Estimating Tax Elasticities along the Income Distribution

Clara von Bismarck-Osten Jakob Wegmann

November 2026

Abstract

The elasticity of earnings with respect to taxation governs the trade-off between redistribution and efficiency, but available designs identify it only for specific ranges of income (at a threshold in bunching designs, for the group a reform affects in reform designs). This paper proposes a strategy to estimate the elasticity along the income distribution. It builds on the observation that, under a tax system that taxes spousal income jointly and gives tax advantages for children, individuals with the same earnings are differentially likely to see a tax change when the composition of the tax unit changes (a child leaves, a couple marries). The strategy formalises such comparisons and isolates the variation that stems from partner earnings. In a panel built from the universe of French tax returns, the earnings elasticity declines from 0.45 in the bottom decile to 0.23 in the middle deciles and to below 0.1 in the top decile. The profile is shifted up for women and at older ages, and the lowest earners' participation response comes from those with higher-earning partners. Accounting for this heterogeneity flattens optimal tax rates at low incomes and alters the predicted effects of reforms towards individual taxation.

We thank Uta Schönberg, Richard Blundell, Christian Dustmann, Kirill Borusyak, Costas Meghir and Emmanuel Saez.

Clara acknowledges the Institute for Fiscal Studies, the UCL Stone Centre, the Chartered Institute of Taxation and the Centre for Research and Analysis on Migration for supporting this project's infrastructure expenses, and the Economic and Social Research Council, the Thomas and Ulla Kolbeck Foundation and the Rockwool Foundation Berlin for supporting her work. Jakob's work was funded by the Rockwool Foundation Berlin and the German Research Foundation (DFG) through CRC TR 224, Project A05.

1 Introduction

How much revenue a tax raises and how much it harms economic activity depend on how elastic (responsive) earnings are with respect to taxation. This elasticity may be heterogeneous, but existing estimates are local. Bunching designs use a threshold of the tax schedule, and tax-reform designs study the bracket a reform affects, e.g., top incomes for a cut to top rates. An estimate at one threshold or for one bracket cannot recover how the elasticity varies along the income distribution, and because it is unclear how it varies (Gruber & Saez, 2002; Saez et al., 2012), it is often assumed not to (Diamond & Saez, 2011; Saez, 2001). Yet the predicted effects of reforms, the calibration of macroeconomic models and the assessment of the trade-off between redistribution and efficiency rest on this assumption and are misleading if it is wrong. With a declining profile of elasticities, the standard model of optimal taxation (Mirrlees, 1971; Saez, 2001) recommends slower withdrawal of transfers at the bottom and steeper progressivity further up; with a rising profile, the opposite.

This paper proposes a strategy to estimate the elasticity of earnings with respect to taxation along the income distribution. The main identification challenge is simultaneity: tax rates are a function of income, and the effort to earn income a function of how much it is taxed, implying that observed changes in tax rates are endogenous and need instrumentation. The instrument proposed in this paper builds on the observation that, under a tax system that taxes spousal income jointly and gives tax advantages for children, otherwise similar individuals are differentially likely to see their marginal tax rate change at "events" that change the composition of the tax unit. Consider two women with the same earnings but different partner earnings. When a child reaches the age at which it can no longer be counted in the tax unit (the event), the woman with the higher-earning partner may be pushed into the bracket above while the other remains in hers. The strategy formalises such comparisons, building on the simulated-instrument (simulated-IV) approach that Gruber and Saez (2002) developed for reforms. Among individuals with the same earnings, we simulate from pre-event incomes who crosses a threshold at the exit and who does not, and the simulated change in the marginal tax rate is the instrument for the realised one. A reform supplies one such change, while the events supply many (because individuals see children exiting their tax unit at every level of own earnings and, depending on partner earnings, sit close to different thresholds). Each

comparison, defined by a threshold, a level of own earnings and a tercile of partner earnings, is analysed on its own, and together the estimates give a profile of earnings elasticities along the income distribution and by partner earnings tercile.

We apply the strategy to an individual-level panel built from the universe of French tax returns, filed at the tax-unit level. With around fifty million individuals a year and 21 years (2003 to 2023), it can be split finely by own and partner earnings and by the other dimensions of heterogeneity we study.

We find the elasticity of earnings with respect to the marginal net-of-tax rate[1] to decline with income, most steeply at the bottom. It falls from 0.45 (standard error 0.11) in the bottom decile of the income distribution to 0.23 (0.03) in the middle deciles and to 0.07 (0.05) in the top decile.[2] Weighted to the French taxpayer population, the average earnings elasticity is 0.23 (0.03) and the average participation elasticity 0.25 (0.02). The decline holds for both genders, and with the same weights the earnings elasticity is 0.29 (0.05) for women and 0.17 (0.03) for men. These estimates can be disaggregated further by partner-earnings tercile to understand which households drive the larger elasticities at low incomes. Among the lowest earners, the participation response comes from those with higher-earning partners, with elasticities of 0.45 and 0.38 in the middle and top partner-earnings terciles, against 0.02 in the bottom one.

The assumptions are relatively weak. The strategy imposes no structure and does not require the event to be exogenous. Its main requirement is that, without the schedule change, earnings growth would not vary with partner earnings around the threshold among individuals with similar own earnings. Since the simulation uses pre-event earnings, the assumption fails if individuals bunch in anticipation, adjusting their earnings to thresholds that are not yet effective. In addition to (i) density tests around the thresholds, it can be inspected through (ii) the earnings growth of the two groups before the event, which should not differ, and (iii) earnings growth in the region of partner earnings where the event changes no rate, which should be constant. All three lend themselves to transparent, graphical inspection.

The variation has two advantages over that of a tax reform. Under a reform the same key assumption must hold between affected and unaffected groups that differ in income, and two

[1] The elasticity of own earnings (work income) with respect to the household net-of-tax rate, one minus the marginal tax rate the household faces at the individual's income. We further refer to it as the earnings elasticity.

[2] Deciles of the 2023 work-income distribution of all earners in France, men and women together.

biases then threaten it: (i) unusually high incomes tend to fall back towards the average (mean reversion), and (ii) growth rates differ along the distribution as inequality changes. Controlling for income is the remedy a large literature has refined in different ways (Graber et al., 2026; Gruber & Saez, 2002; Jakobsen & Søgaard, 2022; Kleven & Schultz, 2014; Kumar & Liang, 2020; Weber, 2014), but it is not straightforward to control for income without absorbing the variation itself when it is *between-income*. By contrast, our variation is *within-income*, in the sense that exposure to the tax change depends on partner earnings and own income can be held fixed. The second advantage is that the tax changes come from *shifts* of the thresholds and not from changes to the schedule. The individual's tax rate is then more likely to follow the direction the event predicts (monotonicity) than under a reform, and the rate change comes without the other provisions of a reform, which may confound the estimate (Kopczuk, 2005; Kumar & Liang, 2020; Saez et al., 2012).

Our results describe a specific population, parents at the time their first-born leaves the tax unit, but the group is broad – broad enough to control for age and to analyse age heterogeneity – and the strategy is transposable. The focus is placed on the child's fiscal exit because it is a purely fiscal event that generally does not coincide with an actual move-out. Parents are 52 on average at the event, but more than 80 percent are between 45 and 59, which allows (i) age fixed effects in every estimation and (ii) sample splits by age as a further heterogeneity analysis (revealing that elasticities are smaller below the minimum retirement age). The strategy is, however, not tied to this one event and can, for example, be used with marriage, which we use as a second design.

Accounting for the estimated heterogeneity matters for both normative and positive assessments. The workhorse model for the design of the tax and transfer system formalises the trade-off between redistribution and efficiency as a mechanism-design problem whose solution gives the optimal marginal rate at each income (Mirrlees, 1971). With the estimated profile of elasticities in place of a single number for the whole income distribution, the optimal rates flatten at the bottom of the distribution and steepen above the median (Brewer et al., 2010; Saez, 2001). With our profiles over own and partner earnings in place of one elasticity per spouse, the assessed welfare gains of reforms towards more individual taxation (Bierbrauer et al., 2024) rise on average, and most for households in which both partners earn little.

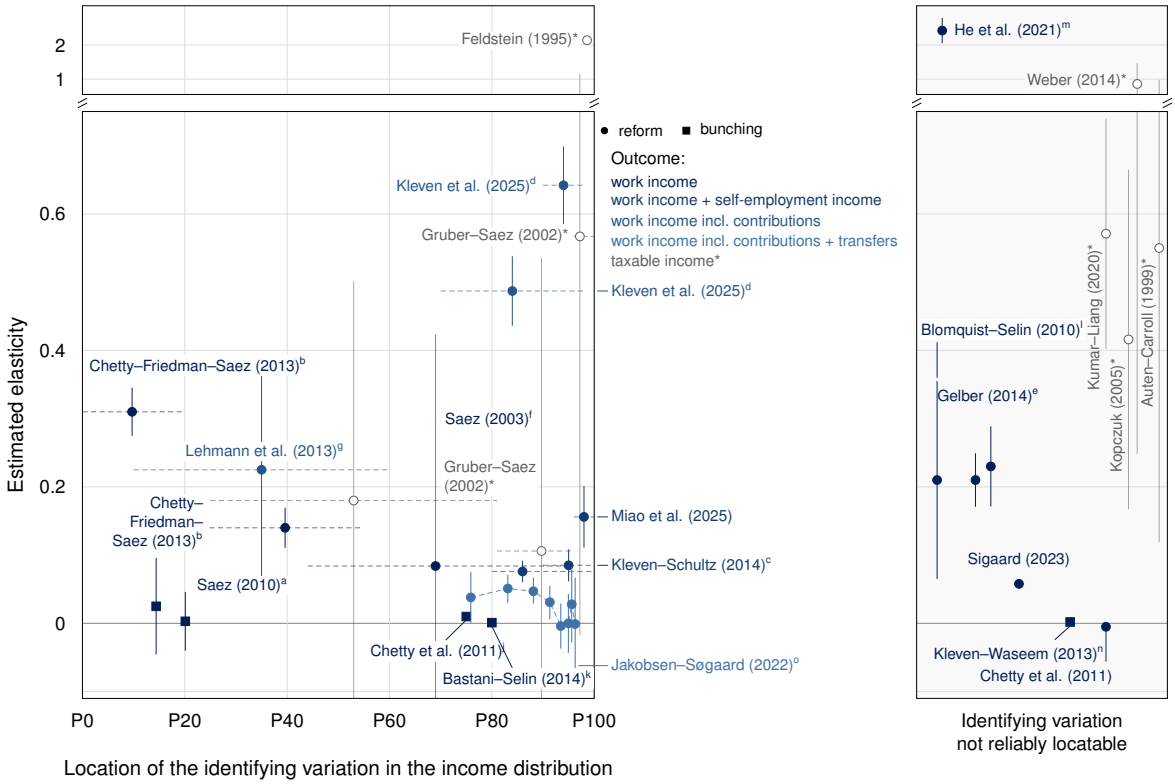


Figure 1: Published estimates of work-income and taxable-income responses

Notes: The figure shows quasi-experimental estimates of work-income tax elasticities from studies published in top field or general-interest journals. Colours denote the outcome (in-figure key); circles denote reform designs and squares bunching designs. Work income incl. contributions is measured gross of employer contributions (payroll contributions in Lehmann et al., employer-administered pension contributions in the Danish studies). Hollow grey markers (labels with a star) show six studies of taxable income[h], a broader concept (earnings plus interest, dividends and business income, net of deductions) estimated on samples not restricted to wage earners. Point markers sit at kinks or local-regression centres; horizontal dashed segments give the income range from which the identifying variation comes. Estimates without a locatable range appear in the right panel, where horizontal position is uninformative. The vertical axis is compressed above the break. Vertical lines are 95% confidence intervals where an elasticity standard error is reported. Percentile ranks refer to each study's ranking universe. Frisch elasticities[i] are not included. Exact populations, horizons, estimands, treated ranges and mapping grades are reported in Table A.1 of the online appendix, whose notes also explain the letters at the markers.

To the literature on heterogeneity in tax responsiveness by income, the paper contributes a profile of tax elasticities along the income distribution, estimated within a single setting, design and income measure. Figure 1 collects the quasi-experimental estimates of earnings responses published in top field or general-interest journals (Bastani & Selin, 2014; Blomquist & Selin, 2010; Chetty et al., 2011; Chetty et al., 2013; Gelber, 2014; He et al., 2021; Jakobsen & Søgaard, 2022; Kleven & Schultz, 2014; Kleven & Waseem, 2013; Kleven et al., 2025; Lehmann et al., 2013; Miao et al., 2025; Saez, 2010; Sigaard, 2023), together with six studies of taxable-income responses (Auten & Carroll, 1999; Feldstein, 1995; Gruber & Saez, 2002;

Kopczuk, 2005; Kumar & Liang, 2020; Weber, 2014). The estimates vary widely (Neisser, 2021); each is either local to a narrow part of the income distribution or one average over a wide range, and most concern the top three deciles. Because the studies differ in design and income measure, it is difficult to establish how much of the variation reflects actual heterogeneity in tax responsiveness.

To the literature on the taxation of couples, the paper contributes evidence across partner-earnings terciles. Normative work gives opposing prescriptions on optimal rates depending on the assumed difference between primary- and secondary-earner elasticities (Boskin & Sheshinski, 1983; Golosov & Krasikov, 2025).[3] Positive work on switches between joint and individual taxation, or on comparable policy variation, finds that joint taxation lowers the employment and earnings of secondary earners but not of primary earners (who are typically men) (Herold & Wallossek, 2024; Herold et al., 2025; Isaac, 2023; Kalíšková, 2014; LaLumia, 2008; Selin, 2014) – consistent with the heterogeneity we map by own earnings and partner-earnings tercile.

The paper also contributes evidence on heterogeneity by income measure, gender and time horizon. Broad earnings elasticities, which add self-employment income to work income, track the earnings elasticities and are shifted up at higher incomes.[4] Women respond more than men in our estimates, but by less than in the literature, which finds elasticities for married women and single mothers several times those of men (Blomquist & Selin, 2010; Eissa, 1996; Eissa & Liebman, 1996). We attribute the smaller gender difference to the fact that the estimates describe parents old enough to have a child leaving the tax unit. Recent work argues that the time horizon is central to the gap between micro and macro estimates (Kleven et al., 2025). We provide estimates by year since the event, up to four years.

The rest of the paper is structured as follows. Section 2 describes the French tax system and the construction of the individual-level panel. Section 3 develops the empirical framework and Section 4 presents the results. Section 5 examines how accounting for heterogeneity affects optimal tax schedules and the evaluation of reforms towards individual taxation. Section 6 concludes.

[3] With a common elasticity, optimal marginal rates on secondary earners are on average *higher* than on primary earners (Golosov & Krasikov, 2025) while the inverse-elasticity rule (rates inversely proportional to elasticities) prescribes the opposite (Boskin & Sheshinski, 1983) (with the secondary earner assumed more responsive).

[4] Our variation does not identify responses of taxable or capital income, the object of a large literature (e.g., Feldstein, 1995; Gruber & Saez, 2002; Neisser, 2021; Saez et al., 2012).

2 Background

2.1 Institutional setting

The main feature of French income tax design on which the strategy proposed in this paper builds is that the step-wise marginal tax schedule *shifts* with the composition of the tax unit (Figure 2, Panel (a)).[5] The system is one of joint family taxation in the sense that the tax liability (equation (1)) is a function of a *family quotient*, which itself is a function of marital status and the number of children. The quotient is 1 for singles, 2 for childless married (or civil-union) couples, 2.5 with one dependent child and 3 with two. Each further child raises it, and the tax advantage a child can confer is capped.[6,7] Because the quotient rises with marriage and with the number of children, the thresholds rise with it in terms of total taxable income. Every child exiting the tax unit moves them down in terms of total taxable income and, in turn, in terms of own and partner income. A married couple's income is taxed jointly in that the marginal tax rate depends on both incomes, and for a given level of own income, the bracket rises with partner income (Figure 2, Panel (b)). The tax liability under individual filing (i) and under joint filing (ii) is

$$(1) \qquad \text{(i)}\ \ T_s\big(z_{is}\big), \qquad\qquad \text{(ii)}\ \ T_s^H(z_{is}+z_{ps};\, Q_{is}) \;=\; Q_{is}\, T_s\left(\frac{z_{is}+z_{ps}}{Q_{is}}\right),$$

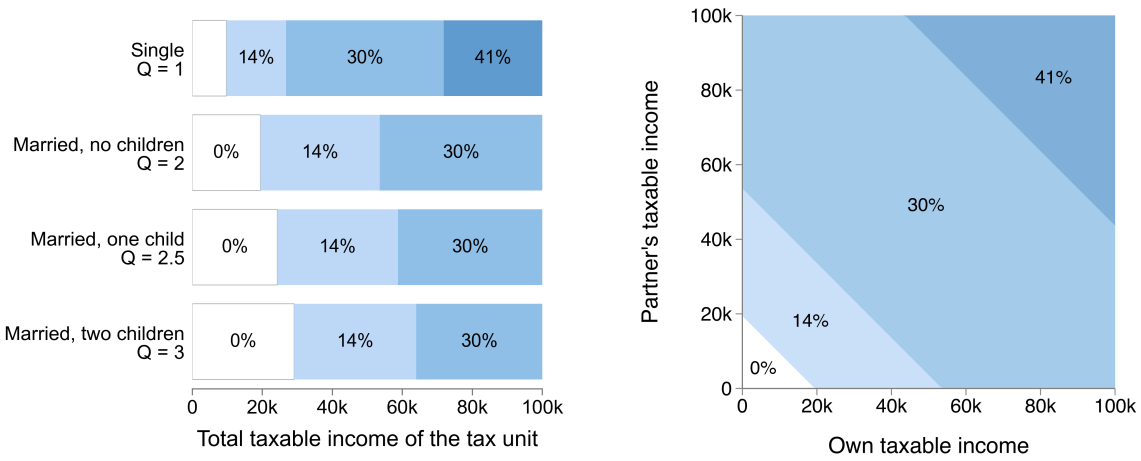
where T_s^H is the household tax function, T_s the single filer's bracket formula in year s, z_{is} and z_{ps} own and partner taxable income, and Q_{is} the family quotient.

The events that change the composition of the tax unit affect the marginal tax rate through the tax base and the quotient (Table 1). The fiscal exit of a dependent child lowers the quotient and leaves the tax base unchanged. Households whose income lies between the old and the new threshold move into a higher bracket. Marriage, in turn, pools the spouses' incomes and doubles the quotient. The bracket formula then applies to half the pooled income, and an

[5] The effective schedule differs from the statutory one through the *décote*, a reduction for small liabilities that is phased out as the liability rises. We take the *décote* into account and analyse effective marginal tax rates at the child's fiscal exit.

[6] The cap on the child advantage limits the benefit from an additional child to a fixed amount.

[7] Other factors also influence the family quotient. For example, it increases for single parents and parents of disabled children. We do not consider these cases.



(a) By household composition (b) Married couple, by own and partner income

Figure 2: Marginal tax rates by composition of the tax unit

Notes: Statutory schedule of 2015, before the *décote*, with the cap on the child advantage applied; Q is the family quotient. Panel (a): by household composition and total taxable income of the tax unit, with the marginal rate printed in each segment. Panel (b): married couple without children ($Q = 2$), by own and partner taxable income. Source: Own calculations based on OpenFisca France.

individual's bracket can rise only if own income is below the partner's and can fall only if it is above.

Table 1: Events and the tax liability function

Event	Tax base	Family quotient Q
Child's fiscal exit	unchanged: the couple's pooled income	$2.5 \to 2$ with one dependent child; $3 \to 2.5$ with two
Marriage	own income $\to$ the spouses' pooled income	1 for each spouse $\to$ 2 for the joint unit

Notes: What each event changes in the tax liability function of equation (1). Households with more dependent children see an analogous fall in the quotient, and the cap on the child advantage limits its effect on the tax. The marriage row refers to two previously single, childless individuals.

2.2 Data

We use an administrative database recording all income tax returns filed in France between 2003 and 2023 (POTE).[8] It includes all variables submitted in the income tax return (Form No. 2042), in addition to tax-relevant characteristics, such as the number of dependent children,

[8]POTE (Panel On The Taxation of Households), INSEE–DGFiP, accessed through the CASD secure data centre (https://www.casd.eu).

marital and filing status, gender and the municipality. An important advantage is that it records individual identifiers in addition to the tax-unit identifier.

From those records at the tax-unit level (the *foyer fiscal*), we build an individual panel of around fifty million individuals a year, which follows each person across tax units. From this panel, we extract the records of all individuals seeing the fiscal exit of their first-born child (or marrying, for the second design of Appendix F) and restructure them into an event-time panel, in which all information is indexed by the year relative to the event. Appendix E.1 describes the construction and characteristics of the two event samples.

3 Empirical framework

This section sets out the model and the identification problem (Section 3.1) and introduces our event-based design in comparison with reform-based designs (Section 3.2). Section 3.3 defines the cells within which comparisons are made, states the assumptions under which the estimands have a causal interpretation and illustrates them with an example cell. Sections 3.4 and 3.5 describe aggregation, estimation and inference.

3.1 The model and the identification problem

We model a jointly taxed couple, extending the standard one-person model (Jakobsen & Søgaard, 2022; Saez et al., 2012) to a unitary household maximising a quasi-linear and iso-elastic utility function, with the costs of earning income additively separable across spouses.

The household solves

$$\max_{z_{is},\, z_{ps}} \quad c_s \;-\; \sum_{j\in\{i,p\}} \frac{n_{js}}{1+\frac{1}{\epsilon_j}} \left(\frac{z_{js}}{n_{js}}\right)^{1+\frac{1}{\epsilon_j}} \qquad \text{s.t.} \qquad c_s \le z_{is} + z_{ps} - T_s^H(z_{is}+z_{ps}),$$

where z denotes income (subscripts i for the individual and p for her partner), c household consumption and T_s^H the household tax function of Section 2.1.[9] The cost of earning income is governed by two individual parameters, the elasticity ϵ_i and potential income n_{is}, the income

[9]The model is written for a couple filing jointly, the tax unit of the child's fiscal exit design, where the event lowers the quotient Q_{is} of equation (1). Marriage instead merges two individually filing units into one joint unit. Before marriage, the household pays $T_s(z_{is}) + T_s(z_{ps})$, and the rate in i's first-order condition, τ_{is}^H, is the own rate $1 - T_s'(z_{is})$.

i would earn at a zero marginal tax rate (Saez et al., 2012). Any change in income that is not a response to the rate runs through n_{is}. Because quasi-linearity rules out income effects, compensated and uncompensated elasticities coincide. Appendix C describes how the proposed strategy can allow for income effects.

Because the couple is taxed jointly on the sum of spousal incomes, the marginal net-of-tax rate is common to the two spouses, $\tau_{is}^H \equiv 1 - T_s^{H\prime}(z_{is} + z_{ps}) = \tau_{ps}^H$. With separable costs, each spouse's first-order condition involves only own income and the common rate, and yields the income-supply function $z_{is} = n_{is} \left(\tau_{is}^H\right)^{\epsilon_i}$ (Appendix A relaxes separability).[10] Taking logs and first differences gives the estimating equation

$$\Delta \ln z_{is} = \epsilon_i \, \Delta \ln \tau_{is}^H + \Delta \ln n_{is}. \tag{2}$$

The parameter of interest is ϵ_i, the earnings elasticity: the elasticity of own earnings with respect to the household net-of-marginal-tax rate.[11] The error term, $\Delta \ln n_{is}$, varies across individuals and years (transitory shocks, life-cycle and secular growth). At an event that changes the composition of the tax unit, it may also respond to the event itself (Section 3.3).

The quantities of interest are averages of the individual elasticities ϵ_i over subpopulations defined by own and partner income (characterised in Section 3.4). The iso-elastic form of the model makes equation (2) hold without approximation and ϵ_i the same at every income and rate. But under any other convex cost of earning income, the equation holds as a first-order (log-linear) approximation around the pre-event optimum, with ϵ_i the local elasticity at the pre-event income and rate (Appendix A). Our estimates are averages of local elasticities for the rate changes the events induce, and the design does not distinguish whether they differ across the income distribution because individuals at different incomes have different preferences or because the same preferences imply different responses at different incomes and rates.

[10] This is the formal counterpart of the standard practice of pairing the one-person model with the marginal rates a household tax system produces (Bick & Fuchs-Schündeln, 2018; Gruber & Saez, 2002).

[11] With separable costs, ϵ_i in equation (2) is the elasticity of own income with respect to the own rate, ϵ_{ii}. Without separability, equation (2) still holds, with ϵ_i now the elasticity of own income with respect to the household rate, $\epsilon_i^H \equiv \epsilon_{ii} + \epsilon_{ip}$ (the sum of its responses to the own and to the partner's rate when the household re-optimises jointly; Appendix A). The child's exit changes the rate common to both spouses, and the design recovers ϵ_i^H. Because spouses are taxed individually before marriage, the event can change the own rate alone, and the marriage design isolates ϵ_{ii} among couples for whom it does.

A naive estimation of equation (2) would not recover ϵ_i. Because the tax schedule depends on income, the marginal net-of-tax rate is a deterministic function of the dependent variable. A change in potential income changes realised income and, when it moves the household across a threshold, the marginal rate as well. In other words, the regressor of equation (2) is correlated with its error term by construction.

The literature has developed two main ways to handle this endogeneity. Bunching designs impose structure to infer the elasticity from the excess mass of taxpayers who locate at kinks of the schedule to avoid the higher rate above them (Kleven, 2016; Saez, 2010). IV designs draw on arguably exogenous variation from tax reforms, using simulated IV (Gruber & Saez, 2002) or grouping estimators (Blundell et al., 1998; Burns & Ziliak, 2017). We build on simulated IV and benchmark our design against it.

3.2 Our variation relative to reform variation

The reform-based instrument of Gruber and Saez (2002) is the individual's predicted rate change from the legislated change. It is a prediction because the realised rate change is endogenous. In the year of the reform, an individual's income moves with potential-income growth and might already reflect her response to the reform's changed tax incentives. To tease out the rate change driven only by the legislation, the prediction applies the post- and pre-reform schedules to the same income of k years earlier,[12]

$$\Delta \ln \tau_{is}^{\text{reform}} \equiv \ln\big(1 - T_s'^{\,\text{post}}(z_{i,s-k})\big) - \ln\big(1 - T_s'^{\,\text{pre}}(z_{i,s-k})\big).$$

The instrument is valid under the assumption that potential-income growth is uncorrelated with the predicted rate change conditional on lagged income,

$$\operatorname{cov}\big(\Delta \ln \tau_{is}^{\text{reform}},\, \Delta \ln n_{is} \,\big|\, z_{i,s-k}\big) = 0.$$

The assumption is demanding because exposure to reforms is usually *between-income*. A reform typically changes rates in one bracket, and the instrument then sorts individuals into predicted treated and predicted controls by lagged income. Because the level of income predicts

[12] See Saez et al. (2012) for an overview of the literature.

income growth for reasons unrelated to taxes, the instrument is correlated with potential-income growth. A high lagged income is in part transitory and reverts. Incomes along the distribution have also grown at different rates, as inequality changed. To alleviate these concerns, the literature has proposed several ways to control for income without absorbing the identifying between-income variation, such as richer income controls (Gruber & Saez, 2002; Kleven & Schultz, 2014), longer instrument lags (Weber, 2014), differencing against placebo periods or cohorts with graphical validation (Graber et al., 2026; Jakobsen & Søgaard, 2022), and allowing responses to vary across individuals (Kumar & Liang, 2020).

A further concern is that reform variation can violate monotonicity (Assumption 3 below). The IV estimand weights individual elasticities by the change the reform induces in their realised net-of-tax rate, and a reform can move these rates in opposite directions. When a reform raises the rate of a bracket, an individual may respond by leaving work or by cutting hours enough to fall below that bracket. In both cases she can face a lower rate where a rise was predicted, and the estimand then no longer aggregates individual elasticities with convex weights. Because the events we use shift the thresholds of the schedule (and not the rates), a violation requires income to fall below the lower threshold of the pre-event bracket. Income that falls back across the shifted threshold only returns to the pre-event bracket. Partner income also places a floor under the rate an individual can reach by reducing own income. Appendix G formalises the argument and illustrates a violation in a calibrated example.

The event-based instrument we propose uses the fact that the bracket thresholds move when the composition of the tax unit changes (Table 1). Our main design uses the fiscal exit of a dependent child, which lowers the quotient by which total taxable income is divided and shifts the thresholds of the schedule down. The tax unit and the schedule are otherwise unchanged (Section 2.1). Our second design uses marriage, which merges two tax units (Appendix F). A reform, by contrast, usually changes the rates of the schedule and not the thresholds, and can come with other, potentially confounding, provisions.[13]

We index the two schedules by $g \in \{0, 1\}$: the individual is taxed under the pre-event schedule ($g = 0$) or the post-event one ($g = 1$). From here on, "schedule" refers to the household tax

[13] A reform is salient, and responses to salient and to unnoticed tax changes differ (Saez et al., 2012). It rarely moves one rate alone and usually changes the tax base with it (Kopczuk, 2005). It also moves rates elsewhere in the schedule, over which individuals optimise too, and an estimate based on the rate at their own income omits them (Kumar & Liang, 2020).

function T^H under g, and "formula" to the calendar-dated bracket formula T_s of equation (1). Dates are expressed in event time, $t \equiv s - m_i$, where s is the calendar year and m_i that of the event. For the fiscal exit, m_i is the year in which fiscal attachment of the first-born ends. Attachment lasts until the child turns 18 and can be extended to 21, or to 25 for a child in education. The schedule indicator in event time, $g_{it} \equiv \mathbb{1}[t \geq 0]$, switches every individual in our designs from the pre-event to the post-event schedule at $t = 0$.[14] Comparisons are made within cells of individuals who all experience the event, and Section 3.3 states what identification requires of it.

Like the reform-based instrument, the event-based instrument is a simulation. It applies the pre-event year's formula twice to pre-event ($t = -1$) income divided by the quotient, under the pre-event ($g = 0$) and the post-event ($g = 1$) schedule, and takes the difference in log net-of-tax rates. Under schedule g, the total taxable income of the tax unit divided by the quotient is $b_{it}(g)$. For an individual in a couple filing jointly, $b_{it}(g) = (z_{it} + z_{pt})/Q_g$. For a spouse taxed individually before marriage, $b_{it}(0) = z_{it}$. The two applications differ only in g, which sets the quotient (and at marriage, the tax unit):

$$Z_{it} \;\equiv\; g_{it} \times \underbrace{\Big[\ln\big(1 - T'_{-1}\big(b_{i,-1}(1)\big)\big) \;-\; \ln\big(1 - T'_{-1}\big(b_{i,-1}(0)\big)\big)\Big]}_{\equiv\ Z_i,\ \text{the } \textit{dose}}, \tag{3}$$

where T_{-1} is the formula of the individual's pre-event calendar year. The dose Z_i is fixed at that year, and the indicator g_{it} switches it on once the post-event schedule applies. Within a cell the dose takes few values (the differences in log net-of-tax rates between brackets), and each dose δ corresponds to a bracket transition. The dose is zero for individuals whose income divided by the quotient is predicted to remain in its bracket.[15] These are the predicted controls, and the others the predicted treated. At the child's fiscal exit the quotient falls from 2.5 to 2 with one dependent child and from 3 to 2.5 with two. With a lower quotient, income divided by the quotient weakly rises at every level of household income, and for almost every individual the event predicts a weakly higher marginal rate, $Z_i \leq 0$.

[14] The design compares the pre-event year $t = -1$ with one post-event year, and the event-time profiles of Section 4.2.3 make this comparison at each horizon.

[15] At the child's fiscal exit, the dose and the realised net-of-tax rate τ_{it}^H of Section 3.5 are computed on the same effective schedule. The hypothetical bracket under the post-event quotient and the realised bracket of every year are assigned from the same set of kinks and effective marginal rates, by year and quotient, so that the *décote* and the cap on the child advantage (Section 2.1) enter instrument and realised rate identically.

The variation in our instrument is *within-income*: conditional on own income and the pre-event bracket, whether an individual is pushed over a bracket threshold is a function of partner income (Figure 2, Panel (b)).[16] Own income can then be held fixed without absorbing the identifying variation.

3.3 Identification and an illustration

To state the identifying assumptions, we let $Y_{it}(g)$ denote potential log income at event time t under schedule $g \in \{0, 1\}$, and $\mathrm{NTR}_{it}(g)$ the corresponding potential log net-of-tax rate. Because the event occurs under both potential outcomes and only the schedule is varied, $Y_{it}(0)$ is not a no-event counterfactual. The schedule change shifts the log net-of-tax rate by $\Phi_{it} \equiv \mathrm{NTR}_{it}(1) - \mathrm{NTR}_{it}(0)$. The construction is stated for earnings, with log work income as the outcome. The participation indicator, $P_{it} \equiv \mathbb{1}[z_{it} > 0]$, is treated in the same way, with its estimation in Section 3.5. The potential outcomes map to observed outcomes through

$$\begin{aligned} y_{it} &= Y_{it}(0) + \mathbb{1}[Z_{it} \neq 0]\,\big[Y_{it}(1) - Y_{it}(0)\big], \\ \mathrm{NTR}_{it} &= \mathrm{NTR}_{it}(0) + \mathbb{1}[Z_{it} \neq 0]\,\big[\mathrm{NTR}_{it}(1) - \mathrm{NTR}_{it}(0)\big]. \end{aligned} \tag{4}$$

In line with the model of Section 3.1, we assume potential log income is linear in the log net-of-tax rate, with the same slope under the pre-event and the post-event schedule (as in Graber et al., 2026; Saez et al., 2012):

$$Y_{it}(\mathrm{NTR}, g) = \epsilon_i \times \mathrm{NTR} + \ln n_{it}(g), \tag{5}$$

where ϵ_i is the individual elasticity and $n_{it}(g)$ the zero-rate income under schedule g (the potential income of Section 3.1). The argument g of n_{it} collects every channel from the schedule change to income other than the marginal rate, and Assumption 2 below rules these channels out.[17]

[16] Jakobsen and Søgaard (2022) state the identifying assumption of reform designs as constant trend differentials and develop its graphical validation, which Section 3.3 follows.

[17] With non-separable costs, the household model of Appendix A delivers the same form locally, with $\epsilon_i = \epsilon_i^H$ at the child's fiscal exit (Section 3.1). The proof of Proposition 1 below does not need linearity. Together with Assumption 2, equation (5) enters it only through $Y_{it}(1) - Y_{it}(0) = \epsilon_i\,\Phi_{it}$. For any model in which the schedule affects income only through the rate and through $n_{it}(g)$, so that under Assumption 2 an unchanged rate means unchanged income, the proposition holds with ϵ_i the ratio of the individual's income change to her

Identification is local in that the assumptions need to hold within "cells" of individuals who are alike in the determinants of income growth, all experience the event and differ in whether it is predicted to change their marginal rate. Each dose δ in the support of Z is compared separately to the predicted controls of its cell. This has two advantages. First, common trends (Assumption 1 below) are easier to defend and to test among such individuals. Second, comparing each dose separately avoids the negative weights that a pooled regression across doses can place on individual elasticities when these are heterogeneous (Callaway et al., 2026; Graber et al., 2026).[18]

Formally, cells restrict own income to a narrow bin and partner income to a tercile, and fix the pre-event bracket, the event year and the event type. Let $X_i \equiv (z_{i,-1},\, z_{p,-1},\, m_i,\, e_i)$ collect these predetermined variables, with e_i the event type (the number of dependent children and single or joint filing before the event). The cells $\mathcal{I}_c$ are the elements of the resulting partition of the support of X, and $c(X_i)$ is the cell of individual i.[19] Throughout, expectations subscripted by c, $\mathbb{E}_c[\,\cdot\,]$, are taken over the individuals of cell $\mathcal{I}_c$, and $\Delta_t V \equiv V_t - V_{-1}$ denotes growth relative to the pre-event year.

Within cells, we impose three assumptions: **Assumption 1** (Conditional common trends). *For every cell, horizon t, and dose δ in the within-cell support of Z,*

$$\mathbb{E}_c\big[\Delta_t Y(0) \mid Z_i = \delta\big] = \mathbb{E}_c\big[\Delta_t Y(0) \mid Z_i = 0\big],$$
$$\mathbb{E}_c\big[\Delta_t \mathrm{NTR}(0) \mid Z_i = \delta\big] = \mathbb{E}_c\big[\Delta_t \mathrm{NTR}(0) \mid Z_i = 0\big].$$

Because Z_i is itself a function of pre-event incomes, Assumption 1 is a restriction on groups defined by pre-event incomes. It holds whenever untreated growth does not vary within cells at all, $\mathbb{E}[\Delta_t Y(0) \mid X_i = x] = f_y^t(c(x))$ for all x, and likewise for $\Delta_t \mathrm{NTR}(0)$. Since the event year and the event type are fixed by the cell, this requires that the exact positions of own income within the cell's bin and of partner income within the cell's partner-income tercile do

rate change, $[Y_{it}(1) - Y_{it}(0)]/\Phi_{it}$. That ratio may then differ across horizons and with the size of the change, which the linear form rules out.

[18] Conversely, if elasticities were homogeneous within cells, the comparison dose by dose and monotonicity (Assumption 3 below) would both be dispensable, because any weighting of identical elasticities recovers the common ϵ. Even then the cells would remain necessary, because the predicted treated and the predicted controls differ unconditionally in own and partner income, event year and event type, and untreated growth can vary along each of these dimensions.

[19] Given X_i, both the pre-event tax bracket and the instrument of equation (3) are determined.

not predict untreated growth. Because $\Delta_t Y(0) = \epsilon_i \, \Delta_t \mathrm{NTR}(0) + \Delta_t \ln n(0)$ by equation (5), the restriction concerns the untreated evolution of net-of-tax rates and of zero-rate income. Zero-rate income includes the direct effects of the event, which occur under both schedules and difference out between the two groups if Assumption 1 holds.

Assumption 1 fails if there is anticipatory sorting around thresholds in the year before they are applicable. Individuals could lower their income before the event (anticipatory bunching), or those predicted to see a rate increase could delay the event, for example by extending a child's attachment. Both would imply excess density below the threshold, among the predicted controls, and less density above it, among those predicted to see a rate increase.

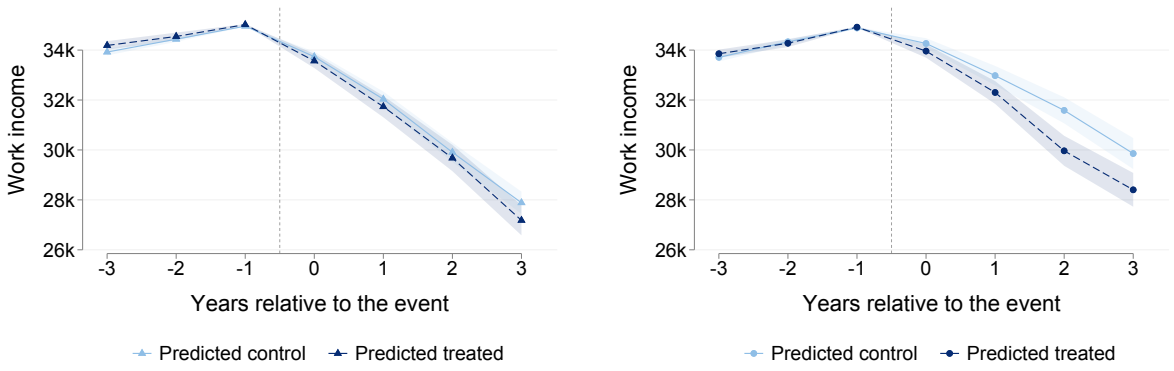
Table 2: The example cell: characteristics by predicted treatment status

	Men		Women	
	Control	Treated	Control	Treated
Work income (€)	34,960	35,022	34,885	34,919
Partner's work income (€)	22,098	28,979	22,592	30,524
Age	54.0	54.7	53.3	53.8
Partner's age	52.1	53.1	54.3	54.8
Marginal rate at $t = 0$				
Predicted	14%	30%	14%	30%
Realised (mean)	16%	24%	16%	24%
N	4,295	2,586	1,687	1,735

Notes: The cell comprises individuals with pre-event own income in [€ 34,000, € 36,000], a pre-event marginal rate of 14% and quotient 2.5, whose first-born exits the tax unit at $t = 0$ in 2006–2020. The two groups are balanced on own income and age, and differ in partner income, which determines whether income divided by the lower quotient crosses the threshold of the post-event schedule. Incomes are measured in the pre-event year and ages at the event. The rate rows refer to $t = 0$, the first year under the post-event schedule.

Assumption 1 can be evaluated with density tests (Figure 5) and by inspecting income growth before the event, as an example illustrates (Table 2 and Figure 3). The example comprises individuals with pre-event own income between € 34,000 and € 36,000, at a pre-event marginal rate of 14%, in tax units with quotient 2.5 whose first-born exits at $t = 0$ (event years 2006–2020). Individuals whose partner earns enough are predicted into the 30% bracket once the quotient falls to 2 ($Z_i = \delta = \ln(0.70/0.86) \approx -0.21$), and the rest are not ($Z_i = 0$). Figure 3 plots the income paths of the two groups, and the assumption is evaluated by their alignment before the event.[20]

[20] Every exhibit of the example applies the civil-status restriction of the estimation sample (Appendix E.1): individuals whose civil status changes at the event are excluded.



(a) Men

(b) Women

Figure 3: Work income around the first-born's fiscal exit, by predicted treatment status

Notes: Mean work income in euro by year relative to the event; bands are 95% confidence intervals; the dashed vertical line marks the event. The means are unadjusted (no matching, no covariates), as in the design's own comparison. The two series are the predicted treated ($Z_i = \delta < 0$) and the predicted controls ($Z_i = 0$). Table 2 describes the sample.

Assumption 2 (Exclusion restriction). $n_{it}(0) = n_{it}(1)$ *with probability one.*

Assumption 2 is that the schedule change does not affect zero-rate income and that it affects income only through the marginal net-of-tax rate.[21] This has two implications. First, the assumption requires that there are no income effects (the fall in the quotient also raises every household's tax liability, and by more the further its income lies above the new threshold). The requirement is the potential-outcomes form of the no-income-effects property of the quasi-linear model of Section 3.1, under which it holds by construction.[22] If leisure is normal, the larger liability raises income while the higher marginal rate lowers it, and our estimates then recover the total response to the schedule change and are lower bounds on the compensated elasticity. Appendix C relaxes the assumption (Figure C.1). Second, the assumption requires that partner income has no channel of its own to income growth arising around the threshold the cell is considering. Such channels could exist: a higher-earning partner may allow the individual to reduce work once the child has left the tax unit, or the individual may adjust own work to changes in the partner's.[23] Either would appear as a gradient of income growth in partner income, whereas the predicted rate change is a step function of partner income.

[21] Formally, decompose $\ln n_{it}(g) = \mu_{it} + \xi_{it}(g)$ with $\xi_{it}(0) \equiv 0$: μ_{it} carries life-cycle and calendar-time growth and the event's direct effects (changed time use, home production, transfers), which occur under both schedules and remain unrestricted. Assumption 2 sets $\xi_{it}(1) = 0$. Reform settings impose the analogous restriction when they assume zero-rate income unaffected by the reform.

[22] Quasi-linearity also implies that ϵ_i is the compensated elasticity, which coincides with the uncompensated one (Graber et al., 2026, Propositions 3–4).

[23] Appendices A and C take up cross-spouse responses and income effects.

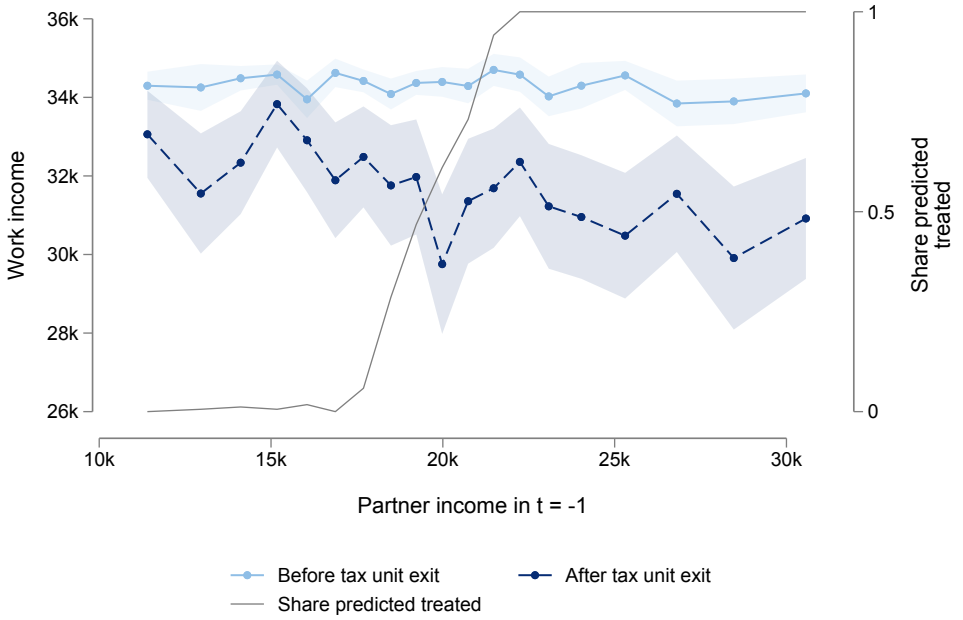


Figure 4: Identifying variation at the first-born's fiscal exit

Notes: Sample: women of the example (pre-event own income in [€34,000, €36,000], pre-event marginal rate 14%, quotient 2.5, first-born exiting the tax unit at $t = 0$ in 2006–2020). Left axis: mean work income at $t = -2$ (before) and, for each individual, the average of $t = +1$ and $t = +2$ (after), by 20 equal-count bins of partner income (about 170 women per bin); bands are 95% confidence intervals. The lowest bin (mean partner income below €10,000) lies outside the plotted range. Right axis: the share of the bin predicted into the 30% bracket. Partner income is the household's taxable income at $t = -1$ net of the individual's own earnings. Because taxable wages carry the 10% deduction for professional expenses, the measure is about nine tenths of the partner's wage less a tenth of own earnings. The validation region comprises the bins with mean partner income above €11,000 in which the predicted-treated share is below 5% (partner income up to about €16,900), and the treatment region those in which it exceeds 95% (from about €22,200). The figure follows the graphical validation approach of Jakobsen & Søgaard (2022). In the validation region the event predicts no rate change. A gradient of income growth (the vertical distance between the two series) in partner income there would indicate a direct partner-income channel and a violation of Assumption 2.

Figure 4 splits the example cell by partner income into a *treatment region*, where the exit is predicted to raise the marginal rate, and a *validation region*, where the exit predicts no rate change. The requirement is evaluated in the validation region: partner income varies within it with no accompanying rate change, and absent such a channel, income growth (the vertical distance between the two series of Figure 4) is flat in partner income.

Assumption 3 (Monotonicity). $Z_i \left[\mathrm{NTR}_{it}(1) - \mathrm{NTR}_{it}(0)\right] \geq 0$ *with probability one, for all* $t \geq 0$.

Assumption 3 is that the schedule change does not move any individual's marginal rate in the direction opposite to the predicted change ($Z_i < 0$ for a predicted rate increase; Section 3.2).

(Rates may still change for other reasons.) This is the standard no-defiers condition, under which Wald estimands are positively weighted averages of individual elasticities (Angrist & Imbens, 1995; Imbens & Angrist, 1994). Appendix G formalises why reform variation can violate the condition, illustrates a violation in a calibrated example, and explains why our events make a violation less likely.

We define the estimands separately for each cell, dose δ and horizon t. The reduced form compares realised income growth, $\Delta y_{it} \equiv \ln z_{it} - \ln z_{i,-1}$ (the observed counterpart of $\Delta_t Y$), between the dose's predicted treated and the predicted controls, and the first stage compares realised net-of-tax rate growth, $\Delta\text{NTR}_{it} \equiv \ln \tau_{it}^{H} - \ln \tau_{i,-1}^{H}$, between the same two groups:

$$RF_c^t(\delta) \equiv \mathbb{E}_c\big[\Delta y_{it} \mid Z_i = \delta\big] - \mathbb{E}_c\big[\Delta y_{it} \mid Z_i = 0\big],$$
$$FS_c^t(\delta) \equiv \mathbb{E}_c\big[\Delta\text{NTR}_{it} \mid Z_i = \delta\big] - \mathbb{E}_c\big[\Delta\text{NTR}_{it} \mid Z_i = 0\big].$$

Our headline estimand is their ratio, the cell–dose elasticity[24]

$$\beta_c^t(\delta) \;\equiv\; \frac{RF_c^t(\delta)}{FS_c^t(\delta)}.$$

Proposition 1 (Causal interpretation). *Under the mapping* (4)*, the income function* (5)*, and Assumptions 1–3, for every cell, dose $\delta \neq 0$, and horizon $t \geq 0$ with $FS_c^t(\delta) \neq 0$:*

$$\beta_c^t(\delta) \;=\; \mathbb{E}_c\big[\,\omega_{it}\,\epsilon_i \mid Z_i = \delta\big], \qquad \omega_{it} \equiv \frac{\Phi_{it}}{\mathbb{E}_c[\Phi_{it} \mid Z_i = \delta]},$$

where Φ_{it} is the rate change induced by the schedule change and the weights $\omega_{it} \geq 0$ average to one. The cell–dose elasticity is then a positively weighted average of individual elasticities among the predicted treated.[25] *(Proof: Appendix B.)*

[24] The ratio $RF_c^t(\delta)/\delta$ instead scales the response by the predicted rather than the realised rate change (the policy elasticity of Hendren (2016)). It needs only Assumption 1 and is unaffected when frictions or threshold-crossing responses attenuate the realised rate change.

[25] Proposition 1 is the event-based analogue of the result of Graber et al. (2026), who characterise when reform-based elasticity estimands have a causal interpretation. Their two conditions are a binary instrument contrast and income controls rich enough to reproduce the conditional mean of the instrument. In our setting, the first is a feature of the design (each estimand compares one dose, $Z_i = \delta$, with the predicted controls), and the second is met because each contrast is taken within a cell.

3.4 Aggregation

We aggregate the cell–dose elasticities in two steps with frequency weights, which are proportional to the sizes of the regression samples.[26] A pooled regression across doses would instead apply implicit weights, which can turn negative when elasticities are heterogeneous (Callaway et al., 2026; Graber et al., 2026). Each estimate $\beta_c^t(\delta)$ rests on its own regression sample (Section 3.5) observed at horizon t. We write $N_c^t(\delta)$ for its size and $\mathcal{D}_c$ for the set of the cell's non-zero doses. In the first step, we combine the cell–dose elasticities of a cell:

$$\beta_c^t \equiv \sum_{\delta \in \mathcal{D}_c} w_c^t(\delta)\, \beta_c^t(\delta), \qquad w_c^t(\delta) \equiv \frac{N_c^t(\delta)}{\sum_{\delta' \in \mathcal{D}_c} N_c^t(\delta')}.$$

In the second step, we aggregate over the cells of a group h (own-income intervals for the income profiles and own-by-partner-income rectangles for the maps of Section 4):

$$\epsilon_h^t \equiv \sum_{c \in h} v_h^t(c)\, \beta_c^t, \qquad v_h^t(c) \equiv \frac{\sum_{\delta \in \mathcal{D}_c} N_c^t(\delta)}{\sum_{c' \in h} \sum_{\delta \in \mathcal{D}_{c'}} N_{c'}^t(\delta)},$$

where $v_h^t(c)$ is the share of cell c in the combined regression samples of the cells of group h. We report ϵ_h^t as an event-time profile in Section 4.2.3 (by gender in Appendix E.3). The headline displays of Section 4 pool the post-event years $t = 0$ to 3 in the same way, each year counting in proportion to the group's combined regression samples in that year.

Because both steps use non-negative frequency weights, the aggregates ϵ_h^t and their averages over the post-event years inherit the causal interpretation of Proposition 1. The aggregation weights $w_c^t(\delta)$ and $v_h^t(c)$ are known sample frequencies, and alternative, policy-relevant weights can be substituted without affecting the identification argument.[27]

3.5 Estimation and inference

Estimation is by two-stage least squares within cells, separately for women and men. The event window spans $t = -3$, -2 and $t = 0, \ldots, 3$, with $t = -1$ the base year. We run one regression for each cell, dose and period. Its sample pools the dose's predicted treated with

[26] The aggregation is stated for earnings and applies to participation in the same way (Section 3.5).

[27] Inverse-variance weights are also non-negative and preserve the causal interpretation, but they weight an estimate by its precision. The headline displays of Section 4 use the frequency weights throughout. Section 4.2.1 reports an inverse-variance-weighted aggregate for comparison.

the cell's predicted controls, and the regression is

$$\Delta y_{it} \;=\; a_c^t(\delta) \;+\; \beta_c^t(\delta)\,\Delta\mathrm{NTR}_{it} \;+\; W_i'\gamma \;+\; u_{it}, \tag{6}$$

with Z_{it} as the instrument for the realised rate change $\Delta\mathrm{NTR}_{it}$. The coefficient $\beta_c^t(\delta)$ is a cell–dose–period estimate. Before the event $Z_{it} = 0$, and equation (6) has no first stage. At $t < 0$ we therefore estimate the reduced form and first stage on Z_i, the placebo contrasts reported alongside the estimates. The controls W_i are fixed effects for own income at $t = -1$, -2 and -3 (in bins), age, marital status, the number of children at $t = -1$, and the event year (the cells of Section 3.3 fix the last two exactly, whereas estimation pools over them). The fixed effects guard against differences in composition between the predicted treated and the predicted controls. The cells fix own income only within a bin, and within a cell the fixed effects absorb the remaining differences in age and in own income before the event (Appendix E.2 reports the balance). In the lowest own-income bin, the cells are formed separately for individuals with at least €400 of work income at $t = -1$ (the earner sub-sample) and for those below (the entrant sub-sample), so that a cell compares individuals of the same pre-event earner status.

Standard errors are clustered by person. Within a cell, the estimates for different doses and periods rest on the same individuals and are correlated, and the standard errors of the aggregates β_c^t and ϵ_h^t account for this correlation. Different cells contain different individuals and are treated as independent.[28]

The outcome is the log difference of earnings between $t = -1$ and t, with earnings measured as work income without self-employment income. The coefficient $\beta_c^t(\delta)$ is the earnings elasticity of Section 3.1, the elasticity of own earnings with respect to the household net-of-marginal-tax

[28] The classical standard error treats every cell–dose–period estimate as independent. The clustered standard error is 1.39 times the classical one at the median bin of the income profiles and 1.45 times at the median tile of the maps, mostly because an aggregate pools four post-event years of the same people. The pre-event placebo contrasts are pooled with the frequency weights of the estimate they accompany and tested with the same covariance. The balanced-sample and matching estimates of Section 4.3 and the marriage estimates of Appendix F treat cells as independent in their standard errors, as do the placebo tests of the marriage design. The standard errors of the totals over both genders and of the gender differences treat the women's and the men's estimates as independent.

rate.[29] Because a log difference is defined only for individuals with positive work income in both years, the earnings elasticity is an intensive-margin response.[30]

For participation, the change in the participation indicator, $\Delta_t P_i \equiv P_{it} - P_{i,-1}$, replaces Δy_{it} in equation (6). The coefficient is the participation semi-elasticity, the change in the participation probability per unit change in the log household net-of-marginal-tax rate. We divide it by the cell's participation rate before the event to obtain the participation elasticity.[31] Analysing participation separately from the intensive-margin earnings elasticity is one of the approaches that Chen and Roth (2024) recommend. Section 4.3 evaluates the robustness of the estimates to the comparison method (nearest-neighbour matching within cells on age and pre-event work income in place of the regression comparison), to the sample (the balanced sample of individuals observed in every year of the event window, and the sample without the 2020–2023 event years) and to a stricter first-stage requirement, which drops the cells with weak first stages (Appendix E.8). It also reports the combined response of earnings and participation, estimated by Poisson regression of the level of earnings, another of the approaches that Chen and Roth (2024) recommend, and applies the strategy to a second event, marriage (Appendix F).

4 Results

Section 4.1 describes the cells, reports descriptive statistics on the population in them and evaluates the identifying assumptions. Section 4.2 then reports the elasticities, first aggregated to the population of French taxpayers (Section 4.2.1), then as profiles along the own-income distribution (Section 4.2.2) and by year relative to the event (Section 4.2.3), and finally as maps over the joint distribution of own and partner income (Section 4.2.4). Section 4.3 evaluates

[29] This is the earnings elasticity of optimal-tax models, taken with respect to one minus the marginal rate and defined on the earnings of those who work rather than on hours (Kleven et al., 2009; Saez, 2001). The reform model of Section 5.2 uses it in the same sense (Bierbrauer et al., 2024).

[30] The reduced form $RF_c^t(\delta)$ and first stage $FS_c^t(\delta)$ are estimated alongside, as regressions of Δy_{it} and $\Delta \mathrm{NTR}_{it}$ on Z_i with the same controls, and the two-stage least squares coefficient equals their ratio.

[31] The participation literature uses a different price, the participation tax rate. It is the share of own earnings that the household pays in additional tax when the individual works, given partner income. The child's fiscal exit moves the two prices differently. The marginal rate changes only where the household crosses a threshold, which is the step our cells isolate, whereas the participation tax rate also rises for the predicted controls. Its contrast between the predicted treated and the predicted controls is about one eighth of the marginal-rate contrast.

the robustness of the estimates and applies the strategy to a second event, and Section 4.4 reports further heterogeneity, by gender, age, income definition and household characteristics.

4.1 Cells, descriptive statistics and validation

We first describe the cells. There are 545 of them, 297 for men and 248 for women (Table 3, Panel A), formed in the event-time panel of all individuals who see the fiscal exit of their first-born child. A cell fixes, in the year before the event, (i) an own-income bin €10,000 wide, (ii) a partner-income tercile within the bin, (iii) the pre-event bracket and (iv) one tax change the exit can induce (e.g., from the 14 to the 30 percent rate), and it compares the individuals predicted to see that change with those of the same bin, tercile and pre-event bracket predicted to remain in it (the predicted controls; Section 3.3). Together the cells hold 1.4 million individuals of each gender. Every estimate below is an average over cells, each weighted by its regression sample (Section 3.4).

Table 3: The cells

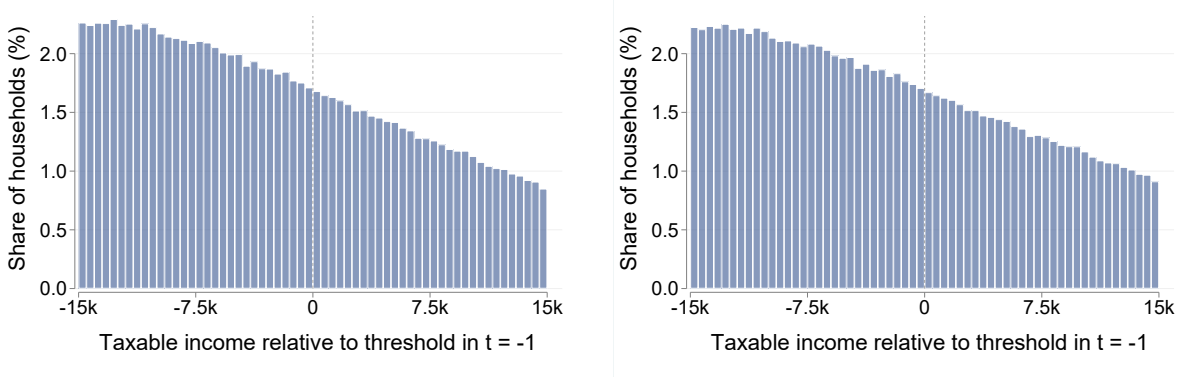
	Men	Women
A. The cells		
Cells	297	248
of which in the €0–10,000 bin	93	102
Pre-event brackets	10	10
Individuals in the cells (millions)	1.39	1.40
predicted treated (share)	0.38	0.39
B. The parents in the year before the event		
Age at the event (mean)	53.3	51.1
Partner's age at the event (mean)	50.9	53.4
Married (share)	0.98	0.98
Dependent children (mean)	1.90	1.92
Own work income (€, mean)	27,021	18,126
Partner's income (€, mean)	21,632	42,481
Positive work income (share)	0.85	0.79
Predicted into a rate change (share)	0.32	0.30

Notes: Panel A counts the cells of the headline estimation (work income, all event years) and the individuals they hold in the year before the event. The pre-event brackets are the distinct marginal rates of the effective schedule (Section 2.1) at which the cells start, over the schedules of 2003–2023. Panel B describes the child's fiscal exit sample below the income cap in the year before the event (ages and marital status at the event), as in Panel B of Appendix Table E.1; partner's income sums work, self-employment and pension income.

Second, Panel B of Table 3 reports descriptive statistics on the individuals in these cells. At the first-born's fiscal exit, women are 51 and men 53 years old on average, 98 percent are

married and they have 1.9 dependent children. Women's partners earn €42,000 on average, against the women's own work income of €18,000, while men's partners earn €22,000, against €27,000. At the event, 30 percent of the women and 32 percent of the men are predicted into a rate change. Appendix E.1 gives the full description of the sample.

Third, we evaluate the identifying assumptions on the pooled sample of all cells, with the exercises that Section 3.3 illustrates for one example cell. Assumption 1 requires that households neither time the event nor bunch below the thresholds on account of its tax consequence. Either would show as a break in the density of household taxable income at the hypothetical post-event threshold, and Figure 5 shows none for either gender. The predicted treated and the predicted controls differ in partner income and are alike in own income (Appendix Table E.2). Assumption 2 requires that partner income has no channel of its own to income growth within the cell. Appendix Figure E.1 pools the validation-region check of Figure 4 across cells and shows that, before the event, income growth has no systematic gradient in household income on either side of the hypothetical threshold for either gender.[32] Appendix D reports the realised rate changes of the two groups in the example cell, a direct check of the mapping of equation (4). The pre-event placebo contrasts of Section 3.5 accompany every estimate below, and an estimate is flagged where they reject.



(a) Men (b) Women

Figure 5: Density tests at the hypothetical threshold, pooled across cells

Notes: Distribution of household taxable income at $t = -1$, measured relative to the hypothetical threshold at which the household would enter the 30 percent bracket after the child's exit (at its post-exit quotient), pooled across cells. The vertical axis is the share of households in each €500 bin, in percent. The estimated log-density discontinuities are −0.001 (s.e. 0.009) for men and −0.002 (s.e. 0.010) for women. Sample: the child's fiscal exit sample below the income cap (Appendix E.1), within €15,000 of the threshold (N = 776,649 men and 782,811 women).

[32]Partner-income groups finer than terciles weaken the requirement, and the matching implementation of Section 3.5 additionally aligns pre-event income paths within cells.

4.2 Main results

4.2.1 Aggregated elasticities

Table 4 reports the aggregate earnings and participation elasticities under four weighting schemes common in the literature, by gender and combined. Each aggregate is a weighted average of the estimates by own-income bin and partner-income tercile (Section 3.4). The weights are the shares of French taxpayers, the shares of parents at the first-born's fiscal exit, the shares of aggregate earnings, or the precision of the estimates, and the first three differ in the mass they place on low incomes.

Table 4: Aggregated earnings and participation elasticities

	Earnings elasticity			Participation elasticity		
Weighting	Total	Men	Women	Total	Men	Women
French taxpayers	0.229	0.169	0.289	0.247	0.212	0.283
	(0.031)	(0.035)	(0.051)	(0.024)	(0.039)	(0.029)
Event population	0.246	0.174	0.322	0.218	0.205	0.230
	(0.029)	(0.033)	(0.048)	(0.023)	(0.033)	(0.033)
Income-weighted	0.171	0.137	0.217	0.158	0.145	0.177
	(0.024)	(0.031)	(0.037)	(0.013)	(0.016)	(0.022)
Inverse-variance weighted	0.114	0.125	0.102	0.111	0.127	0.096
	(0.016)	(0.023)	(0.022)	(0.008)	(0.012)	(0.011)

Notes: Each entry is a weighted average of the cell estimates of the earnings or participation elasticity (Section 3.5), pooled over $t = 0$ to 3; the rows differ in the weights across the groups of own-income bin × partner-income tercile. *French taxpayers*: each own-income bin weighted by its share of French taxpayers (INSEE distribution of annual salary income and employment rates by gender, 2021, with the non-employed in the bottom bin), the bin's weight split equally over the partner-income terciles that contribute an estimate. *Event population*: each group weighted by its regression samples in the event year. *Income-weighted*: each bin weighted by its share of aggregate earnings in the same distribution, split in the same way. *Inverse-variance weighted*: each group weighted by the inverse of its variance. The men's €0–10,000 earnings bin (a single cell) and the women's €60,000–70,000 bin are left out of the aggregates. The Total column combines the genders with equal population shares in the French-taxpayer and inverse-variance rows, with their shares of the event population in the event row and with their shares of aggregate income in the income row. Standard errors in parentheses are clustered by person and treat the groups, and the two genders, as independent.

We treat the aggregate weighted to the population of French taxpayers as the headline. It gives an average earnings elasticity of 0.23 (0.29 for women and 0.17 for men) and an average participation elasticity of 0.25 (0.28 and 0.21). The weights of the event population give similar averages, 0.25 for earnings and 0.22 for participation. The earnings averages are in line with the meta-analysis of Neisser (2021), in which the estimates for income before deductions, the measure closest to our work income, have a median of 0.185 and a mean of 0.287. Compared with the published estimates of work-income responses in Figure 1, our

estimates by own income (Section 4.2.2) are close to the reform-based estimates in the middle of the distribution and to the small elasticities that most designs find in the top deciles.

Women's aggregates lie above men's under every scheme except inverse-variance weighting, for earnings and for participation. Under the French-taxpayer weights, the difference of 0.12 in the earnings elasticities is 1.9 standard errors. Part of it arises because more women than men have low incomes, where elasticities are largest (45 percent of women and 34 percent of men are in the bottom bin, which includes the non-employed). Section 4.4 returns to the differences by gender.

The two remaining schemes give lower averages, because they move mass away from low incomes, where the responses are largest (Section 4.2.2). Inverse-variance weights give the smallest averages (0.11 for earnings and for participation), since the cells at low incomes are the least precisely estimated. Weighting each individual by their contribution to the tax base, the convention in the taxable-income literature, gives an earnings elasticity of 0.17 (0.22 for women and 0.14 for men) and a participation elasticity of 0.16. These differences across schemes are suggestive evidence of heterogeneity in the elasticities along income.

4.2.2 Profiles along own income

The earnings elasticity declines with own income, most steeply at the bottom (Figure 6, men and women pooled, averaged over the four years from the event). It is 0.45 below €10,000, 0.23 and 0.22 in the two bins from €10,000 to €30,000 and between 0.04 and 0.09 from €30,000 to €60,000. The decline holds for both genders (Section 4.4).

The participation elasticity also declines with own income, most steeply at the bottom (Figure 7). It falls from 0.32 in the bottom bin to between 0.12 and 0.17 from €10,000 to €40,000 and to between 0.07 and 0.13 above. Like the earnings elasticity, it declines for both genders (Section 4.4).

4.2.3 Event-time profiles

The earnings response is largest in the event year and fades within two years, while the participation response persists (Figure 8, men and women pooled, and Appendix E.3 by

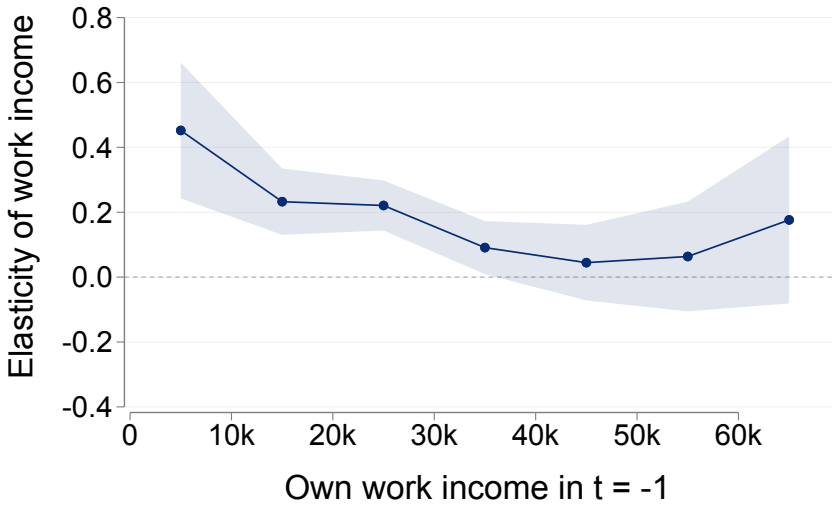


Figure 6: Earnings elasticity along the own-income distribution

Notes: Sample: the child's fiscal exit sample of Appendix E.1, all event years. Outcome: work-income growth relative to $t = -1$. Each point averages the cell estimates of equation (6) within an own-income bin, pooled over $t = 0$ to 3 and over men and women with their frequency weights (Section 3.4). Points are at bin midpoints of own work income at $t = -1$ and shaded areas are 95% confidence intervals from person-clustered standard errors. No bin fails its pre-event placebo. Appendix Figure E.4 reports the profile by gender.

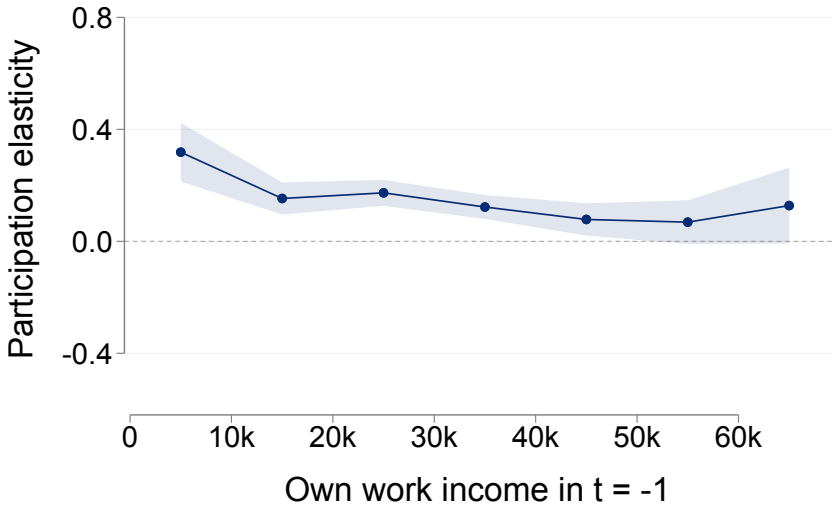
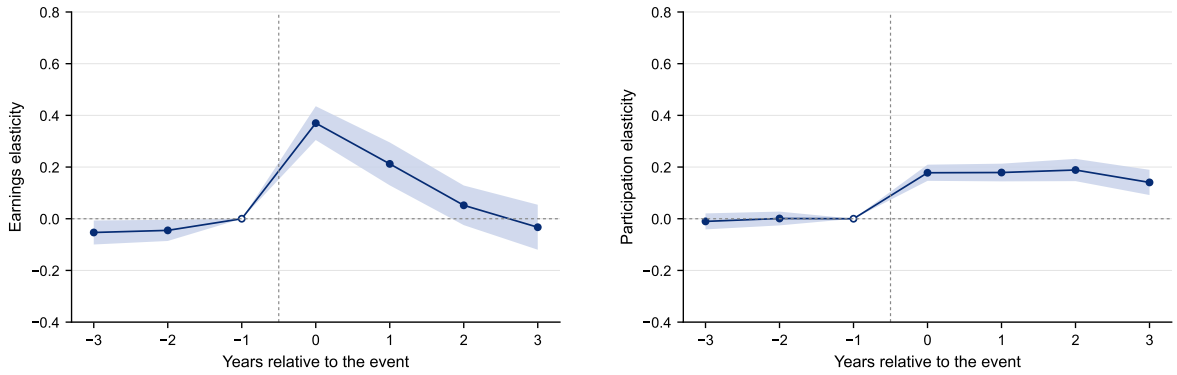


Figure 7: Participation elasticity along the own-income distribution

Notes: As Figure 6, with an indicator for positive work income as the outcome. No bin fails its pre-event placebo. Appendix Figure E.5 reports the profile by gender.

gender). The earnings elasticity is 0.37 in the event year, 0.21 a year later and indistinguishable from zero from $t = 2$. The participation elasticity is between 0.18 and 0.19 for three years

and 0.14 in the fourth. The estimates for $t = -3$ and $t = -2$ are placebos, because in those years the child is still in the tax unit and the household's schedule has not yet changed (Section 3.5).[33]



(a) Earnings elasticity (b) Participation elasticity

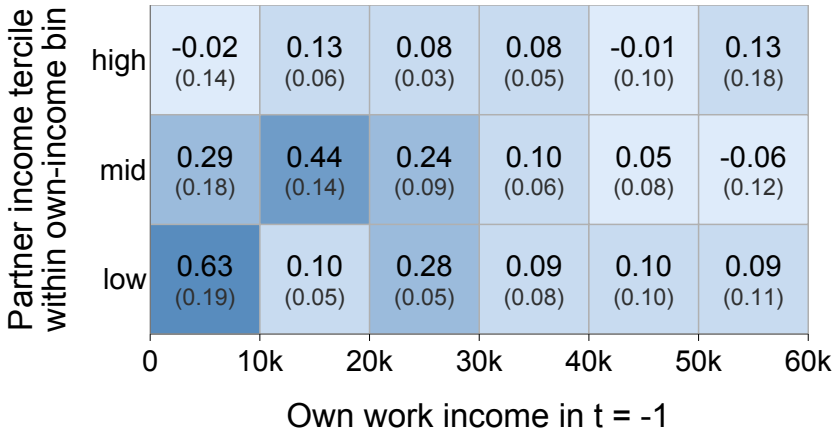
Figure 8: Event-time earnings and participation elasticities, men and women pooled

Notes: Sample and outcomes as in Figures 6 and 7. Each point pools the cell estimates of one year relative to the event over all own-income bins and both genders, with the frequency weights of Section 3.4. The year before the event, $t = -1$, is the reference year and is shown at zero (its hollow marker does not flag a failed placebo); the points at $t = -3$ and $t = -2$ are placebo estimates. Shaded areas are 95% confidence intervals from person-clustered standard errors, with the two genders treated as independent.

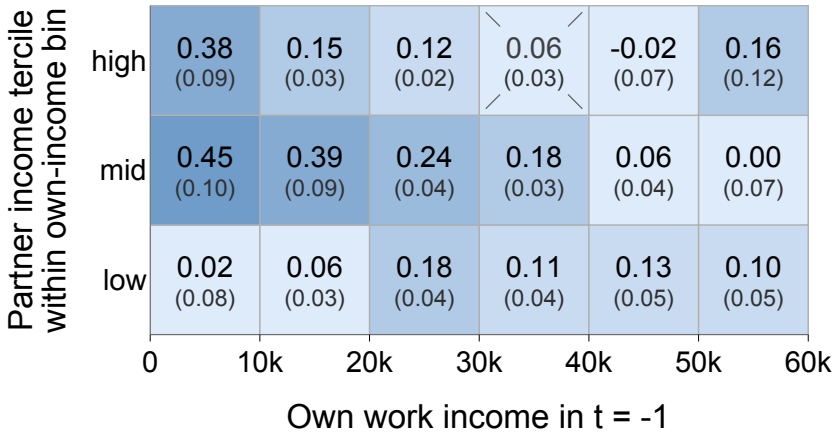
The headline earnings elasticity is the response to a predictable and permanent change in the marginal rate, averaged over the four years from the event. Households can foresee the exit but do not adjust before it (there is no bunching below the hypothetical post-event threshold in Figure 5, and the placebo estimates of Figure 8 are close to zero). A household that foresees the rate change learns nothing new about its lifetime income when the change arrives and responds to the rate alone, the intertemporal, or Frisch, response to anticipated tax changes (Looney & Singhal, 2006). A response to the rate alone is also what Assumption 2 requires, and without income effects the Frisch, compensated and uncompensated elasticities coincide (Section 3.1). The earnings response is concentrated in the event year and the year after, while the participation response persists through the window. The headline earnings elasticity measures a response at the time the schedule changes, which points to adjustments timed around the event, such as the exit from work of parents close to retirement (Section 4.4).

[33]Pre-event points are divided by the pre-event contrast in realised rates, which is smaller than in the event year. On the event-year scale the two earnings placebos are −0.01 to −0.02.

4.2.4 Maps over own and partner income



(a) Earnings elasticity



(b) Participation elasticity

Figure 9: Earnings and participation elasticities by own and partner income

Notes: Earnings and participation elasticities by own-income bin and within-bin partner-income tercile, men and women pooled. Sample, outcomes and estimates as in Figures 6 and 7 (elasticities with respect to the household net-of-marginal-tax rate, pooled over $t = 0$ to 3). Both incomes are measured at $t = -1$. Each tile pools the cell estimates of its own-by-partner-income rectangle with the frequency weights of Section 3.4. Beneath each value is its standard error, clustered by person within gender and with the two genders treated as independent. Where one gender's tile is not read (a dashed tile in Appendix Figure E.6), the pooled tile rests on the other gender's cells; the €0–10,000 earnings column rests on women's cells. Corner ticks and a grey value mark a tile that fails its pre-event placebo. Because the terciles are formed within each own-income bin, a row is not at constant partner income; Appendix Table E.4 reports the tercile boundaries. The grid ends at €60,000, the top of the own-income range supported for both genders. Appendix Figure E.6 shows the maps by gender.

Figure 9 adds the design's second income dimension and reports the elasticities by partner-income tercile within each own-income bin, men and women pooled. The maps show whether

the high elasticities at low own incomes come from households with a high-earning partner or from households in which both partners earn little. Because the terciles are formed within each own-income bin, the comparison across terciles holds own income fixed.

The answer differs between earnings and participation. For earnings, the bottom-bin response is concentrated where the partner also earns little, with an elasticity of 0.63 in the low partner-income tercile against 0.29 in the middle and -0.02 in the top tercile. For participation the pattern reverses, in that the most elastic appear to be those "insured" by higher-earning partners. The bottom-bin elasticity is 0.02 where the partner also earns little, against 0.45 in the middle and 0.38 in the top tercile. Above €20,000 the differences across terciles narrow, and from €30,000 up the elasticities are small in every tercile. Both patterns are mostly driven by women (Section 4.4).

The maps are consistent with partner income acting as insurance. A lower earner whose partner earns considerably more can afford to leave work, while a household in which both partners earn little can afford to give up neither income. Hours may also be easier to adjust at low pay, where part-time work is common, and earnings can then change without leaving work.

4.3 Robustness

In Table 5 we evaluate the robustness of the earnings map to alternative estimation choices and sample restrictions. To do so, we re-estimate the map under each variant and calculate the correlation, rank correlation and mean absolute difference between its tiles and those of the headline map, as well as the gradient of the tile values in own income.

The balanced sample holds its composition fixed over the event window. The resulting map remains highly correlated with the headline map, which indicates that changes in sample composition do not drive the main results.

Another check uses the sample without the 2020–2023 event years. These years follow the introduction of withholding at source in 2019. Excluding them leaves the earnings map essentially unchanged.

We also re-estimate the earnings map with nearest-neighbour matching within cells in place of the regression comparison. The gradient keeps its negative sign, but the rank correlation with the headline map is somewhat lower, which reflects the noisier matching estimates.

The estimates are also robust to a stricter first-stage requirement, which drops the cells with weak first stages, and to weak-instrument-robust inference (Appendix E.8).

Lastly, we estimate the map by Poisson regression of the level of earnings, which gives the combined response of earnings and participation. The combined response is close to the earnings response.

Table 5: Robustness of the earnings map

	Corr.	Rank corr.	mAD	Gradient
Headline	1.00	1.00	0.00	−0.06 (0.02)
Balanced sample	0.94	0.92	0.04	−0.08 (0.02)
Without 2020–2023 event years	0.97	0.89	0.03	−0.07 (0.02)
Nearest-neighbour matching	0.71	0.42	0.13	−0.05 (0.02)
Poisson, level of earnings	0.76	0.69	0.06	−0.03 (0.02)

Notes: Each row re-estimates the earnings map of Figure 9 under one of the estimation variants of Section 3.5, men and women pooled. Corr. and Rank corr. are the Pearson and Spearman correlations of the tile values with the headline map, and mAD is their mean absolute difference. Gradient is the slope of the tile values in own income per €10,000, each tile weighted by its estimation sample, with its standard error in parentheses (tiles treated as independent).

The marriage design applies the strategy to a second event, the merger of the spouses' two tax units at a first marriage (Appendix F). The women's earnings elasticities decline with own income as at the child's fiscal exit, and the men's are not different from zero above €20,000.

4.4 Further heterogeneity

This subsection reports the estimates by gender, by age, for a broader definition of income and by household characteristics.

By gender, the elasticities decline with own income for women and for men alike (Appendix Figures E.4 and E.5), and the two genders differ mainly in the role of partner income. With similar profiles, women's higher aggregates, 0.29 against 0.17 for earnings and 0.28 against 0.21 for participation under the French-taxpayer weights (Table 4), arise in part because more women than men have low incomes, where the elasticities are largest (Section 4.2.1). Over partner income, by contrast, the patterns of the maps are mostly driven by women (Appendix

Figure E.6). Among the lowest-earning women, the participation elasticity is 0.69 with a middle-tercile and 0.35 with a top-tercile partner, against 0.02 where the partner also earns little. Men's maps show less partner-income heterogeneity.

By age, the earnings elasticities are smaller for individuals below the minimum retirement age. For those below 58 in the year before the event (below the minimum retirement age in every year of the data), the French-taxpayer aggregates for incomes from €10,000 up are 0.08 for earnings and 0.13 for participation, against 0.20 and 0.16 at all ages (Appendix Table E.7). The difference is larger for men than for women.

A broader definition of income, which adds self-employment income to work income, shifts the profile up (Appendix Figure E.7). From €10,000 up the shift is small, with broad elasticities within 0.01 of the earnings elasticities between €10,000 and €30,000 and 0.04 to 0.06 above them from €30,000 up.

Appendix E.10 reports the estimates by household characteristics.

5 On the relevance of accounting for heterogeneity

This section gives two examples of calculations into which the earnings elasticity enters and shows what accounting for its heterogeneity along the income distribution changes. The first solves for the optimal tax schedule in the Mirrlees–Saez framework (Mirrlees, 1971; Saez, 2001), with the estimated profile over the income distribution in place of a single elasticity (Section 5.1). The second applies the perturbation approach of Bierbrauer et al. (2024) to reforms towards individual taxation of couples, with the estimated surfaces over own and partner income in place of one elasticity per spouse (Section 5.2). Both calculations use the elasticities of men and women pooled, averaged over the four years from the event, and apply them to the whole income distribution. This assumes that the heterogeneity along income estimated for parents at the first-born's fiscal exit holds for earners of all ages and family situations.

5.1 In optimal tax models

The Mirrlees problem is the workhorse model for the design of the tax and transfer system (Mirrlees, 1971). A government that values redistribution chooses a nonlinear income tax to maximise social welfare subject to a revenue requirement, and because it observes incomes but not abilities, the tax can depend on income only. At each income, the optimal schedule balances the revenue that a higher marginal rate collects from everyone above that income against the fall in earnings of those who face it. Saez (2001) shows that the optimal schedule can be written in terms of a few estimable quantities (the sufficient statistics).

The optimal marginal rate at income z satisfies

$$\frac{T'(z)}{1-T'(z)} \;=\; \frac{1}{e(z)}\,\big(1-G(z)\big)\,\frac{1-H(z)}{z\,h(z)}, \tag{7}$$

where $T'(z)$ is the marginal rate at z, $e(z)$ the earnings elasticity at z, $G(z)$ the average social marginal welfare weight on earners above z, and $H(z)$ and $h(z)$ the distribution and density of income. A common elasticity is not required. In the Mirrlees problem, the incentive-compatibility constraint at each income depends on the elasticity at that income. Elasticities that vary with income alter these constraints but not the structure of the problem, and equation (7) holds with $e(z)$ evaluated at each z. The calibrations in the literature nevertheless assume a single elasticity at every income (Brewer et al., 2010; Saez, 2001). We solve the problem once with a single elasticity of 0.23, the French-taxpayer aggregate of Section 4.2.1, and once with the estimated profile along own income of Section 4.2.2, and compare the two schedules.[34]

The schedule is computed numerically, following Brewer et al. (2010), because the income distribution and the welfare weights on the right-hand side of equation (7) depend on the schedule itself. The computation works in skill space, iterates on equation (7) until the schedule is consistent with the income distribution and the welfare weights it generates, and carries each type's own elasticity in its incentive-compatibility constraint. The distribution of skills is inferred from the 2023 administrative distribution of positive work incomes, men and

[34]The profile is a smooth fit to the seven pooled bin estimates of Figure 6, held fixed below the first and above the last bin midpoint. Each type takes the value at its 2023 income and keeps it as its earnings change.

women weighted equally, with a Pareto tail fitted above €70,000.[35] Preferences are quasi-linear and iso-elastic, as in Section 3.1, and social welfare is the sum of the logarithms of utilities, under which the welfare weights decline with income.

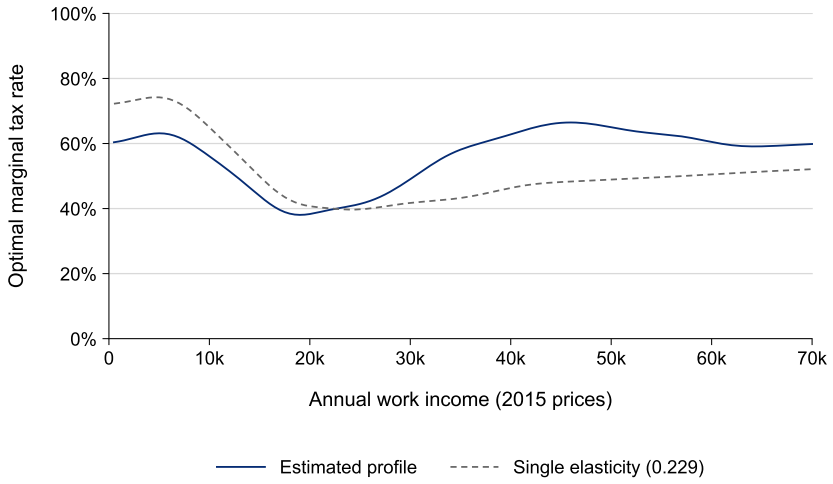


Figure 10: Optimal marginal tax rates under a single elasticity and under the estimated profile

Notes: Marginal rates of the two optimal schedules at each income. The income distribution is that of positive earners in the 2023 administrative data, men and women weighted equally, with a Pareto tail above the last income bin, and social welfare is logarithmic. The grey dashed line uses the single elasticity of 0.229 and the navy line the fitted profile of elasticities along own income. Incomes are in 2015 prices. The displayed rates are €500-bin medians smoothed with a window of half-width €5,000, and the numbers in the text are from the unsmoothed solution. The figure is shown up to €70,000, the range of the estimates.

With the estimated profile in place of the single elasticity, the optimal schedule withdraws transfers more slowly at the bottom and is more progressive above the median (Figure 10). Below €10,000, where the estimated elasticity is largest, the optimal marginal rate is on average 11 percentage points lower (62 against 73 percent), and between €10,000 and €20,000 it is 6 points lower. The two schedules cross at €22,000. Between €30,000 and €70,000, where the estimated elasticities are small, the rate with the profile is higher by 8 to 18 points. A fast withdrawal of the transfer reduces earnings most where responses are largest, and the revenue forgone at the bottom is recovered above the median, where a higher rate loses little in earnings. Beyond the 98th percentile of earners, the upper end of the range of the estimates, the calculation carries the last estimated elasticity forward.

[35] Each observed income is read as the earnings a type would choose under a flat marginal rate of 41 percent. The revenue requirement is €4,000 per person.

5.2 In the assessment of jointness reforms

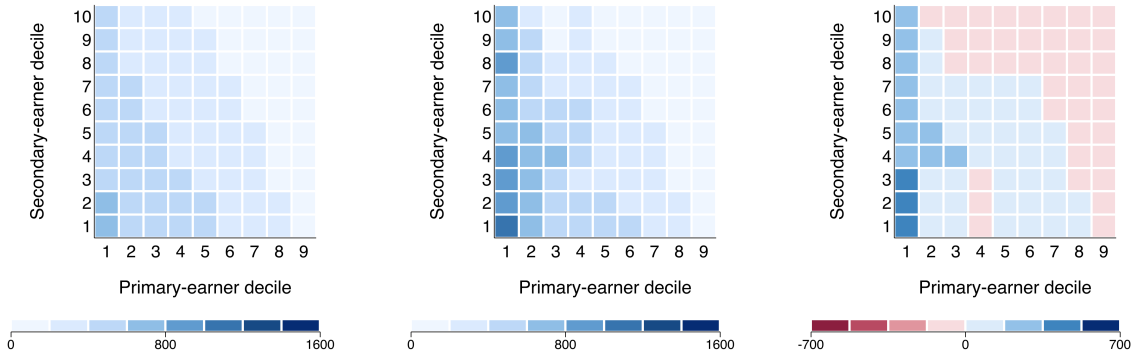
For couples, the optimal tax schedule is a multidimensional screening problem, which is analytically demanding even under strong simplifying assumptions (Golosov & Krasikov, 2025; Kleven et al., 2009). Golosov and Krasikov (2025) show that, with a single elasticity common to all individuals, it is optimal that secondary earners face on average higher marginal rates than primary earners (transfers targeted to couples with low secondary earnings are withdrawn as those earnings rise). Once elasticities differ between spouses, the direction of the optimal differential is open, because the inverse-elasticity rule of Boskin and Sheshinski (1983) prescribes lower rates on the more responsive spouse, typically the secondary earner. It depends on how much the two spouses' elasticities differ and at which incomes, which the maps of Section 4.2.4 estimate.

Rather than solving for the optimal schedule, we ask whether small reforms that bring the French system closer to individual taxation raise welfare, following Bierbrauer et al. (2024) and Bierbrauer et al. (2026). Starting from joint taxation, under which both spouses face the household's marginal rate, they derive sufficient statistics for reforms that lower the rate on secondary relative to primary earnings in a revenue-neutral way. The earnings and participation elasticities of the two spouses enter these statistics. Because each reform is local to couples at a given pair of incomes, its welfare effect depends on the responses of the couples at those incomes, which the estimated surfaces over own and partner income supply.[36]

We run the calculation under two calibrations, both with the statutory 2019 French schedule for a couple with quotient 2 and with equal welfare weights. The first uses the French-taxpayer aggregates of Table 4 for both spouses, an earnings elasticity of 0.23 and a participation elasticity of 0.25. The second uses the estimated earnings and participation surfaces over own income and partner-income tercile.

Figure 11 reports the welfare effect of one local reform for each pair of primary- and secondary-earner deciles, under the single elasticity (Panel (a)) and under the estimated elasticities (Panel (b)), and their difference (Panel (c)). Accounting for the heterogeneity raises the welfare gains of such reforms on average, and most for couples in which both partners earn

[36]The baseline calibration of Bierbrauer et al. (2024) and Bierbrauer et al. (2026) sets the elasticity of primary earnings to 0.25 and that of secondary earnings to 0.75. The single-elasticity calibration below uses one value for both spouses, which isolates the role of the heterogeneity we estimate.



(a) Single elasticity (b) Estimated elasticities (c) Difference, (b) minus (a)

Figure 11: Welfare effects of local reforms towards more individualised taxation

Notes: Each square is a pair of primary- and secondary-earner deciles and reports the welfare effect of one local reform at those incomes, a revenue-neutral reduction of the marginal rate on secondary relative to primary earnings by one percentage point, integrated over a €500-by-€500 range of the two spouses' incomes, following Bierbrauer et al. (2024). The scales are in model welfare units, not euros. Blue denotes gains in Panels (a) and (b). In Panel (c), blue denotes a larger gain under the estimated elasticities than under the single elasticity and red a smaller gain. Panel (a) sets the earnings elasticity to 0.229 for both spouses and the participation elasticity to 0.247. Panel (b) uses the earnings and participation surfaces behind Table 4, over own-income bins and partner-income terciles with men and women pooled equally in each, and averages the two spouses' participation elasticities in the household participation term. Marginal rates follow the statutory 2019 French schedule (Section 2.1) for a couple with quotient 2, after the 10% wage-income deduction and before tax credits and child-related adjustments. The primary earner is the higher-earning spouse, and secondary-earner deciles are formed within each primary-earner decile. A square is shown only when the incomes of both spouses lie within the support of the estimates, which excludes the tenth primary-earner decile. All panels use equal welfare weights.

little, where the maps of Section 4.2.4 place the largest earnings response. The mean welfare effect over the 90 local reforms rises from 308 to 350 model welfare units, and the effect rises in 51 reforms and falls in 39. Under near-Rawlsian weights the mean becomes less negative, from −645 to −603 (Appendix Figure E.9).[37]

6 Conclusion

In a panel built from the universe of French tax returns, this paper estimates the elasticity of earnings with respect to taxation along the income distribution. Averaged over the French taxpayer population, the earnings elasticity is 0.23 (0.29 for women and 0.17 for men) and the participation elasticity is 0.25.

Identification comes from a novel source of variation. Typically, the elasticity is identified from tax reforms, which change the tax rates of some income groups and not of others. Such

[37] Heterogeneous elasticities may also change whether moves towards individualisation are politically feasible (Bierbrauer et al., 2025).

between-income variation requires assumptions on how the incomes of these groups would have evolved without the reform (Jakobsen & Søgaard, 2022). We instead use *within-income* variation in exposure to the tax change that occurs when a child leaves the parents' tax unit.

Under joint taxation, whether a change in the composition of the tax unit moves an individual into another bracket depends on partner income as well as on own income. This permits comparisons, at every level of own income, between individuals who face different predicted tax changes. Own income can therefore be controlled for flexibly without absorbing the identifying variation, which reduces the scope for bias from mean reversion and differential income trends.

A further advantage of the approach is that it estimates the elasticity along the income distribution with the same method, in the same country and for the same income measure. The profile can therefore serve as a benchmark for the local estimates of other studies, whose differences may reflect the income ranges they cover as well as their settings and methods. The earnings elasticity declines with income for both genders, most steeply at the bottom: from 0.45 in the bottom decile to 0.23 in the middle deciles and to below 0.1 in the top decile.

The approach can be applied in other countries and settings in which joint taxation or similar rules make tax exposure differ among individuals with the same own income. This paper studies the child's fiscal exit and applies the approach to marriage as a second event (Appendix F). Other events that change household composition, such as the birth of a child or a divorce, could provide further applications.

References

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, *20*(1), 46–63.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, *90*(430), 431–442.

Auten, G., & Carroll, R. (1999). The effect of income taxes on household income. *The Review of Economics and Statistics*, *81*(4), 681–693.

Bastani, S., & Selin, H. (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, *109*, 36–49.

Bick, A., & Fuchs-Schündeln, N. (2018). Taxation and labour supply of married couples across countries: A macroeconomic analysis. *The Review of Economic Studies*, *85*(3), 1543–1576.

Bierbrauer, F. J., Boyer, P. C., Peichl, A., & Weishaar, D. (2024). *Pareto-improvements, welfare trade-offs and the taxation of couples* (CESifo Working Paper No. 11539). CESifo.

Bierbrauer, F. J., Boyer, P. C., Peichl, A., & Weishaar, D. (2025). *The political economy of joint taxation* (CESifo Working Paper No. 11728). CESifo.

Bierbrauer, F. J., Boyer, P. C., Peichl, A., & Weishaar, D. (2026). *Tax reforms and multi-dimensional screening* (Working paper, April 2026 version).

Blandhol, C., Bonney, J., Mogstad, M., & Torgovitsky, A. (2026). When is TSLS actually LATE? [Advance access]. *The Review of Economic Studies*.

Blomquist, S., & Selin, H. (2010). Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. *Journal of Public Economics*, *94*(11-12), 878–889.

Blundell, R., Duncan, A., & Meghir, C. (1998). Estimating labor supply responses using tax reforms. *Econometrica*, *66*(4), 827–861.

Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, *89*(1), 181–213.

Boskin, M. J., & Sheshinski, E. (1983). Optimal tax treatment of the family: Married couples. *Journal of Public Economics*, *20*(3), 281–297.

Brewer, M., Saez, E., & Shephard, A. (2010). Means-testing and tax rates on earnings. In J. A. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, & J. M. Poterba (Eds.), *Dimensions of tax design: The Mirrlees review* (pp. 90–173). Oxford University Press.

Burns, S. K., & Ziliak, J. P. (2017). Identifying the elasticity of taxable income. *Economic Journal*, *127*(600), 297–329.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2026). Difference-in-differences with a continuous treatment [Forthcoming]. *American Economic Review*.

Chen, J., & Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, *139*(2), 891–936.

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.

Chetty, R., Friedman, J. N., & Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review*, *103*(7), 2683–2721.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy*, *100*(3), 437–467.

Diamond, P., & Saez, E. (2011). The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives*, *25*(4), 165–190.

Eissa, N. (1996). Tax reforms and labor supply. In J. M. Poterba (Ed.), *Tax policy and the economy, volume 10* (pp. 119–151). MIT Press.

Eissa, N., & Liebman, J. B. (1996). Labor supply response to the Earned Income Tax Credit. *The Quarterly Journal of Economics*, *111*(2), 605–637.

Feldstein, M. (1995). The effect of marginal tax rates on taxable income: A panel study of the 1986 Tax Reform Act. *Journal of Political Economy*, *103*(3), 551–72.

Gelber, A. M. (2014). Taxation and the earnings of husbands and wives: Evidence from Sweden. *The Review of Economics and Statistics*, *96*(2), 287–305.

Golosov, M., & Krasikov, I. (2025). The optimal taxation of couples. *The Quarterly Journal of Economics*, *140*(3), 2163–2211.

Goux, D., Maurin, E., & Petrongolo, B. (2014). Worktime regulations and spousal labor supply. *American Economic Review*, *104*(1), 252–276.

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.

Gruber, J., & Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, *84*(1), 1–32.

He, D., Peng, L., & Wang, X. (2021). Understanding the elasticity of taxable income: A tale of two approaches. *Journal of Public Economics*, *197*, 104375.

Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy*, *30*(1), 51–89.

Herold, E., Koch, L., & Neisser, C. (2025, November). *Joint taxation and intra-household inequality: Evidence from same-sex couples* (Working paper).

Herold, E., & Wallossek, L. (2024). The marriage earnings gap. *Working Paper*.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Isaac, E. (2023). Suddenly married: Joint taxation and the labor supply of same-sex married couples after *U.S. v. Windsor*. *Journal of Human Resources*.

Jakobsen, K. M., & Søgaard, J. E. (2022). Identifying behavioral responses to tax reforms: New insights and a new approach. *Journal of Public Economics*, *212*, 104691.

Kalíšková, K. (2014). Labor supply consequences of family taxation: Evidence from the czech republic. *Labour Economics*, *30*, 234–244.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, *8*(1), 435–464.

Kleven, H. J., Kreiner, C., Larsen, K., & Søgaard, J. (2025). Micro versus macro labor supply elasticities: The role of dynamic returns to effort. *American Economic Review*, *115*(9), 2849–2890.

Kleven, H. J., Kreiner, C. T., & Saez, E. (2009). The optimal income taxation of couples. *Econometrica*, *77*(2), 537–560.

Kleven, H. J., & Schultz, E. A. (2014). Estimating taxable income responses using Danish tax reforms. *American Economic Journal: Economic Policy*, *6*(4), 271–301.

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.

Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, *89*(11), 2093–2119.

Kumar, A., & Liang, C.-Y. (2020). Estimating taxable income responses with elasticity heterogeneity. *Journal of Public Economics*, *188*, 104209.

LaLumia, S. (2008). The effects of joint taxation of married couples on labor supply and non-wage income. *Journal of Public Economics*, *92*(7), 1698–1719.

Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid t-ratio inference for IV. *American Economic Review*, *112*(10), 3260–3290.

Lehmann, E., Marical, F., & Rioux, L. (2013). Labor income responds differently to income-tax and payroll-tax reforms. *Journal of Public Economics*, *99*, 66–84.

Looney, A., & Singhal, M. (2006). *The effect of anticipated tax changes on intertemporal labor supply and the realization of taxable income* (NBER Working Paper No. 12417). National Bureau of Economic Research.

Miao, D., Selin, H., & Söderström, M. (2025). Earnings responses to even higher taxes. *The Economic Journal*, *135*(667), 838–860.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, *38*(2), 175–208.

Montiel Olea, J. L., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), 358–369.

Neisser, C. (2021). The elasticity of taxable income: A meta-regression analysis. *The Economic Journal*, *131*(640), 3365–3391.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, *68*(1), 205–229.

Saez, E. (2003). The effect of marginal tax rates on income: A panel study of "bracket creep". *Journal of Public Economics*, *87*(5-6), 1231–1258.

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

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, *50*(1), 3–50.

Selin, H. (2014). The rise in female employment and the role of tax incentives: An empirical analysis of the Swedish individual tax reform of 1971. *International Tax and Public Finance*, *21*(5), 894–922.

Sigaard, H. S. (2023). Estimating labor supply responses to Danish tax reforms. *Journal of Public Economics*, *224*, 104943.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of thomas rothenberg* (pp. 80–108). Cambridge University Press.

Weber, C. E. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. *Journal of Public Economics*, *117*, 90–103.

A The household model and the estimand

This appendix relaxes the separability assumption of Section 3.1. It derives the household estimating equation under a general convex joint cost of earning income, shows that joint filing collapses it to equation (2) with the coefficient read as the elasticity of own income with respect to the household rate, and states which elasticities the two events identify.

We model a household choosing both spouses' incomes subject to a joint budget constraint,

$$\max_{z_{is},\, z_{ps}} \quad c_s - \Psi\big(z_{is}, z_{ps};\, n_{is}, n_{ps}\big) \qquad \text{s.t.} \qquad c_s \le z_{is} + z_{ps} - \mathcal{T}_s(z_{is}, z_{ps}),$$

where z_{ps} is the partner's income, $\mathcal{T}_s$ is the couple's tax function, and Ψ is a convex joint cost of earning income. The form of $\mathcal{T}_s$ depends on the filing regime (Section 2.1), and our design uses this dependence. The joint cost may be additively separable across spouses or non-separable. Non-separability can reflect leisure complementarities, home production, or coordinated retirement (Gelber, 2014; Goux et al., 2014).[38]

The potential incomes n_{is} and n_{ps} index each spouse's cost of earning income; we normalise them so that n_{is} is the income i would earn if both spouses faced marginal tax rates of zero, as in the separable case of Section 3.1.

With $\tau_{is} \equiv 1 - \partial\mathcal{T}_s/\partial z_{is}$ and $\tau_{ps} \equiv 1 - \partial\mathcal{T}_s/\partial z_{ps}$ denoting the spouses' marginal net-of-tax rates, the first-order conditions are $\partial\Psi/\partial z_{is} = \tau_{is}$ and $\partial\Psi/\partial z_{ps} = \tau_{ps}$. Totally differentiating the pair around the pre-event optimum and solving the resulting two-equation system yields the elasticities

$$\text{(A.1)} \qquad \epsilon_{ii} \;=\; \frac{\tau_{is}}{z_{is}}\,\frac{\Psi_{pp}}{D}, \qquad \epsilon_{ip} \;=\; -\,\frac{\tau_{ps}}{z_{is}}\,\frac{\Psi_{ip}}{D}, \qquad D \;\equiv\; \Psi_{ii}\Psi_{pp} - \Psi_{ip}^2 \;>\; 0,$$

[38] Quasi-linearity in household consumption shuts down income effects at the household level, in the sense that neither the partner's net income nor the change in the household's tax *liability* affects income choices. It also makes the intra-household allocation of consumption irrelevant for those choices, because with transferable utility any sharing rule supported by lump-sum transfers leaves the first-order conditions unchanged. For the same reason the model's predictions for incomes coincide with those of an efficient collective household à la Chiappori (1992) with quasi-linear private utilities. We need not take a stance on the household decision process. Gelber (2014) rejects the income-pooling restriction of the unitary model. Because the rejection operates through income effects, which quasi-linearity rules out, the income predictions used here do not rest on the restriction. See Appendix C for the extension with income effects.

where subscripts on Ψ denote second derivatives with respect to the spouses' incomes, evaluated at the optimum, and $D > 0$ by strict convexity. Solving the system also yields the household estimating equation in changes,[39]

$$\Delta \ln z_{is} = \epsilon_{ii} \, \Delta \ln \tau_{is} + \epsilon_{ip} \, \Delta \ln \tau_{ps} + \Delta \ln n_{is}. \tag{A.2}$$

Every coefficient in (A.2) is an elasticity of i's *own* income, and the subscripts index whose rate varies: ϵ_{ii} is the response of i's income to i's own rate, ϵ_{ip} its response to the *partner's* rate. As derivatives of the household's optimal income-supply function $z_i^*(\tau_{is}, \tau_{ps})$, both are jointly re-optimised responses in which the partner's income also adjusts. The cross-derivative Ψ_{ip} is the numerator of ϵ_{ip} and enters ϵ_{ii} through D. Even when i's rate alone changes, the partner's re-optimisation raises ϵ_{ii} above the one-person elasticity $\tau_{is}/(z_{is}\Psi_{ii})$ by the factor $1/(1-\rho^2)$, with $\rho \equiv \Psi_{ip}/\sqrt{\Psi_{ii}\Psi_{pp}}$. Under quasi-linearity the cross-rate response is transmitted entirely through the partner's behavioural adjustment, and it is zero when costs are separable, $\Psi_{ip} = 0$.[40]

Under joint filing, the marginal rate is common to both spouses, $\tau_{is} = \tau_{ps} = \tau_{is}^H$ (Section 2.1). Any change in it is then common as well, and equation (A.2) collapses to the estimating equation (2) of the main text,

$$\Delta \ln z_{is} = \underbrace{\left(\epsilon_{ii} + \epsilon_{ip}\right)}_{\equiv \, \epsilon_i^H} \Delta \ln \tau_{is}^H + \Delta \ln n_{is},$$

with the coefficient read as the *elasticity of own income with respect to the household rate*, $\epsilon_i^H \equiv \epsilon_{ii} + \epsilon_{ip}$. This elasticity includes the component transmitted through the partner's adjustment. From (A.1), $\epsilon_i^H = (\tau_{is}^H/z_{is})\,(\Psi_{pp} - \Psi_{ip})/D$. It nests the one-person elasticity at $\Psi_{ip} = 0$ and exceeds it under complementarity ($\Psi_{ip} < 0$) through both the cross term and the factor $1/(1-\rho^2)$.

[39] For additively separable, iso-elastic Ψ the relation is exact and $\epsilon_{ip} = 0$, while in general the elasticities are local to the pre-event optimum. The error term is composite. With non-separable costs, changes in the *partner's* potential income also shift i's income through the same cross-derivative Ψ_{ip}, and $\Delta \ln n_{is}$ stands for the combined contribution of both spouses' potential-income changes to i's income. Assumption 1 requires that this component does not vary with the dose within cells.

[40] Separability is the assumption underlying most quantitative work on the taxation of couples (Bick & Fuchs-Schündeln, 2018).

Under non-separable costs, all three elasticities are couple-level parameters, because both spouses' cost curvatures enter each of them. They are local to the couple, in the sense that the same individual would have a different elasticity if single. Estimates from couples are not directly comparable to estimates on samples of singles. A gradient of responses in partner income at fixed own income can come from sorting of individual elasticities across partner types, from income effects (Appendix C), from the non-separability studied here, or from a common cost function whose local elasticity varies with the household rate, which differs across partner-income terciles. The design identifies the gradient without separating these sources.

At the child's fiscal exit, the household files jointly before and after the event. Because the event-induced rate change is common to both spouses, the design recovers the elasticity of own income with respect to the household rate, ϵ_i^H.[41]

At marriage, by contrast, spouses are individually taxed before the event. The two spouses' predicted changes generally differ (often only one spouse is predicted to see any change). Among couples for whom the event predicts a change in i's rate only, the event-induced variation moves $\Delta \ln \tau_{is}$ alone, and the reduced form identifies the own-rate elasticity ϵ_{ii}, which includes the partner's re-optimisation (the factor $1/(1-\rho^2)$ above). Among the same couples, the *partner's* income response to i's rate change identifies the cross-rate elasticity ϵ_{pi} directly, as in the design of Gelber (2014).

The gap between ϵ_i^H and ϵ_{ii} is likely small. Exploiting the Swedish tax reform of 1990–1991, in which spouses' marginal rates moved independently, Gelber (2014) estimates compensated cross-rate elasticities of 0.05 for husbands and 0.07 for wives, an order of magnitude below his own-rate estimates of 0.41 and 0.47. Because these cross-rate elasticities are positive, ϵ_i^H weakly overstates ϵ_{ii}. Where the distinction matters, the marriage design can in principle provide internal estimates, or external ones can be imputed: $\epsilon_{ii} = \epsilon_i^H - \hat{\epsilon}_{ip}$.[42]

[41] The income-weighted average of ϵ_i^H across spouses gives a different object, the response of the couple's *combined* income to the household rate. Revenue and Laffer calculations need this response, whichever way it splits across own- and cross-rate channels.

[42] Two routes lead from the identified total to the own-rate elasticity. If costs are separable ($\Psi_{ip} = 0$), then $\epsilon_{ip} = 0$ and $\epsilon_i^H = \epsilon_{ii}$. Otherwise $\hat{\epsilon}_{ip}$ can be imputed from Gelber (2014) and $\epsilon_{ii} = \epsilon_i^H - \hat{\epsilon}_{ip}$ reported as an adjustment or as bounds.

B Proof of Proposition 1

Proof. By (4), (5) and Assumption 2, $\Delta y_{it} = \Delta_t Y_i(0) + \mathbb{1}[Z_{it} \neq 0]\,\epsilon_i \Phi_{it}$ and $\Delta \mathrm{NTR}_{it} = \Delta_t \mathrm{NTR}_i(0) + \mathbb{1}[Z_{it} \neq 0]\,\Phi_{it}$. By Assumption 1, the conditional means of the untreated components are equal for the $Z_i = \delta$ and $Z_i = 0$ groups of the same cell and cancel in the contrasts: $RF_c^t(\delta) = \mathbb{E}_c[\epsilon_i \Phi_{it} \mid Z_i = \delta]$ and $FS_c^t(\delta) = \mathbb{E}_c[\Phi_{it} \mid Z_i = \delta]$. Taking the ratio gives the weighted-average representation, with weights averaging to one by construction. Under Assumption 3, all Φ_{it} share the sign of δ, which is also the sign of their conditional mean, and the weights are non-negative. □

C Model extensions

C.1 Income effects

Relaxing quasi-linearity, we replace linear consumption with a concave utility $v(c_s)$: the household problem becomes

$$\max_{z_{is},\, z_{ps}} \; v(c_s) - \Psi\big(z_{is}, z_{ps};\, n_{is}, n_{ps}\big) \qquad \text{s.t.} \qquad c_s \leq z_{is} + z_{ps} - T_s^H(z_{is} + z_{ps};\, Q_{is}) - \kappa_{is},$$

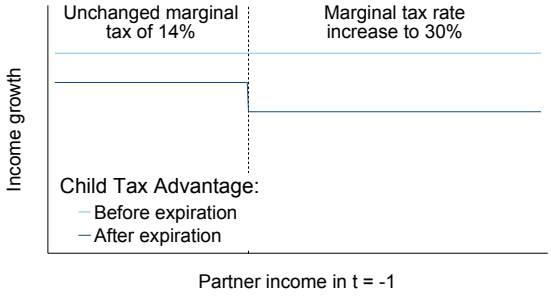
where κ_{is} collects lump-sum household expenses, which may change at the event (expenses for the child). The first-order conditions become

$$\frac{\partial \Psi}{\partial z_{js}} \;=\; v'(c_s)\,\tau_{js}^H, \qquad j \in \{i, p\}:$$

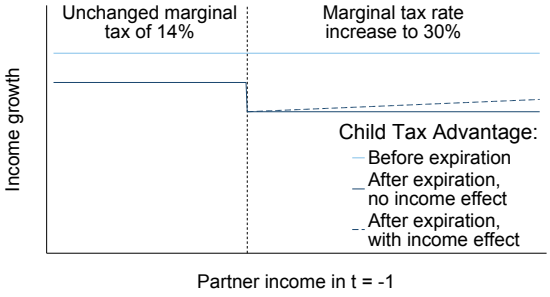
income choices now depend on the marginal utility of household consumption and through it on partner income, the level of the tax *liability* and expenses κ_{is}. Log-linearising adds a virtual-income term to equation (2), as in Gruber and Saez (2002). Two remarks follow for our design. First, households predicted to cross a threshold at the event also experience a larger increase in tax liability than their within-cell controls. Absent a virtual-income instrument, our estimates should be read as total responses to the schedule change. The two channels oppose each other, because the higher marginal rate reduces income while the larger liability raises it when leisure is normal. Under income effects our estimates are lower bounds on

the compensated elasticity. Second, the change in expenses is part of the event, not of the schedule change. Because the child leaves and the expenses change under both schedules (Section 3.3), the income consequence of that change belongs to untreated growth. Even a change in expenses common to every household of a cell moves income through $v'(c_s)$ when v is concave. Since consumption varies with partner income within the cell's partner-income tercile, income effects can convert a common event shock into a partner-income *gradient* in growth. This is the class of violation the validation region detects.[43] Figure C.1 plots income growth against partner income on either side of the threshold, without income effects in Panel (a) and with an income effect that grows with the distance above the threshold in Panels (b) and (c), for a linear and a progressive schedule. Panel (d) replaces the flat growth of Panels (a) to (c) with a constant partner-income gradient in income growth that the event itself does not change.

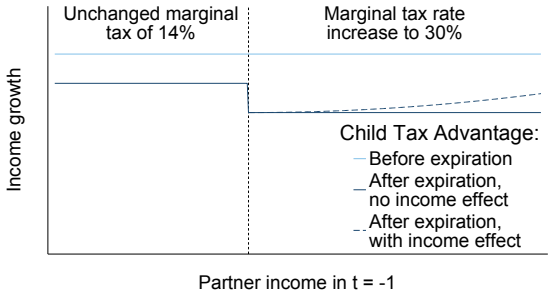
[43] Appendix G discusses income effects as one reason why first-stage monotonicity can fail in reform-based simulated-IV designs and explains why this does not happen in ours.



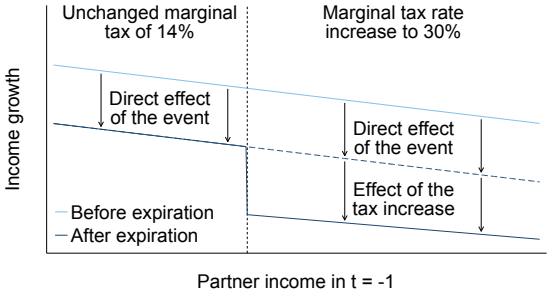
(a) Child's fiscal exit, linear schedule, no income effects



(b) Child's fiscal exit, linear schedule, income effects



(c) Child's fiscal exit, progressive schedule



(d) Child's fiscal exit, constant partner-income gradient

Figure C.1: Income effects

Notes: Stylised income growth against partner income in $t = -1$ for one cell, before and after the event. Above the vertical dashed line the event moves the household into a higher bracket, while below it the marginal rate is unchanged. Panels (b) and (c) add an income effect that grows with the distance above the threshold (dashed curve), under a linear and a progressive schedule. Panel (d) replaces the flat growth of Panels (a) to (c) with income growth that falls linearly with partner income, a constant gradient that the event itself does not change. The dashed curve of Panel (d) is the before-event curve shifted by the event's direct effect alone, and the arrows split the fall above the threshold into that effect and the effect of the tax increase, which includes an income effect.

D Realised rate changes in the example cell

The predicted-control group's realised marginal rate is a direct check of the mapping of Section 3.3. Between $t = -1$ and $t = 0$ this group's household marginal rate rises by 1.9 percentage points on average for women and 1.7 for men, with a median change of zero, against 10.5 and 10.0 points for the predicted treated, whose predicted change is 16 points. In the first stage, the realised change per unit of the predicted change in the log net-of-tax rate is 0.53 (s.e. 0.01) for women and 0.52 (0.01) for men at $t = 0$, and declines to 0.38 and 0.40 by $t = 3$ as incomes move across brackets.

E Further results

E.1 The event samples

Table E.1: The two event samples

	Child's fiscal exit		Marriage	
	Women	Men	Women	Men
Panel A: construction (individuals)				
Event observed (2004–2023)	1,932,140	1,959,441	2,295,222	2,297,699
Marriage: predicted rate increase	–	–	575,517	203,268
Civil status unchanged; marriage: design screens[a]	1,914,489	1,936,085	309,067	121,554
Own work income at $t = -1$ below €70,000	1,884,452	1,758,070	308,483	121,154
Positive work income at $t = -1$	1,497,580	1,490,879	269,731	97,744
Panel B: characteristics of the event sample at $t = -1$				
Annual event cohort size (mean)	95,724	96,804	89,249[b]	79,209[b]
Age at $t = 0$ (mean)	51.09	53.28	33.61[c]	35.49[c]
Partner's age at $t = 0$ (mean)	53.39	50.94	35.65[c]	33.52[c]
Married at $t = 0$ (share)	0.98	0.98	0.53[c]	0.55[c]
Dependent children at $t = -1$ (mean)	1.92	1.90	0.27[c]	0.35[c]
Own work income (€, mean)	18,126	27,021	19,160	22,500
Partner's income (€, mean)	42,481	21,632	27,239	22,363
Positive work income (share)	0.79	0.85	0.93	0.93
Predicted into a rate change (share)	0.30	0.32	0.25	0.11
Mean |predicted ΔMTR| if treated	0.12	0.12	0.12[c]	0.12[c]

Notes: Panel A counts individuals at each step of the sample construction described in the text. The fourth row is the event sample used throughout and the fifth is the sub-sample on which the earnings outcome is defined. The marriage columns follow the spouses for whom the wedding predicts a rate increase. The design compares them with the predicted controls of the same cells (941,102 women and 1,004,916 men below the income cap). The marriage estimation sample includes both groups and holds 1,249,585 women and 1,126,070 men. Panel B describes the event samples below the income cap in the pre-event year ($t = -1$), unless otherwise indicated. The marriage columns describe the estimation sample. Partner's income sums work, self-employment and pension income. Counts respect the data provider's disclosure rules. [a] Both spouses childless and separately taxed in the year before the wedding and filing jointly the year after. [b] Mean over the 2005 to 2017 cohorts; later cohorts are much smaller. [c] All first marriages below the income cap (2,282,283 women and 2,250,589 men), the marriage design's sample before its screens.

Cleaning of the raw panel drops households in the overseas departments and French residents abroad, returns without any individual identifier and individuals aged 75 and over, and resolves the individual-year duplicates that multiple declarations create. Table E.1 describes the two event samples, and the figure notes of Section 4 refer to it. Panel A traces their construction. The child's fiscal exit sample comprises individuals in married or civil-union tax units (i) whose first-born leaves the tax unit between 2004 and 2023 (1,932,140 women and 1,959,441 men); (ii) whose civil status is unchanged at the event, so that the fall in the quotient is the only change to the household's schedule (the restriction removes 0.9 percent of the women and

1.2 percent of the men); and (iii) whose own work income at $t = -1$ lies below €70,000 in 2015 prices (the cap removes a further 1.6 percent of the women and 9.2 percent of the men). The marriage sample comprises first marriages of spouses who were childless and separately taxed in the year before the wedding and file jointly the year after. Within cells, the design compares the spouses for whom the wedding predicts a rate increase (Section 3.2) with the predicted controls.

Each sample is used for two outcomes. The earnings outcome is log work-income growth and requires positive work income at $t = -1$ and in the outcome year. The participation outcome is an indicator for positive work income, and its sample also includes individuals without work income at $t = -1$. Two years before the event, 33 percent of the women and 22 percent of the men in the €0–10,000 comparisons are in work, against 97 percent or more in every bin from €10,000 up. A cell–dose comparison enters estimation if (i) at least 30 predicted-treated individuals carry the dose within the cell at $t = -1$ and (ii) the dose's regression sample (predicted treated and predicted controls together) holds at least 300 earnings outcomes at $t = -3$. The same floors apply to the participation regressions of the dose.

Panel B characterises the samples in the pre-event year. Parents at the first-born's fiscal exit are on average 51 to 53 years old and almost all married. The exiting first-born is aged 19 to 26 in 86 percent of households. Women's partners earn on average €42,481 and men's €21,632, against own work income of €18,126 and €27,021. Between 79 and 85 percent have positive work income at $t = -1$. Of the women 30 percent and of the men 32 percent are predicted into a rate change, with a mean absolute predicted change in the marginal rate of 12 percentage points among them. The marriage sample is younger (34 to 36 years old at the event) and has fewer dependent children and higher participation. A quarter of the women and a tenth of the men are predicted a rate increase at the wedding.

E.2 Balance and the pooled validation-region check

Table E.2 summarises the balance of predetermined characteristics in the pooled validation of Section 4.1. It gives means for the predicted-control and predicted-treated groups and the standardised mean difference averaged across comparisons. The two groups differ more in partner income than in own income. This follows from the construction of the cells, which fix

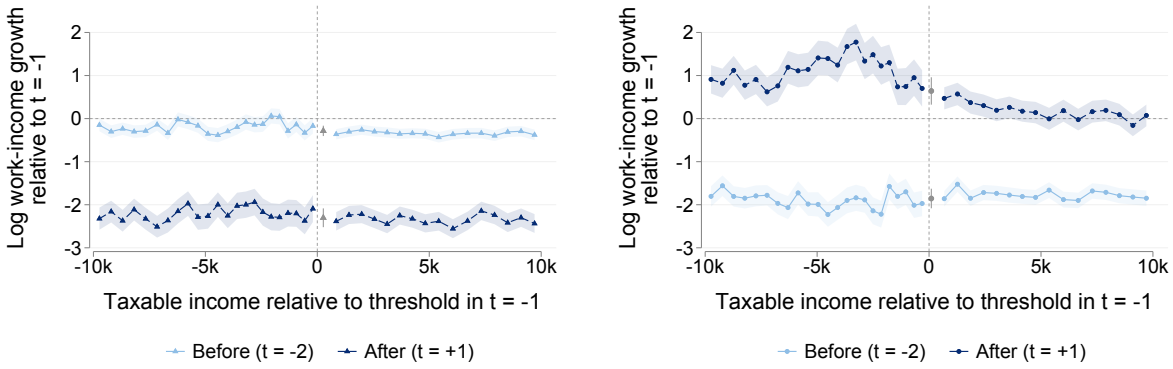
partner income only within a tercile and own income within a narrower interval, and within a cell it is therefore largely partner income that decides who is predicted treated. Own income follows similar paths in the two groups before the event. The differences that remain between the two groups, of 2 to 6 percent in own income, are absorbed by fixed effects for own income, age, marital status and the number of children (Section 3.5). Narrower intervals would reduce these differences but leave less power to include fixed effects. Nearest-neighbour matching, which we report as a robustness check, reduces them further and yields noisier estimates.

Table E.2: Balance across predicted-treated and predicted-control comparisons

	Women			Men		
	Control	Treated	SMD	Control	Treated	SMD
Own income						
Work income, $t = -3$ (€)	16,107	17,032	0.19	26,569	27,387	0.13
Work income, $t = -2$ (€)	16,367	17,131	0.19	26,851	27,595	0.14
Work income, $t = -1$ (€)	16,558	17,167	0.23	27,116	27,710	0.23
Positive work income, $t = -1$ (%)	79.4	81.0	0.03	89.4	89.0	−0.01
Partner income						
Partner work and self-employment income, $t = -1$ (€)	27,167	30,182	0.30	16,594	18,347	0.34
Partner has positive work income, $t = -1$ (%)	84.6	87.2	0.04	77.3	82.3	0.10
Composition						
Age	49.5	50.6	0.25	51.8	52.8	0.22
Partner age	52.1	52.9	0.17	49.4	50.6	0.28
Number of dependent children	2.24	1.97	−0.47	2.29	1.95	−0.56
Fiscal parts, $t = -1$	3.37	3.15	−0.45	3.41	3.13	−0.55
Assignment and first stage (differ by design)						
Household taxable income, $t = -1$ (€)	39,584	45,112	1.06	39,596	44,416	1.04
Predicted change in log net-of-tax rate (×100)	0.00	−11.42	–	0.00	−11.30	–
Realised net-of-tax-rate change, $t = +1$ (pp)	−1.36	−5.56	–	−1.49	−5.59	–
Sample						
Unique individuals	842,950	534,851		836,555	512,991	
Comparison-year cells	964			1,215		

Notes: Each dose is compared with the predicted controls in the same own-income bin, partner-income tercile, pre-event bracket and event year. Means are computed within each comparison-year arm and averaged across comparison-year cells with the cell's predicted-treated count as weight. SMD is the mean of the signed within-cell standardised differences with the same weights. Household taxable income determines assignment and differs between arms by design (standardised difference 1.06 for women, 1.04 for men). The predicted rate change is zero in the control arm by construction, but controls whose income drifts across a threshold realise a change. No standardised difference is reported for either row. The rows for $t = -3$ and $t = -2$ rest on the comparisons observed in those years. Controls may enter several dose comparisons; the counts report unique individuals. Levels are in nominal euros of the event year. Sample: the child's fiscal exit sample of Appendix E.1, in the cells of the design.

Figure E.1 pools the validation-region check of Figure 4 across the cells (Section 4.1) and measures household income from each household's hypothetical post-event threshold. Before the event, work-income growth shows no systematic gradient in household income across the validation region or above the threshold, for either gender.



(a) Men

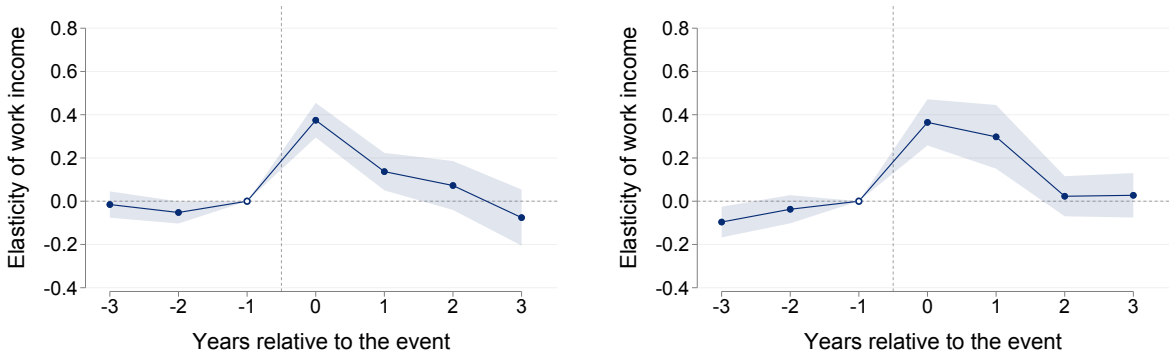
(b) Women

Figure E.1: The validation-region check, pooled across cells

Notes: Log work-income growth relative to $t = -1$ (100 times the change in log work income) at $t = -2$ (before the event) and $t = +1$ (after), against household taxable income at $t = -1$ minus the household's statutory threshold at the post-event quotient, pooled across the cells. Points are means of 40 quantile bins. Growth is computed for individuals with work income above €1,000 in both years and winsorised at the 5th and 95th percentiles of the gender sample at each horizon. Bands are 95% confidence intervals clustered by person. A pre-event profile that is flat across the threshold indicates that household income carries no income-growth gradient of its own before the event (Jakobsen & Søgaard, 2022). Sample: the child's fiscal exit sample of Appendix E.1, in the cells.

E.3 Event-time profiles by gender

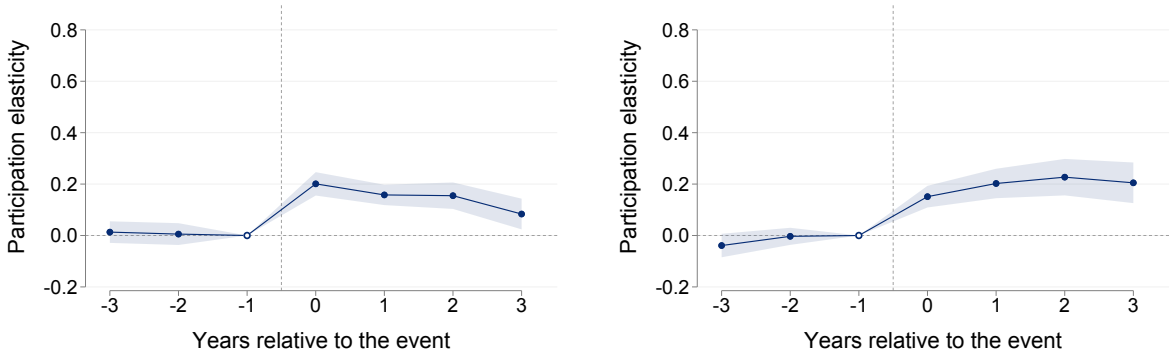
Figures E.2 and E.3 report the estimates by year relative to the event separately for men and women, with the pre-event years as placebo checks (Section 4.2.3 reports the two genders pooled).



(a) Men (b) Women

Figure E.2: Event-time earnings elasticities around the child's fiscal exit

Notes: Sample: the child's fiscal exit sample of Appendix E.1 (all event years). The outcome is work-income growth relative to $t = -1$. Each point pools the cell–dose elasticities of one year relative to the event across all own-income bins and weights each by its regression sample. Shaded areas are 95% confidence intervals from person-clustered standard errors. The year before the event, $t = -1$, is the reference and is shown at zero with a hollow marker (a reference point, not a failed placebo); estimates for $t = -3$ and $t = -2$ are placebo checks. They are divided by the pre-event contrast in realised rates, which is smaller than in the event year. On the event-year scale, the women's placebo at $t = -3$ is -0.03 and the men's at $t = -2$ is -0.02. The placebo reduced form is pooled over all pre-event cells with the weights of the estimates and tested with the person-clustered covariance. The t-statistics are 0.53 for men and -1.96 for women. The participation counterpart is Figure E.3.



(a) Men (b) Women

Figure E.3: Event-time participation elasticities around the child's fiscal exit

Notes: As Figure E.2, for participation: the outcome is an indicator for positive work income and the estimates are participation elasticities. The pooled pre-event placebo reduced form does not reject for either gender (t-statistics -1.08 for women and 1.32 for men).

E.4 Aggregated participation elasticities

Table E.3 repeats the participation aggregates of Table 4 (Panel A) and aggregates them over the own-income bins from €10,000 up (Panel B). In that range the French-taxpayer aggregate

is 0.16 (0.18 for women, 0.15 for men). It is an exit elasticity of established workers, whose participation two years before the event is 0.99.

Table E.3: Aggregated participation elasticities

A. All own-income bins

	Men	Women	Total
Population-weighted: French taxpayers	0.212 (0.039)	0.283 (0.029)	0.247 (0.024)
Population-weighted: event population	0.205 (0.033)	0.230 (0.033)	0.218 (0.023)
Inverse-variance weighted	0.127 (0.012)	0.096 (0.011)	0.111 (0.008)
Income-weighted	0.145 (0.016)	0.177 (0.022)	0.158 (0.013)

B. Own work income at $t = -1$ of €10,000 or more

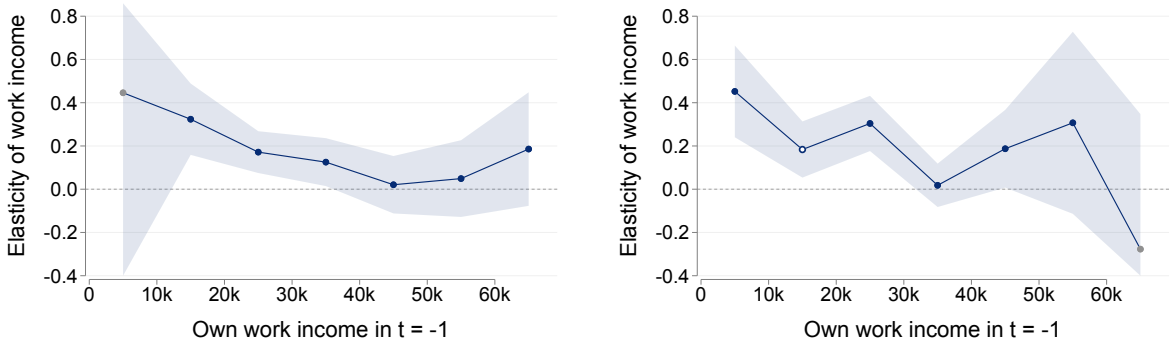
	Men	Women	Total
Population-weighted: French taxpayers	0.147 (0.017)	0.178 (0.024)	0.162 (0.015)
Population-weighted: event population	0.147 (0.018)	0.174 (0.027)	0.158 (0.015)
Inverse-variance weighted	0.124 (0.012)	0.086 (0.011)	0.105 (0.008)
Income-weighted	0.135 (0.015)	0.159 (0.023)	0.145 (0.013)

Notes: As Table 4, for participation: the outcome is an indicator for positive work income, and the entries are participation elasticities with respect to the household net-of-marginal-tax rate. The cell coefficients are divided by participation two years before the event (Section 3.5). Panel B repeats the aggregation over the own-income bins from €10,000 up.

E.5 Profiles along own income by gender

This subsection reports the gender breakdowns of the pooled earnings and participation profiles of Section 4.2.2 (Figures 6 and 7), discussed in Section 4.4.

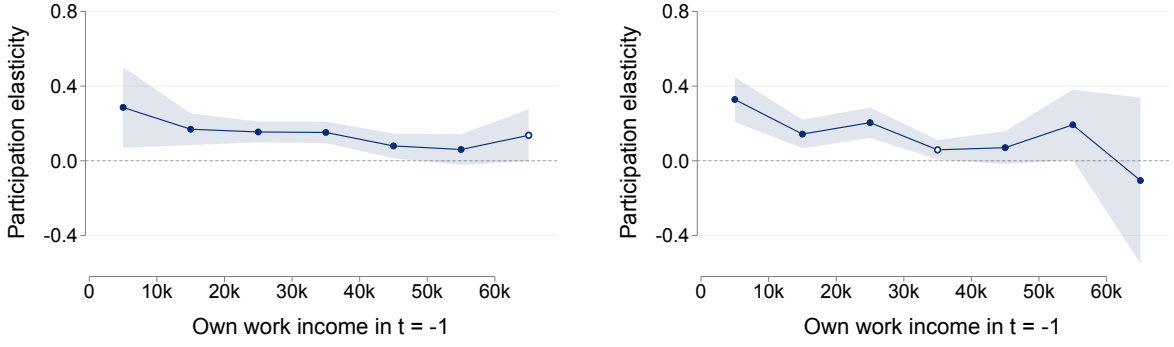
For participation, the outcome is an indicator for positive work income. The bin estimates are participation elasticities (the change in the participation probability per unit change in the log household net-of-marginal-tax rate, divided by participation two years before the event). They are constructed as in Section 3, with the samples and pooling window ($t = 0$ to 3) of the earnings profiles.



(a) Men (b) Women

Figure E.4: Earnings elasticities along the own-income distribution, by gender

Notes: As Figure 6, with the cell–dose–period estimates aggregated within each gender. A hollow marker flags the one bin that fails its pooled pre-event placebo (absolute t-statistic of at least 1.96), the women's €10,000–20,000 bin. Bins resting on a single estimate are drawn in grey with their interval cut at the axis: the men's €0–10,000 bin, 0.45 with interval −0.50 to 1.39, and the women's €60,000–70,000 bin, −0.28 with interval −0.90 to 0.35. Neither enters an aggregate.



(a) Men (b) Women

Figure E.5: Participation elasticities along the own-income distribution, by gender

Notes: As Figure 7, with the estimates aggregated within each gender and displayed as in Figure E.4. Hollow markers flag the two bins that fail their pre-event placebo, the women's €30,000–40,000 bin and the men's €60,000–70,000 bin (t-statistic 1.96). The women's €60,000–70,000 bin rests on three cell–dose–period estimates and 885 individuals and enters no aggregate. No bin rests on a single estimate.

E.6 Maps over own and partner income by gender

Figure E.6 repeats the maps of Figure 9 by gender. The women's maps carry both patterns of Section 4.2.4, with bottom-bin earnings elasticities of 0.63, 0.29 and −0.02 from the low to the top partner-income tercile and participation elasticities of 0.02, 0.69 and 0.35, while the men's earnings map has no clear partner-income gradient within own-income bins (Section 4.4;

its €0–10,000 column is not read) and the men's participation map has sizeable responses in most tiles.



(a) Men, earnings elasticity

(b) Women, earnings elasticity

(c) Men, participation elasticity

(d) Women, participation elasticity

Figure E.6: Earnings and participation elasticities by own and partner income, by gender

Notes: As Figure 9, with the cell estimates aggregated within each gender: earnings in Panels (a) and (b), participation in Panels (c) and (d). Above the value of a dashed tile is the number of underlying individuals. A dashed tile with a bracketed value is too imprecise or rests on too few individuals to be read, n.a. means that none of the rectangle's cells yields a usable estimate, and a blank rectangle has no cell. Grey fill marks an estimate below −0.075. Standard errors are clustered by person.

Table E.4 reports the partner-income boundaries of the terciles behind the maps of Figures 9 and E.6.

Table E.4: Partner-income tercile boundaries by own-income bin

	Women		Men	
Own income	Low to middle	Middle to high	Low to middle	Middle to high
0–10k, earner sub-sample	24.7	38.3	13.8	26.2
0–10k, entrant sub-sample	27.9	51.7	11.6	26.4
10–20k	23.1	34.1	13.6	22.0
20–30k	25.3	36.9	15.4	23.6
30–40k	29.0	42.3	16.5	27.1
40–50k	32.8	50.8	16.6	28.9
50–60k	36.3	60.3	15.8	29.9

Notes: Boundaries between the within-bin partner-income terciles of Figure 9, in thousand euro of partner income at $t = -1$; the lowest tercile includes negative values and the top tercile is open-ended. In the €0–10,000 bin the terciles are formed separately by pre-event earner status.

E.7 Broad earnings

Broad earnings add self-employment income to work income and are reported in Figure E.7.

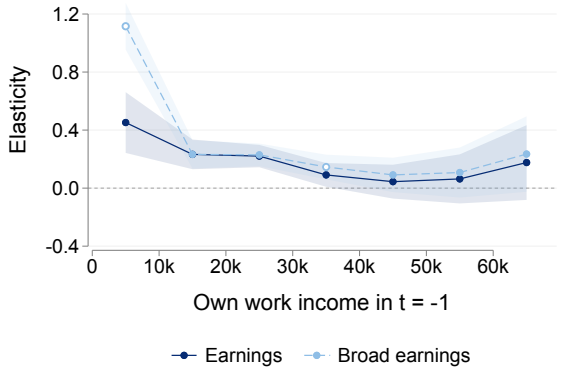


Figure E.7: Broad earnings elasticities along the own-income distribution

Notes: As Figure 6, with broad earnings (work plus self-employment income) as the outcome, on cells defined by own work income. Self-employment income sums the declared industrial, commercial, non-commercial and agricultural profit boxes (turnover before the abatement under the micro regimes, net profit under the real regimes); turnover declared under the micro-entrepreneur flat-rate scheme (*versement libératoire*) is not included. The figure pools men and women as in Figure 6 and overlays that figure's earnings profile (dark navy, solid) with the broad earnings profile (light blue, dashed). The broad earnings profile is within 0.01 of the earnings profile between €10,000 and €30,000 and 0.04 to 0.06 above it from €30,000 up. Hollow markers flag bins of the broad earnings profile that fail their pre-event placebo (the bottom bin and the €30,000–40,000 bin). The broad earnings outcome has no participation counterpart.

E.8 First-stage strength

Because each cell's first stage is just identified, a small predicted-treated group can leave it weak, and a cell enters an exhibit only if its heteroskedasticity-robust first-stage F is at least

9 (with one instrument, the effective F of Montiel Olea and Pflueger (2013)). The weak cells this screen still admits carry almost no weight. Cells with F below 23.1, the Montiel Olea and Pflueger (2013) critical value for a worst-case bias of 10 percent, hold 0.05 percent of the earnings weight and 1.4 percent of the participation weight (Table E.5). Weak-instrument-robust inference agrees with the Wald intervals we report. Anderson–Rubin sets (Anderson & Rubin, 1949) are bounded in every displayed cell and reject zero wherever the Wald intervals do, and the tF correction of Lee et al. (2022), applied to each cell, raises the bin standard errors by at most 1 percent for earnings and 18 percent for participation, and no income-profile bin loses its significance at 5 percent.

Table E.5 reports the robust first-stage F for all post-event cell estimates and for the cells behind each exhibit. Table E.6 shows the income profiles under the stricter requirement $F \geq 104.7$, which moves every bin by less than one standard error except the women's participation bins at €10,000–20,000 and €20,000–30,000.

Table E.5: First-stage strength of the cell estimates behind each exhibit

	Women						Men					
Cells	Number	Median F	Ratio	$F < 16.4$	$F < 23.1$	$F < 104.7$	Number	Median F	Ratio	$F < 16.4$	$F < 23.1$	$F < 104.7$
Panel A: earnings												
All estimates	835	275	0.89	14.7	16.2	33.8	948	353	0.82	15.2	17.9	29.3
Passing the elasticity and s.e. screens	174	1,402	0.87	0.0	0.1	3.3	203	1,016	0.79	0.0	0.0	1.0
Displayed: income and event-time profiles	174	1,402	0.87	0.0	0.1	3.3	203	1,016	0.79	0.0	0.0	1.0
of which in the maps	142	1,825	0.88	0.0	0.1	0.8	154	1,197	0.79	0.0	0.0	0.4
of which in the aggregation table	173	1,402	0.87	0.0	0.1	3.3	202	1,016	0.79	0.0	0.0	1.0
Panel B: participation												
All estimates	939	362	0.86	12.0	13.5	28.8	1,124	311	0.83	13.8	15.9	29.6
Passing the elasticity and s.e. screens	362	694	0.86	0.7	1.4	13.9	422	706	0.79	0.3	1.8	10.9
Displayed: income and event-time profiles	356	697	0.86	0.3	1.0	13.6	419	706	0.79	0.3	1.7	10.9
of which in the maps	307	728	0.86	0.1	0.5	11.7	342	759	0.78	0.2	1.3	8.1
of which in the aggregation table	353	697	0.86	0.3	1.0	13.6	419	706	0.79	0.3	1.7	10.9

Notes: First-stage strength of the cell–dose–period two-stage least squares estimates behind the exhibits of Section 4, $t = 0$ to 3. Each regression has one instrument and one endogenous regressor; F is the heteroskedasticity-robust first-stage F, i.e. the squared robust first-stage t-ratio. With one instrument it equals the effective F of Montiel Olea and Pflueger (2013). Ratio is the median of the robust over the classical F. Median F and Ratio weight each cell by its regression sample. The last three columns give the share of that weight, in percent, in cells below the Stock and Yogo (2005) critical value for a maximal two-stage least squares size of 10 percent (16.38), below the Montiel Olea and Pflueger (2013) critical value for a worst-case bias of 10 percent (23.1), and below 104.7, the value above which the two-stage least squares t-ratio has correct 5 percent size (Lee et al., 2022). Displayed cells pass the robust first-stage screen $F \geq 9$ (an absolute t-ratio of at least 3); the two 'of which' rows count the subsets that enter a displayed tile and the aggregation table.

Table E.6: Income profiles under a stricter first-stage requirement

	Displayed ($F \geq 9$)		$F \geq 104.7$			
Own income	Estimate (s.e.)	Cells	Estimate	Cells	Weight kept	Difference
Panel A: earnings						
Women						
0–10k	0.45 (0.11)	18	0.45	18	100.0	0.00
10–20k	0.18 (0.07)	58	0.18	58	100.0	0.00
20–30k	0.30 (0.07)	57	0.26	48	92.4	-0.04
30–40k	0.02 (0.05)	28	0.02	20	94.8	0.00
40–50k	0.19 (0.09)	9	0.19	8	99.1	0.01
50–60k	0.31 (0.21)	3	0.31	3	100.0	0.00
60–70k	-0.28 (0.32)	1	-0.28	1	100.0	0.00
Men						
0–10k	0.45 (0.48)	1	0.45	1	100.0	0.00
10–20k	0.32 (0.08)	41	0.32	41	100.0	0.00
20–30k	0.17 (0.05)	74	0.18	68	99.3	0.00
30–40k	0.13 (0.06)	44	0.13	36	97.2	0.01
40–50k	0.02 (0.07)	22	0.02	20	99.6	0.00
50–60k	0.05 (0.09)	13	0.05	13	100.0	0.00
60–70k	0.19 (0.13)	8	0.20	7	98.1	0.01
Panel B: participation						
Women						
0–10k	0.33 (0.06)	67	0.33	64	98.3	0.00
10–20k	0.14 (0.04)	102	0.05	72	86.7	-0.09
20–30k	0.20 (0.04)	102	0.13	53	77.1	-0.08
30–40k	0.06 (0.03)	60	0.05	28	85.6	0.00
40–50k	0.07 (0.04)	17	0.08	9	83.7	0.01
50–60k	0.19 (0.10)	5	0.20	4	96.5	0.01
60–70k	-0.11 (0.23)	3	-0.02	1	33.8	0.09
Men						
0–10k	0.29 (0.11)	27	0.29	27	100.0	0.00
10–20k	0.17 (0.04)	87	0.20	59	82.3	0.03
20–30k	0.15 (0.03)	122	0.17	83	91.2	0.01
30–40k	0.15 (0.03)	91	0.17	55	88.6	0.02
40–50k	0.08 (0.03)	49	0.08	26	90.0	0.00
50–60k	0.06 (0.04)	25	0.06	16	89.9	0.00
60–70k	0.14 (0.07)	18	0.08	9	77.6	-0.05

Notes: Own-income bins of the income profiles (Figures E.4 and E.5), frequency weights, $t = 0$ to 3: the displayed estimate with its person-clustered standard error, and the same construction with only the first-stage requirement raised from $F \geq 9$ to $F \geq 104.7$ (Lee et al., 2022). Weight kept is the share of the bin's frequency weight retained, in percent. Difference is the change in the bin estimate. The stricter construction is reported as a point estimate only. Over bins with at least two displayed cells, the largest change is 0.041 for earnings and 0.090 for participation. The two bins that move by more than one displayed standard error are those of women's participation at €10,000–20,000 and €20,000–30,000 (by −0.090 and −0.078).

E.9 Responses below age 58

Table E.7 reports the bins from €10,000 up and their aggregates for individuals below 58 in the year before the event (Section 4.4).

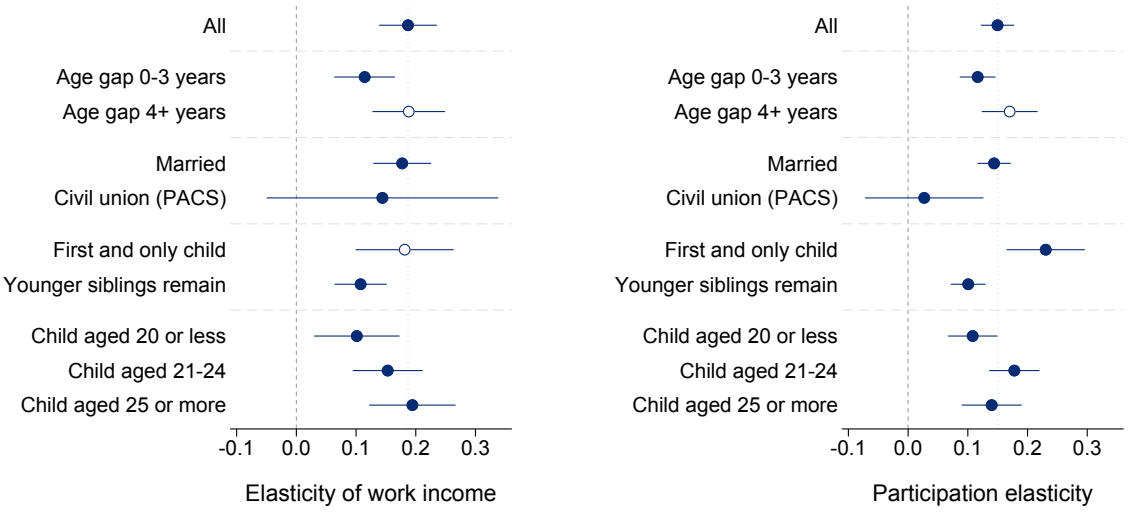
Table E.7: Responses from €10,000 up, all ages and below 58

	Women		Men	
	Headline	Below 58	Headline	Below 58
Own income				
A. Earnings elasticity				
10–20k	0.18†	0.08†	0.32	0.09
	(0.07)	(0.05)	(0.08)	(0.07)
20–30k	0.30	0.18	0.17	0.09
	(0.07)	(0.05)	(0.05)	(0.04)
30–40k	0.02	−0.07	0.13	0.05
	(0.05)	(0.05)	(0.06)	(0.05)
40–50k	0.19	0.13	0.02	−0.11
	(0.09)	(0.08)	(0.07)	(0.06)
50–60k	0.31	0.20	0.05	0.04
	(0.21)	(0.22)	(0.09)	(0.08)
60–70k	−0.28[a]	−0.01[a]	0.19	0.16
	(0.32)	(0.34)	(0.13)	(0.11)
Aggregate[b]	0.23	0.10	0.17	0.06
	(0.05)	(0.03)	(0.03)	(0.03)
Removed (%)[c]		3.5		7.7
B. Participation elasticity				
10–20k	0.14	0.12	0.17	0.06
	(0.04)	(0.04)	(0.04)	(0.04)
20–30k	0.20	0.18	0.15	0.12
	(0.04)	(0.04)	(0.03)	(0.02)
30–40k	0.06†	0.03	0.15	0.11†
	(0.03)	(0.03)	(0.03)	(0.03)
40–50k	0.07	0.06	0.08	0.02
	(0.04)	(0.04)	(0.03)	(0.03)
50–60k	0.19	0.13	0.06	−0.01
	(0.10)	(0.09)	(0.04)	(0.04)
60–70k	−0.11	−0.02	0.14†	0.14
	(0.23)	(0.24)	(0.07)	(0.07)
Aggregate[b]	0.18	0.15	0.15	0.11
	(0.02)	(0.02)	(0.02)	(0.02)
Removed (%)[c]		3.7		8.3

Notes: The *Headline* columns repeat the profiles of Figure E.4 (Panel A) and Figure E.5 (Panel B), and the *Below 58* columns re-estimate them for individuals below 58 in the year before the event, with the same cells, pooling window ($t = 0$ to 3) and weights. The €0–10,000 bin is omitted. With men and women pooled, the aggregates are 0.20 (0.03) in the headline and 0.08 (0.02) below 58 for earnings, and 0.16 (0.02) and 0.13 (0.01) for participation. Standard errors in parentheses are clustered by person. † The bin's pooled pre-event placebo fails (absolute t-statistic of at least 1.96). [a] A single cell–dose–period estimate. [b] Bins entering Table 4 from €10,000 up, with its French-taxpayer weights. [c] Share of the panel's estimation sample from €10,000 up that the restriction removes, within the dose groups estimated both with and without it.

E.10 Household subgroups

Figure E.8 reports the elasticities by the age gap between the parents, by marriage versus civil union, by whether the exiting first-born is the only child and by the age of the exiting child, for own incomes from €10,000 up. The confidence intervals overlap in most comparisons.



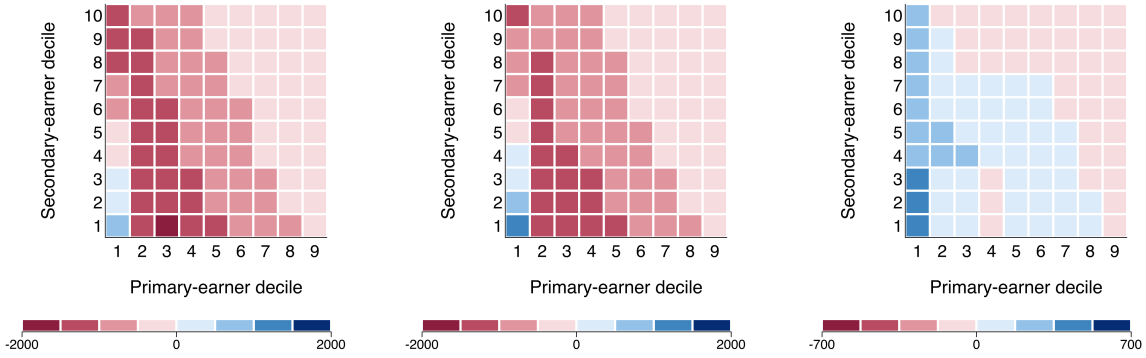
(a) Earnings elasticity (b) Participation elasticity

Figure E.8: Elasticities by household subgroup

Notes: Each row pools the cells of one subgroup with own income from €10,000 up, men and women together, and weights each estimate by its regression sample (Section 3.4). Outcomes and estimates as in Figure 9. Whiskers are 95% confidence intervals from standard errors clustered by person within gender, with the two genders treated as independent. A hollow marker means that the row fails its pooled pre-event placebo. Subgroups are defined at $t = -1$, except the child's age, measured at the event; "First and only child" means that no other child is ever observed in the tax unit. The All row (0.19 for earnings, 0.15 for participation) uses the same weights; with French-taxpayer weights the aggregates from €10,000 up are 0.20 and 0.16.

E.11 Welfare effects under near-Rawlsian weights

Figure E.9 repeats the exercise of Figure 11 with near-Rawlsian welfare weights.



(a) Single elasticity (b) Estimated elasticities (c) Difference, (b) minus (a)

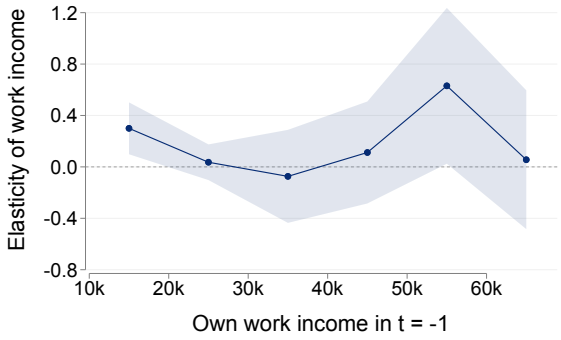
Figure E.9: Welfare effects of local reforms under near-Rawlsian weights

Notes: As Figure 11, using near-Rawlsian welfare weights concentrated on the lowest-income households. Panels (a) and (b) show the single-elasticity and estimated-response calibrations; Panel (c) reports estimated minus single-elasticity. Colour scales are in model welfare units under the €500-by-€500 local-cell and one-percentage-point reform normalisation. These units are not euro equivalents. In Panels (a) and (b), blue denotes gains and red denotes losses. In Panel (c), blue denotes a positive difference and red denotes a negative one.

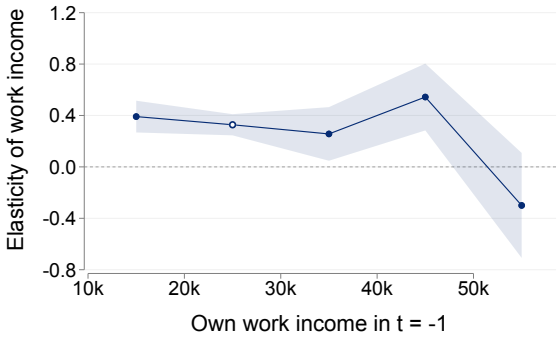
F The marriage design

The marriage design applies the strategy of Sections 3 and 4 to first marriages of spouses who were childless and separately taxed in the year before the wedding and file jointly the year after. The instrument applies the pre-event formula to the two spouses' incomes at $t = -1$ pooled into one unit with quotient 2. As at the child's fiscal exit, the predicted treated are the spouses whom the event moves to a higher marginal rate (25 percent of the women and 11 percent of the men with own work income below €70,000), compared within cells with the predicted controls. This share falls with own income, to under 13 and 6 percent above €30,000. Cells, estimation, the pooling over $t = 0$ to 3 and the weights are as at the child's fiscal exit. The one difference is that the outcomes are residualised on the partner's income, both spouses' ages and own work income at $t = -2$ and $t = -3$, fitted among the predicted controls of each cell. This allows for a constant partner-income gradient in income growth. Standard errors and placebo tests are classical and treat the cells as independent.

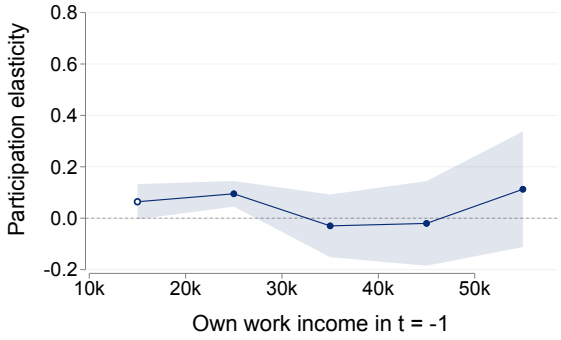
Figure F.1 shows the estimates in the paper's profile format. The profiles start at €10,000, because below it fewer than half of the spouses are in work two years before the wedding, and end at €60,000 (€70,000 for men's earnings). The women's earnings elasticities decline with own income, as at the child's fiscal exit, and the men's are not different from zero above €20,000. For participation, the women's elasticities are largest below €30,000 and the men's are close to zero throughout. The estimates are noisier than at the child's fiscal exit, with wide confidence intervals above €30,000, where few spouses are predicted a rate increase, and two of the 21 bins fail their pre-event placebo.



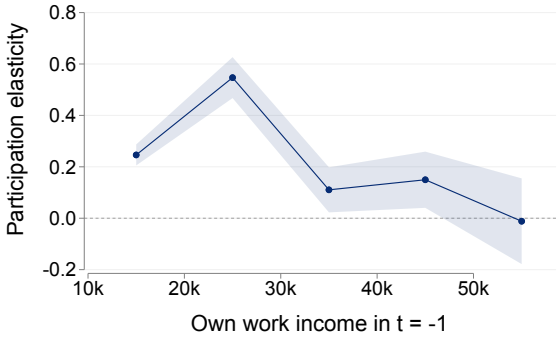
(a) Men, earnings elasticity



(b) Women, earnings elasticity



(c) Men, participation elasticity



(d) Women, participation elasticity

Figure F.1: Earnings and participation elasticities along the own-income distribution at marriage

Notes: Sample: the marriage sample of Table E.1, the spouses whom the wedding moves to a higher marginal rate, compared within cells with the predicted controls. Outcome: work-income growth relative to $t = -1$ in Panels (a) and (b) and an indicator for positive work income in Panels (c) and (d), residualised as described in the text and pooled over $t = 0$ to 3 with frequency weights; the participation estimates are divided by pre-event participation. Each marker pools the cell–dose elasticities of an own-income bin, and the band is the 95 percent confidence interval. Display as in Figure E.4, with own income from €10,000 to €60,000, or to €70,000 for men's earnings. Hollow markers fail their pre-event placebo. Standard errors and placebo tests are classical and treat the cells as independent.

G Monotonicity

This appendix (i) states the monotonicity condition in the standard reform-based simulated-IV setting (Gruber & Saez, 2002; Jakobsen & Søgaard, 2022), (ii) characterises when the condition fails and illustrates a violation in a calibrated example, and (iii) argues that violations are less likely under schedule *shifts* such as the one in our design than under reforms that change the schedule's gradient.

G.1 The condition in the reform-based setting

Consider a reform that raises statutory rates on part of the schedule. Let $g \in \{0, 1\}$ indicate whether the reform is in force, and let $\text{MTR}_i(g)$ denote the marginal tax rate individual i faces under regime g, inclusive of behavioural responses. The simulated instrument is the mechanical prediction of $\text{MTR}_i(1) - \text{MTR}_i(0)$, holding income fixed at its pre-reform value. The IV estimand averages individual responses with weights proportional to the realised first-stage responses $\pi_i \equiv \text{MTR}_i(1) - \text{MTR}_i(0)$ (Angrist & Imbens, 1995; Borusyak et al., 2022). The π_i must have the same sign for the weights to be non-negative and for the estimand to be a convex average that cannot reverse the sign of a uniformly signed effect (Blandhol et al., 2026). First-stage monotonicity requires

$$\text{MTR}_i(1) \;\geq\; \text{MTR}_i(0) \quad \forall\, i, \tag{G.1}$$

with the reverse inequality for a reform that cuts rates. A violation puts negative weight on the responses of individuals whose realised marginal tax rate moves *against* the direction of the reform.

G.2 When the condition fails

Condition (G.1) holds under three strong assumptions standard in the literature: (i) a quasi-linear utility model without income effects, (ii) continuous labour supply, and (iii) no heterogeneity in elasticities. The intuition is that the utility ranking between the kink point and interior solutions in brackets unaffected by the reform remains unchanged, whether the reform increases or decreases the marginal tax rate in a given bracket. Any new interior optimum in a bracket unaffected by the reform is then dominated by the kink point at the reform-affected bracket.

When either of the first two assumptions is relaxed (allowing income effects or hours rigidities) and elasticities or preferences are heterogeneous, first-stage monotonicity can be violated.

We illustrate such a violation using a model with income effects, hours rigidities, and heterogeneity in the curvature of the utility function.[44] Individuals derive utility from consumption c and hours worked h according to

$$u_i(c,h) = \frac{c^{1-\sigma_i}}{1-\sigma_i} - \chi_i \frac{h^{1+\varphi_i}}{1+\varphi_i},$$

where $\sigma_i > 0$ governs the curvature of utility from consumption and the strength of income effects, $\varphi_i > 0$ governs the curvature of the disutility from labour, and $\chi_i > 0$ scales the marginal disutility of hours worked. Labour income is given by $y = wh$. We assume that hours rigidities restrict individuals to four discrete choices: 0, 20, 30, or 40 hours per week.

We use the 2015 French income tax schedule. In the part of the schedule on which we focus, labour income up to €26,791 per year is taxed at a marginal rate of 14%, and income above this threshold at 30%.[45] Because utility is defined over hours worked, we can equivalently express the tax schedule as a function of hours, with a kink at $h = k$ corresponding to the income threshold. We set the hourly wage to €20.6, so that the threshold at €26,791 corresponds to approximately 25 hours of work per week.

We study a hypothetical reform ($g = 1$) that eliminates the 30% bracket by raising the marginal tax rate above the threshold from 30% to 41% (the next statutory rate in the French schedule).

We consider two preference types, Type A and Type B, and calibrate the model such that both share the same discrete optimum of 30 hours per week under no reform. Type A is calibrated with $\sigma_A < 1$, so that the substitution effect dominates, and Type B with $\sigma_B \gg 1$, so that the income effect dominates. For each type, the scale χ_i is set to deliver this common optimum. The resulting preference parameters are $(\sigma_A, \varphi_A) = (0.696,\ 0.300)$ and $(\sigma_B, \varphi_B) = (7.000,\ 0.300)$, with $\chi_A = 0.272$ and χ_B set to deliver the common optimum.

Under no reform, both types choose 30 hours per week and are in the 30% bracket:

$$\mathrm{MTR}_A(0) = \mathrm{MTR}_B(0) = 0.30.$$

[44] Hours rigidities would suffice to illustrate a violation. We add income effects to make the argument more general.

[45] To simplify the exposition, the example uses the statutory schedule, without the tax reduction for small liabilities of Section 2.1; the reduction does not affect the example.

The simulated instrument holds both types' incomes fixed at their 30-hour values and predicts that both face a rise in the marginal tax rate from 30% to 41%. The predicted first-stage response is +0.11.

The optimal hours choices under the reform and under no reform are shown in Figure G.1. Under the reform, the two types' optimal choices diverge:

$$h_A^* = 20\,\text{hours/week}, \qquad h_B^* = 40\,\text{hours/week}.$$

The realised marginal tax rates respond to the same reform in opposite directions:

$$\pi_A = \text{MTR}_A(1) - \text{MTR}_A(0) = 0.14 - 0.30 = -0.16 < 0,$$
$$\pi_B = \text{MTR}_B(1) - \text{MTR}_B(0) = 0.41 - 0.30 = +0.11 > 0.$$

Type A's realised marginal tax rate moves against the direction of the reform. It *falls* although the reform raises statutory rates and the instrument predicts a rate increase for Type A. This violates condition (G.1). Type B's realised marginal tax rate moves with the reform.

The estimand places negative weight on Type A's response. It is not weakly causal in the sense of Blandhol et al. (2026), because with heterogeneous elasticities it can even reverse the sign of a uniformly signed response.

A strong income effect for Type B is not necessary for the violation. If Type B stays at 30 hours under the reform ($\pi_B = +0.11$) while Type A moves to 20 ($\pi_A = -0.16$), the same violation arises.

Models with participation responses can also generate violations of first-stage monotonicity. For example, an individual who leaves work in response to a reform that increases tax rates moves out of the tax schedule altogether (in France, to a marginal tax rate of zero), so that $\pi_i < 0$. If other individuals j remain employed and face the higher rate under the reform ($\pi_j > 0$), realised marginal tax rates move in opposite directions across individuals. This violates condition (G.1). Participation responses to tax increases are common (this paper documents them too), and they are a source of monotonicity violations in reform-based IV strategies.

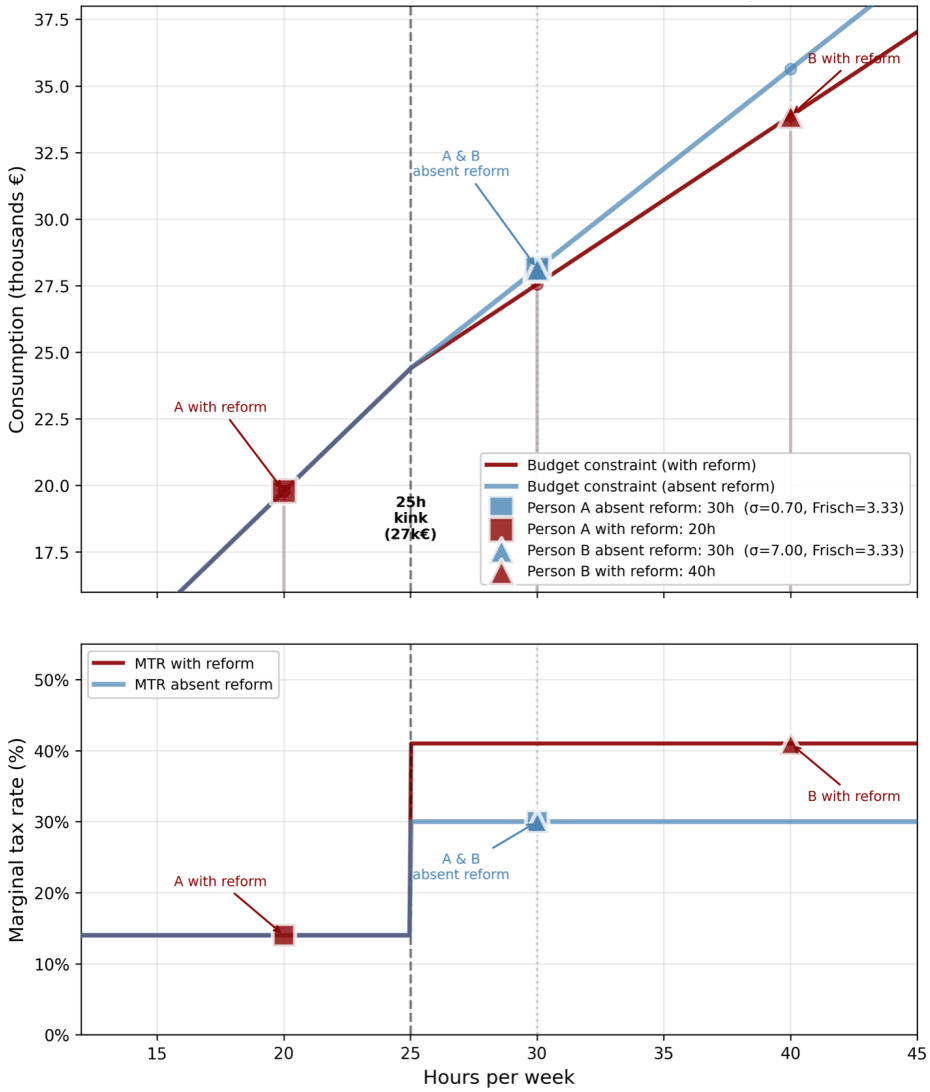


Figure G.1: Violation of first-stage monotonicity under heterogeneous labour-supply responses

Notes: Top panel. Budget constraints under the reform (red) and under no reform (blue). Both types choose 30 h/week under no reform. Under the reform, Type A chooses 20 h/week and is below the kink; Type B ($\sigma_B = 7$) chooses 40 h/week and is above the kink.
Bottom panel. Marginal tax rates under the reform (red) and under no reform (blue), as a function of hours worked, with the same optimal choices indicated.

G.3 The condition in the event-based setting

In the main design the schedule change is the fiscal exit of the first-born, which lowers the quotient from 2.5 to 2 with one dependent child and from 3 to 2.5 with two (Section 3.2). Because household income is divided by the quotient, the marginal rate at every level of household income y satisfies $T_s'(y/Q') \geq T_s'(y/Q)$ for $Q' < Q$: the exit weakly raises the marginal

rate at every point of the budget set. Only the rate-increase direction of condition (G.1) applies, and a violation cannot arise from the shift itself. It requires a household to reduce its income by enough that its realised marginal rate falls strictly below the rate it would have faced under the old quotient. The required fall in income is much larger than under a reform. An individual predicted into the next bracket has income divided by the new quotient at or above the upper threshold of the pre-event bracket, and a violation requires it to fall below the lower threshold. Household income must fall below the ratio of the two thresholds times its pre-event level. Partner income makes even this channel unlikely. Where partner income alone places the household above the tax-free threshold under the post-event quotient, the individual's exit from work leaves the household in a positive bracket (a condition that can be checked cell by cell, since the cells fix the partner-income range).[46] The same floor arises at marriage, where the partner's income is added to the tax base.

With a partner earning €40,000, for example, an individual faces a marginal rate of 14% even at zero own income under the 2019 statutory schedule, because the partner's income alone places the household above the tax-free threshold. For an individual predicted to move from the 14% to the 30% bracket, a violation would require the realised marginal rate to fall strictly below 14%, which the partner's income prevents unless it collapses as well. The same holds for every predicted rate increase. Violations are also unlikely for predicted decreases, because they would require a large increase in labour supply.

[46] Two residual channels remain: (i) a simultaneous collapse of both spouses' incomes below the household tax-free threshold, and (ii) the partner's own exit from work, which lowers household income through the same mechanism.