

The Labour Market in CGE Models[§]

written as a chapter for the “Handbook of CGE Modeling”

(eds. P. Dixon and D. Jorgenson)

Stefan Boeters*

CPB, Netherlands Bureau for Economic Policy Analysis, Den Haag

Luc Savard**

Université de Sherbrooke

October 2011

Abstract

This chapter reviews options of labour market modelling in a CGE framework. On the labour supply side, two principal modelling options are distinguished and discussed: aggregated, representative households and microsimulation based on individual household data. On the labour demand side, we focus on the substitution possibilities between different types of labour in production. With respect to labour market coordination, we discuss several wage-forming mechanisms and involuntary unemployment.

Keywords: computable general equilibrium model, labour market, labour supply, labour demand, microsimulation, involuntary unemployment

JEL Code: C68, D58, J20, J64

[§]We thank Peter Dixon, James Giesecke, John Hutton, Egbert Jongen, Andreas Peichl, Sebastian Siegloch and Peter Sørensen for helpful comments on earlier versions of this chapter.

*Stefan Boeters, CPB, P.O. Box 80510, 2508 GM Den Haag, Netherlands, e-mail: s.boeters@cpb.nl

**Luc Savard, Department of Economics, Université de Sherbrooke, 2500, boulevard de l'Université, Sherbrooke (Québec), J1K 2R1, Canada, e-mail: luc.savard@usherbrooke.ca

1 Introduction

If we look at the body of CGE literature as a whole, the labour market has certainly not been one of the main points of attention. In fact, many of the classical CGE studies in the areas of trade liberalisation, tax analysis and climate policy work with the simplest possible set of assumptions about the labour market: labour supply is fixed and a uniform, flexible, market-clearing wage balances labour supply and demand. The authors of these classical studies apparently did not fear to introduce a serious bias into their analysis when treating the labour market in this simplified manner. And they may have been right. Even if one is convinced that the real labour market is much more complex than our simplifying model, this does not automatically mean that its full complexity must show in every concrete analysis. Engaging in a more detailed modelling of the labour market is only worth the while if it can be made plausible that assumptions about the labour market mechanisms actually change the outcome of a particular study significantly. Given that in a modelling context we are bound to work with simplifications anyway, we see the burden of proof with those claiming that this is the case.

There are two typical – and quite distinct – constellations that motivate researchers to go beyond the basic set-up of a labour market clearing, flexible wage. They either want to analyse a specific change in the labour market institutions, or they are interested in the labour market consequences of a policy measure which is not directly labour-market related. As an example of the first motivation, take the case of in-work benefits, which are supposed to increase labour supply of the low skilled. Addressing this issue in a CGE model requires a mechanism which endogenises labour supply and a labour-market segmentation which separates the low-skilled from other groups. Thus a certain level of labour-market complexity is necessary. The second motivation can be illustrated by the analysis of trade liberalisation. This affects the labour market only indirectly, but we can ask about the aggregate consequences – What are the effects on wages and unemployment? – as well as about distributional effects: Who gains and who loses? We give a more extensive overview of motivating issues in Section 2.

Apart from these question-driven approaches to labour market modelling, there is also an approach that one could call “presentation-driven”. Say, you work in the

field of climate policy analysis – an issue that is neither intrinsically labour-market related nor likely to have significant effects on the labour market. However, when you present your results, they are called into doubt, because “your model doesn’t even allow for involuntary unemployment” (which, as everyone admits, is a worrisome feature of most economies). Trying to convince your audience that including unemployment does not make much of a difference would divert the discussion from the main issue and might be difficult after all. So it can be a sensible strategy to add a more complex labour-market module to the model, if only to prevent people from digressing. A similar constellation can be found “history-driven” when modellers start off from an existing model containing labour market features that are irrelevant for the question at hand, but would require some effort to eliminate. In both presentation- and history-driven contexts, we want to make sure that the labour-market features present in the model do not complicate the interpretation of the results unnecessarily by producing spurious effects.

We fully recognise that model development in practice works under many restrictions that are not strictly academic. This is what we have often experienced in the work with our own models as well. Nevertheless, as economists, our thinking is dominated by the question-driven approach. In this chapter, we want to advocate and support a “question precedes model” strategy. In our view, the ideal set-up of a CGE study is the following. First, we need to formulate a clear question to be answered, preferably more specific than merely “analysing the consequences of policy X”. Often we can get, simply by phrasing the question clearly, a good idea of which model features will be relevant to the outcome, and which will not. It is a question of modelling efficiency to focus on the first, and to disregard the latter.

Stated broadly, this chapter has two objectives: (1) giving an overview of what options there are for labour market modelling in a CGE framework, (2) discussing advantages or disadvantages of these options, depending on the modelling context. The structure of the chapter is derived from the three major parts of any labour market module – labour supply, labour demand and market coordination – and from two directions of model development: refinement of mechanisms and disaggregation of units. With respect to labour supply, we primarily focus on the distinction between the representative-household and microsimulation approaches. Concerning labour demand, substitutability and complementarity of different types of labour

in production are centre stage. Finally, when it comes to labour market coordination, we review different theories of imperfectly competitive labour markets. – The second structuring dimension distinguishes between two strands in the development of labour market modelling: more complex mechanisms and a deeper disaggregation. Starting from the default option of almost all first-generation CGE models – market-clearing wages in a single labour market – we can in principle develop in both directions independently: (a) more complex mechanisms, say endogenous unemployment, at the same level of aggregation, or (b) the same, simple mechanism at a deeper level of disaggregation.

In many cases, however, there are interactions between complexity and disaggregation, which we will explore in this chapter as well. Let us illustrate this by three examples. (1) The introduction of involuntary unemployment confronts us with characteristically different unemployment rates for different groups of workers, which leads us to treat these labour market segments separately. (2) We find that an important mechanism of labour market coordination – collective wage bargaining – is only relevant to particular sectors, which requires us to think about sectoral labour mobility. (3) The differentiation between male and female labour supply, which is necessary if we want to do justice to empirical labour supply elasticities, raises questions about the substitutability of male and female work in production. We must always be aware of the fact that introducing a new model feature, which may be well motivated in a certain context, can create loose ends at other points in our model.

We try to give a comprehensive overview of modelling options, but in some respects we have been selective. First, we mainly focus on issues that can be treated in static or recursively dynamic models. Problems that require a dynamic model with forward-looking agents, such as life-cycle decisions like educational choice and the timing of retirement, are not covered. Second, business-cycle issues, such as the role of sticky wages in the propagation of shocks (new Keynesian features typically covered in DSGE models), are beyond the scope of this chapter. Finally, we concentrate on models at the national or multi-national level and disregard special problems of factor mobility that arise in regional modelling.

2 A classification of labour-market related questions

In the introduction, we have advocated the perspective of seeing models as tools to answer questions. The most straightforward type of question is: “What effect will policy intervention X have on economic indicator Y?” Depending on which kind of policy intervention and which kind of economic indicator we have in mind, the criteria for an assessment of CGE labour market modelling may vary significantly. In this section, we look more systematically at the types of questions that have been addressed with CGE models. This results in a classification that is used for structuring a list of typical CGE studies with a labour market focus.

Location of the initial shock

It is important to be clear about the location where the initial policy shock enters the model. Does it affect the labour market directly or only indirectly through other elements of the model? Let us return to the examples mentioned in the introduction. Take the case that we want to analyse the macroeconomic consequences of a policy encouraging labour supply of low-skilled workers, e.g. some in-work benefit system. Then we need a labour market module that is sufficiently complex for the policy shock to be meaningfully modelled. In this case this means that labour supply must be flexible and the low-skilled must be treated as a separate group.

Many of the big themes of CGE modelling, in contrast, are not directly labour-market related. International trade liberalisation hits export and import markets, climate policy measures affect energy markets, and the impact of both policies on the labour market is only indirect through a shift in labour demand, i.e. in the real wages that can potentially be paid to the workers at a given level of employment. The same applies to tax policy analysis, as far as it is concerned with capital taxation, corporate taxation, intermediate input taxes or consumption taxes. The only exemption is wage taxation, which needs a labour market representation with flexible labour supply and heterogeneous wages for a meaningful analysis.

In all cases in which the labour market is only affected indirectly through labour demand, we are confronted with the following key question: Will the real wage follow the movement of the marginal product of labour one-to-one, as it would

in a perfect labour market? Or is there some sort of wage rigidity that hinders a parallel movement? Put into even more policy-relevant terms: To what extent will higher demand for labour translate into higher wages, and to what extent into more employment (and vice versa for lower labour demand)? What is asked for is a “wage curve”, i.e. a functional relationship between unemployment and the wage, which then in turn determines the wage-employment split. We return to the issue of the wage curve in Section 5.3, where we review different options of modelling imperfect labour markets.

Outcome variables of interest

Just as important as the location of the initial shock is the outcome variable we are interested in. Potential outcome variables of CGE models can be classified according to the level of aggregation they require. At one extreme of the spectrum, we have typical macro variables, which give us an impression of the overall economic effect of the policy measure analysed: GDP, national income, exports and imports, consumption and investment, or a welfare measure such as the Hicksian equivalent variation. At an intermediate level of aggregation, we have sectoral effects: output, employment, productivity, exports and imports by sector, which are prominent in many core issues of CGE modelling (e.g. trade liberalisation, climate policy). In this chapter, however, they are only of interest in so far as they affect the labour market. Labour market variables with a comparable, intermediate level of aggregation are group-specific outcomes such as wages, participation, employment and unemployment by skill group or gender. Finally, if the model allows for full disaggregation, we have the additional option of reporting these variables by socio-demographic attributes that have no functional role in the labour market mechanism modelled, e.g. income class, age, education or number of children. This kind of reporting is normally motivated by distributional concerns.

Unlike the location of the initial shock, the outcome variables used will not directly constitute a classification criterion of studies. The criterion is rather whether a model encompasses both the macro and the micro level. If it does, this will normally also be reflected in the reporting of results.

A classification of typical studies

Let us categorise a number of typical CGE studies with a labour market focus according to the criteria developed in the preceding paragraphs: Where is the policy shock located? Does the model contain a genuine micro level?

We start, in roughly chronological order, with studies that focus on labour market shocks:

- Early attempts of addressing labour market issues in a CGE framework are Gelauff et al (1991) and Dewatripont et al (1991), who analyse labour taxation and social security contributions in the Netherlands and Belgium, respectively.
- Sørensen (1997) studies options of stimulating low skilled employment (tax cut for low incomes and consumption tax relief on low-skilled intensive services) in a model calibrated to the Danish economy.
- Hutton and Ruocco (1999), and Böhringer et al (2005) analyse changes in labour taxation with an aggregated labour market module. The wage generating mechanism is efficiency wages in the first paper and collective bargaining in the latter.
- Bovenberg et al (2000) focus on tax reform as well, but in a model that allows for more dimensions of labour market heterogeneity. A full-fledged version of their model for the Dutch economy is presented in Graafland et al (2001).
- Aaberge et al (2004), Arntz et al (2008) and Boeters (2010) are examples of integrating microsimulation elements with a focus on re-financing the pension system, stimulation of low-skilled employment and tax progressivity, respectively.
- Agénor et al (2007) simulate various labour market policy measures (reduction in payroll taxation, cuts in public sector wages and employment, reduction in trade unions' bargaining power) in a model with a dual labour market and collective wage bargaining for a stylised Middle-East or North-African economy. Agénor and El Aynaoui (2003) is an application of this model to Morocco.

- Cogneau and Robilliard (2008) set up a linked microsimulation-CGE model of Madagascar for analysing poverty alleviation policies such as agricultural subsidies, a workfare scheme and untargeted per capita transfers.
- Dixon et al (2011) study the labour market effects of restricting employment of illegal immigrants in the U.S. by either stricter border controls or higher fines for employers.

A second set of CGE studies, again in roughly chronological order, analyse policy or macroeconomic shocks that do not directly hit the labour market, but nevertheless have effects on employment and distribution that depend on the labour market specification:

- Ballard et al (1985) is a classical study on tax policy. Their discussion of labour-market issues, however, is restricted to choosing an appropriate value of the aggregated elasticity of labour supply.
- The study of de Melo and Tarr (1992) has a seminal status for the analysis of trade liberalisation. As a specific labour market feature, they introduce wage bargaining in the automobile sector, which naturally leads them to a kind of dual-labour-market structure.
- The trade liberalisation issue has been linked to poverty analysis in models that use a full microsimulation-CGE linkage, e.g. Hérault (2007) for South Africa, Bourguignon and Savard (2008) for the Philippines and Bussolo et al (2008) for Latin America.
- Fæhn et al (2009) and Fraser and Waschik (2010) are two studies from the field of energy economics and climate policy analysis that have a special focus on the interactions of energy and labour markets.
- The impact of macroeconomic shocks such as financial or currency crises is analysed in Ferreira et al (2008) for Brazil and in Robilliard et al (2008) for Indonesia.

3 Labour supply

According to the sketch in the introduction, labour supply modelling can develop towards a higher degree of complexity in two ways: (1) more subtle labour supply mechanisms and (2) a lower level of aggregation. The structure of this section is derived from the aggregation dimension. We start at the most aggregated level of a single representative household (Section 3.1) and show how basic calibration tasks can be approached: (a) implementing empirically plausible labour supply elasticities, (b) differentiating labour supply along the intensive and extensive margin, and (c) allocating involuntary unemployment. In Section 3.2, we discuss the changes resulting from the existence of several representative households instead of a single one. Finally, in Section 3.3, we turn to microsimulation, where labour supply is implemented at the lowest possible aggregation level, i.e. at the level of the individual household.

3.1 Labour supply of a single representative household

At the level of a single representative household, unbothered by disaggregation issues, we can concentrate on the task of modelling labour supply in a way that is consistent with given empirical elasticities. In many classical CGE models (e.g. Dervis et al, 1982), in addition to working with a single representative household, it is assumed that the labour supply of this household is fixed. Once we want to model flexible labour supply, we are confronted with a crucial distinction. Labour supply is flexible along two margins: hours of work (intensive margin) and participation (extensive margin). In this section, we show how labour supply of a representative household can be calibrated to a set of three aggregate labour supply elasticities: (a) elasticity of participation with respect to the wage, (b) elasticity of working hours with respect to the wage, (c) elasticity of working hours with respect to (non-wage) income. The calibration is performed by determining the parameters of a conventional utility function comprising material consumption and leisure.¹

¹The following material is adapted to the present context from Boeters and van Leeuwen (2010).

3.1.1 Hours of work

We consider a worker household that must decide on its hours of work under a budget constraint and a time constraint. The budget constraint is

$$p_C (C_D + C_0) = wH (1 - t_a) + Y_0$$

where p_C is the consumption price index, C_D and C_0 are disposable and necessary consumption, respectively, w is the wage rate, H is hours of work, t_a is the average tax rate on labour income, and Y_0 is non-labour income.² The time constraint is

$$F + H = T$$

with leisure F and time endowment T . The choice of the worker household is modelled as the maximisation of a utility function that covers disposable consumption and leisure, $U_e = U_e(C_D, F)$. As our task is to determine concrete functional parameters, we assume a CES utility function³ with parameters θ_C and σ

$$U_e = \left[\theta_C \left(\frac{C_D}{\bar{C}_D} \right)^{\frac{\sigma-1}{\sigma}} + (1 - \theta_C) \left(\frac{F}{\bar{F}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \quad (1)$$

From this utility function, we can derive the following expenditure and demand functions, where variables with an upper bar denote initial (and thus constant) values:

$$\begin{aligned} p_U &= \left[\theta_C p_C^{1-\sigma} + (1 - \theta_C) \left(\frac{w(1 - t_m)}{\bar{w}(1 - \bar{t}_m)} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \\ \frac{C_D}{\bar{C}_D} &= U_e \left(\frac{p_U}{p_C} \right)^{\sigma} \\ \frac{F}{\bar{F}} &= U_e \left(p_U \frac{\bar{w}(1 - \bar{t}_m)}{w(1 - t_m)} \right)^{\sigma} \end{aligned}$$

²Compared to the simplest possible textbook example of labour supply, this formulation contains three extensions that are important for empirical calibration: necessary consumption, non-labour income and a variable average tax rate, which causes average and marginal tax rates to diverge. As we focus on static models, we do not extend the model with savings (see Rutherford (1998) for the joint calibration of labour supply and savings).

³We use the “calibrated share form” of the CES function, see Rutherford (1998). By expressing all quantities and prices as multiples of the initial values, this form clearly conveys the ideas that quantity normalisations are arbitrary and that the essential information is about relative changes. In addition, the value shares in the initial situation can without transformation be used as share parameters of the function.

p_U is the necessary expenditure for one unit of utility and t_m is the marginal labour income tax rate. The utility level can alternatively be calculated as

$$U_e = \frac{Y_D}{p_U}$$

where Y_D is disposable extended income defined as

$$Y_D = w [H(1 - t_a) + (T - H)(1 - t_m)] + Y_0 - C_0$$

evaluating leisure with the marginal after-tax wage rate.

Y_D is also used to calibrate the share parameter of the utility function (1). θ_C is the initial share of disposable consumption in disposable extended income,

$$\theta_C = \frac{\bar{C}_D}{\bar{Y}_D} \quad (2)$$

and, correspondingly,

$$1 - \theta_C = \frac{\bar{w}(1 - \bar{t}_m)\bar{F}}{\bar{Y}_D}$$

In the following two subsections, we describe how labour supply at the hours-of-work margin is calibrated to empirical labour supply elasticities.

3.1.2 Income elasticity of labour supply

In this section, we argue that the disposable time endowment, T , should be calibrated in a way that produces an income elasticity of labour supply in an empirically plausible range. In contrast, an ad-hoc specification of T is likely to result in an unrealistic value of this elasticity. We follow the approach of de Melo and Tarr (1992). Ballard (2000) has highlighted this approach as a means of improving the empirical fit of the model.

Originating from a homothetic CES function, the demand functions are homogeneous of degree one in disposable extended income. We thus have⁴

$$\varepsilon_{FY_D} = 1$$

⁴We denote the elasticity $\frac{\partial \log x}{\partial \log y}$ by ε_{xy} or (if it is an empirical value to be reproduced in the model), η_{xy} .

From this we can derive the income elasticity of labour supply. To be precise, we calculate the per cent change in labour supply with respect to an exogenous variation of the non-labour income, Y_0 , that would increase $Y = wH(1 - t_a) + Y_0$ by one percent, if labour supply did not react.

$$\eta_{HY} = \varepsilon_{HY_0} \frac{Y}{Y_0} = \varepsilon_{HF} \varepsilon_{FY_D} \frac{dY_D}{dY_0} \frac{Y}{Y_D}$$

We have

$$\begin{aligned} \varepsilon_{HF} &= -\frac{T - H}{H} \\ \varepsilon_{FY_D} &= \frac{dY_D}{dY_0} = 1 \end{aligned}$$

and therefore

$$\begin{aligned} \eta_{HY} &= -\frac{T - H}{H} \frac{Y}{Y_D} \\ &= -\frac{T - H}{H} \frac{wH(1 - t_a) + Y_0}{w[H(1 - t_a) + (T - H)(1 - t_m)] + Y_0 - C_0} \end{aligned}$$

We treat η_{HY} as a parameter that we can observe empirically, and we use it to determine T , the (unobservable, disposable) time endowment. Solving for T , as a multiple of initial labour supply, gives

$$\begin{aligned} \frac{T}{H} &= \frac{[wH(1 - t_a) + Y_0] - \eta_{LY} (w[H(1 - t_a) - H(1 - t_m)] + Y_0 - C_0)}{[\eta_{HY} wH(1 - t_m)] + wH(1 - t_a) + Y_0} \\ &= 1 - \frac{\eta_{HY} [wH(1 - t_a) + Y_0 - C_0]}{\eta_{HY} wH(1 - t_m) + wH(1 - t_a) + Y_0} \end{aligned} \quad (3)$$

For small, negative values of η_{HY} , $T > H$ is warranted. At the same time, small absolute values of η_{HY} will result in a small amount of disposable leisure. In a simplified benchmark case with $Y_0 = C_0 = 0$ and proportional taxes ($t_m = t_a = t$), eq. (3) reduces to

$$\frac{T}{H} = 1 - \frac{\eta_{HY}}{1 + \eta_{HY}} = \frac{1}{1 + \eta_{HY}} \quad (4)$$

If we follow Ballard (2000) and set η_{HY} to the empirically plausible value of -0.1 , we arrive at $T/H \approx 1.1$. This may seem overly little: only 4 hours of disposable leisure in relation to a standard work week of 40 hours. In ad-hoc specifications, one often finds a value of 1.75 (e.g. Rutherford, 1998). However, this would lead to income elasticities of labour supply which are far beyond what we empirically observe.⁵

⁵Equivalently to calibrating the time endowment T , we can also set some arbitrary time endow-

3.1.3 Wage elasticity of labour supply

With the relative time endowment, T/H , determined by the income elasticity of labour supply, we proceed with calibrating the value of the elasticity of substitution between material consumption and leisure, σ , using the wage elasticity of labour supply⁶, η_{Hw} , which is calculated as

$$\eta_{Hw} = \varepsilon_{H\tilde{w}} = -\frac{T-H}{H}\varepsilon_{F\tilde{w}}$$

where $\tilde{w} = w(1 - t_m)$. The elasticity of leisure demand with respect to the marginal after-tax wage can be routinely decomposed into a substitution effect and an income effect. The income effect deserves attention, because we need the effect of the wage on the disposable extended income, Y_D .

$$\eta_{Hw} = -\frac{T-H}{H} \left[-\sigma\theta_C - (1 - \theta_C) + \frac{w(1 - t_m)T}{Y_D} \right]$$

Solving for σ gives the calibration equation

$$\sigma = \frac{\eta_{Hw} - \frac{T-H}{H} \left((1 - \theta_C) - \frac{w(1-t_m)T}{Y_D} \right)}{\frac{T-H}{H}\theta_C} \quad (5)$$

To get a feeling for magnitudes, we again consider the special case with $Y_0 = C_0 = 0$ and $t_m = t_a$. Then we have

$$\frac{w(1 - t_m)T}{Y_D} = 1$$

and eq. (5) simplifies to

$$\sigma = \frac{\eta_{Hw} + \frac{T-H}{H}\theta_C}{\frac{T-H}{H}\theta_C} = 1 + \frac{\eta_{Hw}}{\frac{T-H}{H}\theta_C} \quad (6)$$

Further simplification of eq. (6) is achieved by observing that in this case

$$\theta_C = \frac{H}{T},$$

ment (say, 24 hours a day) and calibrate a minimum level of leisure in the fashion of a Stone-Geary utility function. This approach has been taken in Annabi (2003), although without discussion of the consequences for the income elasticity of labour supply.

⁶To be precise, we deal with the elasticity of the hours of work with respect to the *marginal after-tax* wage. Differently specified elasticities require modifying the calculations accordingly.

which yields⁷

$$\sigma = 1 + \frac{T}{T - H} \eta_{Hw}$$

Finally, we insert eq. (4), which leaves us with

$$\sigma = 1 - \frac{\eta_{Hw}}{\eta_{HY}}$$

This shows that the inclusion of η_{HY} in the calibration makes the outcome for σ more volatile. With an exogenous, relatively large T/H ratio, a small value of η_{Hw} would have warranted a small deviation of σ from one. With η_{HY} additionally appearing in the equation, σ can easily assume much higher values. To get a feeling for numerical values, we follow Sørensen (1999) and set η_{Hw} to 0.1.⁸ Together with $\eta_{HY} = -0.1$ (as in Section 3.1.2), this produces $\sigma = 2$.

Alternatively, it would be possible to calibrate the model to the compensated and uncompensated elasticities of labour supply. Ballard and Fullerton (1992) use values of 0.2 and 0 in their benchmark case.

3.1.4 Labour supply: participation

When we proceed to the calibration of labour supply along the extensive margin (participation), we can no longer base our work on the fiction that the representative household represents a large number of identical individuals. The simplest way of implementing the difference between participating and non-participating households is to assume heterogeneity in their fixed cost of taking up work. Those with low fixed costs enter the labour market, whereas those with high fixed costs stay at home.⁹ It is not necessary to specify the precise nature of these fixed costs. They may consist of costs that are caused by the difficulties of family coordination if both partners have a paid job, commuting costs between home and work, or simply some kind of labour market attachment, an inherent utility from interacting with others in a productive environment.

⁷This is also what you get in Rutherford (1998), if you leave out the upper nest with the consumption-savings decision (assuming that the savings rate is zero).

⁸The meta analysis of Evers et al (2008) suggests a somewhat higher elasticity, but it is difficult to distil a core value from this study.

⁹For a general discussion of this approach, see Bourguignon and Magnac (1990), Magnac (1991), Kleven and Kreiner (2006a).

The two-step labour-supply decision (participation, hours of work) is solved backwards: First the individuals determine the optimal choice of hours *assuming* that they participate, then they compare this optimal outcome with the fixed cost of working. Things become slightly more complicated if there is involuntary unemployment. A possible assumption is that individuals draw a comparison between the (unemployment-weighted) expected utility of supplying labour and their respective fixed costs. In the case of unemployment (index u), utility is

$$U_u = \left[\theta_C \left(\frac{C_D^u}{\bar{C}_D} \right)^{\frac{\sigma-1}{\sigma}} + (1 - \theta_C) \left(\frac{T}{\bar{F}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (7)$$

where disposable consumption in the case of unemployment, C_D^u , is income less necessary consumption

$$C_D^u = Y_D^u = cwH(1 - t_a) + Y_0^u - C_0$$

and unemployment benefits are assumed to be a fixed replacement rate, c , multiplied with the after-tax income of the employed workers, $wH(1 - t_a)$. The formulation (7) creates a problem, however. All relevant variables in this equation are fixed, either institutionally (c, t_a) or through the calibration of the labour supply decision of the employed workers (θ_C, σ, T). For a reasonable unemployment model, we must have $U_u < U_e$, which is not automatically warranted. If several factors interact, U_u may turn out to be larger than U_e : As an outcome of the calibration (see Sections 3.1.2 and 3.1.3), $\bar{F}$ is typically only a small share of T , and the elasticity of substitution, σ , is considerably larger than one. Both these facts contribute to a high utility level of the unemployed. On the other hand, we have basic consumption, C_0 , which makes the relative difference between C_D^u and $\bar{C}_D$ larger than simply given by the replacement rate. If the first factor dominates, we end up with a utility reversal.

Finding a solution to this problem would require further exploring the value of involuntarily unemployed time, which seems to be an unresolved question in labour economics. The model can in principle easily be adjusted in order to allow for more flexibility. As it stands, the parameters of the utility function have been calibrated *locally* at the point where the employed workers supply labour. However, there is no strong reason to assume that the outcome of the calibration is also informative with respect to the utility *difference* between two distant points, $U_e - U_u$. We can

approach the utility reversal problem by introducing an additional parameter. This parameter allows for the possibility that unemployed workers cannot consume their total time endowment, T , as leisure. We can think of different reasons for this: Searching for a job requires time, even more so if the unemployed are expected to attend active labour market measures. A correction factor for disposable leisure could also capture effects of the social embeddedness that the work sphere supplies. However, it is particularly difficult to quantify this effect.¹⁰ In general terms, we may assume that a given fraction, δ , of the additional non-working time of the unemployed does not count as “leisure”:

$$U_u = \left[\theta_C \left(\frac{C_D^u}{\bar{C}_D} \right)^{\frac{\sigma-1}{\sigma}} + (1 - \theta_C) \left(\frac{T - \delta \bar{H}}{\bar{F}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

Given U_u , with the implied difficulties, we can calculate the expected utility of supplying labour, U_l ,

$$U_l = (1 - u)U_e + uU_u$$

which is the same for all individuals. They compare it with their idiosyncratic fixed cost of supplying labour, U_0 , and supply labour if $U_l > U_0$.

The distribution of the U_0 ’s over the population must be calibrated. As our empirical basis, we have the actual participation rate and the elasticity of labour supply at the extensive margin. This is sufficient to calibrate the distribution of the fixed costs locally (at the point of actual participation), but not globally. The rest of the distribution must be fixed by some functional assumption. A relatively simple assumption is that costs are uniformly distributed between U_0^- and U_0^+ . For fixing the values of these bounds, we first have to calculate the change in U_l caused by an exogenous variation in the wage. We consider the case of an isolated change in the wage of the respective individuals when they are employed. In this case, the unemployment rate and the utility from unemployment can be considered

¹⁰A possible line of investigation would be whether there are time-use studies that inform us about how much time the unemployed actually spend on searching. Jenkins and Montmarquette (1979) is a coarse trial to find indirect ways for evaluating unemployed time.

constant.¹¹ In terms of elasticities, we then have

$$\begin{aligned}\varepsilon_{U_l,w} &= \frac{(1-u)U_e}{U_l} \varepsilon_{U_e,w} = \frac{(1-u)U_e}{U_l} (\varepsilon_{Y_D,w} - \varepsilon_{p_U,w}) \\ &= \frac{(1-u)U_e}{U_l} \left(\frac{wT(1-t_m)}{Y_D} - \frac{wF(1-t_m)}{Y_D} \right) \\ &= \frac{(1-u)U_e}{U_l} \frac{wH(1-t_m)}{Y_D}\end{aligned}$$

The elasticity of labour supply at the extensive margin can be calculated as

$$\eta_{Nw} = \varepsilon_{N,U_l} \varepsilon_{U_l,w} = h \frac{(1-u)U_e}{N} \frac{wH(1-t_m)}{Y_D},$$

where h is the density of the fixed cost distribution and N is the number of participating individuals. Solving for h , we obtain

$$h = \eta_{Nw} \frac{NY_D}{(1-u)U_e wH(1-t_m)}. \quad (8)$$

Given a particular value for η_{Nw} ,¹² h can be evaluated at the initial point, and then treated as a constant in the counterfactual simulations. This means that the elasticity at the extensive margin is precisely reproduced only for the initial point; after the initial situation, it is endogenous.

The bounds of the uniform distribution for h can be determined as

$$\begin{aligned}U_0^- &= \bar{U}_l - \frac{\bar{N}}{h} \\ U_0^+ &= \bar{U}_l + \frac{N_0 - \bar{N}}{h}\end{aligned}$$

where N_0 is the total population and $\bar{N}$ is initial participation. Finally, counterfactual participation can be calculated as

$$N = \bar{N} + h(U_l - \bar{U}_l) \quad (9)$$

¹¹This would not be the case for a general change in the wage, which applies to all individuals.

¹²Kleven and Kreiner (2006b, p.18-20) survey the current state of empirical evidence on the elasticity at the extensive margin. It is particularly difficult to calibrate a model with a representative agent to these elasticities, because they differ considerably by household type. One might choose a value of 0.2, which is roughly the aggregate average in Kleven and Kreiner's core scenario.

3.1.5 Supply of different labour varieties

So far, we have assumed that the labour supplied by the representative household is homogeneous. If we want to distinguish between different types of labour, but remain in the setting with a single representative household, we can allow for transformation among the different labour supply options.¹³ For concreteness, let us abstract from any further complication associated with the valuation of leisure and focus on the distribution of a fixed endowment of time between two labour supply options. This can straightforwardly be modelled by a constant-elasticity-of-transformation (CET) function with the two options as arguments. Examples for this approach are Hutton and Ruocco (1999), who discuss full-time versus part-time work, Gaasland (2008) for the case of farm versus non-farm work and Cloutier et al (2008) for skilled versus unskilled labour. In formal terms, we have a given amount of labour supplied, $\bar{L}$, as a CET aggregate of two varieties, L_1 and L_2

$$\bar{L} = CET(L_1, L_2) = \left[\beta_1 L_1^{\frac{\tau-1}{\tau}} + \beta_2 L_2^{\frac{\tau-1}{\tau}} \right]^{\frac{\tau}{\tau-1}} \quad (10)$$

where the β_i are share parameters and $\tau < 0$ is the elasticity of transformation. The standard CET approach (analogously to, e.g., the transformation of domestic production into domestically sold and exported varieties) is to maximise earnings

$$\max_{L_1, L_2} Y = w_1 L_1 + w_2 L_2$$

for exogenous wages w_i , subject to the resource constraint (10). This gives first-order conditions

$$w_i = \frac{\partial CET(L_1, L_2)}{\partial L_i}$$

which determine the allocation of the endowment between the options.

The problem with the ordinary CET set-up is that the L_i will in general *not* add up to $\bar{L}$. This makes the interpretation of the result difficult, because it was precisely the aim of the exercise to *distribute* a given amount of labour supply between two options.¹⁴ As a reaction to this problem, other modellers (e.g. Dixon and Rimmer (2003) and Giesecke et al (2011), who deal with occupation-specific

¹³Modelling set-ups with differentiated households are discussed below in Sections 3.2 and 3.3.

¹⁴Magnani and Mercenier (2009), in a model of occupational choice, try to solve this problem by only deriving the *ratios* between the different types of labour supply from the CET set-up

labour supply) have approached the distribution of labour across varieties as an issue of substitutability. They use the ordinary time constraint

$$L_1 + L_2 = \bar{L} \quad (11)$$

and assume that *incomes* from the two varieties of labour are imperfect substitutes in the utility function¹⁵ U

$$U = CES(w_1 L_1, w_2 L_2) = \left[\gamma_1 (w_1 L_1)^{\frac{\sigma-1}{\sigma}} + \gamma_2 (w_2 L_2)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

with distribution parameters γ_i and elasticity of substitution $\sigma > 0$. The distribution of labour between the two varieties is modelled as utility maximisation subject to the resource constraint (11). This solves the additivity problem, but generates new difficulties in the interpretation. Why should incomes from different sources be imperfectly substitutable in generating utility? A possible interpretation is that individual households have varying innate affinity for the different labour supply options. They receive utility not only from income, but also from the closeness to their most preferred option. Households can be ranked according to their innate affinity, with those at the top of the list switching to the respective option first. The higher participation in a certain option, the lower therefore the marginal non-income valuation of this option, creating a smooth transformation from other options. An explicit model of closeness to the intrinsically preferred option has been included in MIMIC (Graafland et al, 2001, p. 84-86) for the choice between discrete hours-of-work options. Implicitly, similar assumptions are also at work in the standard discrete choice modelling of labour supply (see Section 3.3). In any case, it remains a challenge to make an approach of this sort a consistent integral part of a full model. One question to be answered in this context is: How can we account for income effects of the non-income utility from labour supply options on the demand of other goods and leisure?

and then combining these ratios with the time constraint. This cannot completely eliminate the consistency concerns. However, Magnani and Mercenier (2009) show that the resulting expression can be interpreted as the outcome of the aggregation of a mass of agents that are heterogeneous in one dimension, which is an attractive feature of this approach.

¹⁵Giesecke et al (2011) use a CRESH utility function. Here, I simplify with a CES function.

3.2 Labour supply of several representative households

The approach of Section 3.1 can be extended to more than a single representative household. In general, there is no restriction to the number of representative households with which we can work. A common split in only two households is the distinction between low-skilled and high-skilled workers (e.g. Lejour et al, 2006). At the other end of the spectrum, there are models with as much as 100 representative households (e.g. Piggott and Whalley (1985), who differentiate between households by household composition, profession and wage level, or Dixon and Rimmer (1995), who have the marginal propensity to consume as an additional criterion for differentiation). At a certain level of disaggregation, however, the question arises naturally whether one should then not rather switch to microsimulation, where the unit is the individual household (see Section 3.3).

When working with several representative households, we must decide on the criteria of differentiation. This is not always clear-cut, because several types of arguments tend to interfere. First, and foremost, we want the household structure to respond to the research question pursued. For example, when we try to answer distributional questions, a household differentiation by income class is natural. When the research motivation is female labour market participation, we need to differentiate by household composition. Second, however, any disaggregation at the household level requires complementary assumptions with respect to labour demand and labour market coordination. When we distinguish between two households, do we also want to specify diverging labour demand conditions or coordination mechanisms for the two types of labour supply? Or can labour of the two households simply be added up to a homogeneous aggregate? (For details, see Sections 4 and 5 respectively.) Our choice of household disaggregation criteria can be influenced by these follow-up problems. Third, we need data for the calibration of the differentiated households. Depending on the disaggregation criteria, these can be more or less easily available (see Section 3.2.2).

3.2.1 Possible household types

This section contains a list of criteria that have been used for separating representative households. For each criterion we discuss the motivation for the split and possible problems in implementation.

Skill type

The split into skill types, usually understood as level of education, responds to the huge amount of literature on skill-specific wage disparities and their possible reasons (skill-biased technological change and shifts in international trade patterns in the course of globalisation). It acknowledges that wages do not always move in parallel, which becomes relevant in situations where differential effects on labour markets of different skill types are plausible. A typical example of a situation of this sort is trade liberalisation, which changes the exposure of a country with imports from regions with different comparative advantages (e.g. Thierfelder and Robinson, 2002, or Carneiro and Arbache, 2003). Similar effects can be the consequence of sectoral re-allocations due to tax policy or climate policy measures.

Most of the literature on skill-specific labour market effects uses a split into two classes, high- and low-skilled (with a conventional cut-off point analogously to completed college education in the USA). This follows a long tradition of attempts to estimate the substitution pattern between these two skill classes and capital (see Thierfelder and Robinson, 2002). Conceptually, it is easy to extend the skill split to more than two classes. However, the more skill classes, the more challenging labour demand estimation, which becomes more likely to produce implausible substitution patterns (see Section 4). In addition, the more skill classes, the less plausible the implicit claim that skill is an unchangeable attribute of the households (i.e. that individuals cannot switch from one skill class to another). Jung and Thorbecke (2003) and Cloutier et al (2008) are two examples in which the choice of the skill type is endogenous, involving investment in education. Jung and Thorbecke (2003) work in a recursively dynamic context and let the education decision be governed by myopic expectations. The model of Cloutier et al (2008) is static, representing a long-term equilibrium. Transformability of skills is imperfect (CET function), and the choice between skills is driven by contemporary wages.

A data-related issue with skill classes in multi-country models is the problem of comparability of skills data across borders. The larger the differences between educational systems, the higher the obstacles to finding comparable data. Dimaranan and Narayanan (2008) explain how the skills split is implemented in the GTAP context. As detailed data are only available for a subset of the countries covered by GTAP, they estimate a functional relationship between the share of skilled labour

payments, growth of GDP per capita and the average number of years of tertiary education. This is used to generate values for the countries with missing data.

Household composition

The two most important dimensions of household composition are couples versus singles and the number of children. The differentiation between the resulting household types is mainly inspired by labour supply estimates. Labour supply flexibility of singles and couples shows huge differences, and the presence of children is a major factor determining female labour market participation. Most estimations of labour supply are performed at least at this level of disaggregation, and many labour market economists are very reluctant to use more aggregated values. A second motivation for disaggregation by household composition is a fiscal system that varies by household characteristics (e.g. Bahan et al, 2005).

Once we explicitly account for couples as a distinct household type, we are faced with a new problem: cross-effects of the income of one partner on the labour supply of the other. In this case, a full set of elasticities would be more extensive than the one discussed for the representative household in Section 3.1. This is one reason why the intricacies of couple households are normally not approached in the representative household setting, but rather by means of microsimulation (see Section 3.3).

A certain simplification is achieved, however, if we assume that not all labour supply is flexible. A common specification is to assume one of the partners in the household to be the breadwinner, who supplies labour inflexibly. Only the labour supply of the other household members is considered flexible and calibrated to empirical elasticities. This approach is followed, e.g., in the MIMIC model (Graafland et al, 2001). Two options of modelling representative couple households with flexibility of both partners' labour supply are explored and compared in Boeters et al (2005). In models of low-income economies, Fontana and Wood (2000) and Cockburn et al (2007) distinguish between labour supply of partners depending on the home production responsibilities of women.

Occupation

Classification by occupation is a close substitute to classification by skill type. There are three potential reasons for differentiating along the occupation instead of the skill (education) dimension. First, for many countries, labour classification by occupation

is more readily available (or even the only available information), e.g. in the official ILO data.¹⁶ Second, switching from one occupation to another may be more difficult than switching from low-skilled to high-skilled tasks within the same occupation. So labour supply might be better formulated in occupational than in educational terms. Third, there are professional organisations for occupations that limit the access to specific labour markets and thereby create wage differentials. A model that combines occupational and skill classifications (Giesecke et al, 2011) is discussed in Section 4.2.

Sectoral employment

Classification by sectoral employment is another close substitute for skill and occupation. If labour is immobile between sectors and if there are sectoral wage differentials, a sectoral classification of labour supply may be an appropriate way to capture the consequences of sectoral shifts caused by, e.g., trade liberalisation or climate policy. Decaluwé et al (2010) use this approach in a model of the Quebec economy.

In studies on low-income countries the distinction between workers attached to the rural versus the urban sector is important. This combines sectoral and regional aspects. In the short run, workers are attached to their respective sector. In the long run, however, mobility between the rural and the urban sectors must be taken into account (see Section 5.5). Whalley and Zhang (2004) use this household decomposition in a model of China to capture the Hukou system (constraints on movement of workers from rural to urban sectors).

Income class

Household differentiation by income class is usually motivated by distributional analysis, with income deciles as common classification criteria (e.g. Kim and Kim, 2003).¹⁷ There are two aspects to observe here. First, the classification of households with different composition into income classes requires some kind of equivalence scale, the choice of which will always be somewhat arbitrary. Second, income is not an ideal classification criterion because it is not exogenous. It may be the case that through the policy shock analysed, a household switches from, say, the tenth to the

¹⁶In the LABORSTA database at <http://laborsta.ilo.org/>, sectoral data are given in a occupational breakdown (Tables 1E). Educational data are only available at the country level.

¹⁷In Bassanini et al (1999) households are formed according to income classes as well, but then interpreted as skill groups.

ninth decile. Only in the unlikely case that all factor prices move in parallel remains the relative income position of all households unchanged. With relative factor prices changing, it is important to keep the exact interpretation of results with respect to income deciles in mind. We report the change for the decile of households that were in a certain decile *before* the reform. This does not necessarily mean that precisely these households are in the same decile *after* the reform.

Income types

A classification by skill type or occupation is at the same time a classification by income type. In addition, a classification by type of non-labour income might be useful in certain contexts. The distinction between labour, capital and other income (most prominently welfare benefits and old-age pensions) is particularly important for income tax reforms that treat different sorts of income differently (e.g. dual income taxes). In the case of transfers, the recipients of these income type are often a clearly separated group (pensioners or the unemployed). The recipients of labour and capital income, however, are not that clearly separate. Nevertheless there may be practical modelling reasons for forming a distinct household that collects capital income. Often micro data used for household decomposition contain unreliable or no information about capital income. Allocating all capital income to a hypothetical capitalist household may then be preferable to constructing some ad-hoc method of allocating it to the individual worker households (this is the route chosen by Arntz et al, 2008).

In developing countries with a large agricultural sector, by contrast, special attention is paid to the income from agricultural land ownership. A household decomposition by status of land-ownership is used in Boccanfuso and Savard (2008) for analysing the impact of the liberalisation of the groundnut sector in Senegal.

Wage level

Household classification by wage level is an option when we analyse policy measures such as a minimum wage or wage subsidies in the low-income segment. Certainly, the wage level will be correlated with classification criteria discussed above: skill level, occupation or sectoral employment. However, for policy measures that directly target a particular range of wages, these classification criteria may not be sufficient because they leave a large share of wage dispersion unexplained (Lee, 1999).

The problem of wage-targeted policy measures is that they can affect workers with slightly different wages in a qualitatively different way (e.g. those just above or just below the minimum wage). Addressing this problem in the setting of representative households requires us to use precisely the critical wage level as a demarcation criterion for households. This is somehow artificial, however, because a classification dictated by the policy measure in question will only accidentally be useful for labour supply or labour demand. Working with a set of completely disaggregated households (see Section 3.3) becomes preferable, because this does not force us to sacrifice other important dimensions of disaggregation (e.g. household composition). Examples for the representative-household approach to analysing minimum wages are Dixon and Rimmer (2003) and Dixon et al (2010).

Age

Differentiating households by age is mostly a feature of overlapping generations models and therefore not discussed in this chapter.

3.2.2 Sources of elasticity estimates

When we work with a larger number of representative households, the requirements for having elasticities available to calibrate those households increases in proportion. In Section 3.1, we had three elasticities (of hours of work with respect to the wage and with respect to other income, of participation with respect to the wage) to calibrate a single representative household. With n households, we require at least $3n$ elasticities.¹⁸

These elasticities will not normally be available from existing studies.¹⁹ Given the large number of possible classifications of households, it would be a coincidence if there existed a labour-supply estimation with exactly the classification needed. In addition, the results of any empirical study will normally be presented in a condensed form, as average elasticities for large subgroups of the population (e.g. singles and partners in couples) so that the reader does not have access to the original level of disaggregation.

¹⁸Or even more when we take account of the cross-elasticities in couple households.

¹⁹Evers et al (2008) is a comprehensive meta analysis of labour supply elasticities. A comparative estimation of elasticities in a number of OECD countries can be found in Bargain et al (2011b).

This means that working with a large number of representative households leaves us with two options: either assuming that elasticities are identical for large subsets of these households, or estimating the elasticities ourselves. Given that only in a minority of cases do CGE projects encompass resources for estimation, the availability of suitable elasticity estimates easily becomes a binding restraint. Arntz et al (2008), Cogneau and Robilliard (2008) and Bourguignon and Savard (2008) are examples of studies that contain labour supply estimates to be used in a combined micro-macro simulation framework.

3.2.3 Heterogeneity within aggregate households

Distinguishing between a number of representative households is likely to be insufficiently fine for a meaningful distributional analysis. As long as there is homogeneity *within* a representative household, the distributional impacts for large subsets of the population are exactly the same. In poverty analysis, a common indicator is a head-count index, which is defined as the share of the population that is below a poverty line defined in absolute or relative terms. With an indicator of this sort, poverty is bound to remain exactly as before if none of the representative households crosses the critical line. In contrast, as soon as at least one of the households crosses the line, this immediately leads to a large, discrete change in poverty. Normally we want to have continuous model reactions to continuous variation of the model parameters, and therefore such model behaviour is considered to be undesirable. Boccanfuso et al (2008) provide an extensive discussion of this issue.²⁰

As a possible extension, there are examples of models where a given number of representative households is combined with *within*-household inequality. The most important case is wage inequality according to a specific functional distribution (e.g. log-normal) within an aggregate household defined by skill type and household composition (e.g. Dervis et al, 1982, p. 526). The implicit assumptions of this approach are that relative wages do not change in the course of policy shocks and that the individuals within a representative household do not differ in any other respect than the wage. This restrictiveness of the heterogeneity-within-a-representative-household

²⁰There are other definitions of “poverty”, which can produce useful results even with representative households, see, e.g., Johnson and Dixon, 1999.

approach naturally leads to the follow-up step of dispensing with the concept of representative households altogether and switching to individual households instead. This is covered in the following section (3.3).

3.3 Microsimulation of labour supply

One important motivation for microsimulation, i.e. working with microdata of individual households and *not* aggregating them into representative households, is distributional analysis. When performing a distributional analysis, we are often interested in different dimensions: not only income classes, but also household composition, age, regional or sectoral employment. It is not possible to capture all these characteristics adequately in a representative-household approach. Any pre-defined classification of representative households works as a restriction on the available redistributional results. By contrast, in a microsimulation set-up, households can be classified and re-classified at the reporting stage so that classification can flexibly be adjusted to the research question.

A second motivation for microsimulation becomes relevant when we model policies that do not affect all individual households in the same way. In this case, the representative-households approach requires households to be classified according to the degree to which they are affected by a policy. A switch in attention to another policy measure can then mean a revision of the households classification. Again, working with a microsimulation set-up is much more flexible, because simulation is not affected by aggregation issues and aggregation takes place only after simulation results for individual households have been generated.

A third motivation for microsimulation originates in labour supply estimation. Empirical labour supply analysis is done at the micro level (see Section 3.3.1). So the natural outcome is labour supply elasticities that vary by household. The most straightforward approach is using the estimated parameters directly, as it is done in microsimulation, rather than aggregating them by simulating the joint reactions of the individual household represented in the aggregate. At best, the reaction of the calibrated representative household precisely mirrors the reactions of the micro units. However, without an explicit comparison, we can never be sure not to produce

some unwanted aggregation error.²¹

3.3.1 Functional approach to micro labour supply

Since van Soest (1995), labour supply modelling has almost exclusively been performed in a discrete-choice setting. Labour-supply econometricians pre-define a number of labour supply options, encompassing full-time and part-time work as well as non-participation. Then they estimate the parameters of some discrete-choice function (e.g. multinomial logit), which gives the probabilities for each of the options to be chosen. The attractiveness of the options depends on both leisure and after-tax income associated with a particular number of working hours. While leisure is fixed per option, the after-tax income varies across individuals because of differences in the hourly wage and in the local properties of the tax and transfer system. It is this variation that the approach exploits for estimating parameters of the utility function that in turn determine the discrete-choice probabilities.²² Usually, these parameters are estimated separately for different subsets of households (couples and singles), with a number of shift parameters for household characteristics (see Creedy and Kalb (2005a) for an introductory survey).

Before the advent of the discrete-choice approach, labour-supply estimations relied on a set of continuous choices, but ran into problems when confronted with non-linear budget constraints (see Hausman, 1995, for an overview). With a non-linear budget constraint, labour supply can react discontinuously to policy changes, even if the choice set is continuous. This causes problems both for estimation and for simulation. Therefore it has become standard to directly address the discontinuity issue by a discrete-choice set-up.

In the tradition of labour-supply estimations, discrete choice has mainly been implemented for modelling different hours-of-work options (including non-participation).

²¹For a general discussion of the relationship between micro and macro labour supply elasticities, see Keane and Rogerson (2011).

²²In the van Soest (1995) approach, differences in the choice behaviour of observationally identical households are rationalised by household-specific, stochastic preference shocks. A different route is chosen by Dagsvik and Strøm (2003) and Aaberge et al (1995), who base their estimation on a model of varying demand conditions in the labour market.

In the modelling of the labour market in low-income countries, discrete choice has additionally been used to capture the formal-informal and the employment-unemployment switch (Magnac, 1991; Bourguignon et al, 2005). However, compared to the hours-of-work choice, formal versus informal work and employment versus unemployment lend themselves less naturally to the interpretation of a *choice*. A more common interpretation is that these are reactions to changing demand conditions, where the role for choosing is limited. Maintaining the discrete choice framework in these contexts requires conceptualising the allocation of employment and formal work as governed by some kind of intrinsic propensity. If demand conditions change so that there is more opportunity for employment or formal work, the workers with the highest propensity will switch status. Bourguignon et al (2005), Bourguignon and Savard (2008) and Cogneau and Robilliard (2008) have advocated this approach in different contexts. It remains important to keep in mind that this excludes the interpretation that status switches of workers are the consequences of involuntary reactions to changes in demand conditions.

3.3.2 Counterfactual microsimulation

In a microsimulation setting, a counterfactual policy simulation means that the after-tax income for one or several of the discrete labour supply options changes (whereas the amount of leisure remains fixed per option). This affects the relative attractiveness of the different options, and thus the respective probabilities with which they will be chosen. There are two different methods of implementing the simulation, which we will discuss in turn.

First, if the discrete-choice function lends itself to generating explicit expressions for the choice probabilities, we can use these expressions directly in the model. The most important case is the multinomial logit model, in which the unobserved, idiosyncratic error terms, ε_i , that generate heterogeneity between observationally identical individuals are assumed to be extreme-value distributed. The logit approach produces expressions of the following form for the probabilities (p) to prefer

a particular option i over all other options j from the choice set

$$\begin{aligned} p_i &= p(\varepsilon_i > \varepsilon_j + U_j - U_i) \quad \forall j \neq i \\ &= \frac{\exp(U_i)}{\sum_j \exp(U_j)} \end{aligned}$$

where U_i is the estimated deterministic utility per option (McFadden, 1974). The probabilities p_i are in general strictly positive for all options. If we work with these probabilities, we interpret each individual household from the sample as representing a larger number of identical households. These households are distributed among the different labour supply options according to the probabilities. It is sometimes seen as a drawback of this approach that it does not exactly reproduce the aggregate labour supply behaviour of the sample population in the initial situation. In the sample each household chooses exactly one single option, with a probability of zero for all other options. This need not be a problem if one is inclined to adopt the sample perspective. Even if all choices are clear-cut within the sample, this need not be the case for the underlying population. However, if reproduction of the observed choices in the sample is a high priority, other simulation methods are called for.

The second approach, due to Duncan and Weeks (1998), which addresses the problem of the reproduction of the initial sample choices, relies on the drawing of random numbers. We draw a large number of sets of random numbers for the error terms ε_j of the discrete choice options.²³ Of these sets of random numbers, only those are kept that result in the choice observed in the sample. Drawing is repeated until a certain number (10 or 100) of fitting sets per household have been obtained. That is, sets where the random number of the option actually chosen is sufficiently large so that

$$\varepsilon_i > \varepsilon_j + U_j - U_i \quad \forall j \neq i. \quad (12)$$

In the counterfactual simulations the values of the U_j change. This means that the inequality (12) potentially does not hold any more for some sets of random numbers. We then obtain positive probabilities for other options than the one initially chosen. This evaluation of sets of random numbers cannot be done in a simultaneous system of equations, because of its non-continuity. Using a separate microsimulation module

²³These error terms must be consistent with the estimation, e.g. sets of extreme value distributed random numbers in the case of the logit model.

and iterating it with the main CGE model becomes necessary. In addition, the probability steps generated can only be as fine as implied by the number of random number sets (e.g., with 10 random numbers, we have 10 per-cent steps), which means that the reaction of the micro-households become discrete. This can may create problems in the convergence of the model modules when iterated.²⁴

As a variant to the Duncan and Weeks (1998) procedure, Bonin and Schneider (2006) have derived explicit switching probabilities for the multinomial logit model, conditional on the initial choice. This makes it possible to set up a simulation mechanism which, as in Duncan and Weeks (1998), reproduces the initial situation exactly, but without drawing random numbers and using them to evaluate the utility function under counterfactual conditions. Analytical switching probabilities have been applied in some microsimulation studies (e.g. Peichl et al, 2010), but not in a combined micro-macro model yet.

3.3.3 Linkage of the micro and macro modules

Once the working mechanisms of the micro module have been determined, the linkage between the micro and the macro part of the model enters centre stage. In principle, there are three options: (1) One-way linkage: first running one module, then the other. (2) Iteration in a soft link: run both modules alternatingly, until they converge. (3) Integrated model: combine both modules in a single model code and solve in one step. We will discuss these options in turn.²⁵

Depending on the shock analysed and the question asked, a *one-way linkage* can be either a bottom-up linkage (first micro, then macro; e.g. when simulating a change in the taxation of labour) or a top-down linkage (first macro, then micro; e.g. when simulating a trade shock and analysing its distributional consequences). Except for very special conditions, a one-way linkage will produce inconsistent results, because the reactions of the second module are not fed back into the first one. Let us consider a special case of a top-down linkage, where a one-way linkage is consistent indeed.

²⁴The model in Arntz et al (2008) contains an algorithm that identifies individual households which jump back and forth in the iterations, and then smoothes the reaction of these households.

²⁵Assessments of the different linkage options can also be found in Davies (2004) and Peichl (2009).

When labour supply of the individual households is completely inelastic, and all households have the same consumption and savings structure, the microsimulation part of the linkage consists solely in distributing the aggregate changes in factor income among households so that distribution analysis can be performed. There is no feed-back to be transmitted into the macro model. Obviously, these are very restrictive assumptions, which are not strictly valid in any realistic setting. In a somewhat looser sense, it has been argued that feedback effects can be expected to be small so that a one-way linkage provides a sufficient approximation. The problem with such an argument is, as always, that we do not know whether an irrelevance-of-feedback assumption is justified until we actually have performed the iteration and compared the results with and without feedback.

This is the reason why we believe the step to an iteration procedure between the two model parts is advisable. Iteration not only takes account of the feedback, but also forces the modeller to conduct an additional consistency check. If both modules are consistent, then the model has the potential of *convergence through iterations*. Convergence is, however, not assured. There is not much systematic knowledge about convergence-enforcing algorithms. In their “sequential recalibration” approach, Rausch and Rutherford (2010) propose using level information from the micro model to recalibrate a conventional utility function of a representative household in the macro module. The first-order characteristics of this function (marginal reaction to changes of the exogenous parameters) are determined by arbitrarily chosen elasticities. With respect to the final result, the value of these elasticities does not matter, because the share parameters of the function are re-calibrated in each iteration step. The elasticities can then be used to achieve a smooth convergence behaviour of the model.

Another idea is to not only use level information from the micro module for transfer to the macro module, but also first-order information, i.e. the (simulated) derivatives of the endogenous variables with respect to the exogenous variables. However, with a complex microsimulation module, generating this first-order information is not necessarily faster than doing more (because more slowly converging) iterations. This remains a question of trial and error. In Boeters (2010), working with first-order information turned out not to be successful. Instead, introducing damping factors (i.e. transferring not the complete reaction of the micro module to

the CGE model, but only a fraction) was the key to achieving convergence.

Finally, under certain conditions it is also possible to form an *integrated whole of the two modules*. This requires (a) that there be explicit expressions for the micro-reactions (e.g. use of the analytical switching probabilities for the logit model by Bonin and Schneider, 2006), (b) that they be sufficiently simple and (c) that there be not too many micro units. This constitutes a trade-off between ease of handling and detail in heterogeneity. Magnani and Mercenier (2009) strongly advocate an integrated approach, but to us a more liberal attitude seems advisable. As with other model characteristics, the choice of the iteration set-up should be guided by the questions asked and the type of results aspired to. If an extensive sensitivity analysis with many model runs is the goal, an integrated model design may be an enormous asset. If, on the other hand, integration means that the most plausible functional forms for microsimulation are no longer available, or that essential dimensions of heterogeneity are lost, a soft link with iterations is preferable.

3.3.4 Data consistency between micro and macro module

If we opt for an integrated model set-up, data consistency is an obvious issue. The general equilibrium approach requires that all markets clearance conditions (demand equals supply) and all income balances (expenditure equals income) hold. If this is not the case either within the data sets or between them, adjustment is required to get the model afloat.²⁶

Adjustment, however, is no less an important issue if we work with an iterated soft link or a one-way linkage. Even if the consistency requirement does not impose itself as strictly as with an integrated model, lack of consistency can cause serious problems in these cases as well.²⁷ In the soft-link case, failing convergence of the two

²⁶See Robilliard and Robinson (2003) for a general discussion of this point and for a specific proposal how to deal with inconsistencies.

²⁷This kind of problem is usually not well documented. If a consistency problem arises in a concrete study, the data is corrected, but this is normally not an issue focused on. If there is a problem which happens to escape the modellers, inconsistent results are reported, but it is almost impossible for a reader or reviewer to discover this. So all evidence we have are discussions among active modellers who alerted us to possible inconsistencies at this point.

model parts can result at the technical level. At the level of the interpretation of results, both in the soft-link and the one-way-linkage cases spurious effects can be the consequence. Data discrepancies that are somehow transferred from one module to the other (and thus implicitly eliminated) then seem to be scenario results (see Bourguignon and Savard, 2008, for further discussion).

The two most common data inconsistency issues are (1) inconsistency at the level of the individual income balances (income does not equal expenditure plus savings) and (2) inconsistency between micro and macro information about consumption shares, savings rates, income shares (labour income, capital income, transfers) and/or skill (or other labour) types.

In the first case, data must be adjusted at the level of the individual households so that the individual income balances hold. In the second case, in addition to adjustment at the individual level, there is a second option. We can leave the individual household data as they are, but introduce some aggregate entity (“residual household”), which makes up for the discrepancy. For concreteness, consider the case in which the share of capital income in the household data set is smaller than in the national accounts. Then we can either shift some income from labour to capital at the level of the individual households. Or we can inflate the micro module to the point where macro labour income is met and introduce a residual capital income recipient for the rest. This approach has been chosen, e.g., in Arntz et al (2008).

Data adjustment is always highly context dependent so that it is not possible to derive general rules to be followed. This remains a task where expert judgement is required, including profound knowledge both of the background of the data, the model mechanisms and the aim of the modelling project. Some – maybe obvious but still rather vague – guidelines for data adjustments are the following:

- Keep the data you consider most reliable and adjust the rest. Normally, it is reasonable to assume that national account data have been prepared in a way that warrants consistency according to internationally agreed principles. So we will override national account information by information from micro datasets only if there is a strong reason.
- Adjust the data so that the implied changes interfere as little as possible with

the aim of the study. Try to avoid adjustment of variables that are key to the policy shock analysed or to the interpretation of the results.

- Use simple methods of adjustment (straight multipliers) as long as there are no strong reasons for complex mechanisms (cross entropy estimation as in Robilliard and Robinson, 2003).²⁸
- Adjust large quantities rather than small ones, in order to avoid large relative changes.

Needless to say that these guidelines can easily interfere with and contradict one another so that, ultimately, they cannot replace expert judgement.

3.4 Household-specific expenditure structures

Our main interest is in aspects of household disaggregation that are directly related to labour market outcomes. However, households may also differ in other respects, and this must be kept in mind when working with a labour supply module, either at the level of several representative households or at the level of individual households in microsimulation.

3.4.1 Household-specific savings

If micro data sets contain information about income spending, we can work with household-specific savings rates in the CGE set-up, either with representative or individual households. In a static setting, this does not cause any additional complications. The individual savings must only be aggregated to macroeconomic savings. If the model is set up as a (recursively) dynamic one, however, household-specific capital stocks must be traced through time. (See Dixon and Rimmer (1995) for an example with household-specific savings rates.)

²⁸In our initial example of the share of capital income being too low in the household survey, the Robilliard and Robinson (2003) procedure would imply to give the household with high capital income a higher aggregation weight. Just shifting a constant share of labour income to capital income is a reasonable alternative.

The issue becomes considerably more involved once we consider involuntary unemployment in a dynamic setting. As savings (and therefore the dynamically updated capital stock) are dependent on the labour market status (employed or unemployed), current wealth becomes a function of the whole labour market history of the respective household. If unemployment is a stochastic phenomenon with non-zero probabilities for both employment and unemployment, this means that we must keep track of all possible labour market histories. With t periods, this means 2^t possible states per individual, and even 4^t for couples, when unemployment can vary independently for both partners. For a larger number of periods, this quickly becomes intractable.

There are two possible approaches to this problem. The first has been intensively explored in forward-looking dynamic modelling, namely the conditions under which it is possible to pool savings of the individual households and derive an aggregated savings function. Usually these conditions are rather restrictive and therefore not particularly appealing (Benhabib and Bull, 1983; Krebs and Scheffel, 2010). The second option is stochastic dynamic microsimulation. Here, we give up the ambition to track all possible labour market histories and merely assign each individual a labour market state, depending on the drawing of random numbers. Examples of dynamic microsimulation, even if not in a CGE context, are Merz (1993) and Heckman et al (1998).

3.4.2 Household-specific consumption structures

Another dimension where household-specific information on income spending can be used is the consumption structure. This information is often available in micro data sets. A complication implied is that household-specific consumption structures lead to household-specific real wages, which in principle matters when deriving labour supply. However, as long as household consumption structures do not differ dramatically, this effect is unlikely to be large. The error we make by working with a household-independent consumer price index seems tolerable, at least for high-income countries (see Boeters et al (2010) for income-related consumption structures in the context of a reform of the value added tax).

For low-income countries, however, household-specific consumption structures

may be relevant, in particular in the context of distributional analysis. Expenditure structures vary strongly between rural and urban households and between urban poor and urban non-poor (see, e.g. Boccanfuso and Savard, 2011).

3.5 Dynamic labour supply

Most of the labour market aspects discussed in this chapter are independent of whether we work in a static or dynamic model. In dynamic models, both recursive and intertemporal, we are faced with additional complications due to the secular trend in labour productivity. Dynamic models are usually calibrated to a steady increase in productivity and real wages over time. Given the labour supply mechanisms implemented in the model, this will in general have long-term effects on hours of work, participation and unemployment. These must be carefully checked for plausibility.

One of the main difficulties in dynamic labour supply theory is the consistency of short-term and long-term reactions. In cross-sections, we normally observe labour supply elasticities that are small, but significantly positive, both at the intensive and extensive margins. Taken at face value, this would mean a secular increase in labour supply, following productivity and real wage developments. But this is not what we actually observe. Long-term labour supply is almost perfectly stable, and at the intensive dimension, labour supply rather falls than rises with the secular increase in productivity (see Prescott, 1986; Ngai and Pissarides, 2008; Ramey and Francis, 2009, where modern treatments tend to distinguish work, leisure and home production, rather than only work and leisure). One radical answer to this dilemma stems from dynamic modelling, where the use of functional forms is restricted by the requirement to derive explicit expressions. In this context authors often choose a Cobb-Douglas specification of utility, because this produces exactly offsetting income and substitution effects of wage changes (see King et al, 1988, and Kimball and Shapiro, 2008). This effectively means giving up the ambition to calibrate utility functions to empirical short-run labour supply elasticities, as described in Section 3.1. When working with recursively dynamic models, however, we are less restricted in the choice of functional forms and can search for a specification that allows for labour supply effects in a policy counterfactual without implying a long-term labour supply trend.

A suitable framework for counteracting higher labour supply in the long run is a model of household production (e.g. Benhabib et al, 1991). By introducing productivity of work at home, which follows the productivity change in market work, we can let opportunity costs of working increase over time. When we do not have an explicit representation of household production in our model (as in the specification exposed in Section 3.1), we can resort to a shortcut that captures the essence of the home production approach.

For hours-of-work, this shortcut consists in introducing compensatory efficiency increases for leisure, α , in the utility function of the workers, eq. (1), when calibrating labour supply (Section 3.1.1):

$$U_e = \left[\theta_C \left(\frac{C_D}{\bar{C}_D} \right)^{\frac{\sigma-1}{\sigma}} + (1 - \theta_C) \left(\frac{\alpha F}{\bar{F}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

α may then be adjusted over time in the baseline calibration of the model so that the number of supplied hours remains constant (this makes α increasing over time). Similarly, in the calibration of participation (see Section 3.1.4), the distribution of fixed costs of taking up work may be adjusted so as to prevent an endogenous trend in participation. The calibration is repeated for each year, taking the changes in the endogenous model parameters in eq. (8) into account and holding the labour supply elasticity at the extensive margin, η_{Nw} , constant.²⁹ An exogenously given time path of participation rates may be accommodated by adjusting $\bar{N}$ and $\bar{U}_l$ over time in eq. (9). A possible interpretation of the resulting shifts in the distribution of fixed costs is a shift in commuting costs due to a changing valuation of time caused by productivity increases.

4 Labour demand

In comparison to the other main themes covered in this chapter – labour supply and labour market coordination – labour demand offers the least conceptual choices. In CGE models, labour demand is derived from sectoral production functions, which are almost exclusively set up as nested CES functions. The challenge in CGE models

²⁹Otherwise, the secular increase in U_e would cause an ever-increasing elasticity of labour supply.

with a more detailed labour market module is to adjust the production functions in a way that responds to the requirements of the other model components (in particular labour differentiation initiated on the supply side) and does justice to our empirical knowledge about labour demand elasticities.

We have seen in Section 3 that there are a large number of motives for distinguishing different types of labour, often originating from distributional analysis. The basic question is whether these types of labour also need to be distinguished on the labour demand side. It is important to realise that we do not have to do so. Any labour differentiation on the supply side can be made undone on the labour demand side by simply adding up all labour (e.g. male and female labour, labour of different age categories) to a uniform aggregate.³⁰ Even if wages diverge empirically, this need not prevent homogeneous aggregation. We can assume that workers differ by constant labour efficiency factors, which allows us to add up the individually supplied, *efficiency-weighted* hours of work, even if not the raw hours of work. Constant labour efficiency factors mean that wages within an additive labour demand category will change exactly in proportion. The critical question is thus: Do we think that proportional wage movements describe a particular labour market segment sufficiently well, or are there reasons to believe that relative wages of different sub-categories of labour will systematically respond to changes in relative supply or demand? Only in the latter case must we think about how to represent different types of labour in the production function.

4.1 Functional implementation of labour demand³¹

Modelling labour demand for a specific set of labour types, we start from a given estimation of labour demand elasticities. The first decision to be taken is what kind of elasticities we want to work with. Labour demand estimations can be reported either in terms of labour demand elasticities or of substitution elasticities. Both are interconnected, but the relation between them can only be made explicit if we know the input value shares (see, e.g. Hamermesh, 1993, Ch. 2). This creates two follow-up

³⁰E.g. Borjas et al (2011) find that similarly skilled immigrants and natives are perfect substitutes in production.

³¹This paragraph uses material from Boeters and Feil (2009)

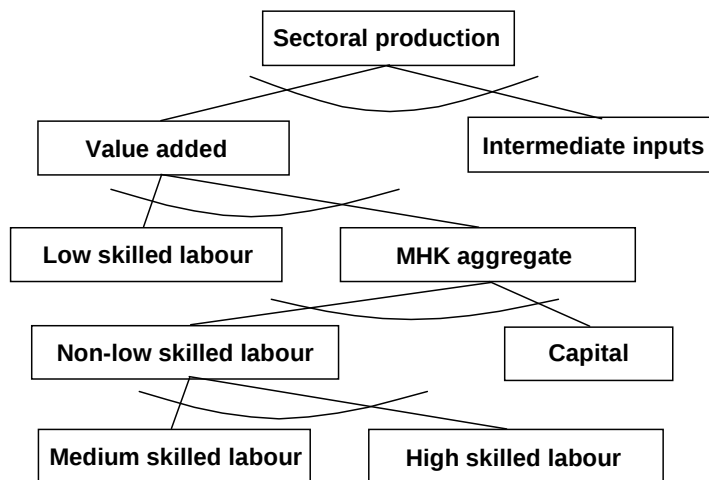
problems: (1) Many labour demand estimations do not report value shares so that we cannot compute substitution elasticities from demand elasticities or vice versa. (2) In general the value shares in the estimation will deviate from the value shares in the data sets used to set up the CGE model. So we either have to use the type of elasticity we can extract from the empirical study, or – if the empirical study leaves the choice open – we must decide which type of elasticity we treat as more fundamental. These fundamental elasticities are then taken from the study, and the other type of elasticity is derived, using the value shares from the model. At the end of Section 4.3, we motivate the choice of labour demand elasticities as the empirical basis for calibrating CGE models.

Take as a concrete example Falk and Koebel (1997), who estimate labour demand elasticities for five production sectors in Germany, using a Translog production function with five inputs: three skill classes of labour (low, medium and high skilled), capital and materials. The result is a full matrix of estimated cross-price elasticities. Assume that we try to approach these empirical results with a conventional, separable, nested CES function as in Figure 1. Value added is split into the contributions of low-skilled labour and an aggregate of non-low-skilled labour and capital in the upper-level nest. At a second level, capital is split off, and finally, medium and high-skilled labour are separated. Such a set-up can account for stylised patterns of substitution elasticities, in particular the well known capital-skill complementarity (Fallon and Layard, 1975), but it is not flexible enough to represent full matrices of estimated cross price demand elasticities.

In the production structure of Figure 1, we end up with four free parameters (elasticities of substitution at various levels of the production tree) to be calibrated. However, a fully flexible structure, such as the one estimated by Falk and Koebel (1997), features at least 10 independent elasticities of substitution: a 5×5 matrix where 10 elements are mirror images of the opposite side and 5 elements are linearly dependent on the other entries in the same row or column.

The NNCES (non-separable, nested CES) approach to production function calibration (Pollak and Wales, 1987; Perroni and Rutherford, 1995; Perroni and Rutherford, 1998) increases the flexibility of the nested CES framework through an extension to more generic forms. Other flexible forms known from econometric tradition (Translog or Diewert) can locally represent arbitrary production or cost functions as

Figure 1: Separable, nested CES function



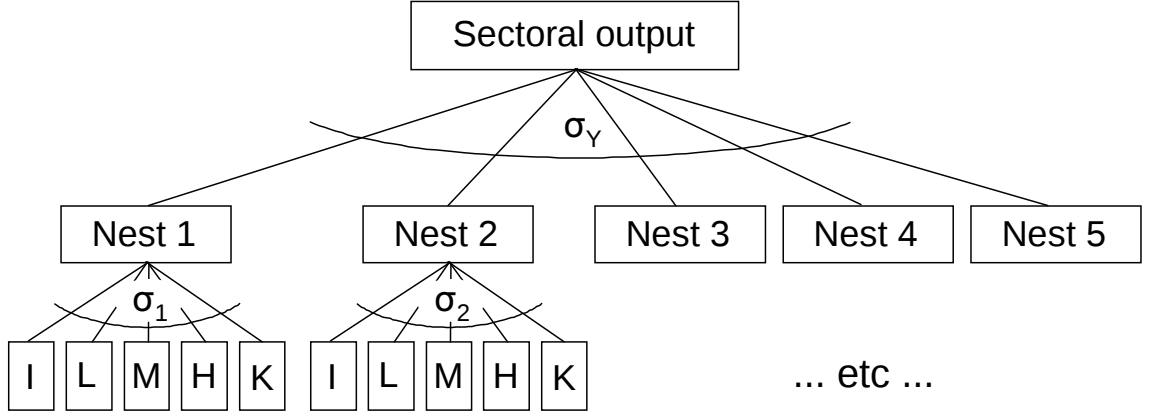
well, but they typically do not exhibit global regularity (the corresponding cost function must be non-decreasing and concave in prices). This can cause computational problems in CGE models (Perroni and Rutherford, 1995).³²

The basic idea behind the NNCES function is that each factor of production may enter the production function at more than one single place (therefore “non-separable”). A typical set-up is depicted in Figure 2: Sectoral output is decomposed into five sub-nests, each of which then in turn contains input from all factors: intermediate inputs (“I”), the three skill types (“L”, “M”, “H”) and capital (“K”). Flexibility is increased not only by a larger number of elasticity parameters (six), but also because the split of each production factor into the individual sub-nests can be chosen freely.

This is now the exact opposition of the problem we had before: Instead of too few parameters, we have too many. We have 26 free parameters at hand (six elasticities and four free share parameters for each factor) for producing a match between the model and the 10 exogenous elasticities. To resolve the resulting indeterminacy, it has been proposed to restrict certain elasticities to zero or one (Pollak and Wales,

³²Nevertheless, other functional forms can be found in the literature. E.g., Dixon et al (1992, p. 133-137) work with an explicit Translog function for labour demand. Apparently, practical problems with calculation are less severe than the theoretical arguments in Perroni and Rutherford (1995) suggest.

Figure 2: NNCES set-up of production function



1987) or to add a penalty function. A plausible goal for the calibration is to limit the dispersion of input factors across several nests and to restrict the absolute value of the elasticities of substitution. With a penalty function of this sort, the approach can be expressed as follows:

$$\begin{aligned}
 & \max \sum_{n,i} (\theta_n^i)^2 - (\sigma_Y)^2 - \sum_n (\sigma_n)^2 \\
 \text{s.t.} \quad & \eta_{ij}/\theta_j^Y = \sigma_Y + \sum_n (\sigma_n - \sigma_Y) \frac{\theta_n^i \theta_n^j}{\theta_n^Y} \\
 & \theta_n^Y = \sum_i \theta_n^i \theta_n^i \\
 & 1 = \sum_n \theta_n^i
 \end{aligned}$$

where i and j are indexes for the factors of production and n is an index for the nests at the intermediate level (“Nest 1” etc. in Figure 2). The θ_n^i are the shares of the individual nests in the total amount of factor i , θ_n^Y is the share of the respective nest in total sectoral output. The own- and cross-price elasticities, η_{ij} , as well as the aggregate value shares, θ_i^Y , are exogenous to the calibration. σ_Y , the σ_n ’s and the θ_n^i ’s must be determined by minimising the penalty function. An implementation of this approach for Germany can be found in Boeters and Feil (2009).

4.2 Dimensions of labour demand heterogeneity

In this section we review the multiple dimensions of labour heterogeneity that have led to specific labour demand categories in applied modelling. Partly, these dimensions overlap with those of labour supply, discussed in Section 3.2.1. However, on the labour demand side, different aspects are relevant. In particular, it is more challenging to justify a particular labour split. In the following list, we report the dimensions of labour demand heterogeneity encountered in the literature, and we check to which extent they can be backed up with empirical estimates of labour demand elasticities. After discussing the dimensions one-by-one, we turn to additional issues that arise once we combine several dimensions in a single production function (Section 4.3).

Skill type

The most prominent decomposition of labour is along the skill (or qualification) dimension. The distinction of two skill types (skilled and unskilled or qualified and non-qualified) has a long tradition in the discussion about raising wage differentials and skill-biased technological change (see, e.g. Berman et al, 1998). This split can be found in many CGE models, even if they do not focus on labour market issues, simply because it is implemented in the GTAP data set (Dimaranan and Narayanan, 2008). With two types of labour, the simplest way of value added modelling is a one-level structure, where capital services and the two types of labour are included with a single elasticity of substitution. An example is the GTAP model (Hertel et al, 2008), where sector-specific elasticity estimates from Jomini et al (1994) are used. Other papers with a two-skills split and a similar specification are Böhringer et al (2005) for Germany, Boccanfuso and Savard (2008) for Senegal and the global WorldScan model (Lejour et al, 2006).

A potential advantage of labour split into the two standard skill types (i.e. low- and high-skilled, with college degree, or a national analogue, as the demarcation line) is the availability of advanced labour demand estimations. We can draw upon an established literature on the substitution possibilities between low skilled labour, high skilled labour and capital (e.g. Griliches, 1969; Fallon and Layard, 1975; Krusell et al, 2000). However, these results are not often used for the specification of existing CGE models. Exceptions are Agénor et al (2003), if only qualitatively, and Rojas-Romagosa (2010), who revises the value-added modelling in WorldScan.

Labour splits into more than two skill categories can be regularly found in existing models, but they are more difficult to back up with elasticity estimates from the literature. Löfgren (2001) uses four skill categories, without any substitution possibilities (Leontief specification). Maisonnave et al (2009) have three categories, with a Cobb-Douglas specification. Models that use an empirical specification are Boeters and Feil (2009) (with the NNCES approach described in Section 4.1) and MIMIC (Graafland et al, 2001, p.111). Interestingly, in the estimations used for the specification of MIMIC, capital-skill complementarity is rejected for the Netherlands.

Occupation

The classification by occupation is a close substitute for skill. The choice between these two options may be driven by data availability, in particular when internationally comparable data are needed (Boeters and van Leeuwen, 2010). In some models, occupation and skill characteristics are combined in the labour segmentation (e.g. Carneiro and Arbache, 2003; Colombo, 2008). Giesecke et al (2011) present an ambitious set-up for the Vietnamese economy, which cross-classifies skill (qualification) and occupation. Labour demand of firms is formulated primarily in terms of occupations, which, in turn, are decomposed by skill. Households, on the other hand, are primarily defined by skill and supply occupation-specific labour according to the transformation mechanism presented in Section 3.1.5.

Full-time and part-time labour

Hutton and Ruocco (1999) are concerned with the prevalence of part-time work among women. This motivates them to distinguish between part-time and full-time work as imperfect substitutes in production. They report no estimation results for the elasticity of substitution between these labour types and use a Cobb-Douglas structure in their model. There is some empirical work trying to determine the substitutability between workers and hours (for an overview of the earlier literature, see Hamermesh, 1993, pp. 127-134), but it has not been used in a CGE context yet.

Formal and informal labour

The distinction between formal and informal work can be found in many CGE models. There are different motivations for making this explicit in a model: productivity differentials between the formal and informal sector, wage differentials between workers in the two sectors, and a differential treatment in taxation (see Section 5.5 for a detailed discussion). Problems of labour demand modelling do, however, not

arise in most cases, because formal and informal labour is used in separate sectors. The question of substitutability between formal and informal work then shifts to the substitutability between formally and informally produced goods and services in consumption. Only models with highly aggregated production sectors have both formal and informal labour in the same production function. An example is MIMIC (Graafland et al, 2001, p. 110), using estimations of Baartmans et al (1986) for calibration.

Rural and urban labour

Similar to the formal/informal distinction, the rural/urban divide is mostly a sectoral one, with different sectors demanding labour in the rural and urban area. An exception is Hendy and Zaki (2010), who allow for both rural and urban labour as input in each sector. The elasticity of substitution is chosen without reference to an empirical study.

Gender

The gender decomposition of labour in CGE models has become more widespread, driven by the rising interest in the gender dimension of inequality and the recognition of the special role women play for economic development. The gender decomposition has been introduced to the literature by Fontana and Wood (2000), followed by Fofana et al (2003), Fontana (2004), Colombo (2008) and Hendy and Zaki (2010).³³ As with other dimensions of decomposition, the empirical foundation is a problem. Fontana and Wood (2000) do not base their specification on empirical estimations, but argue qualitatively for relatively low elasticities “to reflect the rigidity of gender roles” (p.1179). As those roles are highly varying across countries and in many cases also over time, it is not likely that deep substitution parameters can be identified through econometric techniques.

Ethnicity

A more recent and less common decomposition is by ethnic group, where the specific delineation of groups is country specific. Maisonnave et al (2009) use an ethnic decomposition into African, Coloured, Indian and White in a model of the South African economy. In the case of Israel, Flaig et al (2011) use the ethnic/nationality

³³An early predecessor is Dixon et al (1978), who analyse the economic consequences of an exogenous increase of relative female wages.

categories Jewish and non-Jewish, Israeli and non-Israeli, Palestinians and foreigners. In both cases, the substitution elasticities are chosen ad hoc. Maisonnave et al (2009) assume that there is no substitution possible (Leontief specification). The elasticities in Flaig et al (2011) have been set relative to the GTAP skilled/unskilled substitution elasticity, based on expert judgement.

4.3 Labour heterogeneity in several dimensions

In Section 4.2, we have presented possible dimensions of labour demand separately. Many of the papers discussed do not only deal with one dimension, but with labour heterogeneity in two or more dimensions. Once we have more than a single dimension, we must think about the structure of the production function. The following options have been used in the literature:

- Cross-classifying all dimensions, deleting irrelevant combinations (if necessary) and organising all categories of labour in a single CES nest. This approach has been chosen by Colombo (2008), who has a $2 \times 2 \times 2$ classification of skill, occupation (self-employed, waged workers) and gender. Carneiro and Arbach (2003) use a non-exhaustive cross-classification of formal/informal, three skill types, rural/urban and civil servants in a one-level nesting.
- There are a considerable number of examples where multi-level CES structures are used for different dimensions of heterogeneity. Flaig et al (2011) combine two skill types at the upper level with ethnicity/nationality at the lower level. Hendy and Zaki (2010) use a CES production function with three levels: urban and rural workers at the upper level, male and female work at the intermediate level, and skilled and unskilled labour at the lowest level. Maisonnave et al (2009) decompose three skill types at the upper level (Cobb-Douglas specification), and four ethnicities at the lower level (Leontief specification). The dimensions in Dixon and Rimmer, 2003 are occupation and two aspects of entitlement in a low-wage policy proposal, and in Dixon et al (2011) occupation, legal status and place of birth.

Apart from the problem of finding labour demand parameters at all (in most of the studies reviewed, elasticities of substitution are merely guessed), with several

dimensions of heterogeneity, we are faced with the problem of the nesting hierarchy. Which types of labour should be grouped into a nest at a higher level, and which at a lower level? Traditionally, a loose convention has established that factors of production that are close complements are grouped together at a low level of the nesting, whereas factors that are good substitutes are combined at a higher level (i.e. higher levels of the nesting have higher values of the elasticity of substitution than lower levels). There is no empirical reason for such a convention. Ultimately, the only criterion that counts is which nesting fits the data best. In the case of KLEM estimations, there have been studies that compare different CES nestings with regard to their empirical performance (e.g. van der Werf, 2008). We do not know, however, of any comparable study in the field of labour demand estimations. Here, the estimation techniques are rather dominated by flexible functions, which are a candidate for direct implementation in CGE models (see Section 4.1).³⁴

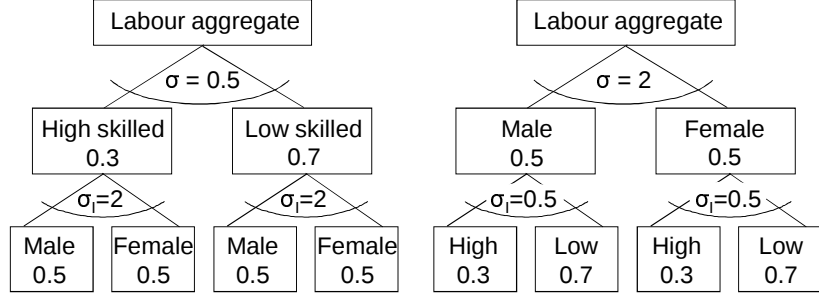
When deciding for a multi-layer CES structure, modellers are left with the question of how to set up the nesting. Let us take an optimistic case and assume that some estimates (or guesstimates) of elasticities of substitution between different varieties in each dimension are available. It is important to realise that the same elasticities of substitution, if implemented in a different nesting structure, can lead to different demand elasticities. In a two-level nesting structure as in Figure 3, the own-price elasticity of factor of production i is

$$\eta_{x_i p_i} = -\sigma_I (1 - \theta_{iI}) - \sigma (\theta_{iI} - \theta_i), \quad (13)$$

where σ and σ_I are the elasticities of substitution at the upper and lower level, respectively, θ_i is the value share of factor i , and θ_{iI} is the value share of factor i in its relevant sub-nest. For concreteness, let us distinguish between high/low skilled and male/female workers. The value share of high skilled work is 0.3, and within each skill class, male and female workers have the same share of 0.5. Elasticities of substitution are 0.5 between skill groups and 2 between genders. The two nesting options are given in Figure 3. For the own-price demand elasticity for low-skilled

³⁴A recent example for a flexible-form labour demand estimation with several dimensions of labour heterogeneity (skill, age and type of employment contract) is Bargain et al (2011a).

Figure 3: Two-level nesting options



male work, we obtain

$$\begin{aligned}\eta_{LM} &= -2 \times (1 - 0.5) - 0.5 \times (0.5 - 0.35) = -1.075 \\ \eta_{LM} &= -0.5 \times (1 - 0.7) - 2 \times (0.7 - 0.35) = -0.85\end{aligned}$$

for the left and right panel of Figure 3 respectively. This is a significant difference.

This is another example of the general phenomenon discussed in Section 4.1: There is no one-to-one relationship between elasticities of substitution and elasticities of demand. We need to answer the question of which of these parameters – elasticities of substitution or elasticities of demand – we want to consider as more fundamental. General equilibrium economists tend to focus on the elasticities of substitution. Researchers working on the substitution possibilities between capital, high and low skilled labour (e.g. Krusell et al, 2000) have tried to identify the relevant elasticities of substitution, not elasticities of factor demand. On the other hand, labour market economists are often interested in labour demand elasticities as an ingredient to a partial labour market model (see Peichl and Siegloch, 2010). In principle, this question of which elasticity is more fundamental could be approached empirically. What is needed are comparative labour demand estimations with different data sets (implying varying value shares). This would allow us to address the question: Which estimated parameters are more stable across data sets, elasticities of substitution or demand elasticities? However, we do not know of any empirical study approaching the issue from this angle.

In the absence of an empirical guideline, we tend to favour elasticities of demand as the more fundamental ones. After all, demand elasticities are the elasticities that are most directly linked to empirically relevant results (e.g. “By how much must

the wage fall to induce absorption of a certain change in labour supply?”). This is a point in favour of treating labour demand elasticities as fundamental and calibrating elasticities of substitution (conditional on the given value shares) so that the demand elasticities are met (as in Section 4.1 in the case of the NNCES function).

5 Labour market coordination

5.1 Scope of the market

Determining the scope of a labour market in a CGE model, i.e. delineating the individuals under a uniform labour market regime, touches upon a number of issues that are hardly ever explicitly discussed in economics texts. This is because they tend to fall in-between the major fields of theoretical and empirical work. In theory papers, it is customary to merely *postulate* how many and which markets are to be distinguished. Empirical labour market economists, in contrast, take multi-dimensional heterogeneity at face value. This becomes apparent in multiple regression analysis, the very point of which is to filter out the effects of as many dimensions of heterogeneity as possible. The question of how to delineate labour markets does not arise naturally in either the purely theoretical or the purely empirical context. In applied modelling, however, this question imposes itself on the model builder. It is often answered in an implicit way, which remains silent about other options and about the consequences of a particular choice. In this section, we want to make the relevant issues more explicit.

By the criteria reviewed in Sections 3 and 4, both labour supply and labour demand are potentially differentiated in several dimensions. The follow-up question is about whether, and to which extent, this differentiation is to be transferred to labour market coordination. There are two extreme options to decide this question: (1) a large number of very narrowly defined labour markets with an independent wage rate in each of them, or (2) a single labour market, in which the sum of all supply and demand is coordinated. In practical modelling, labour market differentiation is most often driven by the supply rather than the demand side. Model builders are motivated, either by distributional concerns or by the design of the policy instrument, to distinguish between different types of workers (e.g. by gender, age or skill

level) and impose this distinction on labour demand and labour market coordination as well.

It is a basic point to be kept in mind, however, that we do not have to differentiate between labour markets simply because we have decided to distinguish between different types of labour. Treating different types of labour as homogeneous inputs and aggregating them in a single market with a uniform wage is a perfectly viable option. Naturally, wage differentials between individuals are an empirical fact, which is, *prima facie*, not compatible with a homogeneous market. However, there are at least two simple mechanisms – both discussed in a CGE context by Boadway and Treddenick (1978) already – that allow us to work with a uniform market model in spite of existing wage differentials. The first of these mechanisms is efficiency weighting. Individuals with higher hourly wages are assumed to supply more effective units of labour per hour than low-wage individuals. Weighting with individual-specific efficiency factors derived from the wages, we can simply add up labour supply of heterogeneous individuals. Assuming efficiency factors to be constant implies that wages of different individuals in one market move precisely in proportion, leaving relative wages unchanged.

The second mechanism for accommodating wage differences without separating markets are compensating wage differentials, typically applied to sectoral wages. Jobs in different sectors are associated with varying non-pecuniary conditions such as riskiness of the job or agreeableness of the work sphere. Under inter-sectoral mobility of workers, differences in observed wages can be rationalised by assuming that lower wages are compensated by non-pecuniary benefits so that, taken as a package, job opportunities are equally attractive for workers across sectors.³⁵ Simple and well-established as this idea is, it may create follow-up problems in modelling: (1) Non-pecuniary benefits contribute to household utility and must be accounted for when performing a welfare analysis. (2) As a contribution to household utility, non-pecuniary benefits might have an income effect on labour supply.³⁶ (3) If there

³⁵It is a considerable empirical challenge in itself to determine what share of observed sectoral wage differentials is attributable to a different composition of the work force, and what share remains as a pure sectoral wage differential. It has repeatedly been shown that neither of these shares is negligible (see, e.g. Slichter, 1950; Krueger and Summers, 1988; Genre et al, 2011).

³⁶In Böhringer et al (2005) this effect is included in the following way: Labour supply responds,

is involuntary unemployment, this is an additional factor contributing to the relative attractiveness of sectors, see Section 5.5.³⁷

The crucial empirical question for labour market differentiation is whether wages of different groups of workers move in parallel (with appropriate constant adjustment factors) or not. This question arises almost exclusively in the modelling context, once we want to determine whether market segmentation is appropriate or not. So it is no surprise that it has not been addressed in existing empirical labour market studies, with the exception of the wage differential between skilled and unskilled workers. For other potential labour market segmentations original empirical research is required. A case in point is gender-specific labour markets. There is a growing literature interested in gender-specific labour-market consequences of various policy shocks (e.g. Fontana and Wood, 2000, Fofana et al, 2003, Hendy and Zaki, 2010). These studies assume that there are separated labour markets for men and women, but they do not provide empirical evidence supporting an independent movement of gender-specific wages. If such evidence exists, it could considerably strengthen the modelling set-up.

Another basic point to be kept in mind is that homogenising forces on *one* side of the market, either supply or demand, are sufficient for the constitution of a homogeneous labour market. If labour is perfectly *transformable* between different categories (see Section 3.1.5), this will create a uniform labour market, because any potential wage differential would provoke labour supply adjustments that cause wages to equalise. The most relevant case is sectoral labour mobility. In almost all other dimensions of labour supply heterogeneity (gender, age, skill), transformation options are limited, if existent at all. – The mirror image of a labour-market homogenising effect can arise at the labour demand side. Once two types of labour are perfectly substitutable in production, any difference in the wages will disappear. While it is standard in labour demand modelling to assume imperfect substitutabil-

not to the wage *per se*, but to the value of a job. The value of a job, in turn, is equal across sectors and composed of three components: wage, non-pecuniary benefits and unemployment risk.

³⁷The approach of Devarajan and Rodrik (1991) remains somewhere half-way in this compensating differential story. They assume sector-specific efficiency factors which are calibrated to match intersectoral wage differentials. However, they do not address the question of what prevents all workers from switching to the highest paying sector.

ity of skill types, the case with other dimensions of labour heterogeneity (gender, age, hours of work) is much more complex. Clearly, the burden of proof is with those that claim imperfect transformability in these cases (see Section 4.2).

A special constellation to be mentioned is one-way transformation or substitution, which causes asymmetry between labour market segments. An example is the dual labour market model of Harris and Todaro (1970), discussed in Section 5.5, where workers who do not find a job in the formal sector can always retreat to the informal sector, but not in the opposite direction. Another asymmetry pattern is plausible in the demand for skills. There are many low-skilled tasks that can without a problem be taken over by high-skilled workers, if necessary – but not inversely.

5.2 Wage forming mechanism

The vast majority of CGE models have worked with one or both of the two extreme assumptions about wage formation: flexible, market clearing wages and institutionally fixed wages. In the market-clearing set-up, competition on both sides of the labour market is assumed so that wages equilibrate demand and supply and there is no involuntary unemployment. In contrast, in the fixed-wages set-up, wages do not respond to changes in supply or demand conditions. Labour supply and demand are determined given a fixed level of wages, and the resulting difference is interpreted as involuntary unemployment.

Simple as it seems, the fixed-wages set-up is not without difficulties in a CGE context. Actual wage contracts are usually fixed in nominal terms. In general, however, nominal prices are undetermined in CGE models, since only relative prices have significance. If we nevertheless fix a nominal wage in our CGE model, what we essentially do is fixing a real wage in terms of the numéraire of the model.³⁸ We consider it to be good modelling practice to make the fixation of real wages explicit and not implicit through the choice of a numéraire.³⁹ If the domestic consumer price index (CPI) is chosen to be the numéraire, the fixation of nominal wages may well

³⁸We have come across CGE papers where the authors apparently confused fixing the wage as an institutional feature of the labour market and using it as the numéraire of the model.

³⁹The numéraire choice is often not reported in the model description, just because it is assumed to be arbitrary. In this case the fixation of nominal wages becomes a black-box phenomenon.

be plausible, because explicit or implicit indexation of wages to consumer prices is a common phenomenon.⁴⁰ However, it is not in all cases that the CPI acts as numéraire. In many applied one-country models of small open economies this role is taken by the export and import prices. Then fixing a nominal wage means that the wage is indexed to foreign prices, which is implausible, except if it is an explicit institutional feature of the economy modelled.

Fixing real wages is the simplest possible way of accounting for labour market rigidities and involuntary unemployment. However, being simple means being restrictive as well. Even if we take it for granted that there are institutions fixing the real wage, this fixation will hardly be absolute. At least in the medium and long run, wages do adjust to labour market conditions, even if subject to longer lasting contracts. So the next plausible step is to think about *which* labour market conditions will drive the wage adjustment and *how* the adjustment takes place.

In general, any in-between modelling of labour market rigidities will have the form of a “wage curve”. Understood in a broad sense, this is any locus off the labour supply curve whose intersection with labour demand determines the wage, thus replacing labour supply in this function. There are two principal approaches to the construction of a wage curve. The first one is empirical. We start from the observed situation and try to determine the factors driving changes in the wage econometrically. The result of the estimation is then implemented as an empirical wage curve in the model (for an overview, see Folmer, 2009). The “Wage Curve” (in a narrow sense) of Blanchflower and Oswald (1995) is the best known example of this approach. Blanchflower and Oswald find an empirically astonishingly stable relationship between the level of unemployment and the wage (see our extensive discussion in Section 5.6). Two examples of models that include the unemployment-wage relationship (and *only* this) explicitly as one of the equations are Hutton and Ruocco (1999) and Maisonnave et al (2009). In the MIMIC model (Graafland et al, 2001, p. 125), the wage depends, in addition, on the average and marginal labour tax rates.

For setting up a wage curve to be empirically estimated, we need to determine

⁴⁰This is explicit, for example, in Dixon et al (1978), Corden and Dixon (1980), Dixon et al (1982).

the variables that are candidates for inclusion in the equation. Usually, this is done based on a structural model of wage determination in imperfectly competitive labour markets, which brings us to the second approach to the wage curve. There are three theories that have been used in this way: search and matching (Pissarides, 1990), efficiency wages (Shapiro and Stiglitz, 1984) and collective wage bargaining (McDonald and Solow, 1981). By establishing a wage that is above the market-clearing level, these theories cover at the same time the phenomenon of involuntary unemployment. As theories of unemployment they are discussed in the following section (5.3).

The attempts to determine the empirical content of these structural theories of wage formation may be classified into two strands. In the first, dominant type of literature, the theories are used to derive empirical wage curve specifications, which are then used to generate elasticity estimates (as reviewed in Folmer, 2009). In the second type of literature (Pissarides, 1998, and Sørensen, 1999), the focus is on the structural parameters in each of the theories. These are identified and given empirically plausible values. Our review in Section 5.3 follows the second approach and pays particular attention to the question of how this fits in with the stylised fact of Blanchflower and Oswald’s wage curve.

5.3 Involuntary unemployment

Pissarides (1998) and Sørensen (1999) show how the most prominent options of modelling unemployment – search and matching, efficiency wages and collective bargaining – are calibrated and implemented in a simple CGE context. We start by briefly reviewing the core ideas of the three options, before going into more detail in Sections 5.3.1 to 5.3.3.

The *search-and-matching model* starts from the core assumption that finding a job is a time- and effort-consuming process, the more so, the less vacancies there are. Similarly, posting a vacancy is costly, and not worth the effort if there are no unemployed workers to be found. Therefore a certain level of unemployment is necessary to keep the process of continuous reallocation of jobs running. The level of unemployment is determined by structural parameters of the labour market, such as the efficiency of the matching process.

Search-and-matching models exist in a vast number of varieties, which is partly due to the fact that they lend themselves conveniently to empirical estimation with micro data. The model varieties differ in several aspects: whether there is on-the-job search or not, how wages are determined once a match materialises, and whether jobs and/or workers are homogeneous or heterogeneous in productivity. However, only the most simple of these varieties have been used in a CGE context.

The core idea of the *efficiency wage model* is that employers can increase the productivity of workers by paying wages that are above the market-clearing level. There are several versions of the precise productivity-increasing mechanism. In the original version the market-clearing wage is below the subsistence level of workers so that only higher wages allow workers to reproduce their working skills (Leibenstein, 1963). In other versions of the story higher wages increase the motivation of workers (“gift exchange”, Akerlof, 1982), or they create unemployment so that firing workers becomes a serious threat that deters workers from shirking (“unemployment as a worker disciplining device”, Shapiro and Stiglitz, 1984). In any case, the formal theory starts from an efficiency curve, which gives work efficiency as a function of the (relative) wage. For the empirical implementation, the story behind this efficiency curve is not strictly necessary.

The *collective wage bargaining model* formalises the idea that wages result from negotiations between firms’ associations and trade unions. The crucial idea is that part of the social costs of higher wages and resulting unemployment are externalities not borne by the bargaining parties in a particular sector, but by society as a whole (through unemployment benefits financed by general tax revenue or social security contributions). This externality induces the bargaining parties to go for wages that are above the market-clearing level.

In the case of bargaining theories, as well, we have quite a few sub-varieties: right-to-manage versus efficient bargaining, insider model versus utilitarian union, bargaining over wages only versus bargaining over wage/hours-of-work packages, monopoly union versus two bargaining parties. However, the differences between these varieties become small, or even irrelevant, once the model is calibrated to the actual unemployment rate. For the case of the insider model versus utilitarian union model this is shown in Boeters (2011).

Even if we tend to use a structural model for the implementation of involuntary unemployment, we are left with the difficult question of which one to use. All mechanisms sketched above have their plausibility, albeit to a different degree in different segments of the labour market. The search-and-matching mechanism is certainly an important element in any convincing explanation of unemployment. But one can doubt whether this is the whole story. It characterises unemployment mainly as a transitory phenomenon and downplays long-term, structural effects. The efficiency-wage approach is right in stressing the disciplining effect of unemployment, but one can doubt whether work effort would actually drop to zero without unemployment, as the shirking theory implies. Finally, the relevance of collective bargaining for wage formation is mainly determined by the concrete institutions of the country in question. It is considerably higher for continental Europe and Scandinavia than for the Anglo-Saxon countries.

In the following paragraphs, we present simple formalisations of the three basic approaches. We also discuss how, and to what extent, they can be given an empirically meaningful interpretation in a CGE context.

5.3.1 Search and matching

The core element of any search-and-matching model is the matching function, which gives matches, M , as a function of the number of unemployed workers, U , and vacancies, V . A Cobb-Douglas specification with efficiency parameter μ and share parameter ε ($0 < \varepsilon < 1$) reads:⁴¹

$$M = \mu U^\varepsilon V^{1-\varepsilon}$$

It is convenient to define “labour market tightness”, θ , as an additional variable:

$$\theta \equiv V/U$$

and express the probabilities of finding a job, a , and filling a vacancy, q , as

$$a \equiv M/U = \mu \theta^{1-\varepsilon}$$

⁴¹The material of this section follows closely the expositions in Pissarides (1998) and Sørensen (1999).

$$q \equiv M/V = \mu\theta^{-\varepsilon}$$

There are two different labour market states for workers: employed (index E) and unemployed (index U). With instantaneous utilities U , discount rate ρ and exogenous separation rate s for existing jobs, the following recursive value equations can be formulated, where J^i is the value of being in state i .

$$\rho J^E = U^E - s(J^E - J^U)$$

$$\rho J^U = U^U + a(J^E - J^U)$$

Analogously, on the part of the firm, jobs can be either occupied (index O) or vacant (index V). With job productivity⁴² p , labour cost w and vacancy cost γ , we have

$$\rho J^O = p - w - s(J^O - J^V)$$

$$\rho J^V = -\gamma + q(J^O - J^V)$$

Finding a match for a vacant job creates a rent, which needs to be shared between firm and worker. A simple sharing rule is Nash bargaining, which is formalised as the maximisation of a Nash function containing the weighted product of the individual shares, $J^E - J^U$ for the worker and $J^O - J^V$ for the firm. In logs and with λ as the relative bargaining power of the worker, Nash bargaining can be formulated as

$$\max_W [\lambda \log(J^E - J^U) + \log(J^O - J^V)]$$

where the bargaining parties treat J^U and J^V as exogenous. This gives the following first order condition:

$$\lambda \frac{dU^E}{dw} (J^O - J^V) - (J^E - J^U) = 0$$

Substitution of the state value variables, observing the free-entry condition, $J^V = 0$, brings us to

$$\frac{dU^E}{dw} \lambda \gamma (\rho + s + a) - q (U^E - U^U) = 0$$

If utility is linear in consumption, we have

$$\frac{dU^E}{dw} = 1 - t^m$$

⁴²Productivity is endogenous in a general equilibrium context, but it remains exogenous for the individual firm.

$$U^E = (1 - t^a) w$$

$$U^U = c(1 - t^a) w$$

where t^m and t^a are the marginal and average tax rates on labour income, and c is the replacement rate (unemployment benefit as a share of after-tax labour income). Then we have

$$(1 - t^m) \lambda \gamma (\rho + s + a) - q(1 - c)(1 - t^a) w = 0$$

and w can be expressed as a mark-up on the cost of a vacancy, γ .

$$w = \frac{(1 - t^m) \lambda (\rho + s + a)}{q(1 - c)(1 - t^a)} \gamma \quad (14)$$

As long as the institutional parameters – λ , c and the tax rates – are constant, the mark-up factor depends positively on the probability of finding a job, a , and negatively on the probability of filling a vacancy, q . In the dynamic equilibrium,⁴³ inflow to and outflow from unemployment must be the same, which means

$$au = s(1 - u)$$

and can be used for replacing a in eq. (15):

$$w = \frac{(1 - t^m) \lambda \left(\rho + \frac{s}{u}\right)}{q(1 - c)(1 - t^a)} \gamma \quad (15)$$

This is a wage curve in the narrow sense, because it gives us the wage, w , as a decreasing function of the unemployment rate, u .

Calibration

The calibration of the search-and-matching model to empirical wage curve elasticities turns out to be a problem. Taking logs on eq. (15) and calculating the partial derivative with respect to u gives the wage curve elasticity

$$\eta_{wu} = \frac{d \log w}{d \log u} = -\frac{s}{u\rho + s} \quad (16)$$

The expression on the right hand side contains only three parameters, the separation, unemployment and discount rates. These are parameters on which in principle

⁴³We focus on the equilibrium version of the search-and-matching theory. An interesting disequilibrium set-up, whose ambition is to cover short-term adjustment as well, is presented in Dixon and Rimmer (2003) and applied to illegal immigration in Dixon et al (2011).

empirical information is available. In contrast, parameters that are difficult to fix empirically, namely the bargaining power λ and the vacancy cost γ , have disappeared. Rather than helping us to determine additional parameters, eq. (16) produces an empirical implausibility. At a yearly basis,⁴⁴ s is of the same order of magnitude as u , and ρ is small. This gives a value of close to -1 for eq. (16), considerably diverging from the prominent Wage Curve elasticity value of -0.1 . (See Section 5.6 for the interpretation of empirical Wage Curve estimates.)

This empirical implausibility is not necessarily the end of the calibration story. It is reasonable to argue that in eq. (15), the probability of filling a vacancy, q , is not to be treated as a parameter, but as an endogenous variable that moves inversely with unemployment. However, deriving the precise functional relationship is not straightforward and involves other equations, which again requires decisions about what is endogenous and what is exogenous.

5.3.2 Efficiency wages

The cornerstone of the efficiency wage theory is some kind of monitoring technology, which is necessary in order to enforce discipline among the workers.⁴⁵ In the simplest case we have discrete effort, i.e. workers decide whether to exert either no effort or a given effort package, which avoids being detected shirking. The detection technology is monitoring in regular time intervals, which uncovers shirking, if present, with probability q . As in the search-and-matching set-up, we work with recursive value functions. This time we have three, for employed workers that shirk (index ES) or do not shirk (EN), and for the unemployed (U):

$$\begin{aligned}\rho J^{EN} &= U^{EN} - s (J^{EN} - J^U) \\ \rho J^{ES} &= U^{ES} - (s + q) (J^{ES} - J^U)\end{aligned}$$

$$\rho J^U = U^U + a (J^E - J^U)$$

⁴⁴The choice of the time unit for flow accounting is relevant to the individual flow rates. However, changing the time unit would affect s and ρ proportionally so that the effect cancels out.

⁴⁵The following formulation of the theory is based on Pissarides (1998), who in turn draws upon Shapiro and Stiglitz (1984).

where the J are state values, U are instantaneous utilities, s is the exogenous separation rate, ρ is the discount rate, and a is the probability of finding a job. In the steady state, where the flow equilibrium must hold, a can be expressed in terms of the unemployment rate, u ,

$$a = \frac{s(1-u)}{u}$$

If they want to prevent shirking, firms must offer a wage that makes workers indifferent between shirking and not shirking, i.e.

$$J^{EN} = J^{ES} = J^E \quad (17)$$

For tracing out the algebraic consequences of this, we must specify the instantaneous utility function. A common specification (e.g. Pissarides, 1998) is to assume linearity in the wage and commensurability of effort, e , with the wage so that we have

$$U^{EN} = w - e, \quad U^{ES} = w \quad U^U = b$$

with unemployment benefits b . Then the non-shirking requirement (17) results in the wage equation

$$w = \rho J^U + \frac{\rho + s + q}{q} e$$

i.e. the wage is determined as a mark-up on the annualised value of unemployment, which increases with the required effort, e , and decreases with the detection probability, q . Further substitution of J^U leads to

$$w = b + \left[\frac{s}{u} + \rho + q \right] \frac{e}{q} \quad (18)$$

If unemployment benefits are defined as a fixed share of the wage (i.e. we have a constant replacement rate, c), we can use $b = cw$, and eq. (18) turns into

$$w = \frac{1}{1-c} \left[\frac{s}{u} + \rho + q \right] \frac{e}{q} \quad (19)$$

Calibration

Eqs. (18) and (19) are wage curves in the strict sense that they contain a negative effect of the unemployment rate, u , on the wage, w . In the calibration context, we have good prospects to obtain information on (or estimation of) the separation rate, s , the discount rate, ρ , the unemployment rate, u , and the replacement rate. Two

unknowns remain, e and q , for which there are no prospects of empirical foundation.⁴⁶ As eq. (19) must hold in equilibrium,⁴⁷ we can formulate an expression for e as a function of q :

$$\tilde{e} = (1 - c) q \left(\frac{s}{u} + \rho + q \right)^{-1} \quad (20)$$

where $\tilde{e}$ is the effort normalised by the wage, $\tilde{e} = e/w$.

From eqs. (18) and (19), we can calculate the respective wage curve elasticities, η_{wu} and η_{wu}^c .

$$\eta_{wu} = \frac{d \log w}{d \log u} = -\frac{s\tilde{e}}{qu} \quad (21)$$

$$\eta_{wu}^c = -\frac{1}{1 - c} \frac{s\tilde{e}}{qu} \quad (22)$$

Using eqs. (20) and (21) to eliminate $\tilde{e}$, we arrive at

$$q = -\frac{s}{u} \left(\frac{1 - c}{\eta_{wu}} + 1 \right) - \rho \quad (23)$$

$$q = -\frac{s}{u} \left(\frac{1}{\eta_{wu}^c} + 1 \right) - \rho \quad (24)$$

This allows us to calculate q . If we choose plausible values for the other parameters in eq. (23), $s = 0.2$, $u = 0.1$, $c = 0.6$, $\rho = 0.05$ and the target wage curve elasticity of $\eta_{wu} = -0.1$, then we end up with q being close to 6. In the case of a fixed replacement rate, this increases to close to 18. Interpreting these values, it is important *not* to interpret q as a probability (where values above one do not make sense), but as an instantaneous probability rate. A rate of above one means that the relevant event (caught while shirking) is expected more than once in the time interval (one year). Values of 6 and 18 for q then mean that a shirker is caught roughly after 2 months and in less than one month respectively.⁴⁸

⁴⁶The calibration of the efficiency wage model, again, closely follows Pissarides (1998).

⁴⁷(19) must hold in both cases, even if unemployment benefits are not defined as a fixed share of the wage.

⁴⁸Pissarides (1998) observes the problem of the indeterminacy of q and e as well. He does not resort to the wage curve for calibration, however. Instead, he uses arguments based on the value of leisure time in the utility function. This produces a consistency problem, because in his model specification, utility is linear in income, as in the model of this section.

5.3.3 Collective wage bargaining

In the model discussed in this section, wage formation is conceptualised as the outcome of collective bargaining between a trade union and a representative firm. There are many varieties of this model, but for a concrete start, and to establish terms, we make more specific assumptions: (i) bargaining is only about the wage, not about employment (“right-to-manage” approach)⁴⁹, (ii) the trade union is only concerned with the utility of its employed members (“insider model”)⁵⁰, (iii) hours of work are exogenous. Formally, wage determination is implemented as the maximisation of a Nash bargaining function, Ω , where trade unions are represented by the utility mark-up over their fallback option, $U_e - U_a$, and firms by profits, π . The relative bargaining power of the trade union, λ , is an unobservable parameter to be determined in the calibration.

$$\max_w \quad \Omega = (U_e - U_a)^\lambda \pi \quad (25)$$

Maximisation is with respect to the *before-tax* wage, w .⁵¹ The fallback option of the union, U_a , is composed of possible employment in another sector, $\tilde{U}_e$, and unemployment (receiving unemployment benefits), U_u , with weights determined by the unemployment rate, u :

$$U_a = (1 - u)\tilde{U}_e + uU_u$$

The fallback option is exogenous to the individual wage bargain so that the first-order condition of the maximisation of the Nash function is

$$\lambda \frac{dU_e/dw}{U_e - U_a} + \frac{d\pi/dw}{\pi} = 0$$

Both firms and employed workers make optimal choices, given the wage. This allows us to apply the envelope theorem and express the first-order condition in terms of

⁴⁹Sørensen (1999) shows that for the type of numerical analysis intended, the choice between right-to-manage and efficient bargaining (where bargaining extends also to the number of employed workers) hardly matters.

⁵⁰Further below in this section, we show that the results are identical to those obtained with a utilitarian union as long as the value shares and the elasticities of labour demand and hours supply are constant.

⁵¹Maximisation with respect to the after-tax wage yields the same results. Nevertheless, the distinction must be kept in mind when comparing the formulas obtained with the literature.

partial effects:

$$\lambda \frac{\partial U_e / \partial w}{U_e - U_a} + \frac{\partial \pi / \partial w}{\pi} = 0$$

where the partial effect for the firm under profit maximisation is

$$\frac{\partial \pi}{\partial w} = -L$$

Finally, using elasticities, we have

$$\frac{\partial \log \Omega}{\partial \log w} = \lambda \frac{\partial \log U_e}{\partial \log w} \frac{U_e}{U_e - U_a} - \frac{wL}{\pi} = 0. \quad (26)$$

General equilibrium

With symmetrical sectors in equilibrium, we set

$$U_e = \tilde{U}_e$$

and eq. (26) can be re-formulated as

$$\frac{\partial \log \Omega}{\partial \log w} = \lambda \frac{\partial \log U_e}{\partial \log w} \frac{U_e}{u(U_e - U_u)} - \frac{wL}{\pi} = 0.$$

This reveals the basic wage curve relation: The higher unemployment, u , the lower the union benefit from a marginal increase of w , and, consequently, the lower the wage in equilibrium.⁵²

Calibrating the bargaining power parameter

The relative bargaining power of the trade union is a parameter notoriously impossible to observe. Its value can be recovered, however, from the observed unemployment rate. Assuming that the initial state is an equilibrium, eq. (26) must hold and the value of λ can be determined as⁵³

$$\lambda = \frac{\bar{w}\bar{L}}{\bar{\pi}} \left(\frac{\partial \log U_e}{\partial \log w} \right)^{-1} \frac{\bar{u}(\bar{U}_e - \bar{U}_u)}{\bar{U}_e} \quad (27)$$

The value of λ depends on the specification of the utility function. In particular, it depends on the modelling of the utility of the unemployed, i.e. how a difference in income compared to the employed translates into a difference in utility (see Section

⁵²See our general discussion of the wage curve as basis for calibration in Section 5.6.

⁵³Variables with an upper bar denote the values of the initial equilibrium used for calibration, and are therefore treated as constant.

3.1.4). However, *given* a particular specification of the utility function, the calculation of λ is straightforward.

The numerical essence of the Nash bargaining first-order condition can then be expressed as

$$\frac{\partial \log U_e}{\partial \log w} \frac{U_e}{U_e - U_a} = \frac{wL}{\lambda \pi} := \zeta \quad (28)$$

If we are prepared to assume that we are in a Cobb-Douglas world (or close to it),⁵⁴ then ζ is a constant, and the Nash FOC boils down to keeping the left-hand side of eq. (28) at its initial level. This does not change when we work with other specifications of the trade union's objective function.

In the case of a utilitarian union, which also takes the level of employment into account when bargaining over the wage, the Nash function is

$$\Omega' = [(U_e - U_a)L]^\lambda \pi,$$

and the corresponding first-order condition

$$\frac{\partial \log \Omega'}{\partial \log w} = \lambda \left(\frac{\partial \log U_e}{\partial \log w} \frac{U_e}{U_e - U_a} - \varepsilon_{Lw} \right) - \frac{wL}{\pi} = 0, \quad (29)$$

where the additional term, ε_{Lw} is the (absolute value of the) elasticity of employment, L , with respect to the wage.

$$\varepsilon_{Lw} = - \frac{\partial \log L}{\partial \log w}$$

As long as this is constant or almost constant, we are essentially back at eq. (28), although λ has a different value now.

$$\lambda' = \frac{\bar{w}\bar{L}}{\bar{\pi}} \left(\overline{\left(\frac{\partial \log U_e}{\partial \log w} \right)} \frac{\bar{U}_e}{\bar{u}(\bar{U}_e - \bar{U}_a)} - \bar{\varepsilon}_{Lw} \right)^{-1}.$$

Compared to eq. (27), the denominator shrinks, and thus $\lambda' > \lambda$. A utilitarian union is roughly equivalent to an insider union that has a (correspondingly) lower relative bargaining power.

⁵⁴The assumption can be formulated still somewhat weaker: Even if the factor shares change, this will not have a systematic effect on wage negotiations. This comes close to denying the assumption, implicit in the Nash bargaining function, that parties bargain over *relative* mark-ups over their respective outcome. So if some shock increases (reduces) the basis they bargain upon, their mark-up will increase (fall) accordingly. This mechanism does not have much intuitive appeal.

The same applies to a more general Nash function with variable weights for individual utility and employment, such as

$$\max_w \quad \Omega'' = (U_e - U_a)^\lambda L^\mu \pi,$$

which includes both the utilitarian union ($\lambda = \mu = 1$) and the insider model of eq. (25) with $\lambda = 1$ and $\mu = 0$ (see Graafland et al 2001, ch. 7).

$$\frac{\partial \log \Omega''}{\partial \log w} = \left(\lambda \frac{\partial \log U_e}{\partial \log w} \frac{U_e}{U_e - U_a} - \mu \varepsilon_{Lw} \right) - \frac{wL}{\pi} = 0, \quad (30)$$

This leaves us with two parameters, μ and λ , which cannot be identified through a single first-order condition without further information. However, we can choose an arbitrary value of μ , solve for λ , and – as long as the value shares and the labour demand elasticities are constant – we still remain with the essential equation

$$\frac{\partial \log U_e}{\partial \log w} \frac{U_e}{U_e - U_a} = \overline{\left(\frac{\partial \log U_e}{\partial \log w} \right)} \frac{\bar{U}_e}{\bar{u} (\bar{U}_e - \bar{U}_u)}$$

Linear utility function

The majority of the existing literature works with the simplifying assumption of a utility function that is linear in net income. If individual hours of work are exogenous and normalised to unity, we have

$$U_e = (1 - t_a)w$$

$$U_u = b$$

where b are unemployment benefits. We then can solve the first order conditions of Nash bargaining explicitly for the wage, which gives us a wage curve

$$w = \frac{\zeta u}{\zeta u(1 - t_a) - (1 - t_m)} b \quad (31)$$

where w is increasing in t_a and b , and decreasing in u and t_m .

The picture is different if we assume that b is not given in absolute terms but as a constant share (“replacement rate”) of after-tax income:

$$b = c(1 - t_a)w$$

Then w drops out of the wage-bargaining first-order condition and we are left with

$$1 = \frac{\zeta u c (1 - t_a)}{\zeta u (1 - t_a) - (1 - t_m)}$$

This can be solved for u

$$u = \frac{(1 - t_m)}{\zeta(1 - t_a)(1 - c)} \quad (32)$$

u is increasing in c and t_a , and decreasing in t_m . We now have an explicit expression for u , but the downside is that w cannot be determined from the first-order condition any more but only in general equilibrium. With exogenous labour supply⁵⁵ $\bar{N}$, employment is determined as $(1 - u)\bar{N}$, and w results implicitly determined as the wage that – given macroeconomic labour demand – creates the desired employment level. There is a monotonous relationship between u and w , but without knowing general equilibrium elasticities, we cannot say anything about the wage curve elasticity.

Empirical implications

Again, we look at two extreme cases: unemployment benefit fixed in real terms and a constant replacement rate. In the case of fixed unemployment benefits, we have an explicit wage curve, eq. (31), which allows us to calculate the implied wage curve elasticity. With proportional taxes ($t_a = t_m$) this is

$$\eta_{wu} = -\frac{1}{\zeta u - 1}$$

Taking into consideration the calibration of ζ , eq. (28), and assuming a Cobb-Douglas world, where factor shares are constant, we arrive at

$$\eta_{wu} = -\frac{1 - c}{c}$$

For a replacement rate of $c = 0.6$, this means a wage curve elasticity of $\eta_{wu} = -0.67$, which is considerably higher than the empirically plausible value of -0.1 .

In the case of a fixed replacement rate, we have the unemployment rate as a function of policy parameters:

$$u = \frac{(1 - t_m)}{\zeta(1 - t_a)(1 - c)}$$

⁵⁵When labour supply is endogenous, even more equations become relevant to the determination of the wage.

i.e. the wage moves independently of u , which makes the derivation of a wage curve elasticity impossible.

The challenge in calibrating the wage curve of the collective bargaining model is finding a specification that leaves enough flexibility for adjustment to empirical parameters. A possible approach is to define a “mixed replacement rate regime”, in which unemployment benefits are a linear combination of a fixed part and an indexed part. With α as the share of the fixed part, we have

$$b = \alpha \bar{b} + (1 - \alpha) cw$$

resulting in a wage curve

$$w = \frac{\alpha \zeta u}{\zeta u(1 - t_a) - (1 - t_m) - (1 - \alpha)c\zeta u} \bar{b}$$

The same substitutions as above, combined with the assumption of a proportional tax schedule, yield the wage curve elasticity

$$\eta_{wu} = -\frac{(1 - t)(1 - c)}{(\alpha - t)c}$$

For $\alpha = 1$, this reproduces the elasticity value $\eta_{wu} = -(1 - c)/c$ derived earlier. Other values for α lead to a shift in η_{wu} . However, as one is the upper bound for α , changing α can only increase the (absolute value of the) wage curve elasticity, and does therefore not solve the calibration puzzle.

5.4 Long-term trends in unemployment

In the dynamic calibration of unemployment, we are faced with a problem similar to those discussed in the context of labour supply (see Section 3.5). A secular upwards trend in labour productivity, and thus in wages, can produce a trend in unemployment, which might well be spurious. This is because the comparison of state values for employment and unemployment, and essentially the instant utility values for these two states, are crucial for all structural models of unemployment. If an upward trend in wages changes the relative values of employment and unemployment, this will have a systematic effect on unemployment.

With unemployment benefits determined as a fixed share of the after-tax income of the employed (constant replacement rate), and utility linear in income, we arrive at a benchmark situation in which the unemployment rate remains constant even if wages change. Comparison with this reference case reveals how a trend in wages can systematically influence unemployment. When unemployment benefits are fixed in absolute terms and do not evolve proportionally with wages, wage increases make employment more attractive compared to unemployment, and thereby reduce the unemployment rate. The inverse effect occurs when we combine a constant replacement rate with basic, non-utility-generating, consumption, e.g. in a linear expenditure scheme. A proportional development in income for the employed and unemployed then does not translate into a proportional development of utility any more. With basic consumption uniform for both groups, relative utility of the unemployed will increase with rising wages, generating an upward trend in unemployment.

Similar non-proportionality effects may be produced by leisure as a utility-generating item (see Sections 3.1.1 and 3.1.4) or by adjustment factors for leisure productivity (see Section 3.5). In the collective bargaining model (see Section 5.3.3), we have a simple option for neutralising all these systematic effects on unemployment in the long run. λ , the relative bargaining power of trade unions, may be treated as a time-varying parameter, adjusted so that unemployment remains constant over time in the baseline. In the search-and-matching and efficiency wage models, the matching efficiency and the parameters of the monitoring function can take on similar roles for calibration.

5.5 Informal sector and dual labour market

In this chapter, we have repeatedly touched upon the issue of an informal sector and the “dual labour market” structure it can give rise to. For distributional and poverty analysis in particular, the distinction between the formal and informal sector is indispensable, given that a large fraction of the poorest workers are located in the latter (Fortin et al, 1997). Here we collect different aspects of this issue and investigate whether the dual labour market structure can be seen as simply one special case of differentiated labour markets (as discussed in Section 5.1).

On the labour demand side, it is usually a sectoral distinction that motivates the formal/informal divide. There is a number of different reasons for distinguishing between these two sectors: differences in technology, in wage formation or in the administrative treatment of firms (in particular taxation), see Fortin et al (1997) for an overview. A typical assumption about wage formation is that in the informal sector, wages are competitive at a full employment level, whereas in formal sector, wages are above the market-clearing level by one of the mechanisms discussed in Section 5.3 or some form of rent sharing. The basic idea of the “dual labour market” set-up of Harris and Todaro (1970) comes in on the labour supply side. Formal and informal labour markets are not seen as completely separated (which would amount to a simple labour market segmentation as discussed in Section 5.1), but as connected through imperfect labour mobility. With the wage higher in the formal than in the informal sector, we need a rationing mechanism that prevents a complete shift of the workforce to the formal sector. According to Harris and Todaro (1970), the crucial element is unemployment. Searching for a job is a time-consuming process (as highlighted in the search-and-matching literature, Section 5.3.1), therefore a critical level of unemployment will stop the migration from the informal to the formal labour market. The marginal worker is indifferent between a secure job in the informal sector and an uncertain job search process in the formal sector (“waiting queues” for jobs).

In the original Harris and Todaro (1970) set-up, migration between the informal and formal sectors is seen as a dynamic phenomenon, with the *change* in the number of workers in the formal sector as a function of the difference between the expected wage in the formal sector and the certain wage in the informal sector. This dynamic specification has been implemented, e.g., in Agénor et al (2003, Section 3.2.1). In most comparative static applications, however, a steady state assumption is used, to the effect that expected wages must be equal for all migration incentives to disappear,

$$w_i = (1 - u)w_f \tag{33}$$

where w_i and w_f are informal and formal wage respectively, and u is the unemployment rate in the formal sector.⁵⁶ Eq.(33), however, has an apparent shortcoming

⁵⁶Harris-Todaro loci of this simplest possible form can be found, e.g., in the models of Gilbert and Wahl (2002) for China, Alzua and Ruffo (2011) for Argentina.

when it comes to calibration. The formal-informal wage differential as well as the unemployment rate are in principle (though not perfectly) observable, and eq. (33) will hold by mere coincidence. Only by introducing some degree of freedom can eq. (33) be changed into a form that can be calibrated. The most straightforward way of creating this degree of freedom is to introduce fixed cost of migration (c_m), which is most plausible when informal and formal sector are thought to be separated spatially (as for example rural versus urban sector).⁵⁷ Even if there is no distance to be bridged, “fixed cost of migration” can be thought of as the cost of becoming accustomed to the rules and conventions of the formal labour market, and of establishing the contacts necessary to find a job. With fixed costs of migration in place, the Harris-Todaro locus becomes

$$w_i + c_m = (1 - u)w_f \quad (34)$$

and c_m can be calibrated so that eq. (34) holds for empirical values of unemployment and the wage differential. In more general terms, c_m can be understood as any utility-generating difference between the sectors (apart from the wage and unemployment), e.g. less flexibility or a higher status in the formal sector.⁵⁸

The next hurdle for the Harris-Todaro locus to be taken is the empirical elasticity of migration. Eqs. (33) and (34) both assume that, in a partial analysis of the formal sector and as a reaction to a change in the formal wage, unemployment must adjust so that the product of the employment rate and the wage remains constant, i.e. the elasticity of unemployment with respect to the formal wage is

$$\varepsilon_{uw_f} = \frac{1 - u}{u}$$

This can be compared to wage curve estimations and checked against expert judgement. If the outcome of this check is that more adjustment is needed, we have the following two possible extensions. First, we can assume that the migration costs are not uniform across workers, but heterogeneous. The workers who have the lowest migration costs switch to the formal market first, so that marginal migration

⁵⁷In many applications, the formal/informal and urban/rural distinction have been used interchangeably. Stifel and Thorbecke (2003), however, make the point that they need to be distinguished, which leads to a “dual-dual” model.

⁵⁸Stifel and Thorbecke (2003, p. 220) encounter the same problem and solve it with some ad-hoc “adjustment parameter”, used as a multiplier for unemployment in (33).

costs are increasing with the size of the formal sector, and the migration elasticity is reduced compared to the case with homogeneous migration costs.⁵⁹ Second, we can go beyond the simple comparison of expected *wages* and proceed to expected *utilities*. This requires specifying the utility of the unemployed in the formal sector. Is there any unemployment compensation? If not, how do the unemployed organise themselves to make their living? In any case, the Harris-Torado locus becomes

$$U(w_i) = uU(w_u - c_m) + (1 - u)U(w_f - c_m) \quad (35)$$

By specifying the elasticity of the utility function with respect to income (which in turn determines the degree of risk aversion), we can change the migration elasticity. With no risk aversion, eq. (35) essentially collapses back into the form of eq. (34). With risk aversion, in contrast, an increase in unemployment is more detrimental to utility than to the expected wage. Therefore, we need a smaller increase in unemployment to compensate for a given wage rise. This means that the migration elasticity is lower than in the case of eq. (34).⁶⁰

The Harris-Todaro approach is not restricted to assuming full employment in the informal sector. In a more flexible setting, with both formal and informal unemployment, we have

$$(1 - u_i)w_i + c_m = (1 - u_f)w_f \quad (36)$$

A formulation similar to eq. (36) has been used in Böhringer et al (2005) to model wage differentials between several formal sectors in Germany. As with the calibration issue discussed above, it turned out that it is impossible to match empirical sectoral wage differentials with a plausible spread in unemployment rates. The additional assumption of compensating non-pecuniary job characteristics is necessary to get the model afloat.

The Harris-Todaro approach can be implemented in a straightforward manner with sectors that are a part of the official input-output tables. However, the attempt

⁵⁹This idea is similar to the model of the extensive margin of labour supply in Section 3.1.4. In an empirical microsimulation setting, Bourguignon et al (2005) use the same idea: They estimate the “closeness” of households to the formal labour market, and, once additional jobs become available, assign them to those households with the lowest values of the distance measure.

⁶⁰This idea is used in the “generalised Harris-Todaro migration function” in a model for Ethiopia by Gelan (2002).

to model the informal sector can run, by its very nature, into data availability problems because informal activities are notoriously difficult to observe. Since the seminal study on Kenya (ILO, 1972), estimates of the size and structure of the informal sector are available for most countries. Given the importance of the informal economy for many developing countries, many statistical offices use methods to make the estimates of the informal sector consistent with the formal economy and publish it as an integral part of the national accounts. When there is reason to doubt the magnitudes officially published, data adjustment may be performed with secondary data sources (see, e.g., Fortin et al, 1997).

Even in countries with a high rate of formal employment, and therefore a low quantitative significance of the informal sector, capturing informal structures can be important as a benchmark for formal wages. In most varieties of unemployment models (Section 5.3) it is assumed that the non-market-clearing wage is formed as a mark-up on the income of the unemployed. When there is a fully fledged social security system, it is reasonable to assume that this benchmark is provided by unemployment insurance or social assistance. If replacement rates from the social security system in the case of unemployment are low, however, we need to specify informal income sources in order to explain how people can survive nevertheless. This is particularly important if the model works with a linear expenditure system that includes some minimum consumption, which must be covered in any case.

The reference point of an informal income that forms the basis of a mark up in wage formation in imperfectly competitive labour markets has also been used in the empirical determination of a wage curve. Graafland and Huizinga (1999) use an informal sector (whose productivity is unknown a priori, but can be estimated or calibrated) to give the model enough flexibility to reproduce the empirical long-run pattern of wages.

5.6 The wage curve as a calibration target

Our discussion of the Harris-Todaro approach to informal sector modelling and migration in Section 5.5 enables us to take a fresh look at the wage curve elasticity as a calibration target for structural models of involuntary unemployment. In Sections 5.3.1 to 5.3.3, we have assumed that structural wage equations derived from one of

the models of involuntary unemployment need to be calibrated to match the results of empirically estimated wage equations. The prime example of a relationship of this sort is the wage curve of Blanchflower and Oswald (1995), which gives a stable relationship between the wage level and unemployment: a wage curve elasticity of -0.1 .

However, the usefulness of wage curve estimations for the calibration of structural models of unemployment needs a second look. After all, empirical wage curve estimations are usually based on regional differences, whereas a regional disaggregation can hardly be found in CGE models (except for explicit regional modelling, which we do not cover in this chapter). Are wage curve estimations nevertheless a useful input for calibration?

For concreteness, assume we have two regions (indexed by i) with efficiency wage formation according to the model in Section 5.3.2, and in both regions the efficiency wage curve, eq. (18),

$$w_i = b_i + \frac{s_i(1 - u_i)}{u_i} \frac{e_i}{q_i} + \frac{\rho + s_i + q_i}{q_i} e_i \quad (37)$$

holds (with the respective sectoral variables). In addition, we have a migration equation that links both sectors. Here, we take the simplest possible Harris-Todaro formulation (see Section 5.5 below), even if it is not fully consistent with the dynamic value-function approach that we have used in Section 5.3.2.

$$w_1(1 - u_1) = w_2(1 - u_2)$$

Taken at face value, the migration equation produces a positive wage curve elasticity, namely

$$\varepsilon_{wu} = \frac{dw_i}{du_i} \frac{u_i}{w_i} = \frac{w_i}{(1 - u_i)} \frac{u_i}{w_i} = \frac{u_i}{1 - u_i}$$

This is the Harris-Todaro idea in a nutshell. Higher wages in one sector must be *compensated* by higher unemployment in order to make potential migrants indifferent between the regions. However, this is not what we observe empirically, namely a *decreasing* locus between unemployment and the wage, i.e. the wage being depressed by higher unemployment.⁶¹ The most straightforward way to make this empirical

⁶¹Kingdon and Knight (2006) discuss the choice between a Harris-Todaro locus and the wage curve in the case of a model for South Africa.

observation consistent with the migration locus is to introduce sector-specific non-labour market amenities, a , e.g. more attractive landscape in a particular region. With normalisation for Sector 1, we then have⁶²

$$w_1(1 - u_1) = w_2(1 - u_2) + a_2 \quad (38)$$

When this is the *only* difference between regions, the relation between unemployment and the wages is solely determined by the sectoral wage curves. Unemployment and the wage share the role of compensating for non-economic regional amenities, with their respective shares derived from the regional wage curves. The resulting elasticity is the one calculated in Section 5.3.2:

$$\eta_{wu} = \frac{d \ln w}{d \ln u} = \frac{dw}{du} \frac{u}{w} = -\frac{s\tilde{e}}{qu}$$

Put otherwise, the system consisting of eqs. (37, 38) produces *different* elasticities depending on which parameter is shocked. For a marginal change in a , we get the elasticity η_{wu} , for a marginal change in a parameter appearing in eq. (37), e.g. the autonomous separation rate s , we get the elasticity ε_{wu} . This makes the interpretation of empirical estimation results a difficult task. While we can derive clean elasticities from the equation system, in reality we will measure a mix of all possible differences between regions, affecting or not affecting the regional wage equation. For getting the clean elasticity η_{wu} needed for calibration, we would ideally correct for factors that enter through the regional wage equation. However, as estimations of the wage curve are usually performed without assuming a concrete model (which would be the prerequisite for identifying factors that directly affect the wage curve) and without the aim of estimating an elasticity that can be used for calibration, this correction is not in the focus of empirical economists. When we use their wage curve estimates for calibration as in Sections 5.3.1 to 5.3.3, we implicitly assume that regional differences exist only in the non-economic amenities a .

6 Welfare analysis

By far the largest part of all CGE studies reviewed in this chapter restrict themselves to a positive economic analysis, i.e. tracing out the consequences of policy shocks for

⁶²Note the similarity to fixed migration costs discussed in connection with eq. (34).

observable economic variables, thus circumventing any potential problems connected to the calculation of welfare measures. Welfare analysis is one of the key assets of CGE analysis, however. In this chapter, we therefore review some complications that may arise once we try to compute welfare measures for several representative households (Section 6.1) or in a microsimulation set-up (Section 6.2).

6.1 Welfare measures for representative households

Welfare computations in the basic CGE model with several representative households are simple. We record the change in the utility functions of the households, evaluate them at the initial prices, add up and report the result as equivalent variation (EV).⁶³ Interpreting EV as a welfare measure is subject to the usual caveat: A positive EV of a policy reform means that, *if lump-sum redistribution between households is possible*, then a Pareto improvement compared to the initial situation *can* be generated. Such redistribution is not necessarily feasible, and, if feasible, cannot be assumed to actually take place. In these circumstances, EV is only valid as a welfare measure if the utility of different households is commensurable and the marginal utility of income is the same for all households.

In a model such as the one discussed in Section 3.1, the computation of EV requires special attention when status changes occur. Status changes to be considered are, first, participation to non-participation (or back), and, second, employment to unemployment (or back). Especially in the latter case, calculation of EV is not trivial, because unemployed individuals are not at their utility maximum. As a particular complication, EV for the switch from employment to unemployment is not the negative of EV for a switch from unemployment to employment, because of the quantity restriction. This is discussed in a formal way below.

When status changes take place, the number of individuals per representative household is not constant. Therefore, EV is first calculated for all possible transitions, and then summed up with appropriate weights.

The first group to be considered are households that are employed and remain

⁶³The following applies, vice versa, to compensating variation (CV), which is obtained if we value utility differences at the after-shock prices.

employed.⁶⁴

$$EV_{EE} = \bar{p}_U (U_e - \bar{U}_e)$$

Second, we have households that switch from employment to unemployment:

$$EV_{EU} = \bar{p}_U (U_u - \bar{U}_e)$$

Third, there are households that are unemployed and remain so. These households are subject to a demand constraint ($F = T - \delta \bar{H}$, see Section 3.1.4). Therefore we can restrict the welfare calculation to the consumption part of the utility function:

$$EV_{UU} = \bar{p}_C (C_D^u - \bar{C}_D^u)$$

Fourth, there are households that switch from unemployment to employment. Here we have no dual formulation of the utility function and must calculate equivalent income directly by solving the following equation for $\tilde{C}_D^u$:

$$U_e = \left[\theta_C \left(\frac{\tilde{C}_D^u}{\bar{C}_D} \right)^{\frac{\sigma-1}{\sigma}} + (1 - \theta_C) \left(\frac{T - \delta \bar{H}}{\bar{F}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

EV results as

$$EV_{UE} = \bar{p}_C (\tilde{C}_D^u - \bar{C}_D^u)$$

Here it becomes clear that treatment of switches into and out of employment is asymmetrical by the very nature of EV calculations. In the case of unemployed individuals, we ask “What amount of money would compensate them for being unemployed, given their labour supply constraint?” and in the case of employed individuals: “Which income loss would make them indifferent to being unemployed, given the possibility of optimal adjustment of labour supply?”

Apart from the switches between employment and unemployment, we must keep track of the switches between participation and non-participation. We start with households that switch from non-participation to participation. They are faced with idiosyncratic fixed costs of taking up work (see Section 3.1.4), and these cost must be accounted for in the welfare calculation. We discuss the two cases of increasing and decreasing participation separately.

⁶⁴For an explanation of the symbols see Section 3.1.

If the expected utility from participation increases, participation does so as well, and we have $N > \bar{N}$. The worker that was the marginal participant in the initial situation reaps the full gain, i.e.

$$EV_{\bar{N}} = U_l - \bar{U}_l = \frac{N - \bar{N}}{h}.$$

The new marginal participant has a welfare gain of zero, because he/she is indifferent compared to his/her previous non-participation. Integrating EV between these two points under the assumption of a uniform distribution gives

$$EV_P = \int_{\bar{N}}^N \frac{N - n}{h} dn = \frac{(N - \bar{N})^2}{2h}. \quad (39)$$

In the opposite case expected utility from participation, and therefore participation itself, increases ($N < \bar{N}$). Then the originally marginal participant has no utility loss, because he/she was indifferent to switching to non-participation. The worker who is indifferent in the new situation, by contrast, suffers the full utility loss.

$$EV_N = U_l - \bar{U}_l = \frac{N - \bar{N}}{h}.$$

Integrating gives

$$EV_P = \int_N^{\bar{N}} \frac{n - \bar{N}}{h} dn = -\frac{(N - \bar{N})^2}{2h},$$

which can be consolidated with eq. (39) to become

$$EV_P = \text{sign}(N - \bar{N}) \frac{(N - \bar{N})^2}{2h}.$$

Finally, all these effects are added up.

$$EV = \min(\bar{N}, N) \cdot [\min(1 - u, 1 - \bar{u}) \cdot EV_{EE} + \min(u, \bar{u}) \cdot EV_{UU}] \quad (40)$$

$$+ \max(0, \bar{u} - u) \cdot EV_{UE} + \max(0, u - \bar{u}) \cdot EV_{EU}] \quad (41)$$

Depending on the set-up, there may be a further complication due to the existence of residual households that originate from the adjustment of data inconsistencies between the macro and micro level (see Section 3.3.4). We can either evaluate their utility change (which requires the assumption of a specific utility function) and add the result to the macroeconomic EV in eq. (41), or we can define a compensating tax instrument that keeps utility of the residual households fixed and shifts all welfare changes to the worker households.

6.2 Welfare measures for microsimulated households

In principle, the concept of equivalent variation (EV) is applicable to individual households in the same way as it is to aggregated households. There are, however, two potential complications. The first one arises from the fact that the most common form of microsimulation involves discrete-choice modelling of labour supply (see Section 3.3.1). Once labour supply cannot be chosen continuously, the determination of EV also involves discrete elements. The second complication joins in if individual households are faced with a stochastic decision, as in the case of involuntary unemployment. We discuss these two problems in turn.

In the deterministic case, we have a utility function U that depends on income from work, Y_i , leisure, F_i , and individual- and option-specific fixed terms, R_i , with option-specific utility, U_i

$$U_i = U(Y_i, F_i, R_i)$$

When the initial choice of the household is option i and the final, optimal choice in the counterfactual situation is j with utility

$$U^+ = U(Y_j, F_j, R_j)$$

the basic idea is to calculate EV so that the following equation holds

$$U(Y_i + EV, F_i, R_i) = U^+$$

However, we do not know a priori whether the household compensated by a lump-sum transfer will remain with the same labour supply choice. It might switch to another option associated with a different value of the EV. Given the discrete choice, this can only be determined by calculating option-specific EVs for all options k , i.e.

$$U(Y_k + EV_k, F_k, R_k) = U^+$$

and then selecting the minimum of the resulting range of values (see Creedy and Kalb (2005b) for an extended discussion of this approach):

$$EV = \min_k EV_k$$

Matters become even more complex when the discrete choice of the households is stochastic. In Boeters (2010) e.g., households decide on labour supply *before* they know whether they will end up employed or unemployed. In this case we have a choice between different values of expected utility, probability-weighted averages over n labour market states (where $n = 2$ for singles and $n = 4$ for couples in the employed/unemployed case):

$$EU_i = \sum_n p_{i,n} U(Y_{i,n}, F_{i,n}, R_i)$$

Again, the basic idea is to calculate EV for all options and then find the minimum. In contrast to the deterministic case, however, utility values are now state-specific as well, whereas EVs are not. We thus obtain

$$\sum_n p_{k,n} U(Y_{k,n} + EV_k, F_{k,n}, R_k) = U^+$$

In Boeters (2010) this has produced the follow-up problem that in a small number of cases, EV takes on a considerably negative value, which is compatible with income from work, but produces negative income when subtracted from unemployment benefits. In the usual case of utility functions that require positive income, this leads to a infeasibility. Boeters (2010) deals with this problem by introducing a state-specific lower bound on EV so that negative values of income are prevented.

6.3 Utility weighting

So far, we have assumed that EV is simply summed up over individuals or representative households. This amounts to evaluating a utilitarian welfare function, where each welfare gain or loss counts the same, disregarding the characteristics of the individual or household where it occurs. Normally, economists engaged in distributional analysis are eager to go beyond this approach, and proceed to policy evaluation under genuine inequality aversion, i.e. giving higher weights to the utility of the poor than to the utility of the rich.

With heterogeneous households this is an ambitious project, and we know of no convincing solution to the resulting problems in the literature. Complications arise due to the combination of two features of heterogeneous labour markets. First, we

look at households that endogenously choose their labour supply. Therefore income is not well suited as an indicator of the distributional position of a household. A better candidate is earnings potential, i.e. income at some standardised value of labour supply. Second, many approaches of inequality aversion assume that utility values of different households are in principle commensurable, and that weighting can be derived from the decreasing marginal utility of income. However, with disaggregated, heterogeneous households, we normally encounter incommensurable utility functions. They result from the calibration to labour supply elasticity values, which means that they have different parameters and the absolute utility levels are not comparable. Similarly, if individual utility functions are econometrically estimated, we end up with different utility functions for different households. The marginal utility of income is no longer available as a natural way of deriving welfare weights.

An interesting approach of dealing with these problems is presented in Aaberge and Colombino (2008). The authors propose a weighting method that involves two diverging utility functions per household, one that determines labour supply, and a different one for meaningful welfare weighting. This obviously produces consistency problems, but Aaberge and Colombino (2008) argue that they can be tolerated.⁶⁵

7 Conclusions

More than other core elements of a CGE model, the labour market lacks a consensus or majority set-up. Which modelling strategy is appropriate strongly depends on the policy shock to be analysed and on the output variables of interest. Therefore our main aim in this chapter has been to present a portfolio of modelling options, together with their advantages and disadvantages in different modelling contexts. We hope that this kind of overview can be a useful guide to help practical modellers in their choice of an appropriate specification.

Although there is a broad menu of choices in labour market modelling, some core decisions must be taken. In our view, the most important choices are:

⁶⁵Decoster and Haan (2010) choose another route by implementing the ideas of Fleurbaey (2006) about welfare weighting in a discrete-choice labour-supply framework.

- In labour supply: whether to work with a small number of aggregated households or with a large set of micro-units.
- In labour demand: whether to follow the disaggregation motivated by the labour supply side, or rather to aggregate to broad labour demand categories.
- In labour market coordination: whether to work with an empirically founded wage curve or rather with a structural model of involuntary unemployment.
- In model organisation: whether to iterate a micro module (if existing) with the macro module or engage in one-way linkage.

Even if we advocate the view that the model structure can only be chosen when the shock to be analysed and the type of result to be generated have been specified, we are not completely neutral with respect to the four points above. We want to close this chapter with some broad recommendations that can serve as guidelines.

- Working with a microsimulation set-up is a forceful instrument, which has a number of clear advantages (direct link to modern labour supply estimation, explicitness in distributional questions) and avoids a number of typical problems that arise in determining the characteristics of representative households. Therefore, in our view, any labour-market related study should carefully check whether the microsimulation set-up can be made use of.
- Labour supply analysis suggests a large number of potentially interesting labour subgroups. Except for a very small subset of these (namely subgroups defined by skill), the econometric basis for formulating labour demand in these categories is weak. We think that in general the assumption of perfect substitutability in demand (implying an efficiency-weighted *additive* treatment of individual labour quantities) is a plausible default. Demand for different categories of labour should only be differentiated if there is evidence that wages do not move in parallel.
- Modelling labour market coordination presents itself as a sharp trade-off. Ideally, we would like to have a theoretically founded, structural model of involuntary unemployment, which contains enough free parameters to be calibrated

to empirical wage curve elasticity parameters. The review of models in Section 5.3 has shown that this is not easily available. Any reasonably simple structural model of unemployment has severe difficulty to be calibrated to empirically plausible wage curve elasticities. Working with these elasticities directly, without a structural foundation, is possible, but reduces our resources of providing an economic interpretation of changes in the wage as a response to policy shocks.

- If microsimulation (or any other micro module) is used, we strongly advocate an iterative modelling framework. Iteration between modules is an indispensable tool for detecting model inconsistencies and for finding clues to the explanation of model mechanisms which should not be dispensed with lightly. It might indeed often be the case that feedback from one module to the other is quantitatively small so that it does not contribute much to the qualitative model outcomes. However, the only way to confirm this is to perform the model iteration and compare the results with the one-way linkage.

References

- Aaberge, Rolf and Ugo Colombino (2008), Designing optimal taxes with a microeconomic model of household labour supply, CHILD Working Paper, 06/2008.
- Aaberge, Rolf, Ugo Colombino, Erling Holmøy, Birger Strøm and Tom Wennemo (2004), Population ageing and fiscal sustainability: An integrated micro-macro analysis of required tax changes, Statistics Norway Discussion Paper, 367.
- Aaberge, Rolf, John K. Dagsvik and Steinar Strøm (1995), Labor supply responses and welfare effects of tax reforms, *Scandinavian Journal of Economics*, 97, 635–659.
- Agénor, Pierre-Richard and Karim El Aynaoui (2003), Labor market policies and unemployment in Morocco: A quantitative analysis, World Bank Policy Research Working Paper, 3091.
- Agénor, Pierre-Richard, Alejandro Izquierdo and Hippolyte Fofack (2003), *Immpa: A quantitative macroeconomic framework for the analysis of poverty reduction strategies*, The World Bank Policy Research Working Paper Series, 3092.
- Agénor, Pierre-Richard, Mustapha K. Nabli, Tarik Yousef and Henning Tarp Jensen (2007), Labor market reforms, growth, and unemployment in labor-exporting countries in the Middle East and North Africa, *Journal of Policy Modeling*, 29, 277–309.
- Akerlof, George A. (1982), Labor contracts as partial gift exchange, *Quarterly Journal of Economics*, 97, 543–569.
- Alzua, Maria Laura and Hernan Ruffo (2011), Effects of Argentina’s social security reform on labor markets and poverty, MPIA Working Paper, 2011-11.
- Annabi, Nabil (2003), Modeling labor markets in CGE models: Endogenous labor supply, unions and efficiency wages, Working Paper, CIRPEE and Université Laval, Québec.
- Arntz, Melanie, Stefan Boeters, Nicole Gürtzgen and Stefanie Schubert (2008), Analysing welfare reform in a microsimulation-AGE model: The value of disaggregation, *Economic Modelling*, 25, 422–439.

- Baartmans, Kristien, Frits Meijer and Andrae van Schaik (1986), Houserepair and the informal sector (in Dutch), University of Delft.
- Bahan, David, Bernard Decaluwé, André Lemelin and Nabil Annabi (2005), Offre de travail endogène et mobilité du capital dans un megc bi-régional: La version statique du modèle d'équilibre général calculable du ministère des finances du québec, Ministère des Finances du Québec, collection Feuille d'argent, Travaux de recherche, 2005-001.
- Ballard, Charles L. (2000), How many hours are in a simulated day? The effects of time endowment on the results of tax-policy simulation models, Discussion Paper, Michigan State University,, URL <http://www.econ.msu.edu/faculty/ballard/HoursPaper.pdf>.
- Ballard, Charles L. and Don Fullerton (1992), Distortionary taxes and the provision of public goods, *Journal of Economic Perspectives*, 6, 117–131.
- Ballard, Charles L., Don Fullerton, John B. Shoven and John Whalley (1985), A general equilibrium model for tax policy evaluation, University of Chicago Press.
- Bargain, Olivier, Herwig Immervoll, Andreas Peichl and Sebastian Siegloch (2011a), Distributional consequences of labor-demand shocks: The 2008-09 recession in Germany, CESifo Working Paper, 3403 (forthcoming in *International Tax and Public Finance*).
- Bargain, Olivier, Kristian Orsini and Andreas Peichl (2011b), Labor supply elasticities in Europa and the US, IZA Discussion Paper, 5820.
- Bassanini, Andrea, Jørn Henrik Rasmussen and Stefano Scarpetta (1999), The economic effects of employment-conditional income support schemes for the low-paid: An illustration from a CGE model applied to four OECD countries, OECD Economics Department Working Paper, ECO/WKP(99)16.
- Benhabib, Jess and Clive Bull (1983), Job search: The choice of intensity, *Journal of Political Economy*, 91, 747–764.

- Benhabib, Jess, Richard Rogerson and Randall Wright (1991), Homework in macroeconomics: Household production and aggregate fluctuations, *Journal of Political Economy*, 99, 1166–1187.
- Berman, Eli, John Bound and Stephen Machin (1998), Implications of skill-biased technological change: International evidence, *Quarterly Journal of Economics*, 113, 1245–1279.
- Blanchflower, David G. and Andrew J. Oswald (1995), *The wage curve*, MIT Press.
- Boadway, Robin and John Treddenick (1978), A general equilibrium computation of the effects of the Canadian tariff structure, *Canadian Journal of Economics*, 11, 424–446.
- Boccanfuso, Dorothée, Bernard Decaluwé and Luc Savard (2008), Poverty, income distribution and CGE micro-simulation modeling: Does the functional form of distribution matter?, *Journal of Economic Inequality*, 6, 149–184.
- Boccanfuso, Dorothée and Luc Savard (2008), Groundnut sector liberalization in Senegal: A multi-household CGE analysis, *Oxford Development Studies*, 36, 159–186.
- Boccanfuso, Dorothée and Luc Savard (2011), The food crisis and its impacts on poverty in Senegal and Mali: Crossed destinies, *Development Policy Review*, 29, 211–247.
- Boeters, Stefan (2010), Optimal tax progressivity in unionised labour markets: Simulation results for Germany, *ZEW Discussion Paper*, 10-035.
- Boeters, Stefan (2011), Optimal tax progressivity in unionised labour markets: What are the driving forces?, *Economic Modelling*, 28, 2282–2295.
- Boeters, Stefan, Christoph Böhringer, Thiess Büttner and Margit Kraus (2010), Economic effects of VAT reform in Germany, *Applied Economics*, 42, 2165–2182.
- Boeters, Stefan and Michael Feil (2009), Heterogeneous labour markets in a microsimulation-AGE model: Application to welfare reform in Germany, *Computational Economics*, 33, 305–335.

- Boeters, Stefan, Michael Feil and Nicole Gürtzgen (2005), Discrete working time choice in an applied general equilibrium model, *Computational Economics*, 26, 183–211.
- Boeters, Stefan and Nico van Leeuwen (2010), A labour market extension for WorldScan: Modelling labour supply, wage bargaining and unemployment in a CGE framework, CPB Document, 201.
- Böhringer, Christoph, Stefan Boeters and Michael Feil (2005), Taxation and unemployment: An applied general equilibrium approach for Germany, *Economic Modelling*, 22, 81–108.
- Bonin, Holger and Hilmar Schneider (2006), Analytical prediction of transition probabilities in the conditional logit model, *Economics Letters*, 90, 102–107.
- Borjas, George J., Jeffrey Grogger and Gordon H. Hanson (2011), Substitution between immigrants, natives and skill groups, NBER Working Paper, 17461.
- Bourguignon, François, Anne-Sophie Robilliard and Sherman Robinson (2005), Representative versus real households in the macroeconomic modeling of inequality, in: Timothy E. Kehoe, T.N. Srinivasan and John Whalley (eds.), *Frontiers in Applied General Equilibrium Modeling*, 219–254, Cambridge University Press.
- Bourguignon, François and Thierry Magnac (1990), Labour supply and taxation in France, *Journal of Human Resources*, 25, 358–389.
- Bourguignon, François and Luc Savard (2008), A CGE integrated multi-household model with segmented labour markets and unemployment, in: François Bourguignon, Luiz A. Pereira da Silva and Maurizio Bussolo (eds.), *The impact of macroeconomic policies on poverty and income distribution: Macro-micro evaluation techniques and tools*, 177–212, Palgrave-Macmillan.
- Bovenberg, A. Lans, Johan J. Graafland and Ruud A. de Mooij (2000), Tax reform and the Dutch labor market: An applied general equilibrium approach, *Journal of Public Economics*, 78, 193–214.
- Bussolo, Maurizio, Jann Lay, Denis Medvedev and Dominique van der Mensbrugghe (2008), Trade options for Latin America: A poverty assessment using a top-down

- macro-micro modeling framework, in: François Bourguignon, Luiz A. Pereira da Silva and Maurizio Bussolo (eds.), *The impact of macroeconomic policies on poverty and income distribution: Macro-micro evaluation techniques and tools*, 61–90, Palgrave-Macmillan.
- Carneiro, Francisco G. and Jorge S. Arbach (2003), *The impacts of trade on the Brazilian labor market: A CGE model approach*, *World Development*, 31, 1581–1595.
- Cloutier, Marie-Hélène, John Cockburn and Bernard Decaluwé (2008), *Education and poverty in Vietnam: A computable general equilibrium analysis*, CIRPÉE Working Paper, 08-04.
- Cockburn, John, Ismail Fofana, Bernard Decaluwé, Ramos Mabugu and Margaret Chitiga (2007), *A gender-focused macro-micro analysis of the poverty impacts of trade liberalization in South Africa*, in: Peter J. Lambert, John A. Bishop and Yoram Amiel (eds.), *Equity*, 269–305, Elsevier.
- Cogneau, Denis and Anne-Sophie Robilliard (2008), *Simulating targeted policies with macro impacts: Poverty alleviation policies in Madagascar*, in: François Bourguignon, Luiz A. Pereira da Silva and Maurizio Bussolo (eds.), *The impact of macroeconomic policies on poverty and income distribution: Macro-micro evaluation techniques and tools*, 213–245, Palgrave-Macmillan.
- Colombo, Giulia (2008), *The effects of dr-cfta in Nicaragua: A cge-microsimulation model for poverty and inequality analysis*, *Proceedings of the German Development Economics Conference*, 6, 1–43.
- Corden, W. Max and Peter B. Dixon (1980), *A tax-wage bargain in Australia: Is a free lunch possible?*, *Economic Record*, 56, 209–221.
- Creedy, John and Guyonne Kalb (2005a), *Discrete hours labour supply modelling: Specification, estimation and simulation*, *Journal of Economic Surveys*, 19, 697–734.
- Creedy, John and Guyonne Kalb (2005b), *Measuring welfare changes in labour supply models*, *The Manchester School*, 73, 664–685.

- Dagsvik, John K. and Steinar Strøm (2003), Analyzing labor supply behavior with latent job opportunity sets and institutional choice constraints, Statistics Norway Discussion Paper, 344.
- Davies, James B. (2004), Microsimulation, CGE and macro modelling for transition and developing economics, WIDER Discussion Paper, 2004/08.
- de Melo, Jaime and David Tarr (1992), A general equilibrium analysis of US foreign trade policy, MIT Press.
- Decaluwé, Bernard, André Lemelin and David Bahan (2010), Endogenous labor supply with several occupational categories in a bi-regional CGE model, *Regional Studies*, 44, 1401–1414.
- Decoster, André and Peter Haan (2010), Empirical welfare analysis in random utility models of labour supply, IZA Discussion Paper, 5301.
- Dervis, Kemal, Jaime de Melo and Sherman Robinson (1982), General equilibrium models for development policy, Cambridge University Press.
- Devarajan, Shantayanan and Dani Rodrik (1991), Pro-competitive effects of trade reform: Results from a CGE model of Cameroon, *European Economic Review*, 35, 1157–1184.
- Dewatripont, Catherine, Simon Erlich, Victor Ginsburgh and Denise van Regemorter (1991), The effects on unemployment of reducing social security contributions: A general equilibrium analysis for Belgium, in: *Applied general equilibrium modelling*, Kluwer.
- Dimaranan, Betina and Badri Narayanan (2008), GTAP 7 data base documentation, Chapter 12.B: Skilled and unskilled labor data, GTAP Resource, 2936.
- Dixon, Peter B., Martin Johnson and Maureen T. Rimmer (2011), Economy-wide effects of reducing illegal immigrants in U.S. employment, *Contemporary Economic Policy*, 29, 14–30.
- Dixon, Peter B., John R. Madden and Maureen T. Rimmer (2010), Linking national and multi-regional CGE models: The effects of an increase in award wage rates in Australia, *Regional Studies*, 44, 1369–1385.

- Dixon, Peter B., Brian R. Parmenter, Alan A. Powell and Peter J. Wilcoxon (1992), Notes and problems in applied general equilibrium economics, North Holland.
- Dixon, Peter B., Brian R. Parmenter and John Sutton (1978), Some causes of structural maladjustment in the Australian economy, *Economic Papers*, 57, 10–26.
- Dixon, Peter B., Brian R. Parmenter, John Sutton and David P. Vincent (1982), ORANI: A multisectoral model of the Australian economy, North Holland.
- Dixon, Peter B. and Maureen T. Rimmer (1995), Macro and sectoral implications of tax increases with differing distributional impacts, *Australian Economic Review*, 28, 111–117.
- Dixon, Peter B. and Maureen T. Rimmer (2003), A new specification of labour supply in the MONASH model with an illustrative application, *Australian Economic Review*, 36, 22–40.
- Duncan, Alan and Melvyn Weeks (1998), Simulating transition using discrete choice models, *Journal of the American Statistical Association: Papers and Proceedings*, 106, 151–156.
- Evers, Michiel, Ruud A. de Mooij and Daniel van Vuuren (2008), The wage elasticity of labour supply : A synthesis of empirical estimates, *De Economist*, 156, 25–43.
- Fæhn, Taran, Antonio G. Gómez-Plana and Snorre Kverndonk (2009), Can a carbon permit system reduce Spanish unemployment?, *Energy Economics*, 31, 595–604.
- Falk, Martin and Bertrand Koebel (1997), The demand of heterogeneous labour in Germany, *ZEW Discussion Paper*, 97-28.
- Fallon, Peter R. and Richard Layard (1975), Capital-skill complementarity, income distribution, and output accounting, *Journal of Political Economy*, 83, 279–302.
- Ferreira, Francisco H.G, Phillippe G. Leite, Luiz A. Pereira da Silva and Paulo Picchetti (2008), Can the distributional impacts of macroeconomic shocks be predicted? A comparison of top-down macro-micro models with historical data for Brazil, in: François Bourguignon, Luiz A. Pereira da Silva and Maurizio Bussolo

- (eds.), *The impact of macroeconomic policies on poverty and income distribution: Macro-micro evaluation techniques and tools*, 119–174, Palgrave-Macmillan.
- Flaig, Dorothee, Khalid Siddig, Harald Grethe, Jonas Luckmann and Scott McDonald (2011), *The integration of Palestinian-Israeli labour markets: A CGE approach*, Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's 2011 AAEA and NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011.
- Fleurbaey, Marc (2006), *Social welfare, priority of the worst-off and the dimensions of individual well-being*, in: Francesco Farina and Ernesto Savaglio (eds.), *Inequality and economic integration*, 225–268, Routledge.
- Fofana, Ismaël, John Cockburn and Bernard Decaluwé (2003), *Modeling male and female work in a computable general equilibrium model applied to Nepal*, Working Paper, Université Laval.
- Folmer, Kees (2009), *Why do macro wage elasticities diverge? A meta analysis*, CPB Discussion Paper, 122.
- Fontana, Mariza (2004), *Modelling the effects of trade on women, at work and at home: Comparative perspectives*, *Economie Internationale*, 99, 49–80.
- Fontana, Mariza and Adrian Wood (2000), *Modeling the effects of trade on women, at work and at home*, *World Development*, 28, 1173–1190.
- Fortin, Bernard, Nicolas Marceau and Luc Savard (1997), *Taxation, wage controls, and the informal sector: An applied general equilibrium approach*, *Journal of Public Economics*, 66, 293–312.
- Fraser, Iain and Robert G. Waschik (2010), *The double dividend hypothesis in a CGE model: Specific factors and variable labour supply*, University of Kent School of Economics Discussion Papers, 10/01.
- Gaasland, Ivar (2008), *Modelling farmers' labour supply in CGE models*, Working paper, Institute for Research in Economics and Business Administration, 03/08.
- Gelan, Ayele (2002), *Trade liberalisation and urban-rural linkages: A CGE analysis for Ethiopia*, *Journal of Policy Modeling*, 24, 707–738.

- Gelauff, George M.M., Frank A.M. van Erp, Johan J. Graafland, Adriaan E. van Hien and André G.H. Nibbelink (1991), Towards an analysis of tax effects on labour market and allocation: A micro/macro approach, in: *Applied general equilibrium modelling*, Kluwer.
- Genre, Véronique, Karsten Kohn and Daphne Momferatou (2011), Understanding inter-industry wage structures in the euro area, *Applied Economics*, 43, 1299–1313.
- Giesecke, James A., Nhi H. Tran, G. Antony Meagher and Felicity Pang (2011), Growth and change in the Vietnamese labour market: A decomposition of forecast trends in employment over 2010-2020, *COPS General Paper*, G-216.
- Gilbert, John and Thomas Wahl (2002), Applied general equilibrium assessments of trade liberalisation in China, *The World Economy*, 25, 697–731.
- Graafland, Johan J., Ruud A. de Mooij, André G.H. Nibbelink and Ate Nieuwenhuis (2001), *MIMICing tax policies and the labour market*, Elsevier.
- Graafland, Johan J. and Frederik H. Huizinga (1999), Taxes and benefits in a non-linear wage equation, *De Economist*, 147, 39–54.
- Griliches, Zvi (1969), Capital-skill complementarity, *Review of Economics and Statistics*, 51, 465–468.
- Hamermesh, Daniel S. (1993), *Labor demand*, Princeton University Press.
- Harris, John R. and Michael P. Todaro (1970), Migration, unemployment and development: A two-sector analysis, *American Economic Review*, 60, 126–142.
- Hausman, Jerry (1995), Taxes and labor supply, in: Alan J. Auerbach and Martin Feldstein (eds.), *Handbook of Public Economics* 1, 213–263, North Holland.
- Heckman, James J., Lance Lochner and Christopher Taber (1998), Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents, *Review of Economics Dynamics*, 1, 1–58.
- Hendy, Rana and Chahir Zaki (2010), Assessing the effects of trade liberalization on wage inequalities in Egypt: A microsimulation analysis, *ERF Working Paper*, 555.

- Hertel, Thomas W., Robert A. McDougall, Badri Narayanan G. and Angel H. Aguiar (2008), GTAP 7 data base documentation – Chapter 14: Behavioral parameters, GTAP Resource, 2937.
- Hérault, Nicolas (2007), Trade liberalization, poverty and inequality in South Africa: A computable general equilibrium-microsimulation analysis, *Economic Record*, 83, 317–328.
- Hutton, John and Anna Ruocco (1999), Tax reform and employment in Europe, *International Tax and Public Finance*, 6, 263–287.
- ILO (1972), Employment, income and equality: A strategy for increasing productivity in Kenya, International Labour Office, Geneva.
- Jenkins, Glenn P. and Claude Montmarquette (1979), Estimating the private and social opportunity costs of displaced workers, *Review of Economics and Statistics*, 61, 342–353.
- Johnson, David T. and Peter B. Dixon (1999), Australian poverty quantified by a family based poverty index, *Economic Record*, 75, 103–114.
- Jomini, Patrick, Robert McDougall, Greg Watts and Philippa S. Dee (1994), The SALTER model of the world economy: Model structure, data base and parameters, Industry Commission of Australia, Canberra.
- Jung, Hong-Sang and Erik Thorbecke (2003), The impact of public education expenditure on human capital, growth, and poverty in Tanzania and Zambia: A general equilibrium approach, *Journal of Policy Modeling*, 25, 701–725.
- Keane, Michael P. and Richard Rogerson (2011), Reconciling micro and macro labor supply elasticities: A structural perspective, NBER Working Paper, 17430.
- Kim, Euijune and Kabsung Kim (2003), Impacts of the development of large cities on economic growth and income distribution in Korea: A multiregional CGE model, *Papers in Regional Science*, 82, 101–122.
- Kimball, Miles S. and Matthew D. Shapiro (2008), Labour supply: Are the income and substitution effects both large or both small?, NBER Working Paper, 14208.

- King, Robert G., Charles I. Plosser and Sergio T. Rebelo (1988), Production, growth and business cycles. I. The basic neoclassical model, *Journal of Monetary Economics*, 21, 195–232.
- Kingdon, Geeta G. and John Knight (2006), How flexible are wages in response to local unemployment in South Africa?, *Industrial and Labor Relations Review*, 59, 471–495.
- Kleven, Henrik Jacobsen and Claus Thustrup Kreiner (2006a), The marginal cost of public funds: Hours of work versus labor force participation, *Journal of Public Economics*, 90, 1955–1973.
- Kleven, Henrik Jacobsen and Claus Thustrup Kreiner (2006b), The marginal cost of public funds: Hours of work versus labor force participation, CEPR Working Paper, 5594.
- Krebs, Tom and Martin Scheffel (2010), A macroeconomic model for the evaluation of labor market reforms, ZEW Discussion Paper, 10-050.
- Krueger, Alan B. and Lawrence H. Summers (1988), Efficiency wages and the inter-industry wage structure, *Econometrica*, 56, 259–293.
- Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull and Giovanni L. Violante (2000), Capital-skill complementarity and inequality: A macroeconomic analysis, *Econometrica*, 68, 1029–1053.
- Lee, David S. (1999), Wage inequality in the U.S. during the 1980s: Rising dispersion or falling minimum wage?, *Quarterly Journal of Economics*, 114, 977–1023.
- Leibenstein, Harvey (1963), The theory of underemployment in densely populated backward areas, in: *Economic backwardness and economic growth*, chap. 6, John Wiley & Sons.
- Lejour, Arjan, Paul Veenendaal, Gerard Verweij and Nico van Leeuwen (2006), WorldScan – A model for international economic policy analysis, CPB Document, 111.
- Löfgren, Hans (2001), A CGE model for Malawi: Technical Documentation, TMD Discussion Paper IFPRI, 70, 1–43.

- Magnac, Thierry (1991), Segmented or competitive labor markets, *Econometrica*, 59, 165–187.
- Magnani, Riccardo and Jean Mercenier (2009), On linking microsimulation and computable general equilibrium models using exact aggregation of heterogeneous discrete-choice making agents, *Economic Modelling*, 26, 560–570.
- Maisonnave, Hélène, Bernard Decaluwé and Margaret Chitiga (2009), Does South African affirmative action policy reduce poverty? A CGE analysis, CIRPEE Working Paper, 09-36.
- McDonald, Ian M and Robert M. Solow (1981), Wage bargaining and employment, *American Economic Review*, 71, 896–908.
- McFadden, Daniel (1974), Conditional logit analysis of qualitative choice behavior, in: P. Zarembka (ed.), *Frontiers in Econometrics*, 105–142, Academic Press.
- Merz, Joachim (1993), Labor supply and recent German tax reform impacts – Behavioral response in a combined dynamic and static microsimulation model, Invited paper prepared for the Special Conference on Microsimulation and Public Policy of the International Association for Research in Income and Wealth, December 1993, University of Canberra.
- Ngai, L. Rachel and Christopher A. Pissarides (2008), Trends in hours and economic growth, *Review of Economic Dynamics*, 11, 239–56.
- Peichl, Andreas (2009), The benefits and problems of linking micro and macro models – Evidence from a flat tax analysis, *Journal of Applied Economics*, 12, 301–329.
- Peichl, Andreas, Hilmar Schneider and Sebastian Siegloch (2010), Documentation IZAΨMOD: The IZA Policy Simulation MODEL, IZA Discussion Paper, 4865.
- Peichl, Andreas and Sebastian Siegloch (2010), Accounting for labor demand in structural labor supply models, IZA Working Paper, 5350 (forthcoming in *Labour Economics*).
- Perroni, Carlo and Thomas F. Rutherford (1995), Regular flexibility of nested CES functions, *European Economic Review*, 39, 335–343.

- Perroni, Carlo and Thomas F. Rutherford (1998), Comparison of the performance of flexible functional forms for use in applied general equilibrium modelling, *Computational Economics*, 11, 245–263.
- Piggott, John and John Whalley (1985), *UK tax policy and applied general equilibrium analysis*, Cambridge University Press.
- Pissarides, Christopher A. (1990), *Equilibrium unemployment theory*, Basil Blackwell.
- Pissarides, Christopher A. (1998), The impact of employment tax cuts on unemployment and wages; The role of unemployment benefits and tax structure, *European Economic Review*, 42, 155–183.
- Pollak, Robert A. and Terence J. Wales (1987), Specification and estimation of nonseparable two-stage technologies: The Leontief CES and the Cobb-Douglas CES, *Journal of Political Economy*, 95, 311–333.
- Prescott, Edward C. (1986), Theory ahead of business-cycle measurement, *Carnegie-Rochester Conference Series on Public Policy*, 25, 11–44.
- Ramey, Valerie A. and Neville Francis (2009), A century of work and leisure, *American Economic Journal: Macroeconomics*, 1(2), 189–224.
- Rausch, Sebastian and Thomas F. Rutherford (2010), Computation of equilibria in OLG models with many heterogeneous households, *Computational Economics*, 36, 171–189.
- Robilliard, Anne-Sophie, François Bourguignon and Sherman Robinson (2008), Examining the social impact of the Indonesian financial crisis using a macro-micro model, in: François Bourguignon, Luiz A. Pereira da Silva and Maurizio Bussolo (eds.), *The impact of macroeconomic policies on poverty and income distribution: Macro-micro evaluation techniques and tools*, 93–118, Palgrave-Macmillan.
- Robilliard, Anne-Sophie and Sherman Robinson (2003), Reconciling household surveys and national accounts data using cross-entropy estimation, *Review of Income and Wealth*, 49, 395–406.

- Rojas-Romagosa, Hugo (2010), Modelling human capital formation in WorldScan, CPB Memorandum, 244.
- Rutherford, Thomas F. (1998), CES preferences and technology, in: GAMS MPSGE Guide, 89–115, GAMS Development Corporation, Washington, URL <http://www.gams.com/solvers/mpsge.pdf>.
- Shapiro, Carl and Joseph E. Stiglitz (1984), Equilibrium unemployment as a worker discipline device, *American Economic Review*, 74, 433–44.
- Slichter, Sumner H. (1950), Note on the structure of wages, *Review of Economics and Statistics*, 32, 80–91.
- Sørensen, Peter Birch (1997), Public finance solutions to the European unemployment problem?, *Economic Policy*, 12, 223–251.
- Sørensen, Peter Birch (1999), Optimal tax progressivity in imperfect labour markets, *Labour Economics*, 6, 435–452.
- Stifel, David C. and Erik Thorbecke (2003), A dual-dual CGE model of an archetypal African economy: Trade reform, migration and poverty, *Journal of Policy Modeling*, 25, 207–235.
- Thierfelder, Karen and Sherman Robinson (2002), Trade and the skilled-unskilled wage gap in a model with differentiated goods, International Food Policy Research Institute (IFPRI), TMD Discussion Papers, 96.
- van der Werf, Edwin (2008), Production functions for climate policy modeling: an empirical analysis, *Energy Economics*, 30, 2964–2979.
- van Soest, Arthur (1995), Structural models of family labor supply: A discrete choice approach, *Journal of Human Resources*, 30, 63–88.
- Whalley, John and S. Zhang (2004), A numerical simulation analysis of Hukou labour mobility restrictions in China, *Journal of Development Economics*, 83, 392–410.