EUROMOD WORKING PAPER SERIES

EM 14/18

Piecemeal modelling of the effects of joint direct and indirect tax reforms

Bart Capéau, André Decoster, Sebastiaan Maes and Toon Vanheukelom

September 2018

Piecemeal modelling of the effects of joint direct and indirect tax reforms*

Bart Capéau, André Decoster, Sebastiaan Maes and Toon Vanheukelom

Department of Economics, KU Leuven

Abstract

This paper offers a framework to establish a micro-based budget and welfare evaluation of a joint reform in personal income taxes, social security contributions and indirect taxes. One often lacks an encompassing model for both labour supply decisions in real world tax and benefit contexts and the allocation of disposable income to commodities. In this paper we therefore elicit the assumptions which allow us to combine different submodels, such that an assessment of a joint reform becomes possible in a consistent conceptual framework. In addition, we characterise households' labour supply decisions by a random utility random opportunity (RURO) model of job choice. This allows us to incorporate effects from the demand side of the labour market into our analysis. We apply this framework to a recently enacted Belgian tax reform which shifts the burden away from labour taxes. We find substantial empirical evidence that, both from a distributional and from a budgetary perspective, it is important to account for indirect taxes, for labour demand-side effects and for unobserved job characteristics, when assessing this kind of joint tax reform. As for the budgetary effects, the cost recovery effects of the tax shift are modest. This is, among other things, explained by a more encompassing income effect in our job choice model, than is found in the more classic discrete choice model of labour supply.

JEL: H31, J22, J24, H23, D63

Keywords: job choice, joint direct and indirect tax reform, microsimulation, welfare analysis

Corresponding author:

André Decoster, andre.decoster@kuleuven.be

---

* We are grateful to Rolf Aaberge, Ugo Colombino, John Creedy, John Dagsvik, Norman Gemmell, Zhiyang Jia, Tom Strengs, and Tom Wennemo, as well as workshop participants at the University of Essex, the World Bank and the New Zealand Treasury for comments on and help with earlier versions of the paper. The paper benefited from financial support from the National Bank of Belgium (Sebastiaan Maes; project 3H170248), the Belgian Federal Science Policy Office BELSPO (Bart Capéau; project BR/132/A4/BEL-Ageing), and the Joint Research Centre Sevilla (Toon Vanheukelom; Contract No. 198961-2015 A10-UK). The results presented here are based on EUROMOD version G4.0+. EUROMOD is maintained, developed and managed by the Institute for Social and Economic Research (ISER) at the University of Essex, in collaboration with national teams from the EU member states. We are indebted to the many people who have contributed to the development of EUROMOD. The process of extending and updating EUROMOD is financially supported by the European Union Programme for Employment and Social Innovation `Easi' (2014-2020). We make use of microdata from the EU Statistics on Income and Living Conditions (EU{SILC) made available by Eurostat (59/2013-EU-SILC-LFS). The results and their interpretation are the authors' sole responsibility.

1 Introduction

In this paper we offer a framework that allows us to make a micro–based budget and welfare evaluation of a joint reform in both the personal tax and benefit system and the indirect tax system. Such joint tax reforms are initiated by governments worldwide in an attempt to shift part of the tax burden from labour to consumption, which is considered to be less detrimental to economic growth (see Myles, 2009a,b, and c for a review). However, there is not so much literature on micro–based empirical policy evaluations of these kinds of joint reforms.[1] Exceptions are Bach et al. (2006), Capéau et al. (2009), Pestel and Sommer (2017), and Savage (2017). From these studies, it can be inferred that such reforms may have substantial distributional effects. These papers however all proceed without an encompassing model for the labour market participation decision and the allocation of disposable income to commodities, which, at first sight, would seem a necessary tool for a consistent analysis of these distributional effects.

The reason for this gap is that such an encompassing model which at the same time sufficiently keeps track of the existing intricacies of the direct and indirect tax–benefit system, often becomes theoretically intractable, let alone useful for empirical implementation. And even if such a model did exist, few available datasets would allow for the estimation of such a model, as detailed information is required on households' gross incomes, their labour market participation, and their expenditures. Therefore, the papers cited above make use of existing disconnected microsimulation models of direct taxes and benefits on the one hand, and of indirect taxes on the other. The former are often connected with a behavioural model of labour supply, the latter with a demand system that allocates expenditures. The cited papers then glue these model pieces together in a rather ad hoc fashion, to arrive at an evaluation toolbox for the joint tax reform.

What the present paper offers is a framework that allows one to underpin how these model pieces can be fit one into another in a consistent way. Thereto, we rely on a two–stage budgeting approach.[2] The first stage models the labour supply decision, which determines households' disposable income. The second stage models the allocation of this disposable income to commodities and saving. It is known that such a two–stage budgeting approach requires the assumption of weak separability between leisure and consumption goods (Gorman, 1971). One of our main contributions is to exploit the fact that such a two–stage approach still entails the necessity of including commodity prices in the first stage decision (contrary to what was done by Bach et al., 2006, and Capéau et al., 2009; Pestel and Som-

[1] On the contrary, numerous macroeconomic evaluation tools have been developed, all investigating primarily employment and growth effects of such a shift. See for example Altig et al. (2001) for the US; Dahlby (2003) for Canada; Böhringer et al. (2005) for Germany; European Commission services (2006) for 15 EU member states; NBB (2017) for Belgium; and de Castro Fernàndez et al. (2018) for France.

[2] Similar ideas have been used to develop empirically tractable models of labour supply and commodity demand over the life cycle, by separating the within period allocation of the budget over different goods from the allocation of life time income over the different periods (See e.g. Browning et al., 1985; Blundell and Walker, 1986; and Blundell et al., 1994).

mer, 2017, who subtract indirect taxes from expenditures in the estimation of their labour supply model, do not motivate why this should be done). We argue that this inclusion of commodity prices in the first stage of the decision process, though limited to entailing an income effect of indirect tax reform, is sufficient to assess budgetary effects of, behavioural reactions to, and welfare implications of a joint tax reform.

Admittedly, from an empirical point of view, the assumption of weak separability between leisure and commodities seems unwarranted and may lead to biased estimates in the second stage, i.e. the allocation of the budget to commodities (Blundell and Walker, 1982, Browning and Meghir, 1991). We argue that it is possible to include control variables for labour market status at this stage, in order to get more consistent estimates. In line with the two–stage budgeting approach, we will however not allow changes in labour market status due to a tax reform to have something other than a pure income effect in the second stage, when simulating tax reforms.

As far as the first step is concerned, we use a random utility random opportunity (RURO) discrete choice model of job choice (see e.g. Aaberge and Colombino, 2014; Dagsvik et al., 2014).[3] This has two advantages. Firstly, contrary to classical discrete choice models of labour supply (e.g. Van Soest, 1995), this model considers a job to consist of a package of attributes: the labour time regime, the wage paid, and other pecuniary and non–pecuniary attributes. As such, offered wages become an aspect of the elements in the choice set. This allows us to capture a number of behavioural reactions which have hitherto received little attention in economic theory. For example, in our framework, jobs with lower gross wages may become more attractive after a reform in the tax–benefit schedule. Secondly, by jointly estimating preferences and opportunities, RURO also allows us to integrate feedback from the demand–side of the labour market into the analysis.[4] Especially when the reform involves modifications in employers' social security contributions, the ability to integrate feedback effects from the labour demand side is a major advantage.

We deliberately kept the modelling of the second step extremely simple. We estimate budget shares for each commodity group by means of parametric Engel curves, and then assume Cobb–Douglas preferences characterised by these estimated shares. For reasons mentioned before, these Engel curves are made dependent on labour market participation of household members. Given the Cobb–Douglas assumption together with the two–stage budgeting approach, we then treat the estimated budget shares as parameters, and keep them fixed throughout our simulations. The constant shares assumption could be relaxed and/or replaced by more complex models, but one might wonder whether this is necessary when it comes to evaluating joint tax reforms. Indeed, despite its restrictive character, this specifica-

[3] Only part of the active population is modelled. The model is thought not to be suitable to capture labour supply decisions of the self–employed or interactions between members of households with complex structures.

[4] This approach is known as top–down approach of integrating macro– and micro–simulation models (see Chen and Ravallion, 2004; Peichl, 2009, 2016; Cockburn et al., 2014).

tion allows us to capture the real income effect of an indirect tax reform on the labour supply decision in a convenient fashion, and this might well be quantitatively the most important impact.

Our approach also suggests some welfare measures to study the distributional impact of the reform, which go beyond disposable income, the latter often being the sole focus of policy makers. At the level of the second stage, we derive a nonparametric upper bound for the welfare change in equivalent variation. One the one hand, this nonparametric approach might attenuate possible bias in the distributional analysis caused by the restrictive character of Cobb–Douglas preferences over commodities, used in the simulation stage. On the other hand, as it is an *upper* bound for any underlying preferences, it tends to evaluate reforms more positively than they actually are, whatever the underlying preferences might be. Since the equivalent variation is derived at the second stage, income gains due to changes in the labour supply decisions as reaction to a tax reform, are considered as lump sum gains. Using estimated preferences underlying the first stage decision process, we also derive welfare measures which take into account the welfare cost of changes in labour time. Because of interindividual differences in the wage rate, every individual faces a distinct price vector, which renders equivalent variation based measures unattractive. We therefore revert to a class of measures proposed by Fleurbaey (2006) that explicitly fix a reference wage across individuals.

To analyse the budgetary effect of the joint reform, we advance a decomposition into first and second order components. The former measures the impact of the reform on government revenues when there is no change in individual behaviour. That is, individuals are not allowed to adapt their bundle of commodities nor their job choice. The second order component, by contrast, collects the change in revenues that can only be ascribed to changes in individual behaviour.

To illustrate the framework proposed in this paper, we perform policy simulations that approximate a tax shift enacted in Belgium in the period 2016–2020. It concerns a multi–year tax reform of which the first measures came into force as of 2016. The reform's principal aim is to lower personal income taxes and social security contributions for both employees and employers. Although the reform is not revenue neutral, the part that is financed comes primarily from increases in VAT (value added tax) rates (e.g. electricity) and excise tax hikes (e.g. alcohol, tobacco, and diesel fuel).

The paper is organised as follows. In Section 2, we present the main components of our framework and explain how these components are linked. Section 3 explains how to measure budgetary effects and how to decompose them into first and second order effects. This section also introduces the various individual welfare measures we will apply. In Section 4 we discuss our empirical modelling strategy. Section 5 contains our simulation results for the recently enacted Belgian tax shift and for an alternative scenario in which a much larger hike of VAT is introduced in order to bridge part of the remaining financing gap. Section 6 concludes the paper. Appendix A.1 discusses the construction of the subsample on which the RURO model operates, Appendix A.2 contains the estimated parameters of the RURO model, the simulated labour supply elasticities, and the model fit, Appendix A.3 explains the procedure we used to simulate with the RURO model, Appendix A.4 gives a brief overview of the recently enacted Belgian tax shift, Appendix A.5 explains how our job choice model allows us to integrate the effects of the reform on the demand side of the labour market, Appendix A.6 discusses why it is often unwarranted to compare differences in welfare across the measures we introduce in Section 3.2, and Appendix A.7 contains additional simulation results.

2 Piecemeal modelling of joint direct and indirect tax reforms

In its most general form, a (static) consumer decision model jointly treats the labour supply decision and the allocation of disposable income to commodities and saving. Formally, let h represent labour time, $\mathbf{x}$ an n-vector of commodities, and $\mathbf{q}$ the associated vector of strictly positive consumer prices, and suppose that $\Omega(\cdot)$ denotes a utility function representing preferences over commodities and labour time.[5] Then, the integrated decision model for labour market participation and consumption is represented by the program

$$\begin{aligned} &\max_{\mathbf{x},h} \; \Omega(\mathbf{x}, h) \\ \text{s.t. } &\mathbf{q}'\mathbf{x} \leq f(w, h; M, \mathbf{z}) \\ &\mathbf{x} \geq 0 \\ &0 \leq h \leq T, \end{aligned} \tag{1}$$

in which $f(\cdot)$ embodies the tax–transfer system, and T denotes total time endowment. Disposable income $y = f(w, h; M, \mathbf{z})$, is a function of gross wages w, labour time h, unearned gross income M, and a vector of individual and/or household characteristics $\mathbf{z}$.

Such models of joint determination have been formulated and successfully empirically implemented in the literature (see e.g. Blundell and Walker, 1982, 1986; Browning et al., 1985; Browning and Meghir, 1991). These contributions, however, refrain from modelling the complexity of tax–transfer systems by assuming that labour income is simply the product of labour time and net wages. This renders these models less suitable for a more detailed assessment of the impact of the tax–benefit system on consumers' behaviour. However, introducing a more detailed description of the tax–benefit system poses a lot of intricate problems, as most existing tax–benefit systems cause the budget set $\{\mathbf{x} \in \mathbb{R}^n_+ \mid \mathbf{q}'\mathbf{x} \leq f(w, h; M, \mathbf{z})\}$ to be non–convex, combined with kinks and jumps in $f(\cdot)$ (see e.g. Hausman, 1981; 1985a,b). Such highly non–linear tax schemes often cause the optimisation program in (1) to be analytically and numerically intractable.

[5] Throughout the paper, we denote vectors by boldface, and the i-th element of a vector $\mathbf{v}$ by v_i.

Many of these issues have been resolved by the introduction of discrete choice modelling into the empirical labour supply literature (popularised by Van Soest, 1995; for overviews, see Aaberge and Colombino, 2014; Blundell and MaCurdy, 1999; Creedy and Kalb, 2005; Blundell et al., 2007; Keane, 2011; Keane et al., 2011). In this approach, the budget constraint $f(\cdot)$ is discretised along the hours margin, yielding a finite number of alternatives from which individuals select the option that delivers the highest utility. In combination with a detailed micro–simulation model, tax–transfer systems of virtually any complexity can be analysed in this framework.

However, the price to be paid for this increased realism on the side of labour supply modelling, is that one reverts to a simple trade–off between disposable income and leisure, irrespective of the allocation of the former to different consumer goods. This independence between the labour supply decision on the one hand and the allocation of the income generated by it on the other, is only warranted if one assumes weak separability between consumer goods and leisure in the preference structure (Gorman, 1971). Unfortunately, this assumption was subject to much criticism when it comes to empirical applications (see e.g. Blundell and Walker, 1982; Browning and Meghir, 1991). Estimates of commodity demand functions can be severely biased when the erroneous assumption of separability between budget allocation and choice of leisure time is maintained. However, as mentioned before, curing the defect would force one to revert to simple, linear budget constraints.

So it seems as if one faces a choice: either using a labour supply model in which real world tax–benefit systems are integrated, but without indirect taxes and detailed consumption decisions integrated in the analysis; or modelling consumption decisions in great detail, but without the possibility to link this with a sufficiently realistic behavioural labour supply model. On top of this, even if a tractable general model for labour supply and the allocation of disposable income to commodities would be available, few datasets contain the information necessary to estimate such a model, as information on both gross labour income and disaggregated expenditures is not available.

In the absence of both a suitable encompassing model and the data to estimate such a model, we therefore propose a *piecemeal modelling* strategy to assess the impact of a joint tax reform at the micro–level. Given the limitations outlined above, our methodology proposes a consistent integration of different submodels, which are allowed to interact to the maximal extent. This interaction takes two forms. First, in the construction of an integrated dataset with both income and detailed expenditures, we rely on parametric Engel curves, estimated on a detailed budget survey, to impute expenditures in an income survey. To attenuate the impact of the assumption of separability between budget allocation and labour supply, we included labour market status variables as covariates in the estimation of the Engel curves. Second, we advance a two–stage budgeting approach in which we allow (changes in) relative consumer prices to impact the labour supply decision. While this impact of prices of goods in the second stage on the first stage decision is well established in theoretical literature, it seems to have been largely overlooked in empirical applications. We first explain the second interaction in Sections 2.1 and 2.2. The interaction through the imputation of expenditure data into the income dataset is explained in Section 4.

2.1 Two–stage budgeting approach

Maintaining the assumption of weak separability, we can rewrite the overall utility function $\Omega(\mathbf{x},h)$ in expression (1) as

$$\Omega(\mathbf{x},h) = H(u(\mathbf{x}),h), \tag{2}$$

where $u(\mathbf{x})$ denotes a subutility function over consumption goods. During the first stage, an individual faces the trade–off between leisure and income, which amounts to solving the following problem:

$$\max_{y,h} V(y,h) \quad \text{s.t. } y \leq f(w,h;M,\mathbf{z}). \tag{3}$$

Further on, we show how preferences specified in the income–labour time space, represented by the utility function $V(\cdot)$, can be constructed from the general model (2).

The second stage of the two–stage budgeting approach consists in the allocation of the budget $y \equiv f(w^*,h^*;M,\mathbf{z})$ determined by the chosen job (w^*,h^*) to the set of commodities $\mathbf{x}$.[6] Thanks to the weak separability assumption between consumer goods and leisure, the decision model for this second stage is summarised by the subprogram

$$\begin{aligned} &\max_{\mathbf{x}} u(\mathbf{x}) \\ \text{s.t. } &\mathbf{q}'\mathbf{x} \leq y. \end{aligned} \tag{4}$$

The solution to this program constitutes a vector collecting the Marshallian demand functions

$$\mathbf{x}^* = \boldsymbol{\xi}(\mathbf{q},y). \tag{5}$$

The indirect utility function is then defined as:

$$v(\mathbf{q},y) \equiv u(\boldsymbol{\xi}(\mathbf{q},y)) \tag{6}$$

Replacing $u(\mathbf{x})$ in Equation (2) with this indirect utility function yields a representation of preferences in the income–labour time space, referred to in Equation (3):

$$V(y,h) = H(v(\mathbf{q},y),h). \tag{7}$$

Notice that with this notation, the functional form of $V(y,h)$ incorporates the dependency on commodity prices. In as far as there is interindividual heterogeneity in preferences over

[6] As we work with a static model, the set of commodities also includes saving. We refer to Section 2.2 for more details on our job choice model.

commodities, omitting these variables, as is usually done in discrete choice models of labour supply, can bias results, even if all individuals faced the same commodity prices, $\mathbf{q}$. It is by rendering this dependency explicit, that it becomes clear that even in a labour supply model resulting from weakly separable preferences over leisure and commodities, relative commodity prices have an effect on labour supply.[7] It is this dependency that we will fully exploit in order to investigate the effects of a joint tax reform.

In particular, assume that $u(\cdot)$ belongs to the class of Cobb–Douglas utility functions,

$$u(\mathbf{x}) = \prod_{i=1}^{n} x_i^{\omega_i}, \tag{8}$$

with $0 \leq \omega_i \leq 1,\ i = 1, \ldots, n$, and $\sum_{i=1}^{n} \omega_i = 1$. As the Marshallian demand functions (5) for these preferences are given by $\xi_i(\mathbf{q}, y) = \omega_i y / q_i$, the parameters ω_i can be interpreted as the budget share of commodity i:

$$\omega_i = \frac{q_i \xi_i(\mathbf{q}, y)}{y}. \tag{9}$$

An indirect utility function for this class of preferences equals

$$v(\mathbf{q}, y) = \frac{y}{Q(\mathbf{q})}, \text{with } Q(\mathbf{q}) = \prod_{i=1}^{n} q_i^{\omega_i}, \tag{10}$$

in which $Q(\mathbf{q})$ is known as a Divisia price index. As a consequence, the indirect utility function in (10) yields a real income concept, using the Divisia index as a price deflator. The deflator is household specific through the budget shares $\boldsymbol{\omega}$.

Plugging expression (10) in the overall utility function (2), we obtain

$$\begin{aligned} \tilde{V}(\mathbf{q}, y, h) &= H\left(\frac{y}{Q(\mathbf{q})}, h\right) \\ &\equiv H(c, h), \end{aligned} \tag{11}$$

in which $c = y/Q(\mathbf{q})$ is a measure for consumption in real terms.[8] In this case, the labour supply model should thus be estimated using deflated disposable income, and since the deflator is household specific, this is more than just a normalisation issue.[9] Furthermore, this specification, though very restrictive, still allows us to feed the impact of indirect tax reforms — that is, changes in consumer prices $\mathbf{q}$ — back into the labour supply decision.

As the Cobb–Douglas case is easy to implement, especially when price effects are difficult

[7]It is known from demand theory that the compensated commodity price effects on labour supply are proportional to an income effect when preferences are weakly separable over leisure and consumption (Barten and Böhm, 1986).

[8]We make the dependency of preferences in the income–labour time space on commodity prices $\mathbf{q}$ explicit by including it as an argument. Hence the notation $\tilde{V}(\cdot)$ instead of $V(\cdot)$.

[9]In this specific case also uncompensated commodity price effects on labour supply are proportional to an income effect.

to estimate using richer specifications, we proceed the discussion for this specific case.[10]

2.2 Job choice model and labour demand feedback

To model the first–stage labour supply decision, we employ a random utility random opportunity (RURO) framework (see Aaberge et al., 1995; Aaberge et al., 1999; and Dagsvik and Strøm, 2006; for surveys, see Aaberge and Colombino, 2014; and Dagsvik et al., 2014). The RURO model differs from the standard discrete choice multinomial logit model for labour supply (McFadden, 1973; Van Soest, 1995) in two ways. First, in contrast to the standard model, an individual chooses a job rather than optimal working hours. A job consists of a wage offer, w, a labour time regime, h, and a number of other pecuniary and non–pecuniary attributes (e.g. fringe benefits, challenge, prestige, …). Second, the RURO model introduces demand–side restrictions in a structural fashion. Job availability is modelled in RURO by an individual specific stochastic process governing the probability that jobs with a specific wage and labour time regime are offered to that individual. The availability of certain jobs may not only depend on an individual's personal characteristics and capabilities, but also on the demand side of the labour market and on macroeconomic fluctuations. A similar reasoning holds for non–market alternatives: their availability depends on certain abilities an individual might possess and on the availability of the infrastructure and institutions that facilitate particular leisure activities. Consequently, the relative availability of job offers versus non–market alternatives in a RURO model may depend both on personal characteristics and macroeconomic circumstances.

From the previous subsection, it turns out that the labour supply model should be specified in real terms, i.e. deflating disposable incomes obtained from a particular job choice, by an individual specific Divisia price index.

Formally, let $\mathcal{B}$ denote the set of all market and non–market alternatives available to an individual. From the econometrist's point of view, the probability that an individual prefers alternative (w_k, h_k) over all other alternatives in this set can then be expressed as follows:

$$P(w_k, h_k \mid \mathcal{B}) = \frac{\exp\left[H\left(f(w_k, h_k; M, \mathbf{z})/Q(\mathbf{q}), T - h_k\right)\right]\varphi(w_k, h_k)}{\int_{(w,h)\in\mathcal{B}} \exp\left[H\left(f(w, h; M, \mathbf{z})/Q(\mathbf{q}), T - h\right)\right]\varphi(w, h)\,\mathrm{d}w\,\mathrm{d}h}. \tag{12}$$

Note that this equation constitutes a weighted version of the likelihood contribution in the standard multinomial logit framework, where the probability to choose an alternative k only depends on its relative attractiveness as embodied by the utility function $H(\cdot)$. In the RURO model each alternative is in addition weighted by a measure $\varphi(w, h)$ that captures the likelihood that an alternative (with specific wage w and labour time regime h) will be available in the individual–specific choice set. If all alternatives are equally available and the

[10] We have only a cross section dataset at our disposal, and our approach stipulates commodity prices to be identical across individuals, so that we cannot exploit price variation to estimate richer price effects. More details on our estimation strategy can be found in Section 4.

wages do not vary over jobs, the weights cancel out and expression (12) would reduce to the standard multinomial logit formula. For more details on the assumptions that underpin our implementation of the RURO model, we refer to Capéau and Decoster (2016) and Capéau, Decoster and Dekkers (2016).

In our modelling strategy, we represent individuals' opportunities and preferences by the following functional forms.

- Opportunities

$$\frac{\varphi(w,h)}{\varphi(0,0)} = \begin{cases} g_1(w)g_2(h)\theta, & \text{if } w,h > 0, \\ 1, & \text{if } w,h = 0, \end{cases} \tag{13}$$

where the distribution of offered wages, $g_1(w)$, is lognormal with a sex, education, and experience specific location parameter and a sex specific scale parameter. Offered hours follow a sex specific piecewise uniform distribution, $g_2(h)$, with peaks at half–time, three quarters and full–time working hours; and θ is a measure for the relative intensity of job offers versus the availability of non–market alternatives, dependent on sex, age, education, region and a type specific unemployment rate. We consider the latter to be a proxy for the macroeconomic impact on individual job offer availability. The measure θ is positively valued and can be converted into a probability measure:

$$\pi_1 = \frac{\theta}{1+\theta}, \tag{14}$$

in which π_1 can be interpreted as the number of job opportunities relative to the total number of market and non–market opportunities for an individual. Dagsvik and Jia (2016) show how the separability between the role of w and h in the specification of the opportunities introduced in (13) is necessary for a partial nonparametric identification of the model.[11]

- Preferences

$H(c,h)$ is a gender specific Box–Cox utility function with marginal rates of substitution dependent on age, education, region and the number of children (these variables are denoted by the vector $\mathbf{r}$):

$$H(c,h) = \beta_c \frac{c^{\alpha_c}-1}{\alpha_c} + \boldsymbol{\beta}_l'\mathbf{r}\left(\frac{\left(\frac{T-h}{T}\right)^{\alpha_l}-1}{\alpha_l}\right). \tag{15}$$

[11] See Capéau and Decoster (2016) for a more detailed treatment of the identification properties of the RURO model.

For couples, an interaction term between spouses' leisure is added:

$$
\begin{aligned}
H_2(c, h_f, h_m) \quad = \quad & \beta_c \frac{c^{\alpha_c}-1}{\alpha_c} + \boldsymbol{\beta}'_{lf}\mathbf{r}_f \left(\frac{\left(\frac{T-h_f}{T}\right)^{\alpha_{lf}}-1}{\alpha_{lf}} \right) + \boldsymbol{\beta}'_{lm}\mathbf{r}_m \left(\frac{\left(\frac{T-h_m}{T}\right)^{\alpha_{lm}}-1}{\alpha_{lm}} \right) \\
& + \beta_{l_f l_m} \left(\frac{\left(\frac{T-h_f}{T}\right)^{\alpha_{lf}}-1}{\alpha_{lf}} \right) \cdot \left(\frac{\left(\frac{T-h_m}{T}\right)^{\alpha_{lm}}-1}{\alpha_{lm}} \right).
\end{aligned}
\tag{16}
$$

The use of the RURO model to simulate the impact of policy changes widens the scope of the analysis in two directions, both not readily available in the standard Random Utility Model (RUM) framework. First, the structural specification above allows us to integrate labour demand–side effects into the analysis. The intricacies of integrating labour demand side effects into models of labour supply have been established by Peichl and Siegloch (2012), and specific issues of this link related to the RURO model have been addressed by Colombino (2013). In Appendix A.5 we explain in detail how we have translated a decrease in employer social security contributions into changes in the choice set for individuals, mediated by changes in the relative intensity of market opportunities, θ, and by a change in the first moment of the gross wage distribution, $g_1(\cdot)$. Second, the job choice model, in which unobserved characteristics of the job are present in the structural specification, allows us to capture behavioural reactions that cannot be simulated within the standard framework. In the RURO model e.g., jobs with lower wages but with more attractive unobserved non–pecuniary characteristics might become more attractive after a reform in the tax–benefit schedule that subsidises low wage jobs.

3 Measuring budgetary effects and welfare evaluation

To examine the budgetary effects of a joint tax reform in more detail, we propose a decomposition that allows us to separate behavioural from non–behavioural effects. We also introduce a nonparametric welfare measure based on the equivalent variation to study the distributional impact of the reform. In addition, we present an analysis that incorporates the welfare cost of changes in labour time. Throughout, we indicate pre–reform variables by a subscript 0, while post–reform variables are subscripted by 1. Recall that we denote the chosen job by (w^*, h^*).

3.1 Budgetary effects

We decompose the budgetary effects of tax reforms into a first and second order component. The former measures the impact of the tax reform on government revenues when there is no change in individual behaviour. That is, individuals are allowed to change neither their bundle of commodities nor their job choice. The second order component, by contrast, only collects the change in revenues that can be ascribed to changes in individual behaviour.

3.1.1 Revenues from indirect taxation

We denote expenditures on good i in pre– and post–reform situation (denoted by subscript $j = 0, 1$) by $e_{j,i}$, and these expenditures are measured at consumer prices $\mathbf{q}$, inclusive of all indirect taxes. As we assume general equilibrium effects on producer prices $\mathbf{p}$ to be absent, these can be treated as fixed. Therefore, consumed quantities $\mathbf{x}$ can be measured in terms of these prices. That is, consumption of good i in situation j, $x_{j,i}$, is measured by the value in euro's, when valued at producer price p_i (independent of j). Indirect taxes on good i in situation j then equal:

$$IT_{j,i} = e_{j,i} - x_{j,i}, \qquad j = 0, 1,\ i = 1, \ldots, n. \tag{17}$$

Thus, indirect taxes $\mathbf{t}_j$ ($j = 0, 1$) can easily be defined as *ad valorem* rates:

$$t_{j,i} = \frac{e_{j,i} - x_{j,i}}{x_{j,i}} \qquad j = 0, 1,\ i = 1, \ldots, n. \tag{18}$$

Pre– and post–reform government indirect tax revenues can be calculated as

$$IT_j = \mathbf{t}_j' \mathbf{x}_j, \qquad j = 0, 1. \tag{19}$$

We obtain the pre–reform quantity of a particular good i, $x_{0,i}$, from data on consumer expenditures (expressed in terms of consumer prices), by dividing this amount by $1 + t_{0,i}$. To recover post–reform quantities, we first simulate the new expenditures. Using the Cobb–Douglas assumption, these amount to

$$e_{1,i} = \omega_i y_1, \qquad i = 1, \ldots, n, \tag{20}$$

where y_1 is the disposable income stemming from the post–reform job choice, and ω_i is the budget share, which is a parameter in the Cobb–Douglas case, and therefore kept constant when simulating the reform. Post–reform quantities are then easily obtained as:[12]

$$x_{1,i} = \frac{e_{1,i}}{1 + t_{1,i}}. \tag{21}$$

The indirect tax rate $\mathbf{t}$ includes all taxes which cause a wedge between producer and consumer prices: value added taxes, excises, and *ad valorem* taxes.[13] The change in revenue from indirect taxes, $IT_1 - IT_0$, can then be decomposed into a first and second order effect

[12] In practice, one can calculate $T_j, j = 0, 1$, immediately from $e_{j,i}, i = 1, \ldots, n; j = 0, 1$ as follows: $T_j = \sum_{i=1}^{n} \frac{t_{j,i} e_{j,i}}{1+t_{j,i}}, j = 0, 1$.

[13] An *ad valorem* tax is a tax which is expressed in terms of the final consumer price q.

as follows:

$$IT_1 - IT_0 = \underbrace{\mathbf{x_0}'\left(\mathbf{t}_1 - \mathbf{t}_0\right)}_{\text{first order effect}} + \underbrace{\mathbf{t}_1'(\mathbf{x}_1 - \mathbf{x}_0)}_{\text{second order effect}} . \tag{22}$$

In this equation, the first order effect measures the change in government tax revenue that would arise if individuals could not alter their consumption bundle during a reform. The residual part embodies the second order effect and captures the change in revenue that is due to individuals' altered consumption behaviour — even with constant budget shares, quantities do change — and job choice.

3.1.2 Revenues from direct taxes and social security contributions

We define government revenues from personal income taxes and social security contributions as the difference between employer labour costs g and employee disposable income y:

$$DT_j = g_j - y_j, \quad j = 0, 1. \tag{23}$$

This definition of direct tax revenues is broader than revenue from personal income taxes. It also consists of social security contributions, paid by both employee and employer, and it is net of benefits paid. Employers' labour cost can be written as

$$g_j = (1 + \sigma_{j,er})\, w_j^* h_j^*, \quad j = 0, 1, \tag{24}$$

in which $\sigma_{j,er}$ denotes the rate of employers' social security contributions, expressed in terms of gross earnings. The latter are inclusive of employee's social security contributions. The change in direct tax revenues can then also be decomposed into a first and second order effect:

$$\begin{aligned} DT_1 - DT_0 = & \underbrace{(\sigma_{1,er} - \sigma_{0,er})\,(w_0^* h_0^*) - (f_1\,(w_0^*, h_0^*; M, \mathbf{z}) - f_0\,(w_0^*, h_0^*; M, \mathbf{z}))}_{\text{first order effect}} \\ & + \underbrace{(1 + \sigma_{1,er})\,(w_1^* h_1^* - w_0^* h_0^*) + (f_1\,(w_0^*, h_0^*; M, \mathbf{z}) - f_1\,(w_1^*, h_1^*; M, \mathbf{z}))}_{\text{second order effect}} . \end{aligned} \tag{25}$$

In this equation $f_0(\cdot)$ and $f_1(\cdot)$ embody respectively the pre– and post–reform schedules for personal income taxes, social security contributions paid by the employee, and benefits received. That is, these functions map gross labour income into disposable income, taking into account other income M, and characteristics $\mathbf{z}$. As before, the first order effect assumes there is no change in individuals' behaviour. That is, the post–reform job choice is identical to the pre–reform (w_0^*, h_0^*), simulated in the baseline.

3.2 Welfare analysis

In this section, we first propose a nonparametric monetary metric based on the equivalent variation to approximate the change in individual welfare induced by the joint tax reform, neglecting the welfare cost of a change in labour time (Section 3.2.1). This allows us to assess the reform's distributional impact in terms of welfare, and enables us to distinguish those who win from those who lose. This money metric operates at the level of the subutility function $u(\cdot)$ in Equation (2). Therefore, any income change triggered by the tax reform — either directly or indirectly through possible changes in labour supply or job choice — is considered as a lump sum welfare effect. In Section 3.2.2, we present a measure which also incorporates the welfare cost of changes in labour time.

3.2.1 Welfare effects of changes in income and prices

The equivalent variation is implicitly defined as the amount of money an individual would have to forego in the baseline — that is, at pre–reform prices $\mathbf{q}_0$ and pre–reform income y_0 — in order to be indifferent between the pre–reform and the post–reform situation:

$$v(\mathbf{q}_1, y_1) = v(\mathbf{q}_0, y_0 - EV). \tag{26}$$

This equivalent variation can be interpreted as a welfare cost measure: when utility in the post–reform situation is lower than in the baseline, it measures how much income a person needs to give up in the baseline to be equally well–off. In order to obtain a measure of the change in welfare, ΔW_{EV} say, it suffices to simply reverse its sign:

$$\Delta W_{EV} = -EV. \tag{27}$$

Using the properties of the expenditure function, $e(\cdot)$, we can derive the following explicit formulation for this change in individual welfare:

$$\Delta W_{EV} = e(\mathbf{q}_0, v(\mathbf{q}_1, y_1)) - e(\mathbf{q}_0, v(\mathbf{q}_0, y_0)). \tag{28}$$

Adding and substracting y_1 yields

$$\Delta W_{EV} = \underbrace{[y_1 - y_0]}_{\text{change in disposable income}} - \underbrace{[e(\mathbf{q}_1, v(\mathbf{q}_1, y_1)) - e(\mathbf{q}_0, v(\mathbf{q}_1, y_1))]}_{\text{welfare cost of a pure price change}}, \tag{29}$$

which decomposes the individual welfare effect into the change in disposable income and the welfare effect of a pure price change.

As the expenditure function associated with our Cobb–Douglas subutility function (8) is not necessarily very realistic, but rather assumed to keep the interaction with other model parts tractable, one might prefer to look for an approximation of equation (29) which is more universally valid. Thereto we start from the last term in expression (29), and note that the utility level $v(\mathbf{q}_1, y_1)$ can be obtained at the bundle $\mathbf{x}_1$, even though this bundle would not necessarily be the cost minimizing bundle at pre–reform prices $\mathbf{q}_0$, and this irrespective of the specific form of the underlying preferences represented by $v(\mathbf{q}, y)$. Consequently,

$$e(\mathbf{q}_0, v(\mathbf{q}_1, y_1)) \leq \mathbf{q}_0'\mathbf{x}_1, \tag{30}$$

which, using equation (29), yields the following upper bound

$$\Delta W_{EV} \leq [y_1 - y_0] - (\mathbf{q}_1' - \mathbf{q}_0')\mathbf{x}_1. \tag{31}$$

Given the assumption of fixed producer prices, we can rewrite this as

$$\Delta W_{EV} \leq [y_1 - y_0] - (\mathbf{t}_1' - \mathbf{t}_0')\mathbf{x}_1, \tag{32}$$

where the upper bound of the individual welfare effect on the RHS is calculated as the change in disposable income, from which the change in indirect taxes paid, calculated at the post reform quantities, is subtracted.[14] Since this upper bound is universally valid for any underlying preference ordering, and in this sense more robust, the effects of the reform will usually be worse than (or at most equal to) this upper bound, whatever the underlying preferences really are.

Now suppose these preferences are indeed Cobb–Douglas. Then an exact measure for $\Delta\, W_{EV}$ can be derived as follows. Inverting equation (10) and using $\mathbf{q}_r$ to denote a reference price vector, the expenditure function $e(\mathbf{q}_r, v(\mathbf{q}, y))$ for the Cobb-Douglas case turns out to be equal to $y \cdot \frac{Q(\mathbf{q}_r)}{Q(\mathbf{q})}$. Using the pre–reform prices $\mathbf{q}_0$ as reference prices, one gets:

$$\Delta W_{EV} = \frac{Q(\mathbf{q_0})}{Q(\mathbf{q_1})} y_1 - y_0. \tag{34}$$

This shows that — for Cobb-Douglas preferences — the difference between real incomes, i.e. $\frac{y_1}{Q(\mathbf{q_1})} - \frac{y_0}{Q(\mathbf{q_0})}$, can be considered to be a rescaling of the change in welfare ΔW_{EV}, viz. by dividing this number by $Q(\mathbf{q_0})$, which boils down to using a different cardinalisation of the underlying preferences to measure the welfare effect.

3.2.2 The welfare cost of changes in labour time

The equivalent variation based measure of welfare change in equation (29) starts from the indirect utility function $v(\mathbf{q}, y)$ of the second stage of the decision process. Besides the change

[14] Using post reform prices $\mathbf{q}_1$ as reference price vector in the construction of the money metric utility, one obtains a lower bound for the welfare change based on the compensating variation:

$$\Delta W_{CV} \geq [y_1 - y_0] - (\mathbf{t}_1' - \mathbf{t}_0')\mathbf{x}_0, \tag{33}$$

where the indirect tax effect is now calculated at the initial quantities $\mathbf{x}_0$.

in consumer prices, this concept also picks up the change in disposable income, caused by the tax change and the eventual change in job choice triggered by it. However, what is not yet included in the concept of Equation (29), is the welfare cost of changing labour time: that is, the changes in disposable income come without a cost. As far as these changes are merely a consequence of changes in the tax–benefit system, such an assumption might be valid. But in a behavioural model, the changes in disposable income are also partially induced by a second order effect, i.e. the change in labour time. This behavioural reaction bears a welfare cost.

We refrain from constructing a measure of welfare change based on the equivalent variation at the level of the overall utility function $\Omega(\cdot)$. Because of interindividual differences in the wage rates, every individual would then face a distinct price vector. This has unattractive properties for making interpersonal welfare comparisons. For example, two persons with equal preferences and who would each obtain a bundle such that they are indifferent among each other's situation, could nevertheless be considered as not equally well off if they faced different wages.

We therefore revert to a class of measures proposed by Fleurbaey (2006), which explicitly fix a reference wage across individuals, to tackle this problem in the context of the leisure consumption trade–off.[15] The idea is to compare individuals not on the basis of utilities, but in terms of their available opportunities. The valuation of available opportunities can however not be determined in purely objective terms (such as disposable income), but does depend on a person's preferences. Persons with more intense preferences for leisure do not necessarily gain as much from an increased remuneration of labour time, as those who rather like to work.

One such measure is *equivalent consumption*, henceforth denoted by W_{EC}.[16] The idea is to look at the real budget (i.e. consumption) an individual would need in order to be equally well–off as in the situation to be assessed, in case she would not work at all. For a given situation with positive labour time, a person with more intense preferences for leisure (represented by R^s in Figure 1) is considered to be worse off by this measure than someone who rather likes to work (represented by R^f in the same figure). Note that for those who do not work, equivalent consumption equals real disposable income.

Since we have estimated preferences over consumption and labour (the function $H(c, h)$ in Equation 11), we can calculate the value of the equivalent consumption measure, say W_{EC_j}

[15] See Decoster and Haan (2015) and Bargain et al. (2013) for empirical applications.

[16] It is called the *rente* criterion in Fleurbaey (2006), Decoster and Haan (2015) and Bargain et al (2013). We prefer the term *equivalent consumption* to safeguard its relation with the egalitarian equivalent solution of Pazner and Schmeidler (1978). 'Equivalent' refers to an equivalent budget, needed to be equally well–off in a counterfactual as in the actual situation. It is therefore also related to the use of 'equivalent' in 'equivalent variation'. It should evidently not be confused with 'equivalised' income or expenditures, which refer to converting these household level variables into comparable magnitudes at the level of individual members of the household. This latter adaptation is done through so–called 'equivalence scales'.

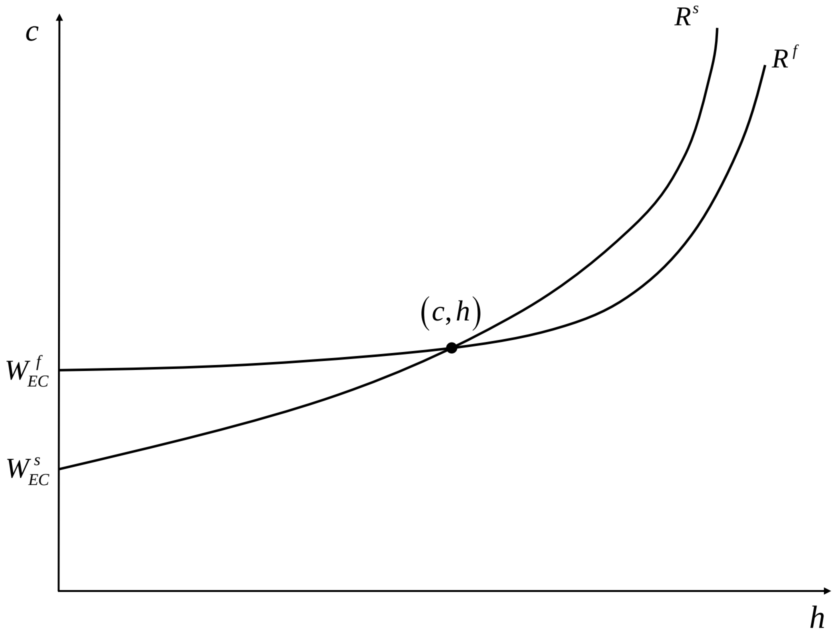


Figure 1: Equivalent consumption and intensity of preferences for leisure

$(j = 0, 1)$, for those individuals that are included in the job choice model, as follows:

$$H\left(W_{EC_j}, 0\right) = H\left(c_j^*, h_j^*\right), \qquad j = 0, 1, \tag{35}$$

where (w_j^*, h_j^*) is the optimal choice in situation j, with $c_j^* = f\left(w_j^*, h_j^*; M, \mathbf{z}\right)/Q\left(\mathbf{q}_j\right)$. The change in welfare from pre– to post–reform is then

$$\Delta\, W_{EC} = W_{EC_1} - W_{EC_0}. \tag{36}$$

To assess the importance of accounting for the welfare cost (gain) of increased (decreased) labour supply within the same class of welfare measures, we also calculate the equivalent consumption for the fictitious reference point (c_1^*, h_0^*). In this point, consumption is set at the post–reform level, whereas hours worked are kept fixed at the pre–reform level. Formally, this welfare measure, say $\overline{W_{EC}}$, is defined implicitly as

$$H(\overline{W_{EC}}, 0) = H(c_1^*, h_0^*). \tag{37}$$

The change in welfare from pre–reform to this reference point in terms of the equivalent consumption amounts to

$$\overline{\Delta W_{EC}} = \overline{W_{EC}} - W_{EC_0}. \tag{38}$$

From Figure 2a it is easily seen that when a person increases labour supply (i.e. $h_1^* > h_0^*$), $\overline{\Delta W_{EC}}$ will exceed $\Delta\, W_{EC}$, as in the former the gain in consumption is acquired effortlessly. Alternatively, when labour supply decreases (i.e. $h_1^* < h_0^*$), the opposite is true and now

ΔW_{EC} will exceed $\overline{\Delta W_{EC}}$ (Figure 2b). Note that when there is no change in labour supply (i.e. $h_1^* = h_0^*$) both measures coincide.

The main advantage of assessing the impact of labour time within the same class of welfare measures, is that one avoids making comparisons across measures with distinct underlying cardinal properties. Section A.6 in the Appendix discusses in more detail why the EV– and EC–based measures are not directly comparable.

4 Data and simulated reforms

We implement our approach on Belgian data. We argued that the income variable in the RURO job choice model needs to be specified in real terms. To arrive at this variable we used the tax benefit microsimulation model EUROMOD (Sutherland and Figari, 2013).[17]

EUROMOD runs on the Statistics on Income and Living Conditions (SILC) survey, which is a micro–level dataset that contains detailed information on income, poverty, social exclusion and other living conditions. For Belgium, the survey's reference population includes all private households and their current members residing in the country. Individuals living in collective households, such as hospitals, youth institutions, and old peoples homes are excluded from the reference population. All of our calculations are performed on the Belgian SILC 2015, which contains 14,145 individuals who live in 6,006 households.

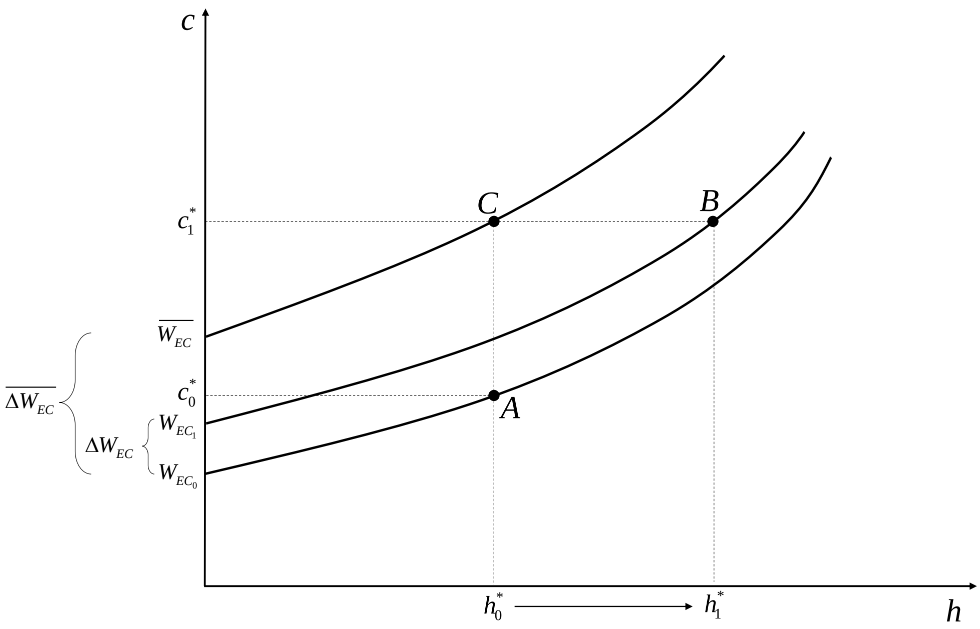
SILC does not contain data on expenditures, which is required to calculate indirect taxes paid by the households and to construct the household specific Divisia price indices, needed to estimate the RURO model. In order to impute budget shares in the SILC data, we estimated parametric Engel curves on the Belgian Household Budget Survey (HBS) 2014.[18] It was shown by Blundell and Walker (1982) and Browning and Meghir (1991) that estimates of commodity demand functions can be severely biased when the assumption of separability between budget allocation and choice of leisure time is erroneously maintained. We therefore included dummies for the household's highest income earner labour market status (working, unemployed, or pensioner), and the number of employed persons in the household, as covariates in these Engel curves.[19] De Agostini et al. (2017) discuss in detail a strategy to impute expenditures, and hence indirect taxes, in SILC from estimated budget shares on the HBS. This strategy is now standardised and implemented in EUROMOD as part of the Indirect Tax Tool.

The RURO model was estimated on a subsample of the SILC data that only contains those households in which the reference person and their partner, if any, are available for the

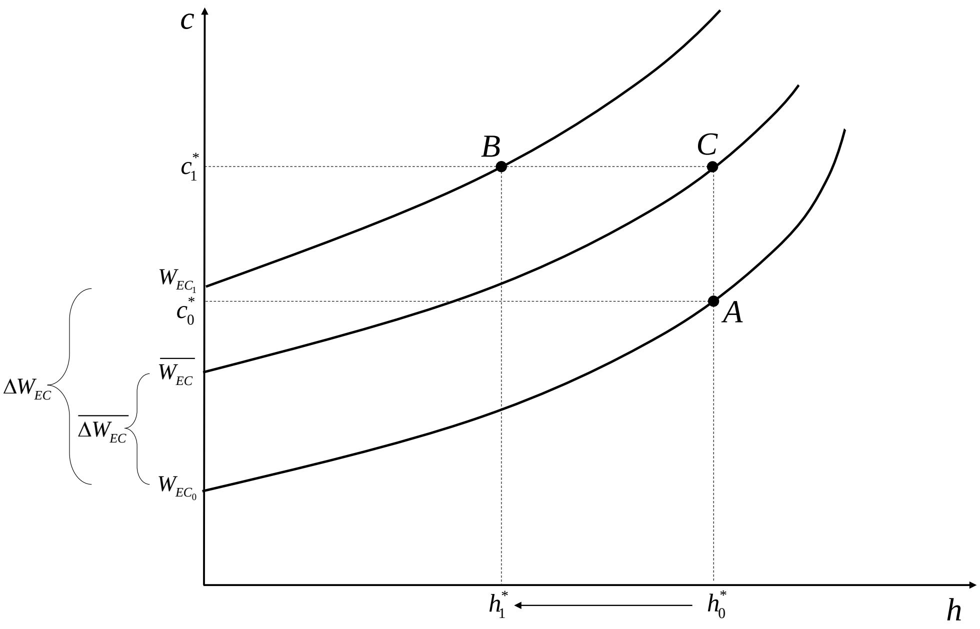
[17]EUROMOD covers the personal income tax code of 27 EU countries for several policy years, and allows us to simulate tax reforms. Recently an Indirect Tax Tool was added for 10 countries, one of which was Belgium (see De Agostini et al., 2017).

[18]The properties of parametric expenditure imputations are studied in Decoster et al. (2007) and Savage (2017).

[19]The HBS contains only limited information on labour market status. In particular, labour time is not available.



(a) Increase in labour supply



(b) Decrease in labour supply

Figure 2: Change in welfare according to equivalent consumption

labour market. Appendix A.1 contains more details on the composition of this subsample. In Section 2.2 we presented our functional forms. Appendix A.2 presents the estimated model parameters, the simulated aggregate wage elasticities of labour supply, and the model fit.[20] The elasticities of our model are broadly in line with the abundant micro–econometric estimates for other countries (see Bargain et al., 2014 and Mastrogiacomo et al., 2017 for recent overviews for several European countries and the US). First, the total own wage elasticity of 0.49 for single females, 0.52 for single males, 0.49 for females in couples and 0.33 for males in couples is mainly determined by the participation elasticities (see Table A.3). Second, the elasticities are declining with the level of the gross wage rate (see Table A.4 and Table A.5). Third, in couples we find substantial negative cross wage elasticities (-0.26 for females and -0.16 for males), which are here mainly driven by reactions at the intensive margin.

After estimating the model parts, we use it to simulate the effects of some scenarios of tax reforms inspired by, or serving as alternatives to a tax shift enacted in Belgium from 2016 onwards. More details on the Belgian tax shift can be found in Appendix A.4.

The (household specific) imputed budget shares from the estimated Engel curves serve as parameters of the Cobb–Douglas preferences in our two–stage budgeting approach. They can thus be used to simulate the effects of indirect tax reforms on expenditures, indirect taxes paid, and the Divisia price indices. The weak separability assumption we exploit in the simulation stage implies that we do not allow the labour market status covariates of the Engel curves to change when labour market status alters as a consequence of a tax reform.

Next, using the new Divisia indices following from the indirect tax reform, and the possibility to simulate direct tax reforms with EUROMOD, the impact of a reform on households' labour supply decisions can be simulated by our estimated RURO model. The simulation procedure is explained more in detail in Appendix A.3. Note that households who do not belong to the subsample on which RURO was estimated, can alter their behaviour only through the expenditures margin in our simulations.

As explained before, the RURO model allows us to separate opportunities from preferences, which renders it possible to incorporate demand–side responses in the analysis. In one of our simulations, we demonstrate this feature by incorporating the demand–side feedback from a decrease in employer social security contributions. This feedback effect is mediated through two channels: the relative measure of market versus non–market opportunities, θ, and the first moment of the gross wage distribution $g_1(\cdot)$. Both these channels impact households' choice sets. Appendix A.5 shows how this impact can be simulated in a tractable way.

The following four tax reform scenarios were simulated and compared with respect to the

[20]In our job choice model individuals are not characterised by one specific wage, rather by the distribution of wage offers determining the choice set from which they make their preferred choice. Therefore we obtain elasticities by shifting the whole distribution of wage offers to the right, increasing the first moment of the individual specific wage distribution by 10%.

pre–reform baseline scenario.

TS1 Scenario 1 constitutes the PIT and SSC components of the wider reform as decided by the Belgian government (see Appendix A.4 for a more detailed description of the policy changes).

TS2 Scenario 2 augments TS1 with the reform in indirect taxes, which is a limited increase in VAT, and some excise rates (see again Appendix A.4 for a more detailed description). This scenario closely resembles the Belgian tax shift.

TS3 Scenario 3 augments TS1 with a much larger increase in VAT of four percentage points. As Scenario TS2 is far from revenue neutral, we try to close at least part of this financing gap in this way.

TS4 Scenario 4 introduces, on top of Scenario TS2, additional effects from the demand side of the labour market by translating the decrease in the social security contributions of the employer into an increase in the job offers and a shift of the wage offer distribution (see Appendix A.5 for a more detailed description of how we implemented this change).

Note that in none of these counterfactual scenarios we mirror the Belgian tax shift of 2016 exactly. This has two main reasons. Firstly, in the baseline scenario we only simulate about half of all indirect taxes paid. This can partly be explained by the under–reporting of alcohol and tobacco consumption, but is mainly due to our inability to observe VAT and excise–payments from transactions between firms. As a result, in modelling TS3, we define the revenue gap to be financed *within* the model, using the rate of partial coverage in the baseline to determine the additional VAT–revenues to be collected. This resulted in an increase in all VAT rates with an additional four percentage points. Secondly, EUROMOD is not at present able to simulate capital income and corporate income taxes. Even though the Belgian tax shift itself aimed at — be it modest — revenue increases from these sources, our simulations currently envisage only the indirect tax increase and cost recovery effects from increased employment as sources of financing the loss in revenue.

5 Results

5.1 Budgetary and employment effects

As discussed in Section 3.1, we decompose the total budgetary effect of a tax shift into two distinct parts: a *first order effect* (or 'impact' effect) and a *second order effect*. The former measures the effect when households could not alter their consumption bundle and their labour supply with respect to the baseline scenario. The latter, by contrast, exclusively captures the impact of these behavioural responses. Changes in labour supply are predicted by the RURO model and we assume that households alter their consumption according to the constant expenditure shares assumption, i.e. assuming Cobb–Douglas preferences. Note that employment effects are by definition always second order effects. Unless otherwise stated, results are grossed up at population level by means of statistical weights.

The upper part of Table 1 displays for each of the four reform scenarios absolute changes with respect to the simulated baseline, whereas the bottom part contains the corresponding percentage changes. The first four rows of each part show the employment effect, the effect on household disposable income, both including and net of indirect taxes, and the effect on government revenue. We start with reform TS1, which only captures the effect of the reduction in personal income taxes and social security contributions. The impact effect of this reform is a net loss in government revenue of €9.9bn (or 2.5% of GDP in 2014), a considerable bill for the treasury coffers.[21] This bill is mainly devoted to the lowering of personal income taxes (€−6.2bn or a decrease of 12.1%) and social security contributions of the employer (€−3.6bn, also a decrease by nearly 12%). Contrary to the social security contributions paid by the employer, the personal income tax reductions and also the limited reduction of social security contributions paid by the employee are immediately reflected in an increase in disposable income of households: it increases with €111 per month per household, or an increase of 3.9%.

The impact effects of simulation TS2, where we add the increase in excises and VAT to simulation TS1 and which thus closely resembles the Belgian tax shift, reveal the largely unfunded character of this reform.[22] This leads to additional revenues of €951 million, pushing down the initial cost from €9.9bn to €8.9bn.[23] This limited character of the revenue raising elements of the tax reform is also revealed in the third line in which we show the impact on disposable income minus indirect taxes paid at the household level. The impact effect goes down from €111 to €94 per month per household when comparing TS1 with TS2, implying an additional indirect tax bill of €17 per month. In percentage terms, the indirect tax part of the tax reform erodes the first order effect in terms of disposable income minus

[21]Nominal GDP in 2014 equals approximately €400bn. This estimate of the impact cost comes close to the most recent estimate by the National Bank of Belgium, which reports a cost of €9.341bn, see NBB (2017, p.2). Yet this resemblance of our total estimated first round cost to the estimate of the National Bank hides non negligible differences in the underlying decomposition into first and second round costs.

[22]As mentioned above, other tax increases complemented the increase in VAT and excises, such as an increase in taxes on income from financial assets. However, the lack of information on capital income prevents us from integrating this part of the tax shift into the simulations. NBB (2017) estimates that the increase in taxation of income from capital, over the whole period, will amount to €661 million (of which €276 million comes from the increase of the withholding tax on income from financial assets from 25% to 27%), and the increase of the corporate income tax to €758 million (in which an increased return of fighting tax fraud and evasion accounts for €150 million).

[23]Without additional incidence assumptions, our household model is unable to pick up the additional revenue from VAT and excises paid by the production sector. Certainly for the excise increases, this is a serious flaw, since additional excise revenue is estimated by NBB (2017) to amount to €1.832 billion. As explained before, this is the main reason why we do not aim at *full* revenue neutrality when simulating a revenue neutral reform.

Table 1: Employment and revenue effects of the four simulated reforms

	baseline	TS1			TS2			TS3			TS4		
		1st order	2nd order	total	1st order	2nd order	total	1st order	2nd order	total	1st order	2nd order	total
							absolute change w.r.t. baseline level						
FTEs (1000 units)	2,607	-	70.6	70.6	-	65.2	65.2	-	43.8	43.8	-	91.9	91.9
disp. inc. (€/mo./hh)	2,846	111	4	115	111	2	113	111	-12	99	111	10	121
disp. inc. - ind. tax (€/mo.)	2,539	111	-9	102	94	-6	88	31	-8	23	94	1	95
net gov. rev. (mio €)	54,943	-9,863	915	-8,948	-8,912	394	-8,518	-5,250	-1,596	-6,845	-8,906	1,340	-7,567
employer SSC (mio €)	30,443	-3,592	92	-3,501	-3,592	29	-3,564	-3,592	-305	-3,897	-3,592	287	-3,305
employee SSC (mio €)	20,823	-111	-18	-129	-111	-58	-169	-111	-286	-397	-111	66	-45
pers. inc. tax (mio €)	51,017	-6,171	-558	-6,729	-6,171	-660	-6,832	-6,171	-1,348	-7,519	-6,171	-349	-6,521
benefits (mio €)	65,208	-12	-642	-653	-12	-611	-623	-12	-548	-560	-12	-781	-793
ind. tax (mio €)	17,868	0	757	757	951	471	1,423	4,613	-205	4,409	957	555	1,511
							relative change w.r.t. baseline level						
FTEs		-	2.7	2.7	-	2.5	2.5	-	1.7	1.7	-	3.5	3.5
disp. inc.		3.9	0.2	4.0	3.9	0.1	4.0	3.9	-0.4	3.5	3.9	0.4	4.2
disp. inc. - ind. tax		4.4	-0.3	4.0	3.7	-0.2	3.5	1.2	-0.3	0.9	3.7	0.0	3.7
net gov. rev.		-18.0	1.7	-16.3	-16.3	0.7	-15.6	-9.8	-3.0	-12.7	-16.1	2.4	-13.6
employer SSC		-11.8	0.3	-11.5	-11.8	0.1	-11.7	-11.8	-1.0	-12.8	-11.8	0.9	-10.9
employee SSC		-0.5	-0.1	-0.6	-0.5	-0.3	-0.8	-0.5	-1.4	-1.9	-0.5	0.3	-0.2
pers. inc. tax		-12.1	-1.1	-13.2	-12.1	-1.3	-13.4	-12.1	-2.6	-14.7	-12.1	-0.7	-12.8
benefits		0.0	-1.0	-1.0	0.0	-0.9	-1.0	0.0	-0.8	-0.9	0.0	-1.2	-1.2
ind. tax		0.0	4.2	4.2	5.4	2.7	8.0	27.7	-1.2	26.5	5.2	3.0	8.2

All effects are with respect to the baseline in the first column of the table. Baseline levels are obtained by simulating the pre–reform scenario. Monthly disposable income is not equivalised.
TS1 is the scenario in which only PIT and SSC-changes are simulated, exclusive of behavioural effects linked to the demand side of the labour market;
TS2 is TS1, augmented with a (limited) change in indirect taxes;
TS3 is TS1, augmented with the much larger change in indirect taxes;
TS4 is TS2, augmented with behavioural effects coming from the demand side of the labour market as explained in Appendix A.5.

indirect taxes from 4.4% to 3.7%.

As this tax reform is largely unfunded, an assessment of the potential cost recovery through additional employment is important. The first line of Table 1 shows the additional employment, expressed in full time equivalents (FTEs), triggered by the tax reform.[24] In simulation TS1, which discards the increase in indirect taxes, the rise in net wages triggered by the reform induces an additional employment of 70600 FTEs, or an increase of 2.7% of baseline employment. Note that this result bears on the full horizon of the measures, rolled out progressively from 2016 to 2020. Not unexpectedly, the slight diminution of the increase in disposable income in TS2 comes with an erosion of the employment effect to 65,200 FTEs. More importantly, when we amplify the VAT increase to move into the direction of more revenue neutrality (TS3), almost one third of the newly created FTEs are lost: additional employment falls back to 43,800 FTEs or 1.7%.[25]

The cost recovery effect of this additional employment is however surprisingly small. In simulation TS1, net additional revenues amount to €915 million (in the column '2nd order'), or a poor 9.3% of the initial bill of the reform. This cost recovery is the net effect of a decrease in benefits to be paid (−€642 million), an increase in social security contributions paid by the employer (+€92 million), an increase in indirect tax payments following the increase in disposable income (+€757 million), but also a surprising decrease in personal income taxes (-558 million). Contrary to what is often raised in the public debate, viz. that a bill of €6,171 million in the personal income tax will finance itself through cost recovery, the final bill exceeds the initial one, at least in the personal income tax. In Section 5.2 below, we will show that this can be attributed to a negative income effect at the top of the income distribution. The income of new entrants is rather moderate, so they generate few additional tax revenues. At the same time, the tax reform allows persons higher up in the income distribution to work less and to afford jobs with somewhat lower gross wages. Since the tax schedule is progressive, this effect excavates possible cost recovery effects.

The rightmost columns of Table 1 show the simulation results which exploit the richness of the RURO model by simulating additional job opportunities. In particular, simulation TS4 translates the lowering of social security contributions of the employer into an increased demand for labour. On top of the increase in net wages coming from lower social security contributions of the employee and lower personal income taxes, this labour demand effect leads to an additional increase in equilibrium wages. This simulation produces the largest positive employment effect: the tax shift would trigger about 91,900 additional FTEs (or

[24]The behavioural model simulates labour supply in hours per week. To transform the additional labour supply into FTEs, we divide by the standard of 38 hours per week.

[25]Simulation TS2 is closest to the actual Belgian tax shift. Our estimate of 65,200 additional FTEs in that case can be compared to the estimate of the macro–model in NBB (2017), which amounts to 52,100 (NBB 2017, p.5). This estimate of 52,100 additional FTEs is the net result of an increase in employment of 85,900 due to increased purchasing power and lower labour costs, and a decrease of 33,800 units, mainly due to higher indirect taxes.

3.5% of employment in the baseline).[26] Yet even this stronger employment effect only leads to a cost recovery of €1,340 million on an initial bill of €8,9 billion.[27] This cost recovery is still only 15% of the initial cost of the reform.

Summing up these aggregate results, we conclude that the actual tax shift does indeed create additional employment, but that the numbers are modest overall. Moreover, removing the largely unfunded character of the tax reform proposal, by inserting some kind of revenue neutrality, further erodes the employment effect. The cost recovery effects are modest too. To deepen our understanding of these aggregate effects, we first look at the employment effects in more detail.

5.2 Labour market effects

Across the different variants of the tax shift simulations, there is only limited variation in the heterogeneity of labour market effects across the population. In Table 2 we therefore only summarise the labour market effects of simulation TS3.[28] The table presents effects of the reform on job choice for the population of individuals included for analysis in our job choice model (Section 2.2), across deciles of the gross wage distribution. To order the individuals, we use the gross wage rate observed in the job choice of the baseline for those individuals who work in the baseline. For the non–participating individuals we impute a gross wage based on their observable characteristics and the wage offer equation estimated in the RURO model. Columns (1) to (4) describe effects for the whole RURO population, both on the extensive (working or not, labelled here as 'participation') and on the intensive (number of hours worked per week) margin. Gross wage being a characteristic of the job, changes in job choice may also trigger an additional effect on earnings, which are the product of gross wage and hours worked. This is presented in the right part of Table 2 (columns (5) to (10)). In that part of the table we limit ourselves to the population being employed in the baseline. The reason is that wages of non–working individuals are not observed. Hence we cannot calculate the percentage change in wages and earnings for these individuals.[29]

Column (2) shows that the increase in employment of 43,800 FTE's, discussed in Section 5.1, mainly results from an increase in participation of individuals characterised by a low gross wage. The increase of participation is most outspoken in the bottom decile of the gross wage distribution (an increase of 8.8 percentage points from a baseline level of 27.4%), and remains above the average up to the fourth decile of the gross wage distribution. Since everybody is already working in the top three deciles of the gross wage distribution, it is no

[26] As mentioned in note 25, the effect of increased purchasing power in combination with lower labour costs led the NBB (2017) to estimate increased employment to be 85,900 units; the same order of magnitude as our estimate.

[27] The small difference in the first order effect of indirect taxes between TS2 and TS4 is caused by the ITT's internal calibration procedure.

[28] Results for the other simulations are available upon request.

[29] It is true that we have imputed a wage for these non–working individuals to assign them a place in the decile ordering based on gross wages, but we choose to limit the use of this imputation for that purpose only.

Table 2: Labour market effects across deciles of gross wages for revenue neutral simulation (TS3)

deciles	whole RURO subpopulation				RURO subpopulation working in baseline					
	participation		labour supply		labour supply		gross wage		earnings	
	basel.	Δ	basel.	Δ	basel.	Δ	basel.	Δ	basel.	Δ
	%	% pts	hours/week		hours/week		€/hour	%	€/month	%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	27.4	8.8	10.6	3.0	38.6	0.1	8.8	5.4	1,483	5.2
2	58.7	5.1	21.4	1.8	36.5	0.0	12.2	0.5	1,920	0.6
3	67.1	4.4	24.2	1.6	36.2	0.0	13.9	0.0	2,179	0.0
4	79.2	2.3	29.7	0.7	37.5	-0.2	15.4	-0.4	2,498	-0.8
5	92.7	0.8	34.8	0.1	37.6	-0.2	17.0	-0.6	2,764	-1.1
6	90.4	0.8	33.5	0.1	37.1	-0.3	18.7	-0.9	3,004	-1.6
7	98.0	0.1	37.0	-0.3	37.7	-0.4	20.8	-1.2	3,403	-2.0
8	100.0	-0.2	37.9	-0.5	37.9	-0.5	23.4	-1.8	3,842	-2.8
9	100.0	-0.3	37.1	-0.5	37.1	-0.5	27.3	-2.7	4,389	-3.6
10	100.0	-0.7	37.3	-0.9	37.3	-0.9	40.2	-7.4	6,471	-8.9
all	81.3	2.1	30.3	0.5	37.3	-0.4	21.5	-2.6	3,477	-3.4

The variables are calculated at the individual level for the two subpopulations mentioned in the top row. Each individual is allocated to the decile based on his or her gross wage (with an equal number of persons by decile). The non–working individuals are assigned a gross wage based on the estimated wage equation of the RURO model. For all columns we keep the allocation of individuals across deciles fixed, i.e. the deciles are always based on all individuals, both working and non–working in the baseline. ‘Participation’ in column (1) is calculated as the ratio of the number of persons working in the baseline to the total population of individuals included for analysis in the RURO model. The average number of hours worked per week in column (3) is also calculated for the whole RURO subpopulation, irrespective of whether the individual was working in the baseline or not. Columns (5), (7), and (9) are averages for the RURO subpopulation of individuals who are working in the baseline.

surprise that the effect we find here, if any, can only be negative. For the top 10% of chosen gross wages in the baseline, the participation decreases with 0.7 percentage points. This negative income effect also pops up in column (4) where we find that for the top four deciles of the gross wage distribution, the average number of hours worked decreases. Overall, the average number of hours worked per week increases from 30.3 to 30.8.

The negative income effect of a significant increase in disposable income following a tax reduction does not come as a surprise. It has been documented frequently in other assessments of tax reforms based on modelling behaviour with a standard discrete choice approach (e.g. Blundell et al., 2000). However, the RURO model is uniquely equipped to unveil a potentially more important, additional, ‘income effect’. It shows up in column (8) as a considerable reduction in gross wages, itself the result of the switch by some individuals to the choice of a wage–hours package with lower gross wages than before the reform. On average, gross wages, following from the choice of jobs after the tax reform, are 2.6% lower than the gross wages of the jobs chosen in the baseline. Like the effect on hours worked, this effect is also mainly found in the upper half of the gross wage distribution.[30] Combined with

[30] In Table A.10 in Appendix A.7 we also show the results for an ordering of the individuals based on their equivalised disposable income. The numbers evidently slightly differ, but the conclusions remain the same. Admittedly, the decrease in wages as represented here might be somewhat exaggerated, as the rather small

the decrease in labour supply this trickles down in a significant reduction of the taxable base of both social security contributions and personal income taxes: gross earnings decrease by 3.4%. Since this effect is predominantly found in the upper half of the distribution (earnings go down by 8.9% in the top decile of gross wages, whereas they increase by 5.2% in the bottom decile), this explains the large negative revenue effect in the progressive personal income tax. In a simulation in which we keep wages constant, employment effects are nearly the same as in simulation TS3: participation goes up by 2.3% instead of 2.1%, average hours per week goes up by 0.6 instead of 0.5, and additional employment is now 53,900 FTEs instead of 43,800. But the large drop in revenues from personal income tax (−€1,348 million) melts away to −€166 million. The same holds, to a lesser extent, for revenues from social security contributions: instead of a decrease of €305 million and €286 million for social security contributions paid by employer and employee respectively, the simulation with fixed wages now turns into additional second round revenues of €153 and €58 million. Overall, the revenue loss of €1,596 million in simulation TS3 (see row 4 in the column "TS3 2nd order" of Table 1) turns into additional net revenues of €635 million.

The crux here is not whether we can produce the right amount of cost recovery. But the structural model of job choice unveils a mechanism which might at least partially explain the often disappointing revenue figures following tax reductions in the form of increases of household disposable incomes. Lowering of personal income taxes and social security contributions allows some individuals, and mainly those who before choose jobs with high gross wages, to afford a new job choice which comes with a lower gross wage, but with less hours to work and preferred unobserved characteristics (less stress, lower commuting time, etc.). Using a structural model which allows for this additional behavioural channel shows that a good empirical estimate of this additional 'income' effect is crucial to produce credible revenue predictions. This effect is expected to occur especially where the reform is targeted at making low wage jobs better paid.

5.3 Distributional analysis

A comprehensive distributional analysis should incorporate all three elements of the reform: the change in disposable incomes, mainly driven by lower social security contributions and personal income taxes for people active on the labour market, increases in indirect taxes to be paid, which are also to be borne by non–active persons, and finally the changes in real disposable income and leisure time induced by changes in behaviour. We again limit the discussion to simulation TS3. The results are summarised in Table 3.[31]

group who quit jobs due to the reform, is registered here as facing a 100% wage loss.

[31] We give a summary of the distributional analysis for all simulation variants in Tables A.11 and A.12 in Appendix A.7.

Table 3: Distributional effects for the revenue neutral simulation (TS3)

deciles	changes in €/month and per household						changes in % of baseline disposable income					
	1st order effect			total effect (incl. behav.)			1st order effect			total effect (incl. behav.)		
	(1) disp. inc.	(2) ind. tax	(3) (1)-(2)	(4) disp. inc.	(5) ind. tax	(6) W_{EV}	(7) disp. inc.	(8) ind. tax	(9) (7)-(8)	(10) disp. inc.	(11) ind. tax	(12) W_{EV}
1	8	53	-45	100	57	47	0.6	4.3	-3.7	8.1	4.6	3.8
2	38	62	-24	65	57	5	2.3	3.7	-1.4	3.9	3.4	0.3
3	69	62	7	77	59	15	3.9	3.5	0.4	4.3	3.3	0.9
4	102	70	32	108	69	37	4.7	3.2	1.5	4.9	3.2	1.7
5	117	76	42	118	75	42	4.7	3.0	1.7	4.7	3.0	1.6
6	144	84	59	135	84	49	4.8	2.8	2.0	4.5	2.8	1.6
7	165	89	77	153	89	63	5.0	2.7	2.3	4.6	2.7	1.9
8	173	95	78	146	93	50	4.6	2.5	2.1	3.9	2.5	1.3
9	169	100	68	113	93	12	4.0	2.4	1.6	2.7	2.2	0.3
10	163	116	47	-7	94	-121	2.9	2.1	0.8	-0.1	1.7	-2.2
all households	111	79	31	99	76	20	3.9	2.8	1.1	3.5	2.7	0.7

All effects are with respect to the baseline simulation. The deciles are constructed by ranking all individuals on the basis of their household equivalised disposable income in the baseline. Each decile contains 10% of the total population of individuals. The values in columns (1), (2), (4) and (5) show averages of *non–equivalised* household incomes and indirect taxes, where the averages are calculated over the population of households. In column (3) we subtract the change in indirect taxes paid at pre–reform quantities from the change in disposable income. The welfare effect in column (6) is calculated on the basis of the RHS of equation (32). All percentages in columns (7) to (12) are with respect to baseline disposable income.

Contrary to Section 5.2 above, where we only used the subpopulation modelled in the behavioural labour supply model, we now present results for the whole population. Each decile of Table 3 contains 10% of the number of persons in the population (including children, pensioners, etc.). The position in the ordering is determined by equivalised disposable income of the household the person belongs to. In the left part of the table we show the effect on monthly household disposable income, and on indirect taxes to be paid in euros per month. Column (6) is the upper bound for the welfare change based on the equivalent variation, defined on the RHS of Equation (32). The total effect on disposable income, column (4), minus the change in indirect taxes at constant quantities, column (2), is a lower bound for a measure of welfare change based on the compensating variation (see Equation (33) in note 14). The amounts in the left part of the table are *non–equivalised.* They should be read as the averages of household incomes or of indirect taxes paid by the household, where averages are taken over these households. The right part of Table 3 expresses the changes in percent of disposable income.

The first observation to be inferred from Table 3 is the big impact of the financing through indirect taxes on the distributional assessment of the reform. Column (7) shows that the first order change in disposable income is more or less flat across the middle of the income distribution (deciles 4 to 8), mildly progressive at the top (deciles 8 to 10) and quite regressive at the very bottom, where the gain in relative terms steeply increases across the bottom three deciles. But the increase in indirect taxes is, not unexpectedly, much more regressive. In terms of disposable income the tax increase is more than twice as large for the bottom deciles than for the top decile. This indirect tax increase not only turns an already modest impact of disposable income at the bottom into a substantial loss (−3.7% for the bottom decile), it also makes the whole first order impact of the reform rather regressive: from the first to the seventh decile, the impact change is increasing from a loss of −3.7% to a gain of 2.3% (in absolute amounts: from a loss of €45 per month to a gain of €77 per month). Since the gains decrease in relative terms at the top three deciles (from 2.1% to 0.8%), one can at best bend the adjective 'regressive' into 'pro middle class' to describe this tax reform, but certainly not into 'pro poor'.[32] The explanation for this 'pro middle class' distributional pattern of the impact effects lies in the character of the tax reform: a decrease in social security contributions paid by the employee only affects people at work, and part of the personal income tax reductions, such as the increase in deductible professional expenses, is also directly linked to labour market status.

As a consequence, when limiting the distributional analysis to the working population,

[32]This conclusion is robust across the different simulations. With the small increase in indirect taxes of simulation TS2, the level of the effect of indirect taxes is of course smaller. But the distributional pattern is similar: the change in indirect taxes is on average 0.6% of disposable income, but it monotonically declines from 1.2% for the bottom decile to 0.3% for the top decile. Combined with the first order changes in disposable income, the total first order effect rises from a loss of 0.6% for the first decile up to a gain of 4.5% in the seventh decile, and then declining again to a gain of 2.6% for the top decile.

one finds a drastically different distributional pattern. Table A.13 in Appendix A.7 shows that across deciles of the gross wage distribution, the first order percentage gain of real disposable income (see column (5)) declines monotonically from 2.7% for the bottom decile to al loss of 0.8% for the top decile.

The second message from Table 3 is the large impact of the behavioural changes on the distributional assessment in terms of monetary gains. Comparing the columns which include the behavioural changes (column (4) in €'s per month and column (10) in percentages) with the ones of the first order effects (columns (1) and (7)), the positive effects on employment at the bottom of the distribution described in the previous section, trigger large positive effects on disposable income for households in the bottom deciles. For all deciles in the bottom half of the distribution, the change in disposable income including behavioural change exceeds the impact effect. Certainly for the bottom two deciles this effect is considerable: the percentage gain for these two deciles jumps from 0.6% and 2.3% to 8.1% and 3.9% respectively. Even after taking into account the increase in indirect taxes, the change in welfare for these bottom two deciles is positive. Even though the welfare gain of the bottom decile amounts to a considerable 3.8% (column (12) reporting the upper bound of ΔW_{EV}), the overall pattern of distributional change in welfare according to the upper bound of ΔW_{EV} does not turn progressive. The percentage gain increases from the second up to the seventh decile. The decline in relative welfare gain for the eight and ninth decile, and the loss incurred by the richest decile is again partly explained by the switch to jobs which earn less but are more attractive in other dimensions, occurring — as seen in Section 5.2 — especially at the upper end of the distribution.

Two additional remarks about this loss at the top of the distribution are worth mentioning. First, in Tables A.11 and A.12 in Appendix A.7, we replicate the distributional analysis of Table 3 for a simulation in which we keep the wages fixed (labelled TS3b). This avoids the welfare loss of the top decile: the upper bound of the equivalent variation is now 0.4% instead of −2.2%. The limited progressivity in the top four deciles remains though: the upper bound of ΔW_{EV} declines from 2.1% for the seventh to 0.4% for the 10th decile. Also, the regressive pattern between decile 2 (welfare gain of 0.4%) and decile 7 (welfare gain of 2.1%) remains. Therefore, we conclude that it is not the model characteristic of being able to choose a job with a different wage which drives the distributional pattern, but still, in the first place, the specificities of the reform. Second, as far as part of these more modest welfare gains at the top of the distribution are indeed explained by behavioural changes, the equivalent variation metric is not really the appropriate one to base an assessment on, as it neglects the welfare cost of changes in labour time. In Section 5.5 we will therefore use the metric introduced in Section 3.2.2 to take that into account in the distributional analysis. First we widen the scope of the distributional analysis to some other relevant characteristics of individuals or households in the next subsection.

5.4 Characteristics of winners and losers

To widen the description of distributional effects beyond monetary metrics, we employ a representation originally advocated by King (1983). It describes characteristics of groups of winners and losers. In the upper part of Table 4 we display the characteristics of five equally sized groups when the gain/loss concept is the one of the impact effect, measured as the first order effect on disposable income minus the first order effect on indirect taxes. We express this as a percentage of disposable income in the baseline (i.e. column (9) in Table 3), and then order the individuals from largest losers to largest winners. In the bottom part of the table we use the gain/loss concept when also behavioural changes are taken into account. In this case, we order the individuals on the basis of the upper bound for ΔW_{EV}, again expressed as a percentage of disposable income (i.e. column (12) of Table 3).

Table 4: Characteristics of quintiles of gainers and losers of reform TS3

	losers				winners	
characteristics	Q1	Q2	Q3	Q4	Q5	all
	quintiles based on gain/loss of impact effect					
age	53.8	60.2	46.1	40.7	38.4	48.3
share of pensioners (%)	32.4	54.2	12.9	3.2	0.8	21.9
share of working in baseline (%)	7.6	23.2	63.0	78.9	89.3	50.9
share of unemployed (%)	53.5	17.7	14.4	10.9	5.4	20.7
disposable income baseline (€/month)	1,538	2,937	3,830	3,502	2,953	2,850
impact effect (%)	-4.2	-0.4	1.2	2.6	4.6	0.3
total effect (%)	0.0	-1.1	0.3	2.1	4.4	0.9
	quintiles based on gain/loss of total effect					
age	52.5	61.0	46.6	40.8	38.3	48.3
share of pensioners (%)	28.4	56.8	13.8	3.2	0.7	21.9
share of working in baseline (%)	14.5	21.1	61.9	79.1	85.2	50.9
share of unemployed (%)	50.3	17.5	14.5	11.0	9.3	20.7
disposable income baseline (€/month)	1,740	2,851	3,789	3,462	2,838	2,850
impact effect (%)	-3.9	-0.5	1.2	2.6	4.1	0.3
total effect (%)	-6.0	-0.5	1.2	2.6	9.8	0.9

Each part of the table displays the value of a characteristic (household size, age, disposable income, etc.) for five groups of relative winners and losers of the reform. The upper part of the table displays the characteristics when individuals are ordered on the basis of the impact effect (i.e. on the basis of column (9) in Table 3 which uses the change in disposable income minus indirect taxes as a percentage of disposable income to order the individuals from largest losers to largest gainers); the second part also takes into account behavioural change (i.e. uses the values of column (12) of Table 3, where ΔW_{EV} expressed as a percentage of disposable income is used). All results in this table are averages calculated for the entire population at the individual level, except for disposable income in the baseline, the impact effect, and the total effect, which are calculated at the level of the household.

From the upper part of Table 4, it is clear that older and non–working people are over-represented in the group of (relative) losers of the reform. The average age being 48.3 years for the whole population of individuals, we find that in the bottom two quintiles of the gain/loss-distribution, the average age is significantly higher: 53.8 in the bottom quintile of largest losers of this reform, and even 60.2 years in the second quintile. The 20% largest gainers of this reform are on average 38.4 years old. Also the share of pensioners, of working people, and of unemployed in the different quintiles of gains and losses, reveals that those who are employed in the baseline are among the biggest winners of this reform, whereas pensioners and unemployed, by contrast, are consistently found in the lower quantiles of the gain/loss distribution. This particular distributional pattern is not surprising, given that the direct tax reform focussed almost exclusively on lowering the burden on employment and had little to offer to those who do not work. Increased indirect taxes magnify this effect even further, as they apply more uniformly to the whole of the population. In that regard, one can interpret a tax shift from direct to indirect taxation as a transfer of the tax burden from those who do work to those who do not work. The row in Table 4 which displays disposable income in the baseline, rephrases the results of Table 3, but from the inverse angle: the biggest losers of this reform have an income in the baseline which is much lower than the income of the gainers or than average income. Note however that the non–monotonic pattern already found in Table 3 is also replicated in this inverse representation: average disposable income of quintile 4 and 5 is below that of quintile 3.

This non–monotonic pattern in terms of disposable income is still much more pronounced in the bottom part of Table 4. The highest disposable incomes of the baseline are predominantly found in the middle of the gain/loss–distribution. The quintile of largest losers of the reform is still predominantly populated by poorer individuals, but the fact that the income of the largest gainers of the reform (quintile 5) is lower than the income in quintiles 2, 3 and 4, reflects the fact that some high income households have moved down in the gain/loss distribution. This is again caused by the effect described above: some high income households lose in terms of disposable income since they now choose to work less and even at lower gross wages. This probably also explains the finding that, comparing the bottom part of the table (which includes behavioural reactions), with the upper part (which is only based on the impact effect), we find that the share of people working in the baseline decreases at the top of the gain/loss distribution. For the the subgroup of 20% largest gainers, it goes down from 89.3% to 85.2%. On the contrary, in the group of largest losers (quintile 1) the share of people working in the baseline goes up, from 7.6% to 14.5%. In the next subsection we enrich the distributional assessment of the reform by also taking into account the welfare cost of the change in leisure time into the welfare metric.

5.5 Taking into account the welfare cost of changing labour time

To account for the change in leisure which goes with the change in labour supply we now turn to the welfare measure based on the equivalent consumption, ΔW_{EC}, which is defined in Equation (36). As discussed in Section 3.2.2, we contrast this measure with $\overline{\Delta W_{EC}}$, which denotes the EC-based welfare change for the fictitious reference point with pre–reform working hours and post–reform real disposable income (see Equation 38). This approach allows us to assess the importance of controlling for changes in labour time within the same class of welfare measures. The advantage is that one does not have to resort to comparisons across different classes of measures, which each have distinct underlying cardinal properties (see also the discussion in Section A.6 in the Appendix).

Table 5 summarises the distributional pattern of these measures across the gross wage distribution of the individuals in the RURO subsample (as in Table 2), and is again limited to simulation TS3. Column (3) shows that the welfare gain according to the EC–criterion is positive up to the 7th decile, reaching its maximum at decile 3 (€41). From the 4th decile onwards, the gain decreases monotonically, which culminates in an average loss of €126 euro for those in decile 10. Note that despite the surge in employment (see Table 2), the average welfare change remains positive in the 1st decile. The values for ΔW_{EC} and $\overline{\Delta W_{EC}}$ are similar for the 5th and 6th decile, whereas larger deviations occur at both ends of the wage distribution. As one would expect, individuals in deciles 1 to 4 would gain more when the cost of increased labour supply is not accounted for. Alternatively, welfare losses for individuals in deciles 7 to 10 would be augmented when we do not account for the decrease in working hours. This comparison illustrates that it is important to account for the welfare implications of changes in leisure time in the assessment of tax reforms.

The relative changes with respect to baseline equivalent consumption are presented in columns (4) and (5). Also in relative terms, the upper deciles are among the biggest losers of the reform. Decile 3 gains the most (3.7%), but also 2 to 3 and 4 to 6 exhibit substantial increases. When discarding the impact of changes in hours worked (column (4)), the gain in the 1st decile (4.9%) approaches that in the 3rd (5.3%).

In Table 6 we provide a comparison between EV– and EC–based welfare measures that does not depend on the particular cardinal properties underlying both concepts. This comparison is inspired by the winners and losers picture provided in Section 5.4 (King, 1977). The upper panel shows the percentage of winners and losers by subcategories of persons who either decrease labour supply, keep their supply fixed, or increase labour time. We first concentrate on the singles within the RURO subpopulation. The number of winners among those who increase labour time is smaller according to the EC–based measure (2.54%) than according to the EV–based measure (3.43%). Reversely, the number of losers among those who have more leisure after the reform, is smaller according to the EC–based measure (1.69%) than according to the EV–based measure (1.98%). Moving to the complete RURO subpopulation, this picture is confirmed when it comes to those who work longer after the reform, but not for those who switch to jobs with lower labour time. This deviation is explained by the interaction between partners' leisure time. We can thus confirm that when correcting for the welfare loss of increased labour time, the number of winners among those who increase labour supply decreases.

The lower panel of Table 6 contains the average baseline welfare level in terms of the EC–measure (thus including welfare from leisure) for the winners and losers according to

Table 5: Comparison of changes in equivalent consumption (W_{EC} and $\overline{W_{EC}}$) across deciles of gross wages for revenue neutral simulation TS3 (only RURO subpopulation)

deciles	baseline in €/month	changes in €/month		changes in %	
of gross wage	W_{EC} (1)	$\overline{W_{EC}}$ (2)	W_{EC} (3)	$\overline{W_{EC}}$ (4)	W_{EC} (5)
1	1,018	50	18	4.9	1.8
2	1,098	46	30	4.2	2.7
3	1,101	58	41	5.3	3.7
4	1,131	41	34	3.6	3.0
5	1,266	29	29	2.3	2.3
6	1,355	22	21	1.7	1.6
7	1,459	5	10	0.3	0.7
8	1,572	-10	-1	-0.6	-0.1
9	1,759	-33	-25	-1.9	-1.4
10	2,234	-140	-126	-6.3	-5.6
all	1,399	7	3	0.5	0.2

The variables are calculated at the individual level; for couples we divide both welfare concepts, $\overline{W_{EC}}$ and W_{EC}, by two. Each individual in the RURO population is allocated to the decile based on his or her gross wage (with an equal number of persons by decile). The non–working individuals are assigned a gross wage based on the estimated wage equation of the RURO model. Columns (2) and (3) show the change in equivalent consumption defined in Equations (38) and (36) respectively. All percentages in columns (4) and (5) are with respect to equivalent consumption in the baseline (the latter are in column (1)).

both criteria. We find that the losers according to the EV–criterion have on average lower overall welfare (1,470) than the losers according to the EC–criterion (1,592). This seems to indicate that the losers when taking into account the welfare cost from losses in leisure, have on average a higher welfare level than what would be the case if one identified the losers without taking into account that cost. It depends however on individual preferences whether these are high income earners with little amount of leisure, or vice versa, or somewhere in between. Summary statistics show that the group identified as losers according to the EC–criterion, both, work on average more, and have higher real income than the corresponding group of losers according to the EV–criterion. There is no obvious explanation for this finding. We suspect that it requires careful examination of changes in the budget set due to the reform among groups with different preferences.

Finally, we find that almost 5% of the households in the RURO subsample lose according to the EC-measure, whereas their real disposable income would increase when they would stick to the same job choice as before the reform. In Appendix A.6 we argue that in principle this cannot occur: if a person obtains a higher real disposable for the observed baseline alternative, the welfare change according to the EC-criterion should always be positive. This departure is caused by unobserved preference variation in the RURO model, which impacts a person's optimal choice behaviour, but is excluded from our definition of the EC-criterion in Equation (35).[33] Some individuals that lose according to our EC-criterion, actually gain

[33]This unobserved variation embodies the relative desirability of unobserved pecuniary and non-pecuniary

Table 6: EV– versus EC–based measures of welfare and changes in labour and welfare

welfare measure	$h_0 > h_1$	$h_0 = h_1$	$h_0 < h_1$
	singles within RURO subpopulation		
$\Delta\, W_{EC}$			
losers	1.69%	38.06%	1.70%
winners	0.34%	55.66%	2.54%
$\Delta\, W_{EV}$			
losers	1.98%	33.91%	0.81%
winners	0.05%	59.81%	3.43%
	complete RURO subpopulation		
$\Delta\, W_{EC}$			
losers	2.65%	23.33%	1.99%
winners	0.68%	69.10%	2.24%
$\Delta\, W_{EV}$			
losers	1.86%	13.72%	0.84%
winners	1.47%	78.71%	3.40%

welfare measure	$\Delta\, W_{EC} < 0$	$\Delta\, W_{EC} \geq 0$	All
W_{EC_0}	1,592	1,324	1,399
	$\Delta\, W_{EV} < 0$	$\Delta\, W_{EV} \geq 0$	All
W_{EC_0}	1,470	1,385	1,399

The welfare variables are calculated at the individual level; for couples we divide the welfare concepts at the household level by two.

in overall welfare when unobserved preference heterogeneity is properly taken into account. The incorporation of variation over this non-systematic source of preference heterogeneity into the welfare evaluation is left for future research (see also Carpantier and Sapata, 2016).

6 Conclusion

In this paper, we propose a *piecemeal modelling* strategy to establish a micro–based budget and welfare evaluation of a joint reform in the personal income and the indirect tax system. Using a two–stage budgeting approach, we allow (changes in) commodity prices to interact with households' labour supply decision. In addition, we employ a random utility random opportunity (RURO) model of labour supply, which enables us to incorporate feedback from the demand–side of the labour market into our analysis.

Applying this framework to a recently enacted Belgian tax shift, we find substantial evidence that it is important to account for indirect taxes and demand–side effects in the assessment of joint tax reforms. Our results show that, despite an overall increase in labour

attributes. See also the references in Section 2.2.

supply, cost recovery effects from personal income taxes and employee and employer SSC are small in each of the four simulated scenarios. These cost recovery effects worsen considerably when we impose more revenue neutrality, by increasing VAT more substantially. In addition, our simulations indicate that a more revenue neutral reform would destroy almost one third of the newly created FTEs, compared to the actual under–financed Belgian tax shift. However, feedback effects from the demand side of the labour market could mitigate this detrimental effect.

If we keep households' consumption bundle and labour supply fixed, our results show that the poorest bottom half of the distribution gain the least from the reform. The gain in percentage of disposable income increases up to the seventh decile. The bottom two deciles even incur a loss. However, if we allow households to alter their behaviour the pictures reverses. The poorest decile now gains the most in relative terms (within this group, employment rises by 14 percentage points among the population included in our job choice model). In line with the literature, we find a negative employment effect at the intensive margin for affluent households. They choose jobs with less working hours on average. Moreover, the RURO model predicts that the latter will also choose jobs with lower gross wages on average. This explains the low cost recovery of the reforms since lower gross wages trickle down into a negative effect on the taxable base for both PIT and social security contributions. From the second to the seventh decile we still obtain an increasing pattern of relative gains, be it to a lesser extent than was the case without behaviour. For the upper three deciles we obtain a steeper decline than in the case without behavioural reactions. If anything, the redistribution is from the rich to the middle class, rather than 'pro poor'.

Finally, when we rank households by means of the equivalent variation, we find that those who are employed in the baseline are predominantly among the *winners* of the reform; those who are unemployed or pensioned, by contrast, are more concentrated among the *losers*. The middle quintile has on average the highest disposable income; the biggest 40% of winners have on average higher incomes than the 40% smallest winners (or losers).

We also find that taking into account the changes in leisure time is important for the distributional assessment of a tax-benefit reform. The welfare gains for individuals with low wages, who increase their labour supply considerably, are smaller when the decrease in leisure time is taken into account. The opposite holds for individuals with high wages, who work less after the reform on average. When exploiting only ordinal properties of the measures based on the equivalent variation and consumption (identifying the group of winners from the group of losers), we can confirm that there are relatively more losers among the group who increases labour supply when accounting for the welfare cost of leisure time. Surprisingly, losers according to a measure including that welfare cost have a higher welfare level than losers according to measures neglecting that cost, they work more, and they have a higher real income. Understanding why requires a careful examination of changes in the budget set due the reform among groups with different preferences. This is an avenue for further research.

Similarly, we showed that a fully consistent of assessment of welfare gains in a discrete choice model requires us to integrate the variation in non–systematic parts of preferences into the welfare evaluation.

7 References

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

Aaberge, R. and Colombino, U. (2014), Labour supply models, in: C. O'Donoghue (ed.), Handbook of Microsimulation Modelling, Contributions to Economic Analysis, Volume 293, Emerald Group Publishing Limited: Bingley, 167–221.

Aaberge, R., Dagsvik, J., and Strøm, S. (1995), Labour supply responses and welfare effects of tax reforms, Scandinavian Journal of Economics, Vol. 97(4), 635–659.

Bach, S., Haan, P., Hoffmeister O., and Steiner V. (2006), Increasing the value-added tax to re-finance a reduction of social security contributions? A behavioral microsimulation analysis for Germany, mimeo, DIW: Berlin.

Altig, D., Auerbach, A.J., Koltikoff, L.J., Smetters, K.A., and Walliser J. (2001), Simulating fundamental tax reform in the United States, American Economic Review, Vol. 91(3), 574–595.

Bargain, O., Decoster, A., Dolls, M., Neumann, D., Peichl, A., Siegloch, S. (2013), Welfare, Labor Supply and Heterogeneous Preferences: Evidence for Europe and the US, Social Choice and Welfare, Vol. 41(4), 789–817.

Bargain,O., Orsini, K. and Peichl, A. (2014), Comparing Labor Supply Elasticities in Europe and the United States, Journal of Human Resources, Vol. 49 (3), 723–838.

Barten, A.P. and Böhm, V. (1986), Consumer theory, in: K.J. Arrow and M.D. Intrilligator (eds.) Handbook of mathematical economics Volume II, North–Holland: Amsterdam, 381–421.

Blundell, R., Browning, M., and Meghir, C. (1994), Consumer demand and the life–cycle allocation of household expenditures, Review of Economic Studies, Vol.61(1), 57–80.

Blundell, R., Duncan, A., McCrae, J., and Meghir, C. (2000), The labour market impact of the working families' tax credit, Fiscal Studies, Vol. 21(1), 75–104.

Blundell, R. and MaCurdy, T. (1999), Labor supply: a review of alternative approaches, in: O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, Volume 3A, North-Holland Elsevier: Amsterdam, 1559–1695.

Blundell, R., MaCurdy, T., and Meghir, C. (2007), Labor supply models: Unobserved heterogeneity, nonparticipation and dynamics, in: J. Heckman, and E. Leamer (eds.), Handbook of Econometrics, Volume 6A, North–Holland Elsevier: Amsterdam, 4667–4775.

Blundell, R. and Walker, I. (1982), Modelling the joint determination of household labour supplies and commodity demands, The Economic Journal, Vol. 92 No.366, 351–364.

Blundell, R. and Walker, I. (1986), A life–cycle consistent empirical model of family labour supply using cross-section data, Review of Economic Studies, Vol. 53(4), 539–558.

Böhringer, C., Boeters S., and Feil, M. (2005), Taxation and unemployment: an applied general equilibrium approach, Economic Modelling, Vol. 22(1), 81–108.

Bourguignon, F., Fournier, M., and Gurgand, M. (2001), Fast development with a stable income distribution: Taiwan, 1979–94, Review of Income and Wealth, Vol. 47(2), 139–163.

Browning, M., Deaton, A., and Irish, M. (1985), A profitable approach to labor supply and commodity demands over the life–cycle, Econometrica, Vol. 53(3), 503–543.

Browning, M. and Meghir, C. (1991), The effects of male and female labor supply on commodity demands, Econometrica, Vol. 59(4), 925–951.

Capéau, B. and Decoster, A. (2016), Getting tired of work, or re–tiring in absence of decent job opportunities? Some insights from an estimated Random Utility/Random Opportunity model on Belgian data, EUROMOD working paper series, EM4/16.

Capéau, B., Decoster, A., and 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, Vol. 9(2), 144–191.

Capéau, B., Decoster, A., De Swerdt, K., and Orsini, K. (2009), Welfare effects of alternative financing of social security. Some calculations for Belgium, in: A. Harding, P. Williamson, and A. Zaidi (eds.) New Frontiers in Microsimulation Analysis, Ashgate: Farnham, 437–470.

Carpantier, J.–F. and Sapata, C. (2016), Empirical welfare analysis: when preferences matter, Social Choice and Welfare, vol. 46(3), 521–542.

Chen, S. and Ravaillon, M. (2004), Welfare impacts of China's accession to the world trade organization, The World Bank Economic Review, Vol. 18(1), 29–57.

Cockburn, J., Savard, L., and Tiberti, L. (2014), Macro–micro models, in: C. ODonoghue (ed.) Handbook of Microsimulation Modelling , Emerald Insight, Bingley, 275–304.

Colombino, U. (2013), A new equilibrium simulation procedure with discrete choice models, International Journal of Microsimulation, Vol. 6(3), 25–49.

Creedy, J. and Kalb, G. (2005), Discrete hours labour supply modelling: Specification, estimation and simulation, Journal of Economic Surveys, Vol. 19(5), 697–734.

Dagsvik, J.K., Jia, Z., Kornstad, T., and Thoresen, T.O. (2014), Theoretical and practical arguments for modeling labor supply as a choice among latent jobs, Journal of Economic Surveys, Vol. 28(1), 134–151.

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

Dagsvik, J.K. and Jia, Z. (2016), Labor supply as a choice among latent jobs: unobserved heterogeneity and identification, Journal of Applied Econometrics, Vol. 31(3), 487–506.

Dahlby, B. (2003), Restructuring the Canadian tax system by changing the mix of direct and indirect taxes, in: H. Grubel (ed.) Tax reform in Canada: Our path to greater prosperity, The Fraser Institute: Vancouver, 77–108.

De Agostini, P., Capéau, B., Decoster, A., Figari, F., Kneeshaw, J., Leventi, C., Manios, K., Paulus, A., Sutherland, H., Vanheukelom, T. (2017), EUROMOD extension to indirect taxation: final report, EUROMOD technical note series, EMTN/3.0.

de Castro Fernàndez, F., Perelle M. and Priffis R. (2018), The economic effects of a tax shift from direct to indirect taxation in France, European Economy Discussion Papers 077.

Decoster, A., De Rock, B., De Swerdt, K., Loughrey, J., O'Donoghue, C., and Verwerft, D. (2007), Comparative analysis of different techniques to impute expenditures into an income data set, Accurate Income Measurement for the Assessment of Public Policies Deliverable 3.4, KULeuven.

Decoster, A. and Haan, P. (2015), Empirical welfare analysis with preference heterogeneity, International Tax and Public Finance, Vol. 22(2), 224–251.

Devriendt, W. and Heylen, F. (2017), Macroeconomic effects of demographic change in an OLG model for a small open economy – The case of Belgium, UGent Faculteit Economie en Bedrijfskunde Working Paper no.2017/931.

Duncan, A. and Weeks M. (2000), Transition estimators in discrete choice models, in L. Mitton, H. Sutherland, and M. Weeks (eds.) Microsimulation Modelling for Policy Analysis. Challenges and Innovations, Cambridge University Press: Cambridge, 292–308.

European Commission services (DG TAXUD) (20006), Macroeconomic effects of a shift from direct to indirect taxation: a simulation for 15 EU member states, paper presented at the 72nd meeting of the OECD Working Part No.2 on Tax Policy Analysis and Tax Statistics, Paris.

Fleurbaey, M. (2006), Social welfare, priority of the worst–off and the dimensions of individual well–being in: F. Farina and E. Savaglio (eds.) Inequality and Integration, Routledge: London & New York, 225–268.

FPB (2015), Macro–economische en budgettaire effecten van de tax shift maatregelen van de federale regering, Federal Planning Bureau.

Gorman, W. M. (1971), Two stage budgeting, mimeo, London School of Economics. Printed in: C. Blackorby and A. Shorrocks (eds., 1995) Separability and Aggregation. The Collected Works of W. M. Gorman, Volume 1. Oxford University Press: Oxford, 22–29.

Hausman, J., (1981), Labor supply, in: H. Aaron and J. Pechman (eds.), How Taxes Affect Economic Behavior, Brookings Institution: Washington DC, 27–83.

Hausman, J., (1985a), The econometrics of nonlinear budget sets, Econometrica, Vol. 53(6), 1255–1282.

Hausman, J., (1985b), Taxes and labor supply, in: A. Auerbach and M. Feldstein (eds.), Handbook of Public Economics, Volume 1, North–Holland Elsevier: Amsterdam, 213–263.

Keane, M. (2011), Labor supply and taxes: A survey, Journal of Economic Literature, Vol.49(4), 961–1075.

Keane, M., Todd, P., and Wolpin, K. (2011), The structural estimation of behavioural models: Discrete choice dynamic programming methods and applications, O. Ashenfelter and

D. Card (eds.), Handbook of Labor Economics, Volume 4A, North–Holland Elsevier: Amsterdam, 331–461.

Kakwani, N.C. (1977), Measurement of tax progressivity: An international comparison, The Economic Journal, Vol.87 No. 345,: 71–80.

King, M.A. (1983), Welfare analysis of tax reforms using household data, Journal of Public Economics, Vol. 21(2), 183–214.

Mastrogiacomo, M., Bosch, N., Gielen, M. and Jongen, E. (2017), Heterogeneity in Labour Supply Responses: Evidence from a Major Tax Reform, Oxford Bulletin of Economics and Statistics, Vol. 79(5), 769–796

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

Myles, G. (2009a), Economic growth and the role of taxation – Aggregate data, OECD Economics Department Working Papers, No. 713, OECD Publishing, Paris.

Myles, G. (2009b), Economic growth and the role of taxation – Aggregate data, OECD Economics Department Working Papers, No. 714, OECD Publishing, Paris.

Myles, G. (2009c), Economic growth and the role of taxation – Aggregate data, OECD Economics Department Working Papers, No. 715, OECD Publishing, Paris.

NBB (2015), Macro–economische impact op de Belgische economie van het door de Regering uitgewerkte taxshiftscenario, National Bank of Belgium, Note DS.15.11.531-NL.docx.

NBB (2017), Macro–economische en budgettaire impact van het door het kabinet van de Minister van Financiën gereviseerde taxshiftscenario en vergelijking met de oefening van november 2015, National Bank of Belgium, Note DS.17.07.320-NL.docx.

Pazner, E. and Schmeidler, D. (1978), Egalitarian equivalent allocations: a new concept of economic equity, Quarterly Journal of Economics, Vol. 92(4), 671–687.

Peichl, A. (2009), The benefits and problems of linking micro and macro models – evidence from a flat tax analysis, Journal of Applied Economics, Vol. 12(2), 301–329.

Peichl, A. (2016), Linking microsimulation and CGE models, International Journal of Microsimulation, Vol. 9(1), 167–174.

Peichl, A. and Siegloch, S. (2012), Accounting for labor demand effects in structural labor supply models, Labour Economics, Vol. 19(1), 129–138.

Pestel, N. and Sommer, E. (2017), Shifting Taxes from labor to consumption: more employment and more inequality, Review of Income and Wealth, Series 63(3), 542–563.

Savage, M. (2017), Integrated modelling of the impact of direct and indirect taxes using complementary datasets, The Economic and Social Review, Vol. 48(2), 171-205.

Sutherland, H. and Figari, F. (2013), EUROMOD: the European Union tax–benefit microsimulation model, International Journal of Microsimulation, Vol. 6(1), 4-26.

Van Soest, A. (1995), Structural models of family labor supply: a discrete choice approach, The Journal of Human Resources, Vol. 30(1), 63–88.

A Appendix

A.1 Selection of the RURO model subsample

We estimate the labour supply model on a subsample of the Belgian Statistics on Income and Living Conditions (SILC) survey of 2015. This RURO subsample contains all households in which the reference person and, in the case of couples, his or her partner are both available for the labour market. That is, we drop households in which at least one partner is younger than 18 or older than 64 years; at least one partner is an employer or self-employed; or at least one partner is disabled, retired, or studying. Households who declare additional live-in adult members that are also available for the labour market were equally removed, as were same-sex couples. The final subsample consists of 644 single females, 526 single males, and 1,424 couples. Table A.1 contains the weighted descriptive statistics for each of these three groups.

Table A.1: Descriptive statistics RURO subsample

Description	Singles		Couples	
	Female	Male	Female	Male
Age (years)	43.0	41.5	40.2	42.6
Dependent children (%)				
0 − 3 years	5.0	0.7	18.2	
4 − 6 years	6.2	1.4	18.6	
7 − 9 years	7.8	0.9	16.2	
Experience (years)	16.6	18.8	16.4	21.0
Education (%)				
Low	22.1	19.6	12.7	15.1
Middle	36.1	39.0	34.7	40.6
High	41.8	41.4	52.6	44.3
Residence (%)				
Brussels	20.2	17.3	9.1	
Flanders	45.4	48.3	60.8	
Wallonia	34.4	34.4	30.1	
Participation rate (%)	67.9	73.7	81.0	91.1
Hours worked (hours/week)				
Unconditional	24.0	29.8	27.0	36.8
Conditional on working	35.3	40.4	33.3	40.4
Hourly wage (euro)	20.4	21.2	20.7	22.8
Disposable income (euro/month)	2,123.1	2,345.9	4,392.5	
Number of observations	644	526	1,424	

Observations are weighted according to their SILC sample weights.

A.2 Parameters and elasticities RURO model

Parameter estimates This appendix contains the estimated parameters of the RURO model explained in Section 2.2 (see Table A.2).

Table A.2: Parameters estimates RURO model

Description	Singles						Couples					
	Female			Male			Female			Male		
	E	S	T	E	S	T	E	S	T	E	S	T
Box-Cox utility function												
constant consumption	5.48	0.33	16.51	5.21	0.32	16.13	5.55	0.31	17.68	5.55	0.31	17.68
exponent consumption	0.34	0.08	4.41	0.43	0.07	6.28	0.61	0.04	15.38	0.61	0.04	15.38
exponent leisure	-5.68	0.66	-8.66	-3.82	0.65	-5.84	-7.89	0.53	-14.94	-7.78	0.55	-14.16
interaction leisure	·	·	·	·	·	·	0.23	0.08	3.03	0.23	0.08	3.03
Leisure covariates												
constant	84.38	35.52	2.38	1.50	30.93	0.05	33.95	12.97	2.62	7.32	8.09	0.90
log(age)	-44.47	19.25	-2.31	1.04	17.00	0.06	-18.84	7.30	-2.58	-4.08	4.47	-0.91
log(age)^2	6.12	2.63	2.33	-0.04	2.33	-0.02	2.81	1.03	2.73	0.66	0.62	1.06
child $0-3$	-0.53	0.54	-0.97	·	·	·	0.46	0.16	2.78	0.15	0.09	1.74
child $4-6$	0.16	0.49	0.32	-1.58	1.02	-1.56	0.58	0.16	3.51	0.10	0.08	1.19
child $7-9$	-0.08	0.39	-0.20	-1.14	1.27	-0.90	0.33	0.15	2.23	-0.05	0.07	-0.63
Brussels	0.39	0.33	1.17	0.66	0.46	1.46	0.21	0.15	1.36	0.32	0.10	3.10
Wallonia	-0.22	0.33	-0.66	0.08	0.46	0.18	0.09	0.20	0.44	0.35	0.14	2.51
low education	0.60	0.55	1.10	-0.13	0.62	-0.21	-0.62	0.30	-2.08	0.16	0.15	1.04
high education	-0.85	0.33	-2.60	-0.76	0.45	-1.71	-0.80	0.18	-4.52	-0.11	0.08	-1.45
Wage rate covariates												
constant	2.21	0.03	79.69	2.23	0.03	78.32	2.21	0.03	79.69	2.23	0.03	78.32
potential experience	2.79	0.23	12.16	2.50	0.24	10.42	2.79	0.23	12.16	2.50	0.24	10.42
potential experience^2	-3.79	0.53	-7.16	-3.20	0.52	-6.12	-3.79	0.53	-7.16	-3.20	0.52	-6.12
low education	-0.15	0.03	-5.85	-0.11	0.02	-5.29	-0.15	0.03	-5.85	-0.11	0.02	-5.29
high education	0.24	0.02	15.71	0.23	0.01	16.32	0.24	0.02	15.71	0.23	0.01	16.32
RMSE	0.27	0.00	61.17	0.27	0.00	63.59	0.27	0.00	61.17	0.27	0.00	63.59
Opportunities covariates												
constant	-58.15	13.54	-4.29	-30.80	18.08	-1.70	-58.15	13.54	-4.29	-30.80	18.08	-1.70
unemployment rate	0.15	0.21	0.72	-0.30	0.29	-1.03	0.15	0.21	0.72	-0.30	0.29	-1.03
log(age)	30.07	7.39	4.07	16.15	9.71	1.66	30.07	7.39	4.07	16.15	9.71	1.66
log(age)^2	-4.18	1.01	-4.14	-2.45	1.31	-1.87	-4.18	1.01	-4.14	-2.45	1.31	-1.87
Brussels	-0.07	0.18	-0.42	0.05	0.23	0.21	-0.07	0.18	-0.42	0.05	0.23	0.21
Wallonia	-1.11	0.20	-5.50	-1.01	0.24	-4.13	-1.11	0.20	-5.50	-1.01	0.24	-4.13
low education	-0.85	0.29	-2.98	0.08	0.37	0.22	-0.85	0.29	-2.98	0.08	0.37	0.22
high education	-0.43	0.21	-2.10	-0.44	0.25	-1.79	-0.43	0.21	-2.10	-0.44	0.25	-1.79
Hours peaks												
interval [18.5; 20.5]	1.75	0.11	16.32	0.85	0.22	3.80	1.75	0.11	16.32	0.85	0.22	3.80
interval [29.5; 30.5]	1.76	0.10	17.04	1.35	0.16	8.26	1.76	0.10	17.04	1.35	0.16	8.26
interval [37.5; 40.5]	1.95	0.07	27.94	2.60	0.06	42.64	1.95	0.07	27.94	2.60	0.06	42.64

E: Estimate S: Standard error T: t-value

Labour supply elasticities Tables A.3 and A.5 contain the counterparts for RURO models for the wage elasticities in a traditional labour supply model. In RURO models, wages are charateristics of job offers, implying that the labour supply reaction on a person's wage is an ill defined concept. Instead we calculate the effect on job choice and, consequently, hours chosen, of a small displacement of the wage offer distribution to the right (by increasing the estimated location parameters of the log normal wage offer distributions). Table A.3 contains the aggregated effects. Contrary to similar exercises with this model in the past, and to other empirical labour supply models, we obtain quite similar intensive margin elasticities for males and females, except for the cross effect in couples.

Tables A.4 to A.5 contain the effects split out in quintiles of the average wage offers. Total own wage elasticities tend to be smaller at higher wages (the picture is less clear for intensive margin elasticities). This is in line with our expectations given the observed negative income effect in our simulations at the higher end of the income distribution. But notice that the effect in the simulations was mainly due to a change in the shape of the net wage distribution in favour of lower wages.

Table A.3: Aggregate wage elasticities of labour supply

	Shift of the female wage distribution			Shift of the male wage distribution		
	Single	Couple		Single	Couple	
	Female	Female	Male	Male	Female	Male
elasticity: total	0.49	0.49	-0.16	0.52	-0.26	0.33
elasticity: intensive margin	0.09	0.17	-0.26	0.05	-0.58	0.12
part in (%)	0.29	0.25	0.09	0.36	0.30	0.19
part out (%)	0.00	0.00	0.19	0.00	0.41	0.00

These results are obtained by increasing the first moment of the distribution of offered wages by 10%.

Table A.4: Total elasticities by wage quintile

	Shift of the female wage distribution			Shift of the male wage distribution		
	Single	Couple		Single	Couple	
	Female	Female	Male	Male	Female	Male
Q1	1.07	1.50	0.00	1.14	0.68	0.63
Q2	0.60	0.52	-0.17	0.40	-0.41	0.38
Q3	0.42	0.37	-0.26	0.55	-0.34	0.18
Q4	0.48	0.36	-0.17	0.56	-0.34	0.31
Q5	0.34	0.31	-0.17	0.23	-0.34	0.20

These results are obtained by increasing the first moment of the distribution of offered wages by 10%.

Table A.5: Intensive margin elasticities by wage quintile

	Shift of the female wage distribution			Shift of the male wage distribution		
	Single	Couple		Single	Couple	
	Female	Female	Male	Male	Female	Male
Q1	0.04	0.12	-0.24	-0.06	-0.97	0.11
Q2	0.12	0.09	-0.30	0.00	-0.83	0.13
Q3	0.06	0.17	-0.28	0.17	-0.53	0.09
Q4	0.18	0.17	-0.24	0.20	-0.49	0.14
Q5	0.02	0.23	-0.21	-0.06	-0.37	0.12

These results are obtained by increasing the first moment of the distribution of offered wages by 10%.

A.2.1 Model fit

Figures 3 to 5 display the in–sample predictions for hours worked, wages, and consumption.[34] The fit of hours worked is rather good, except for single females for whom unemployment is overpredicted and full–time employment underpredicted. The wage rates of males are simulated more accurately than those of females, which are a slightly too high for single females and too low for females in couples. The fit of consumption is acceptable, with small deviations around the modus of each distribution.

One might argue, however, that simulating behavioural transitions — as we do in our policy simulations — is more robust to model misfit than these uncalibrated in–sample predictions (see Section A.3).

A.3 Simulation procedure

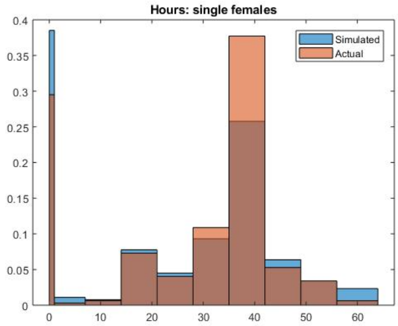
To simulate households' behavioural responses to a tax reform, we make use of the calibration procedure proposed by Duncan and Weeks (2000). This simulation method accounts for both the randomness in the unobserved term of the utility function and the transitional dependence on the observed baseline alternative. Intuitively, this dependence is implied by the modelling assumption that the unobserved terms are left unaltered by the tax reforms under consideration. The procedure is essentially a Monte Carlo approximation and consists of the following three steps:

1. *Calibration*: We draw K distinct vectors of random terms from the Extreme Value Type I distribution, such that the observed alternative is optimal for each household. To avoid the heavy computational burden a brute force search would entail, we draw the random terms from conditional choice distributions, similar to Bourguignon et al. (2001).

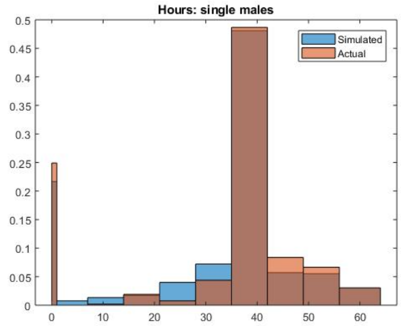
2. *Calculation*: We calculate the effects of a tax reform for each vector of random terms. This yields K different statistics, S_k, of budgetary or distributional effects, as households may find different alternatives optimal after the reform, depending on the specific realisation of the random terms.

3. *Averaging*: We obtain a point estimate for each statistic of interest, S, by calculating the average across these K simulations: $S = \sum_{k=1}^{K} S_k/K$.

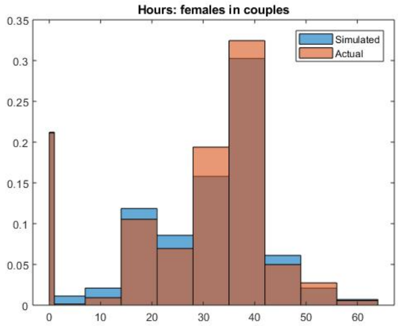
[34] In contrast to our policy simulations, we do not calibrate households' baseline choices to assess the in–sample fit, as this procedure would always entail a perfect correspondence with the data (see also Section A.3).



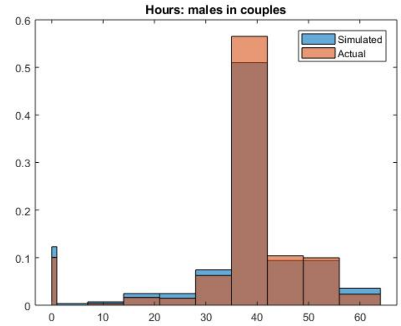
(a) Single females



(b) Single males

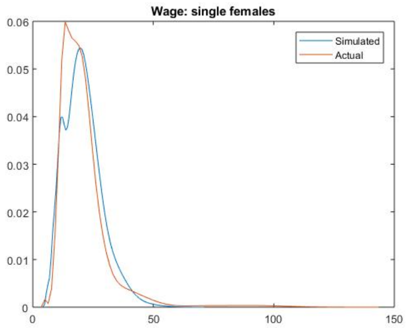


(c) Females in couples

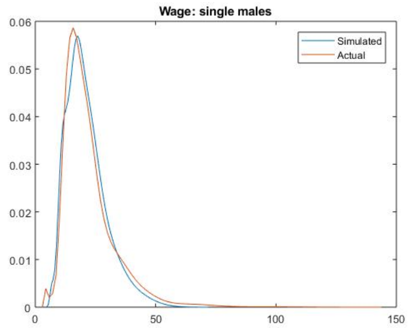


(d) Males in couples

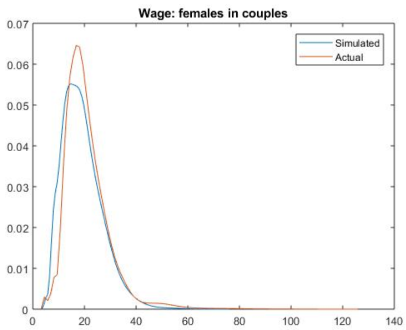
Figure 3: Fit of hours worked (hours/week)



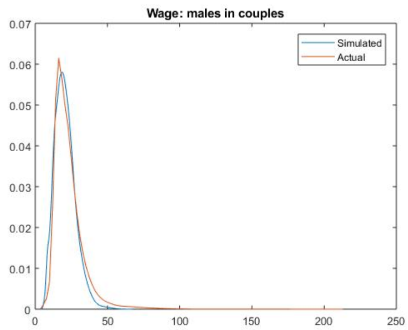
(a) Single females



(b) Single males

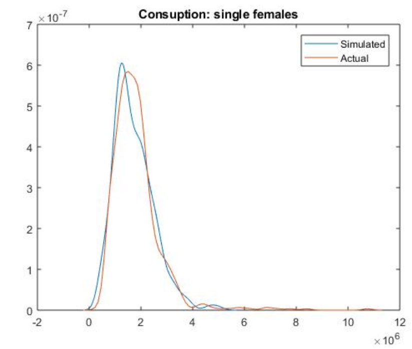


(c) Females in couples

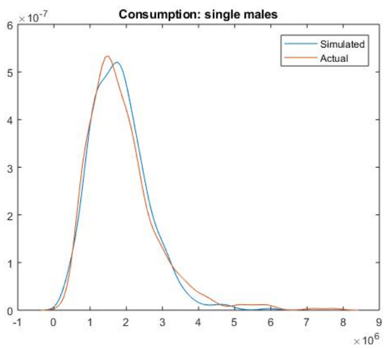


(d) Males in couples

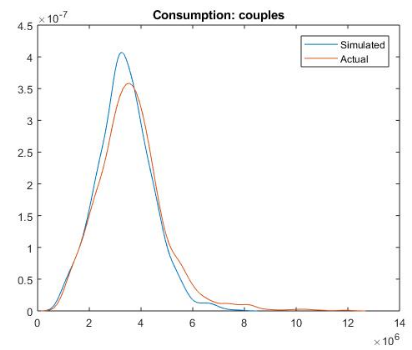
Figure 4: Fit of wage rates (euro/hour)



(a) Single females



(b) Single males



(c) Couples

Figure 5: Fit of consumption (log consumption)

A.4 The Belgian tax shift

We illustrate the framework laid out above by means of a recently enacted Belgian tax shift. This multi-year tax reform was conceived by the federal government which took office in October 2014, and is being rolled out progressively between 2016 and 2020. The reform's principal aim is to alleviate taxes on labour by shifting part of the burden onto non–labour sources. As such, a large portion of the tax shift is to be financed through adjustments to the indirect tax system, like a harmonisation of VAT rates and hikes in excise taxes.

The measures that aim to reduce the tax burden on labour can be divided into two broad categories: those that affect personal income taxation (PIT) and those that affect social security contributions (SSC). The former consists in the abolishment of the 30% bracket and an increase in the upper limit of the 25% and 40% brackets (see Table A.6). The tax free allowance is raised significantly from 7,070 or 7,350 euro in the pre-reform baseline scenario to 8,261 euro in the post-reform scenario.[35] Moreover, the standard deduction of professional expenses for employees is increased and simplified. The current regressive rates are replaced by a uniform rate of 30%.[36] The maximum deductible amount is increased from 3,950 euro in the baseline scenario to 4,495 euro in the post–reform scenario. In addition, the existing in–work tax credits for employees with low earnings are further intensified. These tax credits, however, are capped at 750 euro per year.

Table A.6: Reform of the PIT tax schedule

Tax bracket	Rate (%)	Upper limit of annual net taxable income (euro)	
		Pre–reform	Post–reform
1^{st} bracket	25	8,680	12,360
2^{nd} bracket	30	12,360	abolished
3^{rd} bracket	40	20,600	21,818
4^{th} bracket	45	37,750	37,750
5^{th} bracket	50	∞	∞

Social security contributions are altered for both employees and employers as well as for the self–employed. For employees, the thresholds are increased below which there is a cap (for low–income earners) or a reduction in social security contributions. For employers, the general SSC rate is lowered from 33% to 25%. Meanwhile, the system of structural labour cost reductions, which was intended to boost economic growth is curbed; as such, the effective decrease in employer SSC may be less pronounced than what the drop in rates might suggest. Finally, for the self–employed, the rate applied to annual net earnings below

[35] The tax free allowance is the amount of income that is exempt from taxation. The associated reduction approximately equals the amount of taxes payable when the allowance is applied to the tax schedule as described in Table A.6. The allowance increases with each dependent child as well as for a number of specific household types.

[36] In the pre–reform scenario, the rates are respectively 28.7%, 10%, 5%, and 3% for increasing brackets of gross taxable professional income.

55,577 euro is decreased from 22% to 20.5%.

The PIT and SSC reductions outlined above are financed to a large extent through increases in VAT rates (e.g. electricity) and excise tax hikes (e.g. alcohol, tobacco, and diesel fuel). Other sources that finance the Belgian tax shift are various increases in capital income and corporate income taxes. Nevertheless, the National Bank of Belgium (2015) and the Federal Planning Bureau (2015) estimate the cumulative financing gap of the Belgian tax shift to amount to 6.6 billion euro by 2021 (1.3 % of the projected GDP).

A.5 Integration of the demand side of the labour market

To avoid computationally burdensome iterations, we use an approximate procedure to assess the impact of a reform in employer SSC on labour demand. In this approach we assume that gross wages and job opportunities shift proportionally to those in a competitive labour market. This allows us to derive simple closed form expressions based on elasticities.

Competitive labour market Suppose the labour market is in an equilibrium, possibly after a reform in the PIT and employer SSC. The impact of a change in employer SSC, τ_c, on labour demand h^D and supply h^S can then be written as

$$\begin{aligned} \frac{\Delta h^S}{h^S} &\approx \varepsilon^S \left[\frac{\Delta \bar{w}}{\bar{w}}\right] \\ \frac{\Delta h^D}{h^D} &\approx \varepsilon^D \left[\frac{\Delta \bar{w}}{\bar{w}} + \frac{\Delta \tau^c}{1+\tau^c}\right], \end{aligned} \tag{39}$$

in which ε^D is the elasticity of demand (in terms of the wage cost) and ε^S is the elasticity of supply (in terms of the gross wage rate). If one assumes the labour market to be again in equilibrium after the change in employer SSC occurred, one finds that

$$\begin{aligned} \frac{\Delta \bar{w}}{\bar{w}} &\approx \frac{\varepsilon^D}{\varepsilon^S - \varepsilon^D} \frac{\Delta \tau^c}{1+\tau^c} \\ \dot{\bar{w}} &\approx \frac{\varepsilon^D}{\varepsilon^S - \varepsilon^D} \dot{\tau^c} \\ \\ \frac{\Delta \bar{h}}{\bar{h}} &\approx \frac{\varepsilon^S \varepsilon^D}{\varepsilon^S - \varepsilon^D} \frac{\Delta \tau^c}{1+\tau^c} \\ \dot{\bar{h}} &\approx \frac{\varepsilon^S \varepsilon^D}{\varepsilon^S - \varepsilon^D} \dot{\tau^c}. \end{aligned} \tag{40}$$

These two expressions capture the relative change in equilibrium gross wages $\bar{w}$ and employment $\bar{h}$ as a simple function of demand and supply elasticities.

Link with the RURO model In the RURO model, labour demand and supply are generally not in equilibrium due to market frictions. We assume, however, that the relative change in gross wages and employment after the reform is similar to that of the equilibrium in the competitive labour market. To close the model, we further assume that there is a proportional relationship between offered gross wages w and equilibrium gross wages $\bar{w}$, and between the intensity of opportunities θ and the equilibrium level of employment $\bar{h}$,

$$\begin{aligned} w &= \gamma \bar{w} \\ \theta &= \delta \bar{h}, \end{aligned} \tag{41}$$

such that the constants γ and δ drop out when considering relative changes:

$$\begin{aligned}\dot{w} &= \dot{\bar{w}}\\ \dot{\theta} &= \dot{\bar{h}}.\end{aligned} \tag{42}$$

Implementation We implement the changes in both gross wages and the intensity of opportunities at the individual level. Let i denote an individual of type s and j a choice alternative available to that individual. First, we calculate $\dot{\tau^c}$ for each individual and choice alternative:

$$\dot{\tau^c}_{ij} = \frac{(\Delta\tau^c_{ij} w_{ij} h_{ij})}{(\tau^c_{ij} w_{ij} h_{ij}) + (w_{ij} h_{ij})}. \tag{43}$$

Next, we calculate the change in gross wages and opportunities offered for each individual and choice alternative:

$$\begin{aligned}\dot{w}_{ij} &= \frac{\varepsilon^D_s}{\varepsilon^S_s - \varepsilon^D_s}\dot{\tau^c}_{ij}\\ \dot{\theta}_{ij} &= \frac{\varepsilon^S_s \varepsilon^D_s}{\varepsilon^S_s - \varepsilon^D_s}\dot{\tau^c}_{ij}.\end{aligned} \tag{44}$$

These changes are then implemented in a new simulation run.

The elasticities that we employ are type specific and take the values as described in Table A.7. The supply–side elasticities are calculated at the extensive margin, using our estimated job choice model.[37] The demand–side elasticities are borrowed from an OLG model for Belgium by Devriendt and Heylen (2017). In Tables A.8 and A.9 we present the impact of a 1 percent decrease in employer SSCs on offered gross wages and opportunities, respectively. The raise in gross wages is increasing in individuals' education level and is the largest for males living in couples and the smallest for single females. The effect runs in the exact opposite direction when the impact on opportunities is considered: that is, individuals with low education experience the biggest increase in opportunities.

[37]These elasticities are calculated by increasing the first moment of the distribution of offered wages by 10%.

Table A.7: Elasticities to model demand–side feedback

	Low education	Middle education	High education
Demand	-1.22	-1.12	-1.04
Supply males in couples	0.50	0.21	0.13
Supply females in couples	0.95	0.34	0.22
Supply single males	0.84	0.56	0.32
Supply single females	0.59	0.38	0.38

Source: The demand–side elasticities are borrowed from Devriendt and Heylen (2017). We calculate the supply elasticities at the extensive margin by increasing the first moment of the distribution of offered wages by 10%.

Table A.8: Growth in gross wages induced by a 1% decrease in employer SSCs (%)

	Low education	Middle education	High education
Males in couples	0.71	0.84	0.89
Females in couples	0.56	0.77	0.82
Single males	0.59	0.67	0.76
Single females	0.68	0.74	0.73

These impact effects are obtained by combining the elasticities from Table A.7 as laid out in Equation (44).

Table A.9: Growth in opportunities induced by a 1% decrease in employer SSCs (%)

	Low education	Middle education	High education
Males in couples	0.35	0.18	0.11
Females in couples	0.53	0.26	0.18
Single males	0.50	0.37	0.25
Single females	0.40	0.29	0.28

These impact effects are obtained by combining the elasticities from Table A.7 as laid out in Equation (44).

A.6 Comparison across welfare measures

Throughout the paper we discuss three criteria to measure the change in welfare induced by tax–benefit reforms: (i) a measure based on the equivalent consumption; (ii) a measure based on the equivalent variation; and (iii) a measure based on the difference in real disposable income. In this section, we argue that comparing differences in welfare across these measures is often unwarranted, as they have very different underlying cardinal properties.

Different cardinal properties For those not working both in the pre– and post–reform situation, the change in welfare according to the equivalent consumption criterion is equal to the change in real disposable income,

$$\Delta W_{EC} = \frac{y_1}{Q\left(\mathbf{q}_1\right)} - \frac{y_0}{Q\left(\mathbf{q}_0\right)}, \tag{45}$$

which was shown to be a rescaling of the change in the exact EV–based measure (see Section 3.2.1). The comparison of the EV–based with the EC–based measure is problematic, as the former is expressed in nominal baseline monetary units, while the latter is in real monetary (i.e. consumption) units.

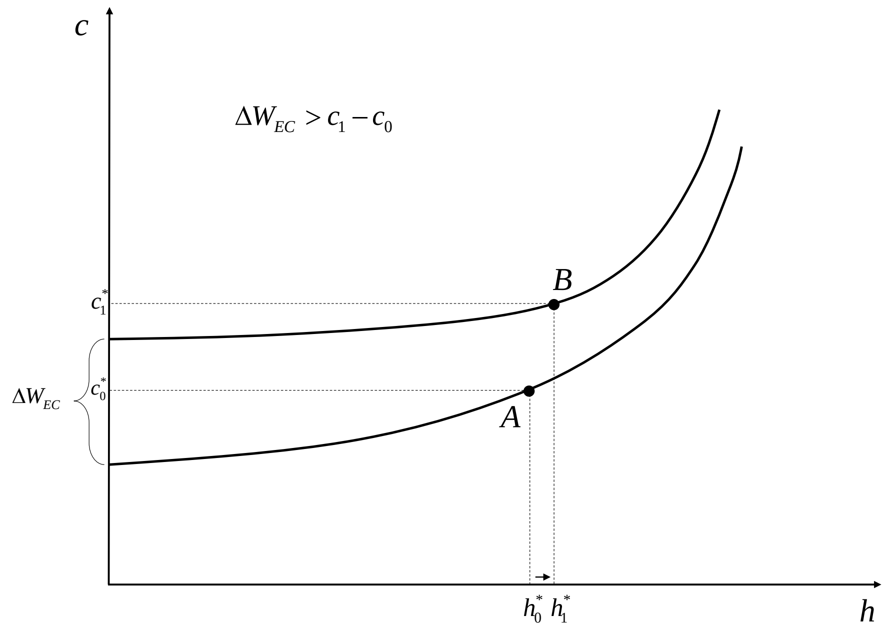
For individuals for who it is optimal to keep labour supply fixed, the difference in real disposable income has a welfare interpretation. That is, the sign of the change in real disposable income coincides with that of ΔW_{EC} and the exact ΔW_{EV}. More general relations between these measures depend heavily on individuals underlying preferences.

For those who switch to jobs with different working hours, there is no obvious connection between the welfare gain according to the EC–criterion and the gain in real disposable income. That is, although the EC–criterion values an individual's leisure time explicitly, an increase in working hours does not necessarily imply that the change in this criterion, ΔW_{EC}, is smaller than that in real income, $c_1 - c_0$. The reverse also holds for a decrease in working hours. The reason is that these measures, while both being expressed in the same real monetary terms (i.e. units of consumption), refer to different cardinalisations of welfare.[38] Figure 6 shows that the comparison depends crucially on the shape of the indifference curves and thus on individuals' preferences.

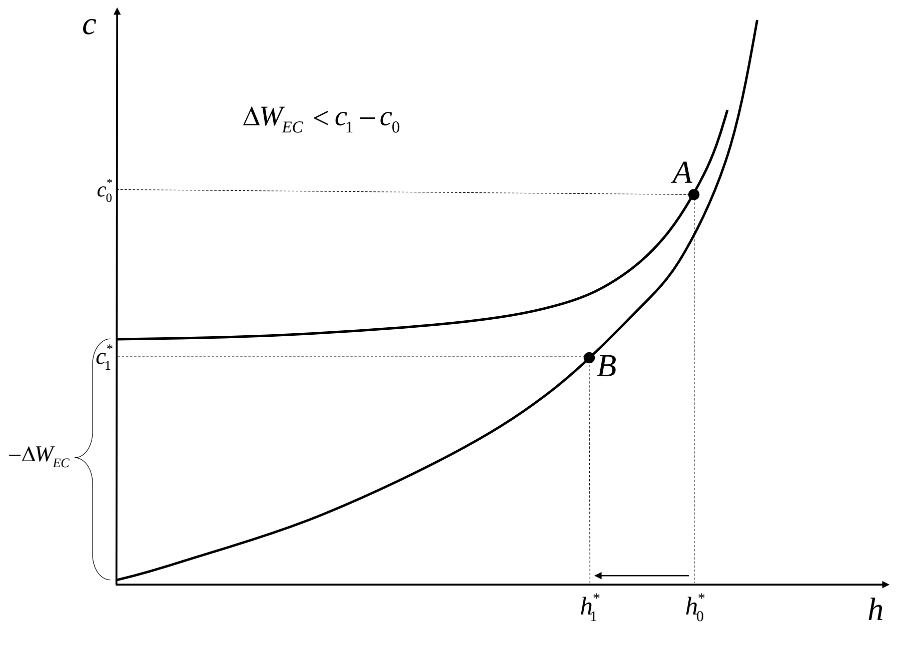
Note, however, that the gain in welfare according to the EC–criterion is always lower than the increase in real income for those entering the labour market in response to the reform. Similarly, for those exiting the labour market, the welfare loss is always less than the loss in real income.

In addition, if real income at the pre–reform job choice increases due to the mere mechanical effect of a tax reform (but at fixed expenditure shares), then the reform cannot

[38] Admittedly, in this case the difference in real disposable income is not an appropriate welfare measure: it can be positive while agents actually lose welfare, or vice versa. The same holds, however, when comparing the EC–based measure with a EV–based measure that takes into account changes in leisure, the latter defined as the difference in expenditure functions at a fixed set of reference prices including the wage rate.



(a) Increase in labour supply



(b) Decrease in labour supply

Figure 6: Change in welfare according to equivalent consumption and real disposable income

incur a welfare loss according to the EC–criterion. In practice, this does not hold for 5% of the RURO households.

This discrepancy is explained by random taste variation that is inherent to the RURO model. When defining the equivalent consumption criterion in Equation (36), only the systematic part of the utility function, $H(c_j, h_j)$, was taken into account. In the RURO model, however, the relative desirability of an alternative j, say U_j, is also partly determined by an additive random component, ε_j:

$$U_j = H(c_j, h_j) + \varepsilon_j. \tag{46}$$

The randomness embodies the relative desirability of unobserved pecuniary and non–pecuniary attributes of the alternative, be it a market or non–market choice option. This might explain why we observe some agents to lose welfare, while overall welfare — taking into account the desirability of non–observed attributes of the new situation — actually improved.

A.7 Additional results

Table A.10: Employment effects for the revenue neutral simulation (TS3) across deciles of equivalised disposable income

deciles	whole RURO subpopulation				RURO subpopulation working in baseline					
	participation		labour supply		labour supply		gross wage		earnings	
	basel.	Δ	basel.	Δ	basel.	Δ	basel.	Δ	basel.	Δ
	%	% pts	hours/week		hours/week		€/hour	%	€/month	%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	11.5	14.2	2.9	5.3	25.7	1.8	12.6	2.1	1379	12.7
2	38.1	5.8	12.2	2.1	32.3	0.2	14.7	0.2	2023	1.3
3	62.5	2.7	21.1	0.9	33.8	0.0	15.5	-0.2	2264	-0.4
4	79.9	1.8	27.7	0.5	34.8	-0.1	16.0	-0.5	2363	-0.6
5	88.4	1.1	31.5	0.3	35.7	-0.1	17.3	-0.9	2619	-1.2
6	95.1	0.4	34.5	0.0	36.2	-0.2	19.0	-1.2	2918	-1.5
7	96.7	0.3	35.2	0.0	36.4	-0.1	19.7	-1.1	3024	-1.4
8	97.9	0.1	36.7	-0.1	37.5	-0.2	21.2	-1.8	3373	-2.1
9	97.1	0.1	38.3	-0.5	39.4	-0.6	23.5	-2.4	3962	-3.4
10	97.2	-0.2	40.0	-0.9	41.1	-1.1	30.6	-6.6	5408	-7.6
all	81.3	2.1	30.3	0.5	37.3	-0.4	21.5	-2.6	3477	-3.4

The variables are calculated at the individual level for the subpopulations mentioned in the top row. Each individual is allocated to the decile based on household equivalised disposable income for the *whole* population (with an equal number of persons by decile). 'Participation' in column (1) is calculated as the ratio of the number of persons working in the baseline, divided by the total population of individuals modelled in the RURO model. The average number of hours worked per week in column (3) is also calculated for the whole RURO subpopulation, irrespective whether in the baseline the individual was working or not. Columns (5), (7), and (9) are averages for the RURO subpopulation of individuals who are working in the baseline. For all columns we keep the allocation of individuals across deciles fixed (as described in the first sentence of this note).

Table A.11: Distributional effects for all simulation variants — whole population; first order effect (= without behavioural changes) on disposable income minus indirect taxes

deciles	changes in €/month and per household					changes in % of basel. disp. inc.				
	TS1	TS2	TS3	TS3b	TS4	TS1	TS2	TS3	TS3b	TS4
1	8	-7	-45	-38	-7	0.6	-0.6	-3.7	-3.1	-0.6
2	38	22	-24	-17	22	2.3	1.3	-1.4	-1.0	1.3
3	69	53	7	13	53	3.9	3.0	0.4	0.7	3.0
4	102	86	32	37	86	4.7	3.9	1.5	1.7	3.9
5	117	102	42	46	102	4.7	4.0	1.7	1.8	4.0
6	144	127	59	62	127	4.8	4.2	2.0	2.0	4.2
7	165	148	77	78	148	5.0	4.5	2.3	2.4	4.5
8	173	155	78	79	155	4.6	4.2	2.1	2.1	4.2
9	169	151	68	68	151	4.0	3.6	1.6	1.6	3.6
10	163	145	47	46	145	2.9	2.6	0.8	0.8	2.6
all	111	94	31	35	94	3.9	3.3	1.1	1.2	3.3

All effects are with respect to the baseline simulation. The deciles are constructed by ranking all individuals on the basis of their household equivalised disposable income in the baseline. Each decile contains 10% of the total population of individuals. The values in the left part of the table show averages of changes in *non–equivalised* household disposable incomes minus indirect taxes, where the averages are calculated over the population of households. The right part of the table expresses these changes as a % of household disposable income in the baseline.

TS1 constitutes the PIT and SSC components of the reform as decided by the Belgian government. TS2 augments TS1 with the reform in indirect taxes and closely resembles the actual tax shift. TS3 augments TS1 with a much larger increase in VAT, to move closer to revenue neutrality. TS3b is reform TS3 but in which we keep the wages fixed. TS4 equals TS2, but with additional effects from the demand side of the labour market by translating the decrease in the social security contributions of the employer into an increase in the availability of market opportunities, and a change in the position of the distribution of offered wages.

Table A.12: Distributional effects for all simulation variants — whole population; total effect including behavioural changes

deciles	changes in W_{EV} in €/month per household					changes in W_{EV} in % of basel. disp. inc.				
	TS1	TS2	TS3	TS3b	TS4	TS1	TS2	TS3	TS3b	TS4
1	127	104	47	45	132	10.3	8.4	3.8	3.6	10.7
2	75	58	5	6	68	4.4	3.5	0.3	0.4	4.0
3	81	65	15	20	70	4.6	3.6	0.9	1.1	3.9
4	110	94	37	41	97	5.0	4.3	1.7	1.9	4.5
5	121	105	42	47	106	4.8	4.1	1.6	1.9	4.2
6	138	121	49	56	122	4.6	4.0	1.6	1.9	4.0
7	159	141	63	70	141	4.8	4.2	1.9	2.1	4.3
8	155	137	50	69	139	4.2	3.7	1.3	1.8	3.7
9	136	116	12	47	123	3.3	2.8	0.3	1.1	2.9
10	65	42	-121	20	61	1.2	0.7	-2.2	0.4	1.1
all	111	94	31	35	104	3.9	3.3	1.1	1.2	3.7

All effects are with respect to the baseline simulation. The deciles are constructed by ranking all individuals on the basis of their household equivalised disposable income in the baseline. Each decile contains 10% of the total population of individuals. The values in the left part of the table show averages of changes in W_{EV}, where the averages are calculated over the population of households. The right part of the table expresses these changes as a % of household disposable income in the baseline.
TS1 constitutes the PIT and SSC components of the reform as decided by the Belgian government. TS2 augments TS1 with the reform in indirect taxes and closely resembles the actual tax shift. TS3 augments TS1 with a much larger increase in VAT, to move closer to revenue neutrality. TS3b is reform TS3 but in which we keep the wages fixed. TS4 equals TS2, but with additional effects from the demand side of the labour market by translating the decrease in the social security contributions of the employer into an increase in the availability of market opportunities, and a change in the position of the distribution of offered wages.

Table A.13: Distributional impact of revenue neutral simulation TS3 for the RURO subpopulation working in the baseline. Comparison of changes in real disposable income and W_{EV} across deciles of gross wages

deciles	baseline	changes in €/month			changes in %		
of		1st order	total effect (incl. behav.)		1st order	total effect (incl. behav.)	
gross wage	disp. inc. €/month (1)	real disp. inc. (2)	real disp. inc. (3)	W_{EV} (4)	real disp. inc. (5)	real disp. inc. (6)	W_{EV} (7)
1	1,749	41	60	76	2.7	3.9	4.4
2	1,748	41	52	69	2.7	3.4	3.9
3	1,833	41	41	57	2.6	2.6	3.1
4	1,992	36	32	48	2.1	1.8	2.4
5	2,135	26	17	33	1.4	0.9	1.5
6	2,297	20	-1	15	1.0	0.0	0.6
7	2,370	18	-10	5	0.9	-0.5	0.2
8	2,494	13	-20	-5	0.6	-0.9	-0.2
9	2,753	1	-65	-53	0.1	-2.7	-1.9
10	3,437	-23	-197	-196	-0.8	-6.6	-5.7
all	2,263	22	-6	8	1.1	-0.3	0.3

The variables are calculated at the individual level: for couples we divide real disposable income and W_{EV} by two. Each individual of the RURO population is allocated to the decile based on his or her gross wage (with an equal number of persons by decile). The non–working individuals are assigned a gross wage based on the estimated wage equation of the RURO model. Columns (2) and (3) show differences in real incomes: $\frac{y_1}{Q(\mathbf{q}_1)} - \frac{y_0}{Q(\mathbf{q}_0)}$ (see Subsection 3.2.1 for its relation with a change in welfare). Column (4) gives the RHS of equation (32). Column (2) shows the effects without change in behaviour. Columns (3) and (4) are effects including changes in labour supply and consumption. All percentages in columns (5) and (6) are with respect to real disposable income in the baseline: $\frac{y_0}{Q(\mathbf{q}_0)}$. Finally, column (7) is with respect to disposable income in the baseline.