The Norwegian Labor Income Tax Credit Experiment — Pre-analysis plan: The effects on labor supply

Simon Bensnes[1], Øystein Hernæs[1], Morten Håvarstein[2], Simen Markussen[1], Magne Mogstad[2], Oddbjørn Raaum[1], Ragnhild Schreiner[3], and Gaute Torsvik[3]

[1]Frisch Centre
[2]University of Chicago / University of Oslo
[3]University of Oslo

July 3, 2026

Contents

1 General Information

1.1 Study Title

The Norwegian Labor Income Tax Credit Experiment: The effects on labor supply.

1.2 Trial Dates

- Trial start date: 2025-12-10
- Trial end date: 2030-12-31
- This pre-analysis plan was finalized and registered on July 3, 2026. At the time of registration, the research team had not had access to any individual-level data on treatment status or post-randomization outcomes. The only post-randomization information received consists of aggregate counts from the Norwegian Tax Administration: cell sizes by treatment arm and aggregate login rates by arm through the end of January 2026.

1.3 Researchers

See the title page.

1.4 Abstract

This pre-analysis plan specifies the analyses we will carry out to evaluate the Norwegian labor income tax credit experiment, including definitions of variables and outcomes, empirical hypotheses and specifications, sample definitions and attrition rules. Any deviations from this plan will be transparently documented. The experiment itself, including the intervention, the three treatment arms (A, B, C), the information treatment, and the randomization, is described in the companion document *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*. We refer to that document for all matters of experimental design and do not repeat them here.

2 Interventions

See *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*.

2.1 Implementation Institutions

The Norwegian Ministry of Finance and the Norwegian Tax Administration.

2.2 Country and Region

Norway.

2.3 Funding Sources

The Norwegian Ministry of Finance.

2.4 Description of Intervention Arms

See *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*.

3 Experimental Design

See *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*. Below we record only the facts and parameters relevant for the analysis plan.

3.1 Overview

The experiment randomly divides the target population into three groups, A, B and C. Group A (8 percent) is under the tax credit[1] rules and receives personalized information on labor tax and resulting employment incentives. Group B (46 percent) receives the same type of personalized information, for the ordinary tax-system without a tax credit. Group C (46 percent) does not receive any personalized tax-incentive information.

To be informed about whether one was randomly drawn to receive the tax credit, one had to log in to the tax-authorities' webpage. We define individuals who logged in and checked the tax credit status as the *informed* sample and those who did not as the *uninformed* sample.

For more details, see *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*.

[1] Formally, the scheme is a deduction (*arbeidsfradrag*) of up to NOK 125,000 in general income, worth 22 per cent of its amount in reduced tax. We refer to the scheme as a tax credit throughout, and use "deduction" when describing the statutory mechanics.

3.2 Full sample and study population

The *full sample*, defined by the Norwegian Tax Administration as everyone who was subject to the randomization, consists of approximately 1.33 million individuals born 1991–2006 with tax liability in Norway. The randomization stratifies on a discrete grouping built from tax-liability status, residence and pre-treatment income/benefit indicators; see the stratification table in *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*. Row 7 of that table collects individuals with *limited* tax liability (*begrenset skatteplikt*).[2]

The *study population* for this pre-analysis plan is the full sample net of row 7, i.e. the units in rows 1–6 of the stratification table. The aggregate cell sizes provided by the Tax Administration are:[3]

	Group A	Groups B+C
Rows 1–6 – ordinary tax liability (**study population**)	102,727	1,180,494
Row 7 – limited tax liability (excluded)	3,742	43,023
Full sample	106,469	1,223,517

The study population thus consists of approximately 1.28 million individuals, of whom 102,727 are in treatment group A. All analyses in this pre-analysis plan are conducted on this study population, or clearly specified subsamples.

Based on preliminary figures we have received from the Norwegian Tax Administration, approximately 63% of the study population had logged in by the end of January 2026 to see whether they were allocated to the tax credit rules (the informed sample). The login rate varies substantially with labor-market status – it is markedly lower among individuals not in employment and markedly higher among those who are – and this composition effect will in turn affect the precision of the informed-sample contrast for outcomes correlated with employment status.

3.3 Unit of Randomization

Individuals.

3.4 Informed sample: Awareness of treatment status and compliance with the information treatment.

The Norwegian tax authorities informed the full population, via SMS or e-mail, that they were part of the experiment, and that they needed to log in to the Norwegian Tax Administration's website to see whether they were eligible for the tax credit or not.

The invitation to log in (SMS / email) was identical across treatment arms, and the specific treatment status was revealed only after logging in. The decision to log in is thus expected to be orthogonal to treatment assignment.

[2] Typically foreign workers under the PAYE / kildeskatt scheme, or otherwise covered by § 6-86 (7) of the Tax Act, which excludes such taxpayers from the credit.

[3] These totals are slightly lower than the corresponding numbers in *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment* due to not including individuals registered as deceased after the tax authorities defined the full population and some individuals who had invalid identification numbers.

After login, all participants are informed about the experiment and whether they are eligible or not for the tax credit. If drawn to groups A or B, they are also exposed to an information treatment; see *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment* for details.

Login error. When the experiment was implemented, an error on the tax administration's webpage caused individuals in group C who logged into the *inbox* ("innboks") – and not *my tax* ("min skatt") – to not be directly offered a link to the experiment information page. This error will be handled in the analyses and did not affect members of A and B.

3.5 Sample Size, statistical precision and inference strategy

Sample size. Number of randomization clusters (strata): approximately 13,300, each containing exactly 100 individuals.[4] Treated individuals within the study population (group A): 102,727. Total study population: approximately 1.28 million.

Original power calculations. The randomized design draws on a pre-study report by the research group, prepared at an early stage before the experiment was politically decided and before the implementing parties were involved.[5] That report used pre-existing register data for the relevant age cohorts to compute approximate confidence intervals under several design alternatives by means of simulations (different deduction sizes, alternative numbers of treated individuals, and individual vs. municipality randomization). The key conclusions relevant for the implemented design were:

- With approximately 100,000 treated individuals under stratified individual randomization, the expected confidence interval on the participation rate effect was on the order of 0.4 percentage points.
- With approximately 100,000 treated and a deduction of size $1\,G$, the expected confidence interval on the participation elasticity was approximately 0.2.[6]
- With approximately 100,000 treated, the expected confidence interval on labor earnings (yrkesinntekt) conditional on a prior-year earnings cutoff was around 5,000 NOK.

These figures supported the recommendation of at least 100,000 treated individuals; the implemented design (8% of the 1991–2006 cohort, $\approx 106{,}500$ treated) reflects that recommendation. Both the recommendation and the subsequent decision to set the trial at this scale also reflected the need to retain enough precision for subgroup and heterogeneity analyses, and a precautionary margin against the substantial uncertainties involved in the practical implementation of a tax-system experiment of this kind.

Updated precision estimates. After the design of the tax credit was finalized, we refined the report's ex-ante power calculations using historical data and the variance properties of the estimator we will use in the analysis. Our baseline estimator regresses the outcome on the treatment indicator and strata fixed effects only (see Section 6). Because strata are built on cohort, gender, residency, tax-liability status, and detailed pre-treatment income and benefit history (see *The Norwegian Labor Income Tax Credit Experiment — Description of the*

[4]The exceptions are 224 residual strata, one within each of the groups defined by the categorical variables (birth year, gender, and indicators for previous incomes), each containing fewer than 100 individuals. See *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment* for the full stratification protocol.

[5]https://frischsenteret.no/wp-content/uploads/2025/10/Utredning-om-forsoksordning-med-arbeidsfradrag.pdf

[6]1G equals 1 "basic income unit" which is used as a basis for calculations of benefits by the Norwegian Welfare Administration. In 2026 1 G = NOK 130,160.

experiment for the full stratification protocol), strata fixed effects absorb a substantial share of the variation in current outcomes. The standard error of the estimated treatment effect from comparing a treated group of size N_1 to a control group of size N_0 is well approximated by

$$SE(\hat{\beta}) \approx \sigma\sqrt{1-R^2}\sqrt{\frac{1}{N_1}+\frac{1}{N_0}},$$

where σ is the unconditional standard deviation of the outcome (for binary outcomes $\sigma = \sqrt{p(1-p)}$) and R^2 is the share of Var(Y) explained by strata fixed effects. To estimate the design's statistical precision, we have, as far as possible, implemented the stratification protocol on historical 2022 register data and estimated the relevant residual variances directly. On that basis we estimate $R^2 \approx 0.79$ for continuous annual earnings (Y^W, equivalently $\sigma_{\text{res}}/\sigma \approx 0.46$) and $R^2 \approx 0.59$ for binary participation ($\sigma_{\text{res}}/\sigma \approx 0.64$).

Table 1 reports the resulting SEs and minimum detectable effects (MDEs at 80% power, two-sided 5% significance) for two illustrative outcomes – participation $P^W = \mathbf{1}[\texttt{wyrkinnt} > 0]$ and earnings $Y^W = \texttt{wyrkinnt}$ – under three estimator alternatives.

	Contrast / Sample		
	A vs. B informed sample (63% login)	A vs. B unconditional	A vs. (B+C) unconditional
$N_1 + N_0$	64,700 + 371,900	102,700 + 590,200	102,700 + 1,180,500
Participation $P^W = \mathbf{1}[\texttt{wyrkinnt} > 0]$, $p = 0.89$:			
$SE(\hat{\beta})$	0.08 pp	0.07 pp	0.06 pp
MDE (80% power)	0.23 pp	0.19 pp	0.18 pp
Earnings $Y^W = \texttt{wyrkinnt}$, $\sigma = 2.89\,G$:			
$SE(\hat{\beta})$	0.006 G	0.004 G	0.004 G
MDE (80% power)	0.015 G	0.012 G	0.012 G

Table 1: Approximate standard errors and minimum detectable effects under three estimator alternatives. SEs use $\sigma\sqrt{1-R^2}$ as the residual standard deviation, where R^2 is the share of variance in Y explained by strata fixed effects: $R^2 \approx 0.79$ for Y^W and $R^2 \approx 0.59$ for P^W, estimated by implementing the stratification protocol on 2022 register data. "Informed sample" assumes a 63% login rate symmetric across treatment arms A and B. Population mean p and outcome standard deviation σ are based on own calculations on register data for residents aged 20–35 in 2022 and 2023. MDE evaluated at 80% power, two-sided 5% significance. $1\,G$ = NOK 130,160 for 2026.

The informed-sample restriction (column 1) raises SE by about 25% relative to the unconditional A vs. B comparison (column 2); pooling B and C as controls (column 3) provides only marginal further precision gains.

The above figures are somewhat narrower than the report's ex-ante calculations because the implemented sample is larger, and the stratification protocol used in the actual design is finer. The calculations do not take into account that the informed sample is not randomly selected within groups A and B.

3.6 Randomization Method

See *The Norwegian Labor Income Tax Credit Experiment — Description of the experiment*.

4 Data Collection and Management

Throughout this section we distinguish between three related income concepts:[7]

- Y^{TC}: the statutory calculation base for the credit (skatteloven § 6-86; defined in detail in Section 4.2.3).
- Y^{W}: active labor income, `WYRKINNT`.
- Y^{KL}: short-term wage proxy, `KONTANT_LONN` from the A-melding.
- Y^{A}: taxable general income, i.e. alminnelig inntekt less the personal allowance. $Y^{A} = \max(0,$ alminnelig inntekt$-$ personfradrag$)$.

The Experiment Description provides an item-by-item comparison of these concepts together with alminnelig inntekt and personinntekt.

4.1 Data Sources

The experiment uses data from two main sources, which are linked by replacing person identifiers with pseudonymized, project-specific identifiers:

- **Statistics Norway (SSB):** provides the longitudinal administrative registries, including demographics, education, the employer-employee registry (A-meldingen), receipt of benefits, the business registry, and annual tax data. These data sources are needed both for the short-term outcomes (A-meldingen) and for reconstructing Y^{TC} and the long-term outcomes (yearly tax data and benefit registries from the Norwegian Labour and Welfare Administration (NAV)). Documentation for available registries: https://www.ssb.no/en/data-til-forskning/utlan-av-data-til-forskere/.
- **The Norwegian Tax Administration:** operates the randomization and provides data on strata, stratification variables, the outcome of the randomization, individual login data, and (after the relevant tax year is finalized) the actual tax credit assigned to each treated individual.

4.2 Outcome Variables

A main motivation for the experiment is to increase labor market participation and employment among young people. We therefore study earnings, participation, hours, and tax-return outcomes. Because finalized tax data are only observable with a substantial lag, we organize the outcomes into a *short-term track* (A-melding data made available for delivery to research projects by SSB approximately 3–4 months after the relevant period) and a *long-term track* (data from the tax assessment, typically available approximately 15 months after the end of the income year). Both primary and secondary outcomes will be studied throughout the experimental period.

4.2.1 Primary outcome: earnings.

Short-term track (A-melding-based): earnings.

$$Y_{i,t}^{KL} = \text{KONTANT\_LONN}_{i,t},$$

[7] Throughout this plan, all paragraph references of the form "§ X-Y" are to the Norwegian Tax Act (*skatteloven*, lov 1999-03-26 nr. 14) unless otherwise specified; references to the National Insurance Act are prefixed "ftrl."

i.e. wage income reported via the A-melding for individual i in month t. The data are reported month by month, but are made available for delivery to research projects by SSB with an approximate lag of 3–4 months; we expect to receive quarterly updates. Y^{KL} is a strictly narrower proxy for the statutory base Y^{TC}: it omits self-employment income (skatteloven § 12-2 g, f) and any wage-replacement benefits not channeled via the employer (e.g. sickness benefits paid directly by NAV, parental benefits paid directly by NAV).

Short-term track (A-melding-based): participation.

$$P_{i,t}^{KL} = \mathbf{1}[Y_{i,t}^{KL} > 0],$$

i.e. the extensive-margin indicator that the individual had any wage income reported via the A-melding in period t.

Long-term track (finalized tax data): earnings (behavioral).

$$Y_{i,t}^{W} = \text{WYRKINNT}_{i,t},$$

the sum of wage income and self-employment income for individual i in income year t, prepared by SSB on the basis of finalized tax-return data. This is a measure of *active labor income* and forms the primary behavioral outcome of the experiment.

Long-term track (finalized tax data): participation.

$$P_{i,t}^{W} = \mathbf{1}[Y_{i,t}^{W} > 0],$$

i.e. the extensive-margin indicator that the individual had positive active labor income (`wyrkinnt`) in income year t.

4.2.2 Secondary outcomes

- **Hours:** weekly work hours (`PERS_SUM_ARBEIDSTID`) from A-meldingen.
- **Tax-return outcomes:** total income (`SAMINNT`), assessed taxes (`FASTSATT_SKATT`), taxable transfers and benefits (`WSKPL_OVERF`), non-taxable transfers and benefits (`WSKFRIE_OVERF`), and income after tax (`IES`).

4.2.3 Incentive characterization: calculation base Y^{TC} and realized credit

The variables in this subsection are not labor-supply outcomes in their own right; they characterize the incentives that treated individuals face and the realized tax saving they receive. They are inputs to the heterogeneity and model-based analyses in Section 6 and to the interpretation of the reduced-form effects on Y^{W} and P^{W}.

Calculation base Y^{TC}. The statutory calculation base under skatteloven § 6-86 third paragraph is the sum of (i) labor income under § 5-10 excluding unemployment benefits (§ 5-10 c nr. 1), (ii) calculated personal income from sole proprietorship under § 12-2 g, and (iii) compensation to participants in partnerships under § 12-2 f. The Experiment Description provides an item-by-item breakdown of Y^{TC} alongside the other income concepts used in this plan, and Section 4.3.1 describes the construction from registry components.

Realized credit and effectively utilized deduction. The granted deduction is a function of Y^{TC} alone and follows the schedule fixed by Stortingets skattevedtak § 6-6:

$$\text{TC}(Y^{TC}) = \begin{cases} Y^{TC} & 0 \leq Y^{TC} < 125\,000, \\ 125\,000 & 125\,000 \leq Y^{TC} < 345\,000, \\ 125\,000 - 0.40\,(Y^{TC} - 345\,000) & 345\,000 \leq Y^{TC} < 657\,500, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

The deduction is granted in alminnelig inntekt and cannot generate a carry-forward loss (skatteloven § 6-86 fourth paragraph). Letting $Y^A_{i,t}$ denote individual i's taxable general income (alminnelig inntekt less personfradrag) in year t, the *effectively utilized* deduction is

$$\tilde{\text{TC}}(Y^{TC}_{i,t}, Y^A_{i,t}) = \min\big(\text{TC}(Y^{TC}_{i,t}),\ \max(0, Y^A_{i,t})\big), \tag{2}$$

and the realized tax saving for treated individual i in year t is

$$\Delta\text{tax}_{i,t} = 0.22 \cdot \tilde{\text{TC}}(Y^{TC}_{i,t}, Y^A_{i,t}) \cdot \mathbf{1}[D^A_i = 1], \tag{3}$$

since the credit reduces alminnelig inntekt and the income tax on alminnelig inntekt is levied at 22%. Maximum realized tax saving is $0.22 \times 125\,000 = 27\,500$ NOK per year. Where available, we use the realized credit as actually assigned in the tax assessment, supplied as a registry variable; otherwise we use the reconstructed value from equation (3).

4.3 Data Processing and Cleaning Rules

We apply the following rules to process the data for analysis:

- **Aggregation:** when using monthly data (e.g. A-meldingen), we aggregate the available months into one observation per person by taking the average over the available months within a year. For expositional reasons, monthly averages may be presented as such or multiplied by 12 as a proxy for annual earnings.
- **Deflation:** all nominal values are deflated using the Basic Amount (*Grunnbeløpet*, G), the reference unit set annually by the Norwegian Labour and Welfare Administration (NAV) on the basis of wage growth and used to index pensions and most social-insurance benefits. We use the average value of G within a given year and apply this to all observations in that year.
- **Nominal vs. real treatment intensity:** The tax credit's parameters (the NOK 125 000 maximum deduction and the phase-out thresholds at NOK 345 000 and NOK 657 500) are set in nominal terms in the Storting's annual tax resolutions. Absent legislative adjustment, the real generosity of the credit (measured in G) declines mechanically over the experimental window. We will note this when interpreting event-study coefficients (equation (7)).
- **Negative values:** `KONTANT_LONN`, `WYRKINNT`, `PERS_SUM_ARBEIDSTID`, and `SAMINNT` are set to 0 for negative values.
- **Winsorizing:** `KONTANT_LONN` and `WYRKINNT` are winsorized at NOK 20 million (also when transformed to annual proxies); we expect this to apply to less than 0.001 percent of observations. Weekly work hours are winsorized at 200 percent of full-time (2 × 37.5 hours/week), expected to apply to less than 0.05 percent.

- **Missing data:** values for income and working hours for persons not present in either the income registry or A-meldingen are assumed to be 0. Missing covariates are replaced with indicator-specific constants (unique dummies not otherwise observed), and an auxiliary indicator equal to 1 if the original value is missing is included as a control.

4.3.1 Construction of Y^{TC} from registry components

The statutory calculation base Y^{TC} is not a single SSB variable. We construct it as the sum of the following components, each drawn from the corresponding registry:

1. Wage and similar compensation (skatteloven § 5-10 a, b, d, e): from A-meldingen reported items, including `KONTANT_LONN`, severance pay, benefits in kind (*naturalytelser*), etc., and from the tax assessment for items not reported via A-meldingen.

2. Wage-replacing benefits in skatteloven § 5-10 c except no. 1: sickness benefits (folketrygdloven kap. 8), caregiver/attendance/training benefits (kap. 9), and parental benefits (kap. 14). These are taken from A-meldingen when paid by the employer and from NAV's benefit registries when paid directly to the recipient. Unemployment benefits (§ 5-10 c nr. 1) are excluded.

3. Calculated personal income from sole proprietorship (skatteloven § 12-2 g): from the tax assessment.

4. Compensation to participants for active participation in companies with participant assessment (skatteloven § 12-2 f): from the tax assessment.

For self-employed taxpayers, any carry-forward of negative calculated personal income from prior years reduces the calculation base, in accordance with Prop. 1 LS (2025–2026) point 4.4.4. The constructed Y^{TC} will be validated against the corresponding information recorded in the tax assessment for treated individuals, and any systematic discrepancies will be reported.

4.4 Attrition

Individuals are defined as attriting if they die or emigrate, as recorded in the SSB Population Register. Such individuals are excluded from the sample from the month of death or emigration onward. Individuals who later re-immigrate are re-included in the sample from the month of immigration. We will run tests for differential attrition, as in the main model.

4.5 Non-Compliance

Compliance with the deduction treatment. Compliance with the deduction is mechanical: the tax system automatically applies the credit to treated individuals when their reconstructed $Y^{TC} > 0$ and they are taxable on general income (i.e. $Y^{A} > 0$). However, individuals may be unaware of the treatment, and the credit only produces a tax saving for those whose general income is positive after all other deductions and the credit itself. As reported in Prop. 1 LS (2025–2026) chapter 4, an estimated 59 % of the treated cohort is expected to derive no benefit from the credit absent a behavioral response (11 % with no labor or business income; 17 % with too low total income to be taxable on general income; 31 % with income above the phase-out endpoint).

5 Estimation sample and balance checks

5.1 Estimation samples

Different combinations of groups A, B, and C, together with the distinction between those who logged in and visited the information page and those who did not, give rise to different estimation samples that are informative on different questions. We refer to the sample of those who logged in and visited the information page as the *informed sample*.

Definition of the informed sample. We define an individual as *informed* if he/she has logged into the Norwegian Tax Administration's webpage and clicked the banner for the tax experiment. According to figures we have received from the Norwegian Tax Administration, approx. 63 percent had logged in before the end of January, both in group A and B. We expect the fraction informed to grow through the year. When constructing the informed sample, we will set a cut-off date such that we include everyone logged in before this date.

The cut-off date will be chosen to make the informed sample as large as possible, subject to informed status remaining balanced across treatment arms. The reason a cut-off is needed is that treatment status itself may affect login behavior over time: treated individuals face lower tax withholding from January 2026 onward, and noticing this in their payslips may prompt logins that have no counterpart in group B. Late logins may therefore be selectively treated.

Main specification. Our main specification compares group A to group B, conditional on the informed sample. This comparison provides the cleanest economic interpretation of the tax incentive: both groups receive the same type of information, so the contrast isolates the effect of the tax credit holding information constant. The restriction to the informed sample also ensures that the login error described in Section 3.4 does not introduce selection bias, since both groups A and B were exposed to the tax-administration's webpage in the same way.

Supplementary specifications. In supplementary analyses we will:

1. **Pool the control groups (B and C).** If the information treatment given to group B has no measurable effect, pooling B and C increases precision. Whenever C is included in the control group, the informed sample must be defined consistently across A, B, and C; we then restrict the informed sample to those who logged into "min skatt" and clicked the banner for the tax experiment, excluding individuals who first entered "innboksen" (the inbox). This avoids contamination from the login error described in Section 3.4.

2. **Re-weight the informed sample to represent the full study population.** Conditioning on the informed sample gives a sample that is not representative of the full study population: those who logged in may differ from those who did not on observable as well as unobservable characteristics. To extrapolate the informed-sample estimate toward a population-wide quantity, we re-weight observations using inverse-probability weights estimated from observables, so that the weighted informed sample matches the marginal distribution of these observables in the full study population. This complements the main A vs. B specification with an estimate that is more representative of the full target population, conditional on the assumption that selection into login is on observables.

3. **Estimate experimental effects without conditioning on login.** We distinguish three unconditional contrasts:

- **A vs. B**: *full-sample credit effect*. The assignment effect of the credit, holding the offer of information about the incentives implied by the tax system constant. Identifies the effect of being assigned the credit, abstracting from the information shock that accompanies any first-time introduction, for both aware and unaware individuals combined.
- **A vs. C**: *rollout effect*. The combined effect of the credit and the initial information shock. This is what would be observed in the short run in a real-world rollout, where the new policy and the news of it reach individuals at the same time.
- **B vs. C**: *information effect*. The behavioral response from providing personalized information regarding the incentives implied by the tax system.

By construction, $\mathrm{A} - \mathrm{C} = (\mathrm{A} - \mathrm{B}) + (\mathrm{B} - \mathrm{C})$. Under the assumption that the information shock is transient ($\mathrm{B} - \mathrm{C} \to 0$ as the credit becomes a known feature of the tax system), the rollout effect converges to the structural credit effect in the long run. **A vs. B** therefore serves as an estimate of the long-run rollout effect from short-run experimental data, contingent on this transience assumption; **A vs. C** is the short-run rollout estimate. Empirically estimating **B vs. C** tests the assumption.

5.2 Randomization into groups A, B, and C

As an implementation diagnostic for the random assignment, we assess balance in pre-assignment characteristics measured before October 2025 – prior to the announcement in the national budget on 15 October 2025 and before the draw was made. Balance tests take the form:

$$\begin{aligned} D_i^A &= \sigma_{s(i)} + X'\beta_1 + \epsilon, \\ D_i^B &= \sigma_{s(i)} + X'\beta_2 + u, \end{aligned} \tag{4}$$

where D_i^A (D_i^B) is a dummy for an individual being in group A (B), $\sigma_{s(i)}$ indicates strata fixed effects, and X is a vector of individual characteristics that should not be correlated with the probability of being drawn into any treatment arm. The tests are against group C. We include the variables used to construct the strata, and additionally include variables not used in stratum construction such as average income in the municipality of residence and educational attainment.

By estimating equations (4) we test separately whether individuals are more likely to be assigned to either treatment arm. We will report joint significance of all variables in X and individual significance for each variable.

5.3 Login / Compliance / Beliefs

The Norwegian Tax Administration provides detailed clickstream data tracking individual logins and specific page views within the personal tax portal. By merging these logs with administrative records, we evaluate whether the intervention influenced portal engagement.

We estimate the effect of the intervention on portal engagement by regressing an indicator for clicking on the link to the information page (CLICK) on treatment assignment and strata fixed effects. Assignment to group A is denoted by the indicator D_i^A and assignment to group B by D_i^B. The individual is indexed by i, the time period by t, and the stratum of individual i by $s(i) \in \{1, \ldots, J\}$. Strata fixed effects are written compactly as $\sigma_{s(i)}$.

$$\text{CLICK}_i = \sigma_{s(i)} + \beta_A D_i^A + \beta_B D_i^B + \varepsilon, \tag{5}$$

where $\sigma_{s(i)}$ are the strata fixed effects.

We estimate equation (5) on three samples and test, in each, whether $\beta_A = \beta_B$:

- All individuals in the study population.
- Individuals who logged in first via "min skatt".
- Individuals who logged in first via "innboks".

The split by entry route is motivated by the login error described above, which affected only individuals in group C entering through "innboks". Reporting the three samples separately allows the reader to see the magnitude of any asymmetry between the groups due to the login error.

We also collect data through two surveys on tax beliefs (perceived tax rates), as well as on whether respondents are aware of being entitled to the tax credit (belonging to group A) or not. In the analysis of the tax credit, these surveys let us test whether the credit actually shifts perceived tax rates and whether behavioral responses are concentrated among those aware of their eligibility. The surveys are conducted by Statistics Norway, both before (November 2025) and after (scheduled October and November 2026) the onset of the experiment. All three groups (A, B and C) are randomly sampled to participate in the survey. See the companion Pre-analysis plan for *Information Frictions in Taxation and Labor Supply* for details.

6 Empirical analysis

Notation. Follows notation in Section 5.3. In regressions comparing group A with B (or with B+C), we write $D_i \equiv D_i^A$ for the credit-assignment indicator. Outcome variables are generically denoted Y for continuous outcomes (earnings, hours, tax liability, etc.) and P for binary participation. The specific income and outcome variables that recur are: Y^{KL} (short-term wage), Y^W (long-term active labor income), Y^{TC} (statutory calculation base), Y^A (taxable general income before the credit), P^{KL} and P^W (short- and long-term participation), and Δtax (realized tax saving). All are defined in Section 4.2.

6.1 Reduced form analysis

Primary confirmatory outcomes. The four primary confirmatory outcomes for this pre-analysis plan are Y^{KL}, P^{KL} (short-term track, available approximately 3–4 months after the relevant period) and Y^W, P^W (long-term track, available approximately 15 months after the end of the income year), all estimated on the A vs. B informed-sample contrast. When data for both short- and long-term tracks are available for a given period, Y^W and P^W take precedence as the primary outcomes for that period.

Main estimates. For each outcome in each period we separately estimate the following regression for group A vs. group B in the informed sample:

$$Y_i = \beta\, D_i + \sigma_{s(i)} + u_i, \tag{6}$$

where $\sigma_{s(i)}$ are strata fixed effects ($s(i)$ denotes the stratum of individual i). We do not include additional pre-period covariates in the baseline specification; their inclusion is reported as a robustness check.

The coefficient β is the average treatment effect of the tax credit on the outcome. We estimate equation (6) for both primary outcomes Y and P, and report both conventional (unadjusted) p-values and p-values adjusted for multiple hypothesis testing based on the approach of Romano and Wolf (2005a, 2005b, 2016), controlling the family-wise error rate at 5%.

We also estimate the model for the secondary outcomes specified in Section 4.2.2 (short-term and long-term earnings and participation; hours; tax-return outcomes). We also estimate it on the variables that characterize the incentives, defined in Section 4.2.3, (Y^{TC} and Δtax); these are mechanical reach-of-the-credit quantities, useful for documenting take-up and the legal coverage of the credit.

Treatment dynamics. The five-year duration of the experiment, together with the fact that behavioral responses to incentives may unfold gradually, motivates an event-study version of equation (6) that allows the treatment effect to vary by time. Our preferred specification uses strata-by-time fixed effects $\sigma_{s(i),t}$, the most flexible feasible setup: it absorbs any time-varying differences across strata and identifies the treatment effect from within-stratum, within-period variation in treatment status. The event-study regression is

$$Y_{i,t} = \sum_{\tau \in \mathcal{T}} \beta_\tau \, D_i \cdot \mathbb{1}[t = \tau] + \sigma_{s(i),t} + u_{i,t}, \tag{7}$$

where $\mathcal{T}$ indexes the periods of available post-treatment data and includes the pre-period for placebo checks. The event-study regression is estimated both at *annual* resolution for the long-term outcomes (Y^W, P^W) and at *monthly* resolution for the short-term A-meldingen outcomes (Y^{KL}, P^{KL}). We will report the path $\{\beta_\tau\}$ both for the experimental window 2026–2030 and, where data are available, for post-experiment years to capture any persistence or reversal once the credit is removed.

In all panel specifications, standard errors are clustered at the individual level to accommodate within-individual correlation across the repeated observations $\{(i,t)\}_t$.

6.2 Heterogeneity and distributional effects

We examine heterogeneity in the credit's effect along three dimensions: pre-specified covariates (6.2.1), the earnings distribution (6.2.2), and incentive groups derived from the credit's heterogeneous incentive structure (6.2.3 and following).

6.2.1 Heterogeneity by other covariates

We estimate equation (6) separately for the following pre-specified subgroups, each defined on covariates measured prior to treatment assignment:

- **Gender:** male, female.
- **Age groups:** four-year birth-cohort bins.
- **Student status:** registered students vs. non-students.[8]
- **Benefit receipt:** any social-insurance benefits in 2025 vs. no benefits.

We report group-specific treatment effects with both conventional (unadjusted) p-values and p-values adjusted for multiple hypothesis testing based on the approach of Romano and Wolf (2005a, 2005b, 2016), controlling the family-wise error rate at 5%. We consider each subgroup as a family of subgroup contrasts (e.g. all gender contrasts as one family, all age contrasts as another).

[8]Student status is defined as pursuing exams equivalent to at least 15 ECTS during the autumn semester of 2025 (i.e. at least 50% course load) with expected degree completion in 2026 or later.

6.2.2 Quantile treatment effects and tests for distributional equality

Following Bitler et al. (2006), we estimate quantile treatment effects of the credit, both unconditional and separately for those who participate and do not participate in the pre-period. We plot the estimated QTE function across quantiles of the outcome distribution.

We additionally perform a Kolmogorov-Smirnov test of the null hypothesis that the marginal distributions of the outcome under treatment and control are equal. The test statistic is the maximum vertical distance between the empirical CDFs; p-values are obtained by bootstrap. We perform the KS test on the full A vs. B (informed) sample and conditional on positive lagged earnings.

6.2.3 Heterogeneity and distributional effects by incentive groups

The tax credit alters work incentives heterogeneously across the earnings distribution. As shown in Section 4.2.3, the realized tax saving Δtax depends jointly on the calculation base Y^{TC} and on taxable general income Y^{A}; together these determine the marginal, average and participation tax-rate changes induced by the credit. Five mutually exclusive incentive regions arise from this dependence, summarized in Table 2.

Region	Y^{TC} range	MTR change	ATR change	PTR change
BELOW	$Y^{A} = 0$ (no taxable general income)	–	–	↓
PHASE-IN	$[0,\ 125\,000)\ \approx\ [0,\ 0.96\,G)$	↓	↓	↓
PLATEAU	$[125\,000,\ 345\,000)\ \approx\ [0.96,\ 2.65)\,G$	0	↓	↓
PHASE-OUT	$[345\,000,\ 657\,500)\ \approx\ [2.65,\ 5.05)\,G$	↑	↓	↓
ABOVE	$\geq 657\,500\ \approx\ \geq 5.05\,G$	0	0	0

Table 2: Incentive regions of the tax credit. The five regions form a partition: BELOW comprises individuals with no taxable general income ($Y^{A} = 0$, so the deduction generates no realized tax saving regardless of Y^{TC}); the rows PHASE-IN, PLATEAU, PHASE-OUT, and ABOVE apply to individuals *with* $Y^{A} > 0$ and use the indicated Y^{TC} range. "MTR change" and "ATR change" are the changes in the marginal and average tax rate on general income induced by the credit. "PTR change" is the change in the participation tax rate (the share of additional gross earnings lost to taxes upon entering employment) for an individual currently at the given region. PTR falls in all four regions where the credit operates (BELOW, PHASE-IN, PLATEAU, PHASE-OUT) because the credit reduces the tax that the individual would owe upon entering employment that lands them in the credit's earnings range; the magnitude is largest in the PHASE-IN region and declines through PLATEAU and PHASE-OUT. $1\,G = 130\,160$ NOK for 2026.

The treatment effect on labor supply is therefore expected to differ across individuals, primarily as a function of where their counterfactual (untreated) position would place them in Table 2. This subsection examines that heterogeneity along three axes: average treatment effects within incentive groups, the credit's effect on the realized earnings distribution (distributional, ex post), and the joint distribution of lagged and realized groups (the transition matrix induced by the treatment). Section 6.2.4 subsequently translates the resulting participation responses into elasticities by incentive group.

We construct groups from 2025 data on Y^{TC} and Y^{A} and estimate effects on earnings for each group. We use two classifications:

1. Deciles of 2025 labor earnings[9], with a separate group for individuals who had zero earnings.

[9] Y^{W} if available, otherwise Y^{KL}

2. The five incentive groups of Table 2 (BELOW, PHASE-IN, PLATEAU, PHASE-OUT, ABOVE), constructed by mapping Y^{TC} and taxable general income Y^A from 2025 to the regions defined there.

We estimate the analogue of equation (6) interacting the treatment indicator with dummies for each group:

$$Y_i = \sigma_{s(i)} + \sum_{d=1}^{K} \alpha_d \, \mathbb{1}[\hat{Q}_i = d] + \sum_{d=1}^{K} \gamma_d \, \mathbb{1}[\hat{Q}_i = d] \cdot D_i + \varepsilon_i. \tag{8}$$

The γ_d coefficients give the average treatment effect within each group. We test joint significance via an F-test of $\gamma_1 = \cdots = \gamma_K = 0$. We report group-specific treatment effects with both conventional (unadjusted) p-values and p-values adjusted for multiple hypothesis testing based on the approach of Romano and Wolf (2005a, 2005b, 2016), controlling the family-wise error rate at 5%.

Treatment effects on the realized income distribution. We map the credit's effect on the entire shape of the realized earnings distribution by estimating a linear probability model for the indicator that individual i's realized earnings fall in earnings bin j. Let $\{\text{bin}_j\}_{j=1}^{B}$ denote a partition of the earnings line; for each j we estimate

$$\mathbb{1}[Y_i \in \text{bin}_j] = \sigma_{s(i)} + \beta_j D_i + \varepsilon_{i,j}. \tag{9}$$

The coefficient β_j captures the change in the probability mass within bin j induced by the credit. By construction $\sum_{j=1}^{B} \beta_j = 0$, since each individual must realize an outcome in exactly one bin. We will use bins corresponding to the incentive regions (BELOW, PHASE-IN, PLATEAU, PHASE-OUT, ABOVE).

6.2.4 Participation elasticities

We compute average participation elasticities as the ratio of the experimental participation response to the mechanical change in the net of participation tax rate (PTR). The PTR is the share of additional gross earnings that an individual loses to the tax-and-transfer system upon entering employment:

- **Participation tax rate** Letting $B(\cdot; d)$ denote the level of social-insurance and means-tested benefits the individual receives under treatment status d as a function of labor earnings, we define the participation tax rate (PTR) as:

  $$\tau^p(y; d) \equiv \frac{T\big(Y^A(y), Y^{TC}(y); d\big) - T\big(Y^A(0), Y^{TC}(0); d\big) + B(0; d) - B(y; d)}{y}.$$

  This captures the full effective participation tax rate, including the loss of benefits due to income-testing when the individual enters employment.

The average participation elasticity is given by[10]

$$\varepsilon^p = -\frac{\mathbb{E}[P \mid D=1] - \mathbb{E}[P \mid D=0]}{\mathbb{E}[P \mid D=1]} \times \frac{1 - \mathbb{E}[\tau^p(Y;1) \mid D=1,\, P=1]}{\mathbb{E}[\tau^p(Y;1) \mid D=1,\, P=1] - \mathbb{E}[\tau^p(Y;0) \mid D=1,\, P=1]}. \tag{10}$$

The PTR averages condition on $P = 1, D = 1$ – i.e. they are taken over treated workers – following the convention that participation elasticities are defined relative to the PTR faced by those who actually work.

Note that $B(0; d)$ is not observed in the data. It will be approximated by applying a tax-transfer calculator to data on individuals' earnings and characteristics. We will examine the accuracy of the approximation on pre-treatment data and the sensitivity of the results to approximation bias.

[10] Y is Y^W if available, otherwise Y^{KL}

6.2.5 Intensive margin and the affected subpopulation

Earnings effect on the affected. Standard labor supply theory predicts that only those who choose earnings in a region where the credit affects their budget constraint should respond to the credit. The "affected" subpopulation consists of individuals who, in either the treated or the control state, locate in the PHASE-IN, PLATEAU, or PHASE-OUT regions. Under the model assumptions detailed in Section 6.3 and Appendix A, we recover the marginal earnings distributions $F_{Y(1)|\text{affected}}$ and $F_{Y(0)|\text{affected}}$ for this subpopulation. We report the average earnings response and the quantile treatment effects within the affected subpopulation.

Test for an intensive-margin response. Following Kline and Tartari (2016), we test whether the experimental earnings response can be entirely explained by changes in participation status (a pure extensive-margin response) or whether an intensive-margin response is required. The test is performed both unconditionally and conditional on participation in the pre-period.

Lee bounds on the intensive-margin earnings response. We estimate the bounds of Lee (2009) on the average earnings effect within the always-participating subpopulation – those who would participate regardless of treatment status. We tighten the bounds by conditioning on pre-determined covariates, in particular pre-period (2025) earnings.

6.3 Model-based analysis

Combining the experimental variation with economic theory is useful for two reasons. First, the model imposes restrictions on how earnings can respond to the tax credit, which can tighten the bounds on the reduced-form treatment effects. Second, it clarifies how the experiment identifies economic target parameters such as compensated and uncompensated elasticities and income effects.

Model. Individuals choose consumption C and labor earnings Y to solve

$$\max_{C,Y}\ u(C,Y) + \xi\,\mathbb{1}[Y=0] \quad \text{subject to} \quad C = Y - T(Y^A, Y^{TC}; d), \tag{11}$$

where $u(\cdot,\cdot)$ satisfies the standard properties, $T(Y^A, Y^{TC}; d)$ is the tax schedule under treatment status $d \in \{0,1\}$ (with $d = 1$ corresponding to the tax credit) as a function of taxable general income Y^A and the calculation base Y^{TC} – both implicit functions of earnings Y – and ξ captures the fixed (utility) cost of working. The model accommodates arbitrary heterogeneity in preferences and in ξ. We will consider how different additional assumptions on the model affect identification and inference; the assumptions we will examine include consumption and leisure being normal goods, quasi-linear preferences, no fixed costs of working, homogeneity restrictions on earnings responses, and parametric specifications of utility.

Bound-tightening via incentive regions. The kinks and phase-out in the tax credit's structure (see the description and Appendix in *The Norwegian Labor Income Tax Credit Experiment*) partition the earnings choice set into regions in which the marginal and average tax rates change in known directions:

- In the PLATEAU region the credit reduces the average tax rate without changing the marginal tax rate. Standard income effects then imply that the credit should not *increase* earnings, i.e. the earnings response is bounded above by zero.
- In the PHASE-OUT region the credit lowers the average tax rate but raises the effective marginal tax rate. Substitution and income effects then both push earnings downward.

- In the PHASE-IN region both the marginal and the average tax rate decline; substitution and income effects work in opposite directions on earnings, but the participation effect is unambiguously positive.

Following Dutz and Håvarstein (2026), we use restrictions of this kind to tighten the reduced-form bounds on the earnings response of individuals affected by the credit.

Identification of target parameters. Following Dutz et al. (2026) and Graber et al. (2026) and Dutz and Håvarstein (2026), we will point or partially identify economic target parameters: the intensive-margin compensated and uncompensated earnings elasticities, the income effect, and functionals thereof, using the variation in marginal and average tax rates induced by the kinks and phase-out in the tax credit's structure.

7 Ethical Considerations

- **IRB approval institution and protocol number:** IRB not applicable. DPIA (data protection impact assessment) August 2025; data protection officer assessment August 2025 attached. Research ethics assessment approved by the Frisch Centre's committee for research ethics (Oct 7, 2025), available online[11] (Bensnes et al., 2025; Hernæs et al., 2025).
- **Informed consent procedure:** Not relevant here. The experiment was implemented by the Norwegian tax administration after it was included in the national budget. The experiment is part of the national tax law for which individuals cannot withhold consent. Individuals cannot withhold consent to participate in registry-based scientific work.
- **Data privacy and confidentiality:** Data management follows the Frisch Centre data rules, which include secure server storage, two-factor authentication, and monitored data transmissions.

References

Bensnes, S., Hernæs, Ø., Håvarstein, M., Markussen, S., et al. (2025, September). *Ethical assessment – pilot scheme with an earned income tax credit for youth* (tech. rep.) (Original title in Norwegian: *Forskningsetisk vurdering – Forsøksordning med arbeidsfradrag for unge*). Frisch Centre for Economic Research.

Bitler, M. P., Gelbach, J. B., & Hoynes, H. W. (2006). What mean impacts miss: Distributional effects of welfare reform experiments. *American Economic Review*, *96*(4), 988–1012.

Dutz, D., & Håvarstein, M. (2026). *Revealed preference analysis of labor supply* (Working Paper).

Dutz, D., Mogstad, M., Håvarstein, M., & Torgovitsky, A. (2026, April). *Identification and estimation of labor supply elasticities from kinked budget sets* (BFI Working Paper No. 2026-53). Becker Friedman Institute for Economics, University of Chicago.

Graber, M., Håvarstein, M., Mogstad, M., Torsvik, G., & Vestad, O. L. (2026). *Substitution and income effects of labor income taxation* (NBER Working Paper No. 34987). National Bureau of Economic Research.

Hernæs, E., Kittelsen, S., & Hauge, K. (2025, October). *Møtereferat frischsenterets forskningsetikkutvalg, møte 24. september 2025* (Meeting minutes) (Accessed: 2025-12-11). Frischsenteret for samfunnsøkonomisk forskning.

Kline, P., & Tartari, M. (2016). Bounding the labor supply responses to a randomized welfare experiment: A revealed preference approach. *American Economic Review*, *106*(4), 972–1014.

[11] Both documents available online at https://www.arbeidsfradrag.no

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, *76*(3), 1071–1102.

Romano, J. P., & Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, *100*(469), 94–108.

Romano, J. P., & Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*(4), 1237–1282.

Romano, J. P., & Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics and Probability Letters*, *113*, 38–40.

Appendices

A Earnings effect on the affected: identification

Standard labor supply models predict that only those who choose earnings in the region where the tax credit affects their budget constraint should respond to the credit. Let $A(d)$ be an indicator for individual i's earnings under treatment status $d \in \{0, 1\}$ falling in a region where the credit alters the tax liability:

$$A(d) \equiv \mathbb{1}\big[T(Y^A(d), Y^{TC}(d); 0) \neq T(Y^A(d), Y^{TC}(d); 1)\big],$$

where Y^A denotes taxable general income, Y^{TC} the calculation base for the credit, and $T(\cdot, \cdot; d)$ is the tax schedule under treatment status d. Since the credit weakly reduces tax liability, we assume:

1. *Outcome equality on the unaffected:* if $A(1) = A(0) = 0$, then $Y(1) = Y(0)$ with probability one.

2. *Monotone affectedness:* $A(1) \geq A(0)$ with probability one.

Under these assumptions, and using random assignment of D, we recover the marginal earnings distributions for the subpopulation with $A(1) = 1$ from

$$\mathbb{P}\big[Y(1) \leq y \mid A(1) = 1\big] = \mathbb{P}[Y \leq y \mid A = 1, D = 1], \tag{12}$$

$$\mathbb{P}\big[Y(0) \leq y \mid A(1) = 1\big] = \mathbb{P}[Y \leq y \mid A = 1, D = 1] - \frac{\mathbb{P}[Y \leq y \mid D = 1] - \mathbb{P}[Y \leq y \mid D = 0]}{\mathbb{P}[A = 1 \mid D = 1]}. \tag{13}$$

Derivation. By the law of total probability,

$$\mathbb{P}[Y \leq y \mid D = 1] = \mathbb{P}[A = 1 \mid D = 1]\ \mathbb{P}[Y(1) \leq y \mid A(1) = 1] + (1 - \mathbb{P}[A = 1 \mid D = 1])\ \mathbb{P}[Y(0) \leq y \mid A(1) = 0].$$

Random assignment and Assumptions 1 and 2 imply $\mathbb{P}[Y(1) \leq y \mid A(1) = 1] = \mathbb{P}[Y \leq y \mid A = 1, D = 1]$, which gives equation (12). Substituting into the equation for $\mathbb{P}[Y \leq y \mid D = 0]$ (which has the same decomposition with treatment effects switched off) and rearranging yields equation (13).

The average earnings response and quantile treatment effects within the affected subpopulation are then obtained from these two recovered marginal distributions.