remain unchanged.

We can further alter the canonical framework by adding optimization errors to the model arising from an unanticipated shock to the initially chosen income. A simple approach to including optimization errors is to consider an error parameter drawn from some distribution $r \in f(r)$ that alters the optimized income z^* into $z^* - r$.[17] These kinds of frictions would typically cause only small deviations in income (depending of course on the size of the shock), but again would not induce individuals located far below the

[16]Note that if the initial location was at the notch, this model would predict those individuals shifting upwards as a response to the reform. This is because individuals would have bunched at the notch in response to the discontinuous incentives, and the reform would have removed the disincentive to be located at higher income levels.

[17]If individuals have an expectation of the shock or are risk-averse, they could respond to the expectation but not to the realized shock.

notch originally to respond to the reform.

In contrast, discrete rather than continuous labor supply can explain large jumps across earnings levels in the distribution. Discrete labor supply could be modeled in a number of ways depending on the institutional reasons for why labor supply might be discrete, as discussed above. Following Saez (2002), we next construct a simple version of a discrete model by including to the canonical model a constraint that an individual must choose her earnings level from a sparse set of available locations. If we modeled discrete labor supply with fixed costs for changing the income level, such as job switching or search costs, we would arrive at similar conclusions as below but the model would be much more specific in terms of the mechanisms inducing discrete labor supply.

More formally, in our stylized model individuals choose from a discrete set of alternative earnings locations $j = 1, ..., N$, yielding utility $u(c_{j-1}, z_{j-1}),\ u(c_j, z_j),\ u(c_{j+1}, z_{j+1})$, but individual preferences and the underlying ability distribution are continuous as before. Thus the choices are discrete even conditional on participating in the labor market. The budget set is now described as $c_j = w_j - T_j + R$, where T_j is the average tax at location j. Considering two locations $j-1$ and j, individual chooses the one which yields the highest utility:

$$u(c_{j-1}, z_{j-1}) \leq u(c_j, z_j) = c_j - \frac{w^i}{1+\frac{1}{e}} \left(\frac{z_j}{w^i}\right)^{1+\frac{1}{e}} \tag{5}$$

The main conceptual difference between this model and the continuous model is that individuals now consider which one of the two distinct earnings levels yields the highest utility. For example, if an individual is located at z_{j-1} and a tax rate cut occurs such that it applies to z_j but not to z_{j-1}, in this model an individual could switch from z_{j-1} to z_j, leading to a potentially large jump in earnings depending on how far apart the two locations are. Note that similarly to above, the tax change in this example does not directly alter the budget set at z_{j-1}. Therefore, a model including a discrete choice component can rationalize much larger jumps in earnings as a response to a local tax rate change compared to any of the continuous models considered above. Consequently, out of the models considered here only the discrete labor supply model can explain our empirical result that the distribution shifted from a wide income range below the original location of the notch as a response to the reform that shifted the location of the notch. Figure A4 in Appendix A illustrates the discrete case graphically. Note that even if one individual faces only a fixed set of discrete locations, the discrete model can produce a smooth aggregate income distribution when the extent of discreteness and available locations are heterogeneous across individuals.

For the empirical applications we consider below, it is useful to review the elasticity concepts associated with the discrete labor supply model. In the discrete model the elasticity concept correctly capturing responses to a change in average taxes is a mobility elasticity, not the standard marginal elasticity (Saez 2002, Kreiner *et al.* 2015). Following Kreiner *et al.* (2015), we can express the mobility elasticity with the following equation:

$$\zeta = \frac{dY}{d(1-m)} \frac{1-m}{Y} \tag{6}$$

where m is the average tax rate difference between two different locations j and $j-1$, $m = \frac{T_j - T_{j-1}}{w_j - w_{j-1}}$, and $Y = \sum_{j=1}^{N}(z_j g_j)$ where g_j is the relative mass of individuals in each earnings location. Equation (6) thus captures the change in earnings inflicted by individuals moving between different earnings locations due to changes in the average tax rate differential between these locations. This elasticity formula captures two important features that are missing from the continuous model when considering discrete responses. First, mobility elasticity captures earnings responses over a broader income range across multiple earnings locations, denoted above by dY. Second, the standard measure for the change in the marginal tax rate does not capture the change in tax incentives across distinct earnings locations, whereas the average tax change across locations (m) does.[18]

Finally, we build a simulation model that enables us to illustrate the income responses to a relocation of an income notch under the different modeling choices considered above. Appendix C presents the results of the simulation analysis graphically. Consistent with the discussion above, we find that none of the variants of the continuous model reproduce

[18]Note that the two tax rate concepts τ and m coincide when the tax system is linear across the distribution.

the widespread earnings responses in the distribution from below the original notch or the shape of the excess mass right below the notch we observed in our empirical case. In contrast, the discrete model produces distribution-wide shifting as a response to the reform and a shape of the excess mass that are at least qualitatively similar to their empirical counterparts. In more detail, assuming 10–15 available earning locations within the distribution for each student yields the closest match with the empirical case.[19] We could potentially produce an even closer match with our empirical observations by assuming further features such as adjustment costs and optimization errors on top of the simple discrete model, but the main point of these simulations is to highlight that we need some kind of a discrete component in the model in order to even qualitatively match the empirical observations.

4.3 Further empirical evidence on discrete responses

Above we provided theoretical reasoning that only the discrete labor supply model is consistent with our main empirical results regarding the widespread changes in the distribution due to the shifting of the notch. Next, we provide further empirical evidence supporting the discrete labor supply model and some thoughts on the external validity of our findings.

We divide individuals in our estimation sample into subgroups based on the industry where they work in. The two groups feature arguably different degrees of discreteness in their labor market based on the typical job characteristics in the industry, such as employment contracts and working hours. In the more discrete group we include the whole public sector, and research, manufacturing and construction industries in the private sector. The less discrete group includes restaurants, bars and cafes, hotels and other accommodation services, cleaning and security services, and retail sales such as supermarkets and gas stations. In the latter group working hours are typically more flexible and wages are more likely to be defined on a hourly rather than monthly basis, creating

[19] When increasing the number of available discrete earnings locations the distributions begin to resemble those in the continuous model, and with only a very small number of available choices the broader changes in the distribution are also small, as can be expected.

less discreteness in their labor market compared to the first group.

Figure 6 shows that the extent of changes in the earnings distribution at the time of the 2008 reform are smaller for those students who work in less discrete labor markets (6.14 (*1.71*)) compared to those working in more discrete labor markets (10.94 (*1.10*)). This difference supports our assertions above that discrete labor supply is a key factor explaining our main result.

Furthermore, this finding does not only support the discrete model in itself, but also suggests that the discreteness in earnings responses at least partly stems from how the labor markets function. Importantly, this result suggests that discrete labor supply is not specifically related to students' behavior. When looking at the labor force as a whole, the observation about the role of discrete labor markets gains more weight, as a significant share of regular wage earners in Finland are employed in the discrete labor market category above (61% in our data including the whole Finnish work force). This suggests that discrete labor supply can induce a relevant constraint for a large share of the regular labor force, at least in developed countries.

Relatedly, we use our data on the entire labor force in Finland to elucidate whether typical individual-level changes in earnings are supported by continuous or discrete models. Figure A5 in Appendix A shows the distribution of one-year changes in earnings for regular wage earners in 1999–2013 in Finland.[20] Figure shows that a large share of these changes are distinctively large. Overall, 44% of all changes in earnings are below 5% in the figure, which could be explained, for example, by changes stemming from centralized wage bargaining. More importantly, as much as one third of all changes are above 10% and one fourth are above 20%. This indicates that a significant proportion of typical annual wage adjustments occur in a manner that is more compatible with discrete changes rather than smooth and continuous adjustments, supporting again the discrete labor supply model. Guvenen *et al.* (2019) find a very similar overall pattern in individual earnings adjustments for US workers, offering further general motivation for the relevance of discrete changes in earnings for wage earners.

[20]The figure includes those wage earners with wage income of at least 30,000 euros in the base year.

4.4 Welfare loss estimation under discrete model and mobility elasticity for students

A key method for empirically analyzing the welfare loss is the sufficient statistic approach (see e.g. Chetty 2009a and Kleven 2020). This approach is attractive because under certain assumptions it is enough to estimate average ETI to approximate the welfare loss of taxes. More formally, following Chetty (2009b) and Slemrod (1998), the welfare (W) effect of a tax change $d\tau$ in a population of taxpayers can be presented as follows

$$\frac{dW}{d\tau} = \frac{dR}{d\tau} - \frac{\partial R}{\partial \tau}|_M \tag{7}$$

where $dR/d\tau$ denotes how total tax revenue (aggregate earnings over all individuals times the tax rate) responds to a change in the tax rate, and $\partial R/\partial \tau|_M$ denotes the mechanical change in revenue absent any behavioral responses. The welfare loss $dW/d\tau$ is defined over the whole population and measures the aggregate change in welfare. As discussed in Chetty (2009a) and Kleven (2020), the underlying earnings model does not change the basic idea of the sufficient statistic approach, for example, assuming discrete rather than continuous labor supply. However, the underlying model can affect how the welfare loss is evaluated and which empirical estimates are sufficient to evaluate it.

In theory, in the continuous model income responses to a tax change are local and apply only to those individuals who directly face changes in tax rates, as discussed above. Thus, under this model the average ETI is a sufficient statistic to capture all income responses and the welfare loss (under certain further assumptions as discussed in Kleven 2020). In contrast, in a discrete model the responses occur over a wide income range even if the reform is local in nature (Saez 2002, Kreiner *et al.* 2015). This is also what we find in our empirical application in Section 3. Therefore, because the aggregate welfare loss estimate should include all effects of taxes to arrive at the welfare loss $dW/d\tau$, we should not only estimate a local average elasticity but also sum this elasticity over all individuals who respond to a particular tax policy under discrete labor supply. In other words, a sufficient statistic to empirically evaluate the welfare loss is the range over which individuals respond in addition to the average elasticity measuring the sensitivity of disutility to taxes. Also, this indicates that even when the average elasticity estimate is smaller in a discrete model compared to a continuous model, the aggregate earnings responses over the distribution capturing the true welfare losses can be similar in both cases. We further illustrate this in our simulation analysis in Section 5.

One practical implication of the broader sufficient statistics concept is that one should observe which parts of the income distribution respond to a local tax reform. This notion favors distributional methods, such as the one we propose in this study, or methods to estimate distributional impacts of programs proposed in the recent literature (Athey and Imbens 2006, Firpo and Pinto 2016 and Cengiz *et al.* 2019) to discover the full extent of responses to taxes. Instead methods that estimate only the average effects, such as differences-in-differences are not that well suited to estimate the distributional aspect of responses to taxes.

Next, we estimate a mobility elasticity for students utilizing the 2008 reform to empirically approximate the welfare losses created by the notch. We apply equation (6) where the necessary ingredients are the estimated earnings responses to the reform and the changes in average tax rates due to the reform.

First, to measure the changes in average tax rates caused by the reform, we simplify the setting by assuming that students on average choose from only two earnings locations: one below the old income threshold and one above it. This simplification facilitates the calculations considerably because we do not observe the counterfactual location for each individual, and should produce a roughly correct estimate on the average changes in incentives over the distribution. Average gross earnings were 6008 euros below the old threshold in 2007 in the income range of 2000–9200 euros, and 11,821 euros above the old threshold in 2008 in the income range of 9201–18,000 euros.[21] By using the actual tax and subsidy rules before and after the 2008 reform, we find that the net earnings difference between these two locations increased from 3534 to 4807 euros due to the shifting of the notch, highlighting the significant impact of the reform on labor supply incentives.

[21]We limit our analysis to the income range of 2000-18,000 euros as there were no changes in the distribution in the area below 2000 or above 18,000 euros between 2007–2008 (see Figure 3).

Next, we define the average earnings response to the reform. Average earnings were 7116 and 7529 euros in 2007 and 2008, and thus the average real earnings within an income range of 2000–18,000 euros increased by 413 euros from 2017 to 2018.[22] To approximate the mobility elasticity, we relate the average change in log gross earnings to the log change in the difference in net incomes between the average earnings locations described above. This delivers a mobility elasticity estimate of 0.183.[23] Therefore, even though the reform caused large and distinctive earnings responses over a wide income range in the distribution, the strong change in incentives caused by the reform implies that the elasticity estimate is nevertheless modest.

Our elasticity estimate is close to what Søgaard (2019) finds for university students in Denmark (0.1) who face shifting of a kink instead of a notch in their study subsidy program, but the income range over which the responses occur are much more modest in the Danish setting. Because the notch creates much stronger incentives compared to a kink in our setup, it is not surprising that the responses occur over a broader income range than in the Danish setting, although the average elasticities are similar. Also, our estimated elasticity is within the range of the average ETI estimates in the literature (see Saez *et al.* (2012) for a survey), but this literature does not typically assess the aggregate income changes over the distribution as in our mobility elasticity estimation. We further discuss the implications empirical welfare loss analysis under discrete labor supply in the next section.

5 Implications of discrete labor supply on elasticity estimation methods and empirical welfare analysis

In this section, we first present our simulation model and then apply it to illustrate how standard reduced-form tax elasticity estimators can be biased when the actual labor supply model is discrete rather than continuous. The benefit of using the simulation

[22] As shown in Figure 4, there were no significant changes in the annual earnings of students in years before or after the 2008 reform, and no changes in the earnings of the control group of young non-student part-time workers.

[23] Table A2 in Appendix A presents the variables used to calculate the mobility elasticity estimate.

approach is that we can track down the impact of different assumptions about labor supply model and econometric methods on the elasticity estimate. Then we discuss how the mechanisms related to labor supply affect the empirical analysis of welfare losses. Also, we discuss how discrete responses affect the bunching method.

5.1 Simulation model

We build our simulation model on the theoretical framework presented in Section 4. The individual utility function is given in equation (3), where the e parameter governs the extent of disutility arising from supplying earnings and would correspond to the elasticity of earnings with respect to taxes in the standard continuous model. The discrete model presented in equation (5) has the same utility function but limits the available earnings locations to a fixed and sparse set as an additional constraint. However, as discussed in Section 4, the model would produce similar results if we had modeled discrete labor supply with an alternative mechanism, for example, job search or job switching costs (see e.g. Gelber *et al.* 2020). We simulate various versions of the budget set arising from the tax and benefit system.

The model assumes an underlying ability distribution from which each individual i receives a predetermined draw w_i, representing earnings in the absence of responses to the tax system. Our parameterized ability distribution is presented in Figure C1 in Appendix C.[24] The available discrete earnings locations for each individual are drawn from the probability distribution presented in Figure C2 in Appendix C. In qualitative terms, our results are not sensitive to the shape of these distributions. The number of choices drawn can be altered in different specifications and the draws vary between different individuals. Therefore, even when the individual-level choices are discrete, the overall earnings distribution is smooth.

In the model, we focus on labor supply conditional on participating in the labor market and, for simplicity, exclude the participation margin from the analysis. Also, this

[24]The distribution is a combination of power distributions and normal distributions, which gives an approximate match for the shape of the empirical earnings distribution of students in our empirical case. Our results are not sensitive to different underlying ability distributions that roughly match the shape of the empirical income distributions of students.

restriction enables us to better compare the simulated estimates to those derived in the ETI literature that focuses on estimating intensive margin responses.

5.2 Bias in tax elasticity estimation

Next, we illustrate how discrete labor supply can lead to a bias in the estimates of the elasticity of taxable income (ETI) when estimated with methods often used in the ETI literature. Typically, panel data ETI estimates are conducted using a differences-in-differences (DiD) style approach involving a treatment group facing a tax reform based on their pre-reform earnings and a control group that was unaffected (or less affected) by the reform (see e.g. Gruber and Saez 2002, Saez *et al.* 2012). Often the control group comprises those with slightly lower pre-reform earnings than the treatment group within the same general population, because finding a control group resembling the treatment group but facing a completely different tax system tends to be difficult.[25]

If labor supply is continuous, the above definitions for the treatment and control groups could deliver an unbiased elasticity estimate. However, under discrete labor supply these definitions are no longer valid. As discussed in Section 4, discrete labor supply expands the group of individuals who are affected by tax reforms, raising two main issues in the standard estimation: first, the group that is potentially affected by the reform is larger than those directly affected based on their pre-reform earnings. Second and more importantly, the typical control group with lower pre-reform earnings can also respond to the reform.

To make the discussion about these issues more rigorous, we need to specify the equations for the standard ETI and mobility elasticity. The standard ETI is estimated by regressing the change in log income on the change in the log of the net-of-tax rate as follows:

$$d\log z_{it} = \varepsilon d\log(1-\tau)_{it} + \epsilon_{it} \tag{8}$$

[25] There are important quasi-experimental tax elasticity studies that do not have this feature. For example, the literature studying the impact of the earned income tax credit (EITC) on earnings in the US typically applies an estimation strategy where women without children are used as a control group for women with children who are eligible for larger tax credits (see Kleven 2019 for a recent review).

where subscript i refers to an individual, t time, z (taxable) earnings, τ the marginal tax rate, ϵ is the error term, and ε is the estimated average earnings elasticity with respect to the net-of-tax rate.

The mobility elasticity formula differs from this specification because it needs to account for the fact that under discrete model a larger group of individuals could respond to a change in the tax rate than those currently at the location of the tax change, as discussed above. Moreover, the correct tax rate measure in the discrete model, denoted by m, differs from the marginal tax rate used in the ETI formula. It is the difference between net earnings relative to gross earnings across different earnings locations, as discussed in Section 4. Empirically, m can be calculated by denoting the initial earnings location as $i-1$ and the new location (pre-reform income plus the change in income) as i. This implies that the change in the tax rate is potentially non-zero for a large part of the earnings distribution, also outside the direct region of the reform. Formally, the mobility elasticity is estimated with the following equation:

$$d \log z_{it} = \zeta d \log(1-m)_{it} + \epsilon_{it} \tag{9}$$

where ζ denotes the average mobility elasticity.

We simulate an income distribution calibrated to match the empirical distribution of students (income range of 0–25,000 euros with mean income of approximately 6000 euros), which has lower average earnings but roughly a similar shape to the earnings distribution for all wage earners. The simulated tax schedule features a marginal income tax rate of 40%, which jumps to 60% above a predetermined income level. We then introduce a reform that repeals this tax kink and reduces the income tax rate to 40% also for the higher incomes above the original kink. This simulated reform cutting income tax rates for higher incomes is similar to many actual reforms that have occurred in developed countries in recent decades (see e.g. Saez *et al.* 2012). In the baseline model, we assume that the original tax kink is located at 10,000 euros, an underlying e parameter of 0.5 and 10 available discrete earnings choices for each individual, which provide a qualitative match for the observed distributions of students (see Appendix C). We later vary these parameters to illustrate the sensitivity of the results to them.

Figure 7 describes the simulated individual-level earnings responses and incurred tax rate changes when starting from different pre-reform income levels using 200 euro income bins in both the discrete and continuous cases. In the continuous case, the change in log earnings and the change in the net-of-tax rate $(1 - \tau)$ are zero for those originally below the kink and a constant 0.2 and 0.4 above the kink, respectively. Instead, in the discrete case the change in the tax rate m does not coincide with τ below the kink; a fraction of these individuals face a change in their tax rate leading to a positive average change in the net-of-tax rate.[26] We also observe from the figure that earnings respond in the discrete case in a wide area, consistent with the predictions of the theory model.[27]

Next, we illustrate graphically the earnings elasticities at different income levels in the continuous and discrete cases in Figure 8. As can be expected, the elasticity in the continuous model is constant 0.5 above the kink point, which equals the chosen e parameter in the simulation model. In contrast, the mobility elasticity estimate in the discrete model is smaller and not constant across the distribution. On average, the mobility elasticity is 0.14, and slightly larger (0.2) above the original kink and smaller (0.12) below it. Thus, Figure 8 provides a simple illustration of the bias in the standard ETI estimation that would arise from defining as treated those who were above the kink before the reform and as control those at lower pre-reform income levels when the labor supply model is discrete instead of continuous. We discuss this bias in more detail below.

Table 3 collects the estimates in simulated microdata. We use two different estimation approaches: the "naive" ETI approach using equation (8) and defining as the treatment group those who were above the kink before the reform and as the control group those in the income range of 5000 euros below the kink, and the mobility elasticity using equation (9), where we consider as treated all individuals within the same income range as in the

[26]The tax rates do coincide above the original kink because of the assumed linear tax system.

[27]Figures A6 and A7 in Appendix A present the underlying simulated earnings distributions before and after the reform in the discrete and continuous cases, respectively. These figures show that the shape of the earnings distribution changes over a broader region below the kink point in the discrete case but no such changes occur in the continuous case. Also, there is significant and sharp bunching in the continuous case just at the kink point but no clear bunching in the discrete case prior to the repeal of the tax rate kink, as can be expected.

naive ETI approach and using a simulated pure control group not affected by the reform.

The results in Table 3 show that the naive ETI estimates are always smaller than the average mobility elasticity estimates, illustrating the downward bias in the standard ETI estimate. A comparison between the estimates suggests that the size of this bias can be notable. In our baseline specification with the original kink at 10,000 euros, the mobility elasticity estimate is more than two times larger than the naive ETI estimate (0.161 vs. 0.07). However, the discrete earnings constraint can be different across various types of wage earners. Tables A3–A5 in Appendix A present the estimates when we vary the number of available earnings choices. The bias is visible in all cases but naturally varies across different specifications, and it is in general larger when earnings responses are more discrete.

Overall, these simulations show that the validity of the methods for obtaining quasi-experimental ETI estimates can be sensitive to the underlying labor supply mechanisms. If the actual earnings responses are discrete instead of continuous, they can create a significant bias to the estimate when using standard estimation methods based on the assumption of the canonical continuous model. As earnings adjustments can be rather constrained for a large share of workers, we believe that these results illustrate what could happen in many actual empirical settings for wage earners.

5.3 Implications for empirical welfare loss analysis

We discussed in Section 4.4 how discrete labor supply potentially affects empirical estimates for the welfare loss of taxes. Next we use our simulation results from above to further illustrate these issues.

First, Table 3 shows that the average mobility elasticity estimates (e.g. 0.16 with 10 discrete choices in Table 3 column (2)) are in general smaller in the discrete model than the unbiased ETI in the continuous model (0.5 in the model and 0.45 if the narrow bunching window is included in the empirical estimate). However, Figure 8 shows that even though the elasticity estimates in the discrete case are lower throughout the distribution, the responses occur over a wider income range compared to the continuous model.

To provide precise numbers, in the discrete model with 10 available locations and the reform taking place at 10,000 euros, the change in average income in the whole distribution is 428 euros. In comparison, in the continuous model with the same tax reform the average change in income is very similar, 492 euros. Given that the averages are calculated over the populations of equal size in the simulations, this implies that the aggregate welfare loss is only slightly smaller in the discrete model. This suggests that the aggregate welfare loss estimates, calculated as changes in income over the whole distribution, are not as different between the models as would be apparent if relying only on average elasticities.

These stylized findings from our simulation model highlight that it is important to take into account all changes that appear in the distribution as a response to a particular tax policy, and not focus, for example, only on the income range where the tax reform occurs. Again, this notion favors distributional methods over local elasticity estimators to capture the total earnings impact of tax reforms, for example, the changes-in-changes methods proposed by Athey and Imbens (2006) and recently applied, for example, by Kottelenberg and Lehrer (2017), methods to capture program effects in distribution, for example by Firpo and Pinto 2016 or changes by bins of the wage distribution as proposed by Cengiz *et al.* (2019).

5.4 Bunching estimation

Discrete labor supply has major implications for the bunching method that has already been criticized for various other reasons (Saez 2010, Kleven and Waseem 2013, Blomquist *et al.* 2019). Discreteness leads to widespread changes in the distribution, which in turn implies that the bunching method focusing on the narrow window around the discontinuity produces downward-biased estimates, because any potential responses occurring outside of this window are ignored. Also, for the same reason the surrounding density outside the bunching region cannot be used to estimate an unbiased counterfactual describing the shape of the distribution in the absence of a local discontinuity. Moreover, using the missing mass at the dominated region just above an income notch cannot deliver relevant information on the extent of short-run frictions (see Kleven and Waseem 2013), because an income range above a notch is no longer dominated when discrete earnings locations are sufficiently far away from each other.

In our empirical application, we observe distinct bunching at and just below the notch. Thus, it would be tempting to use the bunching method in a cross-sectional setting to estimate ETI. Now that we have access to the reform that shifted the location of the notch and observe the widespread changes in the distribution the reform causes, we can use that empirical setup to estimate the mobility elasticity, and to highlight the welfare loss that we would have missed if we had relied solely on the bunching estimate. More precisely, using the reduced-form earnings elasticity formula for income notches presented in Kleven and Waseem (2013), we obtain a local elasticity estimate of 0.065 (*0.007*) for students with 9 subsidy months (see Appendix B for more details). In contrast, our approximation for the mobility elasticity estimate (0.18) presented above in Section 4.4 that captures the whole response to the reform is approximately 2.5 times larger (0.18 vs. 0.065).[28]

More broadly, discrete labor supply could provide one explanation for why numerous previous studies tend not to observe bunching at tax rate kinks among wage earners (see e.g. Saez 2010 and Bastani and Selin 2014). Our simulated distributions for the kinked budget set in Figures A6 and A7 in Appendix A show that there is no visible excess mass at a tax rate kink when assuming discrete labor supply, whereas bunching is very sharp and distinct in the continuous case. These notions could be relevant for many studies considering the bunching approach, as discrete labor supply could be a relevant model for a large share of wage earners.

[28] An alternative approach to measure the difference between the approaches is to relate the broader changes in the distribution to the local excess mass estimate. In Figure 3, we find that the broader changes in the earnings distribution density are approximately three times larger than the bunching estimate (9.81 vs. 2.93).

6 Concluding remarks

In this paper, we provide novel reduced-form evidence of distribution-wide changes in earnings as a response to a reform that shifted the location of a notch in the income tax schedule. We find that the reform caused changes in earnings over a broad income range among Finnish university students. We then show that these patterns are consistent with a discrete labor supply model instead of a continuous model even when the continuous model is augmented with adjustment frictions that merely reduce responsiveness to taxes.

We analyze the broader implications of discrete labor supply on empirical estimation methods and find that standard reduced-form estimation methods used in the literature, such as differences-in-differences and bunching, need to be adjusted when labor supply is discrete. Discrete labor supply expands the group of individuals who are potentially affected by tax rate changes, and consequently renders individuals not facing a reform based on their current earnings an invalid control group. We find that not taking these issues into consideration could result in a significant downward bias in the elasticity estimate.

Also, discrete labor supply can affect the welfare evaluations beyond the bias in the average elasticity estimate. When estimating aggregate welfare losses using the sufficient statistics approach, and labor supply, we should take into account the extent of changes in the overall earnings distribution caused by tax changes in addition to the average elasticity when labor supply is discrete. Thus, the sufficient statistics for estimating aggregate welfare losses would need to include the extent of individuals responding to taxes. For this the distributional quasi-experimental methods are well suited to discover what kind of changes taxes cause in the distribution. We proposed one such method in this paper, and the recent literature contain others suitable for this purpose (for example Athey and Imbens 2006, Firpo and Pinto 2016 and Cengiz *et al.* 2019).

References

[1] Altonji, J. and C. Paxon (1988). Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-offs. *Journal of Labor Economics*, 6(2): 254–276.

[2] Athey, S. and G. Imbens (2006). Identification and Inference in Nonlinear Difference-in-differences Models. *Econometrica*, 74(2): 431–497.

[3] Bastani, S. and H. Selin (2014). Bunching and Non-bunching at Kink Points of the Swedish Tax Schedule. *Journal of Public Economics*, 109: 36–49.

[4] Beffy, M., R. Blundell, A. Bozio, G. Laroque and M. To. (2019). Labour Supply and Taxation with Restricted Choices. *Journal of Econometrics*, 211: 16–46.

[5] Blomquist, S., W. Newey, A. Kumar and C. Liang (2019). On Bunching and Identification of the Taxable Income Elasticity. NBER Working Paper No. 24136.

[6] Blundell, R., M. Brewer and M. Francesconi (2008). Job Changes and Hours Changes: Understanding the Path of Labor Supply Adjustment. *Journal of Labor Economics*, 26(3): 421–453.

[7] Cengiz, D., A. Dube, A., Lindner and B. Zipperer (2019). The Effect of Minimum Wages on Low-wage Jobs. *Quarterly Journal of Economics,* 134(3): 1405–1454.

[8] Chetty, R. (2009a). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics*, 1: 451–488.

[9] Chetty, R. (2009b). Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance. *American Economic Journal: Economic Policy*, 1(2): 31–52.

[10] Chetty, R., J. Friedman, L. Pistaferri and T. Olsen (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *Quarterly Journal of Economics*, 126(2): 749–804.

[11] Chetty, R. (2012). Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, 80(3): 969–1018.

[12] Chetty, R., J. Friedman and E. Saez (2013). Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review*, 103(7): 2683–2721.

[13] Firpo, S. and C. Pinto (2016). Identification and Estimation of Distributional Impacts of Interventions Using Changes in Inequality Measures. *Journal of Applied Econometrics*, 31(3): 457–486.

[14] Dickens, W. and S. Lundberg (1993). Hours Restrictions and Labor Supply. *International Economic Review*, 34 (1): 169–192.

[15] Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics*, 81(4): 674–680.

[16] Firpo, S., N. Fortin and T. Lemieux (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3): 953–973.

[17] Gelber, A., D. Jones and D.W. Sacks (2020). Estimating Earnings Adjustment Frictions: Method and Evidence from the Earnings Test. *American Economic Journal: Applied Economics*, 12(1): 1–31.

[18] Gruber, J. and E. Saez (2002). The Elasticity of Taxable Income: Evidence and Implications. *Journal of Public Economics*, 84: 1–32.

[19] Guvenen, F., F. Karahan, S. Ozkan and J. Song (2019). What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics? Working paper.

[20] Kleven, H. and M. Waseem (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *Quarterly Journal of Economics*, 128: 699–723.

[21] Kleven, H. (2016). Bunching. *Annual Review of Economics*, 8: 435–464.

[22] Kleven, H. (2019). The EITC and the Extensive Margin: A Reappraisal. NBER Working Paper No. 26405.

[23] Kleven, H. (2020). Sufficient Statistics Revisited. *Annual Review of Economics*, forthcoming in volume 13, 2021.

[24] Kreiner, C., J. Munch and H. Whitta-Jacobsen (2015). Taxation and the Long Run Allocation of Labor: Theory and Danish Evidence. *Journal of Public Economics*, 127: 74–86.

[25] Kottelenberg, J., K. and S. F. Lehrer (2017). Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Child Care. *Journal of Labor Economics*, 35(3).

[26] Löffler, M., A. Peichl and S. Siegloch (2018). The Sensitivity of Structural Labor Supply Estimations to Modeling Assumptions. ifo Working Papers 259/2018.

[27] Neisser, C. (2021). The Elasticity of Taxable Income: A Meta-Regression Analysis. Forthcoming, *Economic Journal*.

[28] Rosen, H. (1976). Taxes in a Labor Supply Model with Joint Wage-Hours Determination. *Econometrica*, 44(3): 485–507.

[29] Saez, E. (1999). Do Taxpayers Bunch at Kink Points? NBER Working Papers 7366.

[30] Saez, E. (2002). Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses. *Quarterly Journal of Economics*, 117: 1039–1073.

[31] Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3): 180–212.

[32] Saez, E., J. Slemrod and S. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, 50(1): 3–50.

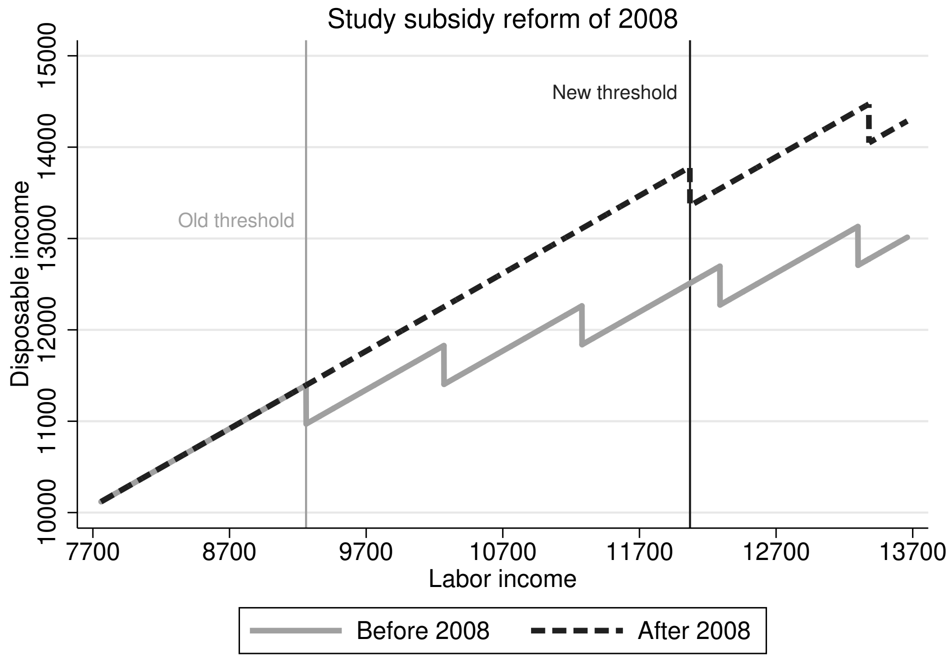
[33] Slemrod, J. (1998). Methodological Issues in Measuring and Interpreting Taxable Income Elasticities. *National Tax Journal*, 51(4): 773–788.

[34] Søgaard, J. (2019). Labor Supply and Optimization Frictions: Evidence from the Danish Student Labor Market. *Journal of Public Economics*, 173: 125–138.

[35] van Soest, A. (1995). Structural Models of Family Labor Supply - a Discrete Choice Approach. *Journal of Human Resources*, 30(1): 63–88.

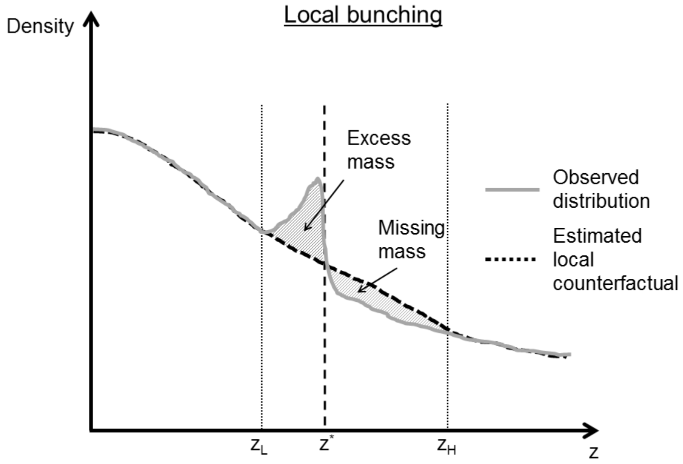
[36] Weber, C. (2014). Toward Obtaining a Consistent Estimate of the Elasticity of Taxable Income Using Difference-in-differences. *Journal of Public Economics*, 117: 90–103.

Figures

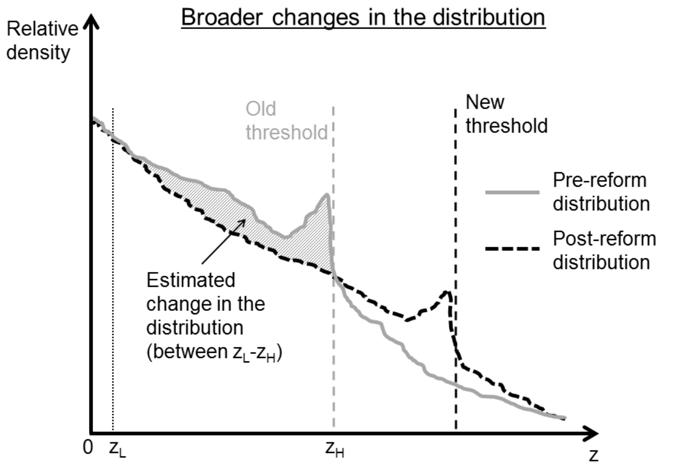


Notes: Figure presents the study subsidy schedule before (gray solid line) and after 2008 (black dashed line) for a student who collects the default 9 subsidy months. The vertical axis denotes disposable income, and horizontal axis labor income. The vertical lines denote the thresholds before (9200 euros) and after (12,070 euros) the 2008 reform. Above the income threshold one month of the study subsidy is reclaimed, resulting in a discontinuous drop in disposable income. Furthermore, an additional month of the subsidy is reclaimed after an additional 1010 and 1310 euros above the threshold before and after 2008, respectively. The figure illustrates the distinctive change in incentives caused by the increase in the income threshold in 2008, highlighting that the reform encouraged students to increase their earnings above the old income threshold.

Figure 1: Disposable income at different income levels for students with 9 subsidy months in 2007 and 2008



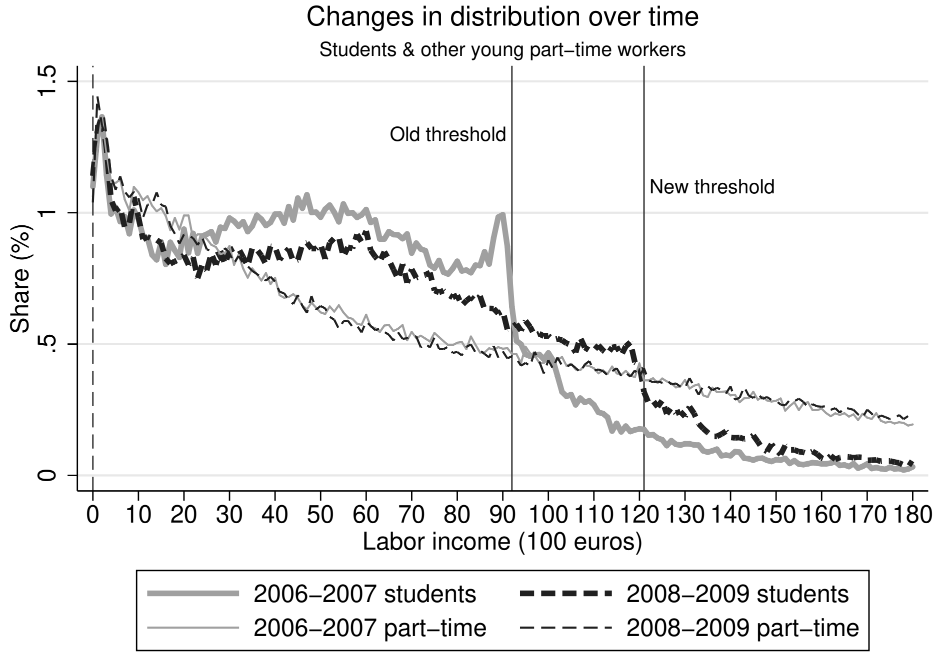
(a) Local bunching



(b) Broader changes in the distribution

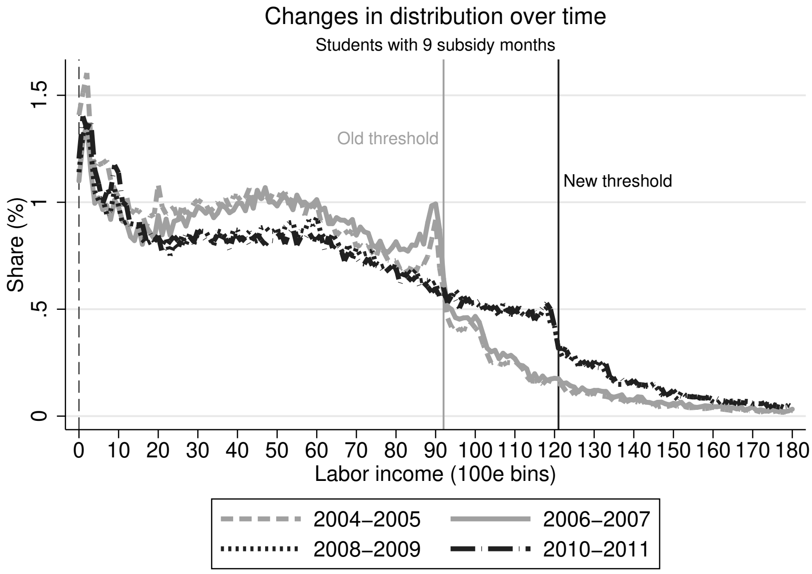
Notes: Graph (a) illustrates the excess bunching at the income threshold in a hypothetical earnings (z) distribution (gray solid line), compared to an estimated counterfactual distribution in the absence of the threshold (black dashed line). In the figure, the threshold is denoted by z^*, and z_L and z_H denote the lower and upper limits of the bunching region. The counterfactual is estimated by fitting a flexible polynomial function to the observed distribution excluding the area close to the notch between z_L and z_H from the regression. Excess bunching is estimated by relating the share of individuals in the bunching region (z_L, z^*) to the counterfactual density. See Appendix B for more details on the bunching estimation.
Graph (b) illustrates broader changes in a hypothetical earnings distribution after an increase in the location of the threshold. The pre-reform distribution is marked with a gray solid line and the post-reform distribution with a black dashed line. z_L and z_H denote the lower and upper limits of the estimation region. Broader changes in the distribution are estimated by relating the observed relative density before the reform to the relative density after the reform in the income range between the lower and upper limits. See Section 2.3 for more details on the estimation method.

Figure 2: Estimating bunching and broader changes in the earnings distribution

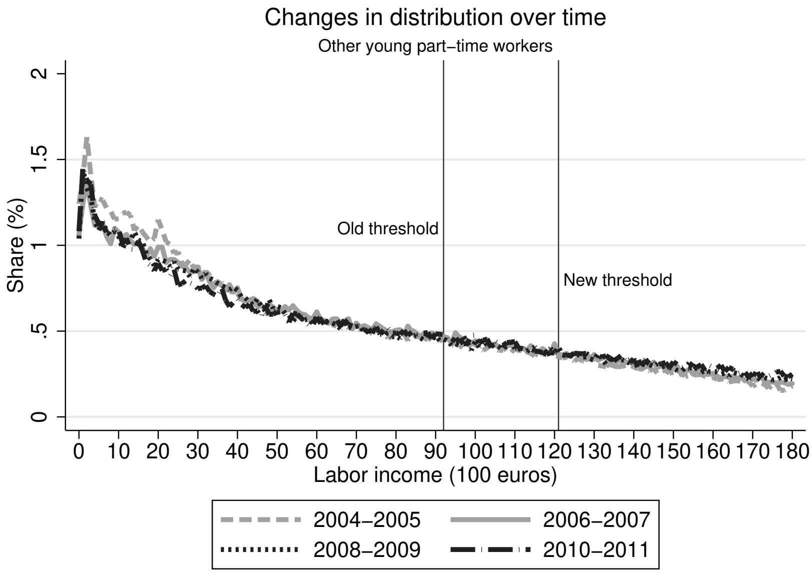


Notes: Figure presents the observed relative earnings distributions before the reform in 2006–2007 (gray solid line) and after the reform in 2008–2009 (black dashed line) within an income range of 0–18,000 euros in bins of 100 euros for students with the default 9 subsidy months in each year, and for young part-time workers who are not students (see Table 2). The first vertical line at 0 denotes the lower limit in the estimation of broader earnings changes in the distribution estimated using equation (2), and the second and third lines denote the pre- and post-reform income thresholds, respectively. The figure illustrates that the earnings distribution after 2008 has a clearly different shape than before the reform, implying that the income threshold affects the shape of the whole labor income distribution, not just the region close to the notch point. The differences-in-differences estimate for broader changes in the distribution within an income range of 0–9200 euros is 9.81 (standard error 1.01). The estimate for broader changes among the student population only is 10.97 (1.85), estimated using equation (1). The bunching estimates at the threshold are 2.93 (0.88) before and 1.71 (0.88) after 2008, respectively. A lower limit of 1100 euros below the threshold is used in the bunching estimation both before and after 2008. See Appendix B for a more detailed analysis of bunching responses.

Figure 3: Earnings distributions of students and non-student part-time workers before and after the 2008 reform



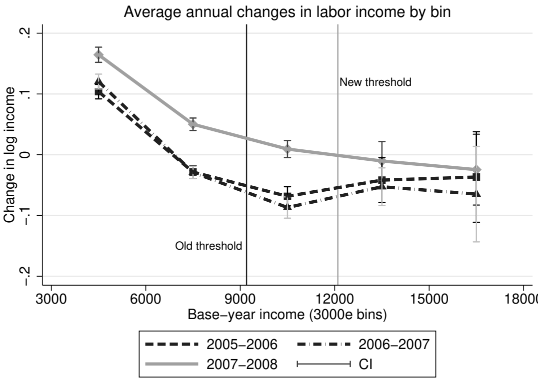
(a) Students



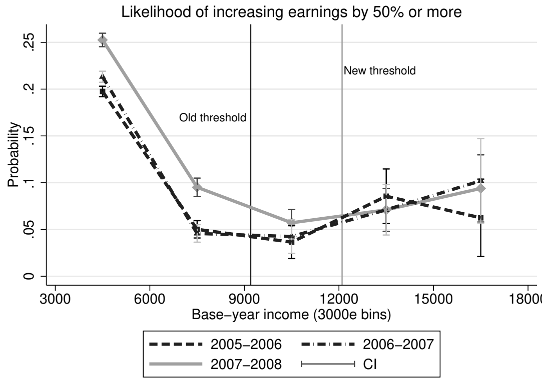
(b) Other young part-time workers

Notes: Figure presents the income distributions of students with 9 subsidy months (graph a) and other young part-time workers (graph b) in 2004–2005 (gray dashed line), 2006–2007 (gray solid line), 2008–2009 (black solid line) and 2010–2011 (black dotted line) within an income range of 0–18,000 euros in bins of 100 euros. The figure shows that the earnings responses of students occurred exactly at the time of the 2008 reform, and that the response is not caused by any gradual changes in the shape of the distribution over time. The distribution for other young part-time workers remained almost unchanged throughout the time period 2004–2011. However, there are similar minor changes at the bottom of distributions of both the treatment and control groups from 2004–2005 to 2006-2007, which further strengthens the case that the distributions develop similarly over time in the absence of the reform.

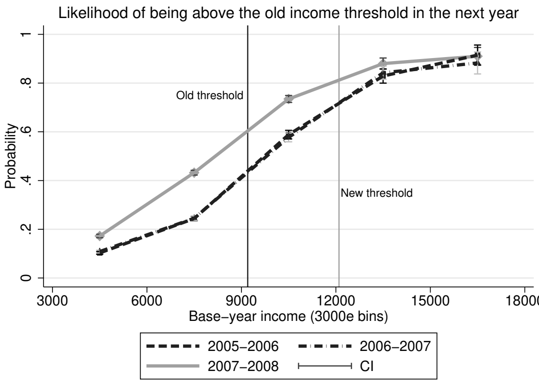
Figure 4: Income distributions of students and other young part-time workers in 2004–2011



(a) Average changes in labor income



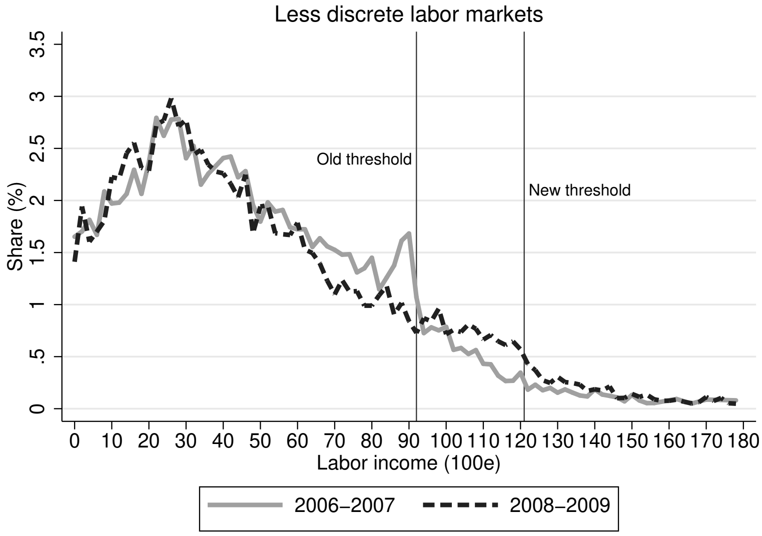
(b) More than 50% earnings increases

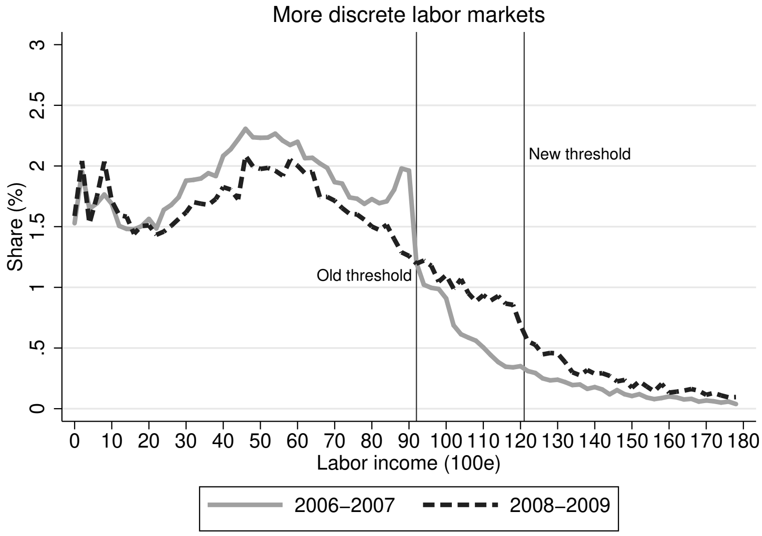


(c) Likelihood of locating above the old income threshold

Notes: Graph (a) presents the average individual-level changes in real log labor income (relative to the 2007 real price index) with 95% standard errors in base-year bins of 3000 euros for students with 9 subsidy months. Gray solid line represents the years 2007–2008, and black dashed lines the pre-reform years 2005–2006 and 2006–2007. The graph shows that earnings increases are more prevalent below the new threshold at the time of the reform compared to previous years, but there are no significant differences above the new income threshold. Graph (b) presents the average likelihood and 95% standard errors for increasing labor income by 50% or more relative to base-year income. The graph illustrates that the likelihood of large income increases is significantly higher below the old threshold at the time of the reform compared to previous years, but there are no significant changes above the old threshold between the years. Graph (c) presents the average likelihood and 95% standard errors for locating above the old income threshold in the next year. The graph shows that this likelihood increased significantly in bins below the new threshold, but there are no significant changes between the years at larger income levels. Overall, these findings support the view that students responded to the relocation of the notch with large intensive-margin earnings increases instead of marginal earnings adjustments along the whole distribution.

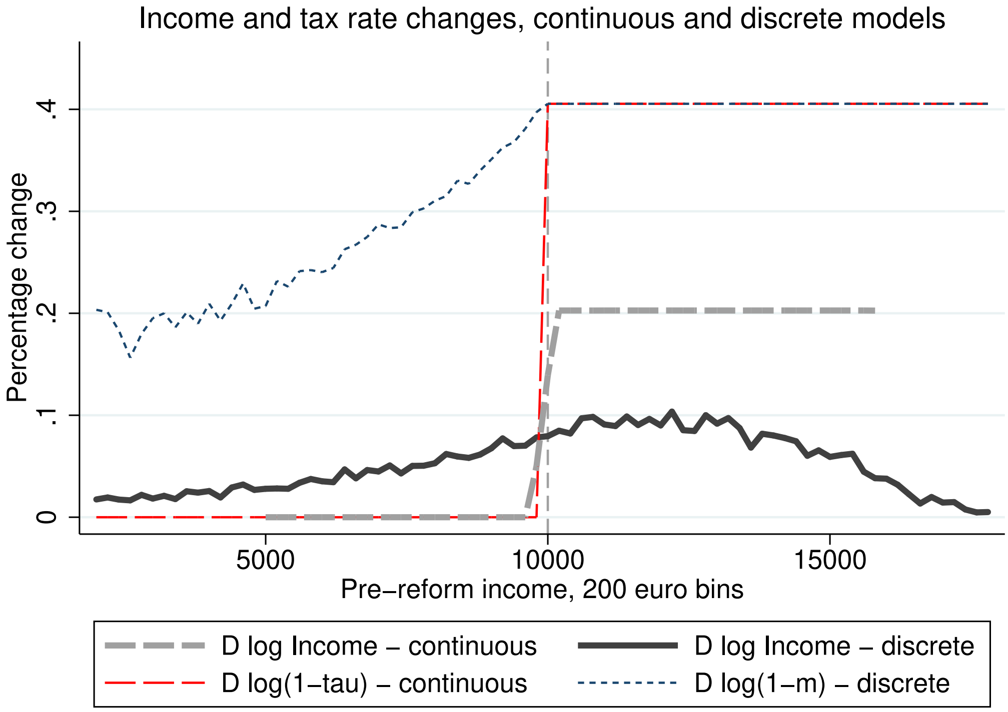
Figure 5: Panel data evidence of individual-level earnings responses





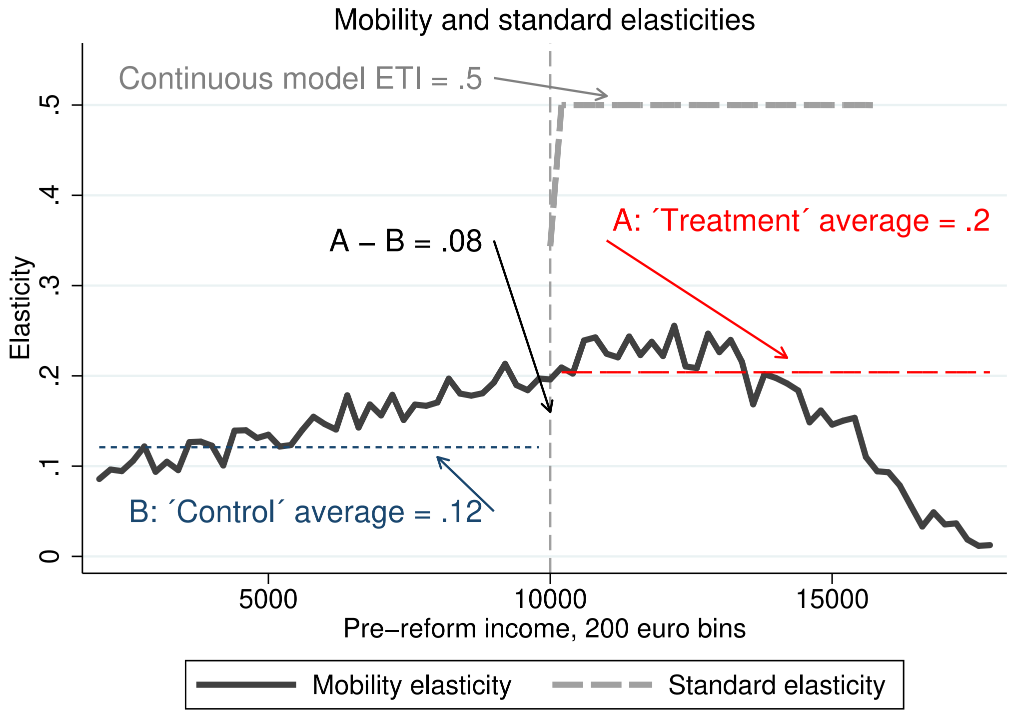
Notes: Figure presents the observed relative earnings distributions before the reform in 2006–2007 (gray solid line) and after the reform in 2008–2009 (black dashed line) within an income range of 0–18,000 euros in bins of 100 euros for students with the default 9 subsidy months in each year working in different types of jobs. Jobs are categorized using firm-level industry classification codes. Less discrete labor markets include restaurants, bars and cafes, cleaning and security services, and retail sales such as supermarkets and gas stations. More discrete labor markets include public sector, and research, manufacturing, construction and maintenance in the private sector. Using equation (1), the estimate for broader changes in the distribution within an income range of 0–9,200 euros for the less discrete group is 6.14(1.71), and for the more discrete group 10.94(1.10), illustrating that broader changes in the distribution are significantly more prevalent for the latter group compared to the first group.

Figure 6: Labor income distributions before and after 2008 for students working in less discrete and more discrete labor markets



Notes: Figure presents the simulated changes in log income and tax rates for the continuous and discrete models after the removal of a tax rate kink at 10,000 euros using an underlying e parameter of 0.5. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink. Horizontal axis denotes pre-reform income in bins of 200 euros. Gray dashed line denotes the change in income in the continuous model, and solid black line in the discrete model with 10 available earnings choices within an income range of 0–25,000 euros. Long-dashed red line denotes the change in the net-of-tax rate in the continuous model, and the short-dashed blue line the average change in the average tax rate between the available earnings locations. Based on the assumptions in the continuous model, only individuals with pre-reform income above the original kink point respond to the reform, as only their net-of-tax rate is assumed to be affected by the reform. In the discrete model, a fraction of individuals in income bins below the kink also respond if their next available earnings location is above the original kink. Consequently, the average tax rates are also affected for those individuals below the kink, but the average change is smaller than in the continuous model above the kink. The average change in tax rate in the discrete model equals the average change in the continuous model above the kink by definition. In the baseline continuous model, all individuals are assumed to respond to the reform with a similar earnings response above the kink. In the discrete model, a smaller share of individuals respond with a discrete jump in earnings depending on the locations of the available earnings choices, constituting a smaller average earnings response compared to the continuous model above the kink.

Figure 7: Changes in log income and net-of-tax rate in continuous and discrete models



Notes: Figure presents the simulated elasticity estimates associated with the responses to the removal of a tax rate kink at 10,000 euros in the discrete and continuous models using an underlying e parameter of 0.5. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink. Horizontal axis denotes pre-reform income in bins of 200 euros. Solid black line denotes the average bin-level mobility elasticity estimates estimated with 10 available discrete earnings choices. Gray dashed line denotes the standard elasticity estimates from the continuous model. The standard continuous average elasticity estimate is 0.5 above the kink, which exactly equals the assumed e parameter in the model. The average mobility elasticity is 0.14, and the average estimate is 0.12 below the kink (B) and 0.2 above it (A), on average. The figure presents an illustration of the bias in the standard differences-in-differences elasticity estimate when using those below the removed kink as a control group. The standard differences-in-differences approach would simply subtract the elasticity (or earnings responses) below the kink from that above it (A–B), producing a downward-biased estimate. However, if the underlying earnings responses are continuous, no such bias emerges when using those below the kink (elasticity=0) as a control group.

Figure 8: Elasticity estimates in discrete and continuous models and the bias in the standard differences-in-differences estimate

Tables

Table 1: Descriptive statistics, all students 1999–2013

	Individual characteristics			
	Age	Female	Labor income	Labor income > 500
Mean	23.8	.58	8446	.80
Median	23	1	6306	1
sd	4.23	.49	8197	.40
N	2,417,517	2,417,517	2,078,538	2,417,517
	One employer	Study subsidy months	9 subsidy months	Years studied
Mean	.55	8.02	.42	2.2
Median	1	9	0	2
sd	.50	2.64	.49	1.80
N	1,863,702	2,417,517	2,417,517	2,098,485
	Field of industry			
	Manufacturing	Hospitality services	Admin. & Public Sector	Other/missing
Mean	.19	.16	.39	.25
sd	.40	.37	.49	.43
N	2,417,517	2,417,517	2,417,517	2,417,517
	Field of study			
	Arts & Humanities	Business & Soc. Science	Tech., Health & Soc. Serv.	Other/missing
Mean	.17	.19	.34	.29
sd	.38	.39	.47	.42
N	2,417,517	2,417,517	2,417,517	2,417,517

Notes: Table presents the descriptive statistics for all students in 1999–2013, excluding first-year students and those who graduate within a given year. Labor income > 500 denotes the share of students with annual labor income above 500 euros. One employer denotes the share of students who we observe to work for only one employer within a year among those with information on the employer in the data. 9 subsidy months denotes the share of students with the default study subsidy choice.

Table 2: Descriptive statistics, non-student part-time workers, 1999–2013

	Individual characteristics				
	Age	Female	Labor income	Labor income > 500	One employer
Mean	21	.56	8318	.93	.62
Median	21	1	6741	1	1
sd	1.710	.496	7229	.25	.48
N	940,786	940,786	932,572	940,786	940,786
	Field of industry				
	Manufacturing	Hospitality services	Admin. & Public Sector	Other/missing	
Mean	.31	.22	.41	.06	
sd	.46	.41	.49	.24	
N	940,786	940,786	940,786	940,786	

Notes: Table presents the descriptive statistics for young, non-student part-time workers used in Figure 3. The group of non-student part-time workers is selected to roughly match students' job and age characteristics. The non-student group comprises individuals who we observe to have less than 12 working months per year in the data, and who are 19–24 years old. The age interval is chosen to match between the 25–75 percentile points of the students' age distribution. Labor income > 500 denotes the share of individuals with annual labor income above 500 euros. One employer denotes the share of individuals who we observe to work for only one employer within a year among those with information on the employer in the data.

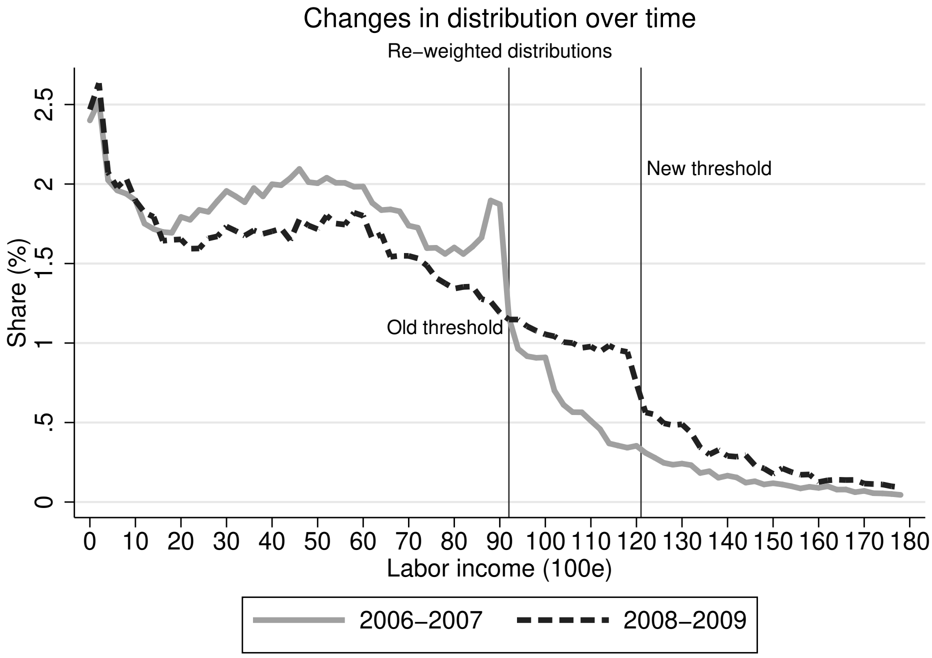
Table 3: Simulated earnings elasticity estimates using different estimation approaches

	(1)	(2)	(3)	(4)
Location of the kink	kink 15k	kink 10k	kink 5k	kink 0
Naive ETI	**-0.043**	**0.07**	**0.051**	-
Standard error	(0.001)	(0.0009)	(0.001)	
Mobility elasticity	**0.083**	**0.161**	**0.294**	**0.551**
Standard error	(0.0005)	(0.0004)	(0.0005)	(0.001)

Notes: Table collects the simulated earnings elasticity estimates using 10 available discrete locations for each individual and an assumed e parameter of 0.5, and varying the location of the original kink. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink. The naive ETI estimates are estimated by regressing the change in the log of earnings on the change in the log of the net-of-tax rate using individuals below the original kink as a control group for those individuals originally located above the kink. In columns (1)–(3) individuals from an income range of 5000 euros below the kink are used as controls. This estimate cannot be measured when the original kink is at zero pre-reform earnings in column (4). The mobility elasticity regresses the observed simulated change in the log income on the change in the net-of-average tax rates between the discrete earnings locations in the same income range but uses a pure control group fully unaffected by the tax reform. Table shows that the naive ETI estimates are downward-biased compared to the underlying unbiased mobility elasticity estimates, stemming from the fact that a fraction of the control group below the kink also increase their earnings after the simulated reform (see Figure 8). Second, the mobility elasticity estimate increases the lower the location of the original kink in the distribution is, as at lower income levels there is, on average, a larger number of available earnings locations above the kink point where individuals can relocate themselves after the reform.

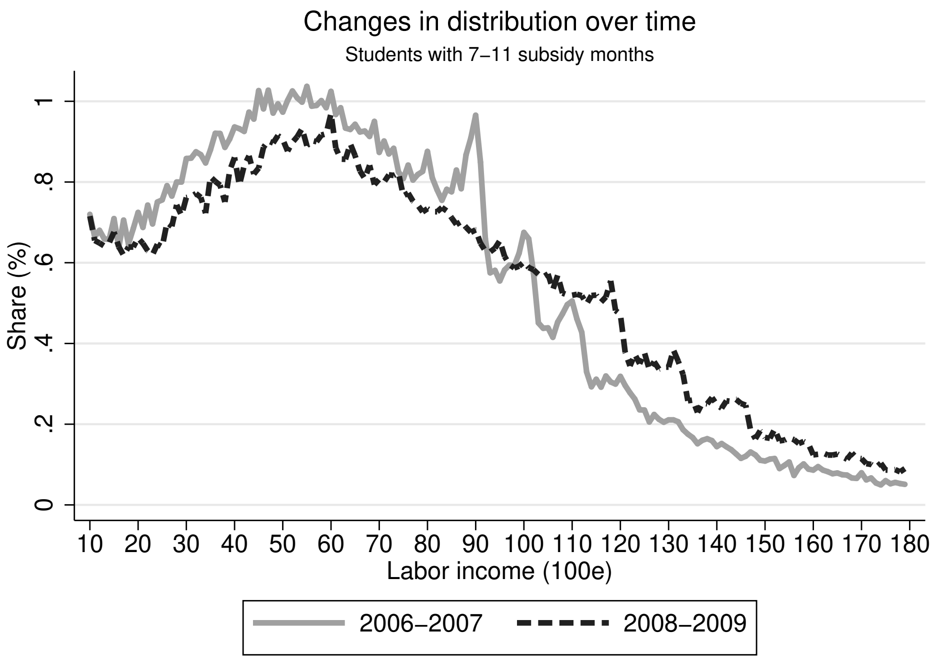
Appendix A

Figures



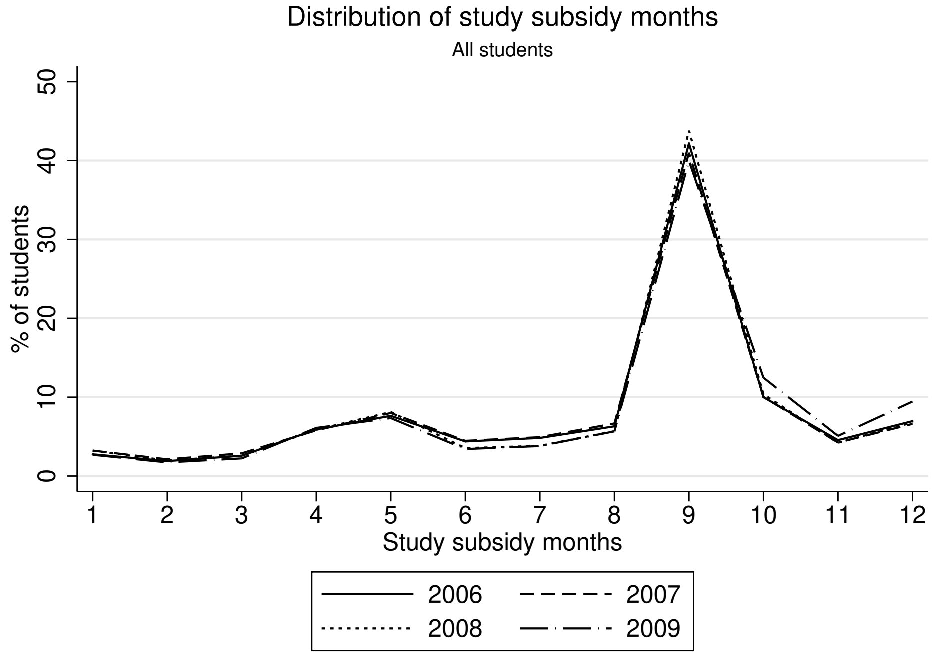
Notes: Figure presents the re-weighted relative earnings distributions before the reform in 2006–2007 (gray solid line) and after the reform in 2008–2009 (black dashed line) within an income range of 0–18,000 euros in bins of 200 euros for students with the default 9 subsidy months in each year. Bin-level inverse probability weighting is used to re-weight the annual distributions using 2006 as the base year. The re-weighting procedure utilizes four groups for both the field of industry and field of study, and three age groups based on age terciles. Using equation (1), the estimate for broader changes in the distribution within an income range of 0–9200 euros is 11.40 (1.01), which is similar to that estimated in the baseline case in Figure 3 in the main text.

Figure A1: Re-weighted income distributions in 2006–2007 and 2008–2009.



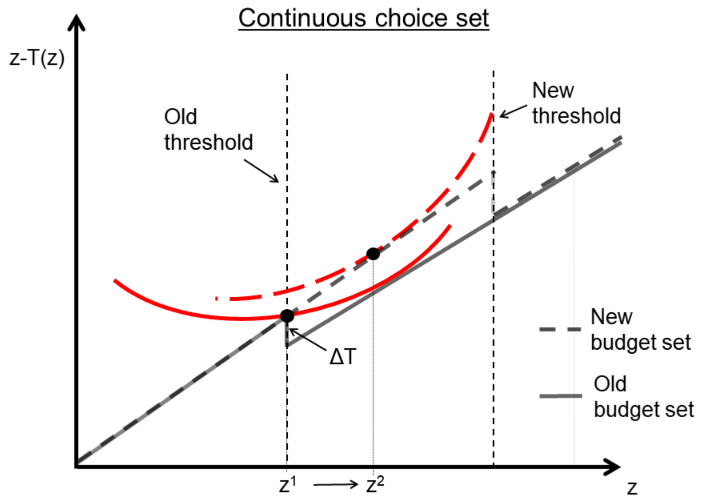
Notes: Figure presents the observed relative earnings distributions before the reform in 2006–2007 (gray solid line) and after the reform in 2008–2009 (black dashed line) within an income range of 0–18,000 euros in bins of 200 euros for students with 7–11 subsidy months. The figure shows that broader changes in the distribution are prevalent when including students who deviate from the default choice of 9 subsidy months. However, as the number of subsidy months defines the location of the income threshold, changes in the distribution are more scattered over the distribution compared to our baseline case with 9 subsidy months in Figure 3 in the main text. Also, a fraction of students who choose other than 9 subsidy months bunch at their associated income thresholds, which appear as additional spikes in the distribution. Relatedly, as the location of the thresholds both before and after 2008 is not constant in this population, we cannot estimate a measure for broader changes in the distribution for this population following the procedures introduced in Section 2.3.

Figure A2: Income distributions in 2006–2007 and 2008–2009, students with 7–11 subsidy months.

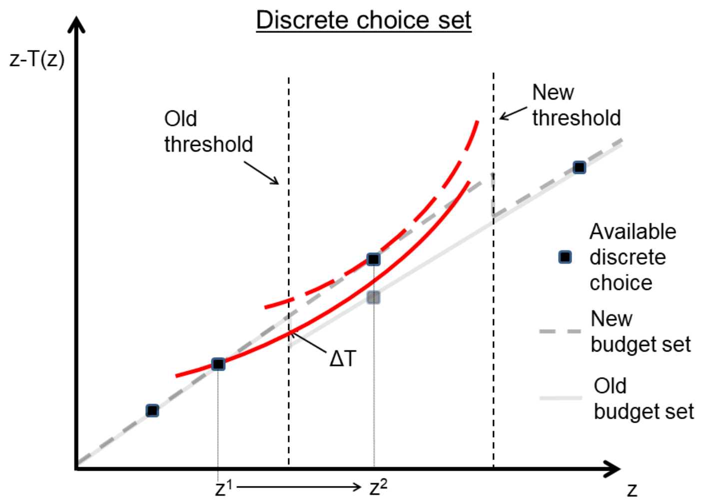


Notes: Figure presents the distribution of study subsidy months in 2006, 2007, 2008 and 2009. In each year, the default 9 months of the subsidy is the most common choice. There are no significant changes in the distribution over time. This indicates that students responded to the reform of 2008 by changing their earnings, but not, on average, by claiming more or less subsidies per year.

Figure A3: Distributions of study subsidy months, 2006–2009.



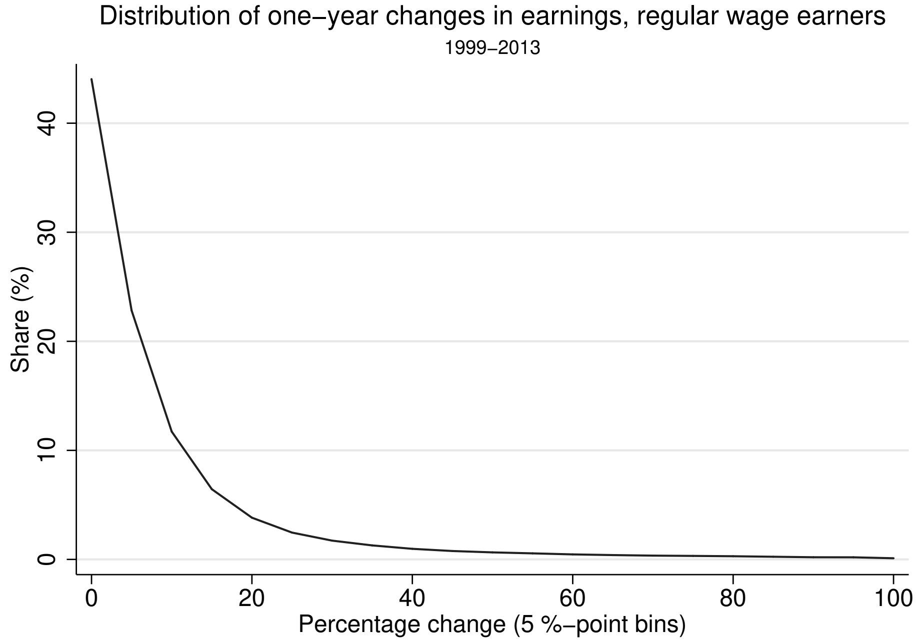
(a) Continuous model



(b) Discrete model

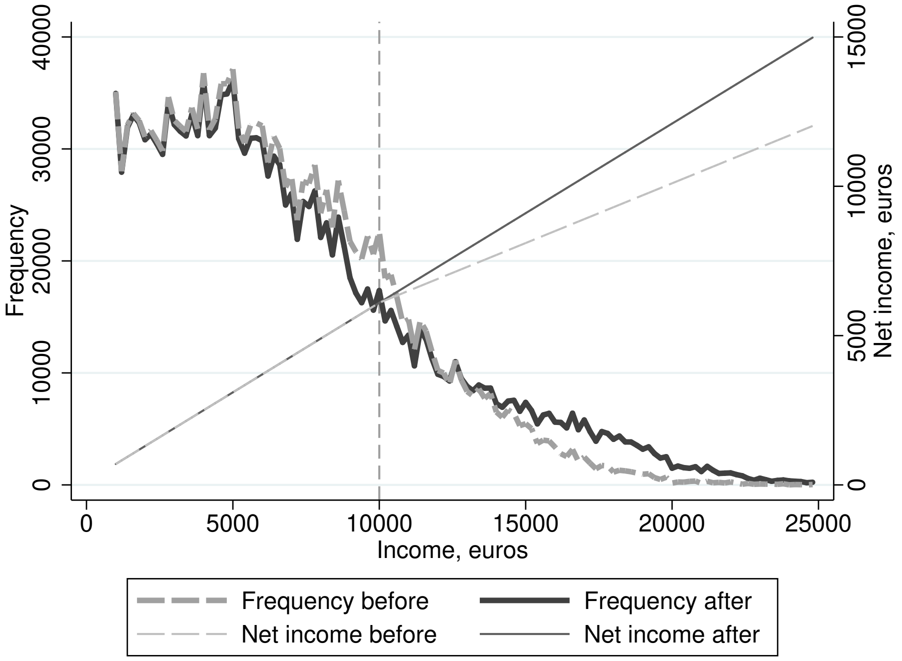
Notes: Figure graphically illustrates the conceptual differences between the continuous and discrete models when an income notch changes its location. The horizontal axis denotes gross earnings (z) and the vertical axis net earnings ($z - T(z)$). In the continuous model in graph (a), the indifference curves are drawn such that an individual would bunch at the original notch, and shifts her location to the right when the location of the notch is increased. By definition, individuals below the original notch are unaffected by the reform.
The discrete model in graph (b) includes the same budget set, but the individual now faces a constraint that only certain discrete earnings locations are feasible, marked with black squares in the figure. Under the old location of the notch, the individual would be located in the first possible earnings level below the notch. When the notch is relocated, the next discrete location above the notch becomes more attractive. The main differences between the continuous and discrete models in the figure are that the earnings response is greater in the discrete model and it also occurs from a region below the original notch.

Figure A4: Relocation of the notch in continuous and discrete models



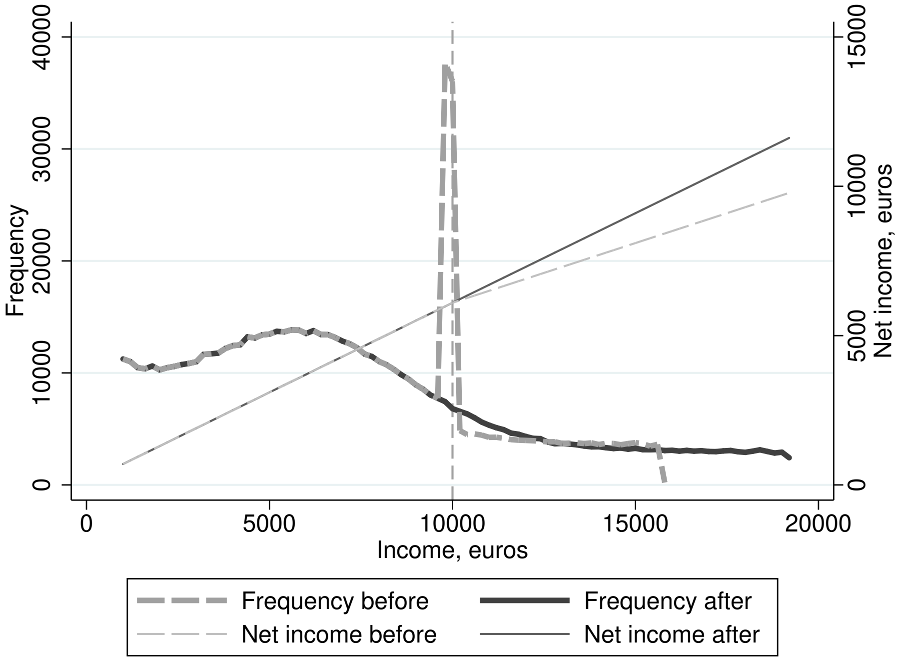
Notes: Figure presents the distribution of one-year changes in wage income in 1999–2013 for wage earners with real wage income (in 2007 terms) of at least 30,000 euros in the base-year. The figure denotes the absolute value of the change, thus including both negative and positive changes in earnings. The figure is restricted to include all changes below 100%, excluding 0.2% of all one-year changes that are larger than that.

Figure A5: The distribution of one-year changes in earnings for regular wage earners



Notes: Figure presents the simulated earnings distributions before (gray dashed line) and after (black solid line) the removal of a tax rate kink at 10,000 euros (dashed vertical line) when assuming a discrete labor supply model. Horizontal axis denotes pre-reform income, and net income denotes the net-of-tax income (right-hand side vertical axis) before and after the removal of the kink. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink. The underlying e parameter of 0.5 and the assumption of 10 available earnings choices within 0–25,000 euros are used in the simulation.

Figure A6: Simulated income distributions with 10 discrete choices before and after the removal of a tax rate kink



Notes: Figure presents the simulated earnings distributions before (gray dashed line) and after (black solid line) the removal of a tax rate kink at 10,000 euros when assuming a continuous labor supply model. Horizontal axis denotes pre-reform income in bins of 200 euros, and net income denotes the net-of-tax income (right-hand side vertical axis) before and after the removal of the kink. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink. The underlying e parameter of 0.5 and the assumption of continuous earnings choices within 0–25,000 euros are used in the simulation.

Figure A7: Simulated income distributions with continuous earnings before and after the removal of a tax rate kink

Tables

Table A1: Income thresholds before and after the 2008 reform

	Before 2008 (academic year 2006/2007)		**After 2008 (academic year 2008/2009)**	
Study subsidy months	Income threshold	Relative income loss at the margin if the threshold is exceeded	Income threshold	Relative income loss at the margin the threshold is exceeded
1	17,340	3.1%	22,550	2.5%
2	16,330	3.2%	21,190	2.7%
3	15,320	3.5%	19,930	2.9%
4	14,310	3.7%	18,620	3.1%
5	13,300	4.0%	17,310	3.3%
6	12,290	4.3%	16,000	3.6%
7	11,280	4.7%	14,690	3.9%
8	10,270	5.2%	13,380	4.3%
9	**9260**	**5.7%**	**12,070**	**4.8%**
10	8250	6.4%	10,760	5.3%
11	7240	7.3%	9450	6.1%
12	6230	8.5%	8140	7.1%

Note: Table presents the annual income thresholds in euros for different subsidy months before and after the 2008 reform. The highlighted 9 months of the subsidy is the default choice. The relative income loss from marginally exceeding the income threshold is calculated using the full study subsidy (461 euros and 500 euros before and after 2008, respectively) plus 15% interest collected by the Social Insurance Institution if the subsidy is reclaimed due to exceeding the income threshold.

Table A2: Variables used in the mobility elasticity estimation for students

	Avg. gross earnings (2000—18,000e)	Avg. net income below notch (2000–9300e)	Avg. net income above notch (9300–18,000)	Differences in net incomes between avg. locations
2007	7116	8693	12,173	3534
2008	7529	8785	13,592	4807
		Gross earnings below: 6008	Gross earnings above: 11,821	

Notes: Table presents the variables used when calculating the mobility elasticity estimate for students in Section 4.4 in the main text. Mobility elasticity is measured by relating the log change in average gross earnings to the log change in the difference in net income between the two average earnings locations below and above the original notch. Net earnings are calculated using the SISU microsimulation model.

Table A3: Simulated earnings elasticity estimates using different discrete earnings choices and the original kink point at 5000 euros

	(1)	(2)	(3)	(4)	(5)
# of discrete choices	5	10	15	20	30
Naive ETI	**-0.133**	**0.051**	**0.172**	**0.244**	**0.324**
Standard error	(0.0015)	(0.001)	(0.0008)	(0.0007)	(0.0005)
Mobility elasticity	**0.259**	**0.294**	**0.30**	**0.295**	**0.272**
Standard error	(0.0007)	(0.0005)	(0.0004)	(0.0003)	(0.0002)

Notes: Table collects the simulated earnings elasticity estimates using different available discrete earnings locations for each individual (5, 10, 15, 20, 30) and e parameter of 0.5. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink located at 5000 euros. The naive ETI estimates are estimated by regressing the change in the log of earnings on the change in the log of the net-of-tax rate using individuals below the original kink in an income range of 5000 euros below the kink as a control group for those individuals originally above the kink. The mobility elasticity regresses the observed simulated change in the log income on the change in the net-of-average tax rates between the discrete earnings locations in the same income range but uses a pure control group unaffected by the tax reform.

Table A4: Simulated earnings elasticity estimates using different discrete earnings choices and the original kink point at 10,000 euros

	(1)	(2)	(3)	(4)	(5)
# of discrete choices	5	10	15	20	30
Naive ETI	**-0.017**	**0.07**	**0.156**	**0.218**	**0.295**
Standard error	(0.0010)	(0.0009)	(0.0008)	(0.0006)	(0.0005)
Mobility elasticity	**0.106**	**0.161**	**0.184**	**0.192**	**0.192**
Standard error	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)

Notes: Table collects the simulated earnings elasticity estimates using different available discrete earnings locations for each individual (5, 10, 15, 20, 30) and e parameter of 0.5. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink located at 10,000 euros. The naive ETI estimates are estimated by regressing the change in the log of earnings on the change in the log of the net-of-tax rate using individuals below the original kink in an income range of 5000 euros below the kink as a control group for those individuals originally above the kink. The mobility elasticity regresses the observed simulated change in the log income on the change in the net-of-average tax rates between the discrete earnings locations in the same income range but uses a pure control group unaffected by the tax reform.

Table A5: Simulated earnings elasticity estimates using different discrete earnings choices and the original kink point at 15,000 euros

	(1)	(2)	(3)	(4)	(5)
# of discrete choices	5	10	15	20	30
Naive ETI	**-0.037**	**-0.043**	**-0.022**	**0.011**	**0.082**
Standard error	(0.0008)	(0.001)	(0.001)	(0.0009)	(0.0008)
Mobility elasticity	**0.044**	**0.083**	**0.108**	**0.122**	**0.136**
Standard error	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0004)

Notes: Table collects the simulated earnings elasticity estimates using different available discrete earnings locations for each individual (5, 10, 15, 20, 30) and e parameter of 0.5. The marginal tax rate below the kink is 40% and 60% above it before the removal of the kink located at 15,000 euros. The naive ETI estimates are estimated by regressing the change in the log of earnings on the change in the log of the net-of-tax rate using individuals below the original kink in an income range of 5000 euros below the kink as a control group for those individuals originally above the kink. The mobility elasticity regresses the observed simulated change in the log income on the change in the net-of-average tax rates between the discrete earnings locations in the same income range but uses a pure control group unaffected by the tax reform.

Appendix B

Bunching estimation

Behavioral responses to local discontinuous changes in the budget set, such as tax rate kinks or notches, are predominantly estimated in the recent literature using a bunching methodology (see Kleven 2016 for a summary). Intuitively, if a discontinuous jump in incentives affects earnings, we should find an excess mass of individuals located just below the threshold in the earnings distribution. The excess bunching thus captures the earnings distortions created by the threshold in the absence of optimization frictions and when earnings choices are continuous. Saez (2010) and Kleven and Waseem (2013) show that under certain restrictions and within the continuous labor supply model, the bunching estimate can be translated into an average earnings elasticity, representing a relevant parameter for the welfare analysis of taxes and income transfers.

We measure local responses to the notch caused by the income threshold following the bunching method presented in Kleven and Waseem (2013). The local counterfactual density is estimated by fitting a flexible polynomial function to the observed distribution, excluding an area around the study subsidy income threshold z^* from the observed income distribution. We group students into income bins of 100 euros and then estimate a counterfactual density by excluding the region $[z_L, z_H]$ around the threshold from the regression:

$$c_j = \sum_{i=0}^{p} \beta_i (z_j)^i + \sum_{i=z_L}^{z_H} \eta_i \cdot \mathbf{1}(z_j = i) + \varepsilon_j \tag{10}$$

where c_j is the count of individuals in bin j, and z_j denotes the income level in bin j. The order of the polynomial is denoted by p. Thus the fitted values for the counterfactual density are given by $\hat{c}_j = \sum_{i=0}^{p} \beta_i (z_j)^i$. The excess bunching is then estimated by relating the actual number of students close to the threshold within (z_L, z^*) to the estimated counterfactual density in the same region:

$$\hat{b}(z^*) = \frac{\sum_{i=z_L}^{z^*} (c_j - \hat{c}_j)}{\sum_{i=z_L}^{z^*} \hat{c}_j / N_j} \tag{11}$$

where N_j is the number of bins within $[z_L, z^*]$.

Following Kleven and Waseem (2013), we set the lower limit of the excluded region (z_L) based on visual observations of the income distribution to represent the point in the distribution where the bunching behavior begins, i.e. when the density begins to increase. We determine z_H such that the estimated excess mass, $\hat{b}_E(z^*) = (\sum_{i=z_L}^{z^*} c_j - \hat{c}_j)$, equals the estimated missing mass above the threshold, $\hat{b}_M(z^*) = (\sum_{i=z>z^*}^{z_H} \hat{c}_j - c_j)$, stemming from individuals who would locate above the income threshold in the absence of it and who respond to the notch by bunching below it, illustrated in Figure 2 in the main text. We apply this convergence condition by starting from a small value of z_H and increasing it gradually until $\hat{b}_E(z^*) \approx \hat{b}_M(z^*)$. This convergence condition also defines the marginal buncher student with income $z^* + \triangle z$, representing the student with highest earnings in the absence of the notch who responds by locating below the income threshold.

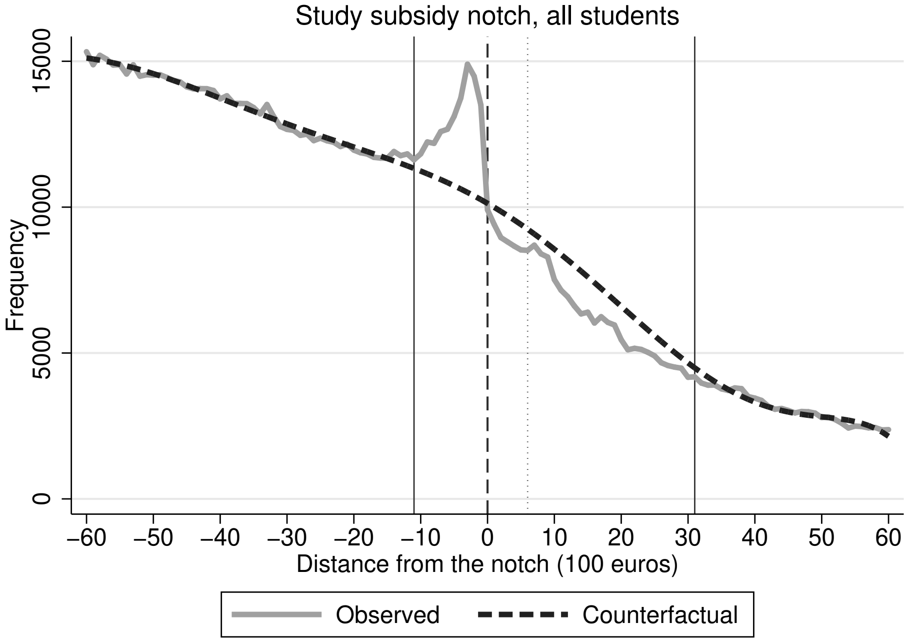
Following Kleven and Waseem (2013), we calculate standard errors by using a residual-based bootstrap procedure. We generate a large number of income distributions by randomly resampling the residuals from equation (10) with replacement, and generate a large number of new estimates of the counterfactual density based on the resampled distributions. Based on the bootstrapped counterfactual densities, we evaluate variation in the bunching estimate. The standard error is defined as the standard deviation in the distribution of the estimate.

Bunching responses

We find clear local responses to the income threshold of the study subsidy program. Figure B1 presents the gross income distribution and the counterfactual distribution relative to the notch in bins of 100 euros in the range of +/- 6000 euros from the notch in 1999–2013. The dashed vertical line denotes the notch point above which a student loses one month of the subsidy. The solid vertical lines denote the excluded range used in the estimation of the counterfactual, which is estimated using a 7th-order polynomial function. The dash-point vertical line above the notch shows the upper limit for the dominated region just above the notch where students can increase their net income by lowering their gross income subject to the income threshold.

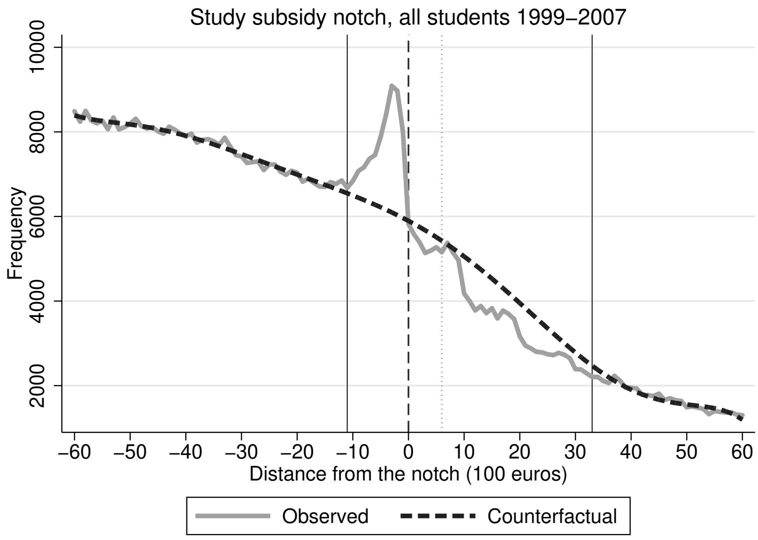
Figure B1 indicates a visually clear and statistically significant excess mass (2.19 (*0.189*)) below the income threshold for all students (standard error in parenthesis). This implies that students are both aware of the notch and respond to the strong local incentives created by it. In addition, there is clear evidence of the existence of some types of frictions. There is an economically and statistically significant mass of students, 0.915(*.027*) of the mass relative to the counterfactual, at the locally dominated region just above the notch where no students should locate in the absence of any types of frictions or constraints and when earnings choices are continuous (Kleven and Waseem 2013). Furthermore, even though the study subsidy schedule ultimately consists of multiple notches, we observe a distinctive response only to the first income threshold they face.

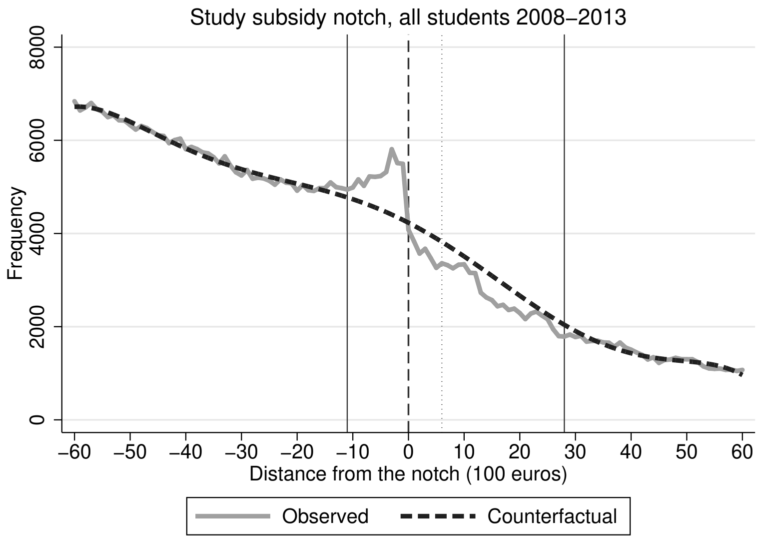
Figure B2 shows the bunching responses before (1999–2007) and after (2008–2013) the 2008 reform. The figure shows that excess bunching is slightly larger before (2.55 (*0.228*)) than after (1.71 (*0.882*)) the reform. One explanation for this is that the incentives not to exceed the notch are somewhat smaller after 2008, since the relative significance of losing one month's subsidy in terms of disposable income is now smaller than before 2008 when the threshold was at a lower income level. However, as discussed in Section 2 in the main text, this standard bunching method is not a valid measure for estimating labor supply responses to tax incentives under the discrete labor supply model, and therefore these estimates need to be interpreted as suggestive.



Notes: Figure presents the observed earnings distribution (gray solid line) and the estimated counterfactual distribution (black dashed line) around the income threshold (denoted by zero in the figure) in bins of 100 euros for all students using pooled data from 1999–2013. The first and second solid vertical lines denote the lower and upper limits of the excluded region when estimating the counterfactual distribution. The counterfactual is estimated using a seventh-order polynomial. The dotted vertical line denotes the upper limit of the region of dominated choice just above the threshold. The estimate for excess bunching at the notch is 2.19 (0.189), and the estimate for the mass at the dominate region is 0.915 (0.027).

Figure B1: Bunching at the study subsidy notch, 1999–2013





Notes: Figure presents the observed earnings distributions (gray solid line) and the estimated counterfactual distributions (black dashed line) around the income threshold (denoted by zero in the figure) in bins of 100 euros for all students before (1999–2007) and after (2008–2013) the 2008 threshold reform. The first and second solid vertical lines in the figure denote the lower and upper limits of the excluded region when estimating the counterfactual distribution. The counterfactual is estimated using a seventh-order polynomial. The dotted vertical line denotes the upper limit of the region of dominated choice just above the threshold. The estimate for excess bunching at the notch before 2008 is 2.55(0.228) and 1.71(0.882) after the reform.

Figure B2: Bunching at the study subsidy notch: Before and after the 2008 reform

Earnings elasticity estimates

We approximate the earnings elasticity at the study subsidy notch using a similar approach as Kleven and Waseem (2013). We derive an upper-bound reduced-form earnings elasticity by relating the earnings response of a marginal buncher student at z^H to the implicit change in tax liability between the notch point z^* and z^H (see Figure 2 in the main text). The marginal buncher represents the individual with the highest income to move to the notch point, compared to a counterfactual state in the absence of the notch. Intuitively, this approach treats the notch as a hypothetical kink which creates a jump in the implied marginal tax rate. More formally, the reduced-form earnings elasticity is calculated with a quadratic formula

$$e(z^*) \approx (z^H/z^*)^2/(\triangle t/(1-t)) \tag{12}$$

where $(1-t)$ is the net-of-tax rate at the notch, and $\triangle t$ defines the change in the implied marginal tax rate for the marginal buncher (Kleven and Waseem 2013). We include all the income tax and benefit rules and use the SISU microsimulation model to calculate the implied marginal tax rates for the students in the estimation.

The implied earnings elasticities are 0.083 (*0.019*) for all students and 0.065 (*0.007*) for students with 9 subsidy months. Nevertheless, as discussed above, the bunching method does not capture all earnings responses when the earnings choices are discrete, and therefore these estimates do not represent the true earnings elasticity of students.

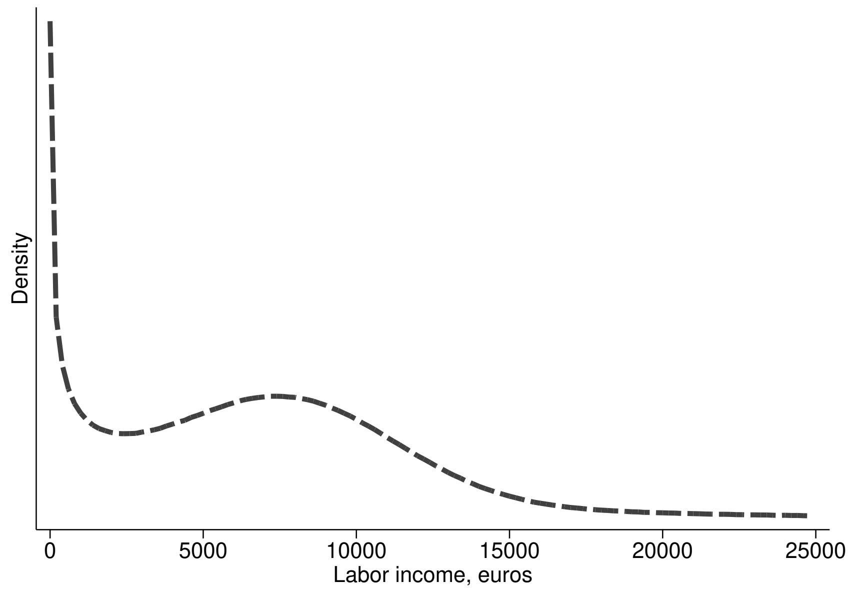
Appendix C

In this Appendix we first present our simulation model we also use in Section 5 in the main text and then present and discuss simulation results that illustrate the mechanisms behind the income responses of students under different modeling assumptions. The results highlight that a discrete model as opposed to any variant of a continuous model qualitatively matches our empirical findings in Section 3 in the main text.

Simulation model

We build our simulation model on the theoretical framework presented in Section 4 in the main text. The individual utility function is given in equation (3), where the e parameter governs the disutility from earnings supply and would correspond to the elasticity with respect to taxes in the continuous model. The discrete model has the same utility function but adds a fixed number of discrete earnings choices to the individual decision problem as an additional constraint. The budget set for individuals arises from the tax system included in the simulations which we discuss below.

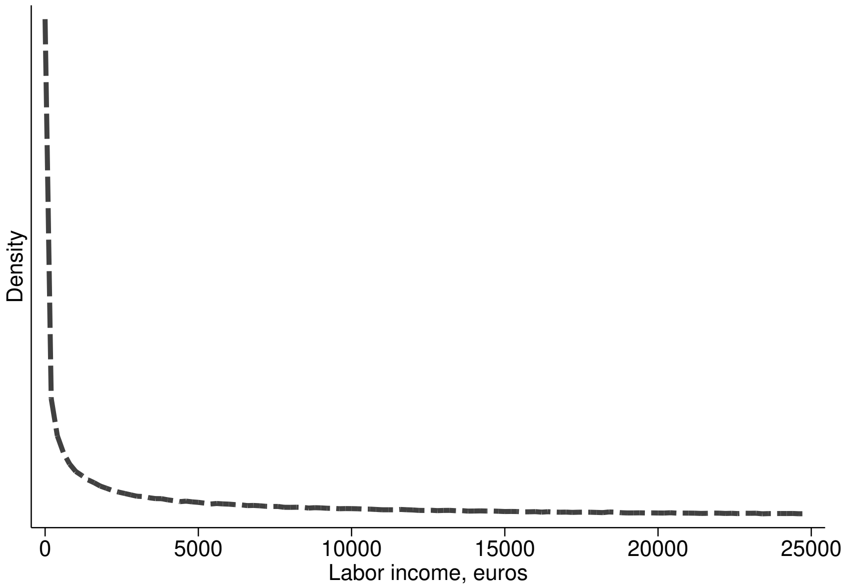
The model assumes an underlying ability distribution from which each individual i receives a predetermined draw w_i. This draw represents earnings in the absence of responses to the tax system. Our parameterized ability distribution is presented in Figure C1. The distribution is a combination of power distributions and normal distributions, which gives an approximate match for the shape of the empirical earnings distribution of students in our empirical case. Our results are not sensitive to different underlying ability distributions that roughly match the empirical income distribution of students.



Notes: Figure presents the underlying earnings distribution used in the simulation model. The distribution is a combination of a power distribution and a normal distribution, which delivers an approximate match for the shape of the empirical earnings distribution of students in our empirical analysis. The simulation results are not sensitive to different underlying ability distributions that roughly match the empirical income distribution of students.

Figure C1: Simulated earnings distribution in the absence of taxes

The available discrete earnings locations for each individual are drawn from the probability distribution presented in Figure C2. The number of choices drawn can be altered in different specifications and the draws vary between different individuals. Therefore, even when the individual-level choices are discrete, the overall earnings distribution is smooth.



Notes: Figure presents the underlying probability distribution of discrete earnings choices utilized in the discrete choice model simulations. The large mass in the probability distribution at small earnings ensures that each individual has at least one available choice that produces positive utility with positive earnings. The thick tail in the distribution ensures that there is another available choice at a higher income level, although the specific location of this choice can vary across different draws. In the simulation procedure, we iterate the model multiple times, and in each round draw new available earnings choices. The resulting earnings distribution for the full population is continuous, although one individual faces only a discrete and limited number of available choices.

Figure C2: Probability distribution of discrete earnings choices

Discrete vs. continuous model

We simulate earnings distributions when the budget set includes an income notch, a scenario that resembles our empirical case for university students. We focus on discrete labor supply conditional on participating in the labor market, and for simplicity, exclude the labor market participation margin from the analysis.

We assume the parameters given in Table C1 below. The marginal income tax rate is set at 22% below the notch and a high marginal tax rate of 61% is applied above the notch, constituting a simplified linear version of the actual budget set for students including multiple notches above the income threshold (see Figure 1 in the main text). The size of the notch, i.e. the size of the drop in disposable income at the income threshold, is 500 euros. The notch is relocated from 9000 to 12,000 euros in the simulated reform.

Table C1: Parameter values in the simulation model for the income threshold reform

Parameter	Value
Marginal tax rate (τ)	
Below the notch	0.22
Above the notch	0.61
Size of the notch	500e
Virtual income (R)	
Before	4100e
After	3600e
Location of the notch (income threshold)	
Before	9000e
After	12,000e

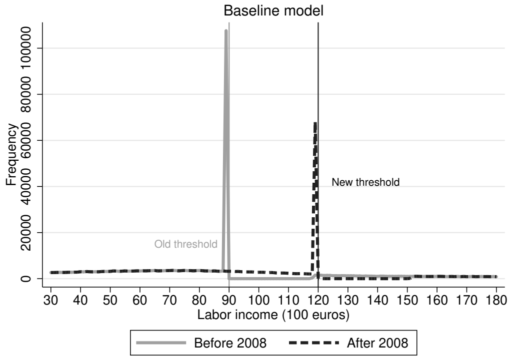
Notes: Table presents the parameter values used in the simulation model. The parameter values are selected to approximate the actual budget set faced by students under the study subsidy program (see Figure 1 in the main text).

Figure C3 shows the simulated earnings distributions within an income interval of 0–25,000 euros using the baseline continuous model and 0.2 as the value for the e parameter.[29] The different panels in the figure present the baseline continuous model (panel a), the continuous model supplemented with adjustment frictions (panel b) and with both adjustment friction and optimization errors (panel c). We assume heterogeneous adjustment frictions represented by a uniformly distributed parameter a in the unit interval. Each individual has a different and independent draw from this distribution. The earnings shocks related to optimization errors are normally distributed mean-zero income shocks with a standard deviation of 800 euros.

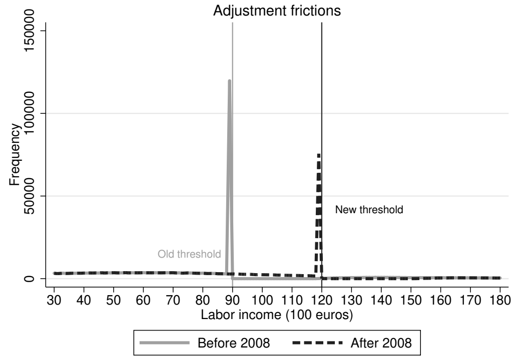
First, we find that bunching at the income threshold is sharp and sizable both before and after the reform in panel (a). Adding adjustment frictions to the model according to equation (4) in the main text in panel (b) only leads to slightly smaller bunching and slightly more individuals being located just above the notch, but no changes in earnings over a broader income range. Adding i.i.d. earnings shocks as optimization errors in panel (c) yields more diffuse bunching, but again no earnings responses over a wider income range below the old income threshold.[30]

[29] Qualitative simulation results are not sensitive to the choice of the e parameter, except that with higher parameter values the densities above the thresholds reduce.

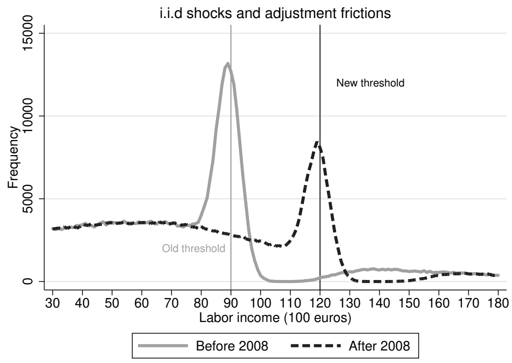
[30] If we were to assume only negative income shocks we would obtain diffuse bunching only below the notch, similarly as in the empirical distribution. However, such asymmetric shocks cannot be not easily justified.



(a) Baseline continuous model



(b) Adjustment frictions

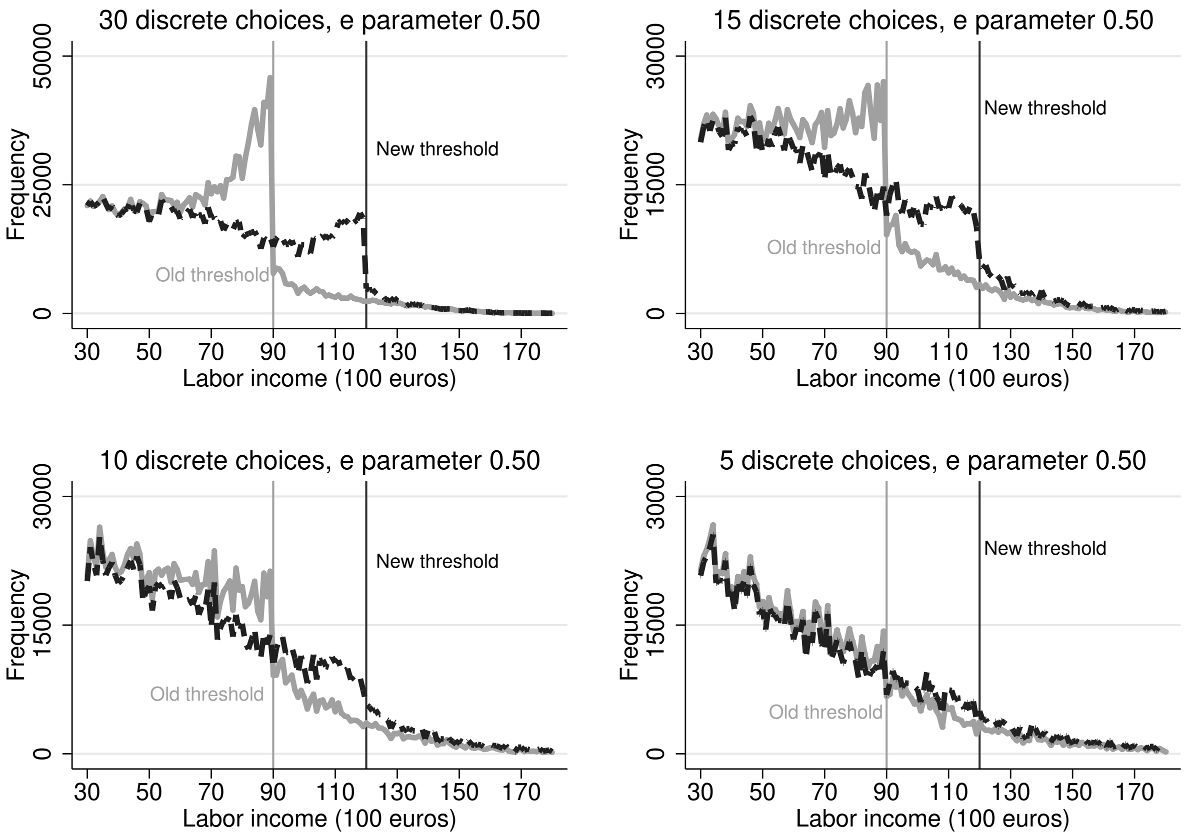


(c) Earnings shocks and adjustment frictions

Notes: Figure presents simulated income distributions before (gray solid line) and after (black dashed line) an increase in the location of the income threshold from 9000 euros to 12,000 euros within an income range of 0–18,000 euros. The underlying e parameter of 0.2 is used in the simulations. Qualitative results are not sensitive to the choice of this parameter value, except that with higher parameter values the densities above the thresholds reduce. Graph (a) presents the standard continuous-choice model. Graph (b) presents the standard model with adjustment frictions and graph (c) includes both adjustment frictions and unexpected i.i.d shocks in earnings to the standard model. The graphs illustrate that these frictions typically discussed in the literature can induce mitigated and scattered bunching around the threshold, but they do not produce broader changes in the earnings distributions we observed in Figure 3.

Figure C3: Simulated income distributions in the baseline continuous model and with different types of adjustment frictions

Next, we add discrete choices as an additional constraint. Figure C4 illustrates the earnings distributions using 30, 15, 10 and 5 available earnings choices for each individual and assuming an underlying e parameter of 0.5. With a discrete choice constraint included in the model, the earnings distributions and the response to the reform qualitatively resemble the empirical distributions (see Figure 3 in the main text). In particular, when the number of discrete choices is set at 10–15, the qualitative shape of the distributions largely resembles their empirical counterparts. First, the earnings distribution shifts to the right from a relatively wide income range below the old threshold. Second, the shape and amount of excess bunching below the threshold are approximately of the same nature and order of magnitude as in our empirical case. Note that scattered bunching below the threshold only results from including discrete earnings choices as a constraint, as this model does not include optimization errors or adjustment frictions. When increasing the number of available discrete locations the distributions begin to resemble those in the continuous model, and with only a very small number of available choices both the broader changes in the distribution and the bunching responses are small, as can be expected.



Notes: Figure presents simulated earnings distributions before (gray solid line) and after (black dashed line) an increase in the location of the income threshold from 9000 euros to 12,000 euros within an income range of 0–18,000 euros using different options for an available discrete earnings choice set. The underlying e parameter of 0.5 is used in the simulations. Using 30 location choices produces distinctive bunching at the threshold, and limited changes in the distribution at lower income levels. In contrast, using 15 or 10 choices produces more limited bunching and more prevalent responses at lower income levels, in a qualitatively similar manner as in Figure 3 in the main text. However, using only 5 available choices reduces both local responses and broader changes in the distribution, which is inconsistent with our empirical observations.

Figure C4: Simulated income distributions with different discrete choice sets

Overall, these findings further support the conclusion in Section 4.2 in the main text that the canonical continuous-choice model cannot explain our empirical findings, not even when complemented with adjustment frictions or optimization errors. In contrast, we need a discrete choice component in the model in order to be able to at least qualitatively explain the earnings responses from far below the old income threshold. Nevertheless, other elements such as optimization errors and inattention could also play a role in the observed responses of students, which might be needed in order to more precisely model the labor supply behavior. However, the main point of this simulation exercise is to highlight that a discrete component is required in order to even qualitatively model the discrete jumps in earnings from below the threshold.