Bunching at Institutional Thresholds: The Cost of Standardization in Higher Education

Raphael Brade[1] , Oliver Himmler[2] , Robert Jäckle[3] , Zouhier Kassaballi[2]

1 ifo Institute, University of Munich, and CESifo, Munich, Germany

2 University of Erfurt, Erfurt, Germany

3 Nuremberg Institute of Technology, Nuremberg, Germany

Abstract. We examine how institutional frameworks shape credit accumulation in higher education, focusing on the European Credit Transfer and Accumulation System. Although designed for a "typical student," the 30-credit benchmark inadvertently functions as a binding constraint for others, creating a salient reference point and a cost kink that makes exceeding the standard load disproportionately difficult. Using administrative and national panel data, we document substantial bunching at this threshold. We identify 40 percent of students earning exactly 30 credits as marginal bunchers, a concentration which implies that institutional constraints hold back 12-15 percent of the total student population from exceeding the standard load. This distortion reflects institutional design rather than ability: bunching intensifies among high-ability students, credit distributions shift with reference points, and acceleration rises when structural frictions are relaxed. Finally, the null result of an informational intervention rules out awareness frictions, suggesting that unlocking student potential requires structural reforms rather than marginal nudges.

JEL: I23, D90, C93

Keywords: Higher Education, Bunching, Reference Dependence

1. Introduction

Higher education systems, whether in Europe or beyond, typically follow a standardized progression plan in which all students are expected to follow the same timeline for timely graduation. These timelines are designed for a "typical" or "average" student and often leave little room for flexibility, particularly for those who could progress faster. In Europe, study programs are structured according to the European Credit Transfer and Accumulation System (ECTS), which defines a full-time semester as 30 credit points (CP), with each CP corresponding to 25 to 30 hours of student work (European Commission, 2015). This system is designed around the expected workload of a “typical” student and serves as a standardized reference for curriculum design and study progression. However, students differ widely in their abilities, motivation, and study preferences. As a result, the effort required to achieve learning outcomes and the capacity to do so vary substantially across individuals, suggesting that credit accumulation should also vary accordingly. In the absence of institutional constraints, this variation would likely yield a somewhat smooth, bell-shaped distribution of earned CP centered around 30, reflecting the normal distribution of abilities and learning preferences within the student population (Johnson et al., 2008).

In practice, the earned credit distribution deviates substantially from a normal distribution due to the standardized design, which is built around a single reference point (RP) of 30 ECTS per semester. This benchmark not only determines how universities structure course offerings, the supply side of credits, but also shapes student decision-making regarding earned credits. Institutional barriers such as limited course availability, scheduling conflicts, and the absence of structured fast-track plans constrain students’ ability and willingness to deviate from the standard pace. As a result, the recommended workload can operate as a binding institutional ceiling, generating a kink in the credit-earning cost function at 30 CP per semester and inducing bunching at this threshold.

This study primarily analyzes the consequences of an institutional RP, such as that of the ECTS, on students' credit attainment behavior and ultimately on student progress through higher education. We analyze the distribution of earned credits per semester in German higher education institutions and test for bunching and missing mass in the distribution at the 30-ECTS threshold. Moreover, we investigate the mechanisms that could lead students to bunch. Finally, a field experiment is conducted to assess the effectiveness of a subtle intervention aimed at encouraging students to exceed their study program’s study plan.

Pursuing a higher-than-average credit load per semester has clear benefits. It allows students to reduce the total duration of their studies, enter the labor market earlier, and signal their high ability to potential employers, thereby improving career prospects and lowering opportunity costs (Aina & Casalone, 2020; Witteveen & Attewell, 2021). Despite these advantages, early graduation remains extremely rare. In Germany, for instance, only 0.03% of students complete their degrees ahead of the official timeline (Statistisches Bundesamt, 2021). This raises the question: To what extent does the institutional setup limit students’ credit accumulation?

Using administrative data from eight cohorts at a German university, supplemented by a nationally representative data set, we find clear evidence of significant bunching in the number of earned CP per semester at the institutional 30-ECTS kink point. Utilizing a bunching estimator, we find significant excess mass at the kink point, with around 40% of the students taking 30 CP considered marginal bunchers. Across both data sources, we find that approximately 13% to 15% of the total student population is constrained by the threshold, reducing their credit loads in response to the 30 CP reference point. For instance, in the administrative data, only about 2% of first-year students take one course beyond the standard plan, far below the share predicted by the counterfactual distribution. The sharp spike in students earning exactly 30 CP, followed by a drop, reflects bunching that deviates from the distribution one would expect if credit loads were shaped only by individual differences in ability, motivation, or preferences. This pattern suggests that the system may unintentionally constrain students who wish to exceed the standard load, raising concerns about untapped academic potential.

We propose two mechanisms through which the 30-credit-per-semester threshold influences student credit accumulation. First, institutional constraints emerge as universities align course schedules, resources, and offerings with the 30-credit benchmark, often neglecting structured pathways for accelerated completion. As a result, students aiming to exceed 30 credits face higher search costs, scheduling conflicts, and reduced course availability. These frictions discontinuously raise the marginal cost of accumulating additional credits beyond the kink point, effectively creating a kink in the credit-cost function. The resulting behavioral response aligns with the kink-based optimization framework developed by Saez (2010), which predicts bunching where marginal costs or incentives change discretely. Second, student behavior may be influenced by psychological responses to the 30-credit RP. According to reference-dependent utility theory (Kahneman & Tversky, 1979; Meng, 2019), students view this benchmark as a salient goal—gaining substantial utility from reaching it, but perceiving little additional benefit from exceeding it. Kleven (2016) further argues that when such a salient reference point coincides with a kink in the cost function, as is plausibly the case in higher education, the combined effect of these mechanisms intensifies the bunching effect.

These mechanisms do not arise in isolation but are primarily shaped by institutional design. While other factors may influence credit accumulation, we provide a robust body of evidence pointing to the ECTS institutional setup being the main factor leading to both bunching behavior and the observed missing mass in credit attainment:

1) In some cases where a study program deviates from the 30 CP-per-semester norm and explicitly advises students to exceed it, we observe a rightward shift in the distribution of

earned credits. This suggests that the 30 CP threshold is not an inherent upper limit for students; higher credit loads are attainable when the institutional setup supports them. In these cases, institutions not only shift the reference point but also provide a higher-quality institutional environment by structurally supporting students through flexible course structures, coordinated planning, and targeted advising, which enables students to exceed the standard pace. This interpretation is consistent with Small and Winship's (2007) broader argument that institutional quality plays a central role in shaping student outcomes.

2) We analyze the behavior of "off-track" students, those who earn fewer credits in their first academic year than the amount specified in their program's credit plan. As a result, these students face relaxed institutional requirements of 30 CP in the following academic year. They have the opportunity to compensate for missed credits by incorporating the outstanding or failed credits into their current study plan, allowing them to recover missed credits while simultaneously maintaining the standard semester load. This added flexibility reduces uncertainty and improves access to credit-earning beyond the standard curriculum, thereby lowering the marginal cost of exceeding 30 credits. Analyzing this group allows us to assess how reduced institutional constraints affect credit accumulation and lessen bunching at the 30-credit threshold. Compared to "on-track" students, we find that off-track students bunch less at the RP and are up to 15 percentage points more likely to exceed 30 CP per semester. Notably, their second-year earned credit distribution exhibits a shift, with the peak in the distribution shifting beyond the 30-credit threshold, while maintaining a similar GPA trend as their on-track peers.

3) Our analysis demonstrates that the observed bunching around the 30 CP threshold is not attributable to students' inherent inability to exceed this limit. Using high school GPA as a proxy for academic ability, we observe an approximately normal distribution of ability within both samples. Notably, the extent of bunching at the threshold intensifies among students with higher academic ability. This pattern indicates that higher-achieving students, despite their greater capacity to surpass the 30 CP limit, encounter institutional or structural constraints that inhibit them from doing so. This produces a ceiling effect that concentrates outcomes of higher-ability students at the institutional reference point rather than allowing performance to spread into the upper tail of the credit distribution. For example, among the top 5% of students by high school GPA, approximately 55% earned exactly 30 CP, while only 1.5% accumulated enough credits to exceed their RP by the equivalent of one additional course. In contrast, among

students near the median of the ability distribution, around 20% earned exactly 30 CP, and just 0.5% earned enough credits to exceed it.

Finally, we find that this phenomenon is not limited to the ECTS. We show a conceptually similar bunching pattern in the US higher education system, which is consistent with the distinct institutional setup and incentives students experience there. In the American system, students face two salient reference points: 12 credit hours, which define full-time status, and 15 credit hours, the standard pace for on-time graduation. Using data from a large U.S. university, we observe a visually pronounced peak in attempted credits at the 15-credit benchmark, followed by a sharp drop, with 4–6% of students exceeding the 15-credit threshold, a pattern strikingly similar to our European findings. This suggests that salient institutional thresholds shape student behavior across diverse higher education systems.

This rigid 30 ECTS setup imposes substantial individual and societal costs. Back-of-the-envelope calculations suggest that delays in graduation due to course load bunching result in approximately €250 million in potential lost earnings annually in Germany. Moreover, the system indirectly prolongs the time to degree by limiting students' ability to build a credit buffer, making them more vulnerable to setbacks and extending delays to a broader group of students[1].

This raises the question of how course load bunching can be reduced. One easily implementable approach is to override the reference point by encouraging students to aim beyond the 30 CP threshold. To test this, we conducted a randomized experiment in which students received information encouraging them to exceed 30 CP. We find no effect on exceeding the 30-CP threshold, suggesting that such low-intensity interventions may be insufficient. Instead, our findings point to the need for more embedded structural measures, such as flexible study plans with multiple reference points or formal fast-track study plans, that better align institutional frameworks with desired student behaviors.

Contribution to the literature. This study contributes to the literature on institutional influences on student learning outcomes by proposing a mechanism through which the structure of academic programs shapes student progress. Prior research has primarily focused on resource constraints, such as increasing student-faculty ratios and limited course availability, as key factors delaying graduation (Bound & Turner, 2007; Huntington-Klein &

[1] In Germany, the average time to degree is 8 semesters, around 43% need an extra year to graduate (Statistisches Bundesamt, 2021).

Gill, 2021). In contrast, we show that institutional design itself, independent of resource limitations, significantly influences student behavior and academic trajectories. Recent studies have highlighted the role of institutional structures in influencing student outcomes. For instance, Bostwick et al. (2022) find that switching from a quarter to a semester system disrupts study progress, leading to lower graduation rates. Moreover, Levin et al. (2013) show that structured fast-track programs in community colleges can accelerate degree completion by removing institutional barriers, yielding a high benefit-to-cost ratio. Our work extends these findings by illustrating how the internal architecture of academic programs, even without resource constraints, elicits behavioral responses that shape student progress.

Second, this study contributes to the economics of education literature by examining how institutional structures affect credit accumulation among high-ability students in higher education. While much research focuses on low- and average-ability students through informational or financial interventions (see Sneyers & De Witte 2018 for a meta-analysis), the constraints faced by high-ability learners remain underexplored in university settings. In contrast, primary and secondary education research has increasingly addressed effective strategies for high-ability (or gifted) learners; see García-Martínez et al. (2021) for a review. A meta-analysis by Kim (2016) finds that enrichment programs in a school setting significantly boost academic achievement and motivation for gifted students, whereas standardized curricula often lead to disengagement and reduced progress due to insufficient challenge. Similarly, Hornstra et al. (2022) show that intrinsic motivation among gifted students declines over time without targeted support. Despite these insights, research in higher education rarely prioritizes high-ability students' progress. This study addresses this gap by analyzing how university program structures shape credit outcomes for high-performing students.

Third, this research further adds a policy-relevant application to the general bunching literature. The bunching approach has been gaining importance as an empirical approach in economics. It has been adopted in a range of settings, such as labor supply responses to tax kinks (Chetty et al., 2011; Saez, 2010), income manipulation around tax notches (Kleven & Waseem, 2013), retirement timing (Burtless & Moffitt, 1986), welfare program participation limits (Hyslop & Maré, 2024), firm behavior near regulatory thresholds (Ewens et al., 2024), and voluntary certification in product markets (Houde, 2022). While in education, the bunching approach has been implemented mainly to reveal teachers' manipulations of student test scores to meet certain graduation and performance thresholds (Dee et al., 2019; Diamond & Persson, 2016). Our study adds to the existing literature by revealing bunching that is driven by student effort choices.

Lastly, this study contributes to the literature on reference-dependent effort provision. Prior work has shown that reference points—such as income targets or goal benchmarks—can shape effort and decision-making across various domains, including labor supply (Camerer et al., 1997), athletic performance (Allen et al., 2017), quota-driven work environments (Gneezy et al., 2011), and educational investments (Meng, 2019; Pope & Simonsohn, 2011). Our setting features a salient institutional reference point that creates both a kink in the cost of earning additional credits as well as a psychological goal that shapes student behavior, thus acting as a complementary mechanism. While this context does not allow us to fully disentangle cost-based frictions from reference-dependent preferences, our findings contribute to a broader understanding of how institutional design and behavioral responses jointly influence effort and academic outcomes.

The paper proceeds as follows. Section 2 describes the institutional background. Section 3 outlines the conceptual framework. Section 4 details the data and identification strategy. Section 5 presents the observational evidence on bunching, including mechanism tests and off-track analysis. Section 6 reports the field experiment results. Section 7 concludes.

2. Institutional Background: Description of the ECTS

The ECTS[2] is almost universally adopted by accredited higher education institutions in the European Union. It is designed to standardize the planning, delivery, and evaluation of study programs across universities and nations in Europe. Academic progression towards a degree is measured through CP earned. Ultimately, the system standardizes student workload per CP and thus makes it easier to compare and convert coursework between universities across Europe. Among ECTS institutions, the most common degree fields are business, administration, or law (25.6%), followed by engineering, manufacturing, and construction (14.9%), health and welfare (13.9%), and arts and humanities (9.5%)[3]. In most higher education institutions across Europe, students do not have to pay tuition fees, as these are largely subsidized by governments. As a result, students incur the same cost regardless of the number of credits they

[2] See: https://education.ec.europa.eu/education-levels/higher-education/inclusive-and-connected-higher-education/european-credit-transfer-and-accumulation-system

[3] See: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics

earn, which eliminates any financial barriers that might affect students' decisions regarding their credit load.

The ECTS sets a RP of 30 CP per semester, around which European universities design and structure their study programs. This alignment restricts planning, making it difficult for students to exceed 30 credits; Section 3.1 further elaborates on how this generates a kinked cost structure. This 30-credit RP, suggested by the ECTS, is based on the ability of a 'typical' student, who would need 25-30 hours of academic work per semester to obtain one CP. The typical student can thus obtain 30 CP per semester if they invest 35-40 hours per week in study time. There will plausibly be students who (i) possess higher ability and could obtain more than 30 CP per semester with the same invested time and effort; (ii) prefer to invest more time and can therefore obtain more credits; and (iii) are a combination of (i) and (ii). However, these students often face substantial implicit costs when attempting to surpass the 30-credit benchmark, which can slow their academic progress.

3. Conceptual Framework

We propose a theoretical framework combining two mechanisms that explain how institutional settings, particularly those found in the ECTS, can contribute to bunching in credit point acquisition. The interplay of both these frameworks provides a plausible theoretical mechanism behind the empirical results.

First, we modify the framework in Saez (2010) to fit the context of higher education. The original framework is concerned with kink points created by discontinuities in marginal tax rates. These kinks then produce bunching in earnings at the threshold. Kleven (2016) outlined that such kinks in cost functions are present in a range of settings, and we argue that the institutional setup in higher education is one of them. The second framework considers a utility function with reference dependence, which will produce bunching in performance at the salient RP. We detail the mechanisms of both frameworks below.

3.1 Bunching Due to Kinks

The first framework outlines how credit attainment changes in response to a discrete change in the cost of earning an additional credit beyond the kink. We utilize a budget set to represent the number of credits students could attain, given the amount of effort they exert. In this model, a student's utility function depends positively on the credits acquired and negatively on the effort exerted (i.e., the cost).

In a scenario where the distribution of costs and preferences is smooth, individual optimization should result in a credit accumulation distribution that is also smooth. Along this distribution, all students would face the same marginal cost *e1* of credit attainment, so any heterogeneity in credit accumulation would be due to differences in preferences.

We argue that students in higher education face a non-linear budget set, characterized by a kink in the cost function at a threshold *k*, corresponding to 30 ECTS per semester. This kink arises from the institutional structure of study programs, which are explicitly designed around this benchmark. Course schedules, teaching capacity, and administrative resources are all optimized to support the standard 30-credit workload, creating a system that implicitly penalizes deviations from this norm. Unlike the standard credit load, there are no formal guidelines or support mechanisms for students seeking to take additional credits. In many cases, attempts to do so are obstructed by logistical barriers—courses beyond the recommended plan often overlap with required classes, are not offered in the same semester, or lack available spots. These constraints generate higher search and coordination costs for students who wish to accelerate their studies, effectively increasing the marginal cost of earning credits beyond the institutional threshold. Consequently, the credit-earning cost function becomes kinked, with a discrete rise in effort and complexity once students attempt to move beyond the 30-ECTS benchmark.

Such a kink at *k* implies the following behavioral predictions:

1) All students acquiring credits below *k* continue to face a marginal cost *e1*. The credit attainment distribution left of the kink *k* is unaffected, as incentives do not change.
2) A higher marginal cost *e2* of extra credits beyond *k*, thus *e2 >e1*. Beyond the threshold *k*, the budget set becomes flatter, indicating more needed effort per credit. This results in an incentive change for all students wanting to attain more than *k* number of credits.
3) A spike in the distribution of earned credits at *k*, resulting from students choosing to decrease the number of credits attained from a point above the threshold *k* to a point at the threshold, due to increasing costs and incentive change.

Figure 1 displays this mechanism. Absent the kink, students locate along the budget set according to their preferences. A kink in the cost function at *k* would flatten the students' budget set after the kink, forcing students to exert disproportionally more effort for a marginal credit thereafter. This influences the decision-making of students who would have exceeded 30 credits absent of the kink, also known as marginal bunchers. Such a kink would shift the marginal bunchers from above the threshold at *e** to bunch at the kink and exert effort *e*. While marginal non-bunchers, students who would in any case not be willing to exceed the threshold, would not be affected by the kink, as they choose to exert an effort level at or left of the kink.

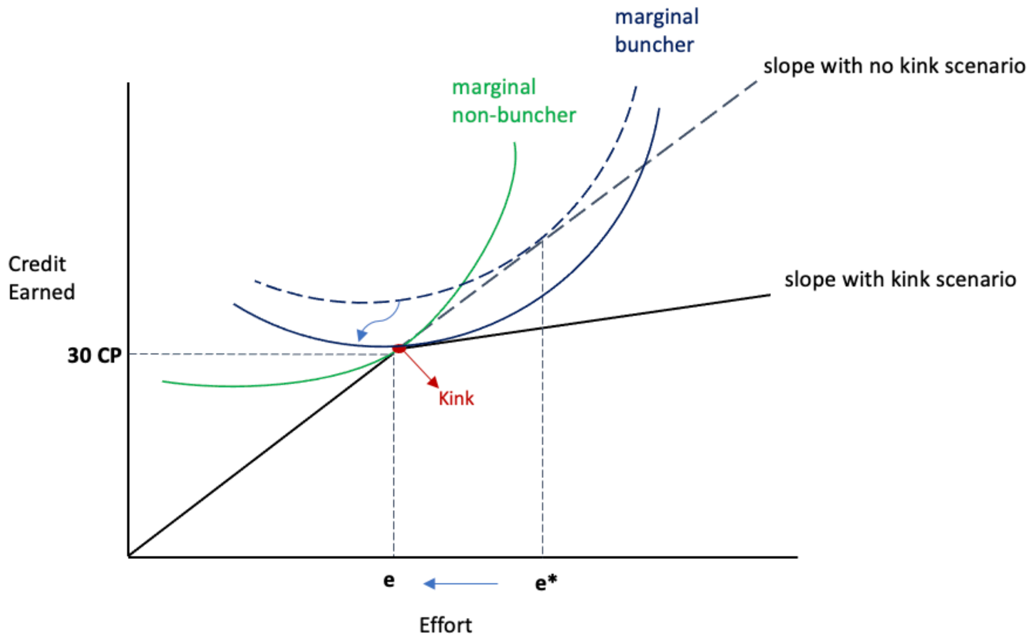


Fig. 1. Budget Set Framework. *Notes:* This graph shows the effect of a discrete increase in the cost of earning a marginal CP beyond the reference point on the number of earned CP. Students affected by the kink are labelled marginal bunchers and represented by the blue indifference curve, while students not affected by the kink are labelled marginal non-benchers and represented by the green indifference curve. Effort e* represents the optimal effort exerted by a marginal buncher, and the kink will decrease their effort level to e. Whereas all students initially located left of or at effort level e are non-bunchers, and the kink will not influence their effort level.

Introducing a kink in the cost function leads, through individual optimization, to a spike in density at the threshold. The number of students bunching at e is the total number of students initially located at the threshold k in addition to the sum of all marginal bunchers, i.e., those students who moved down to the kink. This results in excess mass in the number of students earning exactly 30 CP per semester, along with missing mass in students earning beyond 30 CP per semester.

3.2 Bunching Due to Reference Dependence

Reference points amplify the effects of bunching at kink points (Kleven, 2016). In the reference dependence model, students evaluate their credit acquisition based on a reference point, differing from the standard utility model (Kahneman & Tversky, 1979). The reference point set by the study program significantly influences students' perceptions of success at semester's end. Rather than assessing total net credits, students evaluate their performance based on deviations from this reference point. This creates a kink in the utility function, with a discontinuity at the reference point producing an optimal threshold (Kleven, 2016). Figure 2 illustrates the S-shaped value function of acquired credit points.

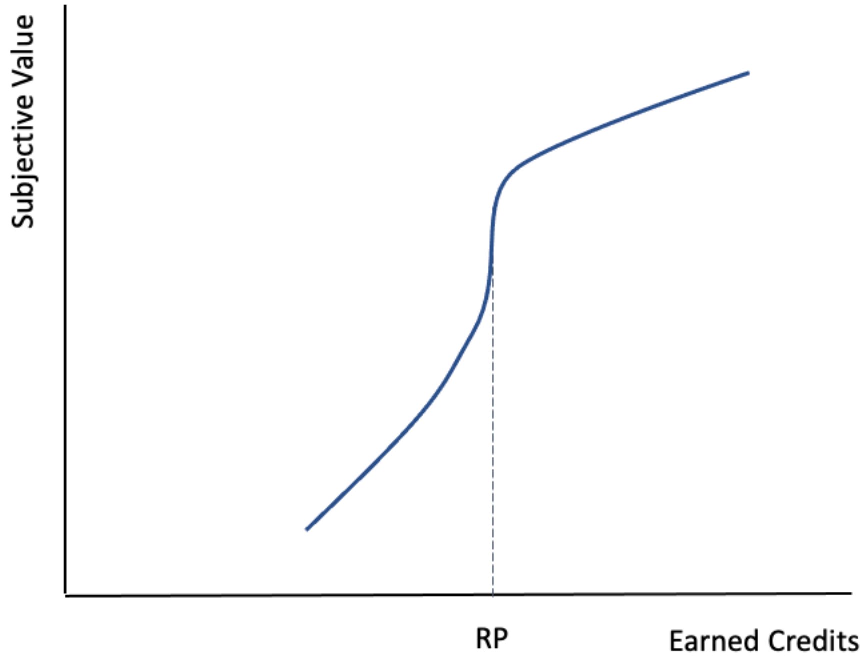


Fig. 2. S-shaped utility function. *Notes:* This graph illustrates an S-shaped value function based on prospect theory. The reference point, set at 30 ECTS credits per semester, serves as a psychological threshold. The function reflects loss aversion and diminishing sensitivity for gains beyond the reference point.

Reference dependence establishes a threshold that acts as an optimal point for students, causing a jump in utility at the reference point and encouraging bunching there. Students exceeding the reference point experience diminishing sensitivity in gains, reducing the marginal utility of additional credits. Consequently, students who might have taken more credits are discouraged, as satisfaction drops after the reference point, making extra effort seem less worthwhile. This mechanism shifts students from above the reference point back to it, contributing to excess mass. Similarly, students experience loss aversion if falling short of the RP, thus, experiencing relatively high marginal utility from earning credits up to the RP. Since the credit number choice occurs at the beginning of the semester, the concepts of diminishing sensitivity and loss aversion have a relatively higher influence on the decision outcome. The reference point affects students across the distribution, preventing missing mass before the reference point by filling gaps with those further back in the distribution (Kleven, 2016). Our two conceptual frameworks are not mutually exclusive: the reference point creates reference dependence, affecting credit demand, and simultaneously introduces a kink in the cost of earning credits. Both lead to bunching through distinct mechanisms.

Bunching at the kink point can be amplified when the kink aligns with a RP, such as the 30-credit RP (Kleven, 2016). When both occur at the same point, it is difficult to disentangle the relative contributions of the cost kink and reference dependence to the observed bunching, making attribution challenging (Kleven, 2016)[4]. Several features of our setting, however, indicate that cost-based responses dominate and that marginal bunchers primarily originate from above the threshold. First, the sharp increase in marginal costs beyond 30 credits strongly discourages students who would otherwise exceed the benchmark. Second, upward adjustments from below require substantial additional effort, making large increases toward 30 credits unlikely. Third, reducing effort is behaviorally easier than increasing it. Finally, the absence of missing mass (hole) just below 30 credits (Figures 3 and 5) rules out substantial upward bunching driven by reference dependence. Together, these patterns suggest that the tendency to bunch at 30 credits primarily reflects downward adjustments made by students who would otherwise earn more than 30 credits.

4. Data and Empirical Approach

We obtained observational data from Germany on earned credits per semester that is representative at both a national and an institutional level. Like most European higher education systems, the German higher education system is structured according to the ECTS. A typical bachelor's degree consists of either 180 or 210 credits, which students are advised to finish in six or seven semesters, respectively. Across most German higher education institutions, tuition fees are largely subsidized by governments, removing financial obstacles that could influence students' credit load decisions.

Germany’s higher education system comprises two types of institutions: universities and universities of applied sciences (UAS). The curriculum at UAS is generally more practice-oriented compared to the traditional, research-focused German universities. Additionally, UAS have more structured study programs, offering students fewer free electives. UAS play a significant and growing role in the German education landscape; in 2020, approximately 40% of freshman students enrolled in a UAS (Statistisches Bundesamt, 2021). Our analysis in this paper includes data from both types of institutions.

[4] One solution is to examine bunching at alternative reference points in the credit distribution that lack associated cost kinks. Comparing the excess mass at these points to that at the kinked reference point (30 CP) allows estimation of how much bunching is due to the kink versus the reference point itself. However, this is not feasible in our context, as the ECTS system features only one salient reference point—30 CP per semester.

We rely on the number of ECTS credits earned as a primary measure of academic progress. The focus of the analysis is on the first academic year, mainly because students usually do not have practical responsibilities like internships and for the sake of comparability with national surveys that usually report on the number of credits earned on a yearly basis. We excluded transfer and dropout students[5] from both datasets since they are not the focus of this research.

The first data source comes from collected administrative student-level data from our partner university's examination office, which is one of the largest UAS in Germany. This includes, among others, data on the number of credits earned per semester from several cohorts and a wide range of study programs from a well-known higher education institution in Germany. Data from this institution included information from 8 cohorts[6] during the period 2014-2021 with a total of around 21,000 students. In this dataset, we include only study programs with an average RP of 30 CP per semester for consistency.

The second data source comes from the National Educational Panel Study (NEPS), which is established in Germany to provide longitudinal data on educational processes and competence development. We specifically make use of the SC5 data provided by the NEPS, which focuses on the German higher education system and collects data from first-time enrolled students in a public or state-approved institution of higher education in Germany in 2010-2011[7]. This dataset included data from 17,910 students from 261 institutions, including students from both UAS and universities. This dataset was intended to be representative of the entire German higher education system.

We argue that the mechanisms leading to bunching are not only present in the European higher education system but are also widespread due to the common use of a single institutional reference point in higher education. Therefore, we also obtained data from a higher education institution in the USA, which operates under a system different from the ECTS. Specifically, we use data from a large university that is part of the 23 California State University (CSU) campuses[8]. The data comprise attempted credit hours for approximately 8,000 full-time students from the 2010 and 2011 cohorts.

---

[5] We excluded dropouts and effective dropouts, those who earn no credits, to make the administrative data more comparable to the national studies, as this group typically does not respond to the NEPS surveys and was found to have a high attrition rate (Zinn et al. 2020).

[6] The cohorts included all students who started in the winter semester of every respective year, and their relevant data on their outcomes were continuously collected till the winter semester of 2022.

[7] At the time of writing this paper, the latest data from NEPS on university students was from 2010-2011.

[8] We sincerely thank professors Nick Huntington-Klein & Andrew Gill for supporting this study and providing the necessary administrative data on the number of attempted credit hours in one of the CSUs.

4.1 Empirical Approach

The general notion of estimating excess mass is to construct a counterfactual distribution in which the influence of the kink on the given distribution is removed. The bulk of the bunching literature has focused on the taxation context, where tests for bunching occur around kink points caused by a sudden change in the marginal tax rate. We rely on similarities between students' responses to abrupt increases in marginal CP cost and individual taxpayer responses to "kinks" in the tax code (e.g., Kleven and Waseem, 2013; Saez, 2010). Similar to how a kink point arises from a change in the marginal tax rate, in our context, a kink point induces an abrupt jump in the cost of attaining a marginal CP after the threshold, which is typically caused by various institution-related costs inflicted on students.

To estimate bunching, we draw upon the methodology proposed by Mortenson and Whitten (2020), which closely follows Chetty et al. (2011)'s methodology of measuring the amount of excess mass around kinks. The adopted methodology displays the frequency of earned credits at each bin while estimating the excess mass, i.e., the extent to which students bunch at the reference point relative to a hypothetical situation without the kink. The number of bunchers is measured by the difference between the observed and counterfactual credit distribution at the kink. To quantify the excess mass in an interval around the kink, the counterfactual distribution is estimated, i.e., the credit attainment distribution that would have been observed absent the kink. This is estimated by grouping the credits into bins of equal size, denoted by j, then fitting a flexible polynomial to the local density of earned credits around the reference point while excluding a small number of observations left and right of the kink. This window of excluded observations is selected to capture all individuals bunching at the kink.

Specifically, the following regression is estimated:

$$C_j = \sum_{i=0}^{p} \beta_i . \left(Z_j\right)^i + \sum_{i=-R}^{R} \gamma_i . 1[\mathrm{j} = \mathrm{i}] + \varepsilon_i \tag{1}$$

Where C_j is the number of students in the credit bin j, Z_j is the credit attainment level in bin j, [-R, R] is the excluded range of observations around the kink, and p is the order of the polynomial.

A potential concern with this specification is that the distribution of earned credits exhibits mechanical heaping at multiples of 5 CP due to the institutional structure of course offerings. These structural spikes may violate the smoothness assumption of the polynomial counterfactual. To address this, we conduct robustness checks (see Appendix Table A1) where we augment the baseline model with modulo-5 indicator variables[9]. These controls absorb the systematic heaping at round numbers, ensuring that the bunching estimate is identified by the underlying smooth distribution rather than mechanical artifacts (Ito and Sallee, 2018).

To estimate the counterfactual distribution of earned credits absent of the kink, a 7th ordered polynomial, p, is fitted, which excludes a small bunching window of observations 1 credit left and right of the threshold (30 CP), [-R, R]. This window was indicated based on a visual inspection of the distribution as recommended by Chetty et al. (2011). To ensure the highest precision, the smallest bin width value of 1 credit point is used.

$$\hat{b} = \frac{\sum_{j=-R}^{R} c_j - \hat{c}_j}{\bar{\hat{c}}_j} \tag{2}$$

The estimated excess mass of credits earned around the kink, denoted by $\hat{b}$, is calculated by the summation of differences between the observed density in the bunching region and the fitted counterfactual density, which is then normalized by the counterfactual values of that region. Therefore, the reported $\hat{b}$ value is equal to the percentage of students at the kink who are classified as bunchers. The standard errors will be calculated by using a bootstrap procedure.

The bunching coefficient b estimates the total excess mass at the threshold. To interpret this magnitude in terms of student behavior, we calculate the average reduction in CP required to generate this excess mass. Following Chetty et al. (2011), we estimate the average behavioral response as $\Delta z = B/h_0$, where the total excess mass (B) is normalized by the counterfactual density (h_0). This ratio represents the average number of CP by which marginal bunchers reduced their workload.

5. Results

5.1 Testing for Bunching in Earned Credit Points

We find substantial bunching at exactly 30 earned CP per semester. Figure 3, Panel 1(a) shows the distribution of earned credits from eight cohorts at our partner university between 2014 and 2021. Approximately 27% of students earn exactly 30 CP, while only 1% exceed their study plan by one course or more (5+ CP). This sharp concentration at the institutional threshold, followed by a steep drop in density, visually demonstrates a behavioral response that deviates

[9] We thank Mortenson and Whitten (2020) for making their estimation code publicly available. We have modified their algorithm to explicitly allow for the inclusion of round-number fixed effects (heaping controls) within the bunching estimation procedure.

markedly from the smooth distribution expected if credit loads reflected only individual differences in ability or motivation.

Formal estimation confirms significant excess mass at the 30 CP threshold. Panel 1(b) presents results from our bunching estimator, which fits a seventh-order polynomial to the credit distribution while excluding observations within one credit point of the threshold. We find that approximately 42% of students earning exactly 30 CP are marginal bunchers, students who would have earned a different amount of credits absent the institutional constraint (b = 0.423, SE = 0.045)[10]. This estimate implies that approximately 1,464 students in our sample chose lower credit loads than they would have absent the constraint, with an estimated average reduction of 2.2 CP per semester. This response creates a substantial missing mass above 30 CP, corresponding to roughly 13% of the total student population. Consequently, only about 2% of first-year students complete one additional course beyond the standard study plan, far below the counterfactual prediction.

National data from the NEPS confirm that this pattern extends across the German higher education system. Panel 2(a) presents the distribution of average earned credits in the first academic year from the NEPS data, distinguishing between Universities and UAS. The distributions exhibit the same characteristic spike at exactly 30 CP, followed by a sharp decline above the threshold. In universities, about 19% of students earn exactly 30 CP, while roughly 6% earn the equivalent of one additional course or more per semester. The concentration at 30 CP is substantially more pronounced in UAS, where 31% of students earn exactly this amount, and only 1.5% exceed it by a full course. This lower rate of exceeding 30 CP is consistent with the more structured study programs and rigid curricula typically found in UAS, which reduce institutional flexibility for accelerated progress.

Panels 2(b) and (c) confirm significant excess mass in both Universities and UAS, identifying approximately 44% and 41% of students at the threshold as marginal bunchers, respectively. These estimates imply that absent the institutional reference point, over 40% of these students would have earned additional credits. While theoretically an upper bound due to potential reference-dependent adjustments from below, the absence of a dip in the distribution immediately preceding the threshold suggests such upward shifts are minimal.

The NEPS data reveal a higher proportion of students exceeding 30 CP compared to the partner university. This discrepancy arises because NEPS encompasses students from various

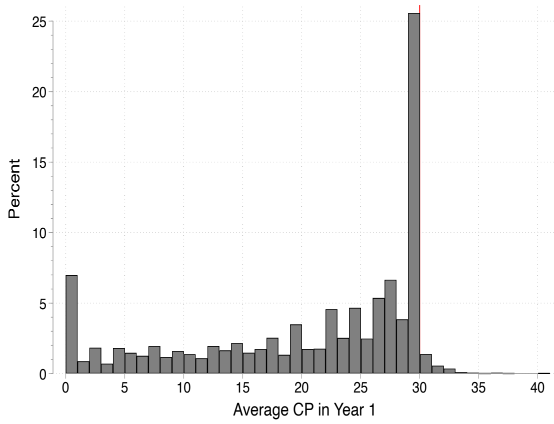
[10] We also find significant bunching in earned credit points in the first two academic years in the partner university (see Figure A2 in the Appendix).

institutions and study programs, some of which set their RPs slightly above 30 CP per semester, deviating from the ECTS standard[11]. Since the NEPS lacks information on the reference points each student is faced with, we cannot account for these observations[12]. Thus, certain programs within NEPS have study plans that encourage students to exceed an average of 30 CP in their first year, explaining the higher number of students earning above 30 CP. Moreover, the NEPS data relies on self-reported earned credits, which could lead to sample bias and, in some cases, overreporting of achieved credits[13], which further explains the discrepancy in the share of students earning above 30 CP. Lastly, as Zinn et al. (2020) discuss that the NEPS suffer from measurable selection bias due to a high attrition rate, leading to inflated reported earned CP.

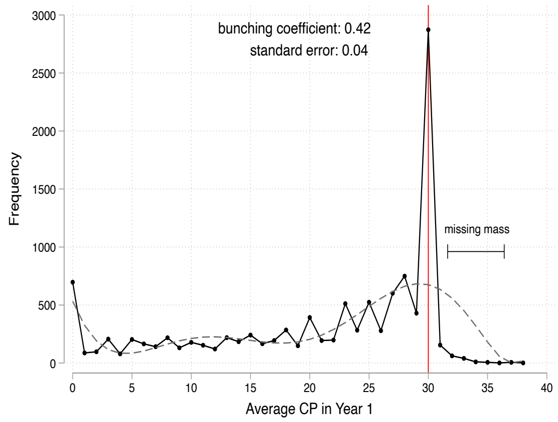
Fig. 3. Empirical and Counterfactual Distribution of Average Earned Credit Points in Year 1.

Panel 1. Partner University Data

(a) Distribution of average earned CP in year 1



(b) Bunching in earned CP in year 1



[11] When comparing the CP distribution of STEM programs from our partner university with those in the NEPS data, we find more closely related patterns. Both datasets show similar proportions of bunchers and students earning above 30 CP. For detailed CP distributions of STEM and non-STEM programs, see Table 1.

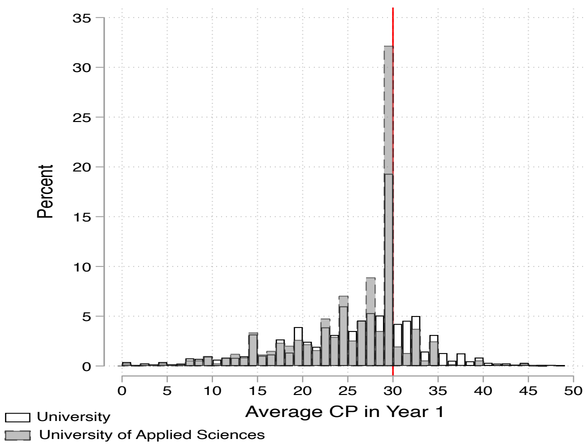
[12] As our partner university's dataset provides RP information for each study program, we exclude programs with an RP other than 30 CP for consistency. Later analysis in the paper uses this variation to show that when a program sets a different RP, the peak of credit accumulation shifts accordingly.

[13] Overestimation bias has been found prevalent when asking students to self-report their learning outcomes, e.g., Mayer et al. (2007) found that 39% of students overreport their SAT scores.

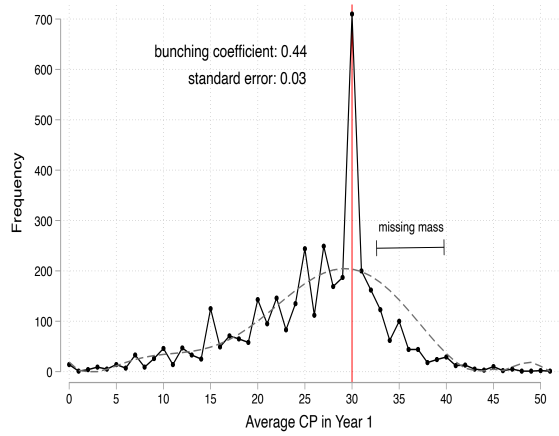
Fig. 3 (Continued). Empirical and Counterfactual Distribution of Average Earned Credit Points in Year 1.

Panel 2. NEPS Data

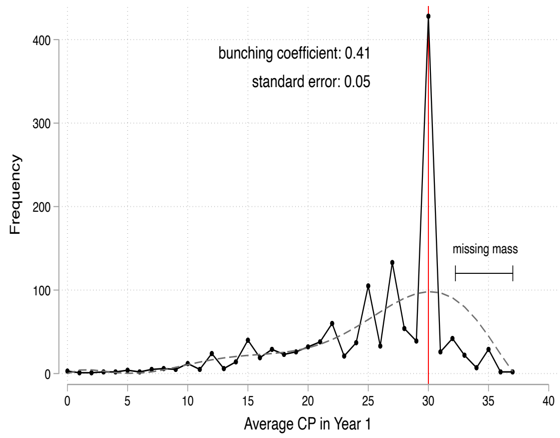
(a) Distribution of average earned CP in year 1



(b) Bunching in earned CP in Universities



(c) Bunching in earned CP in Universities of Applied Sciences



Notes: The figures show the empirical and counterfactual distributions of average earned CP in the first academic year, based on data from the partner university (Panel 1) and the National Educational Panel Study (NEPS) (Panel 2). In the NEPS sample, results are disaggregated by institution type—Universities and Universities of Applied Sciences—to highlight differences in bunching behavior. In the counterfactual panels, black dots indicate the observed number of students earning each specific number of credit points (i.e., the empirical density), while the gray dashed line shows the estimated counterfactual distribution in the absence of bunching. The gap between the observed and counterfactual distributions indicates the extent of excess or missing mass at each credit level. The red vertical line denotes the institutional reference point at 30 earned CP.

Table 1 reports bunching coefficients across program and institution types. The bunching magnitude is large and statistically significant across all fields. In the full samples, 42 to 45 percent of students at the 30 CP threshold are classified as marginal bunchers (Panel A), with consistent results in both STEM and non-STEM fields. This concentration reflects a meaningful distortion of student potential: we estimate that marginal bunchers forego an average of 2.1 to 2.2 CP per semester to align with the threshold. The influence of institutional structure is most clearly visible when comparing Universities and UAS. While both institution types exhibit comparable bunching coefficients, the absolute share of students earning exactly 30 CP is substantially higher in UAS (31% vs. 18%). This demonstrates that highly prescriptive study plans naturally coordinate a larger share of students onto the standard path. In contrast, the flexible structure of traditional universities relies more heavily on active student choice, leading to a more dispersed distribution with lower baseline concentration at the threshold, even though the marginal behavioral response remains comparable.

We explicitly rule out the possibility that bunching is a mechanical artifact of course structures. A potential identification threat arises from the discreteness of course values: because most modules award credits in multiples of 5 ECTS, the credit distribution naturally exhibits systematic "heaping". This structural lumpiness risks distorting the counterfactual, as the polynomial might mistake a mechanical spike for a smooth trend. To address this, we re-estimate the model using modulo-5 indicator variables. These controls effectively absorb the average "round-number effect" across the entire distribution, forcing the estimator to identify excess mass at 30 CP only if it exceeds the structural heaping observed at other multiples of five. As shown in Appendix Table A1, the estimated bunching magnitude remains robust and statistically significant across all specifications. Importantly, allowing for two-sided donor regions, so that excess mass at the threshold may originate from both below and above 30 CP, yields coefficients that are comparable to, and in several cases larger than, the baseline estimates. This pattern indicates that the results are not driven by asymmetric donor assumptions or mechanical heaping, but instead reflect a robust behavioral response to the 30-credit threshold.

The 30-CP threshold effectively acts as a constraint on faster academic progress. We estimate that between 13% and 15% of students in UAS and about 12.4% of students in traditional universities are missing from the distribution above 30 CP per semester because they reduce their workload to align with the institutional norm. Figure A1 in the Appendix visually shows the excess mass at the threshold and the missing mass in the distribution of higher credit loads across all three samples. This distortion is economically meaningful: instead of building credit buffers or graduating early, even high-performing students are constrained to the standard pace. Ultimately, this implies that the current framework does not merely standardize student progress but actively suppresses it, creating an artificial ceiling on human capital accumulation.

Table 1. Estimates of Bunching at the 30 Credit Point Threshold

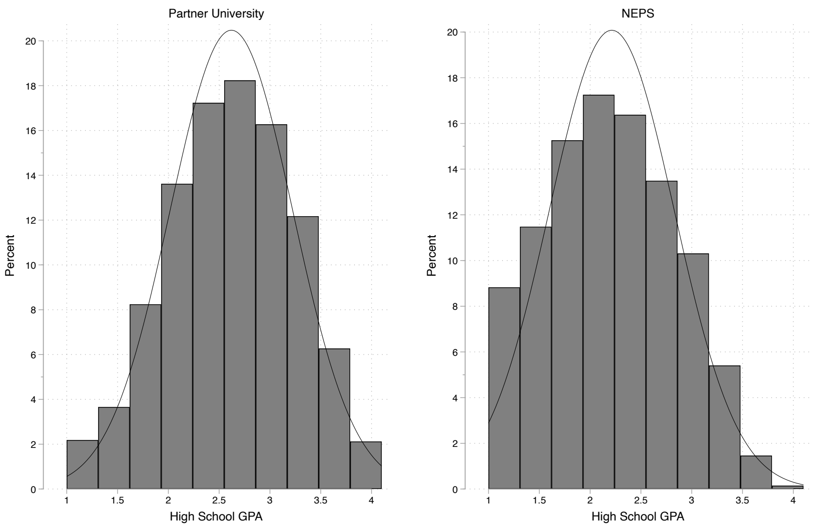
	Partner University (1)	NEPS (Universities) (2)	NEPS (UAS) (3)
Panel A: All Programs			
Share of Bunchers at 30 CP	0.423***	0.430***	0.410***
	(0.045)	(0.041)	(0.064)
Estimated Number of Bunchers	1,464	472	202
Average CP Reduction	2.20	2.27	2.09
Share of Constrained Students	13.0%	12.4%	15.0%
Observed Share at 30 CP	25.5%	18.7%	31.0%
Observations	11,231	3,786	1,339
Panel B: STEM Programs			
Share of Bunchers at 30 CP	0.432***	0.422***	0.388***
	(0.056)	(0.034)	(0.052)
Estimated Number of Bunchers	997	218	87
Share of Constrained Students	12.5%	13.1%	13.2%
Observed Share at 30 CP	22.6%	18.4%	26.0%
Observations	7,967	1,662	659
Panel C: Non-STEM Programs			
Share of Bunchers at 30 CP	0.386***	0.400***	0.407***
	(0.068)	(0.050)	(0.092)
Estimated Number of Bunchers	445	233	110
Share of Constrained Students	13.6%	11.1%	16.4%
Observed Share at 30 CP	32.5%	19.0%	37.1%
Observations	3,264	2,091	669

Notes: "Share of Bunchers (b)" represents the bunching coefficient (estimated excess mass relative to the counterfactual density), derived from equation 1. The bunching coefficient is normalized by a *counterfactual* density, not the observed mass at 30 CP."Estimated Number of Bunchers" is the estimated count of students shifting to the threshold, calculated as the excess mass relative to the counterfactual distribution. "Average CP Reduction" represents the average behavioral response, calculated as the estimated excess mass divided by the counterfactual density. It indicates the average number of credit points by which marginal bunchers reduced their workload to align with the threshold. "Share of Constrained Students " represents the share of the total student population estimated to be constrained by the threshold (calculated as Number of Bunchers divided by Total Observations). "Observed Share at 30 CP" refers to the observed percentage of students earning exactly 30 CP per semester. Standard errors (in parentheses) are calculated using a bootstrap procedure. *** p<0.01, ** p<0.05, * p<0.1.

5.2 Student Ability

It is important to evaluate whether the observed bunching at 30 CP is primarily due to institutional features or whether it is driven by other factors. In this section, we investigate the role of ability in explaining the observed bunching in CP earned.

Panel A. High School GPA



Panel B. First Year University GPA

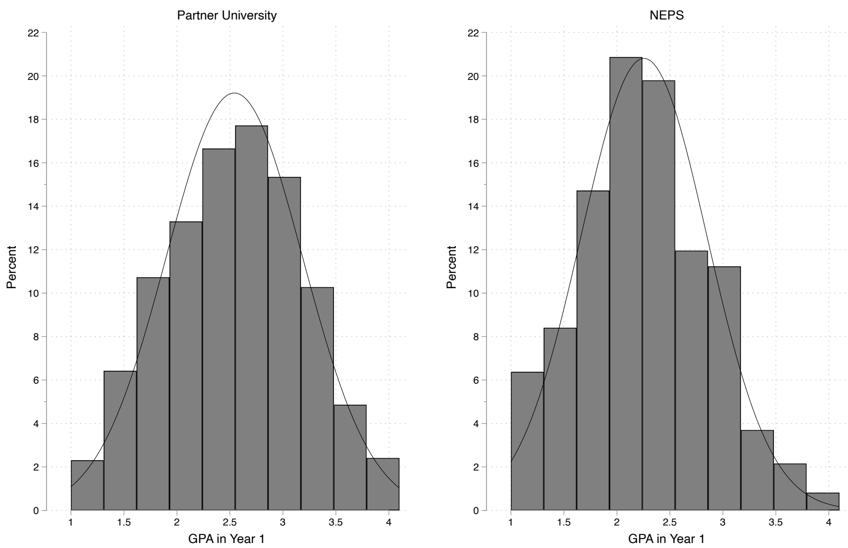


Fig. 4. Distribution of High School and First-Year University Grades. *Notes*: Grades follow the German grading system, where 1 represents the highest grade, and 4 is the lowest passing grade. The left side of each panel corresponds to students from the partner university, while the right side represents students from the NEPS dataset. The overlaid curve in each histogram represents a normal distribution fitted to the empirical mean and standard deviation of the respective group.

Research often shows that cognitive abilities, such as IQ, follow a normal distribution in a population (Johnson et al., 2008). Our datasets, which use high school grades as a proxy for ability, also show a roughly normal distribution for student ability. This demonstrates that bunching does not reflect a concentration of students at a particular ability level (see panel A of Figure 4), but rather a behavioral response to institutional constraints. Credits earned and GPA are educational outputs that depend on the student's ability. Unlike credits earned, university students' GPAs do not bunch at specific values but show a smooth distribution (see panel B of Figure 4), reflecting the normal distribution of student ability. Bunching occurs only in the dimension where institutional constraints plausibly funnel student progress: credits per semester.

5.2.1 Student Ability and Earned Credits Distribution

Since the ECTS 30-CP threshold is set for a 'typical' student, it is plausible that a sizable portion of lower-ability students cannot exceed this threshold. In contrast, higher-ability students should, in principle, be more likely to earn well beyond 30 CP. Figure 5 illustrates the distribution of earned CP categorized by high school GPA quantiles, our proxy for ability. While all quantiles display a visible peak at 30 CP, the shape of the distribution varies systematically with ability. As student ability increases, the entire CP distribution shifts to the right, i.e., higher-ability students tend to accumulate more credits. However, this otherwise positive linear relationship is disrupted at the kink point, where we observe bunching. Notably, the bunching magnitude is positively correlated with student ability, resulting in a more pronounced peak at the kink point. This growing discontinuity at 30 CP suggests that certain constraints, particularly on higher-ability students, result in a ceiling effect driving the concentration at the threshold, rather than continuing to spread across the higher end of the CP distribution.

Panel A. Partner University Data *Panel B. NEPS Data*

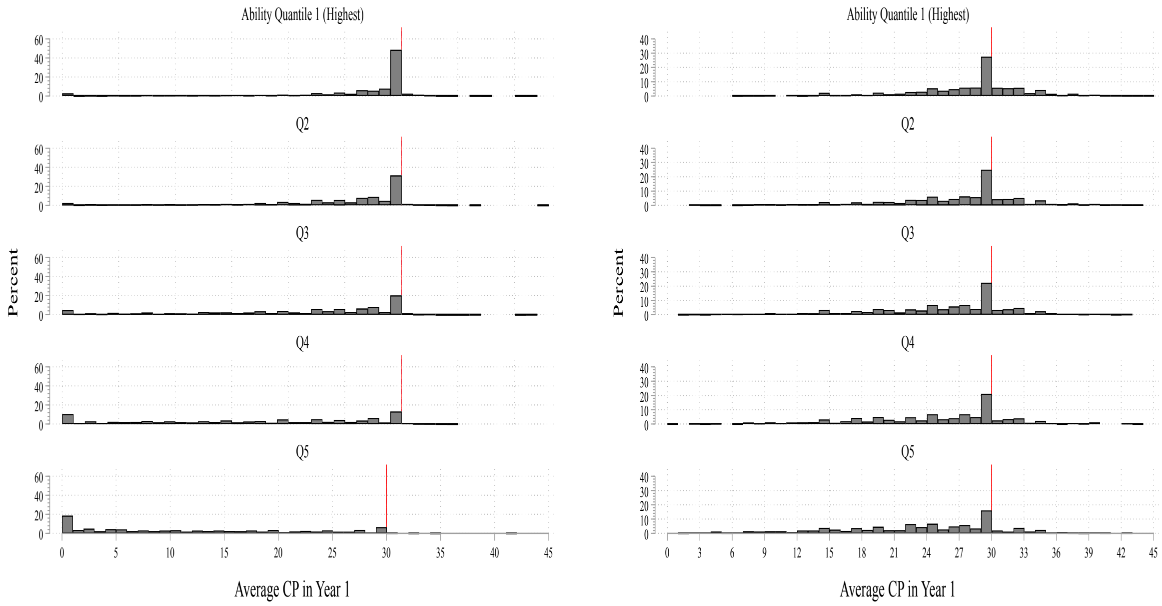


Fig. 5. First-Year Average CP Distribution Split by Ability. *Notes*: The graphs illustrate the distribution of earned credit points among students of varying abilities. Panel A displays data from the partner university, whereas Panel B depicts data from the NEPS dataset. Students are divided into five ability quantiles based on their high school grades, with Quantile 1 representing the highest-achieving students and Quantile 5 the lowest. The red line indicates the institutional reference point.

Figure 6 presents the cumulative distribution of earned credits by ability quantiles, which reinforces this pattern of threshold-driven behavior. For instance, in our partner university data, high-ability students (quintile 1) are notably concentrated just at or below the 30 CP threshold: 50% of them have earned up to 29 CP, and the cumulative share jumps to 95% at exactly 30 CP. In contrast, the 50th percentile of the median ability quintile earns up to 23 CP, with the cumulative share reaching 98% at 30 CP. These distributions suggest that while higher-ability students earn more credits leading up to the threshold, credit accumulation, regardless of ability, converges at 30 CP. Beyond this point, progress plateaus, indicating dominant institutional constraints, particularly on high-ability students. A similar pattern is observed in the NEPS data, though the magnitude is lower.

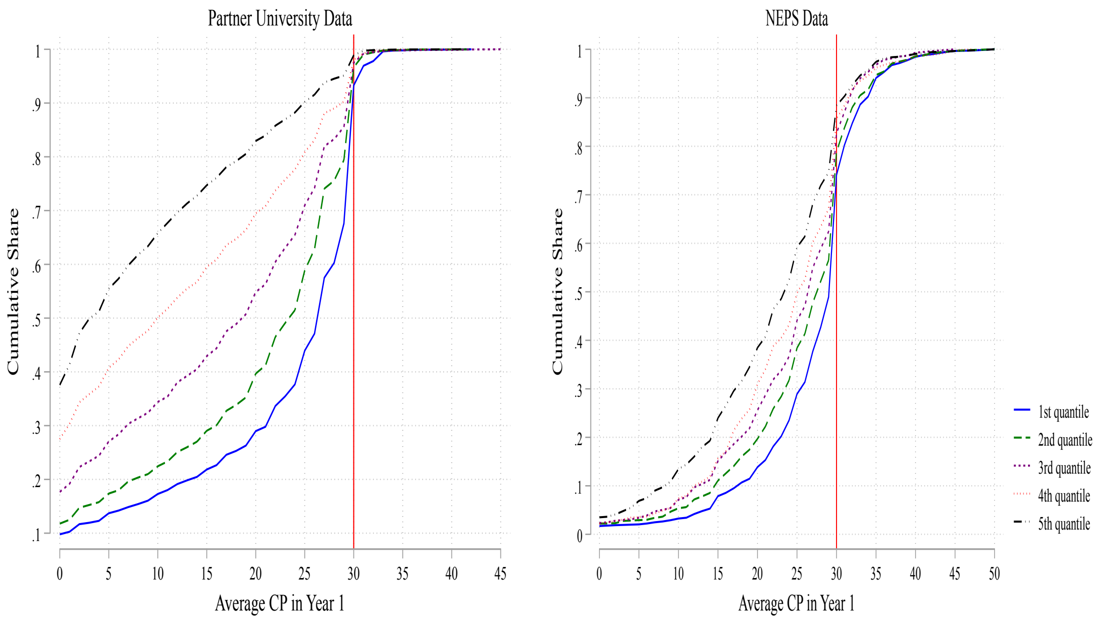


Fig. 6. Cumulative Distribution of First-Year Credit Points by Student Ability Quintile. *Notes*: This figure shows the cumulative distribution of average CP earned in the first year, segmented by student ability quintiles. The left plot displays data from the partner university, and the right plot presents data from the NEPS sample. Students are divided into five ability groups based on high school GPA, with the 1st quintile representing the highest-achieving students and the 5th quintile the lowest. The x-axis indicates the average number of CP earned in Year 1, while the y-axis shows the cumulative share of students within each ability group. The red vertical line marks the institutional reference point at 30 earned CP.

5.3 Study Programs with Non-Standard Reference Points

To further substantiate that institutional reference points drive the observed bunching effects, we exploit the fact that, despite the ECTS prescribing 30 credits per semester, there are some study programs with reference points that differ from the standard 30 credits per semester[14]. By studying these programs, we can observe the effects of changes in the institutional kink point on the distribution of earned credits. Using the partner university data, out of a total of 26 study programs, we have identified 2 study programs with an average RP above the ECTS's recommendation of 30 credits per semester and 2 study programs with a RP below 30 credits during their first academic year.

We categorize the study programs into three groups based on their RP in the first academic year: those with an average RP below 30 CP, those with an average RP above 30 CP, and a reference group comprising study programs with an average RP of exactly 30 credits. Using regression analysis, we examine the relationship between non-standard institutional reference

[14] While deviations from the 30-credit load occasionally occur in specific semesters, the average recommended credit load across the entire program remains 30 CP per semester, i.e., credit load is shifted from one semester to another.

points and the likelihood of students exceeding the 30-ECTS threshold, controlling for several individual characteristics such as age, gender, prior academic performance (high school GPA and diploma type), international student, and transfer status. The results, presented in Table 2, indicate that the RP of a study program significantly affects the probability of students exceeding 30 CP. Compared to the reference group, being in a study program with a RP above 30 credits increases the probability of exceeding an average of 30 CP per semester in the first year by 39 percentage points. In contrast, students in programs with an RP below 30 CP are about 2 percentage points less likely to exceed the 30 CP threshold. This relatively small magnitude reflects a floor effect, as the baseline probability of exceeding 30 CP in the reference group is already small (1.7%), leaving little room for further reduction.

Table 2. Effect of Non-standard RPs on the Probability of Exceeding 30 CP

	Exceeded average 30 CP in year 1	
	(1)	(2)
RP over 30 CP	0.395***	0.390***
	(0.01)	(0.01)
RP under 30 CP	-0.01***	-0.021***
	(0.001)	(0.001)
Controls	No	Yes
Reference group (RP 30 CP)	0.017	0.017
Observations	21,704	21,704

Notes: The table reports OLS regression estimates based on administrative data from eight student cohorts at the cooperating institution. The outcome variable equals 1 if a student earned more than 30 CP per semester, on average. “RP over 30 CP” refers to study programs that set a reference point above 30 CP; “RP under 30 CP” refers to programs with a reference point below 30 CP. We control for age, gender, high school GPA, type of high school diploma, international status, and transfer status. Standard errors are shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Visual evidence reinforces that student effort tracks the institutional benchmark. Figure 7 demonstrates that in programs where the target exceeds 30 CP, the entire distribution shifts rightward. Crucially, a significant proportion of students in these tracks successfully surpass the standard ECTS limit, proving that the 30-credit threshold is not a binding capacity constraint. This responsiveness implies that the widespread bunching observed at 30 CP is a product of the institutional framework rather than limited student ability; when the structural ceiling is raised, students readily scale their effort to match it.

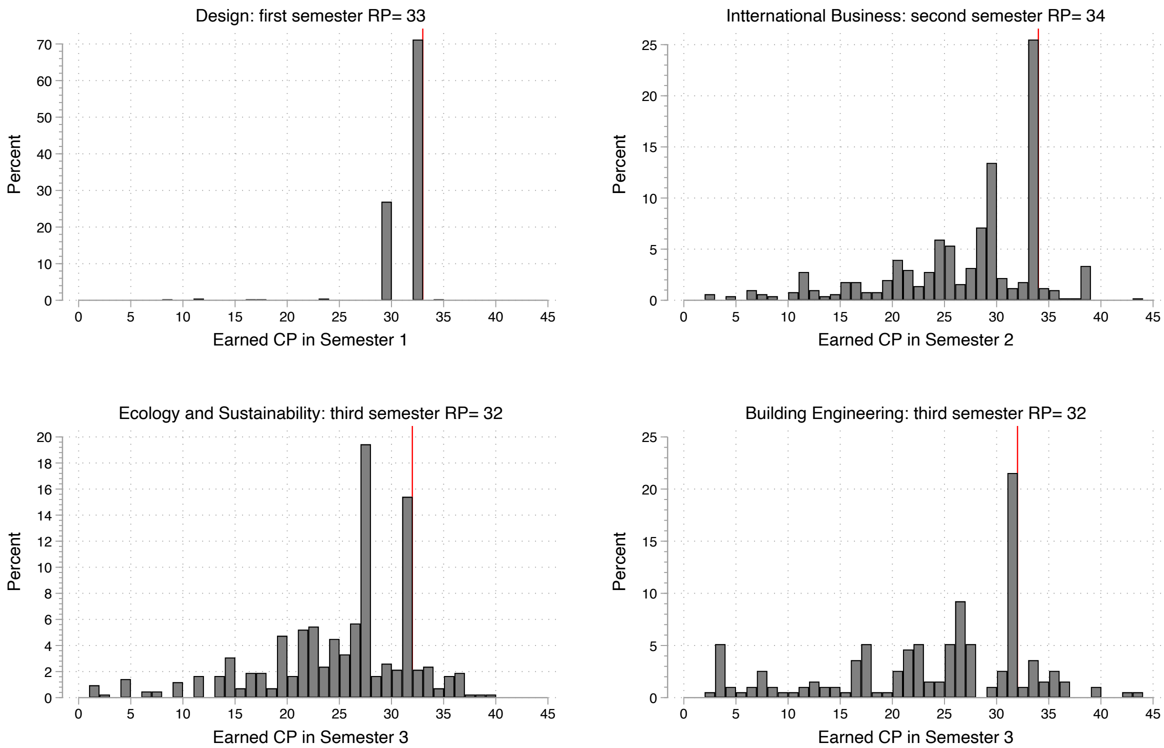


Fig. 7. Average Credit Points Earned Distribution in Study Programs with Non-Standard Reference Points. *Notes*: The figure shows a collection of CP distributions in specific study programs and semesters where the institutional reference point does not match the ECTS's reference point. We remove dropouts, unengaged, and transfer students from this analysis. The red line represents the institutional reference point.

5.4 "On-Track" vs "Off-Track" Students

This section exploits variation in students' exposure to institutional constraints to construct a quasi-counterfactual with reduced barriers at the 30 CP threshold. Our strategy compares "on-track" students, who adhere to the standard plan (averaging 30 CP per semester in Year 1), against "off-track" students who narrowly fall short (averaging 24 to 29 CP per semester in Year 1). While on-track students face high friction to exceed the 30 CP threshold, off-track students operate under a regime where catch-up behavior is explicitly accommodated. We restrict the off-track group to a deficit of just one or two courses to ensure they remain comparable to their on-track peers in baseline ability and motivation. This design allows us to test whether the bunching at 30 CP per semester persists when the marginal cost of exceeding it is structurally reduced.

We argue that institutional rigidity makes it significantly more costly and difficult for on-track students to exceed the standard load. High search costs and rigid scheduling create a sharp kink in the cost function beyond 30 CP per semester, effectively penalizing efforts to accelerate. In contrast, off-track students encounter an institutional framework that offers favorable conditions to exceed 30 CP. Institutions often plan for course repetition by a certain portion of a cohort and adjust accordingly, thereby accommodating students who need to catch up. These logistical adjustments, such as scheduling flexibility and reserved capacities, facilitate easier course access for off-track students. Moreover, off-track students benefit from lower search costs, as they can retake missed courses in addition to orienting their course selection according to their study plan. Such conditions contribute to a smaller kink at the threshold faced by off-track students, making it relatively easier for them to surpass 60 CP total in Year 2[15].

The observed credit distributions reveal stark differences between the two groups. Figure 8 displays the total earned CP per semester distribution in the second year, categorized by students' first-year performance (on-track or off-track) from both datasets. On-track students exhibit a significant bunching coefficient at the 60 CP cumulative threshold, confirming a sustained behavioral response to the standard 30 CP per semester pace. In contrast, off-track students show a rightward-shifted distribution with the peak moving beyond the threshold and considerably less concentration at exactly 60 CP. At our partner university, 29% of off-track students exceed 60 CP in Year 2 compared to only 15% of on-track students, nearly double the rate. This pattern emerges despite off-track students having slightly worse average high school GPA (2.4 vs. 2.2), suggesting that institutional flexibility, not ability, drives the difference. Moreover, these results likely underestimate the potential responsiveness of the highest-ability students, who are naturally best positioned to accelerate when barriers are removed. The NEPS data corroborate this pattern, showing a dispersed, rightward-shifted distribution for off-track students compared to the characteristic spike for their on-track peers.

As off-track students are less constrained by the 30-CP per semester kink, they are expected to bunch less than on-track students on that threshold. Figure 9 tests this by comparing the bunching behavior of on- and off-track students at the threshold in their second year. Our analysis confirms that on-track students are more likely to bunch at an average of 30 CP per semester, exhibiting a significant bunching coefficient, with missing mass in the distribution of students earning more than 30 CP per semester. For instance, our partner university dataset shows that around 56% of on-track students earning 60 CP in year 2 are marginal bunchers. On the other hand, we find no evidence of bunching among off-track students at that kink point,

[15] While retaking courses may differ from new coursework, both require substantial effort. Moreover, some off-track students planned lower initial course loads or attended irregularly, making marginal effort beyond 30 CP per semester broadly comparable across off-track and on-track students.

with less missing mass above the threshold. These results indicate that on- and off-track students have different kink points and face different marginal costs beyond 30 CP per semester. More importantly, we provide evidence that with the right incentives, students choose to exceed the given credit-per-semester reference point, significantly reducing the observed bunching effects.

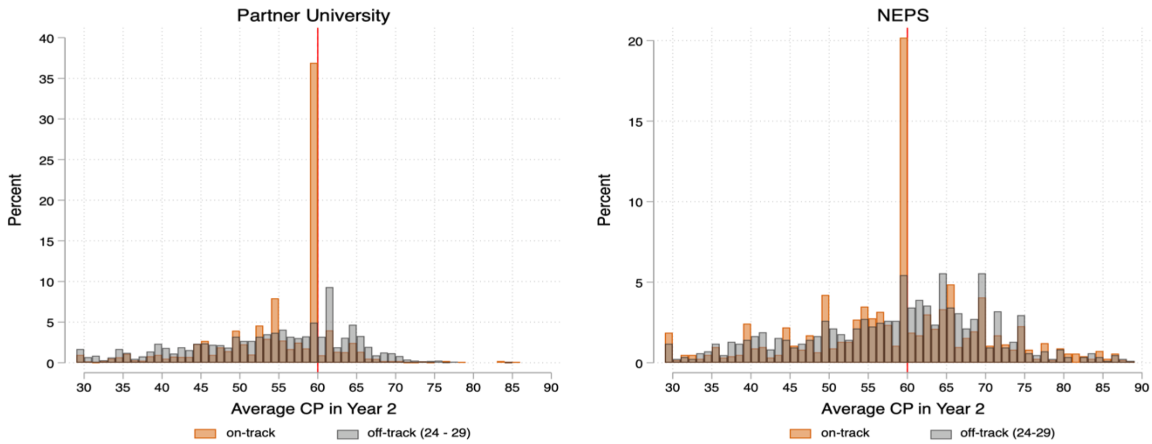


Fig. 8. Second Year Average Earned Credit Points of On-Track vs. Off-Track Students.
Notes: The figure displays the distribution of earned credit points in the second year of study, using data from the partner university (left panel) and the NEPS dataset (right panel). Students are grouped based on their first-year performance: on-track students (orange bars) earned 60 CP or more in Year 1 (averaging ≥30 CP per semester), while off-track students (grey bars) earned between 24 and 29 credit points per semester on average. The red line represents the institutional reference point.

Fig. 9. Bunching in Total Earned Credit Points in the Second Year of Higher Education for On-track and Off-track Students.

Panel A. Partner University Data

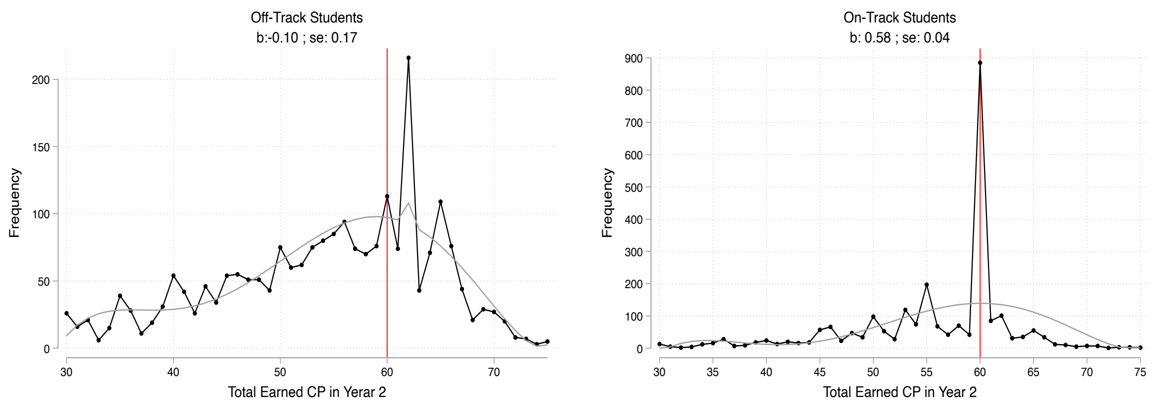
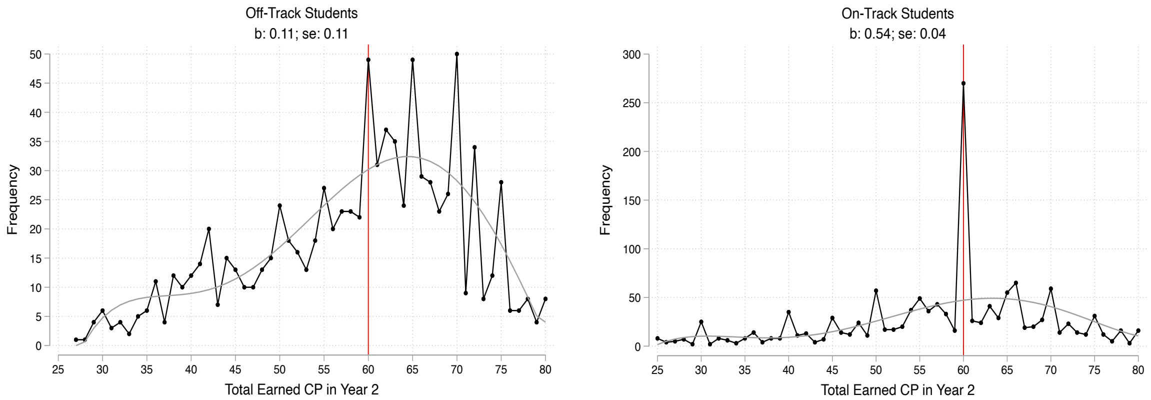


Fig. 9 (Continued). Bunching in Total Earned Credit Points in the Second Year of Higher Education for On-track and Off-track Students.

Panel B. NEPS Data



Notes: The figure presents the bunching analysis of total earned credit points in Year 2 across two datasets. In each panel, the left-hand side displays results for on-track students, while the right-hand side shows results for off-track students. The bunching coefficient is denoted by b, with standard errors (se) in parentheses. We estimate the counterfactual distribution (grey line) using a 5th-degree polynomial due to a smaller sample size, thus ensuring numerical stability. Dropouts and transfer students are excluded from the analysis. The red vertical line marks the institutional reference point.

These visual patterns are confirmed by regression analysis. Table 3 examines how reduced institutional friction (being off-track) affects the probability of earning more than 60 CP in year 2 while controlling for several covariates. This analysis enables us to estimate the number of students who would exceed 30 CP per semester with the right incentives, i.e., when having a smoother credit-earning cost function. Overall, our analysis reveals that off-track students are more likely to exceed 30 CP than on-track students. At our partner university, off-track students who earn an average of 24 to 29 credits in their first year have a 14-16 percentage point higher likelihood of surpassing the 60-credit threshold in their second year compared to on-track students. This pattern is further substantiated by a semester-level analysis: students who were off-track in semester 1 are significantly more likely to exceed 30 CP in semester 2 compared to their on-track peers (see Appendix Table A2). Using the NEPS data, we find similar effects, though lower in magnitude, possibly due to strong selection effects as we must rely on a selective subsample of students who responded twice to the NEPS survey, in their first and second academic year consecutively[16]. Analyzing off-track students suggests that higher

[16] NEPS asks students to report the total number of credits they earned at the end of their first and second academic years. Therefore, to compute the credits earned only in the second academic year, we must subtract the total credits earned at the end of year two from the total credits earned at the end of year one. Thereby, relying on a possibly selective sample of responders who answered this survey question twice in two consecutive years. For example, while 6,046 students reported their total number of credits earned in their first academic year, only 3,757 students managed to report their earned credits two years in a row, thus allowing us to compute their number of earned credits in year two for a selective sample.

marginal costs beyond 30 CP per semester, driven by the institutional setup, hold back up to 16% of students from exceeding this threshold.

Table 3. Effect of Being Off-Track in Year 1 on the Probability of Exceeding 60 Credit Points in Year 2

	University Data Exceeded 60 CP in year 2		NEPS Exceeded 60 CP in year 2	
	(1)	(2)	(3)	(4)
Off-track	0.143***	0.166***	0.079***	0.085***
	(0.01)	(0.01)	(0.01)	(0.01)
Controls	No	Yes	No	Yes
Study program FE	Yes	Yes	Yes	Yes
Constant	0.136	0.141	0.310	0.318
Observations	4,875	4,875	3,088	2,600

Notes: The table reports OLS regression estimates of the effect of being off-track in Year 1 on the likelihood of exceeding 60 credit points on average in the second academic year. The outcome variable is a binary indicator equal to 1 if a student earned more than 60 credit points total in Year 2. Columns (1) and (2) use data from the partner university, while columns (3) and (4) use data from the NEPS sample. Estimations using partner university data include study program fixed effects, while NEPS analyses include faculty fixed effects due to more limited identifiers. Controls include age, gender, high school GPA, type of high school diploma, international status, and transfer student status. Standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Critically, exceeding 60 CP does not compromise grades. Table 4 presents a difference-in-differences analysis comparing GPA trajectories of off-track students who exceeded 60 CP in Year 2 against on-track students. Off-track students maintained stable GPAs (Year 1: 2.32; Year 2: 2.31), nearly identical to on-track students (Year 1: 2.13; Year 2: 2.14). The statistically insignificant difference-in-differences estimate (Column 3) rules out quality-quantity tradeoffs, confirming that students can handle higher credit loads without academic penalty when institutional barriers are reduced. This finding aligns with Attewell and Reisel (2012), who find no adverse effects from heavier-than-recommended course loads, and reinforces that the binding constraint is institutional design, not student capacity.

We find similar results in the NEPS data, where the GPA of off-track students in year two closely follows that of their on-track peers. The insignificant difference in differences between the groups indicates that exceeding the RP does not necessarily compromise the quality of education. This finding is consistent with Attewell and Reisel (2012), who reported no adverse effects on students taking a heavier credit load than recommended by their study programs.

Table 4. Difference-in-Differences Estimates of GPA by Track Status and Academic Year

	Year 1 GPA	Year 2 GPA	Diff. in Diff.
	(1)	(2)	(3)
Panel A: partner university			
Off-track (>30 CP in year 2)	2.32	2.31	0.00
			(0.02)
On-track	2.13	2.14	-
Observations	3,972	3,643	7,615
Panel B: NEPS			
Off-track (>30 CP in year 2)	2.14	2.09	0.00
			(0.01)
On-track	2.05	2.00	-
Observations	2,848	2,276	5,124

Notes: The table reports GPA outcomes in Year 1 and Year 2 for on-track students (those earning ≥30 credit points in Year 1) and off-track students (those earning between 24–29 credit points in Year 1 who accelerated to >30 credit points in Year 2). Estimates are based on a difference-in-differences (DiD) approach, where the difference-in-differences estimate (Column 3) captures whether GPA trajectories differ by initial track status following credit recovery. GPA is measured on the German grading scale, where 1 is the best, and 4 is the lowest passing grade. Panel A presents results from the partner university's administrative data, while Panel B uses data from the NEPS sample. Estimations using partner university data include study program fixed effects, while NEPS analyses include faculty fixed effects due to more limited identifiers. Standard errors for the DiD estimates are shown in parentheses.

These findings establish that institutional design, not student capability, constitutes the binding constraint on credit accumulation. The off-track analysis provides two key insights. First, the substantial 8–16 percentage point treatment effect confirms that credit accumulation above 30 CP is structurally feasible. Reforms removing barriers at the threshold, such as flexible study plans or formal fast-track pathways, would enable a significant share of students to accelerate their progress. Second, the absence of a GPA penalty demonstrates that these higher credit loads do not compromise learning quality. Ultimately, the current system effectively caps the academic potential of capable students not through explicit prohibition, but through implicit institutional friction.

5.5 Welfare Implications: A Back-of-the-Envelope Calculation

The bunching behavior documented in this paper imposes substantial welfare costs through delayed degree completion. To quantify these costs, we analyze the distribution of credits earned during the first five semesters (the structured phase of study) for 4,691 students at our partner university who graduated within five years (see Figure A3). Based on the estimates in

Table 1, approximately 13% of the student population is constrained by the threshold and reduced their credit load in response to the 30 CP reference point. These students forgo an average of 2.2 CP per semester over the structured five-semester phase, accumulating an 11 CP shortfall that delays graduation by roughly one-third of a semester per affected student. Across the full student population, this translates to a potential average delay of 0.05 semesters per student, amounting to 117 cumulative years of delayed labor market entry in our sample.

Extrapolating to Germany's 250,000 annual bachelor's degree graduates (Statistisches Bundesamt, 2021) implies 6,250 student-years of delayed labor market entry nationwide. At a conservative starting salary of €40,000 (Heming et al., 2020), this sums to around €250 million in annual foregone earnings. However, this baseline assumes the loss is limited to entry-level wages. If we instead consider lifecycle earnings profiles, where a delay effectively displaces a peak-earning year prior to retirement—valued at approximately double the starting wage (OECD, 2021)—the aggregate loss could rise to as much as €500 million.

Even these estimates are likely conservative in several respects. First, it excludes the fiscal cost of extended enrollment, which amounts to approximately €73 million in public expenditure (at €11,700 per student-year; OECD 2023). Second, it focuses only on deterministic delays, ignoring dynamic costs: the rigid 30 CP structure prevents students from building "credit buffers" to protect against future setbacks, a friction that likely contributes to the high rate of extended graduations in Germany (43%; Statistisches Bundesamt, 2021). Table A3 in the appendix provides an outline of the calculations.

5.6 Cross-System Evidence: Behavioral Patterns in U.S. Higher Education

The differences between the American and European higher education systems expose students to different incentives, costs, and reference points, which affect their credit per semester choices. In this section, we explore how institutional setups under different systems impact the earned credit per semester distribution.

As in Europe, U.S. study programs define a target credit load per semester to ensure timely graduation. While this target is measured in credit hours rather than ECTS points, the recommended load of 15 credit hours per semester corresponds closely to the 30 CP RP in the ECTS framework[17]. The U.S. system, however, features two salient reference points. The first is administrative: at 12 credit hours, students qualify as full-time, triggering substantial benefits

[17] In the U.S., this target is typically measured in credit hours, while in Europe it is expressed in credit points. One U.S. credit hour is generally equivalent to 2 ECTS credit points.

and a discrete change in incentives[18]. The second is academic: at 15 credit hours, study plans are structured for on-time completion. Beyond this level, the marginal cost of taking additional credits increases due to scheduling constraints and resource allocation, creating a kink in the cost of effort similar to the European setting.

Using data from a large university that is part of the 23 California State University (CSU) campuses, we analyse the distribution of attempted credit hours for two cohorts starting in 2010 and 2011, comprising a total of 8,079 full-time students[19]. Approximately 76% of students at this institution are enrolled full-time, implying that around one quarter of students attempt fewer than 12 credit hours.

Figure 10 presents the distribution of attempted credit hours in the first year for 8,079 full-time students from the 2010 and 2011 cohorts. The distribution exhibits two distinct peaks: a substantial concentration at 12 credit hours (25–30% of full-time students) reflecting the administrative threshold, and a sharp spike at 15 credit hours—the on-time graduation benchmark—followed by a steep decline. We pool 15 and 16 credit hours in the figure because many STEM courses (e.g., chemistry with lab components) award 4 credits, allowing students to reach 16 while adhering strictly to their study plan. The original unpooled distributions appear in Appendix Figure A4.

The right tail of the distribution closely mirrors our European findings. Despite left-censoring at 12 credits, only 4–6% of full-time students attempt credits beyond the 15-credit benchmark. This share is strikingly similar across systems: approximately 3% of students at our German partner university complete 35 CP (one additional course), while a comparable 3–4% of U.S. students attempt 18 credit hours or more. Importantly, the sharp drop beyond 15 credits occurs at a threshold with no discrete change in benefits, unlike the 12-credit cutoff, pointing to a behavioral response to institutional reference points rather than incentive eligibility. Taken together, these patterns indicate that salient reference points embedded in program structures systematically shape student credit-taking behavior across fundamentally different higher education systems.

[18] Full-time students in the U.S. benefit from greater financial aid and tax credits, broader access to campus and extracurricular resources, and eligibility for student discounts and scholarships.

[19] As the sample includes only full-time students, defined as those attempting at least 12 credit hours, the distribution is left-censored at 12 credit hours. Consequently, the graphs reflect only the credit attempted patterns within the full-time student population and cannot be generalized to the entire student body.

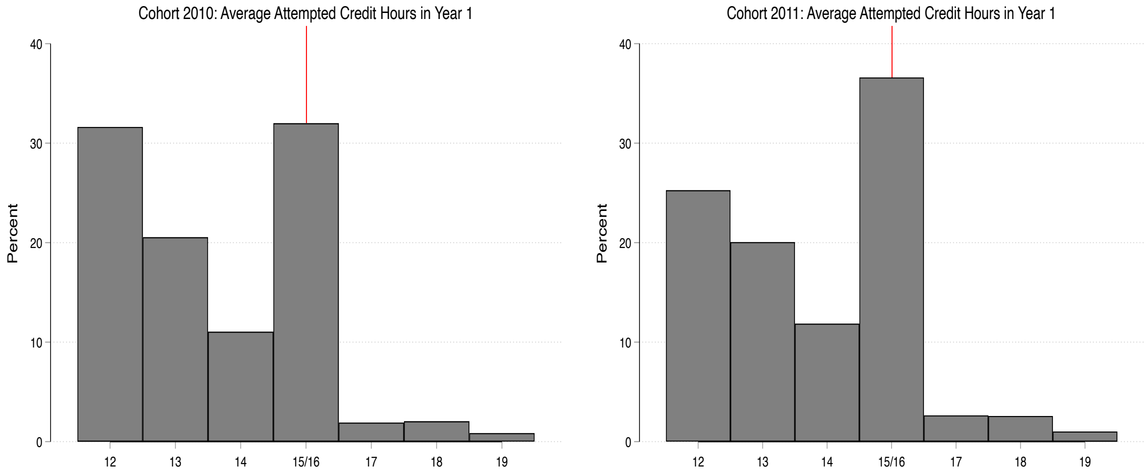


Fig. 10. Average Attempted Credit Hours Distribution in the USA. *Notes*: The figure consists of two histograms, representing the first-year average number of credit hours attempted by full-time students of two cohorts (2010 and 2011) in a large University in the USA. The institutional reference point in the United States is set at 15 credits per semester, while the full-time status threshold is at 12 credit hours.

6. Field Experiment

In the previous sections, we provided evidence of bunching in CP attainment at institutional kink points, leading to students rarely surpassing that threshold. Given that the current institutional setup incurs private and social costs by holding back students, a key question is how these costs can be mitigated. An obvious solution would be to re-evaluate and improve the ECTS, possibly by including several reference points that accommodate the heterogeneity of the student population. However, this requires action at the European level and is unlikely to happen in the short term. Therefore, we investigate whether we can, under the existing ECTS setup, reduce bunching by encouraging students to exceed their study program's RP.

To determine whether this bunching is driven by genuine institutional constraints or merely informational frictions, we conducted a preregistered field experiment at a German university[20]. Leveraging insights from Abeler et al. (2011) on how expectations shape effort, our intervention intends to de-anchor students from the 30 CP default by making the option to accelerate salient and reducing associated search costs. This design allows us to test whether bunching arises from limited awareness rather than binding structural barriers, in which case this low-cost nudge should significantly increase the share of students accelerating their studies.

[20] Pre-registration link: https://doi.org/10.17605/OSF.IO/MYFQS

6.1 Study Background

The field experiment was conducted with a newly enrolled cohort from the winter semester of 2020 in the bachelor's program in Business Administration (BA). The business administration program considered in our study follows a standard curriculum that is structured in accordance with the ECTS. In total, students must obtain 210 credit points within the regular duration of the study program of seven semesters. Freshmen students in the BA program are advised to complete 30 credits in their first semester, as is the norm in most universities across Europe.

In the first week of the semester, randomly selected students in the treatment group received an unannounced letter in the mail, while students in the control group did not receive any information from the research team. The postal letter informed the treatment group that the 30 credits per semester recommendation is aimed at a typical or average student (see Appendix Figure A5 for documentation). Furthermore, the letter suggests that it can be reasonable to attempt more than 30 credits, states the potential advantages of this, and gives a list of courses that are usually scheduled for the second semester or later that can also be taken in the first semester, thus enabling students to exceed 30 CP. Finally, the letter provides basic statistics on the distribution of earned credits by previous cohorts.

6.2 Data and Estimation

We utilized administrative student-level data from the university's examination office. In total, 413 students were randomized into the treatment or control group. Randomization was carried out using stratification (block randomization) and balancing, i.e., re-randomization[21]. We then linked treatment status to administrative data on the credits per semester outcomes and utilized demographic information and pre-treatment outcomes as covariates in our estimations. Overall, all pre-registered covariates are balanced across both groups, indicating a successful randomization (see Appendix Table A4).

As preregistered, the following outcome dimensions are considered: Our key dependent variables are binary variables that indicate whether a student signs up for, attempts, and earns more than 30 credits in the first semester. Earned credits will be measured net of any credits that students transfer from prior studies[22].

---

[21] Stratification variables include high school GPA split into three categories and a dummy indicating whether a student is enrolling for the first time at any university. Based on the interaction of the two stratification variables, we create six strata. Within those strata we re-randomized to achieve balance with respect to age, a female dummy, the high school GPA, a high school degree type dummy, and a high school degree obtained in 2020 dummy.

[22] This is done to ensure consistency with credits signed up for and attempted, as students do not have to sign up for or attempt credits that are transferred from previous studies.

We estimate Intention-To-Treat (ITT) effects based on OLS estimations. We run the following regression analyses:

$$Y_i^k = \alpha_0 + \alpha_1 T_i + \alpha_3 x_i + \alpha_4 z_i + \varepsilon_i, \tag{3}$$

where Y_i^k is the outcome of interest for individual i, T_i is a dummy variable indicating whether individual i was treated. In a further specification, we will control for individual background characteristics vector z_i. This includes age, gender, university entrance certificate grade, type of high school diploma[23], and whether the university entrance certificate was obtained in 2020.

6.3 ITT Effects

Table 5 shows the ITT effects of the intervention, specifically the share of students who signed up for, attempted, and passed more than 30 credits in the treatment group compared to the control group. Across all specifications, treatment effects are small, negative, and statistically insignificant. By the end of the semester, 3% of students in the treatment group earned more than 30 CP, compared to 6% in the control group (difference = −2.8 pp, SE = 2.0 pp). Similar non-significant differences are observed in the number of registered and attempted credits above the 30 CP threshold. These findings demonstrate that purely informational interventions, absent structural support like alternative study plans, are insufficient to motivate students to exceed the institutional RP.

Table 5. ITT Effects on Exceeding 30 Credit Points

	Signup over 30		Attempt over 30		Earn over 30	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.023	-0.022	-0.029	-0.028	-0.029	-0.029
	(0.028)	(0.029)	(0.025)	(0.025)	(0.020)	(0.020)
Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Control mean	0.10	0.10	0.08	0.08	0.06	0.06
Observations	409	409	409	409	409	409

Notes: This table reports OLS estimates of the intention-to-treat effects of the treatment on three outcomes measured as the share of students who sign up for, attempt, and earn more than 30 credits during their first semester. We control for age, gender, university entrance certificate grade, type of high school diploma, date of enrollment, and transferred credits. All specifications include strata fixed effects. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

[23] In Germany, students could qualify to enter university if they either have a high-school degree or some equivalent certification.

The intervention also failed to shift the overall credit distribution or average credit loads. Figure 11 displays distributions of sign-up, attempted, and earned credits by treatment status. The distributions are virtually identical across treatment arms, a pattern confirmed by a Kolmogorov-Smirnov test (D = 0.038, p = 0.999). Consistently, we find no treatment effect on the average net credits signed up for, attempted, and earned (see Table 6). These null effects on both the extensive margin (exceeding 30 CP) and intensive margin (average credits) indicate the intervention did not meaningfully increase credit accumulation.

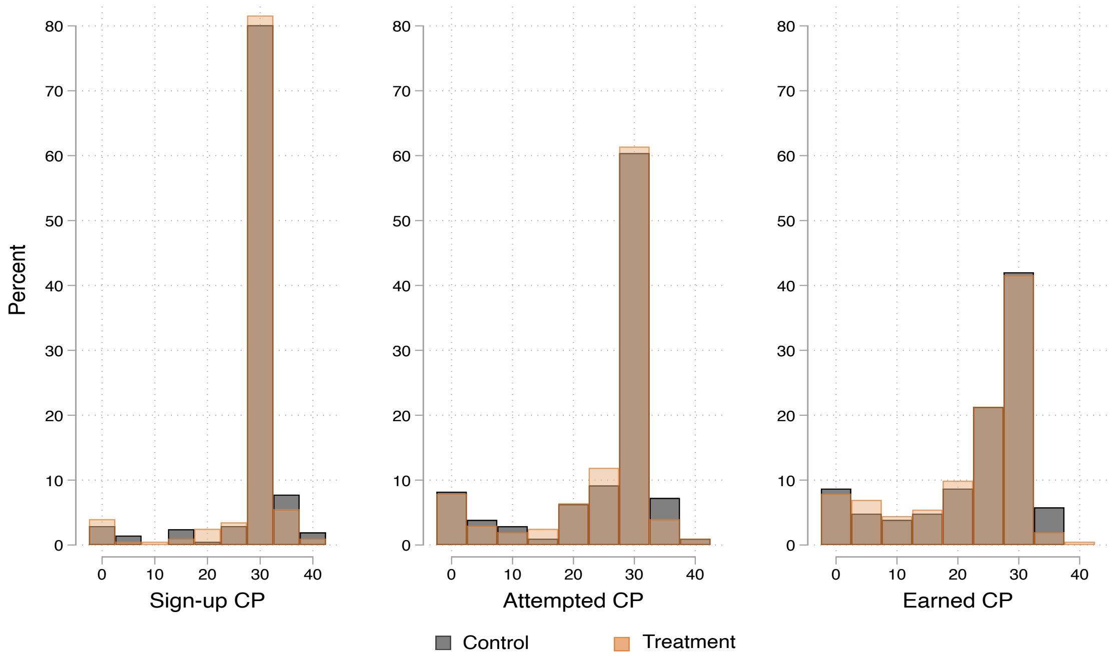


Fig 11. Distribution of Credit Outcomes. *Notes:* This figure presents histograms of student credit outcomes from the field experiment sample in the first semester, comparing the control group (gray) and the treatment group (orange) across three metrics: Sign-up Credit Points, Attempted Credit Points, and Earned Credit Points.

Table 6. ITT Effects of the Field Experiment on Average Credit Outcomes

	Average Signup CP		Average Attempt CP		Average Earned CP	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.426	-0.434	0.011	-0.038	-0.731	-0.764
	(0.709)	(0.701)	(0.973)	(0.949)	(0.996)	(0.964)
Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Control mean	28.97	28.97	25.21	25.21	23.06	23.06
Observations	409	409	409	409	409	409

Notes: Coefficients present the intention-to-treat effect by OLS regression on average credit outcomes at the end of the first semester. The outcomes include average sign-up credit points (credits registered online), attempted credit points (credits for which students sat exams), and earned credit points (credits successfully passed). We control for age, gender, university entrance certificate grade, type of high school diploma, date of enrollment, and transferred credits. All regressions include strata fixed effects. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

6.4 Other Treatment Effects

In Table 7, we investigate any potential secondary effects of the intervention on the number of failed credits, the share of dropouts, and GPA. While we find no effects on the share of dropouts or GPA, the treatment appears to increase the number of failed credits. Compared to the control group, the treatment group has significantly more failed credits of 0.72 CP by the end of the first semester. One plausible explanation is that the intervention, which highlighted the share of students in previous cohorts who failed to earn 30 CP, inadvertently normalized low performance. This likely acted as a negative descriptive norm, subtly communicating that underperformance is common among peers. As shown by Schultz et al. (2007) in their work on social influence, exposure to such a norm can lead students to normalize lower academic achievement, potentially reducing their motivation and effort. This dynamic can even manifest as a boomerang effect, where students who might otherwise have performed better inadvertently adjust their behavior downwards to align with the perceived, less desirable group average, ultimately contributing to an increase in failed credit points.

Table 7. ITT Effects on Secondary Outcomes

	Failed CP		Dropout		GPA	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.741*	0.726*	-0.018	-0.017	0.027	0.025
	(0.416)	(0.418)	(0.026)	(0.026)	(0.047)	(0.045)
Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Control mean	2.15	2.15	0.08	0.08	2.28	2.28
Observations	409	409	409	409	375	375

Notes: Coefficients present the intention-to-treat effect by OLS regression. GPA is reported according to the German grading scale, which ranges from 1 (best) to 4 (lowest passing). Data on GPA is only available for those students who earned credits. Dropout is a binary indicator equal to 1 if the student left the university. We control for age, gender, university entrance certificate grade, type of high school diploma, date of enrollment, and transferred credits. All regressions include strata fixed effects. Standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Conclusion

This paper examines whether the standardized design of the ECTS creates unintended structural barriers to academic progression. Our findings indicate that the 30 CP per semester reference point operates not merely as a guideline but as an effective institutional ceiling. We estimate that 40% of students at the 30 CP threshold are marginal bunchers—representing roughly 13% of the total population, who would have earned more absent institutional constraints. Consequently, only a negligible fraction of students complete even one additional course per semester beyond their plan, despite substantial heterogeneity in ability. This distortion is particularly pronounced in more structured environments, such as UAS, indicating that the rigidity of the academic framework plays a central role in constraining student progress.

We present compelling evidence that this phenomenon is driven by institutional design rather than student capability. Three distinct findings support this conclusion. First, bunching intensifies with better high school GPA, implying that the highest-ability students, those with the greatest capacity to exceed the standard load, are the ones most restricted by the 30 CP threshold. Second, when study programs explicitly set reference points above 30 credits, the probability of exceeding the benchmark rises by 39 percentage points. Third, our analysis of off-track students reveals that when institutional friction is reduced (via the flexibility granted to catch-up students), the probability of exceeding the threshold rises by approximately 15 percentage points without compromising academic performance. These patterns suggest that students possess the capacity to accelerate and readily do so when the cost function is smoothed.

These inefficiencies carry significant economic weight. Our back-of-the-envelope calculations suggest that the graduation delays induced by this bunching generate approximately €250 million in potential annual foregone earnings in Germany alone. Beyond direct costs, the rigidity of the 30 CP model prevents students from building "credit buffers," leaving them vulnerable to future shocks (such as illness or course failure) and increasing the likelihood of extended time to degree. While the ECTS has successfully harmonized higher education across Europe, this standardization appears to come at the cost of flexibility, creating a one-size-fits-all model that suppresses high-ability students.

To investigate the role of informational frictions, we experimentally tested whether reducing search costs and making acceleration options salient would alter student behavior. The null results of our field experiment demonstrate that simply encouraging students to exceed the threshold is insufficient to reduce bunching. The contrast is stark: informational nudges produced no effect, while structural accommodation (as seen in the off-track analysis) produced large behavioral responses. This confirms that the binding constraints are institutional rather than informational, meaning that the solution cannot be found in marginal nudges.

From a policy perspective, higher education institutions must therefore move beyond the single-track model by introducing formal "fast-track" curricula or multiple reference points that legitimize and facilitate acceleration. Such reforms would align institutional frameworks with student heterogeneity, allowing high-performing individuals to signal their ability and enter the labor market earlier. Future research should explore optimal designs for these multi-track systems, such as a fast-track option, and assess their effectiveness in unlocking student potential and reducing the welfare costs of standardized progression.

References

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, *101*(2), 470–492.

Aina, C., & Casalone, G. (2020). Early labor market outcomes of university graduates: Does time to degree matter? *Socio-Economic Planning Sciences*, *71*, 100822.

Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, *63*(6), 1657–1672.

Attewell, P., Heil, S & Reisel, L. (2012). What is academic momentum? And does it matter? *Educational Evaluation and Policy Analysis*, *34*(1), 27–44.

Bostwick, V., Fischer, S., & Lang, M. (2022). Semesters or quarters? The effect of the academic calendar on postsecondary student outcomes. *American Economic Journal: Economic Policy*, *14*(1), 40–80.

Bound, J., & Turner, S. (2007). Cohort crowding: How resources affect collegiate attainment. *Journal of Public Economics*, *91*(5–6), 877–899.

Burtless, G., & Moffitt, R. A. (1986). Social security, earnings tests, and age at retirement. *Public Finance Quarterly*, *14*(1), 3–27.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics*, *112*(2), 407–441.

Chetty, R., Friedman, J. N., Olsen, T., & Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. Macro labor supply elasticities: Evidence from Danish tax records. *The Quarterly Journal of Economics*, *126*(2), 749–804.

Dee, T. S., Dobbie, W., Jacob, B. A., & Rockoff, J. (2019). The causes and consequences of test score manipulation: Evidence from the New York regents examinations. *American Economic Journal: Applied Economics*, *11*(3), 382–423.

Diamond, R., & Persson, P. (2016). The long-term consequences of teacher discretion in grading of high-stakes tests. *National Bureau of Economic Research*.

European Commission. (2015). *ECTS users' guide 2015, Publications Office of the European Union*. https://data.europa.eu/doi/10.2766/87192

Ewens, M., Xiao, K., & Xu, T. (2024). Regulatory costs of being public: Evidence from bunching estimation. *Journal of Financial Economics*, *153*, 103775.

García-Martínez, I., Gutiérrez Cáceres, R., Luque de La Rosa, A., & León, S. P. (2021). Analysing educational interventions with gifted students. Systematic Review. *Children*, *8*(5), 365.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, *25*(4), 191–210.

Heming, J., Stanski, C., & Zimmermann, T. (2020). *Gehaltsreport für Absolventen 2020/21*. StepStone.

Hornstra, L., van Weerdenburg, M., van den Brand, M., Hoogeveen, L., & Bakx, A. (2022). High-ability students' need satisfaction and motivation in pull-out and regular classes: A quantitative and qualitative comparison between settings. *Roeper Review*, *44*(3), 157–172.

Houde, S. (2022). Bunching with the stars: How firms respond to environmental certification. *Management Science*, *68*(8), 5569–5590.

Huntington-Klein, N., & Gill, A. M. (2021). Semester course load and student performance. *Research in Higher Education*, *62*(5), 623–650.

Hyslop, D. R., & Maré, D. C. (2024). Earnings bunching at benefit abatement thresholds: Evidence from recent policy changes. *Motu Economic and Public Policy Research*.

Johnson, W., Carothers, A., & Ian J. Deary. (2008). Sex differences in variability in general intelligence: A new look at the old question. *Perspectives on Psychological Science*, *3*(6), 518–531.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*, 263–292.

Kim, M. (2016). A meta-analysis of the effects of enrichment programs on gifted students. *Gifted Child Quarterly*, *60*(2), 102–116.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, *8*, 435–464.

Kleven, H. J., & Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *The Quarterly Journal of Economics*, *128*(2), 669–723.

Levin, H. M., & Garcia, E. (2013). *Benefit-Cost Analysis of Accelerated Study in Associate Programs (ASAP) of the City University of New York (CUNY)*.

Mayer, R. E., Stull, A. T., Campbell, J., Almeroth, K., Bimber, B., Chun, D., & Knight, A. (2007). Overestimation bias in self-reported SAT scores. *Educational Psychology Review*, *19*, 443–454.

Meng, J. (2019). *To Gain with Less Pain?: Reference-Dependence in Students' Exam Performance and Effort Provision*.

Mortenson, J. A., & Whitten, A. (2020). Bunching to maximize tax credits: Evidence from kinks in the US Tax Schedule. *American Economic Journal: Economic Policy*, *12*(3), 402–432.

Pope, D., & Simonsohn, U. (2011). Round numbers as goals: Evidence from baseball, SAT takers, and the lab. *Psychological Science*, *22*(1), 71–79.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, *2*(3), 180–212.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434.

Small, M. L., & Winship, C. (2007). Black students' graduation from elite colleges: Institutional characteristics and between-institution differences. *Social Science Research*, *36*(3), 1257–1275.

Sneyers, E., & De Witte, K. (2018). Interventions in higher education and their effect on student success: A meta-analysis. *Educational Review*, *70*(2), 208–228.

Statistisches Bundesamt. (2021). *Hochschulen auf einen blick: Ausgabe 2020*. Wiesbaden.

Witteveen, D., & Attewell, P. (2021). Delayed time-to-degree and post-college earnings. *Research in Higher Education*, *62*, 230–257.

Zinn, S., Würbach, A., Steinhauer, H.-W., & Hammon, A. (2020). Attrition and selectivity of the NEPS starting cohorts: An overview of the past 8 years. *AStA Wirtschafts-Und Sozialstatistisches Archiv*, *14*, 163–206.

Appendix

Table A1. Sensitivity of Bunching Estimates to Specification Choices

	Partner University	NEPS (Universities)	NEPS (UAS)
	(1)	(2)	(3)
Panel A: Polynomial Degree 7			
(1) Baseline (Main Text)	0.423***	0.430***	0.410***
	(0.045)	(0.033)	(0.059)
(2) Structural Heaping Correction	0.463***	0.451***	0.415***
	(0.042)	(0.031)	(0.047)
(3) Two-Sided Bunching Window	0.473***	0.431***	0.398***
	(0.050)	(0.036)	(0.057)
Panel B: Polynomial Degree 5			
(4) Baseline	0.533***	0.495***	0.449***
	(0.041)	(0.034)	(0.049)
(5) Structural Heaping Correction	0.561***	0.502***	0.449***
	(0.036)	(0.029)	(0.038)
(6) Two-Sided Bunching Window	0.569***	0.503***	0.490***
	(0.064)	(0.036)	(0.063)
Observations	11,231	3,786	1,339

Notes: This table reports robustness checks for bunching estimates at the 30 credit points threshold across different datasets. Panel A uses a 7th-degree polynomial (main specification), while Panel B uses a 5th-degree polynomial. The bunching coefficient b represents excess mass at the threshold, normalized by the counterfactual density. Rows (1) and (4) report baseline estimates without heaping controls. Rows (2) and (5) include modulo-5 indicator variables to control for mechanical heaping. Rows (3) and (6) estimate a two-sided specification, allowing for mass to originate from both the left (-5, -1) and right of the threshold. Standard errors are calculated using bootstrap replications. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

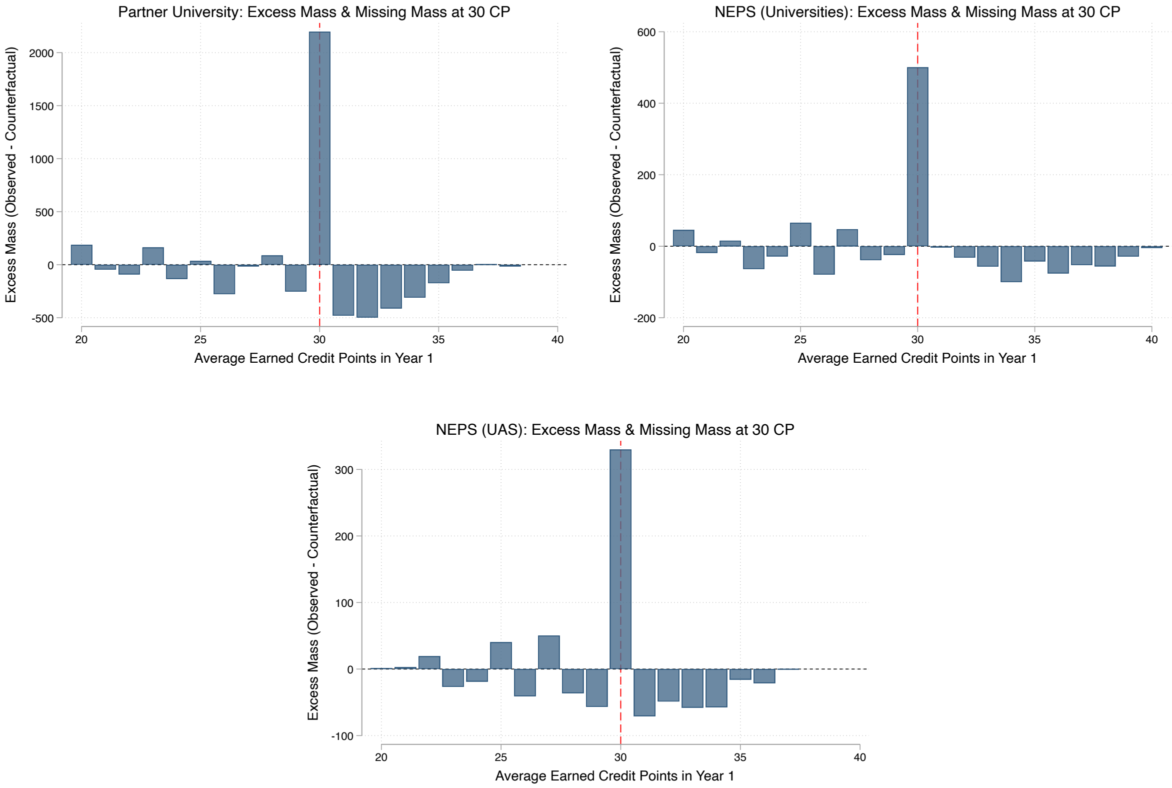


Fig. A1. Excess Mass and Missing Mass at 30 Credit Points Threshold. *Notes*: This figure displays the difference between the observed and counterfactual distributions of earned credit points across three samples: the Partner University (top left), NEPS Universities (top right), and NEPS Universities of Applied Sciences (bottom). The bars represent excess mass (calculated as observed frequency minus counterfactual frequency) as calculated from panel A of Table 1. The large positive value at the red dashed line (30 CP) indicates bunching, while the sequence of negative values to the right indicates missing mass—representing students who would have earned more than 30 CP absent the institutional constraint. The counterfactual distribution is estimated using a 7th-degree polynomial, excluding observations within a [-1, +1] window of the threshold.

Bunching in Earned Credit Points After 2 Years

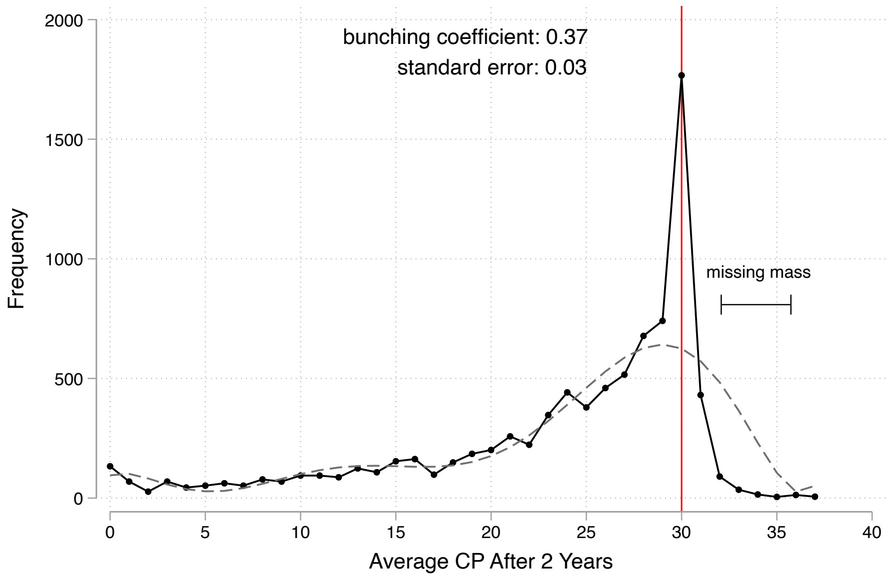


Fig. A2. Empirical and Counterfactual Distribution of Average Earned Credit Points in Year 1 and Year 2. *Notes*: The figures represent the observed and estimated counterfactual distributions of earned credit points in the first two academic years in the partner university. In the counterfactual figures, the black dots represent the observed number of students who earned a given number of credit points (actual density function), while the gray line represents the estimated counterfactual distribution of credit points. The difference between the two curves represents excess or missing mass accordingly. The red line represents the institutional reference point.

Table A2. Effect of Being Off-Track in Semester 1 on the Probability of Exceeding 30 Credit Points in Semester 2

	University Data	
	Exceeded 30 Credit Points in Semester 2	
	(1)	(2)
Offtrack Semester 1	0.105***	0.083***
	(0.005)	(0.005)
Controls	No	Yes
Constant	0.01	0.01
Observations	6,952	6,952

Notes: This table reports OLS estimates of the relationship between being off-track in the first semester (earning fewer than 30 credit points) and the likelihood of averaging more than 30 credit points per semester in the second academic year. The outcome variable is a binary indicator equal to 1 if a student exceeds 30 credit points in semester 2. Column (1) presents estimates without additional covariates; column (2) includes controls for age at enrollment, gender, high school GPA, type of high school diploma, international status, and transfer student status. All regressions are based on administrative data from the partner university, since the NEPS data does not report semester-level credits earned. Standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

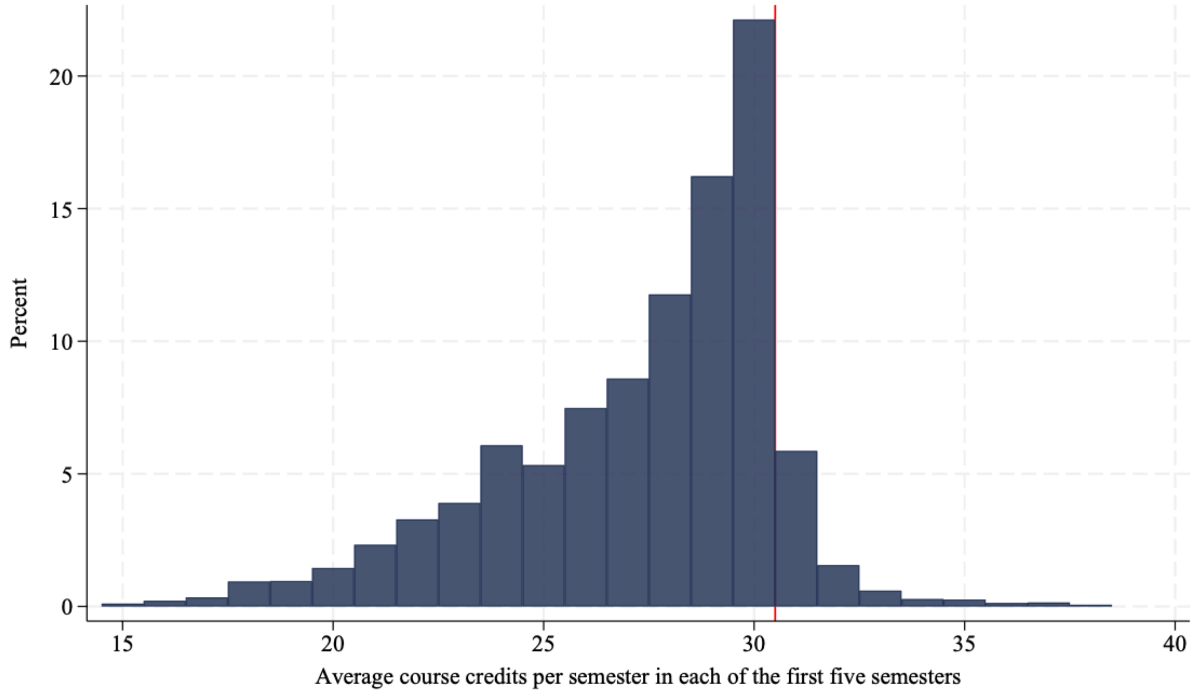


Fig. A3. Distribution of Average Credits Earned During the First Five Semesters.
Notes: This figure presents the distribution of average credits earned per semester over the first five semesters, the structured phase of study, for a sample of 4,691 students at our partner university who enrolled between 2014 and 2017 and graduated within five years.

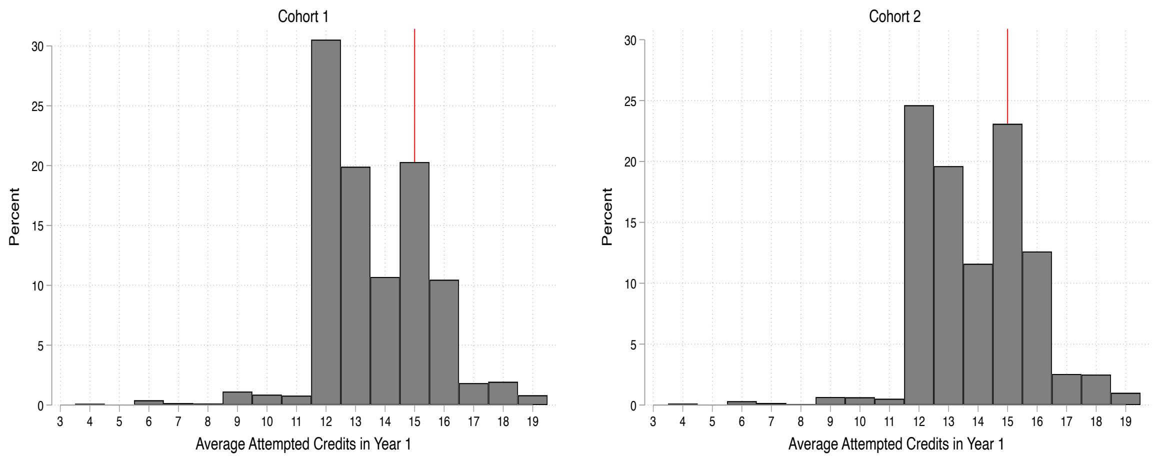


Fig. A4. Average Attempted Credit Hours Distribution in the USA. *Notes*: The figure consists of two histograms, representing the first-year average number of credit hours attempted by full-time students of two cohorts (2010 and 2011) in a large University in the USA. The institutional reference point in the United States is set at 15 credits per semester, while the full-time status threshold is at 12 credit hours.

Table A3. Annual Welfare Cost of Credit Bunching: Foregone Earnings

Component	Calculation	Amount
Share of Constrained Students	1,464 ÷ 11,231 students	13.0%
Behavioral response	From bunching estimation	2.2 CP/semester
Credit shortfall (5 semesters)	2.2 × 5	11 CP
Delay per affected student	11 ÷ 30 CP/semester	0.367 semesters
Average delay (population)	13.0% × 0.367	0.05 semesters
National Foregone Earnings (250,000 graduates)		
Delayed student-years	250,000 × 0.025 years	6,250 years
Median starting salary	Bachelor's degree holders	€40,000
Total annual welfare loss	6,250 × €40,000	€250,000,000

Notes: Calculation based on partner university data (Table 1). Average delay of 0.05 semesters per student reflects 13% affected by threshold × 0.367 semester delay per affected student. Earnings based on the median starting salary for bachelor's graduates.

Table A4. Balancing Table of the Field Experiment

	Control group	Treatment group	Difference
	(1)	(2)	(3)
Age	21.688	21.632	-0.057
	(2.981)	(3.053)	(0.850)
Female	0.444	0.455	0.011
	(0.498)	(0.499)	(0.824)
HS-grade	2.669	2.687	0.018
	(0.490)	(0.471)	(0.704)
Abitur	0.415	0.416	0.000
	(0.494)	(0.494)	(0.994)
First time student	0.386	0.406	0.019
	(0.488)	(0.492)	(0.688)
Enrolment date	10.September.2020	10.September.2020	0.081
	(11.522)	(8.141)	(0.935)
Credits transferred	1.005	0.965	-0.039
	(5.239)	(5.080)	(0.938)
Observations	207	202	

Notes: Columns (1) and (2) display the means in the control and treatment groups, respectively. Standard deviations in parentheses. Column (3) displays the difference between the first two columns. HS-grade refers to high school grade, which is according to the German grading system, where 1 represents the best grade and 4 is the worst grade. Abitur is a dummy variable referring to the share of students with the highest high school diploma type.

[redacted]

Vorname Nachname
Straße Hausnummer
PLZ Wohnort

Informationen für Ihre Studienplanung

Sehr geehrte/r Frau/Herr Vor- und Nachname,

wir begrüßen Sie herzlich an der Fakultät für Betriebswirtschaft. Dieses Schreiben soll Sie bei der Planung Ihres Studiums unterstützen und Ihnen dabei helfen, Ihr Studium gut zu strukturieren.

Der Standard-Studienplan orientiert sich gemäß europäischem ECTS-Referenzrahmen an „typischen" bzw. durchschnittlichen Studierenden und sieht für das erste Semester 30 Leistungspunkte vor. Dementsprechend treten auch nur 17% der Studierenden an unserer Fakultät im ersten Semester Prüfungen im Umfang von weniger als 30 ECTS-Punkten an.[1]

Hinweis der Fakultät BW für [redacted], Studiengang Bachelor BW, Matnr. [redacted]

Da der Studienplan sich an durchschnittlichen Studierenden orientiert, kann es für Sie sinnvoll sein, eine oder mehrere zusätzliche Prüfungen, und damit <u>mehr als 30 ECTS im ersten Semester abzulegen</u>.

Dies kann Ihnen beispielsweise dabei helfen,

- einen Puffer für die Einhaltung der Regelstudienzeit zu schaffen bzw.
- generell die Studiendauer zu reduzieren, oder
- die Arbeitsbelastung in späteren Semestern zu verringern.

Sollten Sie sich dazu entschließen, eine Prüfung vorzuziehen, bieten sich hierfür vor allem die Prüfungen [redacted] und das [redacted] an. Eine Liste dieser Fächer finden Sie auf der umliegenden Seite, ebenso finden sich dort weitere Informationen und Personen, die Sie bei Fragen unterstützen.

Wir wünschen Ihnen viel Erfolg und einen guten Start ins Studium.

Mit freundlichen Grüßen

Fig. A5. Treatment Intervention Letter. *Notes*: This figure presents the full text of the information intervention letter received by participants assigned to the treatment group. The letter was designed to encourage students to exceed 30 credit points in their first semester.