KU LEUVEN

CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS15.25

OCTOBER 2015

Getting tired of work, or re-tiring in absence of decent job opportunities?

Bart CAPÉAU & André DECOSTER

Public Economics

Faculty of Economics
And Business

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.

Bart Capéau & André Decoster†

October 2015

Abstract

We report estimates of a version of the Random Utility and Random Opportunity (RURO) model of job choice (Aaberge, Dagsvik and Strøm, 1995, and Aaberge, Colombino and Strøm, 1999) on Belgian data (SILC 2007). The paper exploits the distinction between preference factors and individual differences in opportunities in such a RURO–model, in order to obtain a more nuanced picture of labour market participation of the elderly. More specifically, the estimated model allows to give a quantitative answer to the question to what extent lower labour market participation of elderly is due to changing preferences (executing a job might become less enjoyable with age) or differences in opportunities (elderly getting less, or less attractive job offers).

† CES–KULeuven, Naamsestraat 69, B3000 Leuven, Belgium,
e–mail: bart.capeau@kuleuven.be, andre.decoster@kuleuven.be.

We are grateful to Rolf Aaberge and Ugo Colombino for an introduction into and many discussions on the RURO–model, to Tom Wennemo for technical assistance with the estimation program and to Bart Cockx and Gijs Dekkers for comments on earlier versions of the paper. Pieter Vanleenhove, Toon Vanheukelom and Rembert De Blander prepared the data from EU-SILC 2007, and used EUROMOD version 5.5 to calculate net incomes for the choice sets. We acknowledge financial support from the BELSPO–projects MIMBEL and BEL–AGEING. The usual disclaimer applies.

1 Introduction

In the present paper we investigate the impact of ageing on job choice behaviour, using a structural discrete choice model. Existing literature has concentrated on the early exit of elderly from the labour market. Emphasis has been put on the role of financial incentives in the social security system to leave or remain active in the labour market. See for example Gruber and Wise (1999) for an overview of these issues in different OECD countries. To that purpose, the *net social security wealth*, that is the present value of future benefits *minus* the present value of social security contributions, and the accrual of that wealth by postponing retirement or exit from the labour market through other channels such as special early leaver schemes, were used as indicators.[1] Later on, the financial implications of such social security schemes were included in a utility based decision model (Stock and Wise 1990).[2] The Stock and Wise approach boils down to the question whether continuing to work one year longer would yield a higher expected utility than retiring now. This departs from a fully forward looking dynamic framework of maximizing expected utility, in which *at each point in time* an optimal decision is taken, for each possible value of future possible states of the world, and expectations are taken over the possible evolution of the state of the world through time. While such fully forward looking models were developed to study the retirement decision (Rust, 1989, Berkovec and Stern, 1991, Rust and Phelan, 1997), they have been criticised for their computational complexity in exchange for a poor gain in predictive validity (Lumsdaine *et al.* 1992, Belloni 2008).

Our point of departure is somewhat different from these studies, in that we do not model the exit decision, but concentrate on the labour market participation decisions of those who decide to continue to be active on the labour market. That is those who are looking for work or are working (as an employee). To set the scene, we summarise in the next table the labour market participation figures and mean labour time of the subjects in the sample that we will use in this paper, by age category (more specifically, 30–49 years compared to 50–64 years). Our sample consists of single males and females, and males and females living together as a couple, who either execute a job as employee, be it full or part time, or are available on the labour market to execute such a job, if they would find a suitable offer.[3]

[1] The use of social security wealth as an indicator for financial incentives of the social security system, can at least be traced back to Feldstein (1974).

[2] See Maes (2011, 2012) and Lefebvre and Orsini (2012) for applications to Belgium. Maes uses the net social security wealth and its accrual as inputs in the decision model, while Lefebvre and Orsini use simulated paths of future wages and benefits. Applications to Italy were made by Belloni and Alessie (2009, 2013).

[3] This excludes self–employed, early retired persons and persons receiving a sickness or invalidity benefit. As such, the definition of participation rate we will use in this paper, that is the number of employed persons relative to those employed and those available for executing a job, differs from both the *employment*

The drop in labour market participation from those aged 30 to 49 years old to those aged 50 to 64 years old ranges between 12 (males in couples) and 31 percentage points (females in couples). The time per week a person spends on jobs drops, comparing the same two age categories, by almost six to more than ten hours per week.

	labour market participation (%)		labour time mean hours per week	
Age Category	30–49	50–64	30–49	50–64
Males in couples	96%	84%	39.6	33.8
Females in couples	85%	54%	27.0	16.5
Single males	85%	62%	33.7	24.6
Single females	73%	58%	26.6	19.7

Source: own calculations on the basis of EU–SILC 2007 Belgium.

We will investigate whether this lower participation of elderly is a matter of preferences (increasing intensity of preference for leisure with age), or whether less jobs are available, suitable for the capacities of elderly. Thereto a model is required which, contrary to the classical discrete choice models of labour supply (Van Soest, 1995), will allow for the impact of both labour supply and labour demand side effects.

More specifically, we used a Random Utility/Random Opportunity (RURO) model of job choice (see amongst others Dagsvik and Strøm, 1992, 2003, Dagsvik, Aaberge and Strøm, 1995, and Aaberge, Colombino and Strøm, 1999).[4] These type of models, first developed in an abstract setting by Dagsvik (1994), try to understand individual heterogeneity in choice behaviour as a combined effect of preference differences and differences in opportunities.

The RURO model is not the only one that embodies restrictions on the choice set into a labour supply model (see, amongst others, Altonji and Paxson 1982, 1992, Van Soest, Woittiez and Kapteyn 1990, Tummers and Woittiez 1991, Dickens and Lundberg 1993, Bloemen 2000, 2008, Ham and Reilly 2002, and Beffy *et al.* 2014). It must be added that the inclusion of dummies for part time and full time work in the discrete choice model of Van Soest (1995), to improve the fit, is in fact a simplified reduced form approach of earlier work with Woittiez and Kapteyn (Van Soest, Woittiez and Kateyn, 1990) which models hours restrictions more explicitly. But the RURO model is the first one that derives these restrictions from an explicit model of a job arrival process, and stresses the individual heterogeneity, both observed and

and *activity* rate in EUROSTAT–statistics. The former is the number of persons working relative to total population or the relevant sub–population, while the latter refers to those working and looking for a job, compared to the total or relevant sub–population.

[4] See also the recent overviews by Aaberge and Colombino (2014), and Dagsvik, Jia, Kornstad, and Thoresen (2014).

unobserved, of the availability of job offers suitable to the capacities of individual agents. This makes this model specially apt to handle the research question we envisage.

Preferences are capturing the extent to which an individual is willing to trade–off leisure for consumption. It may deserve recommendation to take into account also other aspects influencing the choice between alternative leisure activities and available jobs, such as social relations involved, challenge of the tasks, security and health, recognition, and societal relevance. These factors are however not easily observable by the analyst. It was one of the contributions of the development of probabilistic choice and random utility models, as developed by respectively Luce (1959) and McFadden (1973), to integrate these additional determinants of preferences as a non–systematic element, affecting the utility obtained from different available alternatives. By non–systematic it is meant that such factors cannot be traced back to observable characteristics. Random utility models have been applied to labour supply behaviour (Van Soest, 1995) since. They replaced the traditional continuous choice approach to labour supply behaviour (see Hausman, 1985, for a review of the traditional approach), which faced difficulties in deriving tractable closed form solutions of labour supply functions, in the presence of non–linear budget sets. Indeed, many personal income tax systems, such as *e.g.* a minimum income guarantee associated with a linear earned income tax, create non–convexities in the budget set (the available bundles of consumption and labour time a person can chose from), leading to discontinuities and non–uniqueness of the optimal choice in function of wage variations. These phenomena are more easily treated in a discrete choice set–up, which is the approach taken by both, probabilistic choice and random utility models.

The random utility model is however still limited in scope. Interindividual differences in the availability of alternatives from which a person can choose, are exogenous to the model. Applied to job choice, differences in individual budget sets stem exclusively from wage differences and differences in unearned income. In a static model, it is indeed reasonable to assume that unearned income differences are exogenous, and do not depend on individual choices. But in standard random utility models also wages are exogenously fixed individual characteristics, reflecting a person's productive capacities.[5] For several reasons, this is unattractive. Productive capacities can in many cases not be determined appropriately, when considered separately from the specific job in which these capacities are exhibited. Moreover,

[5] Some contributions do allow for unobserved wage heterogeneity, see *e.g.* Van Soest, Das and Gong (2002), Löffler *et al.* (2013), and the second model discussed in Dagsvik and Jia (2014). It is fair to say that Van Soest (1995) already incorporated the problem of imperfect observation of wages for non–participants in the extended version of his model. Besides, there is an earlier literature accounting for the fact that wages are non–linear in hours (See for example Moffitt 1984). However, none of these treats wages as an object of job choice behaviour.

it is quite unnatural that all available job offers, even when perfectly suited to a person's capacities and skills, would pay the same wage. Furthermore, due to organisational limitations of the production process, and social life, it is highly unlikely that persons can completely freely fix the number of hours they will work.

It is exactly these type of frictions in the choice process which are taken into consideration by RURO models, as an additional factor, next to preferences, to understand choice behaviour. Job offers are considered as packages of wages, labour time regimes, and a number of other attributes (Dagsvik and Strøm, 1992, 2003, Aaberge, Dagsvik and Strøm, 1995, and Aaberge, Colombino and Strøm, 1999). These other attributes (challenge, safety and security, esteem and recognition, appreciation of colleagues, responsibility...) are however difficult to observe, especially to the extent that these are important from the viewpoint of the degree of job satisfaction they can provide to a person. Therefore, the individual specific availability of suitable jobs is thought of as the result of a stochastic process of job offers. The impact of explanatory variables on the intensity with which job offers arrive to a person according to that process, is estimated jointly with individual's preference characteristics. Not only the intensity with which job offers arrive, but also the availability of, according to a person's own judgement, attractive non–market alternatives to spend time, is individually specific. Limited physical abilities might impede someone who likes to walk outside, to do so. Choosing under such circumstances between sitting in front of a liquid crystal screen, reading books, or accepting a job, that person might opt for the latter, while the reverse might happen for someone with similar preferences, but in good physical shape. The RURO–model also allows for individual heterogeneity in restrictions on available labour time regimes, even though the effect of these type of restrictions are difficult to identify from the contribution of preferences.

The present paper exploits the distinction of the RURO model between preference factors and individual differences in opportunities, in order to obtain a more nuanced picture of labour market participation of the elderly. More specifically, by estimating a RURO model (applied to Belgium) we are able to give a quantitative answer to the question to what extent lower labour market participation is due to changing preferences (executing a job might become less enjoyable with age) or differences in opportunities (elderly getting less, or less attractive job offers).

For the interest reader, Section 2 presents a self–contained in depth exposition of our account of the RURO model. Section 3 discusses the derivation of the likelihood function resulting from the model, and explains, also in detail, the estimation method. Some technical issues are relegated to an appendix. The data are presented in Section 4. Section 5 contains the estimation results. In Section 6 we investigate the fit of the estimated model, and its behavioural implications. Finally, Section 7 draws some tentative conclusions with respect to the contribution of preferences and opportunities, in the assessment of the age profiles of labour market participation.

2 The RURO model

In the present section, we present an application of the general RURO model (Dagsvik 1994) to job choice behaviour. This job choice model was developed by Dagsvik and Strøm (1992, 2003, and 2006), Aaberge, Dagsvik and Strøm (1995), Aaberge, Colombino and Strøm (1999), Dagsvik, Locatelli and Strøm (2006, 2007), and Dagsvik and Jia (2014) (see also the the surveys of Dagsvik, Jia, Kornstad and Thoresen, 2014, and Aaberge and Colombino, 2014). We first illustrate how the RURO model extends the choice problem of traditional labour supply models from a question of trading off leisure against consumption towards a model of job choice, against other non–market alternatives of time use (Subsection 2.1). Then, we discuss the preference part of the model (Subsection 2.2). Finally, we expose the modelling of opportunities (Subsection 2.3).

2.1 Opportunities and jobs

In general, the RURO model is an economic model of human choice behaviour. Human decision makers are assumed to choose the *best* element from a set of choice possibilities or opportunities, where 'best' is defined in terms of preferences (or, *vice versa*, preferences are derived from observed choice behaviour as that objective which would be maximised given those choices). Applied to job choice, the set of opportunities is to be thought of as a set of possible activities a particular individual might choose to execute. Some of these activities are rendered available through job offers. These job offers will be indexed by j, where the variable j belongs to an index set, $\mathcal{J}$ say. A job offer stipulates an amount of labour time to be supplied when accepting the offer, say h, and pays a wage, w. It is assumed that this wage can be expressed in units of time effectively spent on the job, so that (gross) revenues earned by the job equal the amount of time spent on the job times the wage.[6]

Gross earned labour income is then equal to wh. Gross earned labour income together with some other characteristics, say $\mathbf{x}_f$, among which unearned gross income (exclusive of

[6] This is generally not the case. Output dependent bonuses or piece–rates do not necessarily bear an obvious relation with time spent on the job. More surprisingly, in a regime with fixed monthly wages, the wage per unit of time is variable, since the number of hours a regular (that is: full time) job requires can differ over jobs, and for a specific job, the number of hours to be worked per month is not fixed. Finally, even if there is a fixed hourly wage, and no bonuses, there is no obvious way how to treat paid holidays. Should they be taken into account when calculating an hourly wage or not? When taking these into account, this would result in an increase of the gross hourly wage, as compared to the wage specified in the contract.

transfers), determine the outcome of the gross to net (disposable) income function $c = f(w, h; \mathbf{x}_f)$, where c stands for consumption which in a static model as the present one coincides with gross disposable income.[7] That is, saving is considered as part of consumption. The function f converts gross income components into net disposable income, by subtracting taxes to be paid and adding transfers and subsidies. Usually, the generation of disposable income is constructed from raw data on gross income, labour time and other characteristics, by means of a microsimulation model.

Besides time spent on the job and the remuneration, jobs exhibit a number of other characteristics such as degree of responsibility, variation and challenge of the tasks, safety, healthiness, physical effort, stress, relation with colleagues and superiors. These characteristics will be denoted by $\mathbf{s}$. Preferences over these non–pecuniary attributes affect job choice.

One might also decline all job offers. Evidently, not executing a formal job does not require any time to be spent on the formal labour market ($h = 0$), and is assumed not to pay a wage ($w = 0$).[8] A person who does not work, receives a net transfer (that is after deducting income taxes to be paid from her replacement income) equal to $f(0, 0; \mathbf{x}_f)$. In that case, time is spent on executing some of the available non–market opportunities. However the set of activities[9] one has alternatively available is not the same for all individuals, neither is the extent to which a particular alternative is available. When living in a small town, attending concerts, theatre or visiting museums is certainly not as easy as for big city dwellers. If you are in a wheel chair, hiking is not an option. Which of the available non–market activities will be chosen, in case no job offer is accepted, again depends on preferences (or, *vice versa*, what one chooses to do allows to derive something on the shape of preferences that person supposedly has). Non–market alternatives will be indexed by i, belonging to the index set $\mathcal{I}$. The index set for jobs and non–market alternatives respectively, are disjoint: $\mathcal{I} \cap \mathcal{J} = \varnothing$. We will also use the index variable z to indicate an alternative in general, that is either a job or a non–market alternative. So, $z \in \mathcal{Z} := \mathcal{I} \cup \mathcal{J}$. To be really precise, and index z refers to a set of activities. If this set involves one or more jobs, it the index z will belong to $\mathcal{I}$, while it belongs to $\mathcal{J}$ otherwise. So, an alternative z involves a *set* of activities. We will be somewhat sloppy in the sequel however, by calling an alternative involving one or more jobs in the formal labour market, sometimes simply a job.

[7] We include the gross wage, and the number of hours worked as separate arguments in that function, as some aspects of the tax system, such as the Belgian work *bonus* may depend on the wage, rather than on labour income, wh. We are however aware that this might cause problems for the non–parametric identification of the RURO model.

[8] How to treat informal jobs in this type of framework remains largely an unresolved question.

[9] The word 'activity' will be used here in a broad sense, including occupations which are not very 'active' such as sleeping and day dreaming. A certain type of agency or control is however presumed, since otherwise it would be difficult to talk about choice behaviour.

2.2 Random utility

In the model, preferences are defined over the number of hours h spent on jobs (which is zero if one chooses not to accept any job offer), consumption, c, and a set of other attributes, say $\mathbf{s}$, that a job or certain non–market activities possess, and that a person might care for. These other attributes are not observed by the researcher.

The observable, and thus from a behaviour theoretic point of view relevant, bundle of characteristics an alternative $z \in \mathcal{Z}$ exhibits, is denoted by $(C(z), H(z))$, where $C(z)$ refers to the individually specific net disposable income associated with executing activities involved by z, and $H(z)$ to the labour time involved by the set of activities involved by z. The utility derived from these observable characteristics is denoted by $V(C(z), T - H(z); \mathbf{x}_V)$, where $\mathbf{x}_V$ are the specific values of a set preference shifters for the individual under consideration, and T denotes the number of time units available in the period over which labour time h is registered (*e.g.* 168 hours a week, if labour time is expressed in hours worked per week). Alternatively, one may define preferences over consumption and leisure. The latter, say ℓ, is equal to the time left over for non–market activities, after subtracting labour time, $\ell := T - h$. It is assumed that the econometrician can derive some evidence on the shape of the function V on the basis of observations on $(c, T - h)$ and $\mathbf{x}_V$. So, no individual preference differences apart from those explained by observable characteristics $\mathbf{x}_V$, are allowed for in this part of the utility function, and V is therefore called the *systematic* part of the utility function, and, hence, of preferences.

Since the other attributes besides disposable income and labour time are not observable, their contribution to utility will be specified as a random term. Thus, when a set of activities z bears attributes $\mathbf{s} = s(z)$, the utility these attributes generate, is denoted by the random variable $\varepsilon(s(z))$. So, each individual derives a specific utility from an activity z with attributes $\mathbf{s} = s(z)$, and this utility is considered as the realisation of a random variable. Different realisations of these terms (one for each value $\mathbf{s}$ could take) thus incorporate unobserved individual preference heterogeneity. It is assumed that this utility from non–pecuniary attributes, $\varepsilon(s(z))$, enters overall utility of an alternative z, in a multiplicatively separable way from the systematic part of the utility function. In order to make sense, this requires both, the systematic part of the utility, and the random term, to be non–negatively valued functions.

In summary, the total utility derived from picking an alternative $z \in \mathcal{Z}$, denoted by $U(C(z), H(z), s(z); \mathbf{x}_V)$, equals:

$$U(C(z), H(z), s(z); \mathbf{x}_V) := V(C(z), T - H(z); \mathbf{x}_V) \cdot \varepsilon(s(z)). \tag{1}$$

Now, since $c = f(w, h; \mathbf{x}_f)$, the systematic part of the utility function, $V(c, T - h; \mathbf{x}_V)$, implicitly defines a utility function, say Ψ, defined over hours worked on the formal labour market, h, and wage, w:

$$\Psi\left(w, h; \mathbf{x}_V, \mathbf{x}_f\right) := V\left(f\left(w, h; \mathbf{x}_f\right), T - h; \mathbf{x}_V\right). \tag{2}$$

Consequently, we can define preferences also in the space of hours of work, wage, and other attributes, as follows:

$$U\left(f\left(W\left(z\right), H\left(z\right); \mathbf{x}_f\right), H\left(z\right), s\left(z\right); \mathbf{x}_V\right) := \Psi\left(W\left(z\right), H\left(z\right); \mathbf{x}_V, \mathbf{x}_f\right) \cdot \varepsilon\left(s\left(z\right)\right), \tag{3}$$

where $W\left(z\right)$ is the wage paid by activity z. More in particular, for someone not accepting any job offer, and choosing to an alternative $i \in \mathcal{I}$ which exhibits characteristics $s\left(i\right)$, the utility equals:

$$U\left(0, 0, s\left(i\right); \mathbf{x}_V\right) = \Psi\left(0, 0; \mathbf{x}_V, \mathbf{x}_f\right) \cdot \varepsilon\left(s\left(i\right)\right). \tag{4}$$

The domain of the systematic part of the utility function in the hours–wage space, $\Psi\left(\cdot\,; \mathbf{x}_V, \mathbf{x}_f, I\right)$, is $[0, T] \times [0, \infty)$.

2.3 Random opportunities

Both, jobs and non–market activities, are not equally available to all individuals. This is captured by the notion of intensity with which alternatives are rendered available to a specific individual. The probability to receive a job offer as a civil engineer, for someone who has only completed secondary school, is *e.g.* zero. Something similar holds for non–market activities: they are not all equally available to all agents. Someone having lost her legs will not be able to run (or not in the same fashion as before), though she might continue to be fond of it. In Appendix I we provide a brief introduction to the type of stochastic process that describes the degree to which job offers and non–market alternatives become available to an individual, and which is known as an inhomogeneous spatial Poisson process.
The intensity with which a job is offered to an individual depends on a number of personal characteristics, such as skills, education, experience, and on the characteristics of the job itself, more specifically, the wage, the labour time regime of the job, and its other attributes. In equation (3), preferences were defined over the *continuous* set of all possible amounts of time spent on the jobs. In the real world, however, jobs requiring a non–rational number of hours a week, are not available. Be it alone to organise the production process, it might sometimes be required to have a number of people working together during a fixed number of hours. So, in practice, full–time, three–quarter time and half time, one–quarter, or 20% jobs are more densely offered. Let $g_2\left(h; \mathbf{x}_{g_2}\right)$ be the intensity with which jobs requiring h hours of labour supply, are rendered available to an individual with characteristics $\mathbf{x}_{g_2}$. Similarly, jobs pay different wages, and personal characteristics can co-determine the intensity of job offers

that pay on average higher wages. Let $g_1(w\,|h\,;\mathbf{x}_{g_1})$ be the intensity with which, among the job offers requiring h hours of work, those paying a wage equal to w, are rendered available to a person with characteristics $\mathbf{x}_{g_1}$.
Persons do not only care for the wages a job pays, and the number of hours to be worked, but also for the other attributes of a job. From a behavioural theoretic point of view (see equation 1), these are only important in as far as they yield a specific value for the multiplication factor in the utility function for those alternatives. Two jobs, j_1 and j_2 say, paying the same wage and requiring the same amount of hours, with attributes yielding the same value of the multiplication factor in the utility function for those alternatives, that is $\varepsilon(s(j_1)) = \varepsilon(s(j_2))$, are thus, according to the behavioural model of equation (1), equivalent to each other in the present model, and therefore will be considered as the same opportunity. The intensity with which job offers arrive which yield a value for the multiplication factor in the utility function equal to ϵ_1, is denoted by $\lambda^1(\epsilon_1;\mathbf{x}_q)$. In this function, $\mathbf{x}_q$ refers to a set of personal characteristics that determine intensity with which job offers arrive to that person.[10] We explicitly included age in these characteristics, in order to assess potential age differentials in the intensity of job offers. This job offer intensity is reflected in a function π_1, which is assumed to have range $(0,1)$. We assume the following functional form for λ^1:

$$\lambda^1(\epsilon_1;\mathbf{x}_q) = \frac{\pi_1(\mathbf{x}_q)}{(\epsilon_1)^2}. \tag{5}$$

This functional form implies that attributes which are particularly disliked (yield a zero or very small value for ϵ_1) are excessively abundant, while those that are particularly liked (yielding a very high value of ϵ_1), are extremely scarce, irrespective of personal characteristics $\mathbf{x}_q$. These affect the level of λ_1 for all values of ε_1.
The distinguishing value of different potential non–market activities, is completely absorbed by the different values of the multiplication factor in the utility function they generate. Indeed, the systematic part of the utility function is for all non–market alternatives equal to $\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)$. As for jobs with the same wage and required labour input, two non–market activities, i_1 and i_2 say, with attributes yielding the same value of the multiplication factor in the utility function for those alternatives, that is $\varepsilon(s(i_1)) = \varepsilon(s(i_2))$, are from a behavioural theoretic point of view equivalent to each other, and will therefore be considered as one and the same opportunity. Similarly as for jobs, it will be assumed that leisure activities which are particularly disliked, are abundantly available, while those that are intensely desired, are

[10] The use of the sub–index q for those characteristics will become clear later when we introduce the q–function as the proportion of the intensity of job offers relative to the degree of availability of non–market alternatives. In short, job offer arrivals depend on personal capacities and skills which are subdivided in those apt to execute formal jobs, and those more suited for performing leisure activities. Next there may be personal characteristics on the basis of which discrimination in job offers by employers might take place.

rather difficult to obtain. Also, personal characteristics, the same as those having an impact on the intensity with which jobs are offered, $\mathbf{x}_q$, are influencing the relative availability of non–market alternatives, and this is measured by the function $\pi_0(\mathbf{x}_q)$. The intensity with which non–market activities yielding a multiplication factor equal to ϵ_0, denoted by $\lambda^0(\epsilon_0;\mathbf{x}_q)$, are accessible to an individual with characteristics $\mathbf{x}_q$, is thus assumed to be equal to:

$$\lambda^0(\epsilon_0;\mathbf{x}_q) = \frac{\pi_0(\mathbf{x}_q)}{(\epsilon_0)^2}. \tag{6}$$

It will be assumed that the effect of these characteristics on the availability of non–market runs in the reverse direction as that on job offer intensity.[11]

Consequently, the intensity with which non–market activities yielding a utility level larger than or equal to u_0, are rendered available to an individual with characteristics $\mathbf{x}_q$, denoted by $\Lambda^0(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\,\epsilon_0 \geq u_0;\mathbf{x}_q)$, is equal to:

$$\begin{aligned}\Lambda^0(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\,\epsilon_0 \geq u_0;\mathbf{x}_q) &= \int_{u_0/\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)}^{\infty} \frac{\pi_0(\mathbf{x}_q)}{(\epsilon_0)^2}\,\mathrm{d}\,\epsilon_0 \\ &= \pi_0(\mathbf{x}_q)\,\frac{\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)}{u_0}.\end{aligned} \tag{7}$$

If one assumes that Λ^0 is the intensity measure of a Poisson process, then the number of non–market alternatives that yield a utility level of at least u_0, available to a person with characteristics $(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_q)$, is Poisson distributed (see Appendix I). Let the number of available non–market activities yielding a utility level of at least u_0, be denoted by $N(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\,\epsilon_0 \geq u_0)$. The probability that $N(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\,\epsilon_0 \geq u_0) = n$, is, according to the Poisson distribution, equal to:

$$\begin{aligned}P(N(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\,\epsilon_0 \geq u_0 = n;\mathbf{x}_q)) &= \frac{\left(\Lambda^0\left(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\epsilon_0 \geq u_0;\mathbf{x}_q\right)\right)^n \exp\left[-\Lambda^0\left(\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)\epsilon_0 \geq u_0;\mathbf{x}_q\right)\right]}{n!} \\ &= \frac{\left(\pi_0(\mathbf{x}_q)\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)/u_0\right)^n \exp\left[-\pi_0(\mathbf{x}_q)\Psi(0,0;\mathbf{x}_V,\mathbf{x}_f)/u_0\right]}{n!}.\end{aligned} \tag{8}$$

The higher the intensity measure Λ^0, the more skewed to the right this distribution becomes, that is, the higher the probability that the number of available non–market alternatives yield a utility level of at least u_0 is relatively big.

In the present case, the process is inhomogeneous, since the degree of availability of non–market alternatives depends on the utility value they yield, while in an ordinary Poisson

[11] As a matter of normalisation it will thus be imposed that $\pi_0(\mathbf{x}_q) + \pi_1(\mathbf{x}_q) \equiv 1$.

process this is a constant parameter. The probability that all available non–market alternatives yield a utility level lower than u_0, equals the probability that the number of available alternatives with utility larger than or equal to u_0, is zero:

$$\begin{aligned} P\left(\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)\epsilon_0 < u_0;\mathbf{x}_q\right) &= P\left(N\left(\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)\epsilon_0 \geq u_0\right)=0;\mathbf{x}_q\right) \\ &= \exp\left[-\Lambda^0\left(\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)\epsilon_0 \geq u_0;\mathbf{x}_q\right)\right] \qquad (9) \\ &= \exp\left[-\pi_0\left(\mathbf{x}_q\right)\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)/u_0\right]. \end{aligned}$$

From the last equation, it can be concluded that the utility that can be derived from the available non–market alternatives, which is a stochastic variable equal to $v_0 := \Psi\left(0,0;\mathbf{x},\mathbf{y},I\right)\epsilon_0$, is Fréchet distributed with location parameter $\mu = 0$, scale parameter $\sigma_0\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_q\right) = \pi_0\left(\mathbf{x_q}\right)\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)$, and shape parameter $\alpha = 1$.[12] This will prove useful when deriving the likelihood function in the next section (Section 3.1).

Let $\mathcal{H}$ be the set of all possible labour time regimes of jobs offered in the market, and $\mathcal{W}$ the set of possible wage offers. Wages can obtain any positive value. Let $\mathcal{B} := \mathcal{B}_h \times \mathcal{B}_w$ be the Cartesian product of a measurable subset of labour time regimes $\mathcal{B}_h \subseteq \mathcal{H}$, and wage offers $\mathcal{B}_w = (0,w)$, for some positive w. Analogously to the modelling of the availability of non–market opportunities, the arrival of job offers[13] to a person exhibiting characteristics $\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q\right)$, is modelled by an inhomogeneous spatial Poisson process. Events are job offers that are characterised by a labour time regime, a wage offer, and the utility that can be derived from other attributes. The intensity parameter of this Poisson process will be denoted by $\Lambda^1\left(\mathcal{B},u;\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q\right)\right.$. This parameter reflects the intensity with which jobs paying a wage r lower than w (that is $r \in \mathcal{B}_w$), and specifying a working time regime t in $\mathcal{B}_h$, and which will yield a utility level at least equal to u, are rendered available to a person with characteristics $\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q\right)$, and it is defined as:

$$\begin{aligned} \Lambda^1(\mathcal{B},u;\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q) &:= \int\limits_{t\in\mathcal{B}_h} g_2\left(t;\mathbf{x}_{g_2}\right)\int\limits_{r\in\mathcal{B}_w} g_1\left(r\,|t\,;\mathbf{x}_{g_1}\right)\int\limits_{u/\Psi\left(t,r;\mathbf{x}_V,\mathbf{x}_f\right)}^{\infty}\frac{\pi_1(\mathbf{x}_q)}{\left(\epsilon_1\right)^2}\,\mathrm{d}\epsilon_1\,\mathrm{d}r\,\mathrm{d}t \\ &= \frac{\int_{t\in\mathcal{B}_h} g_2\left(t;\mathbf{x}_{g_2}\right)\int_{r\in\mathcal{B}_w} g_1\left(r|t\,;\mathbf{x}_{g_1}\right)\pi_1(\mathbf{x}_q)\Psi\left(t,r;\mathbf{x}_V,\mathbf{x}_f\right)\mathrm{d}r\,\mathrm{d}t}{u}. \end{aligned} \qquad (10)$$

Let $N\left(\mathcal{B},u\right)$ be the number of job offers with a wage r belonging to $\mathcal{B}_w$, the number of hours to be worked in $\mathcal{B}_h$, and yielding a utility level larger than or equal to u. The probability

[12] In general, the class of Fréchet distributions is defined as: $F\left(x;\mu,\sigma,\alpha\right) := \exp\left[-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}\right]$, where μ is a location parameter, σ a scale parameter, and α is a shape parameter. We could alternatively argue that the multiplier in the utility function, stemming from the attractiveness of the non–pecuniary attributes of non–market alternatives to a particular individual with characteristics $\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_q\right)$, is Fréchet distributed with location parameter $\mu = 0$, scale parameter $\sigma = \pi_0\left(\mathbf{x}_q\right)$, and shape parameter $\alpha = 1$.

[13] As mentioned before, a 'job offer is a short hand for 'an alternative containing at least one job offer'.

for an individual with characteristics $(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q)$ to be offered n such jobs is then equal to:

$$P\left(N\left(\mathcal{B}, u\right) = n; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) = \frac{\left(\Lambda^1\left(\mathcal{B}, u; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)\right)^n \exp\left[-\Lambda^1\left(\mathcal{B}, u; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)\right]}{n!}. \tag{11}$$

Let $P\left(U_{\mathcal{B}} \leq u\right)$ be the probability that all job offers with wages in $\mathcal{B}_w$, and number of hours to be worked in $\mathcal{B}_h$, yield a utility level *less than* u. This is equal to the probability that there are no job offers available that pay wages and specifying a working time in these ranges, and which yield a utility level of at least u:

$$\begin{aligned} P\left(U_{\mathcal{B}} \leq u; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) &= P\left(N\left(\mathcal{B}, u\right) = 0; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) \\ &= \exp\left[-\Lambda^1\left(\mathcal{B}, u; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)\right] \\ &= \exp\left[-\frac{\int_{t\in\mathcal{B}_h} g_2\left(t; \mathbf{x}_{g_2}\right) \int_{r\in\mathcal{B}_w} g_1\left(r|t; \mathbf{x}_{g_1}\right) \pi_1(\mathbf{x}_q) \Psi\left(t, r; \mathbf{x}_V, \mathbf{x}_f\right) \mathrm{d}r\, \mathrm{d}t}{u}\right]. \end{aligned} \tag{12}$$

The utility level that can be obtained from jobs with working time and wage combinations in $\mathcal{B}$, which is a stochastic variable, denoted as $U_{\mathcal{B}}$, and which is governed by the stochastic process of job offers arriving to an individual, is thus Fréchet distributed with location parameter $\mu = 0$, scale parameter

$$\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) = \int\limits_{t\in\mathcal{B}_h} g_1\left(t; \mathbf{x}_{g_1}\right) \int\limits_{r\in\mathcal{B}_w} g_2\left(r\,|t\,; \mathbf{x}_{g_2}\right) \pi_1\left(\mathbf{x}_q\right) \Psi\left(t, r; \mathbf{x}_V, \mathbf{x}_f\right)\, \mathrm{d}\,r\, \mathrm{d}\,t,$$

and shape parameter $\alpha = 1$.

The derivation of this distribution is equally valid for any (measurable) subset $\mathcal{B}$ of the space of possible working times and wage combinations job offers might exhibit. More in particular, it holds for the complement of $\mathcal{B}$ in the set of all possible working time wage combinations, defined as $\mathcal{B}^c := \mathcal{B}_h^c \times \mathcal{B}_w^c = \mathcal{H} \backslash \mathcal{B}_h \times [w, \infty)$. It follows that the random variable $U_{\mathcal{B}^c}$, denoting the utility level derivable from possible job offers with working time wage combinations in $\mathcal{B}^c$, is Fréchet distributed with location parameter $\mu = 0$, scale parameter

$$\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) = \int\limits_{t\in\mathcal{B}_h^c} g_2\left(t; \mathbf{x}_{g_2}\right) \int\limits_{r\in\mathcal{B}_w^c} g_1\left(r\,|t\,; \mathbf{x}_{g_1}\right) \pi_1\left(\mathbf{x}_q\right) \Psi\left(t, r; \mathbf{x}_V, \mathbf{x}_f\right)\, \mathrm{d}\,r\, \mathrm{d}\,t,$$

and shape parameter $\alpha = 1$.

3 Likelihood function, functional form, and estimation

3.1 Derivation of the likelihood

Now we turn to the behavioural implications of the model explained in Section 2. From the available job offers and non–market opportunities, a person will choose that alternative she likes most. The probability that this will be an alternative including a job offer with a working time wage combination in the set $\mathcal{B}$, is equal to the probability that $U_{\mathcal{B}} \geq \max\{U_0, U_{\mathcal{B}^c}\}$. This can be calculated as follows. From the in the previous section derived distribution of the utility levels resulting from the stochastic process governing the arrival of job offers and non–market opportunities, it follows that the probability that $U_{\mathcal{B}^c}$ (the utility from a job offer with a labour time and wage combination in $\mathcal{B}^c$) will be smaller than some non–negative number u, is equal to:

$$\begin{aligned} P\left(U_{\mathcal{B}^c} \leq u; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) &= \int_0^u \frac{\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{(\mathrm{v})^2} \exp\left[-\frac{\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{\mathrm{v}}\right] \mathrm{d}\,\mathrm{v} \\ &= \exp\left[-\frac{\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{u}\right]. \end{aligned} \tag{13}$$

Similarly (see equation 9), the probability that U_0 (the utility from the available non–market alternatives) will be smaller than u, is equal to:

$$\begin{aligned} P\left(U_0 \leq u; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right) &= \int_0^u \frac{\sigma_0\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right)}{(\mathrm{v})^2} \exp\left[-\frac{\sigma_0\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right)}{\mathrm{v}}\right] \mathrm{d}\,\mathrm{v} \\ &= \exp\left[-\frac{\sigma_0\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right)}{u}\right]. \end{aligned} \tag{14}$$

As the processes governing the arrival of job offers and non–market opportunities are assumed to be independent, the probability that $U_{\mathcal{B}}$ (the utility from a job offer with a labour time and wage combination in $\mathcal{B}$) is equal to or greater than $\max\{U_0, U_{\mathcal{B}^c}\}$ is equal to the product of the probability that $U_{\mathcal{B}}$ is greater than or equal to U_0 and the probability that $U_{\mathcal{B}}$ is greater than or equal to $U_{\mathcal{B}^c}$. That is: the product of the probability that U_0 is smaller than a certain non–negative value u and the probability that $U_{\mathcal{B}^c}$ is smaller than that same value u, evaluating that joint probability weighted by the likelihood that $U_{\mathcal{B}}$ takes on the value u, and integrating this over all possible values u that $U_{\mathcal{B}}$ could assume:

$$P\left(U_{\mathcal{B}} \geq \max\left\{U_0, U_{\mathcal{B}^c}\right\}; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) =$$

$$\int_0^\infty \frac{\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{(u)^2} \exp\left[-\frac{\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{u}\right] \exp\left[-\frac{\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{u}\right] \exp\left[-\frac{\sigma_0\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right)}{u}\right] \mathrm{d}\, u$$

$$= \int_0^\infty \frac{\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{(u)^2} \exp\left[-\left(\frac{\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)+\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)+\sigma_0\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right)}{u}\right)\right] \mathrm{d}\, u$$

$$= \frac{\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)}{\sigma_{\mathcal{B}}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)+\sigma_{\mathcal{B}^c}\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)+\sigma_0\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_q\right)}$$

$$= \frac{\int_{t\in\mathcal{B}_h} g_2\left(t; \mathbf{x}_{g_2}\right) \int_{r\in\mathcal{B}_w} g_1\left(r|t; \mathbf{x}_{g_1}\right) \pi_1(\mathbf{x}_q) \Psi\left(t, r; \mathbf{x}_V, \mathbf{x}_f\right) \mathrm{d}\, r\, \mathrm{d}\, t}{\pi_0(\mathbf{x}_q)\Psi\left(0,0; \mathbf{x}_V, \mathbf{x}_f\right) + \int_{t\in\mathcal{H}} g_2\left(t; \mathbf{x}_{g_2}\right) \int_{r\in\mathcal{W}} g_1\left(r|t; \mathbf{x}_{g_1}\right) \pi_1(\mathbf{x}_q) \Psi\left(t, r; \mathbf{x}_V, \mathbf{x}_f\right) \mathrm{d}\, r\, \mathrm{d}\, t}. \tag{15}$$

In a similar fashion it can be derived that the probability to choose a non–market alternative, is equal to the probability that U_0 is equal to or greater than $U_{\mathcal{B}\cup\mathcal{B}^c}$, which is equal to:

$$P\left(U_0 \geq U_{\mathcal{B}\cup\mathcal{B}^c}; \mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right) = \frac{\pi_0(\mathbf{x}_q)\Psi\left(0,0; \mathbf{x}_V, \mathbf{x}_f\right)}{\pi_0(\mathbf{x}_q)\Psi\left(0,0; \mathbf{x}_V, \mathbf{x}_f\right) + \int_{t\in\mathcal{H}} g_2\left(t; \mathbf{x}_{g_2}\right) \int_{r\in\mathcal{W}} g_1\left(r|t; \mathbf{x}_{g_1}\right) \pi_1(\mathbf{x}_q) \Psi\left(t, r; \mathbf{x}_V, \mathbf{x}_f\right) \mathrm{d}\, r\, \mathrm{d}\, t}. \tag{16}$$

We noted before that the model assumes that the characteristics $\mathbf{x}_q$ which influence the intensity with which job offers arrive to a person endowed with those characteristics, are the same as those affecting the degree of availability of non–market alternatives, but acting in the opposite direction. Equations (15) and (16) make clear that this is necessary for identifying the model. Indeed, the same result would be obtained when dividing through the numerator and denominator by $\pi_0(\mathbf{x}_q)$. We therefore introduce the notion of *relative* intensity with which job offers arrive, as compared to the degree to which non–market alternatives are available:

$$q(\mathbf{x}_q) := \frac{\pi_1(\mathbf{x}_q)}{\pi_0(\mathbf{x}_q)}. \tag{17}$$

Estimating the parameters of this function, and assuming that $0 \leq \pi_i(\mathbf{x}_q) \leq 1$ for all $\mathbf{x}_q$, and $i = 0, 1$, and $\pi_1(\mathbf{x}_q) + \pi_0(\mathbf{x}_q) = 1$, will allow then to recover both $\pi_1(\mathbf{x}_q)$ and $\pi_0(\mathbf{x}_q)$. An additional assumption for identification is the independence of the wage offer distribution from the hours specified by the job offers. That is $g_1(w|h; \mathbf{x}_{g_1}) = g_1(w; \mathbf{x}_{g_1})$, $\forall\, h \in \mathcal{H}$.

The likelihood[14] that a person with characteristics $(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q)$ will choose one particular job offer requiring labour time h, and paying a wage w, can thus be written as:

$$\varphi\left(w, h; (\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q)\right) = \frac{\Psi\left(w,h;\mathbf{x}_V,\mathbf{x}_f\right)q(\mathbf{x}_q)g_1\left(w;\mathbf{x}_{g_1}\right)g_2\left(h;\mathbf{x}_{g_2}\right)}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\int\limits_{s\in\mathcal{W}}\int\limits_{t\in\mathcal{H}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)q(\mathbf{x}_q)g_1\left(s;\mathbf{x}_{g_1}\right)g_2\left(t;\mathbf{x}_{g_2}\right)\,\mathrm{d}t\,\mathrm{d}s}, \quad (18)$$

Similarly, the likelihood her most preferred non–market alternative is preferred to any of the job offers, equals:

$$\varphi\left(0, 0; (\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q)\right) = \frac{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\int\limits_{s\in\mathcal{W}}\int\limits_{t\in\mathcal{H}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)q(\mathbf{x}_q)g_1\left(s;\mathbf{x}_{g_1}\right)g_2\left(t;\mathbf{x}_{g_2}\right)\,\mathrm{d}t\,\mathrm{d}s}. \quad (18')$$

Up to now, we exclusively concentrated on individual decision makers. The model is easily extended to the case of households consisting of couples (with or without children), if one is willing to assume a unitary decision making model. More specifically, it is assumed that both partners have identical tastes specified over household consumption, each of the partners' leisure time, and other attributes associated with the activities the partners executed. If one assumes that each partner's process of job offer arrivals and availability of non–market alternatives is independent of that of the other, one arrives at the following expressions for the likelihood that both partners, respectively, only one of them, or none of both, will accept a job offer:

$$\begin{aligned}
\varphi\left(w_1, h_1, w_2, h_2; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1^j}, \mathbf{x}_{g_2^j}, \mathbf{x}_{q^j}; j=1,2\right)\right) &= \frac{\Psi\left(w_1,h_1,w_2,h_2;\mathbf{x}_V,\mathbf{x}_f\right)\prod\limits_{j=1,2} q^j\left(\mathbf{x}_{q^j}\right)g_1^j\left(w_j;\mathbf{x}_{g_1^j}\right)g_2^j\left(h_j;\mathbf{x}_{g_2^j}\right)}{\Psi\left(0,0,0,0;\mathbf{x}_V,\mathbf{x}_f\right)+A+B+C}, \\
\varphi\left(w_1, h_1, 0, 0; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1^j}, \mathbf{x}_{g_2^j}, \mathbf{x}_{q^j}; j=1,2\right)\right) &= \frac{\Psi\left(w_1,h_1,0,0;\mathbf{x}_V,\mathbf{x}_f\right)q^1\left(\mathbf{x}_{q^1}\right)g_1^1\left(w_1;\mathbf{x}_{g_1^1}\right)g_2^1\left(h_1;\mathbf{x}_{g_2^1}\right)}{\Psi\left(0,0,0,0;\mathbf{x}_V,\mathbf{x}_f\right)+A+B+C}, \\
\varphi\left(0, 0, w_2, h_2; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1^j}, \mathbf{x}_{g_2^j}, \mathbf{x}_{q^j}; j=1,2\right)\right) &= \frac{\Psi\left(0,0,w_2,h_2;\mathbf{x}_V,\mathbf{x}_f\right)q^2\left(\mathbf{x}_{q^2}\right)g_1^2\left(w_2;\mathbf{x}_{g_1^2}\right)g_2^2\left(h_2;\mathbf{x}_{g_2^2}\right)}{\Psi\left(0,0,0,0;\mathbf{x}_V,\mathbf{x}_f\right)+A+B+C}, \\
\varphi\left(0, 0, 0, 0; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1^j}, \mathbf{x}_{g_2^j}, \mathbf{x}_{q^j}; j=1,2\right)\right) &= \frac{\Psi\left(0,0,0,0;\mathbf{x}_V,\mathbf{x}_f\right)}{\Psi\left(0,0,0,0;\mathbf{x}_V,\mathbf{x}_f\right)+A+B+C},
\end{aligned} \quad (18'')$$

[14] We use the term likelihood, though, in econometrics, the likelihood would express this as function of the parameters of the model to be estimated.

with:

$$A := \int_{s_1 \in \mathcal{W}} \int_{t_1 \in \mathcal{H}} \Psi\left(s_1, t_1, 0, 0; \mathbf{x}_V, \mathbf{x}_f\right) q^1\left(\mathbf{x}_{q^1}\right) g_1^1\left(s_1; \mathbf{x}_{g_1^1}\right) g_2^1\left(t_1; \mathbf{x}_{g_2^1}\right) \, \mathrm{d}\, t_1 \, \mathrm{d}\, s_1,$$

$$B := \int_{s_2 \in \mathcal{W}} \int_{t_2 \in \mathcal{H}} \Psi\left(0, 0, s_2, t_2; \mathbf{x}_V, \mathbf{x}_f\right) q^2\left(\mathbf{x}_{q^2}\right) g_1^2\left(s_2; \mathbf{x}_{g_1^2}\right) g_2^2\left(t_2; \mathbf{x}_{g_2^2}\right) \, \mathrm{d}\, t_2 \, \mathrm{d}\, s_2,$$

$$C := \int_{s_1 \in \mathcal{W}} \int_{t_1 \in \mathcal{H}} \int_{s_2 \in \mathcal{W}} \int_{t_2 \in \mathcal{H}} \Psi\left(s_1, t_1, s_2, t_2; \mathbf{x}_V, \mathbf{x}_f\right) \prod_{j=1,2} q^j\left(\mathbf{x}_{q^j}\right) g_1^j\left(s_j; \mathbf{x}_{g_1^j}\right) g_2^j\left(t_j; \mathbf{x}_{g_2^j}\right) \, \mathrm{d}t_2 \, \mathrm{d}s_2 \, \mathrm{d}t_1 \, \mathrm{d}\, s_1.$$

where q^j, g_i^j $(i = 1, 2)$ are the relative intensity of job offers, the wage offer density, and the labour time density of partner j $(j = 1, 2)$. For couples, the non–market alternative is that alternative in which none of both partners is engaged in the formal labour market.

3.2 Functional forms

In the present section we present the functional forms of the different components of the model that will be used in the empirical application in Section 5.

At the preference side,

- the systematic part of the log utility function for singles is of the Box–Cox type[15]: $\ln V\left(c, T-h; \mathbf{x}_V\right) = \beta_c \cdot \left(\frac{c^{\alpha_c}-1}{\alpha_c}\right) + \left(\boldsymbol{\beta}_h' \mathbf{x}_V\right) \cdot \left(\frac{((T-h)/T)^{\alpha_h}-1}{\alpha_h}\right)$, with $\alpha_c, \alpha_h < 1$. Intensity of preferences for leisure is increasing (decreasing) in an element of $\mathbf{x}_V$, if the associated parameter of $\boldsymbol{\beta}_h$ is positive (negative).[16] The exponents, α_c and α_h, determine the curvature of the indifference curves in terms of labour time and consumption. The lower these are, the less substitutable leisure and consumption are;

- for couples, a unitary decision model is assumed, but spouses' leisure time is considered to be an assignable good. So, preferences are defined over consumption and each spouse's leisure time. Partner's time endowments are equal. An interaction term capturing potential complementarities between partners' leisure time is added to the utility function:
$$\ln V\left(c, T-h_1, T-h_2; \mathbf{x}_V\right) = \beta_{c,g} \cdot \left(\frac{c^{\alpha_{c,g}}-1}{\alpha_{c,g}}\right) + \sum_{i=1,2} \left(\boldsymbol{\beta}_{h_i}' \mathbf{x}_V\right) \cdot \left(\frac{((T-h_i)/T)^{\alpha_{h_i}}-1}{\alpha_{h_i}}\right) + \beta_{h_1,h_2} \cdot \prod_{i=1,2} \left(\frac{((T-h_i)/T)^{\alpha_{h_i}}-1}{\alpha_{h_i}}\right),$$
with $\alpha_{c,g}, \alpha_{h_i} < 1$ $(i = 1, 2)$. The interpretation of the exponents and the $\boldsymbol{\beta}_{h_i}'$ $(i = 1, 2)$ remains the same as for singles; in addition, $\beta_{h_1,h_2} > (<) 0$ means that partners' leisure are complements (substitutes).

[15] For a justification, see Dagsvik and Røine Hoff (2011), and Dagsvik (2013).

[16] More details are provided in Section 5.1.

At the opportunity side,

- the log of the intensity of job offers relative to the availability of non–market alternatives is linear in the covariates: $\ln q(\mathbf{x}_q) = \boldsymbol{\eta}_q' \mathbf{x}_q$. The vector $\mathbf{x}_q$ should contain a constant term, and the associated coefficient is denoted by $\eta_{q,0}$;

- the wage density $g_1(w; \mathbf{x}_{g_1})$ is assumed to be lognormal:
$g_1(w; \mathbf{x}_{g_1}) = \frac{1}{w \cdot \sigma \cdot \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln w - \boldsymbol{\delta}_{g_1}' \mathbf{x}_{g_1}}{\sigma}\right)^2\right)$; and

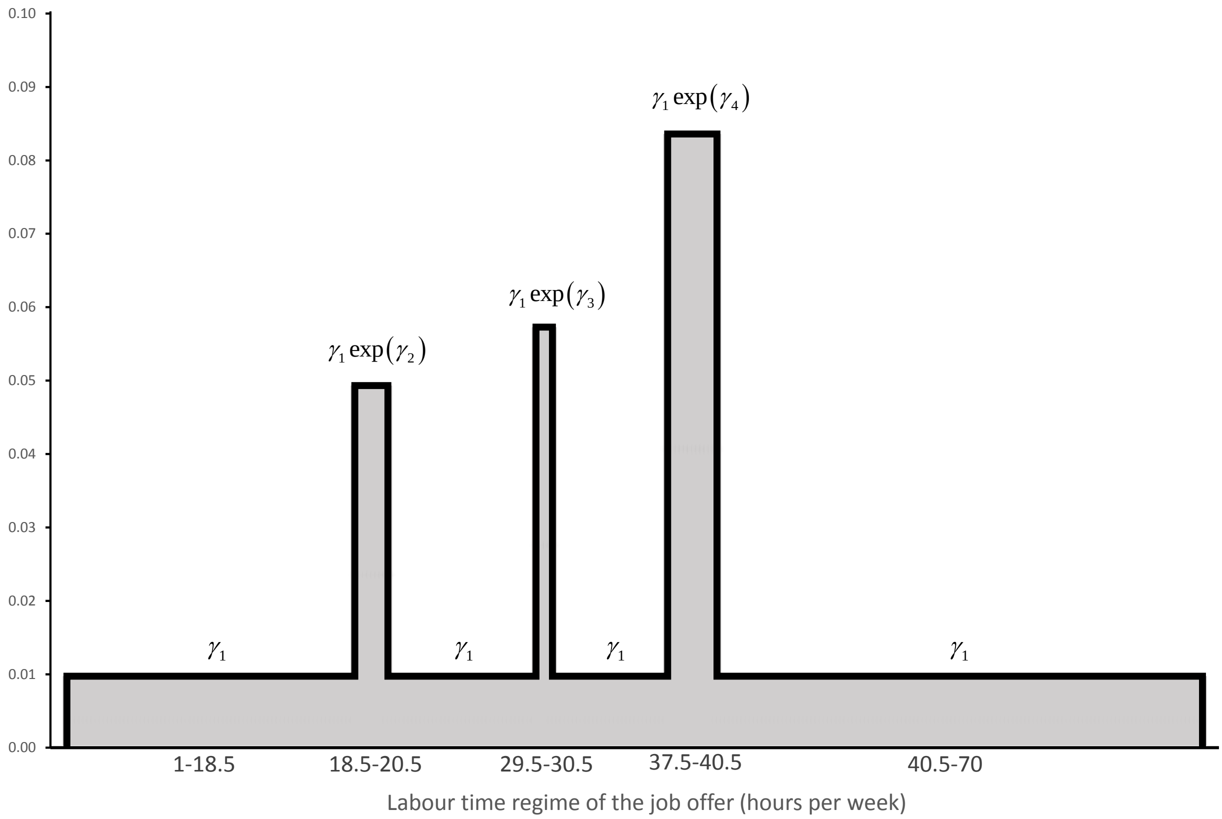
- the distribution of the labour time regimes offered, is piecemeal uniform.[17] There are a number of, say K, peaks, indexed by $k = 1, 2, \ldots, K$, around which the bulk of the job offers' labour time regimes are concentrated (typically around half time, that is 18.5 to 20.5 hour a week in our application, three quarter time, or 29.5 to 30.5 hours a week, and full time, or 37.5 to 40.5 hours). The lower and upper bound of peak k ($k = 1, 2, \ldots, K$) are denoted by respectively $\underline{H}_k$ and $\overline{H}_k$. There is a lower limit, $H_{\min}$, below which job offers are not considered to belong to the formal labour market (fixed at one hour a week in the application below); and an upper limit of labour time spent on formal jobs, denoted by $H_{\max}$, and fixed at 70 hours per week in our application. This results in the following density function:

$$g_2(h; \mathbf{x}_{g_2}) = \begin{cases} \gamma_1 & \text{if} \quad h \in [H_{\min}, \underline{H}_1[\,, h \in \left[\overline{H}_k, \underline{H}_{k+1}\right[, \text{ or } h \in \left[\overline{H}_K, H_{\max}\right[, \\ & \qquad k = 1, 2, \ldots, K-1, \\ \gamma_1 \exp \gamma_{k+1} & \text{if} \quad h \in \left[\underline{H}_k, \overline{H}_k\right[, \quad k = 1, 2, \ldots, K. \end{cases}$$

The only covariate influencing this function will thus be the sex of the person. An example of such a distribution function is given in Figure 1.

[17] This makes that hours restrictions in the job offers have formally a similar implication as including dummies for part time and full time work in a discrete choice labour supply model, as was proposed already in the seminal paper of Van Soest (1995).

Figure 1: Peak distribution for labour time regimes



3.3 Estimation

To estimate the parameters governing preferences, the relative intensity of market over non–market alternatives, and the distribution of wage offers and labour time regimes, a likelihood function, say $\mathcal{L}$, is constructed on the basis of equations (18), (18') and (18"). The individual contributions of a single to that likelihood function are indeed composed of the likelihood that the observed choice is the most preferred one, reflected in equations (18), or (18'), depending on whether the observed choice involves participation on the formal labour market executing a job (or set of jobs) requiring h hours of work, and paying a wage w, or whether it is the non–market alternative. In these expressions, the numerator is thus evaluated at the actually observed choice, when constructing the likelihood function. Similarly, for couples the first, second, third, or fourth equation in (18") applies, dependent on whether both partners, only partner j ($j = 1, 2$), or none of both actually are engaged in formal jobs.
In practice, we do not observe the set of wage offers, $\mathcal{W}$, nor the offered labour time regimes, $\mathcal{H}$. Therefore, a set of alternatives in the space of wages and labour time regimes is sampled from a prior density function, say $\mathbb{P}(w, h; \mathbf{x}_{\mathbb{P}})$. This prior may be individually specific through its possible dependence on the covariates $\mathbf{x}_{\mathbb{P}}$. Denote the set of sampled combin-

ations of wage offers and labour time regimes, possibly including the non–market alternative, by $\mathcal{D}$. The observed choice $\left(w^{\text{obs}}, h^{\text{obs}}\right)$ is to be always included in the sampled choice set. From the sampling densities $\mathbb{P}\left(w, h; \mathbf{x}_{\mathbb{P}}\right)$, the likelihood to sample a set $\mathcal{D}$ given that the observed choice equals $\left(w^{\text{obs}}, h^{\text{obs}}\right)$ can be constructed. [18] It is denoted by $\mathcal{P}\left(\mathcal{D} \,\middle|\, \left(w^{\text{obs}}, h^{\text{obs}}\right)\right)$, and it equals:

$$\mathcal{P}\left(\mathcal{D} \,\middle|\, \left(w^{\text{obs}}, h^{\text{obs}}\right)\right) := \prod_{i:(w_i,h_i)\in\mathcal{D}} \frac{\mathbb{P}\left(w_i, h_i; \mathbf{x}_{\mathbb{P}}\right)}{\mathbb{P}\left(w^{\text{obs}}, h^{\text{obs}}; \mathbf{x}_{\mathbb{P}}\right)}. \tag{19}$$

Recall that the probability (density) that a job paying a wage w, and requiring a number of h hours to be worked, would be optimal given a choice set $\mathcal{C} := \{0, 0\} \cup \mathcal{W} \times \mathcal{H}$, was derived in equations (18), and (18') if the non–market alternative would be the most preferred option. The unconditional probability to sample a choice set $\mathcal{D}$, denoted by $\Pi\left(\mathcal{D}\right)$, can thus be written as[19]:

$$\Pi\left(\mathcal{D}\right) = \sum_{i:(w_i,h_i)\in\mathcal{D}} \mathcal{P}\left(\mathcal{D} \,|\, (w_i, h_i)\right) \varphi\left(w_i, h_i; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)\right). \tag{20}$$

Using Bayes' law, the probability (density) to observe an agent choosing a job offer that pays a wage w_i and requires h_i hours of labour time from the sampled set $\mathcal{D}$, thus equals:

$$\widetilde{\varphi}\left(w_i, h_i; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q, \mathbf{x}_{\mathbb{P}}\right) | \mathcal{D}\right) = \frac{\mathcal{P}\left(\mathcal{D} \,|\, (w_i, h_i)\right) \varphi\left(w_i, h_i; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)\right)}{\Pi\left(\mathcal{D}\right)}. \tag{21}$$

Using equations (19) and (20), we can thus reformulate the simulated likelihood to observe someone choosing an alternative (w, h) from a choice set $\mathcal{D}$ sampled according to the prior $\mathbb{P}\left(w, h; \mathbf{x}_{\mathbb{P}}\right)$, as:

$$\begin{aligned}
&\widetilde{\varphi}\left(w, h; \left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q, \mathbf{x}_{\mathbb{P}}\right) | \mathcal{D}\right) \\
&\qquad = \frac{\Psi\left(w,h;\mathbf{x}_V,\mathbf{x}_f\right) q(\mathbf{x}_q) g_1\left(w;\mathbf{x}_{g_1}\right) g_2\left(h;\mathbf{x}_{g_2}\right) / \mathbb{P}(w,h;\mathbf{x}_{\mathbb{P}})}{\frac{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)}{\mathbb{P}(0,0;\mathbf{x}_{\mathbb{P}})} + \sum_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}} \frac{\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right) q(\mathbf{x}_q) g_1(s;\mathbf{x}_{g_1}) g_2(t;\mathbf{x}_{g_2})}{\mathbb{P}(s,t;\mathbf{x}_{\mathbb{P}})}} \\
&\qquad = \frac{\Psi\left(w,h;\mathbf{x}_V,\mathbf{x}_f\right) q(\mathbf{x}_q) g_1\left(w;\mathbf{x}_{g_1}\right) g_2\left(h;\mathbf{x}_{g_2}\right) \frac{\mathbb{P}(0,0;\mathbf{x}_{\mathbb{P}})}{\mathbb{P}(w,h;\mathbf{x}_{\mathbb{P}})}}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right) + \sum_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}} \Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right) q(\mathbf{x}_q) g_1\left(s;\mathbf{x}_{g_1}\right) g_2\left(t;\mathbf{x}_{g_2}\right) \frac{\mathbb{P}(0,0;\mathbf{x}_{\mathbb{P}})}{\mathbb{P}(s,t;\mathbf{x}_{\mathbb{P}})}}.
\end{aligned} \tag{22}$$

[18] The issue of sampling choice sets for estimating the RURO model is discussed more in detail in Appendix II, and in McFadden (1978), Ben–Akiva and Lerman (1985), Aaberge, Columbino and Wennemo (2009), Train (2009), and Lemp and Kockelman (2012).

[19] For simplicity of notation, we drop the arguments $\left(\mathbf{x}_V, \mathbf{x}_f, \mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}_q\right)$ that might influence the shape of $\Pi\left(\mathcal{D}\right)$.

The corresponding expression for choosing the non–market alternative equals:

$$\widetilde{\varphi}\left(0,0;\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q,\mathbf{x}_{\mathbb{P}}\right)|\mathcal{D}\right)$$

$$= \frac{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\sum\limits_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)q(\mathbf{x}_q)g_1\left(s;\mathbf{x}_{g_1}\right)g_2\left(t;\mathbf{x}_{g_2}\right)\frac{\mathbb{P}\left(0,0;\mathbf{x}_{\mathbb{P}}\right)}{\mathbb{P}\left(s,t;\mathbf{x}_{\mathbb{P}}\right)}}. \tag{22'}$$

Some further issues are in order. First, note that the constant term of the $q\left(\mathbf{x}_q\right)$–function, $\exp\left(\eta_{q,0}\right)$, occurs in any term of the likelihood where γ_1 appears (that is, in those terms of the sum pertaining to a job offer on the formal labour market, $(w,h): w,h>0$), and each time these terms appear as a product. Therefore both, $\eta_{q,0}$ and γ_1, cannot be estimated separately. But γ_1 is still identified by the definition:

$$\gamma_1\left(H_{\max}-\overline{H}_K+\sum_{k=1}^{K-1}\left(\underline{H}_{k+1}-\overline{H}_k\right)+\underline{H}_1-H_{\min}+\sum_{k=1}^{K}\left(\overline{H}_k-\underline{H}_k\right)\exp\gamma_k\right)\equiv 1. \tag{23}$$

This means in practice that one does not estimate all the parameters in the likelihood (22) (respectively 22'), but rather reduces these equations to:

$$\widetilde{\varphi}\left(w,h;\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q,\mathbf{x}_{\mathbb{P}}\right)|\mathcal{D}\right)$$

$$= \frac{\Psi\left(w,h;\mathbf{x}_V,\mathbf{x}_f\right)\widetilde{q}(\mathbf{x}_q)g_1\left(w;\mathbf{x}_{g_1}\right)\mathscr{g}_2\left(h;\mathbf{x}_{g_2}\right)\frac{\mathbb{P}\left(0,0;\mathbf{x}_{\mathbb{P}}\right)}{\mathbb{P}\left(w,h;\mathbf{x}_{\mathbb{P}}\right)}}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\sum\limits_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)\widetilde{q}(\mathbf{x}_q)g_1\left(s;\mathbf{x}_{g_1}\right)\mathscr{g}_2\left(t;\mathbf{x}_{g_2}\right)\frac{\mathbb{P}\left(0,0;\mathbf{x}_{\mathbb{P}}\right)}{\mathbb{P}\left(s,t;\mathbf{x}_{\mathbb{P}}\right)}}, \tag{24}$$

where:

$$\begin{array}{lcl} \mathscr{g}_2\left(h;\mathbf{x}_{g_2}\right) & = & \frac{g_2\left(h;\mathbf{x}_{g_2}\right)}{\gamma_1}, \\ \widetilde{q}\left(\mathbf{x}_q\right) & = & \gamma_1 q\left(\mathbf{x}_q\right). \end{array}$$

For the likelihood to choose a non–market alternative, this becomes:

$$\widetilde{\varphi}\left(0,0;\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q,\mathbf{x}_{\mathbb{P}}\right)|\mathcal{D}\right)$$

$$= \frac{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\sum\limits_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)\widetilde{q}(\mathbf{x}_q)g_1\left(s;\mathbf{x}_{g_1}\right)\mathscr{g}_2\left(t;\mathbf{x}_{g_2}\right)\frac{\mathbb{P}\left(0,0;\mathbf{x}_{\mathbb{P}}\right)}{\mathbb{P}\left(s,t;\mathbf{x}_{\mathbb{P}}\right)}}, \tag{24'}$$

Secondly, as far as (some parts of) the probability to sample a job offer, that is an alternative (w,h) such that $w,h>0$, is (are) independent of the specific value of either w, or h, or both, one can drop that part from the likelihood, and it will be absorbed by the constant term of the $q\left(\mathbf{x}_q\right)$–function. For example, assume wages are sampled from a lognormal distribution with parameters μ and ς, labour time is sampled from the uniform distribution on the $[H_{\min},H_{\max})$–interval, and the probability to sample non–market alternatives is the observed inactivity degree in the sample (that is the relative number of persons in the sample being engaged in formal jobs for less than one hour a week), say π_0^{obs}.[20] That is:

$$\begin{aligned}
\mathbb{P}\left(w,h;\mathbf{x}_{\mathbb{P}}\right) &= \pi_0^{\text{obs}} && \text{if } (w,h)=(0,0)\,, \\
&= \left(1-\pi_0^{\text{obs}}\right)\frac{\left(w\varsigma\sqrt{2\pi}\right)^{-1}\exp\left(-\frac{(\ln w-\mu)^2}{2\varsigma^2}\right)}{H_{\max}-H_{\min}} && \text{if } w>0, h\in[H_{\min},H_{\max}[\,, \\
&= 0 && \text{otherwise.}
\end{aligned} \tag{25}$$

The factor $\frac{\pi_0^{\text{obs}}}{\left(1-\pi_0^{\text{obs}}\right)}\left(H_{\max}-H_{\min}\right)$ occurs then in all terms of the likelihood for which $w,h>0$. Suppose now that one drops this factor from the likelihood and estimates on the basis of:

$$\widetilde{\varphi}\left(w,h;\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q,\mathbf{x}_{\mathbb{P}}\right)|\mathcal{D}\right)$$

$$= \frac{\Psi\left(w,h;\mathbf{x}_V,\mathbf{x}_f\right)\mathscr{q}(\mathbf{x}_q)\mathscr{g}_1\left(w;\mathbf{x}_{g_1}\right)\mathscr{g}_2\left(h;\mathbf{x}_{g_2}\right)\varsigma\exp\left(\frac{(\ln w-\mu)^2}{2\varsigma^2}\right)}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\sum\limits_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)\mathscr{q}(\mathbf{x}_q)\mathscr{g}_1\left(s;\mathbf{x}_{g_1}\right)\mathscr{g}_2\left(t;\mathbf{x}_{g_2}\right)\varsigma\exp\left(\frac{(\ln s-\mu)^2}{2\varsigma^{2^2}}\right)}, \tag{26}$$

where:

$$\begin{aligned}
\mathscr{g}_1\left(w;\mathbf{x}_{g_1}\right) &= w\sqrt{2\pi}g_1\left(w;\mathbf{x}_{g_1}\right), \\
\mathscr{g}_2\left(h;\mathbf{x}_{g_2}\right) &= \frac{g_2\left(h;\mathbf{x}_{g_2}\right)}{\gamma_1}, \\
\mathscr{q}\left(\mathbf{x}_q\right) &= \gamma_1\left(H_{\max}-H_{\min}\right)\frac{\pi_0^{\text{obs}}}{\left(1-\pi_0^{\text{obs}}\right)}q\left(\mathbf{x}_q\right) = \left(H_{\max}-H_{\min}\right)\frac{\pi_0^{\text{obs}}}{\left(1-\pi_0^{\text{obs}}\right)}\widetilde{q}\left(\mathbf{x}_q\right).
\end{aligned}$$

For the non–market alternative, we obtain:

$$\widetilde{\varphi}\left(0,0;\left(\mathbf{x}_V,\mathbf{x}_f,\mathbf{x}_{g_1},\mathbf{x}_{g_2},\mathbf{x}_q,\mathbf{x}_{\mathbb{P}}\right)|\mathcal{D}\right)$$

$$= \frac{\Psi\left(0,0h\mathbf{x}_V,\mathbf{x}_f\right)}{\Psi\left(0,0;\mathbf{x}_V,\mathbf{x}_f\right)+\sum\limits_{(s,t)\in\mathcal{D}\setminus\{(0,0)\}}\Psi\left(s,t;\mathbf{x}_V,\mathbf{x}_f\right)\mathscr{q}(\mathbf{x}_q)\mathscr{g}_1\left(s;\mathbf{x}_{g_1}\right)\mathscr{g}_2\left(t;\mathbf{x}_{g_2}\right)\varsigma\exp\left(\frac{(\ln s-\mu)^2}{2\varsigma^2}\right)}. \tag{26'}$$

This means that using equations (26)–(26') in the constructon of the likelihood function, would yield exactly the same estimates as using equations (24)–(24'), except for the (multiplicative) constant term of $\widetilde{q}\left(\mathbf{x}_q\right)$ in (24)–(24'), which provides an estimate of $\gamma_1\exp\eta_{q,0}$, while in equations (26)–(26'), the (multiplicative) constant of $\mathscr{q}\left(\mathbf{x}_q\right)$ is an estimate of
$\gamma_1\left(H_{\max}-H_{\min}\right)\frac{\pi_0^{\text{obs}}}{\left(1-\pi_0^{\text{obs}}\right)}\exp\eta_{q,0}$.
In both cases, one is able to back out an estimate of $\eta_{0,q}$ by using equation (23). In the first case, one subtracts $\ln\widehat{\gamma}_1$ (with $\widehat{\gamma}_1$, the estimated value of γ_1 from applying equation (23) using the estimates for γ_{k+1}, for $k=1,2,\ldots,K$) from the estimated constant of $\ln\left(\widetilde{q}\left(\mathbf{x}_q\right)\right)$. In the second case, $\ln\left(\widehat{\gamma}_1\left(H_{\max}-H_{\min}\right)\frac{\pi_0^{\text{obs}}}{\left(1-\pi_0^{\text{obs}}\right)}\right)$ needs to be subtracted from the estimated constant of $\ln\left(\mathscr{q}\left(\mathbf{x}_q\right)\right)$.

[20] That is the procedure that we will actually follow, with $\pi_0^{\text{obs}}=.104$, $\mu=2.71$, and $\varsigma=.308$ for males, and the corresponding number for females are .246, 2.63, and .297.

The estimates reported below in Table 3 are inclusive of γ_1, but do not contain the sampling correction terms. That is, we used specification (24)–(24').

3.4 Simulation

In order to evaluate the fit of the estimates, or the estimated model's prediction of behavioural reactions to changes in explanatory variables, a simulation method is used. A choice set is drawn (possibly capturing changes in the intensity with which certain alternatives become available to certain persons) and then it is determined what an agent's best choice would be within this simulated choice set, according to the estimated preferences of that person. If simulation is used for evaluating the fit of the estimated model, then the choice set is drawn according to the model estimates (the relative intensity of job offers, the wage offer distribution, and the labour time regime offers), and the simulated choices from that set are to be compared with actual ones, as observed in the data. This is done in Figures 7–12 below (Section 6.1). Next we will use the simulation method also for calculating elasticities, and to evaluate some counterfactuals (see Sections 6.2 and 7).
Actually, simulating with the estimated model can be done along two lines. Either one uses the estimated measure of intensity with which alternatives (w, h) are offered to an agent, that is, using the estimates of the q–function, the estimated wage offer distribution, g_1, *and* the estimated hours distribution, g_2, to sample a choice set $\{(w_r, h_r)\,; r = 1, 2, \ldots, R\}$. Next, one draws for each of the sampled alternatives, (w_r, h_r), a random variable from the Extreme Value Type I distribution[21], say $\epsilon\,(w_r, h_r)$. Next, it is evaluated which of the drawn alternatives yields highest utility: $\ln \widehat{V}\left(f\left(w_r, h_r; \mathbf{x}_f\right), T - h_r; \mathbf{x}_V\right) + \epsilon\left(w_r, h_r\right)$. The alternative (w_r, h_r) thus yielding the highest utility is considered to be the agent's optimal choice according to the model.
Alternatively, one draws from a prior, $\mathbb{P}_1\left(w, h; \mathbf{x}_{\mathbb{P}_1}\right)$, and looks which alternative has the highest likelihood to be chosen, that is one looks for the maximum of $\widehat{V}\left(f\left(w_r, h_r; \mathbf{x}_f\right), T - h_r; \mathbf{x}_V\right) \widehat{q}\left(\mathbf{x}_q\right) \widehat{g}_1\left(w_r; \mathbf{x}_{g_1}\right) \widehat{g}_2\left(h_r; \mathbf{x}_{g_2}\right) \frac{\mathbb{P}_1\left(0,0;\mathbf{x}_{\mathbb{P}_1}\right)}{\mathbb{P}_1\left(w_r,h_r;\mathbf{x}_{\mathbb{P}_1}\right)}$ over the drawn elements (w_r, h_r) for which $w_r, h_r > 0$. Denote the alternative yielding the maximum by (w_{r^*}, h_{r^*}). The alternative chosen according to this second simulation method coincides with this maximum, (w_{r^*}, h_{r^*}), *if*

$$\widehat{V}\left(f\left(w_{r^*}, h_{r^*}; \mathbf{x}_f\right), T - h_{r^*}; \mathbf{x}_V\right) \widehat{q}\left(\mathbf{x}_q\right) \widehat{g}_1\left(w_{r^*}; \mathbf{x}_{g_1}\right) \widehat{g}_2\left(h_{r^*}; \mathbf{x}_{g_2}\right) \frac{\mathbb{P}_1\left(0,0;\mathbf{x}_{\mathbb{P}_1}\right)}{\mathbb{P}_1\left(w_{r^*},h_{r^*};\mathbf{x}_{\mathbb{P}_1}\right)} > \widehat{V}\left(f\left(0,0;\mathbf{x}_f\right), T; \mathbf{x}_V\right).$$

Else, the non–market alternative is the simulated optimal choice.
The practice in Statistics Norway, by Aaberge and his collaborators, has been to follow a go between. Wages are drawn from the estimated wage distribution $\widehat{g}_1\left(w; \mathbf{x}_{g_1}\right)$. Hours and the

[21] The log of a Fréchet distributed random term is Extreme Value Type I distributed.

proportion of market versus non–market alternatives are drawn from the same priors as for the sampling of the choice set for the estimation: the uniform distribution on $[H_{\min}, H_{\max}[$ for the labour time, and the observed participation degree in the sample, $\left(1-\pi_0^{\text{obs}}\right)$, for the number of market alternatives. Then, the simulated optimal choice, $(w_{r^{**}}, h_{r^{**}})$, say, is determined as follows:

$$\begin{aligned}
&(w_{r^{**}}, h_{r^{**}}) := \\
&\underset{(w_r,h_r)\in\{\{(0,0)\}\cup\{(w_{r^*},h_{r^*})\}\}}{\arg\max} \left\{\mathcal{A}, \ln\left(\widehat{\Psi}\left(w_{r^*}, h_{r^*}; \mathbf{x}_V, \mathbf{x}_f\right)\widehat{q}\left(\mathbf{x}_q\right)\widehat{g}_2\left(h_{r^*}; \mathbf{x}_{g2}\right)\right) + \epsilon\left(w_{r^*}, h_{r^*}\right)\right\}, \\
&\text{where} \quad \mathcal{A} \;:=\; \left\{\ln\left(\widehat{\Psi}\left(w_r, h_r; \mathbf{x}_V, \mathbf{x}_f\right)\right) + \epsilon\left(w_r, h_r\right), \forall\, r : (w_r, h_r) = (0,0)\right\}, \text{ and} \\
&(w_{r^*}, h_{r^*}) := \underset{(w_r,h_r):w_r,h_r>0}{\arg\max} \left\{\ln\left(\widehat{\Psi}\left(w_r, h_r; \mathbf{x}_V, \mathbf{x}_f\right)\widehat{q}\left(\mathbf{x}_q\right)\widehat{g}_2\left(h_r\right)\right) + \epsilon\left(w_r, h_r\right), \forall\, r : w_r, h_r > 0\right\}.
\end{aligned} \tag{27}$$

Notice that if one would have formulated the likelihood including the correction terms for sampling hours from the uniform distribution, and non–market opportunities from the within sample observed probability of being non–active, the equivalent of the expressions over which the maximum is taken in the bottom line of equation (27), would be:

$$\ln\left(\widehat{\Psi}\left(w_r, h_r; \mathbf{x}_V, \mathbf{x}_f\right)\widehat{\widetilde{q}}\left(\mathbf{x}_q\right)\widehat{g}_2\left(h_r; \mathbf{x}_{g2}\right)\right) + \ln\left(H_{\max} - H_{\min}\right) + \ln\left(\frac{\pi_0^{\text{obs}}}{1-\pi_0^{\text{obs}}}\right) + \epsilon\left(w_r, h_r\right).$$

It can be seen from the definition of $q\left(\mathbf{x}_q\right)$ below equation (26), and the definition of $\widetilde{q}\left(\mathbf{x}_q\right)$ below equation (24), that both expressions are equivalent. Furthermore,

$$\ln\left(\widehat{\widetilde{q}}\left(\mathbf{x}_q\right)\widehat{g}_2\left(h_r; \mathbf{x}_{g2}\right)\right) = \widehat{\boldsymbol{\eta}}_q'\mathbf{x}_q + \ln\widehat{\gamma}_1 + \sum_{k=1}^{K}\chi\left(h_r \in \left[\underline{H}_k, \overline{H}_k\right[\right)\widehat{\gamma}_{k+1},$$

where $\chi\left(\texttt{condition}\right)$ is the indicator function, which equals one if the condition serving as its argument is satisfied, and equals zero otherwise.
Another in between simulation method is continuing to draw hours from the uniform distribution on $[H_{\min}, H_{\max}[$, but using the estimated rate of market to non–market opportunities, $\widehat{q}\left(\mathbf{x}_q\right)$, to determine the intensity of sampling formal labour market jobs. The criterion to determine the most preferred job offer from the sampled set, (w_{r^*}, h_{r^*}), that replaces the bottom line of equation (27) in this case, becomes:

$$\begin{aligned}
&(w_{r^*}, h_{r^*}) := \\
&\underset{(w_r,h_r):w_r,h_r>0}{\arg\max} \left\{\ln\left(\widehat{\Psi}\left(w_r, h_r; \mathbf{x}_V, \mathbf{x}_f\right)\widehat{\gamma}_1\left(H_{\max} - H_{\min}\right)\widehat{g}_2\left(h_r; \mathbf{x}_{g2}\right)\right) + \epsilon\left(w_r, h_r\right), \forall\, r : w_r, h_r > 0\right\},
\end{aligned} \tag{28}$$

irrespective of whether the sampling correction terms has been included in the likelihood or not.

In the results presented below in section 6.1 the first method of simulating was used. That is: market alternatives were drawn from the estimated wage offer and offered labour time regimes, and the relative number of job offers tot non–market alternatives is determined by the estimated q-function. A similar approach is followed for constructing the counterfactuals in the simulation exercise of section 7.

4 Data

The model is estimated on the Belgian database of the European Union Statistics on Income and Living Conditions (EU–SILC). We use the data that were collected in 2007. The entire dataset consists of 6348 households or 15493 individuals. It is representative for the Belgian population of private households. Persons living in collective households or institutions are excluded from the target population. The survey provides detailed information on earnings as well as on socio–demographic characteristics of each household.

In order to estimate the model, we relied on three different sub–samples of couples, single females and single males. Each sub–sample consists exclusively of individuals that are available for the labour market; *i.e.* aged between 16 and 64 year and not being sick, in education, disabled or (pre)retired. Self–employed are excluded due to the lack of reliable information on hours worked and income earned. Mixed households in which only one of the partners is available for the labour market are also excluded from the estimation sample. Finally, we drop households whose children are already available for the labour market but are still living with their parents. It is reasonable to assume that their labour supply decisions are different from those of a normal household without working children because it is not clear whether these households consider their labour supply decision as a collective or as an individual process. Given this data selection, we are able to estimate the labour supply model on 1457 couple households, 571 single females, and 449 single males. In total, they represent approximately 41% of the total population in Belgium.

EUROMOD is used as microsimulation tool for the calculation of net disposable income for each element in the opportunity set of households.[22] Gross household labour income is equal to the sum of labour earnings of all household members. The income tax and employee's social security contributions are deducted from gross income, and social transfers such as social assistance, unemployment benefits, child benefits, education benefits and housing benefits are added. We assume full take–up of social assistance if the eligibility criteria are fulfilled.

Descriptive statistics for the selected sub–samples can be found in Table 1. In the wage

[22] Version f5.5 was used. More information about EUROMOD can be found at their website: `https://www.iser.essex.ac.uk/euromod`.

Table 1: Descriptive statistics for the estimation sample

	Singles		**Couples**	
Description	*Female*	*Male*	*Female*	*Male*
Age (years)	41.1	39.9	38.1	40.2
% hh having 0-3 year old children	5.78%	0.45%	18.67%	18.67%
% hh having 4-6 year old children	9.46%	0.89%	17.16%	17.16%
% hh having 7-9 year old children	10.16%	1.78%	18.19%	18.19%
Potential experience (years)	21.3	20.6	18.0	20.4
Education:				
Low educated	22.8%	24.5%	16.8%	19.8%
Secondary education	34.6%	41.9%	38.5%	39.0%
High educated	42.6%	33.6%	44.7%	41.2%
Residence:				
Brussels	19.8%	21.2%	9.3%	9.3%
Flanders	44.1%	45.2%	58.5%	58.5%
Wallonia	36.1%	33.6%	32.3%	32.3%
Participation rate (%)	68.12	78.84	79.40	93.20
Hours worked/week:				
Conditional on working	35.88	39.69	32.50	40.84
Unconditional	24.45	31.29	25.81	38.06
Hourly wage (euro)	14.91	15.20	14.73	16.25
Disposable income (€ /month)	1567	1588	3143	
Number of observations	571	449	1457	

Source: Own Calculations, EU-SILC 2007

offer equation an indicator for experience is used. Since we do not have information on the number of years a person has actually been working since she entered the labour market, *potential* experience is used. It is defined as the number of years since the person entered the labour market. That is age *minus* 15 years for a lowly educated person, age *minus* 19 years for a middle educated person, and age *minus* 23 years for a highly educated person. As this variable is highly correlated with age, age will not be included as a separate variable in the wage offer equation.

Besides information from the EU–SILC questionnaire, we also used external information on type specific unemployment. This variable should serve as a proxy for job availability, and may help to identify the distinction between the contribution of opportunities and preference factors in the model. Appendix III contains more information on this variable.

5 Estimation results

Table 2 specifies the covariates that have been used in the different parts of the model. In Table 3 we report estimated parameters for the model. Next, we investigate the impact of age on preference intensity for leisure (as compared to consumption), and opportunities.

Table 2: Model specification

	Preferences	Opportunities		
	$\mathbf{x}_V$	$\mathbf{x}_q$	$\mathbf{x}_{g_2}$	$\mathbf{x}_{g_1}$
variable		job offers	hours	wages
Regional dummies[a]	yes	yes	no	no
Education dummies[b]	yes	yes	no	yes
Age	yes	yes	no	no
Group specific unemployment rate	no	yes	no	no
Number of children	yes	no	no	no
Gender	yes	yes	yes	yes
Potential experience	no	no	no	yes

[a] Bxl=Brussels, Fl=Flanders, Wal=Wallonia.
[b] Low, Middle, High.

Table 3: Estimation results: Preferences couples

Procedure to sample choice set: importance sampling[a] with replacement and expected number of non–market alternatives equal to π_0^{obs}			
Log likelihood	-8427.3013		
Description	Estimate	Standard Error	t–value
1.a) Consumption & leisure interaction M&F			
Consumption Couples exponent	0.579	0.054	10.69
Consumption Couples constant	4.763	0.310	15.35
Leisure interaction M&F.in couples	0.128	0.052	2.45
Consumption single M exponent	0.261	0.132	1.98
Consumption single M constant	4.517	0.401	11.25
Consumption single F exponent	-0.151	0.166	-0.91
Consumption single F constant	4.239	0.337	12.58
1.b) Leisure coefficients males in couples			
Leisure M in couples exponent	-9.029	0.707	-12.77
Leisure M in couples constant	11.153	5.533	2.02
Leisure M in couples ln(age)	-5.890	3.023	-1.95
Leisure M in couples ln(age)^2	0.821	0.418	1.97
Leisure M in couples ch03	-0.003	0.047	-0.06
Leisure M in couples ch36	0.056	0.051	1.10
Leisure M in couples ch69	-0.011	0.047	-0.24
Leisure M in couples dum region Walloon[b]	0.106	0.055	1.93
Leisure M in couples dum region Brussels	0.131	0.090	1.45
Leisure M in couples dum education LOW[c]	-0.127	0.067	-1.89
Leisure M in couples dum education HIGH	-0.070	0.048	-1.47
1.c) Leisure coefficients females in couples			
Leisure F in couples exponent	-7.673	0.552	-13.90
Leisure F in couples constant	-7.855	11.558	-0.68
Leisure F in couples ln(age)	4.599	6.583	0.70
Leisure F in couples ln(age)^2	-0.492	0.932	-0.53
Leisure F in couples ch03	0.437	0.162	2.70
Leisure F in couples ch36	0.468	0.171	2.74
Leisure F in couples ch69	0.354	0.172	2.06
Leisure F in couples dum region Walloon	0.243	0.145	1.67
Leisure F in couples dum region Brussels	0.014	0.208	0.07
Leisure F in couples dum education LOW	0.673	0.293	2.29
Leisure F in couples dum education HIGH	-0.713	0.162	-4.42

[a] Importance sampling is explained in Appendix II.
[b] Flanders region is reference category.
[c] Middle education level is reference category.

Table 3: Estimation results ctd.: Preferences singles

Procedure to sample choice set: importance sampling with replacement and expected number of non–market alternatives equal to π_0^{obs}			
Description	Estimate	Standard Error	t–value
1.d) Leisure coefficients single males			
Leisure single M exponent	-6.169	1.172	-5.26
Leisure single M constant	-5.752	17.286	-0.33
Leisure single M ln(age)	3.567	9.722	0.37
Leisure single M ln(age)^2	-0.403	1.350	-0.30
Leisure single M ch36	-0.474	0.903	-0.52
Leisure single M ch69	-0.906	0.634	-1.43
Leisure single M dum region Walloon	0.835	0.395	2.12
Leisure single M dum region Brussels	0.211	0.303	0.70
Leisure single M dum education LOW	-0.356	0.321	-1.11
Leisure single M dum education HIGH	-0.507	0.290	-1.75
1.e) Leisure coefficients single females			
Leisure single F exponent	-9.328	1.259	-7.41
Leisure single F constant	22.335	12.697	1.76
Leisure single F ln(age)	-11.839	6.938	-1.71
Leisure single F ln(age)^2	1.637	0.959	1.71
Leisure single F ch03	0.769	0.420	1.83
Leisure single F ch36	0.061	0.158	0.39
Leisure single F ch69	-0.054	0.133	-0.41
Leisure single F dum region Walloon	0.121	0.136	0.89
Leisure single F dum region Brussels	-0.140	0.122	-1.14
Leisure single F dum education LOW	0.213	0.247	0.86
Leisure single F dum education HIGH	-0.569	0.186	-3.05

Table 3: Estimation results ctd.: Relative intensity of market alternatives and peaks hours

Procedure to sample choice set: importance sampling with replacement and expected number of non–market alternatives equal to π_0^{obs}

Description	Estimate	Standard Error	t–value
2.a) Estimated coefficients opportunities & peaks males			
Opportunity M constant	-63.435	18.410	-3.45
Opportunity M unemployment rate	-0.491	0.420	-1.17
Opportunity M dummy region Walloon	-0.645	0.228	-2.83
Opportunity M dummy region Brussels	-1.351	0.289	-4.67
Opportunity M dummy LOW education	-0.256	0.363	-0.71
Opportunity M dummy HIGH education	-0.103	0.274	-0.38
Opportunity M ln(age)	35.200	9.878	3.56
Opportunity M ln(age)^2	-5.152	1.330	-3.87
Peaks M $<18.5,20.5>$ interval	0.654	0.228	2.86
Peaks M $<29.5,30.5>$ interval	0.874	0.189	4.61
Peaks M $<37.5,40.5>$ interval	2.694	0.060	45.21
2.b) Estimated coefficients opportunities & peaks females			
Opportunity F constant	-68.299	13.705	-4.98
Opportunity F unemployment rate	-0.568	0.200	-2.83
Opportunity F dummy region Walloon	-0.494	0.157	-3.14
Opportunity F dummy region Brussels	-0.851	0.219	-3.89
Opportunity F dummy LOW education	0.285	0.256	1.11
Opportunity F dummy HIGH education	0.212	0.209	1.02
Opportunity F ln(age)	38.169	7.460	5.12
Opportunity F ln(age)^2	-5.597	1.019	-5.49
Peaks F $<18.5,20.5>$ interval	1.647	0.100	16.45
Peaks F $<29.5,30.5>$ interval	1.797	0.108	16.63
Peaks F $<37.5,40.5>$ interval	2.177	0.070	31.20

Table 3: Estimation results ctd.: Wage distribution

Procedure to sample choice set: importance sampling with replacement and expected number of non–market alternatives equal to π_0^{obs}			
Description	Estimate	Standard Error	t–value
3. Estimated coefficients wage equations			
3.a) Wage equation males			
Wage M σ	0.266	0.004	59.48
Wage M constant	2.074	0.029	70.87
Wage M potential experience	2.233	0.250	8.92
Wage M potential experience^2	-2.844	0.560	-5.08
Wage M low education	-0.155	0.019	-8.03
Wage M high education	0.267	0.015	17.45
3.b) Wage equation females			
Wage F σ	0.263	0.005	58.44
Wage F constant	2.053	0.027	77.02
Wage F potential experience	2.303	0.242	9.50
Wage F potential experience^2	-3.273	0.603	-5.43
Wage F low education	-0.107	0.024	-4.54
Wage F high education	0.299	0.016	18.81

5.1 Preferences

The marginal rate of substitution between consumption and labour time for singles equals:

$$\mathrm{MRS}_{c,h_j} = \frac{\left(\boldsymbol{\beta}'_{h,j}\mathbf{x}_V\right) \cdot \left((T-h_j)\right)^{\alpha_{h,j}-1} / T^{\alpha_{h,j}}}{\beta_{c,j} \cdot c_j^{\alpha_{c,j}-1}}, \qquad j = 1,2, \tag{29}$$

and for couples it is equal to:

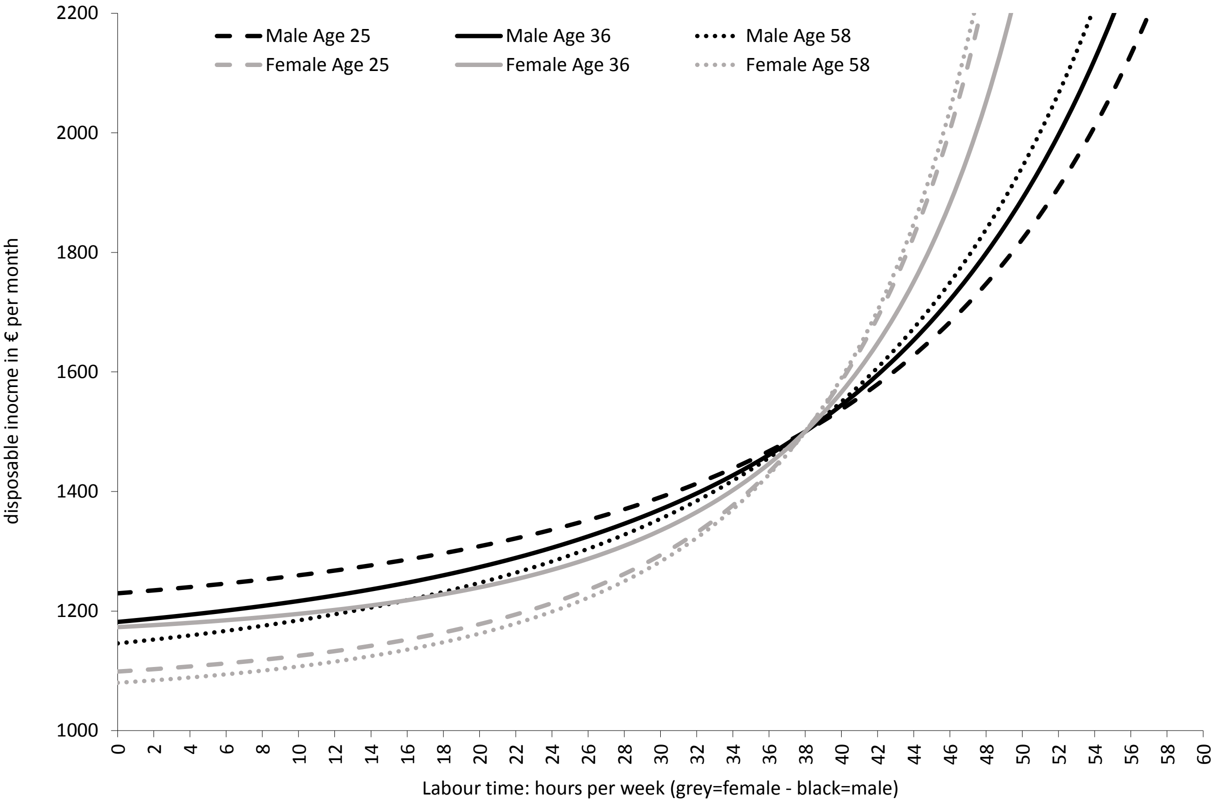
$$\mathrm{MRS}_{c,h_j} = \frac{\left(\boldsymbol{\beta}'_{h_j}\mathbf{x}_V\right) \cdot (T-h_j)^{\alpha_{h_j}-1} / T^{\alpha_{h_j}} + \beta_{h_1,h_2}\left(\frac{((T-h_i)/T)^{\alpha_{h_i}}-1}{\alpha_{h_i}}\right)\left((T-h_j)\right)^{\alpha_{h_j}-1} / T^{\alpha_{h_j}}}{\beta_{c,g} \cdot c^{\alpha_{c,g}-1}}, \tag{30}$$

$$i,j = 1,2;\ i \neq j.$$

Notice that the covariates influencing preferences affect the marginal rate of substitution only through their influence on $\left(\boldsymbol{\beta}'_{h_j}\mathbf{x}_V\right)$. More specifically, as $\left(\boldsymbol{\beta}'_{h_j}\mathbf{x}_V\right)$ increases in one of the covariates, the marginal rate of substitution in any point (c, h_j) becomes higher for persons with a larger value on that variable. Hence, a person exhibiting a higher value on that covariate will have relatively steeper indifference curves. That is, she will exhibit a more intense preference for leisure.

We illustrate this in Figures 2 and 3 for the case of age. For single males, age has a concave parabolic impact on $\left(\boldsymbol{\beta}'_{h_j}\mathbf{x}_V\right)$ with the top situated at 93 years. So, the marginal rate of

Figure 2: Impact of age on steepness of indifference curves: singles

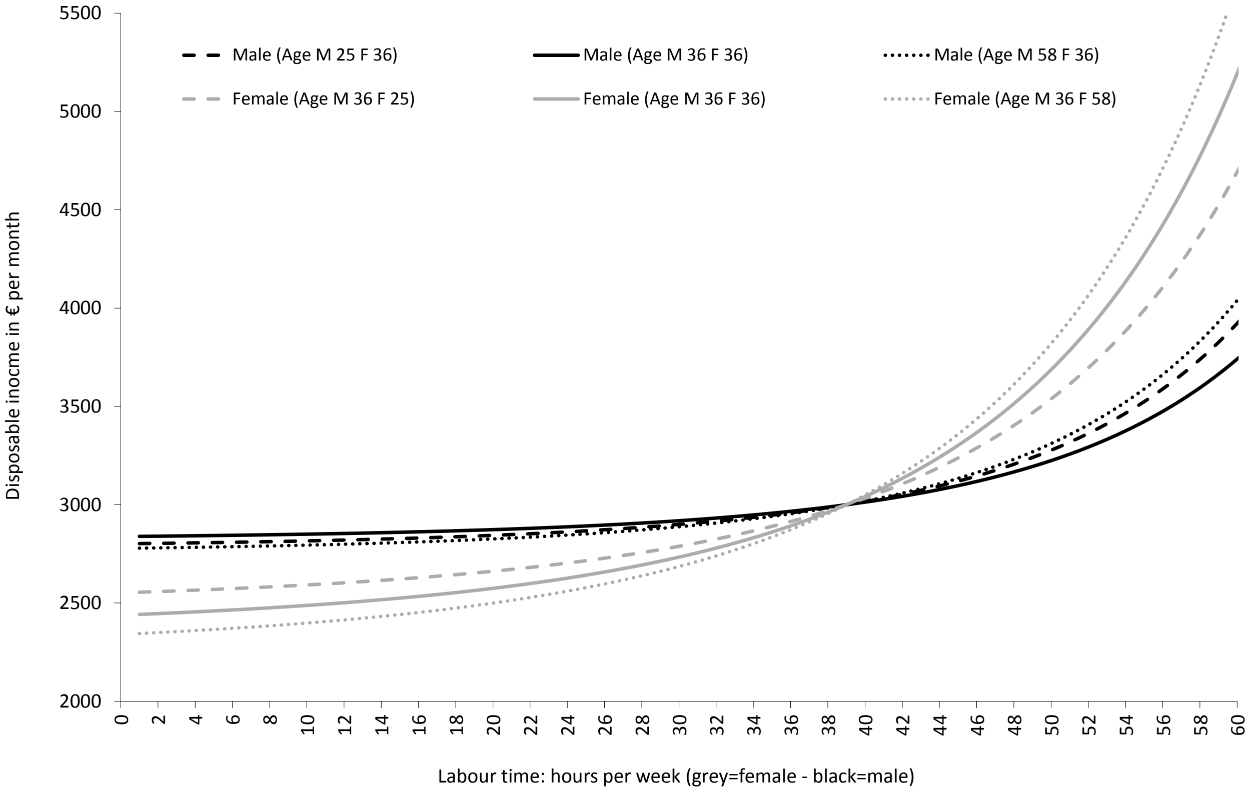


substitution between consumption and labour time, and thus intensity of preferences for leisure, increases with age in the range of the values obtained for age in our data (16–64 years). This is illustrated by the black indifference curves in consumption labour time space in Figure 2: the dashed curve applies to a single male of 25 years, the steeper full curve to someone of 36 years old, and the steepest, dotted curve to someone of 58 years old. The curves become steeper with age.

For single females, the influence of age on $\left(\boldsymbol{\beta}'_{h_j}\mathbf{x}_V\right)$ is convex parabolic, with the bottom at the age of 37 years. So, the indifference curves for consumption and labour time become flatter from the age of 16 until that of 37 years, and their slope starts increasing when becoming older then 37 years. This is illustrated by the gray curves in Figure 2. The dashed (age 25 years) and dotted curves (age 58 years) are both steeper than that of a single female at the age of 36 years. Preference intensity for leisure of single females is *lowest* at the age of 37 years.

At first sight it seems that females have more intense preference for leisure than males (compare the grey and black curves in Figure 2), but note that the indifference curves are not single crossing, so that a classification of intensity of preferences for leisure with respect to sex is not easily made.

Figure 3: Impact of age on steepness of indifference curves: couples



For males, respectively females in couples the situation is reversed as compared to the case of singles (see Figure 3). Female spouses have increasing intensity of preference for leisure with age until the age of 107 years. For males the least intense preference for leisure is reached at the age of 36 years.

As far as the significance of these effects is concerned, the precision of the estimates of the age effect of females in couples and that of single males is poor. Those of single females and males in couples are reasonable.

As for the impact of education, the intensity of preference for leisure relative to consumption of females is decreasing with education level, irrespective whether they are single, or live in a couple. The effect of a lower education level is small though for singles and not really significantly different from that of middle education. For males, the situation is again different, with both higher and lower educated men having less intense preference for leisure relative to consumption, irrespective whether they are single or live in couples. The precision of the estimate for lowly educated single men is poor, while the other education coefficients are moderately well estimated.

5.2 Opportunities

Figure 4 represents the estimation of the wage offer distributions differentiated by sex and education level. Notice that this is a wage *offer* distribution. Simulated and observed wages are discussed in Section 6.1. The estimates of these distributions are very precise, and robust to different specifications for other parts of the model.

From Figure 4 it can be seen that higher education shifts the wage offer distribution to the right, both for males and females. A similar effect is obtained for most of the range of potential experience. Table 4 reports *e.g.* the means of the wage offer distribution for 10 and 25 years of potential experience at different education levels, both for males and females. The mean of the wage offers increases with 1.7 to 2.7 euro over this range. However, as potential experience surpasses 35 years for females, and 39 years for men, the effect of additional experience becomes slightly negative. Potentially, the variable takes up an age effect here. However, age and potential experience were by construction too much correlated to include them separately in this part of the model.

The differences between males and females are small (dashed lines in Figure 4 apply to females, the full lines to males) as compared to the impact of the other covariates, and not always in the disadvantage of females. The latter is the case for persons with a middle education level.

Table 4: Mean of the wage offer distributions by sex, education and experience

Gross wage (euro per hour)

	Male		Female	
	Years of potential experience			
Education level	10 year	25 year	10 year	25 year
Low	8.58	10.33	8.83	10.50
Middle	10.02	12.06	9.83	11.69
High	13.08	15.74	13.25	15.76

Figure 5 represents the distributions of offered labour time regimes by sex.[23] Again, these are not actual labour time regimes nor the ones chosen according to the model. The most salient observation is that this distribution is different for males and for females, the latter receiving more part time, and less full time job offers.

[23] Admittedly, this part of the model is not non–parametrically identified. So, if one would like to explain this peak pattern by differences in preferences, we cannot tell this to be wrong on pure empirical grounds.

Figure 4: Estimated wage offer distributions differentiated by sex and education

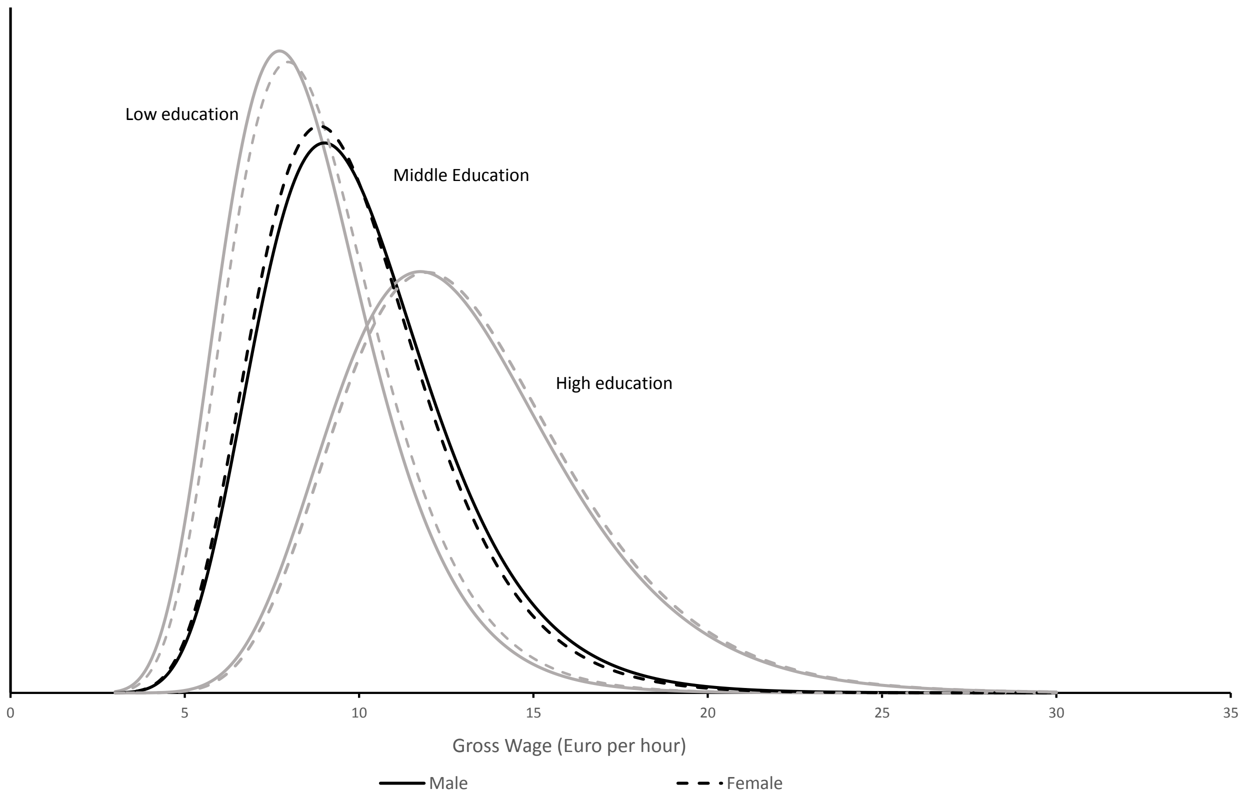


Figure 5: Estimated distribution of offered labour time regimes

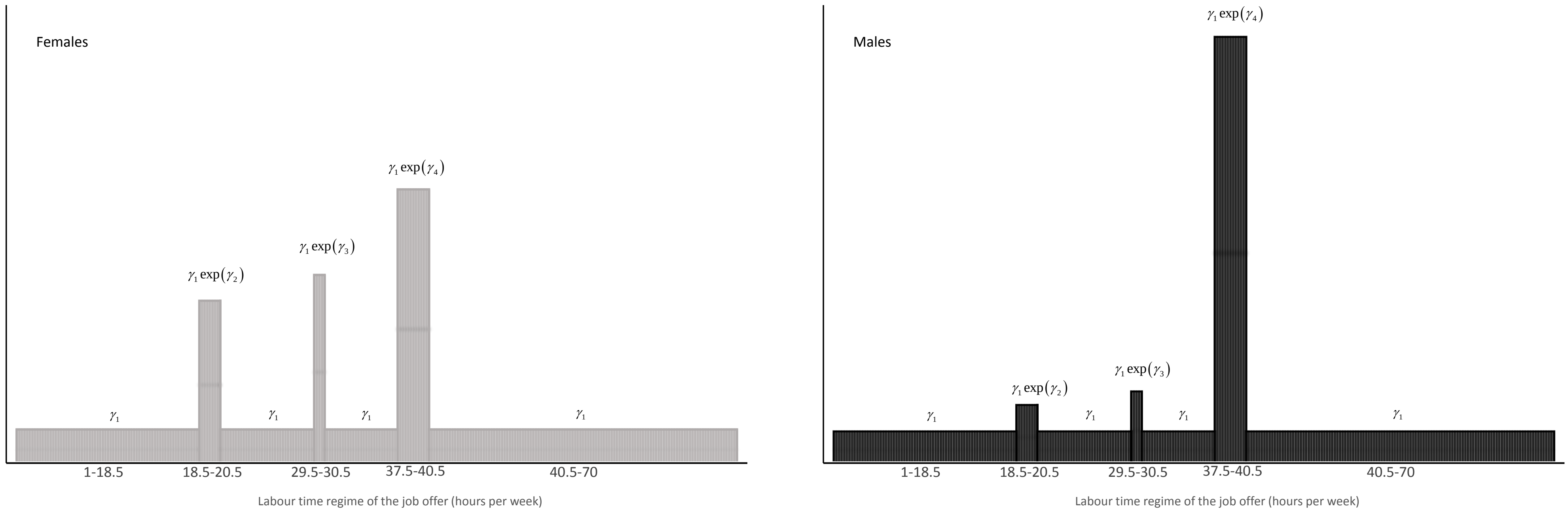
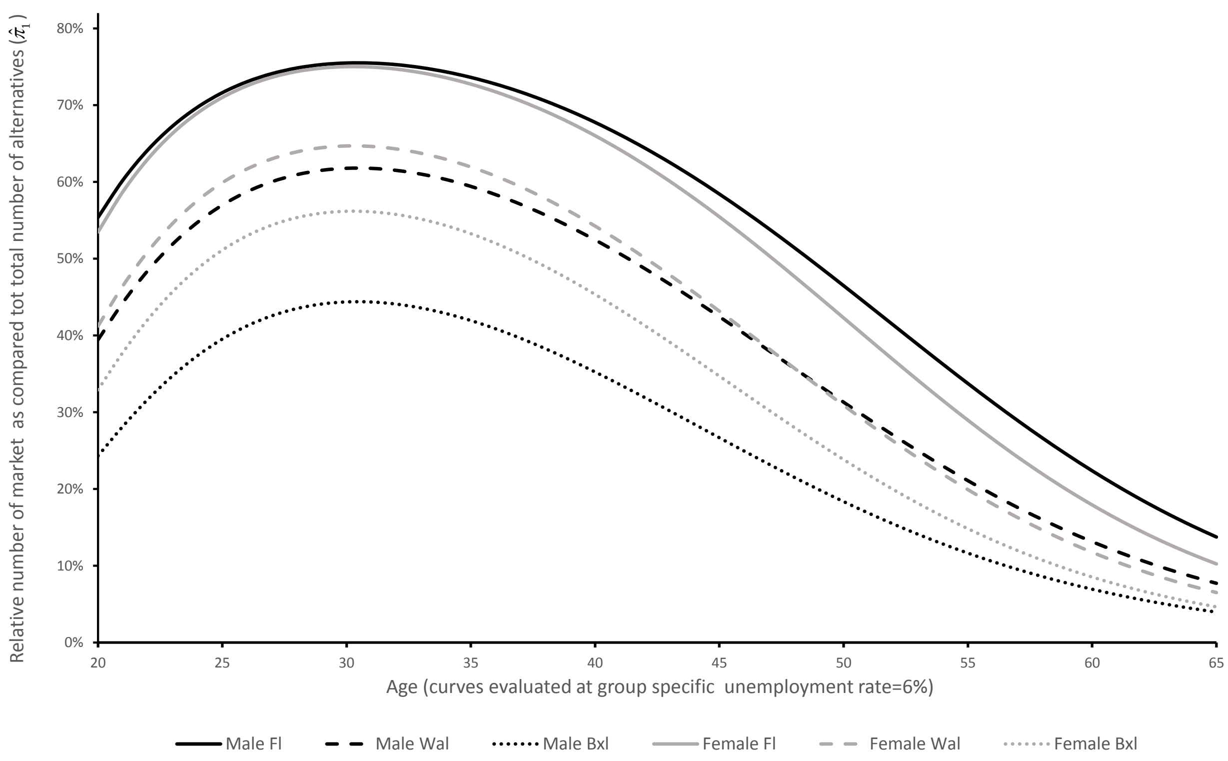


Figure 6 and Table 5 report the impact of the covariates $\mathbf{x}_q$ on the intensity of job offers (the π_1–function). This variable measures the extent to which vacancies posted on the labour market are suited for the capacities a person is endowed with. As RURO is a static model, the total stock of capacities of a person is assumed to be fixed. As such, the estimated π_1 can be interpreted as the percentage of alternatives available to an individual which include a job offer.

Grey curves in Figure 6 apply to females, black ones to males. The reference category (full lines) applies to a person living in Flanders region (Fl) with middle education level. Job offer intensity increases with age until the age of about 30 years, while it (quite drastically) decreases afterwards, a little more so for females than for males.

Dashed and dotted curves reflect the impact of region when compared to the corresponding full lines. Job offer intensity is lower in Walloon (Wal, dashed lines), and especially in Brussels capital region (Bxl, dotted lines).

Figure 6: Intensity of demand for individual capacities on the labour market in function of age, and differentiated by sex and region



The sign of the effect of education on π_1 does not depend on the level of type–specific unemployment rate nor on the age of a person. We report therefore in Table 5 the impact of

Table 5: Intensity of demand for individual capacities on the labour market
Evaluated at age=30 years, and type specific unemployment rate=6%

	Male			Female		
			Region			
Education level	Bxl	Fl	Wal	Bxl	Fl	Wal
Low	38.2%	70.5%	55.6%	63.0%	80.0%	70.9%
Middle	44.4%	75.5%	61.8%	56.2%	75.0%	64.7%
High	41.8%	73.5%	59.3%	61.3%	78.8%	69.4%

education at fixed values for age (30 years) and type specific unemployment rates (6%). Surprisingly, for males high education lowers job offer intensity, while for females low education raises job offer intensity. However, the effects are small, certainly in the light of the rather large standard errors of the estimated coefficients for these variables.

6 Fit and behavioural response

6.1 Fit

We now evaluate the fit of the estimates reported in Table 3 using the first simulation method discussed in Section 3.4. We first report results for couples (Figures 7–9), and then for singles (Figures 10–12).
We compare the marginal distributions of 'observed' and simulated disposable income, which equals the value for consumption, c, in the model[24] (Figures 7 and 10), observed and simulated wages of males and females (Figures 8 and 11), and observed and simulated hours of work (Figures 9 and 12). The curves labelled by `observed` refer to observed values, while `estimated` refers to simulated values by the estimated model. Observed and simulated consumption and wages are compared by means of Gaussian kernel densities of both distributions. For wages, the distributions are conditional on being positive (labour market participants only). For the labour time regimes we construct a histogram with bins coinciding with the peaks and troughs of the offered labour time regimes distributions. In this way, one can get a first assessment of the extent to which differences in offered labour time regimes are reflected in actual choices. We come back to this distinction between offered alternatives and actual choices while discussing Figure 12.
Of course, even if the observed and simulated values would perfectly coincide, there might still occur large differences between simulated and observed values for each individual separately.

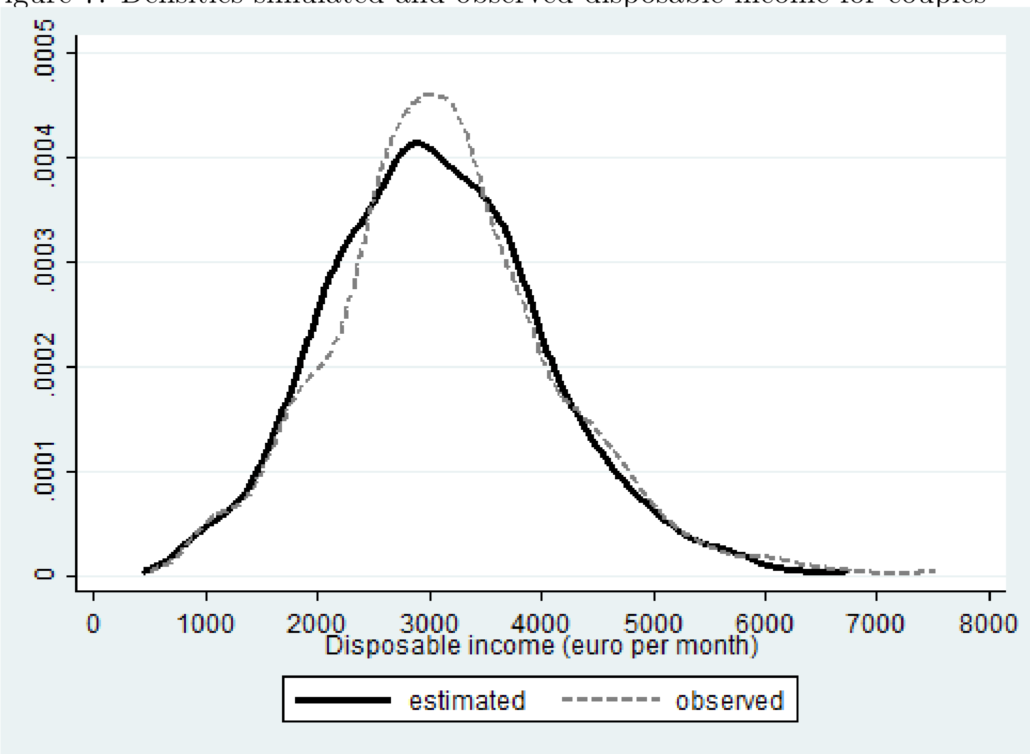
[24] Disposable income is actually not an observed value, but calculated from gross income by EUROMOD.

The transition matrices between observed and simulated labour time regimes are reported in Appendix IV.

1. Couples.

   The mean estimated consumption within couples is 3083 € per month. Compare that with the observed mean of 3143 € per month reported in Table 1. The simulated distribution (black curve in Figure 7) slightly overestimates the number of households with lower incomes, at the expense of those with modal disposable incomes (compare the simulated values, represented by the black curve, with the observed values represented by the dashed grey one).

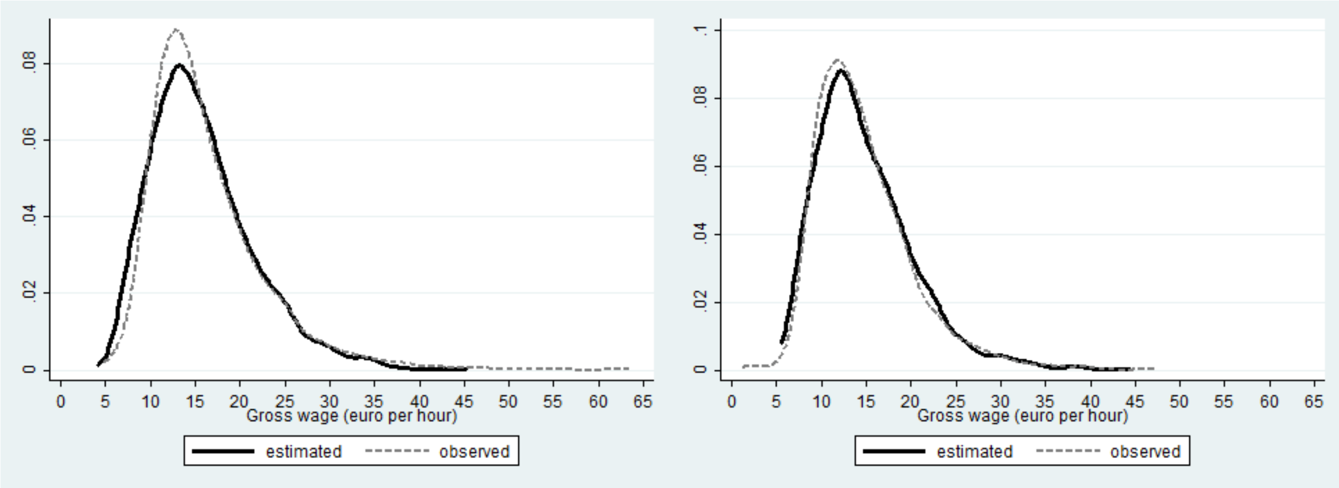
Figure 7: Densities simulated and observed disposable income for couples



Curves labelled by `estimated` refer to the simulation obtained from the estimated model, while `observed` refers to observed values.

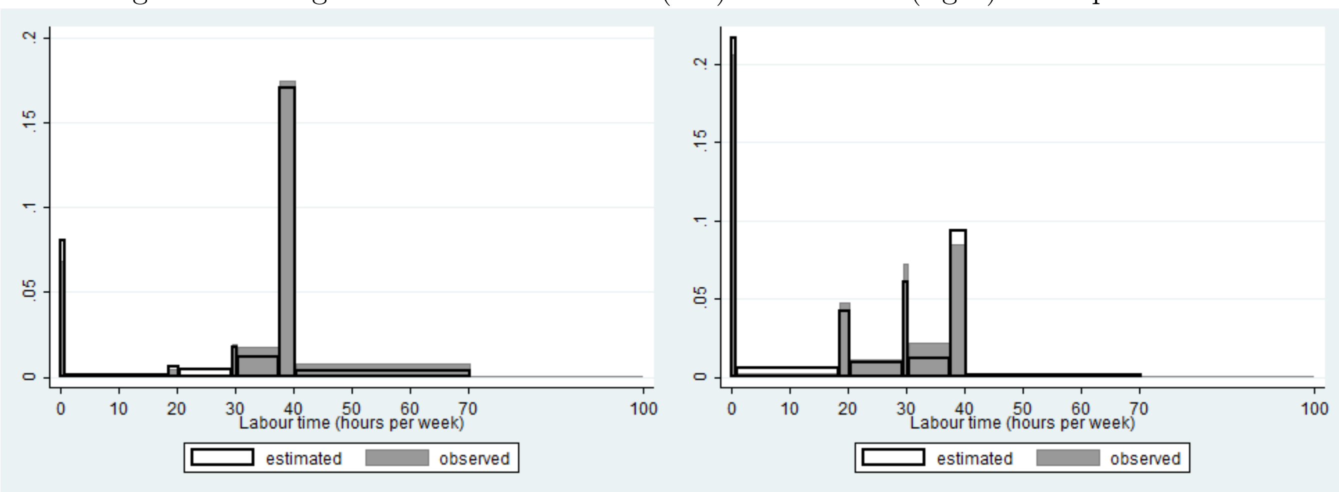
The simulated (conditional) distribution of female wages (black curve on the RHS panel of Figure 8) is fatter at moderately low and moderately high wages, and less densely populated at modal incomes, as compared to the observed one (grey dashed curve on the RHS panel of Figure 8). Similarly, the simulated wage distribution of the males (black line on the LHS panel of Figure 8) is more populated at lower wages than the observed one (dashed grey line on the LHS panel of Figure 8), at the expense of a smaller occurrence of modal wages.

Figure 8: Densities simulated and observed wages males (left) and females (right) in couples



The number of non–participants in the labour market is overestimated. Compare thereto the filled grey (observed) and unfilled black bordered (simulated) left most spikes in both panels of Figure 9. Still, the number of cases in which none of both partners work, is underestimated (2.6% simulated as compared to 3.9% in the sample). The estimated peaks are in general somewhat underestimated, except for the number of females accepting a full time job, which is overestimated. More than full time jobs by males, and jobs between three quarter and full time are underestimated. For females the occurrence of jobs less than half time is overestimated.

Figure 9: Histograms labour time males (left) and females (right) in couples



2. Singles.

   Figures 10–12 represent the fit of the model for singles. The occurrence of moderately low consumption levels of single females is underestimated, while that of modal incomes is overestimated (RHS of Figure 10). Mean estimated consumption of females is 1575 € per month, to be compared with the observed value of 1567 € per month observed (see Table 1). For males these figures are respectively 1571 € per month fitted *versus* 1588 € per month observed. The empirical distribution is somewhat less skewed to the right than the fitted one. That is to a lesser extent also the case for the single females.

Figure 10: Densities disposable income single males (left) and single females (right)

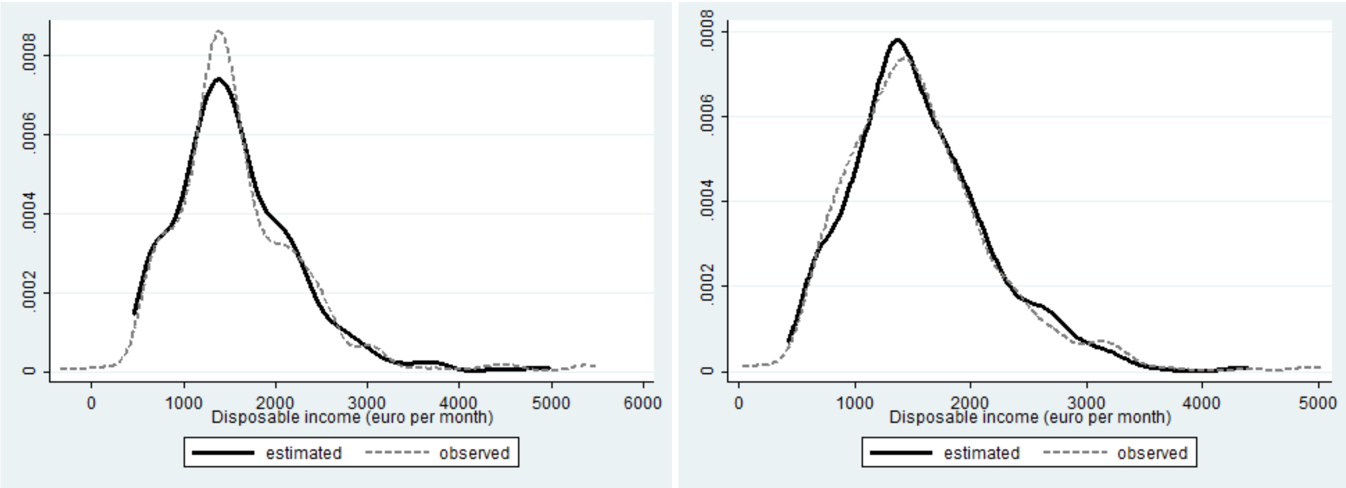
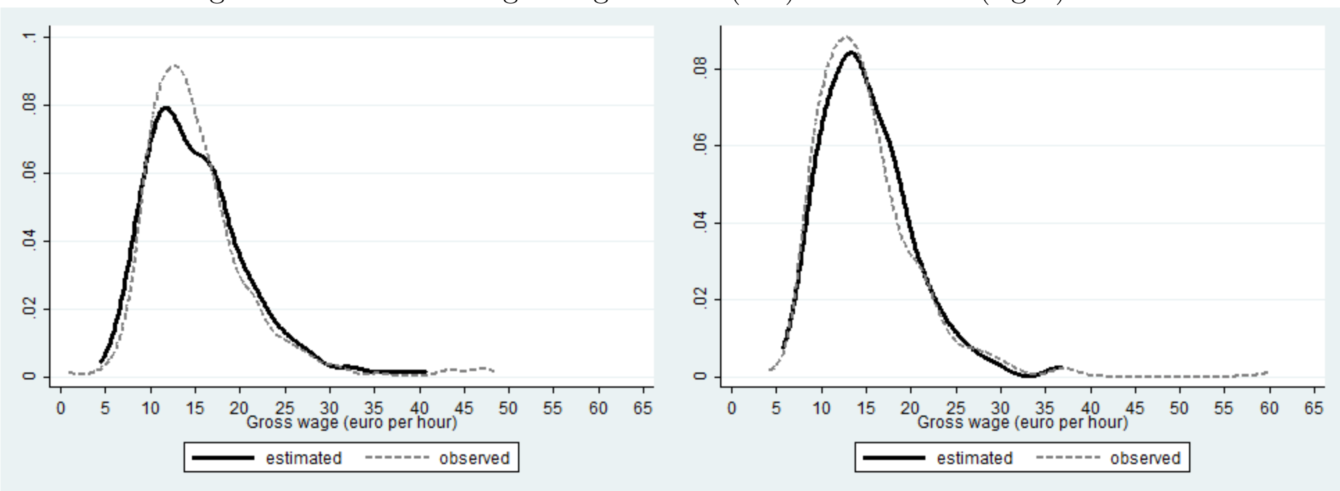
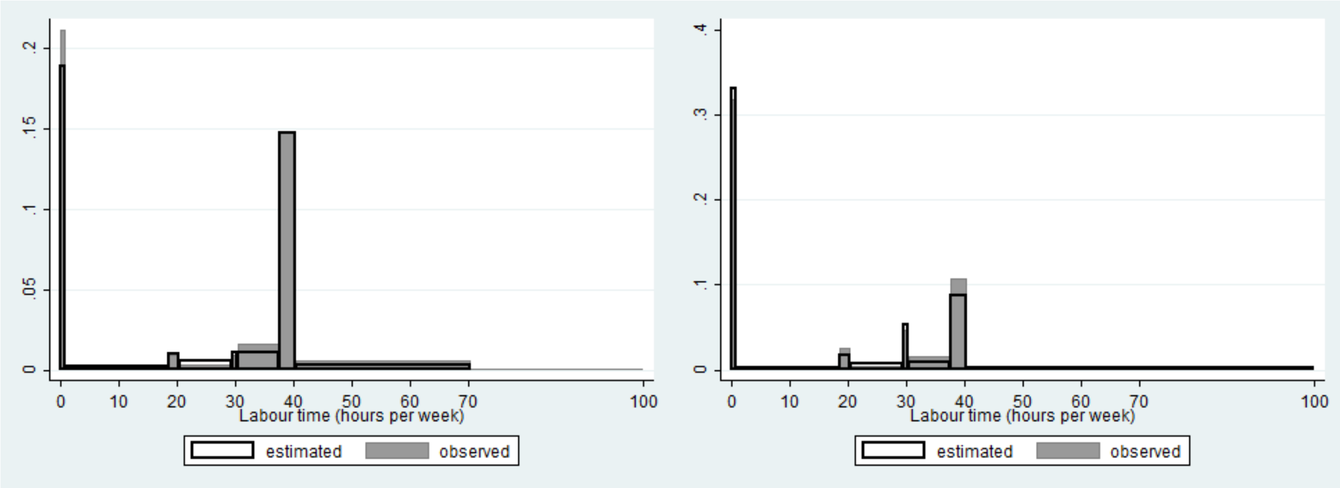


Figure 11: Densities wages single males (left) and females (right)



The wage distribution of single females (RHS of Figure 11) is better fitted than that of males (LHS of Figure 11). The simulated distribution for females is somewhat more skewed to the

Figure 12: Histograms labour time regimes single males (left) and single females (right)



right than the observed one. The simulated distribution of males has slightly fatter tails than the observed one.

Labour market participation of single males (LHS of Figure 12) is overestimated, while that of single females is somewhat underestimated (RHS of Figure 12). The observed occurrence of half, three quarter and full time jobs for males is well fitted. Occurrence of half time and full time jobs among single females is underestimated, that of three quarter jobs slightly overestimated.

Finally, the simulated and observed distribution of labour time regimes for single males (LHS of Figure 12) illustrate that these should be distinguished from the availability of job offers specifying certain labour time regimes, as was mentioned at the beginning of this section. Estimates of the relative availability of half time and three quarter time jobs as compared to jobs between three quarter time and full time were represented by the corresponding peaks and through in Figure 5. So, despite our estimates reveal that half time and three quarter time jobs are more intensely offered than jobs between three quarter time and full time, single males do not choose, neither according to the data, nor according to the simulated behaviour, the former type of jobs more often than the latter.

6.2 Elasticities

Table 6 reports the total reaction in terms of labour supply and the effect on participation to the labour market (extensive margin), following a shift of the density of the males', respectively females', wage offer distribution to the right by 10% (augmenting the estimated location parameter with $\ln 1.1$). Additionally, we report intensive margins (effect on labour supply conditional on participating in the base line, inclusive of the labour market leavers). The variable 'part in' gives the percentage of entrants into the labour market, while 'part out' represents the percentage of leavers.

Table 6: Aggregate wage elasticities of labour supply

	Shift of *female* wage distribution			Shift of *male* wage distribution		
	Couple		Single	Couple		Single
	Female	Male	Female	Female	Male	Male
Total elasticity	0.5034	−0.1054	0.4786	−0.2200	0.3104	0.2858
Intensive margin	0.1660	−0.1342	0.0363	−0.2559	0.1331	0.0280
Part in	2.883%	0.275%	2.977%	0.412%	1.647%	2.004%
Part out	0.000%	1.098%	0.000%	1.510%	0.000%	0.000%

Compared to Marshallian elasticities in the literature estimated by static models using micro data, the total elasticity estimates produced here are rather large (Compare *e.g.* with the figures reported in Tables 6 and 7 of Keane, 2011). However, as far as these total elasticities include the extensive margin, and are calculated as the proportional change in total labour time for the whole sub–sample, these need to be compared with macro elasticities, which are usually much larger, even as compared to the figures obtained here.[25] Still, it should be stressed that the figures reported here are conceptually of a different nature, in that actually obtained wages in the present model are the result of choosing the most attractive job offer according to the persons' preferences. Therefore, a reaction to an exogenous change in that wage cannot be conceived of in the framework we used. What was, alternatively, done, is to shift the entire distribution of the wages included in the job offers, to the right. This cannot be considered the same as a change in an exogenously given wage. As the RURO–model incorporates frictions due to restrictions in the labour market opportunities an agent faces, this might account for the lower values of the elasticities reported here, compared to values obtained for macro figures.

7 Age profiles of labour market participation: the contribution of preferences *versus* opportunities

In the present section we want to explore to what extent the lower participation figures of and the decrease in the number of hours worked by older persons in the dataset (*cf.* the introduction) can, according to the model, be ascribed to an increasing intensity of preference

[25] On the controversy about micro *versus* macro estimates, see amongst others Chetty *et al.* (2011), Chetty (2012), Fiorito and Zanella (2012), Keane and Rogerson (2012), Jäntti *et al.* (2015), and the references therein.

for leisure, or rather to the lower intensity of job offers suitable to the capacities of older people.

Thereto we performed two counterfactual simulation exercises, and compare them with the baseline simulation. In the first, henceforth labelled as 'EO', we have changed the intensity of job offers π_1 . Every individual in the sample got the maximal value for π_1 in function of age. Figure 6 showed that this corresponds to the age of 30. Hence, in this 'equality of opportunity' simulation we calculate labour market choices as if all individuals would get the same number of suitable job offers as someone of age 30. In the second simulation, we leave opportunities unaffected, but modify preferences (labelled 'EP'). Every individual now is endowed preferences exhibiting the lowest intensity of preferences for leisure according to age, that is yielding the flattest indifference curves. For single females this is 37 years, and for males in couples 36 years. For single males and females in couples, the indifference curve is uniformly steepening with age in the relevant age range of the sample. In these cases we used 21 years as the 'lowest' age.

Table 7 contains the results of these simulations. For both counterfactuals we calculated the mean participation rate (denoted by 'part' in the table) and mean labour supply (denoted 'hours' in the table) by age category. For hours we both calculated the unconditional mean, and the average hours conditional on participating. It seems as if 'equalising' differences in opportunities with respect to age, has in the first place an impact on the extensive margin, and much less so on the intensive margin of the number of hours worked. Overall, that is, the fours subgroups combined, the participation rate goes up with 6.9 percentage points (from 82.0% to 88.9%) in the 'EO'–counterfactual. This effect is substantial for all subgroups, ranging from +3.1 percentage points (ppt) for males in couples, to +8.9 ppt for females in couples, and +6.7 and even +11.6 ppt for respectively single males and females. The breakdown of this effect according to age groups reveals that the rise in participation rates increases with age for singles, while for persons in couples it mirrors the concave shape of π_1 with respect to age. Equalising opportunities has no perceptible effect on the average number of hours worked, once we condition on participation. For the whole sample the effect is even slightly negative (the average number of hours decreases from 36.4 to 36.0. This is the result of a small decrease for all subgroups, save for males in couples where we find a small increase (from 39.8 to 39.9 hours worked per week). The second counterfactual shows that a decrease in the intensity of preferences for leisure both has an impact on the participation rate and on the number of hours worked. The overall participation rate increases with 2.6 ppt from 82.0% in the baseline to 84.6% under this 'equalising preferences' scenario. This is a much smaller increase than under the 'EO'–scenario which yielded an increase of 6.9 ppt. Also in contrast with the previous 'EO'–counterfactual, we do find an effect on the intensive margin now. On average, the number of hours conditional on participation increases with 1.9 (from

Table 7: Participation and avg. labour time by age class in baseline and counterfactuals

Age	15<age<=30	30<age<=40	40<age<=50	50<age<=64	all
Category			Couples: males		
n obs.	250	537	416	254	1457
			participation		
part base	93.2%	94.2%	94.5%	81.9%	92.0%
part EO	94.4%	94.4%	97.1%	94.1%	95.1%
part EP	94.0%	94.6%	94.5%	81.1%	92.1%
		average hours of labour time per week – unconditional			
hours base	35.1	38.0	38.2	32.6	36.6
hours EO	35.4	38.1	39.8	36.8	37.9
hours EP	36.3	38.8	39.1	33.6	37.6
		average hours of labour time per week – conditional on part.			
hours base	37.7	40.3	40.5	39.8	39.8
hours EO	37.6	40.4	41.0	39.2	39.9
hours EP	38.6	41.0	41.4	41.4	40.8
Category			Couples: females		
n obs.	365	524	402	166	1457
			participation		
part base	82.5%	85.1%	77.6%	49.4%	78.3%
part EO	86.0%	86.8%	88.8%	86.7%	87.2%
part EP	84.1%	88.5%	84.8%	63.9%	83.6%
		average hours of labour time per week – unconditional			
hours base	27.6	26.0	25.0	16.1	25.0
hours EO	28.2	26.4	28.3	26.3	27.4
hours EP	30.1	30.2	31.6	23.5	29.8
		average hours of labour time per week – conditional on part.			
hours base	33.4	30.5	32.2	32.5	31.9
hours EO	32.8	30.4	31.9	30.4	31.4
hours EP	35.7	34.1	37.2	36.8	35.6

36.4 to 38.3), compared to the small decrease in the 'EO'–scenario. This effect is largest for females in couples (+3.7 hours). For males and females in couples and for single males the effect on hours is more uniform across age classes than in the 'EO'–counterfactual.

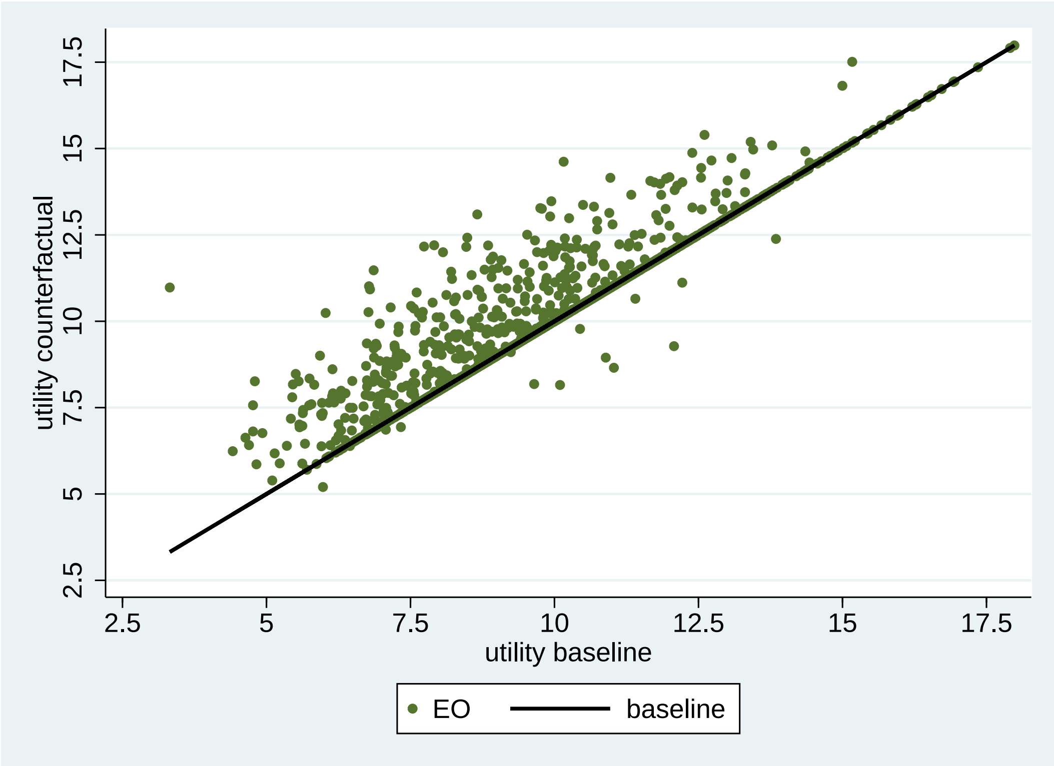
In the Figures 13–15, we report comparisons of utility obtained in the baseline with utility in the 'equalised opportunity' counterfactual. We do not reproduce a similar picture for

Table 7: Participation and avg. labour time by age class in baseline and counterfactuals *ctd.*

Age	15<age<=30	30<age<=40	40<age<=50	50<age<=64	all
Category			Singles: males		
n obs.	106	135	119	89	449
			participation		
part base	77.4%	91.1%	84.9%	65.2%	81.1%
part EO	78.3%	92.6%	90.8%	87.6%	87.8%
part EP	79.2%	92.6%	87.4%	71.9%	84.0%
			average hours of labour time per week – unconditional		
hours base	29.5	35.3	33.1	24.7	31.3
hours EO	30.3	35.7	35.3	32.8	33.7
hours EP	30.6	37.3	35.2	28.8	33.5
			average hours of labour time per week – conditional on part.		
hours base	38.1	38.7	39.0	37.9	38.6
hours EO	38.7	38.5	38.9	37.4	38.4
hours EP	38.7	40.3	40.2	40.1	39.9
Category			Singles: females		
n obs.	102	180	171	118	571
			participation		
part base	63.7%	70.0%	73.1%	55.9%	66.9%
part EO	66.7%	73.3%	82.5%	90.7%	78.5%
part EP	66.7%	70.6%	74.3%	59.3%	68.7%
			average hours of labour time per week – unconditional		
hours base	22.4	26.7	25.0	18.5	23.7
hours EO	23.6	27.9	28.4	30.0	27.7
hours EP	23.8	26.9	25.6	22.5	25.1
			average hours of labour time per week – conditional on part.		
hours base	35.1	38.1	34.2	33.2	35.5
hours EO	35.3	38.0	34.4	33.0	35.3
hours EP	35.7	38.2	34.4	38.0	36.5

the 'EP'–counterfactual, since preferences have been changed thereto. Consequently, utility comparisons between baseline results and counterfactual for the same individual have no clear meaning. This difficulty does not arise in 'equalised opportunity' case. One of our next steps will be to calculate welfare indices apt to a case in which preferences have been changed, and which we think can be more meaningfully compared. The black line in the figure is the

Figure 13: Utility couples: baseline and counterfactual



45°–line and reflects utility in the baseline. The dots are the utility levels obtained by the optimal choice in the 'EO'–counterfactual, *i.e.* when job offer intensity would have reached the same value as when a person was thirty years old. Dots above the black line imply a gain in utility, dots below the black line, a loss. Since for all individuals, the job offer intensity in the counterfactual is at least as large as in the baseline, and for most individuals it is larger, one would at first sight expect that no one would be worse off in the counterfactual compared to the baseline. Yet, for all three sub–samples, we do find some cases in which the final choice is less preferred when confronted with an environment in which the opportunity set has been changed in the direction of more job offers. Hence, a bit surprisingly maybe, increasing the intensity of job offers does not imply a Pareto improvement. The explanation for having dots below the black line, is that increasing the job opportunities for persons with a low preference for jobs (elderly), might be harmful to these people. Indeed, since in this model, the job offer intensity is expressed as a relative number of market *versus* non–market opportunities, a higher job offer intensity, is tantamount to loosing some non–market opportunities. The latter might have been especially valuable for those with relatively more intense preference for leisure. It might be considered an unattractive property of the model that increasing the intensity of job offers is, by definition, mirrored by lowering the degree of availability of non–market opportunities. Indeed, why would I loose the opportunities to do what I liked most, when more jobs were offered to me? The worst that could happen is that none of these additional opportunities is more attractive than what I am currently doing, keeping the level of utility constant. Still, in this static model, the total stock of capacities an individual is endowed with, is considered as fixed. Increasing job offer intensity then means that capacities which were originally only apt for performing non–market activities, now become valuable on the market. In such a case, increasing job offer intensity is not a bless for all.

Figure 14: Utility single females: baseline and counterfactual

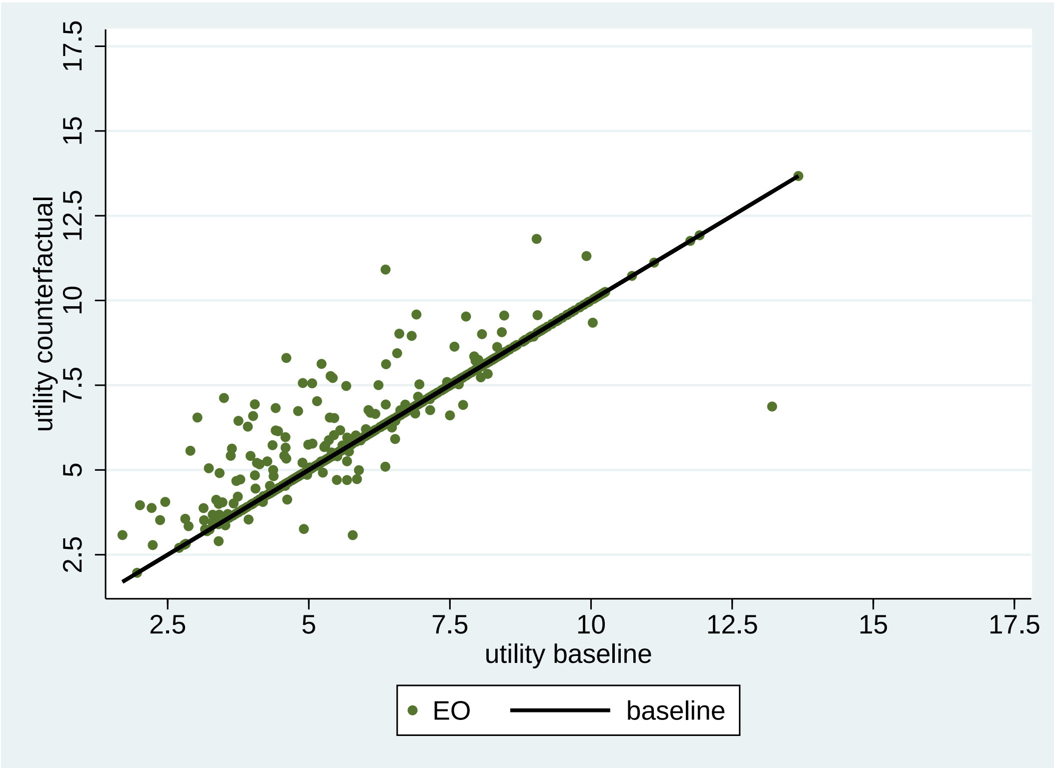
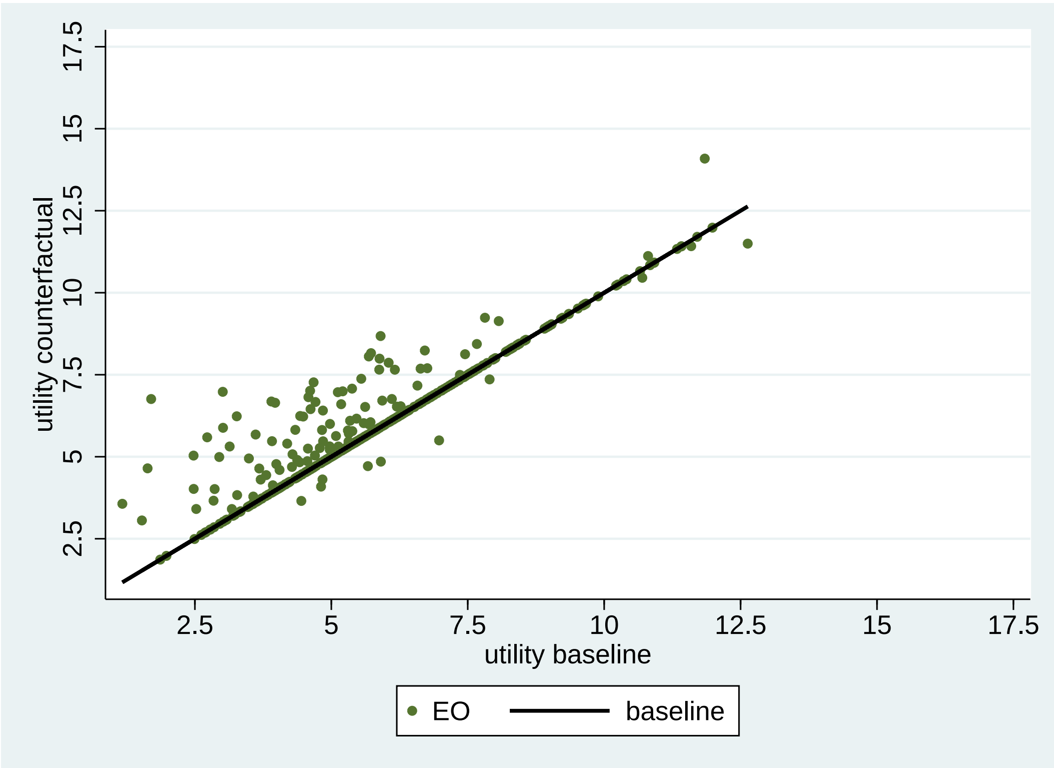


Figure 15: Utility single males: baseline and counterfactual



8 Conclusion

This paper has tried to shed some light on the lower labour market participation of the elderly by exploiting a rich structural model of labour supply, the so–called RURO–model. Contrary to standard random utility models of labour supply, this model adds substantially more heterogeneity in individual opportunities by integrating an individual specific job offer intensity as an explanatory variable for observed behaviour. We have estimated this model on Belgian EU–SILC data of 2007, which allowed us to quantify to what extent lower labour market participation is due to changing preferences (executing a job might become less enjoyable with age) or differences in opportunities (elderly getting less, or less attractive job offers).

The estimates indicate that the relation between job offer intensity and age has an inverted U–shape with top at about 30 years. Job offer intensity after that age decreases sharply, so that the avaialbility of suitable jobs after the age of 50 is lower than that for youngsters. This effect is a bit more outspoken for females than for males. We also found important regional variation in the job offer intensity, being lower in Wallonia, and especially in the Brussels capital region, as compared to Flanders. As for the distribution of wage offers, we found, unsurprisingly, that, on average, higher wages were offered to persons with higher education level. The differences between males and females, however, were small.

The effect of age on preferences was less clear–cut than the effect on opportunities. We found two cases where the intensity of preference for leisure, or the distaste for paid work, monotonically increases with age in the relevant age range (i.e. 16 to 64 years for our data): for females in couples, and for single males. For both these groups, indifference curves in the labour time versus consumption space, become steeper with age. For the other two subgroups, *i.e.* single females and males in couples, the pattern is non–monotonic. For single females and males in couples, preference intensity for leisure is lowest at the age of respectively 37 and 36 years. Their indifference curves become flatter from the age of 16 until that of 37 (respectively 36) years, and their slope starts increasing when becoming older than 37 (respectively 36) years.

To get an idea of the relative importance of both forms of heterogeneity (impact of age *via* preferences and opportunities), we conducted a simulation of two counterfactuals. First, we removed part of heterogeneity in opportunities by giving all individuals the maximal job offer intensity in terms of age (that of a 30 years old person). Second, we removed part of the heterogeneity in preferences by endowing all individuals with preferences at the subgroup specific age at which the intensity of preference for leisure was at its minimum. Our tentative conclusion from a comparison of labour market behaviour in these two counterfactuals with the baseline situation is that opportunities which decline with age are at least as an important factor in explaining low participation rates for the elderly, as is the fact that preference change with growing older. More specifically, the effect of opportunities seems to work primarily through the extensive margin, whereas the effect of preferences is more outspoken in the intensive than in the extensive margin. From a policy point of view this might be relevant. We feel that the rich specification of the RURO–model is promising. Therefore, we think we should try to estimate with data which allow for a better identification of the important π_1–function. This might *e.g.* be obtained by re–estimating the model with more exogenous variation in job availability, either by *e.g.* using cross-country variation or variation through time. Needless to say that the corresponding work to produce choice sets for the estimation in this framework might be labour intensive (one needs a cross–country microsimulation model and/or a microsimulation model which can model the gross–net trajectory for several years). Secondly, structural models like the one presented in this paper are obvious candidates as suppliers of essential information for normative analyses, based on — in this case, revealed — preferences. After careful scrutiny of how sensitive estimated preferences are to the inclusion and specific specification of additional constraints from the demand side of the labour market, we plan to construct welfare measures in line with our first attempt in Decoster and Haan (2015). This is especially important when trying to make welfare comparisons between countries or at different moments in time, when the respective populations to be compared do not necessarily have the same preferences.

References

[1] Aaberge R. & U. Colombino (2014), "Labour supply models," in: C. O'Donoghue (ed.) *Handbook of Microsimulation Modelling*. Contributions to economic Analysis vol. 293. Emerald Group Publishing Ltd.: Bingley, 167–221.

[2] Aaberge R., U. Colombino, & S. Strøm (1999), "Labour supply in Italy: an empirical analysis of joint decisions, with taxes and quantity constraints," *Journal of Applied Econometrics*, vol. 14(4), 403–422.

[3] Aaberge R., U. Colombino, & T. Wennemo (2009), "Evaluating alternative representations of the choice sets in models of labour supply," *Journal of Economic Surveys*, vol. 23(3), 586–612.

[4] Aaberge R., J.K. Dagsvik & S. Strøm (1995), "Labor supply responses and welfare effects of tax reforms," *Scandinavian Journal of Economics*, vol. 97(4), 635–659.

[5] Altonji J.G. & C.H. Paxson (1982), "Labor supply preferences, hours constraints, and hours–wage trade–offs," *Journal of Labor Economics*, vol. 6(2), 254–276.

[6] Altonji J.G. & C.H. Paxson (1992), "Labor supply, hours constraints, and job mobility," *Journal of Human Resources*, vol. 27(2), 256–278.

[7] Beffy M., R. Blundell, A. Bozio, G. Laroque & M. To (2014) "Labour supply and taxation with restricted choices," IFS WP W15/02.

[8] Belloni M. (2008) "The option value model in retirement literature: the trade–off between computational complexity and predictive validity," ENEPRI Research Report No. 50, AIM–WP6.

[9] Belloni M. & R. Alessie (2013) "Retirement choices in Italy: What an option value model tells us," *Oxford Bulletin of Economics and Statistics*, vol. 75(4), 499–527.

[10] Belloni M. & R. Alessie (2009) "The importance of financial incentives on retirement choices: new evidence for Italy," *Labour Economics*, vol. 16(5), 578–588.

[11] Ben–Akiva M. & S.R. Lerman (1985), *Discrete Choice Analysis*, MIT Press: Cambridge, MA, 261–269 (Ch. 9.3).

[12] Berkovec J. & S. Stern (1991) "Job exit behavior of older men," *Econometrica*, vol. 59(1), 189–210.

[13] Bloemen H.G. (2000), "A model of labour supply with job offer restrictions," *Labour Economics*, vol. 7(3), 297–312.

[14] Bloemen H.G. (2008), "Job search, hours restrictions, and desired hours of work," *Journal of Labor Economics*, vol. 26(1), 137–179.

[15] Chetty R. (2012) "Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply," *Econometrica*, vol. 80(3), 969–1018.

[16] Chetty R., A. Guren, D. Manoli & A. Weber (2011) "Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins," *American Economic Review*, vol. 101(3), 471–475.

[17] Dagsvik J.K. (1994) "Discrete and continuous choice, max–stable processes, and independence from irrelevant attributes," *Econometrica*, vol. 62(5), 1179–1205.

[18] Dagsvik K.J. (2013) "Justification of functional form assumptions in structural models: a correction," *Theory and Decision* 75(2), 79–83.

[19] Dagsvik K.J. & S. Røine Hoff (2011), "Justification of functional form assumptions in structural models: applications and testing of qualitative measurement axioms," *Theory and Decision* 70(2), 215–254.

[20] Dagsvik J.K. & Z. Jia (2014) "Labor supply as a choice among latent jobs: unobserved heterogeneity and identification," *Journal of Applied Econometrics*, first online 6 Jan. 2015, `DOI: 10.1002/jae.2428`.

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

[22] Dagsvik J.K., M. Locatelli & S. Strøm (2006) "Simulating labor supply behaviour when workers have preferences over job opportunities and face non–linear budget constraints," Centre for Household, Income, Lambour and Demographic Economics (CHILD, Torino, Italy)), Discussion paper no. 01/2006.

[23] Dagsvik J.K., M. Locatelli & S. Strøm (2007) "Evaluation of tax reforms when workers have preferences over job attributes and face latent choice restrictions," Centre for Household, Income, Labour and Demographic Economics (CHILD, Torino, Italy)), Discussion paper no. 13/2007.

[24] Dagsvik J.K. & S. Strøm (1992) "Labour supply with non–convex budget sets, hours restriction and non–pecuniary job attributes," Central Bureau of Statistics Norway, Discussion Paper No 76.

[25] Dagsvik J.K. & S. Strøm (2003) "Analyzing labor supply behavior with latent job opportunity sets and institutional choice constraints," Statistics Norway, Research Department, Discussion Paper No 344.

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

[27] Decoster A. & P. Haan (2015) "Empirical welfare analysis with preference heterogeneity," *International Tax and Public Finance*, vol.22(2), 224–251.

[28] Dickens W.T. & S.J. Lundberg (1993) "Hours restrictions and labor supply," *International Economic Review*, vol. 34(1), 169–192.

[29] Feldstein M. (1974) "Social security, induced retirement, and aggregate capital accumulation," *The Journal of Political Economy*, vol. 82(5), 905–926.

[30] Fiorito R. & G. Zanella (2012) "The anatomy of the aggregate labor supply elasticities," *Review of Economic Dynamics*, vol. 15(2), 171–187.

[31] Gruber J. & D.A. Wise (eds.) (1999), *Social Security and Retirement around the World*, NBER Book Series – International Social Security, University of Chicago Press: Chicago.

[32] Ham J.C. & K.T. Reilly (2002) "Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro data," *American Economic Review*, vol. 92(4), 905–927.

[33] Hausman J.A. (1985) "Taxes and labor supply," in: A.J. Auerbach & M. Feldstein (eds.), *Handbook of Public Economics*, Elsevier Science Publishers B.V. (North-Holland): Amsterdam, 213–263.

[34] Jäntti M., J. Pirttilä & H. Selin (2015) "Estimating labour supply elasticities based on cross–country micro data: A bridge between micro and macro estimates?" *Journal of Public Economics*, vol.127, 87–99.

[35] Keane M.P. (2011) "Labor supply and taxes: a survey," *Journal of Economic Literature*, vol.1 49(4), 961–1075.

[36] Keane M. & R. Rogerson (2012) "Micro and macro labor supply elasticities: a reassessment of conventional wisdom," *Journal of Economic Literature*, vol. 50(2), 464–476.

[37] Lefebvre M. & K. Orsini (2012) "A structural model for early exit of older men in Belgium," *Empirical Economics*, vol. 43(1), 379–398.

[38] Lemp J.D. & K.M.Kockelman (2012), "Strategic sampling for large choice sets in estimation and application," Transportation Research Part A 46(3), 602–613.

[39] Löffler M., A. Peichl and S. Siegloch (2014) "Structural labor supply models and wage exogeneity," IZA DP No.8281 Forschungsinstitut zur Zukunft der Arbeit: Bonn.

[40] Luce D.R. (1959, repr. 2005) *Individual Choice Behavior: A Theoretical Analysis*, Dover Publications Inc.: Mineola, New York.

[41] Lumsdaine R.L., J.H. Stock & A.D. Wise (1992) "Three models of retirement: computational complexity versus predictive validity," in: D.A. Wise (ed.) *Topics in the Economics of Aging*, University of Chicago Press: Chicago, 21–60.

[42] Maes M. (2011) "Will the dismantlement of early retirement schemes increase older unemployment? A competing–risk analysis for Belgium," *Labour*, vol. 25(2), 252–267.

[43] Maes M. (2012) "Financial and distributional implications of early retirement in Belgium," *Reflets et perspectives de la vie économique*, vol. 51(3), 29–41.

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

[45] McFadden D. (1978), "Modelling the choice of residential location," in A. Karlqvist, L. Lundqvist, F. Snickars & J. Weibull (eds.), *Spatial Interaction Theory and Planning Models*, North-Holland: Amsterdam, 75–96.

[46] Moffitt R. (1984) "The estimation of a joint wage–hours labor supply model," *Journal of Labor Economics*, Vol. 2(4), 550–566.

[47] Rust J. (1989) "A dynamic programming model of retirement behavior," in: D.A. Wise (ed.) *The Economics of Aging*, University of Chicago Press: Chicago, 359–404.

[48] Rust J. & C. Phelan (1997) "How social security and medicare affect retirement behavior in a world of incomplete markets," *Econometrica*, vol. 65(4), 781–831.

[49] Stock: J.H. & D.A. Wise (1990) "Pensions, the option value of work, and retirement," *Econometrica*, vol. 58(5), 1151–1180.

[50] Tummers M.P. & I. Woittiez (1991) "A simultaneous wage and labor supply model with hours restrictions," *Journal of Human Resources*, vol. 26 (3), 393–423.

[51] Train K.E. (2009, 2nd ed.), *Discrete Choice Methods with Simulation*, Cambridge University Press: Cambridge, pp.64–66.

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

[53] Van Soest A., M. Das & X. Gong (2002) "A structural labour supply model with flexible preferences," *Journal of Econometrics*, vol. 107(1–2), 345–374.

[54] Van Soest A., I. Woittiez & A. Kapteyn (1990) "Labor supply, income taxes, and hours restrictions in the Netherlands," *Journal of Human Resources*, vol. 25(3), Special Issue on Taxation and Labor Supply in Industrial Countries, 517–558.

Appendices

Appendix I Poisson processes

Originally, a Poisson process is a stochastic process describing the probability of the number of occurrences of a particular event during a certain time spell. More specifically, a Poisson process assumes that distribution of the time between each pair of consecutive events is independent from the moment at which the first of these two events occurred, or from any other event in the past, and that these inter–arrival times are exponentially distribution with parameter λ. This parameter λ measures the *intensity* with which such events occur. Under these assumptions, the probability that a certain event occurs n times within a given unit of time, where $n \in \{0, 1, 2 \ldots\}$, equals:

$$P\left(N\left(t+1\right)-N\left(t\right)=n\right)=\frac{\lambda^{n}\exp\left[-\lambda\right]}{n!}, \tag{A.1}$$

where $N(t)$ is the number of events that occurred in total after t units of time. A Poisson process is *inhomogeneous* if the intensity parameter depends on the moment of measurement, $\lambda(t)$ say. In that case, the probability that n events occur within a time interval $[t, t+\tau]$, equals:

$$P\left(N\left(t+\tau\right)-N\left(t\right)=n\right)=\frac{\left(\Lambda\left(\tau\right)\right)^{n}\exp\left[-\Lambda\left(\tau\right)\right]}{n!}, \tag{A.2}$$

where $\Lambda(\tau) := \int_t^{t+\tau} \lambda(s)\ \mathrm{d}\, s$, which is, because of the time independence property of Poisson processes, independent of t.

A Poisson process can also be *spatial*. Let an event be described as a point in an m–dimensional space. A spatial Poisson process determines the probability that n events occur within a subset of the m–dimensional space. Let, for example, $\mathcal{B} \subset \mathbb{R}^m$, and let $N(\mathcal{B})$ be the number of events occurring in $\mathcal{B}$. If the occurrence of such events obeys a Poisson process, then the probability that there occur n events in $\mathcal{B}$, equals:

$$P\left(N\left(\mathcal{B}\right)=n\right)=\frac{\lambda^{n}\exp\left[-\lambda\right]}{n!}. \tag{A.3}$$

Again, such a process is said to be inhomogeneous if the intensity of occurrence depends on the points $x \in \mathbb{R}^m$. To describe that process, assume that there exists a measure ρ defined on (measurable) subsets of the space $\mathbb{R}^m$, and that $\rho(\mathcal{B}) = 1$. The probability that there occur n events in the subset $\mathcal{B}$, is then:

$$P\left(N\left(\mathcal{B}\right)=n\right)=\frac{\left(\Lambda\left(\mathcal{B}\right)\right)^{n}\exp\left[-\Lambda\left(\mathcal{B}\right)\right]}{n!},. \tag{A.4}$$

where $\Lambda(\mathcal{B}) := \int_{x\in\mathcal{B}} \lambda(x)\ \mathrm{d}\,\rho(x)$.

Job offers and the availability of non–market activities each are described by an inhomogeneous spatial Poison process in ruro models. These processes are independent. However, given that the ruro model is static, the stock of capacities an individual is endowed with, is assumed to be fixed. If demand for these capacities (by means of job offers) intensifies, a relatively smaller amount of these capacities serves exclusively for executing non–market activities.

Appendix II Sampling choice sets

The McFadden approach.

The method of sampling choice sets for estimating discrete choice models was originally developed by McFadden (1978, section 7) in order to handle cases where the number of choices is so large that the true likelihood function would become intractable. A summary of this procedure can be found in Train (2009, pp. 64–66).

Assume that the true choice set, $\mathcal{C}$, consists of a very large, but still discrete, number of alternatives, indexed by $j, k \in \mathcal{C} := \{1, 2, \ldots, C\}$, where C is a natural number. Assume also that the choice behaviour can be reflected by a multinomial LOGIT model. The systematic part of the utility of an alternative $j \in \mathcal{C}$ is denoted by V_j. That is, V_j is a shorthand for a function of a set of covariates whose values change across alternatives, say $\mathbf{x}_j$, and a set of parameters to be estimated, $\boldsymbol{\beta}$ say. Total utility derived from $j \in \mathcal{C}$ is:

$$U_j := V_j + \epsilon_j, \tag{A.5}$$

where ϵ_j is a to the researcher unobserved factor determining that person's preferences. Assuming that this term is drawn from an Extreme Value Type I distribution, which has distribution function $\exp\left[-\exp\left(-\epsilon_j\right)\right]$, the probability that this person will opt for alternative j from the set of available alternatives $\mathcal{C}$, $P_{j,\mathcal{C}}$ say, is equal to:

$$P_{j,\mathcal{C}} = \frac{\exp\left(V_j\right)}{\sum_{k \in \mathcal{C}} \exp\left(V_k\right)}. \tag{A.6}$$

For some reason, it might be impossible to collect information on all the alternatives in $\mathcal{C}$ that are available to a particular individual (*e.g.* because this set is too large). Therefore the researcher might sample a set of alternatives $\mathcal{D}$ from $\mathcal{C}$. The observed alternative is included in $\mathcal{D}$, since it is observed, and thus, by definition, must have been one of the possible alternatives from which the person has chosen. One might for example partition the set $\mathcal{C}$ into M subsets (with M a natural number such that $M < C$). Each member of the partition is denoted by $\mathcal{K}_m$ ($m \in \mathcal{M} := \{1, 2, \ldots, M\}$). One can then sample the chosen alternative j with certainty from the subset that contains the chosen alternative, $\mathcal{K}_m : j \in \mathcal{K}_m$, and sample at random an alternative from the remaining subsets, $\mathcal{K}_m : j \notin \mathcal{K}_m$.

Sampling a choice set induces a probability to select a subset $\mathcal{D}$ from the true choice set $\mathcal{C}$, given that the person is observed to have selected j. This probability will be denoted by $\pi\left(\mathcal{D} \,|\, j, \mathcal{C}\right)$. For example, when using the sampling procedure discussed in the previous paragraph, and letting K_m be the number of elements in the subset $\mathcal{K}_m$ of the partition (for each $m \in \mathcal{M}$), then $\pi\left(\mathcal{D} \,|\, j, \mathcal{C}\right) = \left(K_n / \prod_{m=1}^{M} K_m\right)$, where n is the index of the subset to which the chosen alternative j belongs, $n := m \in \mathcal{M} : j \in \mathcal{K}_m$.

By sampling a choice set $\mathcal{D}$, the researcher can only retrieve some information on factors leading a person to choose alternative j from $\mathcal{D}$ rather than from $\mathcal{C}$, while she is actually interested in the latter. Denote the probability to choose j from $\mathcal{D}$ by $P_{j,\mathcal{D}}$. Train (2009) shows that:

$$P_{j,\mathcal{D}} = \frac{\exp(V_j)\,\pi(\mathcal{D}\,|j,\mathcal{C}\,)}{\sum_{k\in\mathcal{D}}\exp(V_k)\,\pi(\mathcal{D}\,|k,\mathcal{C}\,)}. \tag{A.7}$$

The essence of the proof lays in the observation that the probability to sample a set $\mathcal{D}$ from $\mathcal{C}$, and observing an agent to choose j, say $\theta_{\mathcal{C}}(\mathcal{D},j)$, is equal to $\pi(\mathcal{D}\,|j,\mathcal{C}\,)\cdot P_{j,\mathcal{C}}$. Similarly, by reversing the conditioning, this probability to sample a set $\mathcal{D}$ from $\mathcal{C}$, and observing an agent to choose j, $\theta_{\mathcal{C}}(\mathcal{D},j)$, is also equal to $P_{j,\mathcal{D}}\cdot\Pi(\mathcal{D})$, where $\Pi(\mathcal{D})$ is the unconditional probability to select a subset $\mathcal{D}$ from $\mathcal{C}$ according to the sampling procedure, that is $\Pi(\mathcal{D}) := \sum_{k\in\mathcal{C}}\pi(\mathcal{D}\,|k,\mathcal{C}\,)\cdot P_{k,\mathcal{C}} = \sum_{k\in\mathcal{D}}\pi(\mathcal{D}\,|k,\mathcal{C}\,)\cdot P_{k,\mathcal{C}}$. The equality follows from observing that the conditional probability to select a set $\mathcal{D}$, given the observed choice is k, while k does not belong to the sampled set $\mathcal{D}$, is zero, by definition of the allowed sampling procedures: $\pi(\mathcal{D}\,|k,\mathcal{C}\,) = 0$ if $k\notin\mathcal{D}$. Equating both expressions for $\theta_{\mathcal{C}}(\mathcal{D},j)$ and solving for $P_{j,\mathcal{D}}$ gives equation (A.7).
McFadden (1978) has shown that the parameters $\boldsymbol{\beta}$ of the utility function, $V_j := v(\boldsymbol{\beta};\mathbf{x}_j)$, can be estimated consistently by maximising the sampled log–likelihood function $\mathcal{L}(\boldsymbol{\beta};\mathbf{X})$ based on the corrected probabilities (A.7), where $\mathbf{X}$ is a data set containing for each observation t values of choice k's attributes, $\mathbf{x}_{t,k}$ say, for all $k\in\mathcal{C}$. Let $j(t)$ denote observation t's observed choice from the set $\mathcal{C}$. Then, this sampled log–likelihood function becomes:

$$\mathcal{L}(\boldsymbol{\beta};\mathbf{X}) := -\sum_t \ln\left(\sum_{k\in\mathcal{D}_t}\left[\exp\left[v(\boldsymbol{\beta};\mathbf{x}_{t,k}) - v\left(\boldsymbol{\beta};\mathbf{x}_{t,j(t)}\right)\right]\frac{\pi(\mathcal{D}_t\,|k,\mathcal{C}\,)}{\pi(\mathcal{D}_t\,|j(t),\mathcal{C}\,)}\right]\right), \tag{A.8}$$

McFadden's result relies upon imposing the *positive conditioning property* upon the choice set sampling probabilities $\pi(\mathcal{D}\,|k,\mathcal{C}\,)$, which reads as:

$$\text{for all } j,k\in\mathcal{D}:\ \text{if } \pi(\mathcal{D}\,|j,\mathcal{C}\,) > 0 \quad \text{then } \pi(\mathcal{D}\,|k,\mathcal{C}\,) > 0. \tag{A.9}$$

A stronger condition which is also sufficient for consistency (and which implies a further simplification for the sampled likelihood function) is the *uniform conditioning property*:

$$\forall\, j,k\in\mathcal{D} : \pi(\mathcal{D}\,|j,\mathcal{C}\,) = \pi(\mathcal{D}\,|k,\mathcal{C}\,). \tag{A.10}$$

In that case the correction factor in the sampled log–likelihood function (A.8), $\pi(\mathcal{D}_t\,|k,\mathcal{C}\,)/\pi(\mathcal{D}_t\,|j(t),\mathcal{C}\,)$, drops out. One could wonder why then the uniform conditioning property has not been unanimously followed, for reasons of simplicity. Note however that this would result in sampling alternatives that for some, or all, agents are relatively unlikely to be chosen. Sampling such alternatives would yield poor information on the factors explaining a person's actual choice. This method is therefore generally less efficient (yielding larger standard errors for the estimates of $\boldsymbol{\beta}$) than sampling alternatives on the basis of some prior knowledge or evidence on the alternatives having been chosen, or likely to be chosen. This is the idea behind importance sampling, which will be discussed later.

The Aaberge–Colombino–Wennemo approach.

Aaberge, Colombino and Wennemo (2009) propose a similar sampling procedure for estimating the RURO model. *In abstracto*, the individual contributions to the likelihood function of the RURO model equal (compare with equation (18) of the main text):

$$Q_{j,\mathcal{C};f} = \frac{\exp\left(V_j\right) f\left(j\right)}{\int_{k \in \mathcal{C}} \exp\left(V_k\right) f\left(k\right) \, \mathrm{d}\, \rho\left(k\right)}, \tag{A.11}$$

where $f(k)$ is a measure for the intensity with which alternative k is rendered available to the individual decision maker, and j is the chosen alternative. The set of alternatives $\mathcal{C}$ might be continuous, and $\rho(\cdot)$ is a measure defined over the space of alternatives. So, f is a shorthand for a function of a set of explanatory variables whose value may be varying across alternatives, say $\mathbf{x}_{f,j}$ ($j \in \mathcal{C}$), and a set of parameters to be estimated, $\boldsymbol{\delta}$. In practice, it might be impossible to observe or use the whole set of alternatives $\mathcal{C}$ in the estimation, and a sample $\mathcal{D}$ is drawn. Assume the probability to sample an alternative k from $\mathcal{C}$ to be equal to $\phi(k)$. Then, the correction proposed by Aaberge, Colombino, and Wennemo (2009, see equation (9), p.593) is equal to:

$$\widehat{Q}_{j,\mathcal{D};f,\phi} = \frac{\exp\left(V_j\right) f\left(j\right) / \phi\left(j\right)}{\sum_{k \in \mathcal{D}} \exp\left(V_k\right) f\left(k\right) / \phi\left(k\right)}. \tag{A.12}$$

The intuitive rationale is that estimates of $\exp(V_k) f(k)$ will be affected by the sampling procedure. More specifically, the term referring to an alternative k in a person's simulated contribution to the likelihood function, that would be sampled relatively more often than the intensity with which it is really rendered available to that person, will get too big a weight. Therefore, its true value, $\exp(V_k) f(k)$, will be underestimated. Dividing true by its sampling weight, $\phi(k)$, would correct for that.

Relating both approaches.

The connection between both approaches seems rather vague at first sight. The formal connection between both, the McFadden (1978) approach and the Aaberge–Colombino–Wennemo (2009) approach, is explained in Ben–Akiva and Lerman (1985).
Recall that the probability to choose an alternative j from the sampled choice set $\mathcal{D}$, $P_{j,\mathcal{D}}$, can be written as:

$$P_{j,\mathcal{D}} = \frac{\theta_{\mathcal{C}}\left(\mathcal{D}, j\right)}{\Pi\left(\mathcal{D}\right)}. \tag{A.13}$$

Now, in the model of Aaberge, Colombino and Strøm (1999), $\theta_{\mathcal{C}}(\mathcal{D}, j)$ is equal to $\pi(\mathcal{D} \,|j, \mathcal{C}) \, Q_{j,\mathcal{C};f}$ (see equation A.11). The unconditional probability to select a subset $\mathcal{D}$ from $\mathcal{C}$ according to the sampling procedure, $\Pi(\mathcal{D})$, equals $\sum_{k \in \mathcal{C}} \pi(\mathcal{D} \,|k, \mathcal{C}) \cdot Q_{k,\mathcal{C};f} = \sum_{k \in \mathcal{D}} \pi(\mathcal{D} \,|k, \mathcal{C}) \cdot Q_{k,\mathcal{C};f}$. Using equation (A.11), this results in:

$$P_{j,\mathcal{D}} = \frac{\theta_{\mathcal{C}}\left(\mathcal{D}, j\right)}{\Pi\left(\mathcal{D}\right)} = \frac{\exp\left(V_j\right) f\left(j\right) \pi\left(\mathcal{D} \,|j, \mathcal{C}\right)}{\sum_{k \in \mathcal{D}} \exp\left(V_k\right) f\left(k\right) \pi\left(\mathcal{D} \,|k, \mathcal{C}\right)}. \tag{A.14}$$

One class of sampling choice sets is *importance sampling.* Each alternative j from the choice set $\mathcal{C}$ gets a (prior) probability weight $\phi(j)$ to be sampled from $\mathcal{C}$. So, the probability to sample the set $\mathcal{D}$ from the set of all alternatives $\mathcal{C}$, $\Pi(\mathcal{D})$, is equal to $\prod_{j\in\mathcal{D}} \phi(j) \prod_{j\notin\mathcal{D}} (1-\phi(j))$.[26] One always includes the chosen alternative according to the data, as an element in the sampled choice set. This implies that the probability to sample $\mathcal{D}$, given the chosen alternative is k, denoted earlier as $\pi(\mathcal{D}\,|k,\mathcal{C})$, equals $\prod_{j\in\mathcal{D}} \phi(j) \prod_{j\notin\mathcal{D}} (1-\phi(j))/\phi(k) = \prod_{\substack{j\in\mathcal{D}:\\ j\neq k}} \phi(j) \prod_{j\notin\mathcal{D}} (1-\phi(j)) = \Pi(\mathcal{D})/\phi(k)$.

So, equation (A.14) reduces to:

$$P_{j,\mathcal{D}} = \frac{\exp(V_j)\, f(j)\, \Pi(\mathcal{D})/\phi(j)}{\sum_{k\in\mathcal{D}} \exp(V_k)\ f(k)\, \Pi(\mathcal{D})/\phi(k)} = \frac{\exp(V_j)\, f(j)/\phi(j)}{\sum_{k\in\mathcal{D}} \exp(V_k)\ f(k)/\phi(k)}, \tag{A.15}$$

which is exactly equation (A.12) we were looking for.

Importance sampling for the RURO *model.*

There are several methods of importance sampling. Ben–Akiva and Lerman (1985, 265–267) mention for example three variants. They all however result in the same expression for the corrected likelihood (A.12).

In the RURO model the set of possible alternatives consists of the union of, on the one hand, the so–called non–market alternatives (that is possible sets of activities when not accepting any possible job offer), and, on the other hand, the set of possible packages of a wage, w, and a number of hours (per week) to be worked, say h, both as specified in a particular job offer. A specific wage–labour time regime is thus denoted by (w, h). As far as it concerns a job offer, it is assumed that with $w > 0$ and $H_{\max} > h \geq H_{\min}$, with $H_{\min} > 0$. Then, non–market alternatives are denoted as wage–labour time regime packages (w, h), such that $(w, h) \equiv (0, 0)$. So, we have a mixed distribution of the discrete variable $(0, 0)$ that indicates the option not to accept any job offer, and the continuous variable $(w, h) \in \mathbb{R}_{++} \times [\,H_{\min}, H_{\max}\,[$.[27]

Furthermore, in the RURO model, the probability (density) $f(k)$ specified in the previous section, is a measure for the intensity with which an alternative k characterised by a wage–labour time bundle (w_k, h_k), is rendered available to the agent. In the sequel, the log normal density with location parameter μ_1, and scale parameter σ of a random variable z, is abbreviated by $n(z; \mu_1, \sigma)$.[28] That is $n(z; \mu_1, \sigma) := \frac{1}{z\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln z-\mu_1)^2}{2\sigma^2}\right)$ if $z > 0$, and it equals zero otherwise. Then, the

[26] See Ben–Akiva and Lerman (1985, p.265) on independent importance sampling.

[27] But the realisation of the stochastic process that describes the intensity with which suitable job offers are rendered available to an individual, results in a discrete set of job offers, each specifying combination of a wage offer and a labour time regime, $(w, h) \in \mathbb{R}_{++} \times [\,H_{\min}, H_{\max}\,[$.

[28] That is the concrete specification we assumed for the wage offer distribution function g_1, as notified in Section 3.2.

probability (density) $f(k)$ will be further specified as:

$$f(k)=\begin{cases}\pi_0(\boldsymbol{\eta}_q;\mathbf{x}_q) & \text{if } (w_k,h_k)=(0,0)\,,\\ (1-\pi_0(\boldsymbol{\eta}_q;\mathbf{x}_q))\,n(w_k;\boldsymbol{\delta}'_{g_1}\mathbf{x}_{g_1},\sigma)\,\gamma_1\exp(\gamma_k) & \\ \qquad\qquad \text{if } (w_k,h_k)\in\mathbb{R}_{++}\times[\,\underline{H}_{k-1},\overline{H}_{k-1}[, \text{ for } k=2,\ldots,K+1, & \\ (1-\pi_0(\boldsymbol{\eta}_q;\mathbf{x}_q))\,n(w_k;\boldsymbol{\delta}'_{g_1}\mathbf{x}_{g_1},\sigma)\,\gamma_1 & \text{if } (w_k,h_k)\in\mathbb{R}_{++}\times\bigcup_{k=1,\ldots,K+1}[\overline{H}_{k-1},\underline{H}_k\,[,\\ 0 & \text{else,}\end{cases}\tag{A.16}$$

where $\overline{H}_0 := H_{\min} < \underline{H}_1 < \overline{H}_1 < \ldots < \underline{H}_k < \overline{H}_k < \ldots < \underline{H}_K < \overline{H}_K < \underline{H}_{K+1} := H_{\max}$, define the bins of a piecewise uniform distribution for labour time regimes.[29]The intensity with which suitable job offers are rendered available to an individual with characteristics $\mathbf{x}_q$, is specified by $(1-\pi_0(\boldsymbol{\eta}_q;\mathbf{x}_q)) := (1+\exp(-\boldsymbol{\eta}'_q\mathbf{x}_q))^{-1}$.[30] The parameters $(\boldsymbol{\delta}_q,\boldsymbol{\delta}_{g_1},\sigma,\gamma_1,\gamma_2,\ldots,\gamma_{K+1})$ form together the parameter vector $\boldsymbol{\delta}$ determining the intensity with which alternatives are rendered available to individuals with characteristics $(\mathbf{x}_q,\mathbf{x}_{g_1})$, as discussed earlier, and which has to be estimated jointly with the preference parameters $\boldsymbol{\beta}$.[31] Where necessary, the dependency of $f(k)$ on $\boldsymbol{\delta}$ will explicitly referred to in the notation as follows: $p(\boldsymbol{\delta};\mathbf{x}_q,\mathbf{x}_{g_1},w_k,h_k) := f(k)$.

The suitable job offers rendered available to an individual are however not observed. But, the econometrician can revert to the method of sampling a choice set, and constructing a sampled likelihood, in order to try to estimate the relevant parameters of the model, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$. The alternatives to be sampled are wage labour time regime combinations.

The sampling methods described here make a fixed number of draws n_s form the set $\{(0,0)\}\cup\mathbb{R}_{++}\times[\,H_{\min},H_{\max}\,[$. With an *a priori* fixed probability, say π_0^{obs}, a draw picks a non–market alternative (that is $(w,h)=(0,0)$). Else, a wage–labour time regime is fixed by independently sampling the wage from a log normal distribution with *a priori* determined location and scale parameters, μ and ς, and the labour time regime from a uniform distribution on the interval $[\,H_{\min},H_{\max}\,[$. The probability (density, in case $(w,h)>0$) to draw a wage–labour time bundle (w,h), denoted by $\mathbb{P}(w,h)$, is thus[32]:

$$\mathbb{P}(w,h)=\begin{cases}\pi_0^{\text{obs}} & \text{if} \quad (w,h)=(0,0)\,,\\ (1-\pi_0^{\text{obs}})\cdot n(w;\mu,\varsigma)\cdot\frac{1}{H_{\min}-H_{\max}} & \text{if} \quad (w,h)\in\mathbb{R}_{++}\times[\,H_{\min},H_{\max}\,[,\\ 0 & \text{else.}\end{cases}\tag{A.17}$$

Four options are considered: either one implements the probability of drawing the non–market alternatives by fixing the number of times the non–market alternative is to be drawn, say n_0, as being equal to the integer nearest to $\pi_0^{\text{obs}}\cdot(n_s+1)$. Or, one treats the number of times the non–market alternative is drawn, to be a random number, say K_0, with expected value $\mathrm{E}\{K_0\}=\pi_0^{\text{obs}}\cdot(n_s+1)$, by letting the probability of each draw to be the non–market alternative to be equal to π_0^{obs}.

[29] See Section 3.2. An illustration of the shape of that distribution function is given in Figure 1.

[30] See Section 3.2.

[31] The vector of covariates determining the value of the density $f(k)$ referred to earlier, and denoted as $\mathbf{x}_{f,k}$, thus consists in this case of $(\mathbf{x}_q,\mathbf{x}_w,w_k,h_k)$.

[32] See equation (25).

Secondly, one can sample with or without replacement. Lemp and Kockelman (2012) argue that sampling with replacement is inefficient since the same alternative may appear several times in the sampled choice set while not yielding additional information on choice behaviour. But on the other hand they warn that figuring out the selection probabilities of the sampled choice set $\mathcal{D}$, denoted earlier as $\Pi(\mathcal{D})$, might be cumbersome (with larger datasets) if the sampling takes place without replacement.
One way to implement sampling without replacement, is selecting n_s alternatives according to the *a priori* determined probability (density) $\mathbb{P}(w,h)$ (see equation A.17), and then remove the repeated draws from the sampled set. If one samples without replacement, the size of the sampled choice is a random number, while it is fixed when drawing takes place with replacement.
Finally, to reassure that the actually chosen alternative belongs to the sampled choice set, one adds it to the sampled set if it was not selected yet, or else, an additional randomly chosen alternative is drawn according to the probability (density) $\mathbb{P}(w,h)$.
In case draws are with replacement, this renders the size of the sampled choice set to be equal to n_s+1. Otherwise, when drawing is without replacement, the expected size of the sampled choice set is equal $n_s+1-\left(\pi_0^{\text{obs}}\cdot(n_s+1)-1\right)$.[33]
In the present application we opted for sampling with replacement according to equation (A.17), and treating the number of non–market alternatives as a random variable. In that case, the probability to draw a choice set $\mathcal{D}$, $\Pi(\mathcal{D})$, equals:

$$\Pi(\mathcal{D})=\left(\pi_0^{\text{obs}}\right)^{k_0}\left(\frac{1-\pi_0^{\text{obs}}}{H_{\min}-H_{\max}}\right)^{n_s+1-k_0}\prod_{(w_k,h_k)\in\mathcal{D}\setminus\{(0,0)\}}\mathfrak{n}\left(w_k;\mu,\varsigma\right), \tag{A.18}$$

where k_0 is the actual realisation of K_0 for a specific draw of the choice set.
The conditional sampling probabilities, $\pi(\mathcal{D}\,|k,\mathcal{C})$, equal $\Pi(\mathcal{D})/\mathbb{P}(w_k,h_k)$ for all $(w_k,h_k)\in\mathcal{D}$. So, the positive conditioning property is satisfied.
The sampled log–likelihood function thus reduces to[34]:

$$\mathcal{L}(\boldsymbol{\beta},\boldsymbol{\delta};\mathbf{X}):=-\sum_t\ln\left(\sum_{k\in\mathcal{D}_t}\left[\exp\left[v\left(\boldsymbol{\beta};\mathbf{x}_{t,k}\right)-v\left(\boldsymbol{\beta};\mathbf{x}_{t,j(t)}\right)\right]\frac{p\left(\boldsymbol{\delta};\mathbf{x}_{t,q},\mathbf{x}_{t,g_1},w_k,h_k\right)}{p\left(\boldsymbol{\delta};\mathbf{x}_{t,q},\mathbf{x}_{t,g_1},w_{j(t)},h_{j(t)}\right)}\frac{\mathbb{P}\left(w_{j(t)},h_{j(t)}\right)}{\mathbb{P}\left(w_k,h_k\right)}\right]\right). \tag{A.19}$$

For couples, the elements of the choice set to be sampled consist of quadruples $(w_1,h_1,w_2,h_2)\in[\{(0,0)\}\cup\mathbb{R}_{++}\times[H_{\min},H_{\max}[\,]^2$, specifying a wage labour time regime for each partner i ($i=$

[33] Of course, if the number of times the non–market alternative is drawn, was fixed, the size of the sampled choice set will be fixed too, and equals $n_s+1-(n_0-1)$.

[34] In the RURO model, the systematic part of the utility function depends solely on the wage and labour time characteristics of a job, but parameters of the utility function may depend on individual characteristics, say $\mathbf{x}_{t,V}$. Therefore $v(\boldsymbol{\beta};\mathbf{x}_{t,k})$ is in fact a shorthand for $v\left(\beta\left(\mathbf{x}_{t,V}\right);w_{t,k},h_{t,k}\right)$, with $\beta\left(\mathbf{x}_{t,V}\right)$ a function of the variables $\mathbf{x}_{t,V}$. See Section 2.2 for the relation of preferences defined over the space of consumption leisure time bundles, and preference in the wage–labour time regime space.

1, 2). Sampling of these alternatives is done by independently sampling for each of the partners a wage labour time regime bundle (w_i, h_i) according to the same prior as for singles, $\mathbb{P}_i(w_i, h_i)$ (see equation A.17). The probability (density) $\mathbb{P}_i$ is indexed on the partner i ($i = 1, 2$) since their characterising parameters, $\left(\pi_{i,0}^{\text{obs}}, \mu_i, \varsigma_i\right)$ need not to be the same for both partners.
So, importance sampling of alternatives (w_1, h_1, w_2, h_2) takes place according to according to the probability (density) function:

$$\mathbb{F}(w_1, h_1, w_2, h_2) = \mathbb{P}_1(w_1, h_1) \cdot \mathbb{P}_2(w_2, h_2). \tag{A.20}$$

Denote the sampled choice set for couples by $\mathcal{D}^c$, and let $\mathcal{D}_i^c$ be the set of combinations of wages and labour time regimes in the sampled choice set $\mathcal{D}^c$ that pertain to partner i. That is, $\mathcal{D}_1^c := \{(w_1, h_1) \,|\, \exists\, (w_2, h_2) : (w_1, h_1, w_2, h_2) \in \mathcal{D}^c\}$, and $\mathcal{D}_2^c := \{(w_2, h_2) \,|\, \exists\, (w_1, h_1) : (w_1, h_1, w_2, h_2) \in \mathcal{D}^c\}$. Furthermore, let k_{00}, k_{01}, k_{10} be respectively the actual number of times the alternative in which both do not accept any job offer is drawn, the number of times an alternative in which the first, respectively second, partner does not accept a job offer while the second, respectively first does, is drawn, for a specific outcome of the sampling procedure. We then obtain the following sampling probability for a choice set $\mathcal{D}^c$, $\Pi(\mathcal{D}^c)$:

$$\begin{aligned}\Pi(\mathcal{D}^c) = \;& \left(\pi_{1,0}^{\text{obs}} \cdot \pi_{2,0}^{\text{obs}}\right)^{k_{00}} \cdot \left(\pi_{1,0}^{\text{obs}} \cdot \left(1 - \pi_{2,0}^{\text{obs}}\right)\right)^{k_{01}} \cdot \left(\left(1 - \pi_{1,0}^{\text{obs}}\right) \cdot \pi_{2,0}^{\text{obs}}\right)^{k_{10}} \cdot \\ & \left(\left(1 - \pi_{1,0}^{\text{obs}}\right) \cdot \left(1 - \pi_{2,0}^{\text{obs}}\right)\right)^{n_s+1-k_{00}-k_{01}-k_{10}} \cdot \\ & \left(\frac{1}{H_{\min} - H_{\max}}\right)^{2(n_s+1-k_{00})-k_{01}-k_{10}} \cdot \prod_{i=1,2} \left(\prod_{(w_{i,k}, h_{i,k}) \in \mathcal{D}_i^c \setminus \{(0,0)\}} n\left(w_{i,k}; \mu_{i,0}, \varsigma_{i,0}\right) \right).\end{aligned} \tag{A.21}$$

The conditional sampling probabilities, $\pi(\mathcal{D}^c \,|k, \mathcal{C})$, equal $\Pi(\mathcal{D}^c) / \left(\prod_{i=1,2} \left(\mathbb{P}_i(w_{i,k}, h_{i,k})\right)\right)$ for all $(w_{1,k}, h_{1,k}, w_{2,k}, h_{2,k}) \in \mathcal{D}^c$. So, the positive conditioning property is satisfied.
The part of the sampled log–likelihood function that pertains to observations on couples, becomes:

$$\begin{aligned}&\mathcal{L}(\boldsymbol{\beta}_c, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2; \mathbf{X}) := \\ &-\sum_t \ln\left(\sum_{k \in \mathcal{D}^c{}_t} \left[\exp\left[v_c(\boldsymbol{\beta}_c; \mathbf{x}_{t,k}) - v_c\left(\boldsymbol{\beta}_c; \mathbf{x}_{t,j(t)}\right)\right] \prod_{i=1,2} \left(\frac{p\left(\boldsymbol{\delta}_i; \mathbf{x}_{t_i,q}, \mathbf{x}_{t_i,w}, w_{i,k}, h_{i,k}\right)}{p\left(\boldsymbol{\delta}_i; \mathbf{x}_{t_i,q}, \mathbf{x}_{t_i,w}, w_{j(t_i)}, h_{j(t_i)}\right)} \frac{\mathbb{P}_i\left(w_{j(t_i)}, h_{j(t_i)}\right)}{\mathbb{P}_i(w_{i,k}, h_{i,k})} \right)\right]\right).\end{aligned} \tag{A.22}$$

Here, v_c is the household utility function which is defined over quadruples (w_1, h_1, w_2, h_2), and parameters $\boldsymbol{\beta}_c$ may depend on household characteristics or individual household members' characteristics, as further specified in Section 3.2.

Appendix III Type specific unemployment rates

The next table contains the type specific unemployment rates that were used as additional identifying variable in the q–functions (the rate of availability of market *versus* non–market alternatives). These unemployment rates are gender, age, and educational attainment specific.

Table A.I: Type specific unemplyoment rates (%)

	Male			Female		
	Education level			Education level		
Age group	Low	Middle	High	Low	Middle	High
15 to 24 years	26.4	14.0	12.3	33.6	22.1	11.0
25 to 29 years	19.0	7.6	6.9	29.7	13.1	4.8
30 to 34 years	18.0	6.6	3.1	23.5	9.3	3.3
35 to 39 years	11.6	5.3	2.0	21.2	6.9	3.2
40 to 44 years	9.5	4.2	2.9	12.2	6.2	3.0
45 to 49 years	7.4	2.8	2.7	9.3	5.8	2.4
50 to 54 years	7.0	3.7	2.3	10.1	7.0	3.5
55 to 64 years	4.7	3.0	3.0	5.8	7.8	5.3[a]

[a] The exact figure is lacking. The average across all education levels for that age class is taken.
Source: Eurostat Unemployment rates by sex, age and educational attainment level (%), Belgium 2007, downloaded in October 2013.

Appendix IV Transition matrices

The row totals of the next tables represent the number of persons in the sample observed in the different labour time regimes (no participation, one to 18.5 hours, half time (18.5–20.5), 20.5 to 29.5 hours, three quarter time (29.5–30.5), 30.5 to 37.5 hours, full time (37.5–40.5), and more than full time), while the cells for each row reflect how these are distributed across the labour regimes according to the simulation. The column totals compose then the simulated marginal distribution. The histograms in the main text (Figures 9 and 12) are thus based on a comparison of the row and corresponding column totals.

Table A.II: Observed versus simulated labour time: males in couples

Observed	Simulated no part.	1.0–18.5	18.5–20.5	20.5–29.5	29.5–30.5	30.5–37.5	37.5–40.5	40.5–100	row tot.
no part.	38	1	2	2	1	9	28	18	99
1.0–18.5	0	0	0	1	0	0	5	1	7
18.5–20.5	0	0	0	0	0	0	9	5	14
20.5–29.5	2	0	0	0	0	1	13	3	19
29.5–30.5	4	0	0	1	2	1	16	4	28
30.5–37.5	10	8	2	11	3	21	88	34	177
37.5–40.5	46	16	11	34	12	54	407	183	763
40.5–100	17	5	2	14	8	35	178	91	350
col. tot.	117	30	17	63	26	121	744	339	1457

Table A.III: Observed versus simulated labour time: females in couples

Observed	Simulated no part.	1.0–18.5	18.5–20.5	20.5–29.5	29.5–30.5	30.5–37.5	37.5–40.5	40.5–100	row tot.
no part.	130	15	26	27	18	17	53	14	300
1.0–18.5	20	7	2	0	2	4	25	6	66
18.5–20.5	34	17	7	8	12	13	37	11	139
20.5–29.5	23	18	15	17	7	14	37	17	148
29.5–30.5	14	9	13	11	7	11	32	8	105
30.5–37.5	37	24	18	22	14	17	75	18	225
37.5–40.5	48	44	33	26	24	32	113	52	372
40.5–100	10	7	9	7	5	11	36	17	102
col. tot.	316	141	123	118	89	119	408	143	1457

Table A.IV: Observed versus simulated labour time: single males

Observed	Simulated no part.	1.0–18.5	18.5–20.5	20.5–29.5	29.5–30.5	30.5–37.5	37.5–40.5	40.5–100	row tot.
no part.	47	3	0	1	1	3	30	10	95
1.0–18.5	0	0	0	0	0	0	2	0	2
18.5–20.5	3	1	0	0	1	0	3	0	8
20.5–29.5	1	0	1	1	0	0	7	2	12
29.5–30.5	0	0	0	0	0	0	1	3	4
30.5–37.5	4	1	2	1	1	4	25	11	49
37.5–40.5	24	8	4	16	1	17	90	39	199
40.5–100	6	1	2	2	1	9	41	18	80
col. tot.	85	14	9	21	5	33	199	83	449

Table A.V: Observed versus simulated labour time: single females

	Simulated								
Observed	no part.	1.0–18.5	18.5–20.5	20.5–29.5	29.5–30.5	30.5–37.5	37.5–40.5	40.5–100	row tot.
no part.	103	11	4	10	11	8	26	9	182
1.0–18.5	3	0	0	0	0	0	6	0	9
18.5–20.5	11	3	1	0	1	2	8	3	29
20.5–29.5	4	2	1	1	3	1	6	7	25
29.5–30.5	7	1	1	1	1	2	13	1	27
30.5–37.5	18	4	5	3	2	4	22	6	64
37.5–40.5	35	8	6	21	11	12	62	29	184
40.5–100	8	1	1	2	1	5	8	25	51
col. tot.	189	30	19	38	30	34	151	80	571