Marginal Personal Income Tax Changes: Tax Revenue, Welfare, and Labour Supply Responses*

Rory Green[a,†] and David Sonnewald[a,‡]

[a]UCLouvain

November 23, 2024

Abstract

We use a massive synthetic data set representative of the universe of taxpayers in Belgium to assess workers' behaviour in response to four reform scenarios that entail marginal changes in the Personal Income Tax (PIT). We employ a novel tool for fiscal policy simulation, the Belgian arithmetic microsimulation model (Beamm), to derive individuals' disposable income after the PIT and examine inequality and welfare indicators. A Random Utility Random Opportunity (RURO) model is estimated to calculate labour supply adjustments to modifications in the tax structure. We find that all the reforms considered are effectively inequality neutral and yield welfare changes of small magnitude, with only two reforms generating a welfare improvement. From a behavioural perspective, workers seem to be more sensitive to adjustments of marginal tax rates, as opposed to the restructuring of income tax brackets.

Keywords: Microsimulation, RURO Models, Tax Reforms.

JEL Classification: C63, H31, J22.

***Acknowledgements:** We are very grateful to our supervisor Koen Declercq (UCLouvain Saint-Louis Bruxelles) for his constant guidance and support. We are thankful for the precious comments and suggestions received from Michael Smart (University of Toronto), participants at the Doctoral Meeting (UCLouvain Saint-Louis Bruxelles), the Micro Breakfast Seminar (McGill University), the Department of Applied Economics internal seminar (HEC Montréal), the 9th World Congress of the International Microsimulation Association (University of Vienna), the IAAEU Workshop on Labour Economics 2024 (Trier University), the IDEAS 2024 Conference (Institut d'études européennes, ULB), the 2024 RCEA International Conference in Economics, Econometrics, and Finance (Brunel University London), the 19th CIREQ PhD Students Conference (Concordia University), the Spring Doctoral Workshop in Economics 2024 (UCLouvain), the 58th Annual Conference of the Canadian Economics Association (Toronto Metropolitan University), the COMPIE Conference (Vrije Universiteit Amsterdam), and the 23rd LAGV Conference (Aix-Marseille). The contribution of Antoine de Mahieu (UCLouvain Saint-Louis Bruxelles) to our RURO `R` code is also acknowledged.

†Email: rory.green@uclouvain.be. Personal website: https://roryhgreen.netlify.app/.

‡Email: david.sonnewald@uclouvain.be. Personal website: https://davidsonnewald.com.

1 Introduction

Taxation, income redistribution, and labour supply are closely interconnected. On the one hand, taxation is the most direct mean with which a State is empowered to level off income distribution. On the other, the effects of tax reforms on labour supply are among the primary interests of policy makers. In fact, changes in tax policies influence individuals' decisions to work, the number of hours they work, and the overall supply of labour in the economy (Keane (2010)). Finding the perfect trade-off among tax revenue, redistribution, and labour supply is thus the objective of every State's fiscal policy.

In the present study, we aim to tackle this challenge in the context of Belgium, the country with the highest tax wedge among OECD countries ("OECD" (2023)). Along the same lines of Creedy et al. (2018)'s exercise based in New Zealand, we look at marginal changes in the Personal Income Tax (PIT) scheme, *i.e.,* very small changes in the tax rates and in the taxable income brackets' thresholds. Hence, we answer the question of what the effects of reforms implementing these changes[1] are on tax revenue, welfare, and labour supply.

We address the optimal direction of tax changes, rather than optimal taxation, using a social welfare function defined in terms of disposable income and inequality. We obtain changes in tax revenue, disposable income, and inequality indicators from the Belgian arithmetic microsimulation model (Beamm), which runs on a completely novel synthetic data set representative of the universe of taxpayers in the country.[2] In this study, we go beyond the standard static version of Beamm by integrating it with a structural labour supply model to account for how individuals adjust their labour supply in response to our reforms of the PIT. Specifically, we use a Random Utility Random Opportunity (RURO) model, expanding on de Mahieu (2021) and providing new evidence on its application to labour supply estimation.

Our analysis shows that all the proposed tax reforms lead to welfare changes of similar (small) magnitude without significantly affecting income inequality, as evidenced by minimal variations in the Gini index. However, only two reforms generate a welfare improvement, and none of them represents an optimal directional tax change. From a behavioural perspective, individuals change their worked hours remarkably only when all marginal tax rates are increased or decreased. In contrast, this change is negligible in the case of income tax bracket variations. Therefore, our findings suggest that, depending on the State's objective, either adjusting the brackets or modifying the rates could be effective strategies for tax policy.

[1] Our package of reforms is thoroughly presented in Section 2.1.

[2] Tax-benefit microsimulation models (as Beamm) are considered to be static when they compute the so-called "day after effect" of a reform, *i.e.,* the straight effect on a certain output without the chance for the individuals to adjust their behaviour accordingly, as if the reform were implemented overnight.

By looking at the effects of fiscal adjustments on inequality, welfare and labour supply, combining a microsimulation model with a structural behavioural model, this article contributes to four main strands of literature. First, it speaks to that line of research that investigates the use of taxation as redistributive tool. This literature grew exponentially in the recent past in response to the increasing wealth and income concentration at the top-end of the income distribution, which many countries have been experiencing since the beginning of the 21st century. Saez and Diamond (2012), for example, claim that high earners should be subject to high and rising marginal tax rates on earnings, and that capital income should be taxed. Piketty et al. (2011) and Saez and Zucman (2020) estimate that the optimal tax rate for the top 1% of US earners' total wealth and incomes is between 75% and 80%. Stephenson (2018) exploits the differences in the type of income taxation in five countries of the European Union (Belgium, Bulgaria, Germany, Lithuania and Poland) to measure the impact of progressivity of taxation on inequality. In fact, he claims that the income tax structure with a differentiated rate, as it is in Germany and Belgium, seems to be the most redistributive. Dianov et al. (2022) explore a similar research question widening the group under analysis to all 27 European Union countries and the United Kingdom. Yet, they find a milder effect of progressive taxation in decreasing inequality. While they do not question the redistributive role of taxation, their findings suggest that the problem of increasing income inequality is more complex and of a multifaceted nature. The present study aims to complement this body of literature using a detailed microsimulation model to examine the effects of increasing (or decreasing) marginal progressivity of the personal income tax to address redistribution and inequality.

Second, this paper is related to that strand of research that uses structural models to estimate the effects of changes in labour incomes due to tax policies on labour supply. Müllbacher and Nagl (2017) use a structural discrete choice framework to explore the fiscal effects of the Austrian tax reform of 2016. They find a total increase in working hours by 0.71%, with supply effects that are stronger at the intensive margin, for females and for low-income earners. Similarly, Bosch et al. (2017) estimate labour supply responses for a large number of subgroups after a major tax reform in the Netherlands. In particular, they find strong differences in labour supply responses between households with and without children. As for Belgium, Decoster et al. (2010) investigate the effects of the introduction of a flat tax, based on a microsimulation model that includes a labour supply model. Hence, they find that there are positive effects both on labour supply and on the tax base. With a similar methodology, de Mahieu (2021) assesses the effects of the extension of Belgian in-work benefits on labour supply and welfare. He finds that further increasing the benefits would slightly increase labour supply and welfare of low-to-middle income deciles, but at very high net cost per job created.

The primary objective and main contribution of this paper is to bridge the gap between studies on the effects of taxation on redistribution and inequality and those examining the impact of taxation on labour supply. In fact, by combining the results of the static microsimulation model with the behavioural response model,[3] impacts on both inequality and labour supply are investigated. Moreover, the employment of Beamm as methodology represents itself a novelty in the current research. Capéau et al. (2018) also develop an encompassing model (including budget and welfare evaluations as well as labour supply decisions)[4] to study a tax reform in Belgium. However, their microsimulation outputs are based on EUROMOD.[5] This model is not tailored to any specific country and it runs on Statistics on Income and Living Conditions (SILC) surveys, which contain much less observations and information compared to the data used in Beamm. In fact, this is the first model that is able to replicate the Belgian fiscal system to such an extent and deliver highly granular and precise results both at the individual and aggregate levels.

From a methodological standpoint, we make two additional contributions. First, we speak to that line of research on optimal tax models. On the one hand, by adopting a structural approach, we align with studies rooted in the seminal paper of Mirrlees (1971). On the other, foregoing overall social welfare maximization, we partly complement the literature employing reduced form models (*e.g.*, Kleven et al. (2009), Piketty and Saez (2013a), Piketty et al. (2014), Saez and Stantcheva (2016)). In fact, while building upon Creedy et al. (2018) and Bierbrauer et al. (2023), we develop our own welfare metric to evaluate the optimal direction of tax changes. Second, this study adds on previous uses of RURO models. For both early and more recent derivations we refer to Aaberge et al. (1995), Aaberge et al. (1999), Dagsvik and Strøm (1995), Dagsvik and Strøm (2006), Capéau et al. (2016), Aaberge and Colombino (2018), and Capéau et al. (2018).

The remainder of the article is structured as follows. Section 2 illustrates the institutional setting of the reforms simulated within this study's framework and the welfare metric that we adopt to evaluate the optimal direction of tax changes. In Section 3, we present the synthetic data used for the empirical analysis. Section 4 provides a brief explanation about Beamm. In Section 5, we elaborate the RURO model and its integration with the outcomes from the microsimulation model. Section 6 evaluates the optimal direction of the reforms under analysis.

[3] We refer to Aaberge and Colombino (2018) for a detailed review about structural labour supply models and microsimulation.

[4] As we do in the present paper, they also model labour supply using RURO.

[5] EUROMOD is a tax-benefit microsimulation model for the 27 countries of the European Union.

2 Marginal PIT Reforms

2.1 Institutional Context

The progressive PIT structure in Belgium has changed a number of times across the last decade, seeing a shift in both the income thresholds and tax rates. The most notable change was the removal of the 30% tax bracket and a broadening of the tax thresholds for low income earners, agreed to in the 2016 Budgetary Agreement by the Belgian Federal Government.[6] Despite these changes to the tax system, for more than two decades Belgium has continued to have the highest tax wedge among OECD countries, which the European Commission (EC (2022)) notes may be a large driver in a lack of long term participation in labour when average income earners are subject to 45% and 50% tax rates. Furthermore, according to the OECD (2022) Economic Surveys on Belgium, Belgian tax revenue contributes to around 43% of GDP, which is strikingly higher than the OECD average of 33.5% and it is currently ranked as the third highest in the world. As such, Belgium presents itself as an interesting case study in the multidisciplinary relationship between taxation, labour supply, and welfare, as the country continuously battles to maintain progressiveness in its tax policies while meeting budgetary requirements.

Table 1: PIT scheme in Belgium (2020)

TI (€) from	**TI (€) to**	**Rate (%)**	**Max tax on bracket (€)**	**Cumulative tax (€)**
0	13,250	25	3,312.5	3,312.5
13,250	23,390	40	4,056	7,368.5
23,390	40,480	45	7,698.5	15,067
40,480	and above	50		

Notes. The reference year is 2020 (income year 2019), which is the year for which the data were collected. Tax brackets and rates are applicable to net taxable income after the deduction of social security charges and professional expenses.

The paper at hand simulates PIT reforms with the goal of analysing the behavioural responses of Belgium taxpayers in addition to the equality and redistributive qualities that those reforms entail. While calculating an optimal tax structure à la Mirrlees (1971) is naturally a policy goal, the complexity of the tax-benefit scheme in behavioural microsimulation models makes this highly impractical. Yet, deriving the optimal direction of tax changes through marginal reforms to the fiscal system can still bring to light a number of beneficial recommendations for policy makers (Creedy et al. (2018)). For this reason, we simulate marginal changes across the Taxable Income (TI) scheme of Belgium (Table 1) in order to better guide the orientation of future reforms. As Creedy et al. (2018), we examine marginal changes in tax

[6] See OECD (2023) for a detailed look at all the changes in the reform.

rates and thresholds' brackets. In the interest of clarity, we define four reforms, whose changes carried out with respect to the base PIT scheme are summarised in Table 2 (we refer to them in the remainders of the paper as numbered in the table). Hence, Reform 1 (2) increases (decreases) all marginal tax rates by 1 percentage point (pp). Reform 3 (4) increases (decreases) all income brackets by 1000 €.

Table 2: Summary of simulated reforms

	Base		**Reform 1**		**Reform 2**		**Reform 3**		**Reform 4**	
	Income	**Rate**	**Income** (/)	**Rate** (+1pp)	**Income** (/)	**Rate** (-1pp)	**Income** (+1000)	**Rate** (/)	**Income** (-1000)	**Rate** (/)
First bracket	0-13,250	25%	/	26%	/	24%	0-14,250	/	0-12,250	/
Second bracket	13,251-23,390	40%	/	41%	/	39%	14,251-24,390	/	12,251-22,390	/
Third bracket	23,391-40,480	45%	/	46%	/	44%	24,391-41,480	/	22,391-39,480	/
Fourth bracket	40,481-∞	50%	/	51%	/	49%	41,481-∞	/	39,481-∞	/

Notes. Below each reform, the changes that they make to the 2020 (income year 2019) PIT scheme are specified. Changes apply to all brackets and thresholds. *E.g.*, Rate (+1pp) means that all rates are increased by 1pp. "/" means no change from the Base case. Income is expressed in Euros (€).

2.2 Evaluating optimal direction of tax changes

Optimality of taxation depends on the distributional value judgments adopted. Although we undertake a structural approach by considering individuals' behavioural responses,[7] we do not evaluate the properties of an optimal tax structure. Therefore, we do not define a welfare function that the State must maximise. Yet, to address the optimal direction of tax changes, we need to specify a form of welfare metric.

Creedy et al. (2018) use money metric utility and a social welfare function based on constant relative inequality aversion. When they marginally change the tax parameters, they look at the obtained values of welfare and revenues changes, and they indicate the direction of an optimal reform by relative orders of magnitude of these ratios (*e.g.*, a reform that is both revenue neutral and welfare improving).

We adopt a similar specification to maintain inequality aversion and to express the equity-efficiency trade-off. However, we use average disposable income instead of money metric utility.[8]

[7] Two main categories of optimal tax models exist: structural models and reduced form models. In the structural approach, which dates back to Mirrlees (1971), the government generally maximizes an explicit welfare function under a government budget constraint. Individuals, in this framework, seek to maximize utility functions defined in terms of net income and leisure. On the other hand, the reduced form approach relies on a minimal set of parameters, placing particular emphasis on the elasticity of taxable income. The optimality condition, in this approach, is usually expressed in relation to the marginal benefits and costs of a tax change (see: Kleven et al. (2009), Piketty and Saez (2013a), Piketty et al. (2014), Saez and Stantcheva (2016)).

[8] Creedy et al. (2018)'s money metric utility is constructed based on the concept of "full income", which is defined as the net income which could be obtained if all endowment of time were devoted to work at the going wage rate. We consider disposable income to be a less restrictive measure.

Therefore, our social welfare function is given by:

$$\mathcal{W} = f(d, g) = \bar{d}(1 - g) \tag{2.2.1}$$

where $\bar{d}$ is the average disposable income and g is the Gini Index, which measures inequality.

For the sake of studying the optimal direction of tax changes, we compare welfare gains (or losses) with the impact on net tax revenue (difference between collected taxes and provided benefits) and labour supply. We follow a comparable approach to Bierbrauer et al. (2023), who evaluate the Pareto-improving direction of tax reforms relative to a tax revenue function.[9] Specifically, we define optimal direction a reform's outcome that improves welfare ($\Delta\mathcal{W} > 0$) without exacerbating net tax revenue and labour supply. The following definition clarifies it formally.

Definition 1. Given $\mathcal{W}_0$ and $\mathcal{W}_r$ the welfare values before and after a tax reform, respectively, let NTR be the net tax revenue and l^s the labour supply. If $\Delta\mathcal{W} = \mathcal{W}_r - \mathcal{W}_0 > 0$, then the reform is an optimal direction of a tax change if the two following conditions are satisfied simultaneously:

- $\Delta NTR = NTR_r - NTR_0 \geq 0$;
- $\Delta l^s = l^s_r - l^s_0 \geq 0$.

By nature of the simulation at hand, we are able to compute the state revenue lost or gained and the resulting disposable income from each potential reform, as well as their effects on inequality and labour supply. Therefore, in light of Equation 2.2.1 and Definition 1, we can quantify welfare changes and identify reforms that entail the optimal direction of tax changes.

To corroborate the robustness of our results, we also evaluate welfare changes in Definition 1 by means of a stochastic dominance criterion (instead of the social welfare function in Equation 2.2.1). We relegate the description of such methodology, as well as its application, to Appendix A.

[9] One strong assumption of Bierbrauer et al. (2023), however, is that the tax revenue is "redistributed lump sum after those whose tax burden goes up (if any) have been compensated".

3 Data

In this study, we use a synthetic data set that is created *ad hoc* for Beamm. This is an absolute novel database, as no existing representative Belgian data cover all the information that Beamm needs.

As different parts of the required information are contained in several data sets, we use statistical matching to merge these various data sets into one single synthetic data set. This technique was pioneered by Anderson (1995), but we refer to Rässler (2002), D'Orazio et al. (2006) and Annoye et al. (2024) for more recent and detailed reviews.

Concretely, we combine three data bases (see Table 3). The main source is the data from the Personal Income Tax declarations (IPCAL) of 2020[10]. These administrative data encompass the universe of taxpayers in Belgium, for a total of almost 7 million observations. The second data set is DEMOBEL (census data), which includes mainly socio-demographic variables from 2022. Finally, to account for labour market characteristics (most notably, number of hours worked) we include also the European Union Statistics on Income and Living Conditions (EU-SILC) data from 2019.

Table 3: Data bases for the statistical matching

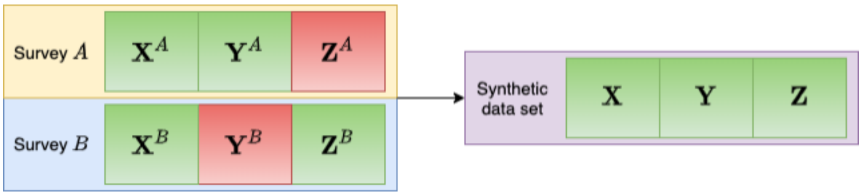
	IPCAL	**DEMOBEL**	**SILC**
Type of data	Administrative	Census	Survey
Type of variables	Income/Revenues	Socio-demographic	Labour market and living conditions
Number of variables	1700	31	98 individual + 64 household
Number of observations	6,927,006	119,059	16,105 individuals in 7,035 households
Individual Identifier	Yes	Yes	Yes
Household Identifier	Yes (fiscal)	Yes (social)	Yes (social)
Year	2020	2022	2019

Notes. Fiscal households file the Personal Income Tax declaration jointly. Social households consist of people living together, regardless of taxation.

Our final synthetic data set is made up of the same observations from the baseline source (IPCAL), with only the missing information (socio-demographic and labour characteristics) being complemented using the other available data bases. It is worth noting that this process does not attempt to reconstruct the missing information at the level of an individual citizen or household. All the information in our data sets is anonymised, so that we cannot match exactly individuals or households across data sources. Hence, statistical matching connects all the information in a way that the final data set is accurate at the level of the entire distribution, *i.e.*, at the level of the entire population (see Figure 1).

[10] 2020 is the fiscal year that refers to incomes of 2019.

Figure 1: Schematic representation statistical matching



Notes. Neural networks are trained on the available information (Y^A and Z^B) to the common variables X to fill in the gaps (marked in red). *Source.* D'Orazio et al. (2006).

We assess the quality of statistical matching by comparing the distributions of our synthetic data with those of our data sources. These robustness checks are relegated to Appendix B. Descriptive statistics for our synthetic data are summarised in Table 4. For RURO's parameters estimation, we take a random sample of 2,500 active individuals aged between 18 and 65.[11]

Table 4: Breakdown by individuals characteristics and labour status

	Employed	**Self-Employed**	**Unemployed**
Gender			
Men	50.04%	50.30%	50.06%
Women	49.96%	49.70%	49.94%
Age			
18-24	12.92%	12.05%	12.99%
25-34	29.53%	29.96%	28.73%
35-44	19.83%	20.49%	19.70%
45-54	14.82%	15.03%	14.78%
55-65	22.90%	22.47%	23.81%
Marital Status			
Single	75.51%	71.95%	78.47%
Couple	24.49%	28.05%	21.35%
Education level			
Low	30.08%	16.84%	61.44%
Middle	36.79%	24.13%	27.92%
High	33.13%	59.02%	10.63%
Country of birth			
Belgium	77.97%	82.36%	69.90%
EU27	7.03%	9.31%	8.95%
Other	14.99%	8.33%	21.15%
Labour market			
Hours worked weekly (mean)	33.53	33.94	
Hourly gross wage (mean)	39.46	39.59	
Gross yearly labour income (mean)	63690.08	64798.32	
Share	71.04%	18.89%	10.07%
Total Observations	2,465,321	655,443	349,530

Source. Synthetic data generated through statistical matching.

Notes. The table is structured vertically (shares sum up to 100% in every category, *e.g.,* gender). "Education Level": Low = up to primary education; Middle = secondary education; High = university degree or higher.

[11] Due to the high demand for computing power, RURO models are typically estimated using samples or subsamples of the actual population. The self-employed are also excluded from the sample for RURO's parameters estimation.

4 Belgian Arithmetic Microsimulation Model

The Belgian arithmetic microsimulation model (Beamm) is a microsimulation model for the tax-benefit system in Belgium.[12] For a detailed explanation of the motivation for such project and a divulgative illustration of Beamm's scope, possible uses and potentialities, we refer to Truyts et al. (2023). In this section, we review briefly the model's structure and how we employ it in the present study.

Microsimulation means that the model employs micro data, which are described in Section 3. As for the rules of the tax-benefit system, these were translated into `R` code such that each covered tax and benefit is calculated for every individual or household. Altogether, this constitutes a comprehensive simulator of the Belgian fiscal system, with its current parameters.[13] Precisely, as of writing, Beamm calculates: child benefits, income support, investment income tax, maternity leave, real property tax, personal income tax, vat and excise duties, and wealth tax. In addition, a number of aggregate outputs are delivered. These include: household disposable income, state budget, tax burden, tax wedge, and inequality, poverty and redistribution indexes. On the contrary, although designed to be integrated at a later stage,[14] the following components are not included: car taxation, gift taxes, inheritance taxes, pensions, social security contributions, and unemployment benefits.

When Beamm is run with the current rules of the fiscal system, it depicts the state of the art of Belgian taxation, and its "returns" in terms of state budget, tax burden, redistribution indexes, etc. However, when we change some parameters according to a potential reform, Beamm allows to study the effects of this policy by comparing the new values in the aggregate outputs with the pre-reform scenario ones. In the literature, this is referred to as the "day after effect", *i.e.*, the straight effect on a certain output without the chance for the individuals to adjust their behaviour accordingly, as if the reform were implemented overnight. Tax-benefit microsimulation models that compute the day after effect of a reform, as Beamm, are considered to be static.[15]

In this study, we go beyond Beamm's static form integrating it with a labour supply model to account for how individuals adjust their labour supply in response to four reforms (see Section 2). Specifically, to model labour supply, we use a Random Utility Random Opportunity model, which is elaborated in Section 5. As this model is designed to derive the labour supply based

[12] For a thorough review of microsimulation modelling (both generic and applied to tax-benefit systems) see O'Donoghue (2014). Beamm was also simplified to make it available as an open-access version online.

[13] Technically, all `R` codes calculating taxes and benefits were put together in a unique function, which constitutes *de facto* the simulator itself. Several `R packages` were also built for its smoother functioning.

[14] At the moment, the available data miss this information.

[15] On the contrary, models that incorporate behavioural reactions are referred to as dynamic.

on workers' disposable income, Beamm is employed in the estimation of RURO's parameters, as well as in the simulation of our reform scenarios. In the latter case, the usual other outputs (tax wedge, inequality indexes, etc.) are also computed and compared to study the effects of these policies in terms of redistribution, inequality and poverty.

5 RURO's Theoretical Framework

The general setting is that of an agent who faces a choice, or a series of choices over time, among a set of options. Specifically, in our context, agents are workers and the choice that they have to make is a bundle of hours to work and the respective wage at which they are remunerated (*i.e.*, their labour supply). In our analysis, we exclude the inter-temporal case. Denote the outcome of the decision in any given situation as y, where y indicates the chosen combination of hours to work and hourly wage,[16] our goal is to understand the behavioural causal process that leads to the agent's choice. That is, the agent's choice is determined, or caused, collectively by a number of factors. Some of these factors (x) are observed and some (ϵ) are not. We assume that the factors determine the agent's choice through a function $y = h(x, \epsilon)$, which is called the behavioural process. This behavioural process is deterministic, *i.e.*, given x and ϵ, the choice of the agent is fully determined. However, since ϵ is not observed, the agent's choice is not deterministic and cannot be predicted exactly. Yet, if we consider that the unobserved factors are random with a certain density $f(\epsilon)$, we can predict the probability of a particular outcome y. In particular, the probability that the agent chooses a particular outcome from the set of all possible outcomes is simply the probability that the unobserved factors are such that the behavioural process provides that outcome: $P(y|x) = Prob(\epsilon \; s.t. \; h(x, \epsilon) = y)$.

To compute this probability, we need to model the two sides of the choice. On the one hand, we need to define a utility function that explains the agents' preferences, *i.e.*, the form of the behavioural process. That is, through the utility function, we can observe what outcome y is chosen by the agents within the set of choices. The design of the utility function's form includes also the assumption on the density function $f(\epsilon)$ of the unobserved factors. On the other hand, the set of choices (opportunities) among which each agent can choose her preferred option (that delivers a certain outcome y) must be defined. More precisely, in RURO, jobs' availability can also depend on the demand side of the labour market, besides individual characteristics and skills.[17] This constitutes a great advantage compared to standard discrete choice models, as

[16] The "nature" of y determines the type of the model. If y is assumed to be discrete, as in the present case, we talk about discrete choice models. This also distinguishes remarkably RURO models from standard discrete choice multinomial logit models for labour supply (*e.g.*, McFadden (1973), Van Soest (1995)), where only optimal working hours are chosen.

[17] The same applies to non-market alternatives.

we are able to account for demand-side restrictions and macroeconomic circumstances.

Our RURO model derivation builds on de Mahieu (2021) and Capéau et al. (2016). While we develop it thoroughly in Sections 5.1, 5.2 and 5.3, we refer to these two studies for further details.

5.1 Random Utility

The opportunities available to the agents are combinations of hours to work and their respective hourly wage, *i.e.*, jobs. Therefore, the utility function that governs the behavioural process of agent i for job j is $U_i(d_j, l_j, \epsilon_{ij})$, where d_j is the disposable income, l_j are the weekly hours of leisure (equal to time endowment minus the number of working hours required for the job opportunity), and ϵ_{ij} is a taste shifter associated to the job choice that captures the effect of the unobserved factors. We assume that couples act as a unique agent, with a joint utility function that is defined by the combined disposable income and leisure time of both individuals who make up the couple, *i.e.*, $U_i(d_j^m + d_k^f, l_j^m, l_k^f, \epsilon_{ijk})$, where m and f represent male and female individuals within the couple, respectively.[18] Given that the behavioural process results from the interaction of both a deterministic and a random component, the utility function of individual (couple) i for job(s) j (and k) can be written simply as the sum of the two:

$$U_i(d_j, l_j, \epsilon_{ij}) = V_i(d_j, l_j) + \epsilon_{ij}$$

$$U_i(d_j^m + d_k^f, l_j^m, l_k^f, \epsilon_{ijk}) = V_i(d_j^m + d_k^f, l_j^m, l_k^f) + \epsilon_{ijk}$$

where V_i is the deterministic part and ϵ_i is the random part. In particular, we assume the two parts to be defined as follows (we drop the indexes for individual (couple) i and job(s) j (and k) to lighten the notation).

- $V(d, l)$ has a Box-Cox structural specification:[19]
  $V(d, l) = \alpha_d(\frac{d^{\alpha_1}-1}{\alpha_1}) + \alpha_l(\frac{l^{\alpha_2}-1}{\alpha_2})$, for singles; and
  $V(d^m+d^f, l^m, l^f) = \alpha_d(\frac{(d^m+d^f)^{\alpha_3}-1}{\alpha_3}) + \alpha_l^m(\frac{(l^m)^{\alpha_4}-1}{\alpha_4}) + \alpha_l^f(\frac{(l^f)^{\alpha_5}-1}{\alpha_5}) + \alpha_l^{mf}(\frac{(l^m)^{\alpha_4}-1}{\alpha_4})(\frac{(l^f)^{\alpha_5}-1}{\alpha_5})$, for couples.

  Heterogeneity in the marginal rates of substitution between leisure and income is allowed by introducing linearly a number of individual-specific covariates (vector X) into the leisure parameters:

[18] We exclude homosexual couples due to data availability.

[19] The deterministic part of the utility function can have different structural forms. The choice of the Box-Cox is standard in the literature (*e.g.*, de Mahieu (2021)), as it guarantees positive marginal utilities.

$$\alpha_l = \alpha_{l0} + \alpha_l' X$$
$$\alpha_l{}^m = \alpha_{l0}^m + \alpha_l{}^{m'} X^m$$
$$\alpha_l{}^f = \alpha_{l0}^f + \alpha_l{}^{f'} X^f$$

- ϵ is a random variable distributed as a Gumbel distribution with location parameter 0 and scale parameter 1, *i.e.*, $f(\epsilon) = e^{-\epsilon} e^{-e^{-\epsilon}}$.[20]

Regardless of how opportunities are created, which is elaborated in Section 5.2, we can predict the probability that an individual (or a couple) i chooses job(s) j (and k). For the sake of clarity in notation, in the remainder of this section, we show the calculation of this probability solely for single individuals. The process for couples is analogous, employing the corresponding utility function. Therefore, the utility maximizer agent i will prefer job j over job k whenever $U_{ij}(d_j, l_j, \epsilon_{ij}) > U_{ik}(d_k, l_k, \epsilon_{ik})$, $\forall j \neq k$. That is, whenever $V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik}$ (we abandon the terms in parenthesis and the repetition that this is valid $\forall j \neq k$ for the sake of clarity). Hence, the probability that agent i chooses job j, is:

$$\begin{aligned} P_{ij} = Prob(V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik}) \\ &= Prob(V_{ik} + \epsilon_{ik} < V_{ij} + \epsilon_{ij}) \\ &= \int_\epsilon I(\epsilon_{ik} - \epsilon_{ij} < V_{ij} - V_{ik}) f(\epsilon_i)\, d\epsilon_i \end{aligned} \tag{5.1.1}$$

where $I(\cdot)$ is the indicator function, equaling 1 when the term in parentheses is true and 0 otherwise, and V_{ij} and V_{ik} have a Box-Cox structural specification as described here above. This is a multidimensional integral over the density of the unobserved portion of utility $f(\epsilon_i)$. Given the assumption that we have made on the form of $f(\epsilon)$, the integral in Equation 5.1.1 can be simplified as follows (see Train (2003) for the full proof).[21]

$$P_{ij} = \int_\epsilon I(\epsilon_{ik} - \epsilon_{ij} < V_{ij} - V_{ik}) f(\epsilon_i)\, d\epsilon_i = \frac{e^{V_{ij}}}{\sum_k e^{V_{ik}}} \tag{5.1.2}$$

[20] This is another standard choice in the literature, mostly driven by its easier form when it comes to the probability's computation.

[21] "Different discrete choice models are obtained from different specifications of this density, that is, from different assumptions about the distribution of the unobserved portion of utility. The integral takes a closed-form only for certain specifications of $f(\cdot)$. Logit and nested logit have closed-form expressions for this integral. They are derived under the assumption that the unobserved portion of utility is distributed *iid* extreme value and a type of generalized extreme value, respectively" (Train (2003)). In the present study, the logit model is obtained providing $f(\epsilon)$ with a Gumbel distribution, which is a particular case of the generalized extreme value distribution (type I). Therefore, the integral has also a closed-form expression.

5.2 Random Opportunities

As clarified in Section 5.1, the choices available to the agents (the opportunities) are a set of jobs, where a job is defined by a number of hours to work and their respective hourly wage. First, we assume that wages are independent of hours worked. In particular, as in de Mahieu (2021), we assume that hourly wages are drawn from a log-normal distribution $g_1(w)$.

$$g_1(w) = \frac{1}{w\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{ln(w) - \gamma' Y}{\sigma}\right)^2\right) \tag{5.2.1}$$

where σ and the vector γ are the parameters of the distribution, and Y is a vector of covariates that might affect the median of the wage distribution. On the other hand, we assume that the average weekly working hours of job opportunities are distributed as uniform-with-peaks distribution. The peaks are chosen in correspondence to typical part-time and full-time regime, while the parameters allow to calibrate their height.

$$g_2(h) = \begin{cases} exp(\alpha^g_{h0}) & : h \in H \backslash \{[18.5, 20.5], [29.5, 30.5], [37.5, 40.5]\} \\ exp(\alpha^g_{h0} + \alpha_{h1}) & : h \in [29.5, 30.5] \\ exp(\alpha^g_{h0} + \alpha_{h2}) & : h \in [18.5, 20.5] \\ exp(\alpha^g_{h0} + \alpha_{h3}) & : h \in [37.5, 40.5] \end{cases} \tag{5.2.2}$$

where the domain H represents the possible values, and it is assumed to range from 0 to 70.

It is veritable to assume that an individual could also decide to stay out of the labour market ($h = 0$). That is, within the opportunities' set there is also a number of "out-of-market" job opportunities that can be available to the individuals. Specifically, we allow the intensity of job offers relative to out-of-market opportunities to vary across individuals according to a set of covariates Z_o. In terms of income, we assume that people opting for an out-of-market opportunity receive unemployment benefits.[22]

$$g_0 = exp(\alpha_o + \alpha_o' Z_o) \tag{5.2.3}$$

For an opportunity to enter in the utility function, it must be expressed in the form of disposable income, d, and weekly hours of leisure, l. For each job opportunity requiring an amount of hours h, the gross wage is computed as a multiplication of the number of hours h and the hourly wage w. The amount of taxes and benefits to which an individual is subject

[22] The selection of the out-of-market option is necessarily simplistic. Non-labour income is inherently individual-specific and cannot be accurately modeled with heterogeneity. Given that all individuals in our dataset are active in the labour market, those who are not employed are, by default, classified as unemployed. Therefore, we assume their non-labour income to be unemployment benefits, despite acknowledging that some individuals may not be eligible due to not meeting certain requirements.

and/or entitled to, based on the gross wage associated with a certain job, is calculated using Beamm. The disposable income is then obtained by simply subtracting and/or adding this amount to the gross income. This transformation is denoted $d_i(l, w)$.

5.3 Closing the model: Estimation of the MLL function

Be T the total time available to an agent in a working week, we introduce to simplify the notation, for any agent i, the function $\Psi_i(h, w) = exp(V_i(d_i(T-h, w), T-h)) = exp(V_i(d_i(l, w), l))$, where $d_i(l, w)$ is the disposable income as obtained through the transformation illustrated in Section 5.3 and $V_i(d, l)$ is the utility function's deterministic part of the agent, as defined in Section 5.1.

Given the probability P_{ij}, as in Equation 5.1.2, and the distributions of wages and working hours defining the opportunities' creation process (Equations 5.2.1, 5.2.2 and 5.2.3), we can write the likelihood that an individual i will choose a job offer j, which requires labour time $h = T - l$ and pays a wage w, as follows:

$$P_i(w, h) = \frac{\Psi_i(h, w) g_{0j} g_{1j}(w) g_{2j}(h)}{\Psi_i(0, 0) + \int_{r \in W} \int_{t \in H} \Psi_i(r, t) g_{0j} g_{1j}(r) g_{2j}(t) dr dt} \tag{5.3.1}$$

which, in case of an out-of-market opportunity, simplifies to:

$$P_i(0, 0) = \frac{\Psi_i(0, 0)}{\Psi_i(0, 0) + \int_{r \in W} \int_{t \in H} \Psi_i(r, t) g_{0j} g_{1j}(r) g_{2j}(t) dr dt} \tag{5.3.2}$$

However, we do not observe the actual job offers that an agent receives (neither the salary or the time component). On the contrary, for every individual, a set of job offers (D_i) is created from a prior density function, denoted $\mathbb{S}$, conditional on the observed choice being included.[23] To account for this, we condition the likelihood that an individual i chooses a job offer j on this same job offer to be in the drawn set of opportunities. That is:

$$P_i(w, h | D_i) = \frac{\Psi_i(h, w) g_{0j} g_{1j}(w) g_{2j}(h) / \mathbb{S}(w, h)}{\sum_{r,t \in D_i} \Psi_i(r, t) g_{0j} g_{1j}(r) g_{2j}(t) / \mathbb{S}(r, t)} \tag{5.3.3}$$

which, in case of an out-of-market opportunity, becomes:

$$P_i(0, 0 | D_i) = \frac{\Psi_i(0, 0) / \mathbb{S}(0, 0)}{\Psi_i(0, 0) / \mathbb{S}(0, 0) + \sum_{r,t \in D_i} \Psi_i(r, t) g_{0j} g_{1j}(r) g_{2j}(t) / \mathbb{S}(r, t)} \tag{5.3.4}$$

Finally, to compute the likelihood, L, that our observed sample is indeed observed we multiply

[23] This is an application of the Bayes law (see Capéau et al. (2016)). We use uniform distributions for the hours (from 0 to 70) and hourly wages (from 0 to 60). The prior probability to draw an out-of-market offer is set at 0.10.

the conditional agents' probabilities to choose a certain job over all N observations:

$$L = \prod_{i=1}^{N} P_i(w, h|D_i) \tag{5.3.5}$$

Our estimation consists into finding the parameters that maximize Equation 5.3.5. To achieve this, first, we take the logarithm of our likelihood function. Second, we use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization algorithm (Broyden (1970), Fletcher (1970), Goldfarb (1970), Shanno (1970)), which is a quasi-Newton method for solving unconstrained nonlinear optimization problems, to find the parameters that maximize the log-likelihood. The estimated parameters are those in the vectors α_l (from the utility function's deterministic part), γ (from Equation 5.2.1), α_h (from Equation 5.2.2), and α_o (from Equation 5.2.3). The covariates used for parameters' estimation, as well as the preliminary estimates, are illustrated in Appendix C.

6 Assessment of Reforms

Individuals' preferences are commonly designed such that workers' utility would increase with additional income and/or leisure time, all other things being equal. Given that our utility function includes also a random component (see Section 5.1), to determine if our model aligns with the general understanding of preferences for leisure and income, we verify that the marginal utilities related to these variables are positive. The proportion of individuals for whom these marginal utilities are positive, based on their observed choices, is presented in Table 5. In fact, our estimates suggest that for all individuals, utility rises with both disposable income and leisure time.

Table 5: Marginal Utilities

	% Observations		
Marginal Utilities	**Single Male**	**Single Female**	**Couples**
U'_d	100	100	100
U'_l	100	100	Male: 100 Female: 100

Notes. Numbers represent the percentage of people for whom marginal utilities are positive at the observed choices.

The effects of each reform (Table 2) on our main outcomes of interest (net tax revenue, labour supply, and welfare), as well as on some intermediate outputs (*e.g.*, the Gini Index), are expressed as variations with respect to the base (pre-reform) case. While all results are thoroughly summarised in Table 6, some general remarks can be drawn.

First, all reforms yield welfare changes of similar (small) magnitude.[24] However, only Reform 2 and Reform 3 generate a welfare improvement. This suggests that, should welfare gains be the only objective of the State, the current fiscal system could be improved either by a marginal decrease in all tax rates or by a marginal increase in all income brackets. That said, none of the reforms meets the criteria outlined in Definition 1 for an optimal tax change, which requires simultaneous improvements in both net tax revenue ($\Delta NTR \geq 0$) and labour supply ($\Delta l^s \geq 0$). In fact, positive changes in the total labour supply go together with decreases in the net tax revenue, corroborating the evidence that the tax burden acts as a hindrance to work.

Second, from a comparison of the two welfare-gaining reforms, Reform 3 emerges as the most effective. This reform, which focuses on changes to the tax brackets rather than marginal rates, achieves the largest welfare improvement with only a modest reduction in tax revenue (and a small increase in labour supply). In contrast, Reforms 2, which focuses on altering the marginal tax rates, leads to a more significant loss in tax revenue without a substantial increase in labour supply. Therefore, Reforms 3 appears to be more effective in achieving welfare improvements with fewer adverse effects on labour supply and tax revenue.

Third, all four reforms are effectively inequality neutral. The changes in the Gini Index are minimal, with variations of a maximum of 0.001 percentage points (pp), and the average disposable income ($\bar{d}$) soars or falls similarly regardless of the parameter changed. This result confirms that marginal reforms do not significantly alter the income distribution. On the contrary, they impact primarily the tax revenue.

Finally, from a behavioural perspective, workers seem to be more sensitive to adjustments of marginal tax rates, as opposed to the restructuring of income tax brackets. Individuals change their worked hours remarkably only when all marginal tax rates are increased (Reform 1) or, although to a lesser extent, decreased (Reform 2). On the other hand, this change is negligible in the case of income tax bracket variations. One possible mechanism at play is that workers may perceive marginal tax rate adjustments as more directly affecting their take-home pay, leading to a stronger behavioural response. When marginal tax rates are altered, individuals may reassess their incentives to work, as they have an immediate perception of the financial benefits or costs of their labour. This heightened sensitivity could be attributed to the fact that changes in marginal rates directly influence the amount of income retained for each additional unit of work. Since bracket changes affect a wider range of income levels and may not translate to immediate, noticeable changes in net pay for all workers, individuals might feel less compelled to alter their working hours in response.

[24] Welfare is measured as in Equaiton 2.2.1.

Table 6: Effects of the reforms

Δ Post-Reform	Reform 1	Reform 2	Reform 3	Reform 4
NTR (Million)	226 (5.49%)	-218 (-5.29%)	-138 (-3.35%)	146 (3.54%)
Labour supply	-138,659 (-0.19%)	9144 (0.01%)	3239 (0.004%)	-3260 (-0.004%)
Average disposable income ($\bar{d}$)	-74 (-0.78%)	68 (0.71%)	49 (0.51%)	-52 (-0.55%)
Gini Index, g (pp)	-0.001	0.001	0.0006	0.0006
Welfare ($\mathcal{W}$)	-17 (-0.50%)	12 (0.33%)	14 (0.37%)	-16 (-0.41%)

Notes. Every cell is the difference between the value after the reform and the base one (Δ Post-Reform). The net tax revenue (NTR) is measured in million of Euros. Labour supply is measured in number of hours worked. Average disposable income is measured in Euros. The Gini Index is defined between 0 and 1 and its variations are measured in percentage points (pp). Welfare is computed as in Equation 2.2.1.

7 Conclusions

In this paper, we use a comprehensive synthetic dataset representative of the Belgian taxpayer population to evaluate behavioural responses to marginal changes in the Personal Income Tax (PIT) structure. By employing the Belgian arithmetic microsimulation model (Beamm) alongside a Random Utility Random Opportunity (RURO) model, we derive the impact of four distinct tax reform scenarios on tax revenue, welfare, and labour supply.

Our findings reveal that none of the proposed reforms results in an optimal directional change as defined by our welfare metric. Specifically we observe that reforms leading to an increase in total labour supply are accompanied by declines in net tax revenue, thereby highlighting the detrimental impact of the current tax burden on workforce participation. However, we do observe welfare gains in two reforms, Reform 2 and Reform 3, with the latter emerging as the most advantageous approach. In fact, Reform 3, which increases all income brackets by 1000 €, achieves a more substantial welfare improvement with only a modest reduction in tax revenue and a slight increase in labour supply.

On the other hand, it is noteworthy that all four reforms exhibit marked neutrality with respect to income inequality, as evidenced by minimal fluctuations in the Gini Index and comparable changes in average disposable income across all scenarios. This underscores the conclusion that marginal reforms exert a limited influence on income distribution, primarily affecting tax revenue dynamics.

The sensitivity of workers to marginal tax rate adjustments suggests that reforms focused on altering these rates elicit more significant changes in labour supply. Conversely, changes in income tax brackets appear to have a negligible impact on working hours. This result has important policy implications. If the State's objective is a smoother balance between fiscal responsibility and labour market incentives, reforms that adjust the brackets could be considered a more optimal strategy for tax policy. To enhance a concrete behavioural response in terms of labour supply, however, modifying the rates could be more effective.

This study contributes to the broader literature on the interplay between taxation, income redistribution, and labour supply by providing empirical evidence from a high-tax context such as Belgium. The integration of a structural labour supply model with a detailed microsimulation framework allows for a more comprehensive assessment of tax policy impacts. It highlights the importance of considering both behavioural responses and distributional outcomes in tax reform evaluations, and it advances the understanding of tax policy implications.

References

Aaberge, R., & Colombino, U. (2018). Structural labour supply models and microsimulation. *International Journal of Microsimulation*, *11*(1), 162–197.

Aaberge, R., Colombino, U., & Strøm, S. (1999). Labor supply in Italy: An empirical analysis of joint household decisions, with taxes and quantity constraints. *Journal of Applied Econometrics*, *14*(4), 403–422.

Aaberge, R., Dagsvik, J. K., & Strøm, S. (1995). Labour supply responses and welfare effects of tax reforms. *Scandinavian Journal of Economics*, *97*(4), 633–659.

Anderson, T. W. (1995). Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *J Am Stat Assoc*, *52*, 200–203.

Annoye, H., Beretta, A., & Heuchenne, C. (2024). Statistical matching using kernel canonical correlation analysis and super-organizing map. *Expert Systems with Applications*, *246*, 123–134.

Apostel, A., & O'Neill, D. (2022). A one-off wealth tax for Belgium: Revenue potential, distributional impact, and environmental effects. *Ecological Economics*, *196*.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, *49*(1), 3–71.

Bach, S., Corneo, G., & Steiner, V. (2012). From bottom to top: The entire income distribution in Germany, 1992-2003. *The Review of Income and Wealth*, *58*(2), 211–232.

Bierbrauer, F., Boyer, P., & Hansen, E. (2023). Pareto-improving tax reforms and the earned income tax credit. *Econometrica*, *91*(3), 1077–1103.

Bosch, N., Gielen, M., Jongen, E., & Mastrogiacomo, M. (2017). Heterogeneity in labour supply responses: Evidence from a major tax reform. *Oxf Bull Econ Stat*, *79*(1), 769–796.

Bourguignon, F., Fournier, M., & Gurgand, M. (2001). Fast development with a stable income distribution: Taiwan, 1979-94. *Review of Income and Wealth*, *47*, 139–163.

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M., & Piketty, T. (2020). Predistribution vs. Redistribution: Evidence from France and the U.S. *World Inequality Lab Working Paper*, *22*.

Broyden, C. (1970). The convergence of a class of double-rank minimization algorithms: 2. The new algorithm. *IMA Journal of Applied Mathematics*, *6*, 222–231.

Capéau, B., Decoster, A., & Dekkers, G. (2016). Estimating and simulating with a random utility random opportunity model of job choice presentation and application to Belgium. *International Journal of Microsimulation*, *9*(2), 144–191.

Capéau, B., Decoster, A., Maes, S., & Vanheukelom, T. (2018). Piecemeal modelling of the effects of joint direct and indirect tax reforms. *Faculty of Economics and Businsess, KU Leuven, Discussion Paper Series*, *18*(10).

Capgemini. (2022). The World Wealth Report. *World Report Series Wealth Management.*

Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2022). World Inequality Report. *World Inequality Lab.*

Cheng, J., Yang, Y., Tang, X., Xiong, N., Zhang, Y., & Lei, F. (2020). Generative Adversarial Networks: A Literature Review. *KSII Transactions on Internet & Information Systems*, *14*(12).

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

Cingano, F. (2014). Trends in income inequality and its impact on economic growth. *OECD Social, Employment and Migration Working Papers*, *163*.

Creedy, J., Gemmell, N., Hérault, N., & Mok, P. (2018). Microsimulation analysis of optimal income tax reforms. An application to New Zealand. *Working Paper 08/2018.*

Dagsvik, J. K., & Strøm, S. (1995). Labor supply with non convex budget sets, hours restriction and non pecuniary job attributes. *Technical report, Manuscript, Statistics Norway.*

Dagsvik, J. K., & Strøm, S. (2006). Sectoral labour supply, choice restrictions and functional form. *Journal of Applied Econometrics*, *21*(6), 803–826.

Davidson, R., & Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, *68*(6), 1435–1464.

Decoster, A., De Swerdt, K., & Orsini, K. (2010). A Belgian flat income tax. effects on labour supply and income distribution. *Review of Business and Economic Literature*, *0*(1), 23–54.

de Mahieu, A. (2021). In-work benefits in Belgium: Effects on labour supply and welfare. *International Journal of Microsimulation*, *14*(1), 43–72.

Dianov, S., Koroleva, L., Pokrovskaia, N., Victorova, N., & Zaytsev, A. (2022). The Influence of Taxation on Income Inequality: Analysis of the Practice in the EU Countries. *Sustainability*, *14*(9066).

D'Orazio, M., Di Zio, M., & Scanu, M. (2006). *Statistical Matching: Theory and Practice*. John Wiley & Sons.

EC. (2022). 2022 Country Report - Belgium. *2022 Stability Programme of Belgium*.

Federal Public Service FINANCE. (2023). Tax Rates. Retrieved May 25, 2023, from https://finance.belgium.be/en/private-individuals/tax-return/tax-rates-income/tax-rates#q1.

Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal*, *13*(3), 317–322.

Fourrier-Nicolaï, E., & Lubrano, M. (2019). *The Effects of Aspiration on Inequality: Evidence from the German Reunification using Bayesian Growth Incidence Curves* [PhD Thesis].

Goldfarb, D. (1970). A family of variable metric updates derived by variational means. *Mathematics of Computation*, *24*(109), 23–26.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Nets*. Curran Associates, Inc.

Høj, J. (2009). Comment réformer le système fiscal belge afin de renforcer l'expansion économique. *Documents de travail du Département des Affaires économiques de l'OCDE*, *741*.

IMF. (2022). Central Government Debt: Percent of GDP. Retrieved November 7, 2023, from https://www.imf.org/external/datamapper/CG_DEBT_GDP@GDD/BEL.

Keane, M. (2010). The Tax-Transfer System and Labour Supply. *MPRA Paper, University Library of Munich*, *55167*.

Kleven, H., Kreiner, C., & Saez, E. (2009). The optimal income taxation of couples. *Econometrica*, *77*, 537–360.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *in: P. Zarembka (ed.), Frontiers in Econometrics, Academic Press: New York*, 105–142.

Meghir, C., & Phillips, D. (2010). *Labour supply and taxes*. Oxford: OUP.

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

Müllbacher, S., & Nagl, W. (2017). Labour supply in Austria: An assessment of recent developments and the effects of a tax reform. *Empirica*, *44*(3), 465–486.

O'Donoghue, C. (2014). *Handbook of Microsimulation Modelling*. Emerald Insight.

OECD. (2023). Wealth. Retrieved February 24, 2023, from https://stats.oecd.org/Index.aspx?DataSetCode=WEALTH.

OECD. (2023). Tax wedge (indicator). Retrieved December 15, 2023, from 10.1787/cea9eba3-en.

OECD. (2023). Income inequality (indicator). Retrieved February 24, 2023, from https://data.oecd.org/inequality/income-inequality.htm.

OECD. (2022). *Oecd economic surveys: Belgium 2022*. https://doi.org/https://doi.org/https://doi.org/10.1787/01c0a8f0-en.

OECD. (2023). *Taxing wages 2023*. https://doi.org/https://doi.org/https://doi.org/10.1787/8c99fa4d-en.

Piketty, T. (2014). *Capital in the twenty-first century*. Harvard University Press.

Piketty, T., & Saez, E. (2013a). *Optimal labor income taxation* (In A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds), Handbook of Public Economics, Vol. 5). Elsevier.

Piketty, T., & Saez, E. (2013b). A theory of optimal inheritance taxation. *Econometrica*, *81*(5), 1851–1886.

Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy*, *6*(1), 230–271.

Piketty, T., Saez, M., & Stantcheva, S. (2011). Taxing the 1%: Why the top tax rate could be over 80%. *VOXEU CEPR*.

Rässler, S. (2002). *Statistical Matching: A Frequentist Theory, Practical Applications, and Alternative Bayesian Approaches* (Springer Science & Business Media, Vol. 168). Lecture Notes in Statistics.

Ravallion, M., & Chen, S. (2003). Measuring pro-poor growth. *Economics Letters*, *78*, 93–99.

Saez, E., & Diamond, P. A. (2012). The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives*, *26*(4), 43–56.

Saez, E., & Stantcheva, S. (2016). Generalised Social Marginal Welfare Weights for Optimal Taxation. *American Economic Review*, *106*(1), 24–45.

Saez, E., & Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, *131*(2), 519–578.

Saez, E., & Zucman, G. (2020). *The triumph of injustice*. WW Norton.

Shanno, D. (1970). Conditioning of Quasi-Newton methods for function minimization. *Mathematics of Computation*, *24*(111), 647–656.

Son, H. H. (2004). A note on pro-poor growth. *Economics Letters*, *82*(3), 307–314.

Stephenson, A. V. (2018). The Impact of Personal Income Tax Structure on Income Inequality for Belgium, Bulgaria, Germany, Lithuania, and Poland. *Atlantic Economic Journal*, *46*(4), 405–417.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

Trading economics. (2022). List of Countries by Personal Income Tax Rate. Retrieved February 24, 2023, from https://tradingeconomics.com/country-list/personal-income-tax-rate.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge University Press.

Truyts, T., Sheremeta, V., & Sonnewald, D. (2023). *Beamm.brussels: An online, open-access tax-benefit microsimulation model for the Brussels Capital Region* [Working Paper].

Van Soest, A. (1995). Structural models of family labor supply: A discrete choice approach. *The Journal of Human Resources*, *30*(1), 66–88.

Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., & Wang, F. Y. (2017). Generative adversarial networks: Introduction and outlook. *IEEE/CAA Journal of Automatica Sinica*, *4*(4), 588–598.

The World Bank. (2021). GDP per capita (current US$). Retrieved February 24, 2023, from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2021&most_recent_value_desc=true&start=2021.

Appendices

A Stochastic Dominance

A.1 Growth Incidence Curves and Poverty Growth Curves

The function that we use as welfare metric (Equation 2.2.1) is constructed along the one used by Creedy et al. (2018). Although it is designed to account for both equity and efficiency, it requires some assumptions on the weights of each component involved in this trade-off.[25] To verify that our evaluations are not sensitive to the metric chosen, we also assess welfare changes using stochastic dominance, which does not depend on a pre-defined function.

The general rule for stochastic dominance is formalised in Davidson and Duclos (2000). In the present study, we use the simplified formulations of Ravallion and Chen (2003) and Son (2004) for first- and second-order dominance, which employ growth incidence curves and poverty growth curves, respectively. Given our reforms' simulation (see Table 2), let y_0 and y_r be the income distributions before and after a reform r, respectively, $p \in [0, 1]$ any quantile, $L'(p)$ the slope of the Lorenz curve $L(p)$, and $\gamma_r = \left(\frac{\mu_r}{\mu_0}\right) - 1$ the growth rate in μ (mean). The growth incidence curve (GIC) is defined as:

$$g_r(p) = \frac{L_r'(p)}{L_0'(p)}(\gamma_r + 1) - 1 \tag{A.1.1}$$

The poverty growth curve (PGC) is defined as:

$$G_r(p) = \Delta Ln(\mu) + \Delta Ln(L(p)) \tag{A.1.2}$$

where $\Delta Ln(\mu) = Ln(\mu_r) - Ln(\mu_0)$ and $\Delta Ln(L(p)) = Ln(L_r(p)) - Ln(L_0(p))$.

Therefore, there is stochastic dominance of the post-reform distribution (y_r) over the pre-reform one (y_0):

- of the first-order (FOD), if the GIC is positive for every quantile: $g_r(p) > 0$, $\forall p \in [0, 1]$;
- of the second-order (SOD), if the PGC is positive for every quantile: $G_r(p) > 0, \forall p \in [0, 1]$.

A.2 Assessment of Reforms

We use stochastic dominance to assess whether our optimal directional changes' evaluation carried out in Section 6 is robust to a different measure of welfare. That is, we verify whether

[25] Average income ($\bar{d}$) has a weight of 1; the Gini Index (g) has a weight of $-\bar{d}$.

reforms that entail optimal directional changes according to the rule $\Delta\mathcal{W} = \mathcal{W}_r - \mathcal{W}_0 > 0$ based on Equation 2.2.1, do so also according to a stochastic dominance criterion. From Definition 1, while we keep the conditions on non-negativity of the differences in the net tax revenue and labour supply unchanged, we let the welfare improving condition $\Delta\mathcal{W} > 0$ to be determined by a stochastic dominance of the post-reform income distribution (y_r) over the pre-reform one (y_0), either of the first- or second-order.[26]

B Synthetic Data

To assess the quality of statistical matching, we compare the distribution of individuals in our synthetic data with that in the EU-SILC data for 2019 (one of the data sources used for statistical matching). Table B.1 replicates Table 4 for the EU-SILC data. Figures B.1 and B.2 compare weekly hours worked and age groups across the two data sets.[27]

[26] Stochastic dominance of the n-order implies automatic dominance of all subsequent orders.

[27] These figures are preliminary, as our final synthetic data set is still being fine-tuned. With the current calibration, asymmetries are particularly pronounced for the extra-time.

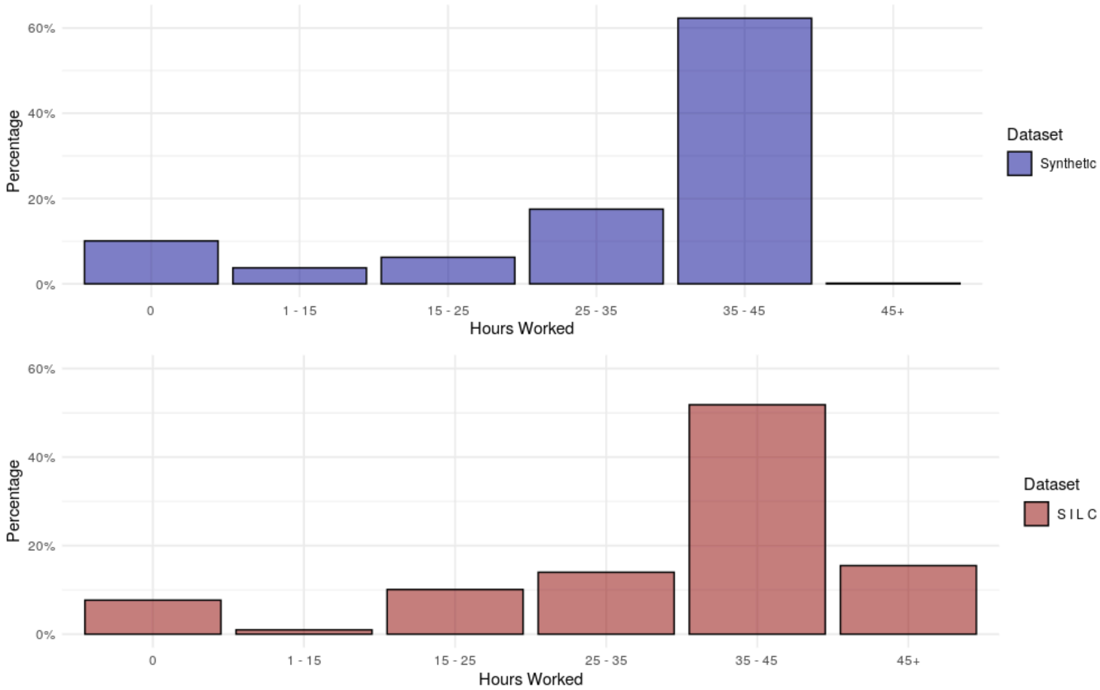
Table B.1: Breakdown by individuals characteristics and labour status

	Employed	**Self-Employed**	**Unemployed**
Gender			
Men	49.91%	65.15%	51.84%
Women	50.09%	34.85%	48.16%
Age			
18-24	4.34%	2.06%	9.98%
25-34	24.02%	17.79%	21.91%
35-44	26.38%	24.26%	16.27%
45-54	26.42%	30.00%	21.48%
55-65	18.84%	25.88%	30.37%
Marital Status			
Single	51.20%	46.03%	72.77%
Couple	48.80%	53.97%	27.23%
Education level			
Low	12.28%	13.09%	31.67%
Middle	33.10%	33.24%	40.35%
High	54.62%	53.68%	27.98%
Country of birth			
Belgium	82.52%	78.24%	66.81%
EU27	8.28%	13.24%	8.68%
Other	9.20%	8.53%	24.51%
Labour market			
Hours worked weekly (mean)	36.44	49.96	
Hourly gross wage (mean)	50.93	44.39	
Gross yearly labour income (mean)	84750.00	81535.00	
Share	82.23%	10.59%	7.18%
Total Observations	5,280	680	461

Source. SILC data 2019.

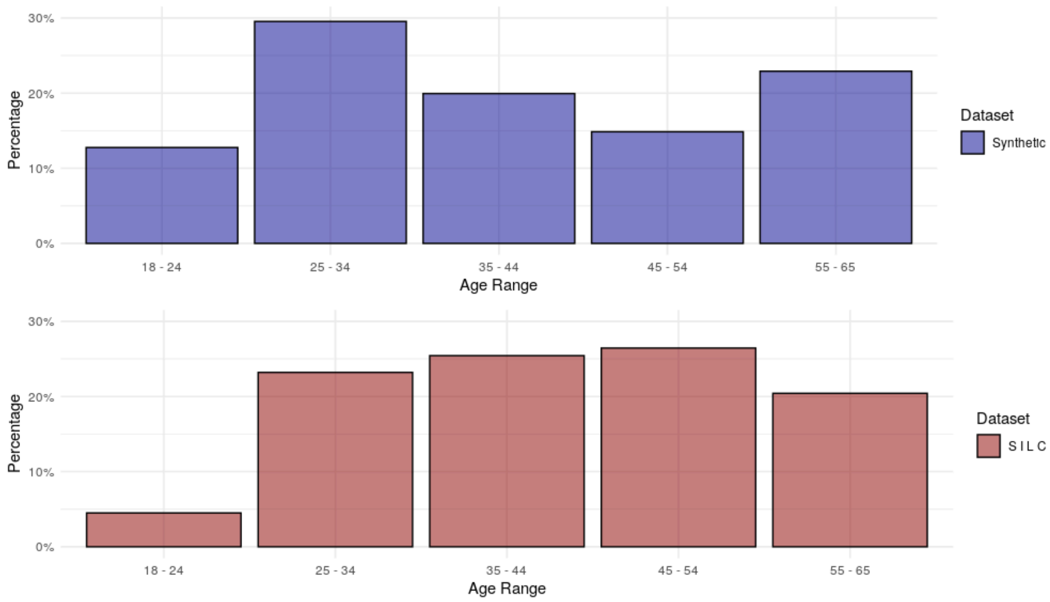
Notes. The table is structured vertically (shares sum up to 100% in every category, *e.g.,* gender). “Education Level”: Low = up to primary education; Middle = secondary education; High = university degree or higher.

Figure B.1: Distribution by weekly hours worked: Synthetic Data Vs. SILC



Source. Synthetic data generated through statistical and EU-SILC data.

Figure B.2: Relative frequency of age groups: Synthetic Data Vs. SILC



Source. Synthetic data generated through statistical and EU-SILC data.

C RURO specification

Table C.2: Covariates

	Preferences	Opportunities		
		intensity	job offers	wages
Variable	X	Z_o	Z_h	Y
Region of residence	yes	yes	yes	no
Level of education (low, middle, high)	yes	yes	yes	yes
Age	yes	no	no	no
Group-specific unemployment	no	yes	yes	no
Number of children	yes	no	no	no
Gender	yes	yes	yes	yes
Experience	no	no	no	yes
Marital Status (single vs. couple)	yes	yes	yes	yes
Hours worked	no	no	no	no
Hourly net wage	no	no	no	yes
Type of contract (part-time vs. full time)	no	yes	yes	no

Notes. Columns 2 to 5 indicate in what equations of the RURO model the variables are employed for the estimation. "Group-specific unemployment" is the unemployment rate by age group (we consider five categories: 18-24, 25-34, 35-44, 45-54, 55-65). X is the vector of covariates used for the estimation of α_l, in the utility function's deterministic part. Z_o is the vector of covariates used for the estimation of α_o, Equation 5.2.3. Z_h is the vector of covariates used for the estimation of α_h, Equation 5.2.2. Y is the vector of covariates used for the estimation of γ, Equation 5.2.1.

Table C.3: Parameter estimates - Preferences singles

Preferences	Estimate	Standard Error	t-value
Single male			
Leisure			
Constant	0.270	0.505	0.535
Log(age)	-5.399	0.456	-11.829
Log(age)^2	2.691	0.234	11.515
Number of children	0.006	0.014	0.416
Brussels	0.018	0.034	0.522
Wallonia	-0.021	0.021	-1.018
Low education	-0.021	0.035	-0.591
High education	-0.045	0.023	-1.964
Exponent	-12.254	0.530	-23.131
Income			
Constant	0.005	0.019	0.279
Exponent	0.819	0.505	1.623
Single female			
Leisure			
Constant	0.259	0.506	0.511
Log(age)	-15.732	0.456	-34.460
Log(age)^2	7.905	0.235	33.599
Number of children	0.019	0.029	0.631
Brussels	-0.045	0.513	-0.095
Wallonia	-0.051	0.125	-0.407
Low education	0.094	0.505	0.187
High education	-0.023	0.138	-0.169
Exponent	-9.158	0.526	-17.397
Income			
Constant	0.185	0.014	13.371
Exponent	0.305	0.012	26.101

Notes. Estimates of parameter in vector α_l (utility function's deterministic part). The Log-Likelihood at the maximum found is 11,266.11. The reference categories are: middle education (for education level) and Flanders (for the Region). The effects of these covariates are measured comparatively to middle-educated Flemish people.

Table C.4: Parameter estimates - Preferences couples

Preferences	Estimate	Standard Error	t-value
Leisure Male			
Constant	0.251	0.506	0.496
Log(age)	-3.008	0.467	-6.436
Log(age)^2	1.538	0.309	4.974
Number of children	-0.018	0.032	-0.580
Brussels	0.011	0.523	0.022
Wallonia	-0.092	0.119	-0.768
Low education	-0.089	0.404	-0.221
High education	-0.074	0.295	-0.250
Exponent	-11.256	0.510	-22.069
Leisure Female			
Constant	0.051	0.107	-0.472
Log(age)	-13.301	0.155	-85.659
Log(age)^2	6.664	0.274	24.337
Number of children	0.000	0.013	0.004
Brussels	-0.013	0.113	-0.114
Wallonia	-0.005	0.011	-0.462
Low education	-0.005	0.134	-0.040
High education	0.003	0.052	0.065
Exponent	-17.967	0.507	-35.384
Income			
Constant	0.077	0.212	0.364
Exponent	0.539	0.128	4.219
Cross-effect	0.001	0.056	0.016

Notes. Estimates of parameter in vector α_l (utility function's deterministic part). The Log-Likelihood at the maximum found is 11,266.11. The reference categories are: middle education (for education level) and Flanders (for the Region). The effects of these covariates are measured comparatively to middle-educated Flemish people.

Table C.5: Parameter estimates - Opportunities

Opportunities	Estimate	Standard Error	t-value
Male			
Intensity			
Working	-3.899	0.509	-7.655
Group specific unemployment rate	0.011	0.509	0.021
Wallonia	-0.626	0.509	-1.230
Brussels	-1.084	0.509	-2.128
Low education	-0.894	0.509	-1.755
High education	0.484	0.509	0.950
Offered time regimes			
Part-time 1	0.618	0.509	1.213
Part-time 2	0.909	0.509	1.784
Full-time	2.543	0.509	4.993
Offered wages			
Constant	4.198	0.505	8.304
Low education	-0.226	0.078	-2.902
High education	-0.079	0.047	-1.675
Log(age)	0.896	0.457	1.958
Log(age)^2	-0.013	0.243	-2.530
Standard deviation	0.475	0.036	13.006
Female			
Intensity			
Working	-3.237	0.509	-6.355
Group specific unemployment rate	-0.021	0.509	-0.041
Wallonia	-0.540	0.509	-1.060
Brussels	-0.879	0.509	-1.725
Low education	-0.677	0.509	-1.329
High education	0.565	0.509	1.109
Offered time regimes			
Part-time 1	1.496	0.509	2.937
Part-time 2	1.731	0.509	3.398
Full-time	2.055	0.509	4.035
Offered wages			
Constant	-33.016	0.561	-58.834
Low education	-1.042	0.638	-1.632
High education	4.821	0.549	8.778
Log(age)	4.389	0.464	9.446
Log(age)^2	-0.455	0.270	-1.682
Standard deviation	9.005	0.501	17.951

Notes. Estimates of parameter in vectors α_o (Equation 5.2.3), α_h (Equation 5.2.2), and γ (Equation 5.2.1). "Group-specific unemployment" is the unemployment rate by age group (we consider five categories: 18-24, 25-34, 35-44, 45-54, 55-65). The Log-Likelihood at the maximum found is 11,266.11. The reference categories are: middle education (for education level) and Flanders (for the Region). The effects of these covariates are measured comparatively to middle-educated Flemish people.