

Mismatches on the Labour Market and Structural Unemployment

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, Prof.mr. M.J. Cohen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen
op donderdag 31 oktober 1996 om 14.00 uur

door

Henricus Bernardus Augustinus Bierings

Promotor:

Prof.dr. J. Muysken

Beoordelingscommissie:

Prof.dr. J.A.H. Maks (voorzitter)

Prof.dr. F.C. Palm

Prof.dr. H.R. Sneessens (Université Catholique de Louvain)

Mismatches on the Labour Market and Structural Unemployment/

H.B.A. Bierings.-

Rijswijk: Bierings, 1996.-

ISBN 90-9009835-6

Subject headings: structural unemployment/

macroeconomic disequilibrium models/

disequilibrium econometrics

© Harry Bierings, Rijswijk 1996

Wirklicher Fortschritt ist,
was Fortschreiten ermöglicht oder erzwingt
Brecht

To my parents and Birgit

Contents

| | |
|---|-----------|
| Acknowledgments | ix |
| 1 Scope and significance | 1 |
| 1.1 Introduction | 1 |
| 1.2 Theoretical notions..... | 8 |
| 1.2.1 Disequilibrium labour market analysis | 9 |
| 1.2.2 Job search theory | 12 |
| 1.2.3 Other theoretical notions | 13 |
| 1.3 Structure of the thesis..... | 15 |
| 2 The analytically and simulation based macro methods to measure mismatches | 21 |
| 2.1 Introduction | 21 |
| 2.2 The analytically-based variant | 22 |
| 2.3 The simulation-based variant | 34 |
| 2.4 Identification | 38 |
| 2.5 Review of empirical applications..... | 39 |
| 3 The validity of the (log-)normal density to describe micro market supply and demand | 49 |
| 3.1 Introduction | 49 |
| 3.2 Sensitivity of the aggregate employment function with respect to the stochastic distribution..... | 51 |
| 3.2.1 Evidence from the theory of stochastic processes: Gibrat's Law against Pareto's Law | 51 |
| 3.2.2 Other distributions..... | 55 |
| 3.3 Empirical results of nested testing | 58 |
| 3.4 Normal and log-normal distribution against general form..... | 66 |
| 3.4.1 Introduction | 66 |
| 3.4.2 Weighted normal employment function..... | 67 |
| 3.4.3 Weighted log-normal employment function | 68 |
| 3.4.4 Empirical results..... | 68 |
| 3.5 Concluding remarks | 71 |
| ANNEX 3A Derivation of the aggregate employment function for the bivariate Pareto distribution transformed to the origin | 75 |
| ANNEX 3B Derivation of the aggregate employment function for the bivariate logistic distribution | 77 |
| ANNEX 3C The aggregate minimum condition | 83 |
| ANNEX 3D Statistical series, definitions and sources | 85 |

| | |
|---|----------------|
| 4 The C.E.S. approach to measuring mismatches on the labour market and structural unemployment | 89 |
| 4.1 Introduction..... | 89 |
| 4.2 The one-level three-variate C.E.S. | 91 |
| 4.3 Unrestricted mismatches and the two-level C.E.S. employment function..... | 96 |
| 4.3.1 Introduction..... | 96 |
| 4.3.2 Relaxation of mismatch structure..... | 97 |
| 4.4 Empirical formulation, data and results..... | 102 |
| 4.5 Concluding remarks..... | 111 |
| ANNEX 4A Data..... | 115 |
| 5 The performance of the monte carlo pseudo maximum likelihood method | 117 |
| 5.1 Introduction..... | 117 |
| 5.2 Model specification and estimation..... | 118 |
| 5.3 Performance testing..... | 121 |
| 5.4 Monte Carlo evidence..... | 124 |
| 5.5 Concluding remarks..... | 127 |
| ANNEX 5A Computation of derivatives..... | 129 |
| ANNEX 5B Computation of covariance matrix for PML_1 and PML_2 (general results)..... | 133 |
| ANNEX 5C The bias-corrected MCPML variant..... | 135 |
| 5C.1 Theoretical results..... | 135 |
| 5C.2 Empirical results..... | 138 |
| 6 An econometric disequilibrium model for the Dutch economy estimated by Monte Carlo pseudo maximum likelihood | 141 |
| 6.1 Introduction..... | 141 |
| 6.2 The model..... | 142 |
| 6.2.1 Structure of the model..... | 142 |
| 6.2.2 Specification of the model..... | 143 |
| 6.2.3 Identification..... | 149 |
| 6.3 Estimating the model..... | 150 |
| 6.3.1 Data..... | 150 |
| 6.3.2 The practice of optimising..... | 152 |
| 6.3.3 Estimation results..... | 154 |
| 6.4 Concluding remarks..... | 162 |
| 7 Mismatches on the regional labour market (Micro evidence) | 169 |
| 7.1 Introduction..... | 169 |
| 7.2 Structure of unemployment by region: some facts..... | 171 |

| | |
|---|------------|
| 7.3 Job entrants and regional labour mobility: operational definitions and measurement | 172 |
| 7.3.1 Job entrants..... | 172 |
| 7.3.2 Regional mobility..... | 174 |
| 7.4 Data and estimation method..... | 180 |
| 7.5 Results | 183 |
| 7.6 Concluding remarks | 190 |
| ANNEX 7A Plausibility of mobility and job entrants data (Labour force survey)..... | 193 |
| 7A.1 Introduction | 193 |
| 7A.2 Quality of data on internal migration and causes of measurement error ... | 193 |
| 7A.3 Consequences of underestimation for the measurement of mobile job entrants..... | 196 |
| 7A.4 Reliability of job entrants data | 197 |
| ANNEX 7B Job entrants and mobility across RBA-regions by commuting distance and other characteristics..... | 199 |
| 8 Summary and conclusions | 201 |
| References | 209 |
| Nederlandse samenvatting..... | 221 |
| Curriculum vitae | 227 |

Acknowledgments

I find myself deeply indebted to a number of people who became friends in the process, for their contribution in finishing this book. I mention only a few. First of all, I want to express my gratitude to Joan Muysken, my thesis supervisor, for his expert guidance and never refraining commitment. Special thanks are also due to Christian Kerckhoffs and Koos Sneek who substantially contributed to Chapter 4, and to Chapters 5 and 6, respectively. I would like to thank Bernard Salanié for introducing me to the principles of the simulation method.

A number of people provided constructive criticisms: Hans Maks, Franz Palm, Henri R. Sneessens. Discussions with Wim Kloek, Peter Kooiman, Guy Laroque, Gerard Pfann, Erik de Regt, Jacques Thijssen, Maarten Vendrik and Arjen van Witteloostuyn also proved valuable. I am grateful to Mark Sanders and Jos Beckers for their computational assistance with respect to Chapters 4 and 7, and to Bart Pijls for his improvements on the lay-out.

This work could not have been finished without the support of Statistics Netherlands. Special thanks go to Cornelis van Bochove and Toon Theunissen. I, of course, remain solely responsible for the opinions expressed and the mistakes made. Rita Gircour corrected my English grammar. Jan de Loeff designed the cover.

Last but not least, I thank my parents and Birgit for their (emotional) support.

1 Scope and significance

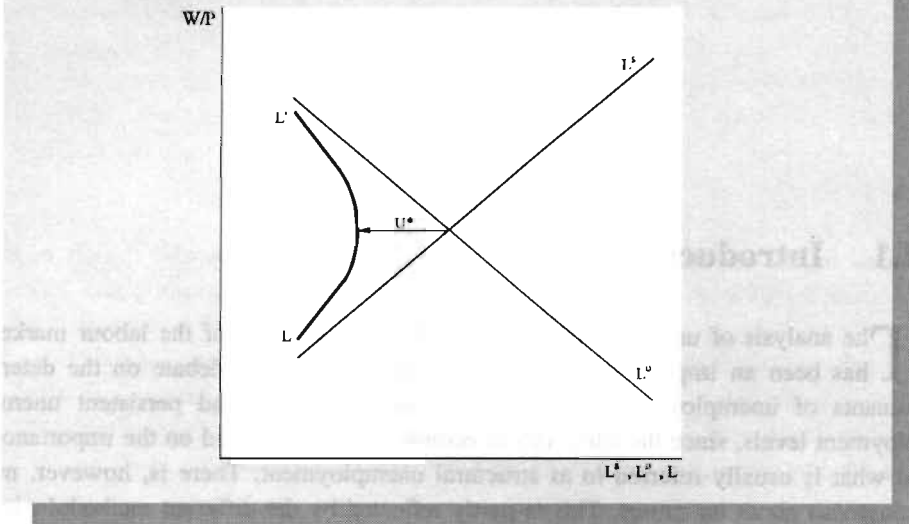
1.1 Introduction

The analysis of unemployment due to the malfunctioning of the labour market has been an important element in the macroeconomics debate on the determinants of unemployment. Following the years of high and persistent unemployment levels, since the early 1980s, economists have focused on the importance of what is usually referred to as structural unemployment. There is, however, no consensus about its causes. This is partly reflected by the different methodologies applied to study the topic. These differences in methodology can be traced back to the fundamental controversy in the profession of whether unemployment should be seen as an equilibrium or disequilibrium phenomenon. Before I will go into this matter any further, let me first address the question of the definition of structural unemployment.

Structural unemployment is that part of unemployment which is not the result of cyclical fluctuations: its causes lie in the structure of the economy, both on the supply side and on the demand side. More precisely, structural unemployment is defined as the level of unemployment which is consistent with (stock) equilibrium on the labour market, *i.e.* when supply equals demand. The latter condition should indicate the cyclical neutrality.

At the aggregate level, structural unemployment (U^*) reveals itself in the simultaneous occurrence of unemployment (U) and job vacancies (V). Besides the “natural” frictions which are due to job search, the inadequacies between supplied and demanded qualifications (also called mismatches) are at the root of the structural unemployment phenomenon. As I will show later on, these micro causes are linked to the macro concept of structural unemployment by way of the so-called transactions (or employment) function. Figure 1.1 shows the transactions function $L-L'$. This curve is situated at the left of the labour supply (L^s) and labour demand (L^d) curves. At each real wage rate (W/P), the transactions curve defines the level of actual employment. Structural unemployment, indicated by U^* then results as the

Figure 1.1 Labour supply, labour demand, employment and structural unemployment



distance between the employment curve, L and the cross-point at which labour supply, L^s equals labour demand, L^d .¹

Research on structural unemployment has moved in two directions, which differ with respect to the interpretation of the causes of structural unemployment. The two causes that may be distinguished, are:

1. The “job search component” which mainly reflects a temporary inadequacy between supply and demand; it is the result of workers and firms engaging in search activities to find or fill a job.
2. The “mismatch component” which is regarded as the more persistent inadequacy between supply and demand; it reflects the imbalance between labour supplies and demands at the micro market level, whatever the nature (skill, professional location *etc.*) of the imbalance.

¹ Another type of representation takes the form of the U (unemployment)- V (vacancies) curve. Taking $L^d = L + V$ and $L^s = L + U$, it is easily shown that when varying real wages the employment function in Figure 1.1 also reflects a hyperbolic U - V relationship. This relationship directly shows the simultaneous occurrence of unemployment and vacancies in the aggregate.

The first explanation is in the spirit of **equilibrium** labour market analysis. At the core of equilibrium analysis is the assumption that prices and wages are **perfectly flexible**. Equilibrium is associated with market clearing: supply and demand are coordinated such that either side of the market can sell or buy the quantity (of labour) that is associated with the equilibrium price or wage level. A situation of imbalance between supply and demand only occurs if the workers and firms fail to reach an optimum in a given period. The imbalance is, however, temporary since prices adjust to restore the equality between supply and demand. In the classic equilibrium model by Lucas and Rapping (1970), unemployment is due to job search and erroneous wage-price expectations. The second explanation is in line with the Keynesian disequilibrium paradigm. Disequilibrium is associated with **non-market clearing**. The basic assumption is that wages and prices are **fixed**, preventing markets to restore to a situation of full employment. The “equilibrium” and “disequilibrium” paradigms concerned with the explanation of unemployment have set the stage for studying the problem of structural unemployment.

Proponents of the equilibrium theory, such as Layard, Nickell and Jackman (1991), mainly approach the problem of structural unemployment from the job search point of view. In contrast to the initial Lucas and Rapping model, Layard *et al* do not think of mismatches as a negligible phenomenon, they only reject an explanation in the Keynesian sense of an arbitrarily rigid price, preventing competitors to sell and buy all they want to (p. 22). The alternative approach they suggest is based (in their own words) on a “exploratory” analysis and takes the form of a model with partially flexible wages. They assume that wage behaviour in a specific sector is caused primarily by unemployment in that sector, rather than unemployment in some other “leading” sector. Under these assumptions a mismatch indicator results which is proportional to the variance of the relative unemployment rates. However, when studying the causes of the huge unemployment rise in Europe, they argue that the explanation must not be sought in a rise of mismatches but more in a decline of **search effectiveness** of the unemployed and firms. Search effectiveness accounts for:

“all that affects the speed with which the unemployed find jobs - that is, the efficiency with which information about vacancies is transmitted, the time and effort the unemployed devote to job search, their “choosiness” with regard to vacancies and job offers, and the recruitment practices of employers.”
(Layard, Nickell and Jackman, 1991, p. 216).

Specific issues they address is **why** workers have become more choosy about taking jobs, or firms about filling vacancies. They find empirical support for the search argument in explaining the rise of unemployment in Europe. The results indicate that the increase in structural unemployment since the 1970s should be attributed to a decline in job search effectiveness and not to increasing mismatches.

On the contrary, researchers advocating the disequilibrium analysis, attribute a key role to mismatches. Major contributors in this field such as Lambert (1990) and Mehta and Sneessens (1990) study structural unemployment from the mismatches perspective. Their method of analysis starts from the perception that the labour and the goods markets are heterogeneous aggregates, composed of a multitude of actually unobserved micro markets, exhibiting different supply and demand schedules. These micro markets operate efficiently apart from a search component: on each market, supply or demand is fully absorbed by transactions. By implication a micro labour market either has vacancies or unemployment. It reflects the so-called *min-condition* stating that the level of transactions (here employment) is equal to the minimum of supply and demand.

In the standard application to the model, there are no interactions between micro labour markets. Hence, an excess supply or demand on one market is not absorbed by another market. Taking supply and demand across micro markets to satisfy some kind of density function, the method allows the computation of a mismatch indicator of micro demands and supplies by adding up across micro markets. This indicator equals the standard deviation of micro excess supplies across markets. It is directly linked to the rate of structural unemployment. The method known as the "smoothing by aggregation approach" was developed by Muellbauer (1978) and Malinvaud (1980).² Innovations to the standard model account for spillover effects from the goods market onto the labour market in a specific way. With it another typical Keynesian element is included in the analysis. It allows unemployment to arise as a result of the inability of firms to sell all the goods they want. Other relatively new contributions have integrated elements of job search into the analysis. Models along the lines of "smoothing by aggregation" have been developed by Kooiman and Kloek (1979), Kooiman (1984), Gouriéroux and Laroque (1985), Sneessens and Drèze (1986), Meersman (1987), Gagey, Lambert and Ottenwaelter (1988), Bierings and Muysken (1988), Lambert (1988), Bierings and Sneek (1989), Laroque and Salanié, (1989), Stalder (1989), Franz and König (1990), Lambert (1990), Padoa Schioppa (1990), Sneessens and Mehta (1993), Muysken, Bierings and De Regt (1994), Muysken and De Regt (1994) and Franz and Smolny (1994).

This study examines the problem of "structural mismatches" in relation to the measurement of structural unemployment. The method of analysis stays within the Keynesian framework of disequilibrium models. I exclude the also interesting, more equilibrium-oriented approaches (see Layard *et al*, 1991) of the discussion. In view of the sluggish price and wage responses, and because disequilibrium seems to be the normal state of affairs in the real world, I believe that the disequilibrium framework promises a better understanding. This, however, does not imply that I ignore

² The word "smoothing" refers to the fact that the aggregate transactions are continuous in L^s and L^d , as opposed to the canonical model which defines transactions to be the minimum of these variables.

all aspects associated with the equilibrium notion. Especially, elements of job search will be incorporated in the analysis. It is indeed possible to have the two logics in one model without having to compromise on the idea of micro markets in disequilibrium.

The purpose of the study is to analyse the importance of mismatches and structural unemployment, and the relation between them. This will be done theoretically as well as empirically. To examine mismatches and structural unemployment one can distinguish a **macro-oriented** approach along the lines of “smoothing by aggregation”, and a **micro-oriented** approach for analysing structural problems on the spot. Since the approaches are complementary they are both used in this study.

Two types of model approaches are identified within the **macro approach**. They differ in the rigor with which the aggregation over micro markets is done. The literature shows that expressions of aggregate transactions are analytically not tractable when micro labour markets modelled together with spillovers from the goods market are added up. One solution is to choose some convenient approximation to arrive at analytical expressions. This is one way of dealing with the problem. Another is to insist on the rigor in modelling, and thus to accept analytically intractable outcomes by simulating the aggregation over micro markets.

Both methods have advantages and disadvantages. The analytical model uses some convenient approximations to make the derivation of analytically tractable expressions of transactions. An example is the specific distribution of micro market supply and demand which has to be assumed (in most cases the log-normal). It is unknown, however, to what extent the assumptions generating these expressions influence the measurement of the mismatches and of structural unemployment. This leaves room for suspicion. Against this view it could be argued that the analytical expressions of transactions have a clear and intuitively appealing economic interpretation; the mismatch indicator is simply a parameter in the aggregate function and the link with structural unemployment is easily established. Moreover at the expense of rigor in modelling, the simple fact that analytical expressions are also obtained creates some kind of flexibility. Since the mismatch indicator is an identified parameter in the aggregate function, it is easily endogenised, allowing the causes of mismatches and structural unemployment to be studied. Another interesting application is the incorporation of “job search” elements.

The simulation-based approach, developed by Laroque and Salanié (1989) uses a technique to estimate a model where a complicated network of endogenous forces at the micro level operates on the mismatch parameters. Although the approach makes the estimation of a complete micro model with spillover effects possible, it has the disadvantage that the influences on the mismatch variances cannot be strictly identified. It is also clear that flexibility in specification on the one hand requires parsimony on the other hand, otherwise the model estimation would become econometrically unfeasible. A more practical problem is that the estimation of the model with the aforementioned features is rather intricate. It poses large

burden on performance testing. The testing is, however, necessary because simulation is not exact. All in all, the simulation-based approach entails a relatively laborious way to get estimates of mismatches and structural unemployment in the end. A final disadvantage of the method is that formally testing the assumptions is undeveloped for this type of model. Nevertheless the advantages are very clear. The simulation approach allows the most general model specification, that is, with labour supply and labour demand modelled simultaneously with goods demand and goods supply, all at the micro market level. Moreover, the method can *a priori* fit any distribution (log-normal or others) of micro supply and demand, which makes it potentially powerful. The treatment of supply and demand as latent unobserved variables is consistent with the underlying economic theory, which is also in terms of unobservable variables. Another advantage of the simulation-based method is that it allows in theory the identification of the variances of the aggregate supply and demand variables. These aggregate quantities operate as common factors on micro market supply and demand. In the analytically-based methods, these variances are not separately identified.

Research in the field of the analytical model broadly moves along two different empirical lines within which sub-variants may be distinguished:

- A. Single equation estimation of the transactions function (variant type 1);
- B. Estimation of transactions function within a model (n-equation model approach):
 - using business survey data as regime classification information (variant type 2);
 - no regime classification information used, *i.e.* supply and demand endogenous (variant type 3).

In most applications of variant type 1, the aggregate transactions function is of the constant elasticity of substitution (C.E.S.) type.³ Within this variant, aggregate labour supply and labour demand are observed from the data; labour supply is measured as the sum of employment and unemployment and labour demand as employment plus vacancies. Critical assumptions of this data-based method refer to the data quality and the treatment of goods market influences. As to the data quality: earlier studies used unemployment and vacancy data from the labour exchange office. By now it is clear that these data contain systematic measurement biases. Particularly data on job vacancies fail because firms only partly register their job vacancies and they often neglect to report the filling or cancellation of the vacancies. From the literature it is known that measurement errors generate biases in the parameter estimates and standard errors. More recent contributions use data that are corrected for these measurement biases, thereby considerably improving on the quality of the vacancy data. Yet, the problem remains that vacancies are hard to measure (van Bastelaer and Laan, 1994, p. 14-15), so that the “good quality of data”

³ “C.E.S.” refers to constant elasticity of substitution, which is a familiar concept in production function theory. In the present context this elasticity has no special meaning.

assumption still constitutes the critical part of the method. The second aspect of the method is its preoccupation with the labour market only: there is no rigorous treatment of goods market influences on the labour market in the sense of spillovers. Recent examples of the variant type 1 studies are Bierings and Muysken (1988), Muysken, Bierings and De Regt (1994), Muysken and De Regt (1994) and Franz and Smolny (1994).

The n-equation model approach defines a simultaneous system of (aggregate) supply, demand and transactions relationships for the goods and labour market. In the variant type 2 method, a system of reduced form equations is obtained by substitution. The transactions function again is of the C.E.S. type. The model equations are log-linear and contain the so-called regime proportions besides transactions and the variables explaining labour supply and labour demand. These regime proportions represent the percentage of firms in a specific regime. In theory there are four regimes, with each regime characterised by its typical configuration of excess supply or excess demand on the goods and labour market (Malinvaud, 1977). When the regime proportions are observed from business survey data the resulting model can be estimated using full information maximum likelihood techniques. The business survey data gathered in a number of European countries contain information about the number of firms claiming that their production plans are hindered by "lack of demand", "lack of capacity", "lack of sufficient labour" or "other reasons". Applications of this method can be found in Kooiman (1984), Gagey, Lambert and Ottenwaelter (1988), Lambert (1988), Stalder (1989), Franz and König (1990) and Lambert (1990).

The variant type 3 method treats the supply and demand variables as endogenous and unobserved, and a simultaneous system is estimated. Some assumptions are made to keep the model analytically tractable. See Gouriéroux and Laroque (1985), Sneessens and Drèze (1986), Padoa Schioppa (1990), Sneessens and Mehta (1993). The simulation-based variant does not distinguish different subvariants. Below, it will be indicated as variant type 4.

A critical assumption which applies to all the macro-oriented methods is that micro markets supply and demand are treated as constructs for which no data are available.⁴ Supplies and demands across micro markets are simply described by some kind of density function. Although in most applications the log-normal density function is advocated because of its "reasonable" properties (Lambert, 1988, p. 17), the basis for such an assumption is rather volatile. As a matter of fact in the litera-

⁴ The statistical construct as such has proved its relevance for the study of mismatches and structural unemployment in aggregate time series analyses. Often it is also an unavoidable assumption when data on individuals fail. For most researchers this is the normal state of affairs which naturally follows from the obligation of statistical offices to protect the privacy of individual respondents. In the Netherlands, the statistical office has offered researchers the opportunity to use the data files under highly regulated conditions ensuring the privacy of respondents.

ture it has not been scrutinised empirically and theoretically. An important question remains how robust the results on mismatches and structural unemployment are with respect to the distribution function chosen. The issue will be dealt with in this study.

In addition, the mere fact that micro market supply and demand is considered as a statistical construct calls for **micro-oriented** studies augmenting the results of the macro approach. Besides interesting possibilities of comparing the results between the approaches, the micro approach is complementary. I mention one issue which cannot be examined within the macro approach (analytical model), while it can within a micro approach. It concerns the mobility behaviour - the mirror side of mismatches - of people when currently facing an excess supply situation on one market. Do they move to another market where excess demand prevails? In a micro study one may attempt to identify for instance, markets for a given profession or region to tackle the problem. In the macro approach as such this is impossible. As a matter of fact, when the standard micro model is extended with intertemporal spillovers, the aggregate outcome is observationally equivalent to the standard model; the spillover coefficients are not identified (see Lambert, 1988). Later on I will go into the details of this issue.

I know of only a few empirical micro studies analysing mismatches and their causes from individual (mobility) behaviour, for the situation of the Netherlands. They mainly concentrate on the regional dimension, that is on regional labour market mobility (see for instance, Evers and van der Veen, 1985, Hartog, Mekkelholt and van Ophem, 1987 and the references quoted in Gorter, 1993). My contribution goes in this direction too. I will explore the factors hampering regional mobility of the unemployed workers. I distinguish job characteristics, but also economic factors such as search time, and the relative vacancy situation.

Now that the scene has been set for this study, I will discuss the theoretical aspects of disequilibrium theory and other theoretical notions contributing to the understanding of mismatches and structural unemployment in the next section. The final section describes how this thesis is organised.

1.2 Theoretical notions

This section will place the present study in its proper perspective. It should, however, be stressed that not all contributions are equally relevant for this study. Some of them are discussed merely for the sake of completeness. The disequilibrium theory of non-market clearing provides the theoretical foundation of this study. Certain elements from job search theory are borrowed to integrate in the analysis of markets in disequilibrium. In the empirical application, the combined approach of disequilibrium and job search arguments allows to test the importance of mismatches and search to explain structural unemployment. As such this touches

on the controversy in the profession about the role of mismatches and search unemployment. The human capital notions on structural unemployment are envisaged in the micro study of this thesis. It concerns the influence of education on mobility behaviour.

The institutional and radical considerations described below are of fairly marginal interest to this study, except that they contribute to the understanding of structural unemployment.

1.2.1 Disequilibrium labour market analysis

The key feature of the macro approach is the assumption of non-clearing micro markets on which wages and prices are fixed. Wages and prices do not work as equilibrating forces and a deficit on one market is not filled by a surplus on the other market. Hence, disequilibria and thus mismatches are the natural or chronic state of affairs. These assumptions are at the core of Keynesian disequilibrium analysis.⁵ Let me provide a brief historical sketch of the main ideas of what has become known as the microfoundations of Keynesian macroeconomics.

The central message of Keynes' *The General Theory* (1936) is that the general Walrasian equilibrium is a special rather than a general case.⁶ Market clearing prices occur only by accident. The subsequent theoretical contributions started with the reappraisal of Keynesian macroeconomics and proceeded to general equilibrium analysis with fixed prices. One of the predecessors who studied disequilibrium macro models with microeconomic foundations was Patinkin (1956). He recognised the importance of spillover effects in the context of firms' behaviour: if goods markets are expected not to clear, that is firms are rationed on the goods market, firms take into account the associated quantity constraints in their decision on labour demand. Clower (1965) works out this idea by introducing the "dual decision hypothesis" for households and the associated distinction between notional (Walrasian market clearing) and effective (non-Walrasian rationing) quantities.

An agent's notional demand on a given market is the demand which he would express on that market in the absence of any quantity constraints on his transactions in any other markets. But agents encounter quantity constraints on their

⁵ Throughout this book "disequilibrium" is used in the fix-price non-market clearing interpretation. It is not referred to as a state in which forces are at work to restore the system of equilibrium, which interpretation would be in the equilibrium theory flavour.

⁶ The concept of a Walrasian equilibrium is a sequel to Walras' (1874) seminal contribution to economic theory. Its characteristic feature is that prices are assumed to be perfectly flexible. That is, a Walrasian equilibrium is associated with the assumption of market clearing. The Walrasian equilibrium method is based upon the assumption of price tâtonnement. That is, transactions only occur after the declaration of the equilibrium price vector.

transactions. An agent's effective demand on a given market is the demand which he will express taking into account the constraints on his transactions in other markets. The "dual decision hypothesis" states that households take the quantity constraints they face in the labour market into account while deciding on the demand for commodities. Leijonhufvud (1968) elaborated on the non-Walrasian argument by pointing out that the essential feature of non-Walrasian equilibrium is that quantities move faster than prices. That is, in the short-run quantity adjustments dominate price movements. Essentially, this is the assumption of fixed prices.

The work of Clower and Leijonhufvud implied a denial of Walras's Law which is the cornerstone of Walrasian general equilibrium analysis. It states that if the values of excess demands sum to zero, so that a negative excess demand in one market must have a counterpart in a positive excess demand in some other market, there will be a tendency for relative price adjustment in the direction of eliminating the excess supplies and demands. The assumed Walrasian auctioneer who knows all the "notional" excess demands can change the prices accordingly before any transactions are executed. The Keynesians emphasise that no *tatōnnement* on prices and transactions takes place at prices other than those of a Walrasian equilibrium, then in any market where there is non-zero excess demand, agents on the "long" side of the market will face quantity constraints. Inter-market spillovers from goods to labour markets and *vice versa* are central to the microfoundation of macroeconomics debate between Walrasian and Keynesian economics. It is exactly the aspect of inter-market spillovers to which this study attaches great importance. In particular, spillovers are believed to be crucial in describing mismatches and structural unemployment.

The first full formalisation of the non-market-clearing approach was in Barro and Grossman (1976). Barro and Grossman brought Patinkin's firms and Clower's household rationing together in one model. Spillovers can occur from the labour to the goods market and *vice versa*. In this way they were able to explain the pro-cyclical pattern of real wages and employment, which cannot be derived from Walrasian economics. This is also envisaged in the application of the simulation-based method.

The general perspective of the Barro and Grossman's model introduces the opportunity to focus on the macroeconomic implications of microeconomic responses to quantity constraints. Barro and Grossman's set up, in common with most of their immediate successors is the Hicksian fix-price method (Hicks, 1965). The follow up of Barro and Grossman's temporary non-Walrasian equilibrium model is particularly manifest in Benassy's (1975, 1977) and Drèze's (1975) generalisations to a multi-market setting. In an n -market model, effective demand in any one market typically depends on constraints encountered in all others, so an agent's effective demand vector is determined component by component. They have shown the existence of quantity constrained equilibria (or the fixed point of multilateral spillover effects). Malinvaud (1977) elaborates on Barro and Grossman's two-market model by indicating that the features of the temporary

equilibrium in the goods and labour market determine specific macroeconomic regimes, depending on whether buyers or sellers face rationing in the goods and the labour market.

Disequilibrium theory enables to understand macroeconomic phenomena from the concepts of effective demand and involuntary unemployment. However, it is impossible to deny the obvious problem with Keynesian fix-price theory, that is the assumption of exogenous prices. This assumption can only apply to a transitory state. The subsequent theoretical contributions are concerned with the question of how to make prices endogenous. Although prices are straightforwardly modelled as endogenous in a mathematical sense in a supply and demand model, it is very difficult to model these as the outcome of an economic process (see Dixon and Rankin, 1995, p. 3).⁷ There are contributions by, for instance, Grandmont and Laroque (1976), Negishi (1978) and Hart (1982), introducing imperfect competition (sellers can manipulate prices not anticipated by buyers) into the macroeconomic model. Hart's work can explain how the monopolist will set his price in excess of marginal cost, and the union will restrict employment in order to raise wages above the marginal disutility of labour, so that involuntary unemployment can arise. The explanation of endogenous prices that falls into the category of imperfect competition is associated with "implicit contract theories" (Azariadis, 1975 and Bailey, 1974), "efficiency wages theories" (Yellen, 1984, Shapiro and Stiglitz, 1984). Implicit contract theories aim to explain the rigidity of real wages, and hence the existence of an equilibrium level of unemployment. The argument is based on the assumption that workers are more risk-averse than firms, and so optimal risk sharing dictates that the firm ensures the workers by keeping wages fixed over the business cycle. The efficiency wage model belongs to the general class of monopsony labour market models. Under the assumption that firms set wages under conditions of asymmetric information, turnover costs and so on, it provides an equilibrium. Other explanations originate from "insiders-outsiders" theory and "dual markets theory" (Salop, 1979) (see also Section 1.2.3). Despite the promising theoretical developments of imperfect competition theories to explain prices, a rigorous treatment of endogenous prices in the empirical models is still complicated. This study will not solve the problem and maintain exogenous fixed prices. I accept this as a shortcoming.

⁷ This problem is also inherent in the Walrasian model: it is based on the assumption that all agents act as price-takers, though it requires an auctioneer to make the prices, ensuring that prices adjust to bring supply and demand into balance.

1.2.2 Job search theory⁸

Job search is usually considered the other component of structural unemployment. In the view of job search theorists most unemployment results from a voluntary action of an individual seeking for work. In the standard job search model, the unemployed worker first decides on whether or not (s)he will search, and second with what intensity. On the other hand, a firm aiming to fill a vacant job decides on the (intensity of) search behaviour using a comparable cost-benefit analysis.⁹

A crucial element in search models is that an unemployed worker has incomplete information on job vacancies and is unaware of the distribution of the wages offered. An unemployed worker, however, has some idea of the distribution of wage rates which are within reach, but the specific wage-job offer combinations are nonetheless unknown. Workers are well-informed about the probability of locating a job vacancy and the search cost involved. The optimal search intensity is determined by confronting the marginal cost of an additional search effort with its marginal returns. This strategy implies that a wage offer is accepted if and only if the wage is higher than (or equal to) their cut-off or reservation wage. This is the wage at which a worker is indifferent between accepting a job or remaining unemployed. Search continues if the wage offer is below the reservation wage.

Changes in structural unemployment result from changes in search effectiveness, which in turn mainly stem from changes in the cost-benefit trade-off. In empirical work, several factors that may influence search effort are considered. The most important are the **replacement** (or benefit-income) **ratio** and **unemployment duration**. Personal characteristics such as age, marital status educational qualification are usually used as controls on search effectiveness. In the present study, these variables are incorporated. *i.e.* in the applications to the analytical model and the micro approach.

In the "standard" job search model, the benefit effects on search effectiveness come about through the reservation wage. The reservation wage rises with the level of unemployment benefits with the prediction that increases in unemployment compensation lead to a reduced probability of making the transition from unemployment to employment. However, evidence (measured by the benefit

⁸ The description of search theory below is supply-side oriented, with the role of employees search captured in the concept of search effectiveness. The literature, however, provides some rare theoretical contributions on search theory from the demand side (Mortenson, 1970 and Jovanović, 1979). Here the expected marginal product of a worker takes over the role of the expected marginal wage in the supply-side models. An integrated approach that confronts supply and demand side search has not been undertaken as far as I know. For the study of structural unemployment, this would provide a more comprehensive framework.

⁹ An extensive survey of job search models can be found in Lippman and McCall (1976) and Pissarides (1985).

unemployment duration elasticities) suggests that this effect is rather small (Layard *et al.*, 1991). Further evidence shows that the unemployed receive and reject very few job offers (Jones, 1989 for Britain, Holzer, 1988 for the USA and van den Berg, 1990 for the Netherlands). Unemployed workers remain unemployed, mainly because they get so few job offers and not because they reject so many. Another school of thought views the effect of unemployment benefits on job offers from another perspective. In this view unemployment compensation increases the resources (and thus search effectiveness) devoted to search, and hence it increases the probability of return to work (see Ben-Horim and Zuckerman, 1987). The number of applications by a worker depends positively on the level of income support. Moreover, benefits could increase the ties to the labour force through information or incentive effects. For example, benefit payments are contingent on recipients ability to demonstrate active job search. For Britain, the influence of unemployment benefits on search effort finds little support for a benefit effect on the number of search methods within a given week. As to **unemployment duration**, Layard *et al.* (1991) consider three main effects on search effort. That is, the effect on job search, the effect on worker's skill, motivation and morale, and the effect of job screening and employer perceptions. It is common to assume that search activity declines with unemployment duration ("discouragement effects"). Related concepts are "hysteresis" effects in unemployment, insider-outsiders effects as in Blanchard and Summers, 1986) or skill-loss as in Pissarides (1992). Employers discriminate against the long-term unemployed, because they believe that they are unmotivated and lack relevant skills and working habits. In this study this aspect is encompassed in the application of the analytically-based method through the modelling of labour supply. Another type of argument explaining why employers are reluctant to hire people arises from employment protection schemes. As to the empirical relevance of duration on structural unemployment, Layard *et al.* (1991) attribute a large role to the effect of unemployment duration. For Britain (period: 1978-1986) they find that the rise in unemployment duration explains one-third of the increase in structural unemployment (p. 39). In the application of the analytically-based method in the present study, I consider the related question of how much the long-term unemployed contribute to structural unemployment.

1.2.3 Other theoretical notions¹⁰

For reasons of completeness, I will describe two other theoretical notions contributing to the understanding of structural unemployment. These notions are not very well suited to be incorporated in an empirical model. The two notions are the

¹⁰ See: Bierings and van Witteloostuyn (1989).

radical and institutional notions and human capital considerations. They are mainly concerned with the mismatch part of structural unemployment.

Radical and institutional notions are in line with Keynesian thinking, because of their reliance on persistencies of market imperfections. Human capital theory has both Keynesian and Walrasian aspects. Radical and institutional economists attribute persistencies of market imperfections to institutional and social structures and emphasise the role of chronic inequalities and discrimination. The notion of fragmented labour markets is common. This is in line with Keynesian thinking, although price rigidities are not considered to be of great importance in explaining market imperfections. Four strands of thought may be distinguished: the institutional, dual, segmented and Marxist labour market theories. They emphasise the role of government regulations, unions, large industries, secondary and internal labour markets *etc.* as possible causes of low mobility.

Institutional theories focus on the consequences of institutional factors on wage setting and employment. Special attention is devoted to the factors that prevent the perfect functioning of the labour market. Kerrs's (1954) notion of the "balkanisation" of labour markets is an early example of institutional theory. Government regulations and unions are held responsible for the transformation of full competition labour markets into an organised power system.¹¹

While institutional theories consider a wide range of obstructing factors, the other segmentation notions concentrate on the supposed dichotomy on the labour market. The dual labour market theory is related to the study of large companies (Galbraith, 1967). Heterogeneity of the labour market results from qualitative differences between leading and peripheral industries (Bluestone, Murphy and Stevenson, 1976). Döringer and Piore (1971) distinguish primary from secondary labour markets. The Marxist view emphasises the struggle between the classes (Wright, 1979). What the three approaches have in common is their denial of the presence of mobility between two compartments of the labour market. While taking the evident conceptual differences into account, the implications of the supposed dichotomy and non-mobility in segmented labour market theory is quite uniform. Jobs in the primary sector are characterised by employment stability and promotion from within (an "internal" labour market). Moreover, they are covered by statutory employment protection, and are part of the legal economy. Jobs in this sector offer some prospect of promotion and involve a substantial element of general or specific training. Jobs in the secondary sector involve less stability, little training and poor promotion opportunities. These jobs may be temporary or casual, dead-end jobs or part of the informal economy; they may be homeworkers uncertain of continuing work, lacking control over work in the absence of worker protection. The dichotomy on the labour market is expressed in the limited upward mobility from secondary to primary labour market. Besides, labour market entrants mainly start in secondary jobs, in particular due to the internal labour market existence.

¹¹ As to the role of union power see Hirsch and Addison (1986) for an overview.

Since Oi (1971) introduced the concept of quasi-fixed labour, the distinction between the internal and the external labour market has been commonly used within labour economics. Internal labour markets may arise from efficiency considerations (such as adequate information exchange). Döringer and Piore (1971) note that the internal labour market restricts mobility from one to the other labour market segment because internal discriminating selection processes impede the entrance to the primary market.

Human capital theory has its origin in the theory of firm's investment.¹² It is concerned with decision behaviour of (future) suppliers of labour in relation to long-term variables (Becker, 1964). Agents optimise their utility by maximising the net discounted value of returns of investment in human capital. In general the optimising procedure takes account of restrictions in time, budget and skills. The decision is about investments in general skills and investments in specific skills. A specific investment is only productive for one firm, while general investments have a wider use. Future labour supply is concerned with investments in human capital before and after labour market entrance. Before entrance, general investment decisions are made with respect to profession and length of education. The cost of investment are the expected income losses, whereas benefits relate to future income. Future income depends on future market conditions such as future occupational and educational demand. After entering the labour market, investments in additional education schemes are usually required. Future labour supply should also decide on the allocation of time for work, education and leisure (Killingsworth, 1983, p. 305).

Human capital theory yields some insights that could explain structural unemployment: educational unemployment arises from supply-demand mismatches by occupation and educational level. Suppliers of labour may have insufficient education or old skills. Changes in structural unemployment could be attributed to changes in the factors determining supply of and demand for human capital. One example is the employer's demand for upgrading skill levels due to technological progress. Upgrading skills requirements give rise to mismatches because it is not satisfied by supply. A more indirect effect results from the increase of cost in human capital investment. The educational aspect will be encompassed in the micro study of the present work.

1.3 Structure of the thesis

This thesis brings together three candidates for explaining structural mismatches (and structural unemployment) on the labour market due to the existence of heterogeneous markets. They include the earlier mentioned macro variants type 1 (based on single equation estimation with observed data) and type 4 (based on simulation) assuming aggregate data, and a micro (cross-sectional) approach based on individ-

¹² For an overview of human capital theory see Blaug (1976) and Siebert (1985).

ual data.¹³ For the variants type 1 and 4, a macro model based on micro markets in disequilibrium is formulated and tested empirically for the Dutch labour market. The micro study concerns the estimation of a cross-sectional model on labour market mobility behaviour in the Netherlands. Combined, the three variants provide a coherent view of the **extent** and the **nature** of mismatches (and the implied structural unemployment) for the Netherlands. More specifically, the two macro approaches shed light on the extent of mismatches on the labour market. The variant type 1 and the micro approach provide information on the factors influencing labour mismatches, in a time series or cross-sectional context, respectively. The variant type 1 and the micro approach relate to the **nature** of mismatches. Estimation of the corresponding models serves the empirical goal of this thesis. The evidence should be of interest to policymakers dealing with the problem of unemployment.

The empirical results also provide insight in the methodological issue of the sensitivity of the estimation results on the extent of mismatches (and structural unemployment) for the macro method chosen. Or to put it differently: Is the measurement of the level of mismatches effectively influenced when the assumptions underlying the variant type 1 method are relaxed, resulting in the simulation-based method - that is allowing for:

- labour supply and demand to be unobserved quantities,
- spillovers at the micro market level and
- aggregation (through simulation) without imposing simplifying assumptions.

Let me explain the contents of the thesis in more detail. It is divided into eight chapters, schematically shown in Figure 1.2. Chapter 2 gives a description of the analytically and the simulation based approach. Preceding Chapter 4, which will elaborate on the C.E.S. method of the variant **type 1** approach, Chapter 3 scrutinises its main underlying assumptions. This is done in a theoretical as well as an empirical way. In particular, the sensitivity of the distribution assumption to describe micro markets is considered. I use several candidate distributions to investigate the consequences for the aggregate functional form and for the measurement of labour market mismatches and the implied structural unemployment.

In Chapter 4, an extension of the C.E.S. transactions function of variant type 1 is presented. It has been developed earlier, parallel by Gagey, Lambert and Ottenwaelter (1988) and Bierings and Muysken (1988). The generalisation relates to the allowance for goods market spillovers on the labour market. As a result a C.E.S.-type transactions function is obtained which describes aggregate employment in terms of two labour demand variables (Keynesian and capacity demand) and labour supply, and two mismatch parameters. The mismatch parameters reflect the so-called “mismatch on the labour market” and “mismatch on the goods market” (*i.e.*

¹³ Of the analytically based methods, I consider neither variant type 2 using extraneous information from business survey data on the extent of disequilibrium nor variant type 3. Below, I will explain why.

between the composition of the demand for goods and the composition of the available production capacity). The chapter provides new theoretical and empirical insights. In the literature, the relationship between the C.E.S. mismatch parameters and the underlying supply/demand dispersions has not yet been established. This is a very uncomfortable situation if one wants to endogenise these mismatches. In practise the choice of factors influencing the mismatch parameters has been led by *ad hoc* arguments. Chapter 4 attempts to fill this gap, as the relationship between micro dispersions and the C.E.S. parameters is made explicit. This paves the way for endogenising the mismatch parameters for the empirical analysis more adequately. Finally, several candidate factors to endogenise the mismatch parameters in the empirical analysis are considered such as capacity utilisation rate, competitiveness, the replacement ratio, search duration and a regional mismatch indicator.

The model specification further relies on Muysken and De Regt (1994). A very important feature of the model is that search elements are integrated in the framework of micro markets in disequilibrium. To be more precise, at the micro market level, the model specification of Muysken and De Regt allows for the level of transactions to be smaller than the minimum of supply and demand as a result of job search duration. This is the so-called matching inefficiency **within** markets which is to be distinguished from the mismatch between markets. The aggregate C.E.S. transactions function then is not simply a function of aggregate supply and demand and a mismatch parameter, but also of search unemployment which in turn depends on job search duration. At this point it is worth noting that most of the literature has dealt with the static model, which has proved to be a shortcoming in the empirical analysis. Muysken and De Regt cope with this problem by introducing dynamics through job search. Another feature of the model is the correction of labour supply for the long-term unemployed. By referring to the some well-known hysteresis arguments, they assume that the matching efficiency of the long-term unemployed is lower than for short-term unemployed; in the C.E.S. employment function this is accounted for by a correction factor on labour supply. The higher the rate of long-term unemployment the smaller effective labour supply is. The possible implications of allowing for a correction for the long-term unemployed become clear in a study of Bean and Gavosto (1989) for the UK. When estimating the model without renormalising labour supply for the long-term unemployed they find a growing mismatch over the sample. When the model is estimated without the correction, mismatches between markets are constant over time. Of particular interest in the analysis of Muysken and De Regt is that it allows for determining the relative importance of mismatch in structural unemployment. This touches on the debate about the relative importance of the rise in mismatch on the labour market since the 1970s. The empirical results presented in this thesis for the Netherlands quantify the importance of (i) search duration and (ii) mismatch in a strict sense (that is only the qualification dimension) and (iii) the long-term unemployed.

The weaker aspects of the data-based method are its reliance on the “good quality of data” assumption and its preoccupation with the labour market, that is a less systematic treatment of spillovers. Besides, simplifying assumptions are made to preserve the simple closed form of the C.E.S. structure.

In this thesis the analytically-based methods which use business survey results to identify the extent of disequilibrium are not considered. They fall into the variant **type 2**, using business survey data as regime classification information within a n -equation model with transactions. The main reason for not considering this variant stems from the disappointing findings by Kooiman (1984) for the Netherlands. He considers the information content of the business survey data for the Netherlands too weak to yield reliable estimates. Moreover, the interview data do not adequately reflect the theoretical entities they should measure. A related problem, which is specific for the Netherlands¹⁴, is that firms are allowed to indicate whether they experienced “no constraints” on production. This requires a transformation of the common four theoretical regimes as in Malinvaud (1977), and additional assumptions about the conditions under which a firm chooses a specific answer. Laroque and Salanié (1992) have similar reservations as well as other ones:

“It requires some faith to assimilate the reported statistics (the business survey data) to the proportions of demand or supply constrained micro markets of the theory ... There are a host of reasons why these regime proportions may be affected with measurement errors, and/or do not fit into the theoretical model. First, labour being a quasi-fixed factor, at least in European countries, the decision to hire is a long-run decision, whereas the determination of employment has a short-run flavour in the model... Finally, the available surveys bear on the manufacturing sector, but they have sometimes been used in models of the whole economy.”

(Laroque and Salanié, 1992, p. 10-11, parentheses mine).

The variant **type 3** variant is not considered here since it has essentially the same point of departure as the variant **type 4**, that is the simulation-based approach using latent supply and demand variables and no use of extraneous information. Compared to the simulation-based approach, it is more restrictive in the specification of the model and the estimation technique used.

Essentially, the simulation-based method developed by Laroque and Salanié (1989) is an estimation procedure which is capable of estimating the macro model specified at the micro level with the above mentioned properties (supply and demand are latent variables, spillovers from the goods on the labour market and *vice versa*). The method which uses approximate estimators from pseudo maximum likelihood estimation is regarded as suitable when exact methods are not tractable, as is the case with the likelihood function and the transactions function (see Laroque and Salanié, 1989). It is worth mentioning that the estimation of a macro model allowing for disaggregation into micro markets was considered an untractable

¹⁴ For Belgium, for instance, this is not the case.

problem until recently. The analytical and statistical problems are formidable even for the simple canonical aggregate disequilibrium model.

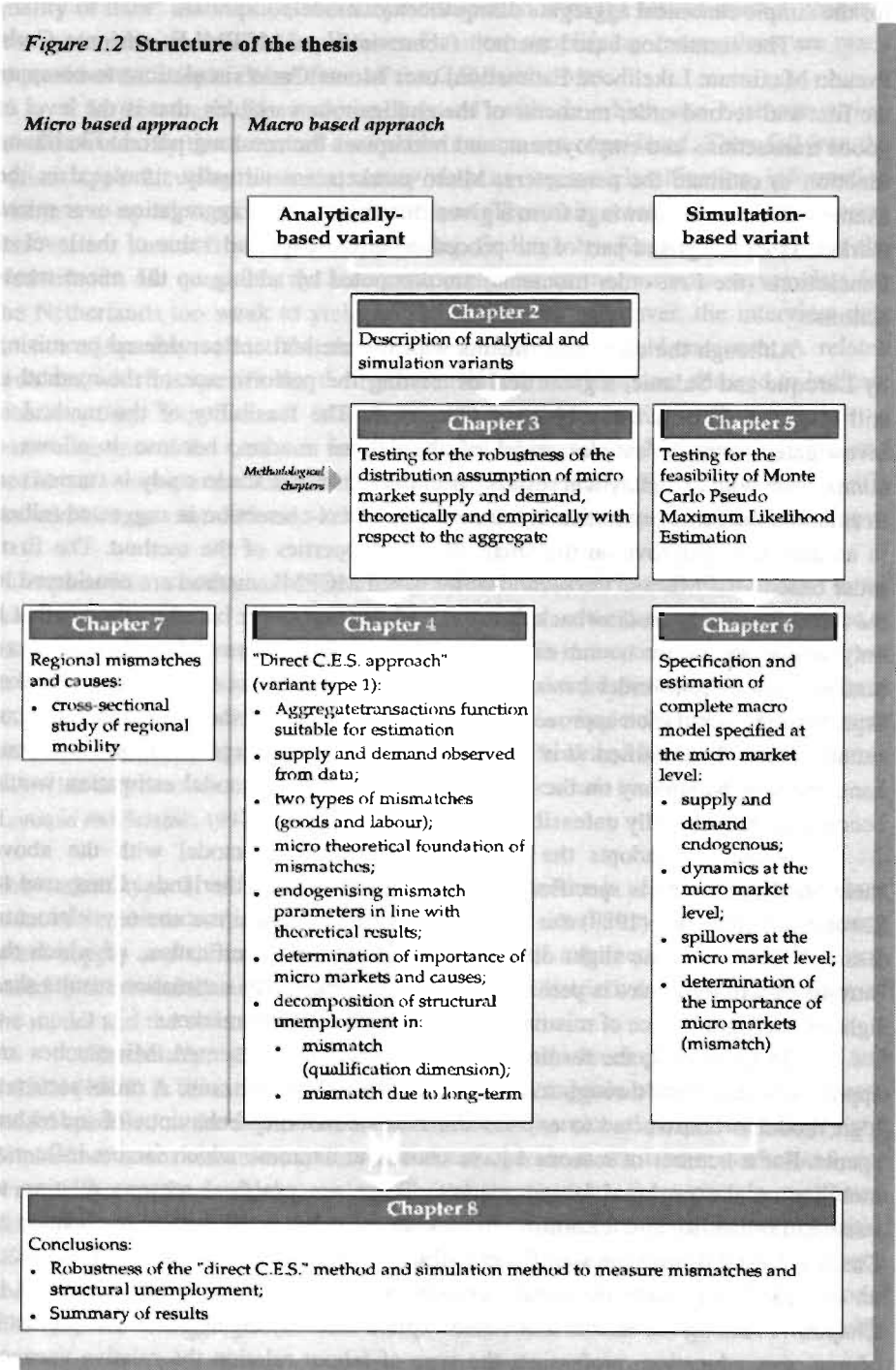
The simulation-based method (abbreviated as MCPMLE = Monte Carlo Pseudo Maximum Likelihood Estimation) uses Monte Carlo simulations to compute the first and second order moments of the endogenous variables, that is the level of goods transactions and employment, and maximises the resulting pseudo likelihood function to estimate the parameters. Micro markets are virtually simulated in the Monte Carlo part as drawings from a given distribution. The aggregation over micro markets is an integrated part of the procedure as the expected value of the level of transactions (the first order moments) are computed by adding up the micro transactions.

Although the early experiments with the method are considered promising by Laroque and Salanié, a great deal of testing the performance of the method is still required. This is the subject of Chapter 5. The feasibility of the method is investigated using a simple model of the labour market, because it allows a comparison with the analytical results. Moreover, a Monte Carlo study is carried out to examine the finite properties of the method. A bias-correction is suggested (albeit in an annex) to improve on the small sample properties of the method. The first-order based MCPML and the second-order based MCPML method are considered in the feasibility study. A drawback, however, of the first-order based method is that it only allows an upper bound estimate of the level of mismatches due to micro markets. The second-order based method does not have this shortcoming. Weaker aspects of the simulation approach are that the influences on the mismatch variances cannot be strictly identified. It is also clear that flexibility in specification on the one hand requires parsimony on the other hand, otherwise the model estimation would become econometrically unfeasible.

Chapter 6 adopts the technique and a macro model with the above mentioned properties is specified and estimated for the Netherlands. Compared to Laroque and Salanié, (1989) the optimising procedure contains some new elements. Besides there are some slight differences in the model specification, of which the introduction of dynamics is perhaps the most important. The estimation results shed light on the significance of mismatches on the Dutch labour market.

In Chapter 7, the results of a micro study are presented. Mismatches are approached indirectly through mobility, the other side of the coin. A cross-sectional logit model is constructed to explain the lack of mobility behaviour of individual agents. For a number of reasons I have chosen to examine which factors influence mobility on the **regional** labour market. There are practical reasons relating to statistical reliability and feasibility. In another respect it is the logical implication of Chapter 4 establishing the significant influence of the regional mismatch indicator on the C.E.S. mismatch parameter. Details on this matter are further discussed in Chapter 7. Among the factors that could explain mobility behaviour are the controls of age, sex, education, profession, the type of labour relation the relative vacancy situation of the region, and search duration. Chapter 8 contains the conclusions.

Figure 1.2 Structure of the thesis



2 The analytically and simulation based macro methods to measure mismatches

2.1 Introduction

This chapter describes the essentials of the two macro methods of markets in disequilibrium to measure mismatches and structural unemployment. As explained in Chapter 1, both methods fit the tradition of “smoothing by aggregation”.¹⁵ They take as a starting point that aggregate markets, like the labour and the goods market, consist of a multitude of efficient operating micro markets with different supply/demand schedules. Micro market supply and demand are stochastic concepts, in the sense that they follow from a random drawing of a given distribution. It is furthermore assumed that markets are efficient, *i.e.* transactions fully absorb demand or supply (the minimum rule). Under these conditions, this generates aggregate transactions (output or employment) which are always smaller or equal to the minimum of aggregate supply and demand, and thus allowing vacancies and unemployment to coexist. In equilibrium, the inefficiency in the

¹⁵ Other methods to measure mismatches or structural unemployment are outside the scope of this book for different reasons. These are, for instance, the earlier mentioned variance method of relative unemployment rates across sectors, professions, etc. (Layard, Nickell and Jackman, 1991) and traditional U-V analysis (see for a review of empirical U-V analysis de Neubourg, 1987). On the one hand, traditional U-V analysis which defines an empirical convex relationship between aggregate unemployment and aggregate vacancies is not suited to identify the extent of mismatches: the derived level of structural unemployment (= equilibrium unemployment at $U=V$) is a composite measure of mismatches and job search. Besides, there are two problems with U-V analysis which should be mentioned here. First, the specification of the U-V curve seems rather *ad hoc*: “In no sense is the vacancy rate the determinant of the unemployment rate” (Holden and Peel, 1975, p. 251). Second, inaccurate estimates are obtained when unemployment systematically exceeds vacancies for a longer period (Muysken and Meijers, 1986). On the other hand, the variance method to measure mismatches does not fit into the framework of fix-price markets in disequilibrium: Layard *et al* (1991) derive the relevant mismatch indicator from a model where price and wage decisions are taken separately on different segments of the economy (see also Chapter 1).

aggregate determines the degree of structural mismatch and thus of structural unemployment. Naturally, the situation of an efficient aggregate market (zero mismatch) evolves as a special case. In its basic variant, the analytically-based model describes aggregate transactions as a continuous non-linear macro function in terms of aggregate supply and demand and a mismatch parameter. I discuss this variant in Section 2.2. In the simulation-based variant, simulation experiments provide estimates of the aggregate variables and the mismatch indicator that otherwise could not be evaluated. Section 2.3 explains this method in more detail.

Before going into the algebra of the methods, let me illustrate by way of introduction, the principle of “smoothing by aggregation” graphically Figure 2.1 displays the situation for two markets. In each market employment (l_1 and l_2) is determined by the minimum of demand and supply. At real wage rate, w^* , the first market exhibits vacancies (v_1), the second unemployment (u_2). Aggregate supply and demand follow from adding up supply and demand at w^* . Aggregate employment, l being the sum of l_1 and l_2 is below aggregate labour supply and demand. Varying w^* and computing the corresponding aggregate supply, demand and employment levels yields the aggregate curves as shown in the figure. With the number of markets increasing, the employment curve increasingly takes the form of a smooth function; the number of kinks are proportional to the number of markets. In the analytically-based variant, this relationship is made explicit, whereas in the simulation-based variant it is implicit.

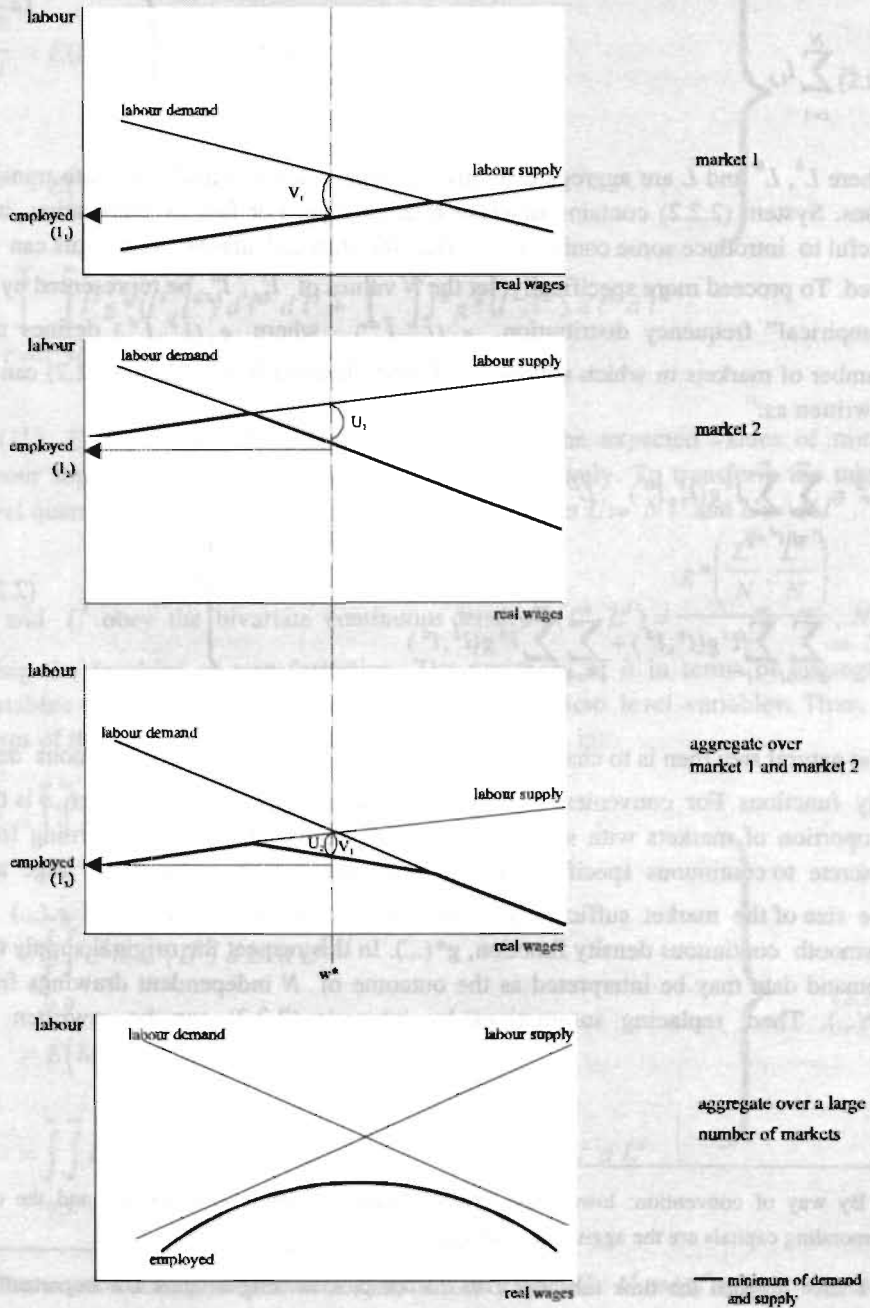
2.2 The analytically-based variant

The analytically-based variant is in effect an empirical strategy which involves searching for a readily interpretable and relative simple functional form for aggregate transactions. In this section, the analytically-based variant is worked out formally for a simple model of the economy. First, an outline of the method in a formal theoretical setting is given. Then, the broad lines are described along which the empirical applications have developed.

The theoretical description concentrates on the labour market only. Consider a model where there are N different micro markets, indexed by $i, i=1, \dots, N$. At date t , supply and demand on a micro market are $l_{i,t}^s, l_{i,t}^d$. On each market, the minimum rule is assumed, so that transactions (=employment) on market i should satisfy:

$$l_{i,t} = \text{Min} \left(l_{i,t}^s, l_{i,t}^d \right) \quad (2.2.1)$$

Figure 2.1 Aggregation over markets



At the aggregate level, supply and demand and the quantity traded simply follow by adding up:¹⁶

$$\left. \begin{aligned} L_t^S &= \sum_{i=1}^N l_{i,t}^s & L_t^D &= \sum_{i=1}^N l_{i,t}^d \\ L_t &= \sum_{i=1}^N l_{i,t} \end{aligned} \right\} \quad (2.2.2)$$

where L^S , L^D and L are aggregate supply, aggregate demand and aggregate transactions. System (2.2.2) contains **discrete** relationships. For further elaboration, it is useful to introduce some continuity, so that the standard algebra apparatus can be used. To proceed more specifically, let the N values of $l_{i,t}^s$, $l_{i,t}^d$ be represented by an "empirical" frequency distribution, $g_t(\bar{l}^s, \bar{l}^d)$, where $g_t(\bar{l}^s, \bar{l}^d)$ defines the number of markets in which supply is \bar{l}^s and demand is \bar{l}^d . Then, (2.2.2) can be rewritten as:¹⁷

$$\left. \begin{aligned} L^S &= \sum_{l^s=0}^{\infty} \sum_{l^d=0}^{\infty} l^s g(l^s, l^d) & L^D &= \sum_{l^d=0}^{\infty} \sum_{l^s=0}^{\infty} l^d g(l^s, l^d) \\ L &= \sum_{l^s=0}^{\infty} \sum_{l^d > l^s}^{\infty} l^s g(l^s, l^d) + \sum_{l^d=0}^{\infty} \sum_{l^s \geq l^d}^{\infty} l^d g(l^s, l^d) \end{aligned} \right\} \quad (2.2.3)$$

The natural step then is to change from frequency distributions to continuous density functions. For convenience, first normalise $\bar{g}(\dots) := g(\dots)/N$, with $g(\dots)$ is the proportion of markets with specified levels of supply and demand. Turning from discrete to continuous specifications, N is assumed to be sufficiently large and the size of the market sufficiently small to justify the approximation of $\bar{g}(\dots)$ by a smooth continuous density function, $g^*(\dots)$. In this respect the original supply and demand data may be interpreted as the outcome of N independent drawings from $g^*(\dots)$. Then, replacing summations by integrals (2.2.3) can be rewritten as:

¹⁶ By way of convention: lower case symbols reflect micro level quantities and the corresponding capitals are the aggregate equivalents.

¹⁷ I have dropped the time subscript t in this chapter, as long as it is not important for understanding.

$$\begin{aligned}
 \frac{L^S}{N} &= E(l^s) = \int_{l^s=0}^{\infty} \int_{l^d=0}^{\infty} l^s g^*(l^s, l^d) d l^d d l^s \\
 \frac{L^D}{N} &= E(l^d) = \int_{l^s=0}^{\infty} \int_{l^d=0}^{\infty} l^d g^*(l^s, l^d) d l^s d l^d \\
 \frac{L}{N} &= E(\text{Min}(l^s, l^d)) = \\
 &= \int_{l^s=0}^{\infty} \int_{l^d > l^s}^{\infty} l^s g^*(l^s, l^d) d l^d d l^s + \int_{l^d=0}^{\infty} \int_{l^s \geq l^d}^{\infty} l^d g^*(l^s, l^d) d l^s d l^d
 \end{aligned}
 \tag{2.2.4}$$

$E(l^s)$, $E(l^d)$ and $E(\text{Min}(l^s, l^d)) (= E(l))$ are the expected values of micro labour supply and demand and employment, respectively. To transform the micro level quantities into the macro quantities, L^s and L^d , let $L^s := N l^s$ and $L^d := N l^d$.¹⁸

L^s and L^d obey the bivariate continuous density $h(L^s, L^d) = \frac{g^*\left(\frac{L^s}{N}, \frac{L^d}{N}\right)}{N^2}$, N^2 being the Jacobian of transformation. The variances of h in terms of aggregate variables are N^2 times the variances of g^* of the micro level variables. Thus, in terms of the macro variables system (2.2.4) transforms in:

$$\begin{aligned}
 L^S &= \int_0^{\infty} \int_0^{\infty} L^s h(L^s, L^d) d L^d d L^s \\
 L^D &= \int_0^{\infty} \int_0^{\infty} L^d h(L^s, L^d) d L^s d L^d \\
 L &= E(\text{Min}(L^s, L^d)) = \\
 &= \int_0^{\infty} \int_{L^s}^{\infty} L^s h(L^s, L^d) d L^d d L^s + \int_0^{\infty} \int_{L^d}^{\infty} L^d h(L^s, L^d) d L^s d L^d
 \end{aligned}
 \tag{2.2.5}$$

¹⁸ The notation distinguishes between L^s and L' (and L^d and L^d'): L^s and L^d refer to aggregate supply and demand, whereas L' and L^d' are the micro levels transformed to the aggregate.

To facilitate further derivations, a tractable function for $h(\dots)$ needs to be specified. Most applications chose the log-normal density, for its evidently attractive properties (see Lambert, 1988). Denying for the moment the appropriateness of this distribution assumption (this is subject of Chapter 3), in the descriptive part here, it suffices to proceed with the log-normality assumption. For expository purposes, a change in representation of the working of a micro market is useful (see Lambert, 1988):¹⁹

$$\left. \begin{aligned} \ln l_i^s &= \lambda^s + \eta_i^s & \ln l_i^d &= \lambda^d + \varepsilon_i^d \\ \ln l_i &= \text{Min}(\ln l_i^s, \ln l_i^d) \end{aligned} \right\} \quad (2.2.6)$$

$$\text{with } \begin{pmatrix} \eta^s \\ \varepsilon^d \end{pmatrix} \approx N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_{\eta^s}^2 & \sigma_{\eta^s \varepsilon^d} \\ \sigma_{\eta^s \varepsilon^d} & \sigma_{\varepsilon^d}^2 \end{bmatrix} \right] = N(0, \Sigma^*) \text{ and}$$

the “average” supply and demand are:

$$\overline{l^s} = e^{\lambda^s + \frac{1}{2} \sigma_{\eta^s}^2} \quad \overline{l^d} = e^{\lambda^d + \frac{1}{2} \sigma_{\varepsilon^d}^2}$$

This representation captures the concept of a micro market, more clearly. In the model, the natural logarithm of supply and demand on market i is expressed as a drawing from the bivariate normal distribution of (η^s, ε^d) corrected for the fixed scalars, λ^s, λ^d . Thus, the only source of variation in the disturbance terms is due to micro markets.²⁰ Notice that in model (2.2.6) the covariance matrix, Σ^* , is unrestricted. It is, however, to be expected that a larger than average supply goes hand in hand with a larger than average demand ($\sigma_{\eta^s \varepsilon^d} > 0$).

As before, rescaling to aggregate quantities, the model (2.2.6) becomes:

¹⁹ I deliberately use two symbols η and ε for the error terms to have consistency in presentation with Chapters 5 and 6.

²⁰ These disturbances may therefore be called **spatial**, or **idiosyncratic** (or individual) shocks (Laroque and Salanié, 1992). In the empirical model where aggregate supply and demand are functions of other variables, it is useful to distinguish in addition, errors with respect to these in principle unobserved variables (see Section 2.3). They may be called **aggregate** shocks. Of course, when aggregate supply and demand are correctly observed from the data, one can abstract from this type of error. This is exactly what is done in Chapter 4 and is quietly assumed here.

$$\left. \begin{aligned} \ln L_i^s &= \Lambda^s + \eta_i^s & \ln L_i^d &= \Lambda^d + \varepsilon_i^d \\ \ln L_i &= \text{Min}(\ln L_i^s, \ln L_i^d) \end{aligned} \right\} \quad (2.2.7)$$

with $\Lambda^s = \lambda^s + \ln N$, $\Lambda^d = \lambda^d + \ln N$, (η^s, ε^d) are unchanged and:

$$E(L^s) = e^{\Lambda^s + \frac{1}{2} \sigma_{\eta^s}^2} = N \bar{l}^s = L^S$$

$$E(L^d) = e^{\Lambda^d + \frac{1}{2} \sigma_{\varepsilon^d}^2} = N \bar{l}^d = L^D$$

The variables in the model are rescaled to the aggregate, with L^S and L^D , the expected values of labour supply and labour demand with respect to the micro disturbances.

The specification for aggregate employment is obtained by evaluating the expectation of $E(\text{Min}(L^S, L^D))$. It shows:²¹

$$L = \Phi^S L^S + \Phi^D L^D \quad (2.2.8)$$

with: $\Phi^S := \Phi(x - \frac{1}{2} \sigma_{d-s})$ (Φ is the standard cumulative density)

$$\Phi^D := \Phi(-x - \frac{1}{2} \sigma_{d-s})$$

$$x = \frac{1}{\sigma_{d-s}} \ln \frac{L^D}{L^S}$$

$$\sigma_{d-s}^2 = \text{var}(\varepsilon^d - \eta^s) = \sigma_d^2 + \sigma_s^2 - 2\sigma_{ds} = \sigma^{*2}$$

It is clear from (2.2.8) that only σ^* is identified from the three independent elements of Σ^* . This parameter is crucial in this book, since it represents the mismatch (or dispersion) between supply and demand across micro markets and defines a one-to-one relationship with the structural unemployment rate. Recall that structural unemployment, U^* , is defined as the unemployment level consistent with

²¹ See also Kooiman and Kloek (1979). The subscripts η and ε in the σ 's are dropped for convenience.

(stock)equilibrium on the labour market, *i.e.* at $L^D = L^S$. Then from (2.2.8) the rate of structural unemployment, u^* ($=U^*/L^S$) follows as:

$$u^* = 1 - 2\Phi\left(-\frac{1}{2}\sigma^*\right) \quad (2.2.9)$$

Note that when the mismatch, σ^* , approaches zero, u^* approaches to zero. This is the situation that corresponds with the aggregate minimum model in which aggregate transactions are determined by the minimum of aggregate supply and demand.

The employment function in (2.2.8) expresses L as a highly non-linear function in terms of aggregate supply, demand and a mismatch parameter. To arrive at a simpler expression, the usual way to proceed is to approximate the standard cumulative normal densities in (2.2.8) by the proportions of markets in excess demand or excess supply (Sneessens, 1983 and Lambert, 1988). These proportions, respectively denoted by P^j with $j=S,D$ may be calculated as follows:

$$P^S = P(L_i^d > L_i^s) = \int_0^\infty \int_{L^s}^\infty \left(\frac{L^s}{L}\right) h(L^s, L^d) dL^d dL^s \quad (2.2.10)$$

$$P^D = P(L_i^s \geq L_i^d) = \int_0^\infty \int_{L^d}^\infty \left(\frac{L^d}{L}\right) h(L^s, L^d) dL^s dL^d \quad (2.2.11)$$

$$\text{with } P^S + P^D = 1 \quad (2.2.12)$$

The factors $\left(\frac{L^s}{L}\right)$ and $\left(\frac{L^d}{L}\right)$ are the weighting factors. In a particular regime, they represent the share of market specific transactions in total transactions.²² Combining (2.2.10) and (2.2.11) with (2.2.8), P^j (with $j=S,D$) may be rewritten as:

$$P^S = L^S \frac{\Phi^S}{L} \quad (2.2.13)$$

$$P^D = L^D \frac{\Phi^D}{L} \quad (2.2.14)$$

The choice of an approximation formulae for Φ^j by some function of P^j takes advantage of functional similarities between Φ^j and P^j . More specifically, the following applies (Sneessens, 1983):

²² In the unweighted variants of eq. (2.2.10) and eq. (2.2.11) these factors are equal to one.

- Since L is smaller than $\text{Min}(L^S, L^D)$, it follows from (2.2.13) and (2.2.14) that P^j is always larger than the associated $\Phi^j (j=S,D)$;
- For limiting behaviour of L^S and L^D , P^j and Φ^j converge:

$$\lim_{\frac{L^j}{L^k} \rightarrow 0} \Phi^j = \lim_{\frac{L^j}{L^k} \rightarrow 0} P^j = 0$$

$$\lim_{\frac{L^k}{L^j} \rightarrow 0} \Phi^j = \lim_{\frac{L^k}{L^j} \rightarrow 0} P^j = 1$$

for $j,k=S,D$ and $j \neq k$.

- The first (partial) derivatives of Φ^j and P^j with respect to L^S and L^D have the same sign over the entire range.

Given these properties, it therefore seems reasonable to approximate Φ^j by the following function of P^j (Sneessens, 1983):²³

$$\Phi^j \approx P^j \beta_1^{1+\frac{1}{\beta_1}} = G^j(P^j) \quad (j = S, D; \beta_1^j > 0) \quad (2.2.15)$$

where the restriction on β_1 implies that P^j is always larger than G^j , G^j and P^j converge when P^j approaches 0 or 1. If it is subsequently required that this approximation should be exact in $L^S=L^D$, where $\Phi^S=\Phi^D=\Phi(-\frac{1}{2}\sigma^*)$ and hence $P^S=P^D=1/2$ then:

$$\beta_1^j = \beta_1 = \frac{-\ln 2}{\ln 2 + \ln \Phi(-\frac{1}{2}\sigma^*)} \quad (j = S, D) \quad (2.2.16)$$

Substituting P^j implied by (2.2.15) and (2.2.16) for Φ^j in (2.2.8) taking account of (2.2.10) - (2.2.12), the following well-known Constant Elasticity of Substitution (C.E.S.) form is easily obtained:²⁴

²³ This book distinguishes three different types of C.E.S. employment functions: the one-level two-variate specification (this chapter), the one-level three-variate specification (Chapter 4) and the two-level three-variate specification (also Chapter 4). I will use the subscripts in β to identify these specifications from one another. Subscript 1 is for the first and subscript 2 for the second variant. The subscripts 3 and 4 are used for the third (two-level) variant mentioned.

²⁴ Only far from the equilibrium level $L^S=L^D$ the approximation will slightly understate the true level of aggregate transactions, L (Kerckhoffs, 1992, p. 9). It is worth mentioning that it is possible to derive the C.E.S. transactions function as an exact result if supply and demand across micro markets satisfy identical, mutually independent Weibull distributions (see Den

$$L \approx \left((L^S)^{-\beta_1} + (L^D)^{-\beta_1} \right)^{-\frac{1}{\beta_1}} \quad (2.2.17)$$

The corresponding relationship for P^S (is also P^D) shows:

$$P^S = \frac{1}{1 + \left(\frac{L^D}{L^S} \right)^{-\beta_1}} \quad (2.2.18)$$

From (2.2.17) and (2.2.18), it is simply verified that for finite β_1 , the proportions, P^j ($j=S,D$) are always in the range $<0,1>$ and transactions L are always below the minimum of aggregate supply and demand.

For empirical studies which aim at explaining the mismatch between supply and demand in aggregate transactions (2.2.17) the theoretical relationship between β_1 and σ^* (c.f. eq. 2.2.16) is crucial. Usually, we have some theory about the factors that influence mismatches and in which direction. But how can these influences be modelled through the C.E.S. exponent β_1 in the transactions equation (2.2.17)? From (2.2.16) it follows that when σ^* approaches zero, β_1 approaches infinity and C.E.S. transactions collapse to the aggregate minimum model (or the Leontieff model, to use the production function terminology).²⁵ When σ^* approaches infinity, β_1 approaches zero. Evidently, there is some kind of inverse relationship between mismatches, σ^* , and β_1 . To make this relationship more explicit, let β_1 be the relevant variable to be approximated by a Taylor expansion around $\sigma^*=0$ in (2.2.16). Then, the following expression between β_1^{-1} and σ^* obtains:

$$\frac{1}{\beta_1} \approx \frac{\sigma^*}{\ln 2\sqrt{2}\pi} \approx 0.57556 \sigma^* \quad (2.2.19)$$

This implies that β_1^{-1} is almost perfectly proportional to the underlying structural mismatch, σ^* , for any reasonable (see Table 2.1 below) value of σ^* . Thus, one can

Broeder, 1983). The “mismatch” in this case, however, is a rather obscure concept. The consequence is that mismatches cannot be endogenised adequately on theoretical grounds (Kerckhoffs, 1992, p. 11).

²⁵ The proportions, P^j ($j=S,D$) approach either to zero or one; for example, P^S approaches zero if $L^S > L^D$, and to one if $L^S < L^D$.

simply endogenise mismatches by replacing β_1^{-1} in (2.2.17) by a function of the variables which seem relevant in explaining mismatches.²⁶

Because of the relationship between σ^* and the rate of structural unemployment, u^* , the relationship between u^* and β_1^{-1} resembles (2.2.19). More specifically, from (2.2.17) follows:

$$u_1^* = 1 - 2^{-\frac{1}{\beta_1}} \approx \frac{1}{\beta_1} \ln 2 \quad (2.2.20)$$

So, finally through (2.2.19) it appears that the structural unemployment rate is defined in terms of the mismatch between supply and demand, σ^* . In percentage changes the interactions between β_1 , σ^* and u_1^* are simple: an x% decline in β_1 implies an x% rise in σ^* and an x% rise in u_1^* (albeit for reasonable values of σ^* : see Table 2.1 below).

This completes the discussion of the theoretical part of the analytically-based method. Before turning to the empirical applications, however, one point should be made. Equation (2.2.20) measures the structural unemployment rate as derived from the C.E.S. transactions function. The latter, however, is an approximation of "exact" transactions as expressed by (2.2.8). Therefore, it is important to know to what extent possible errors due to approximation influence the measurement of u_1^* due to mismatches. Table 2.1 shows the values for u_1^* computed from the exact and the approximation formula. It follows that for values of σ^* up to 0.15, the approximation is exact; for higher values of σ^* - which are unrealistically high given the results reported in empirical studies (see Section 2.4) - they are somewhat less exact. Table 2.1. also shows the values computed from (2.2.16) and its approximation (2.2.20). Here again the differences are negligible. Hence, the approximation formulas may be safely used in empirical applications.

I now come to review the empirical applications within the analytically-based method. To get the essentials of the different contributions right, the description is limited to the standard model with only labour supply and demand. Actual applications might be more general, as due account is taken of both the labour market and the commodity market and its interrelations (spillovers). Section 2.4 contains a more or less complete overview of the specific contributions. Three estimation strategies appear in the literature (recall Chapter 1). They differ to the extent in which extraneous information on the particular regime (excess supply or

²⁶ In the extended (more realistic) alternative version of eq. (2.2.17) of Chapter 4 (Section 4.3), the relationship between the C.E.S. parameter(s) and mismatch(es) is not as straightforward.

Table 2.1 Structural unemployment rate, u_1^* and C.E.S. parameter β_1 for different values of the mismatch indicator σ^* and alternative specifications

| | u_1^* (eq. 2.2.9) | u_1^* (eq. 2.2.20) | β_1 (eq. 2.2.16) | β_1 (eq.2.2.19) |
|--------------|---------------------|----------------------|------------------------|-----------------------|
| σ^* | | | | |
| 0.001 | 0.0004 | 0.0004 | 1733 | 1737 |
| 0.010 | 0.0040 | 0.0040 | 173 | 174 |
| 0.050 | 0.0200 | 0.0200 | 34 | 35 |
| 0.100 | 0.0398 | 0.0398 | 17 | 17 |
| 0.150 | 0.0598 | 0.0598 | 11 | 12 |
| 0.500 | 0.1974 | 0.1990 | 3 | 3 |
| 1.000 | 0.3830 | 0.3990 | 1 | 2 |

demand) is processed. The three strategies, have one thing in common. In a way, they are all attempts not to get involved in the complicated estimation of a system in which supply and demand are latent variables and stochastic functions of other variables and transactions, of course, observed with functional form as above. Then the main problem is that all information on the regime has to be inferred from transactions only (see for instance, Kooiman and Kloek, 1979 for the problems met in trying to estimate such a system)²⁷.

Variant 1 uses data on aggregate labour supply and demand.²⁸ As such it relies on the good quality of data on L , U and V , and thus on L^S and L^D . Under these conditions, the prevailing regime is known at each date and by adding a disturbance term to L , the mismatch parameter (or β_1^1) can be relatively easily and consistently estimated by standard non-linear least squares. The more recent contributions in the field fully recognise the vulnerability of the "good quality of data" assumption (Muysken, 1994). Hence, a great deal of attention is devoted to correct, in particular, the vacancy series for systematic biases.

In the variant type 2 studies, the regime classification information is obtained in an indirect way from the proportions of markets in excess supply or demand, P^j ($j=S, D$). These regime proportions, usually observed from business surveys, are also assumed to be perfect data. Following Lambert (1988), one can combine the closely related approximations for L , P^S and P^D in order to eliminate one of

²⁷ An additional complication arises due to the non-linearity of aggregate transactions in the aggregate supply and demand errors, precluding the analytical computation of the likelihood function (see Section 2.4 about identification).

²⁸ For the sake of simplicity, the contributions estimating eq. (2.2.8) directly, with observed data on supply and demand are placed in this category.

the latent variables (L^S or L^D) and thus allowing estimation to be carried through with standard techniques. To explain the procedure, first note that the C.E.S. transactions function (2.2.17) may be rewritten as:

$$L = L^D (P^D)^{\frac{1}{\beta_1}} \quad (2.2.21)$$

and a similar expression for L obtains in terms of L^S and P^S . Then it follows immediately that:

$$\ln L = \ln L^S + \frac{1}{\beta_1} \ln P^S + \eta_a^s \quad (2.2.22)$$

$$\ln L = \ln L^D + \frac{1}{\beta_1} \ln P^D + \varepsilon_a^d = \ln L^D + \frac{1}{\beta_1} \ln (1 - P^S) + \varepsilon_a^d \quad (2.2.23)$$

The expressions (2.2.22) and (2.2.23) are log-linear and each involve one latent variable only. They map the latent variable L^S and L^D onto the observable variables L and P^D . The disturbance terms η_a^s and ε_a^d stem from the stochastic specifications of L^S and L^D (these are the aggregate shocks referred to in footnote 20). This system constitutes the core of the existing econometric studies in this line of research. In Lambert (1988), which is one of early and well-known studies in this area, the deterministic parts of L^S and L^D are substituted into (2.2.22) and (2.2.23). For the typical situation at hand, (2.2.22) transforms in an expression of P^S in terms of observable variables only (through elimination, $\ln L$ disappears). Both equations are in terms of observable variables only and can be estimated by FIML to account for cross-equation restrictions and simultaneity. The mismatch parameter (or β_1^{-1}) is estimated consistently.

The third variant follows the tradition of Sneessens and Drèze (1986). It takes supply and demand to be unobserved latent variables. They are, however, modelled non-stochastically (a greatly simplifying assumption). In fact a reduced-form estimation on L is carried through, with L according to (2.2.17) and a disturbance term added. This amounts to assuming that the aggregate supply and demand errors are equal (or perfectly correlated) as are the corresponding standard errors (note the similarity with the variant type 1 empirical work). These assumptions allow the analytical computation of the likelihood function. Besides, the ML estimator still yields consistent estimates in this model, despite the fact that the likelihood is not differentiable with respect to the parameters (Laroque and Salanié, 1992).²⁹ Although the likelihood function in this case is unbounded on the

²⁹ This result does not require uniform boundedness of the regressors and as such generalises the earlier proof obtained by Hartley and Mallela (1977) establishing consistency of the FIML estimator in the case of the canonical disequilibrium.

parameter space (it approaches $+\infty$ in regions of the parameter space where β_1^1 approaches zero, and either the standard error of aggregate supply or demand approaches zero), this property of the model, however, did not prevent the successful implementation in a number of cases. Other possible limitations relate to the use of the limited information (two-stage least squares) procedure which is adopted to circumvent the problems of full maximum likelihood (see also Mehta and Sneessens, 1990 and Draper and Kooiman, 1991 for similar applications). As a result, simultaneity and cross-equation restrictions have not been dealt with in an efficient way. Moreover, the treatment of the time disturbances is questionable. Since the two-stage least squares technique requires additive disturbances, some “tricks” have been used to make the highly non-linear time disturbances in the model which is estimated, additive.

2.3 The simulation-based variant

The fourth empirical strategy falls into the category of the simulation-based methods, *viz.* Monte Carlo Pseudo Maximum Likelihood (MCPML) estimation.³⁰ The method may be applied as an alternative to the above-mentioned variants type 1 and 2 **with** regime classification information, and the variant type 3 **without** such information though **with** simplifying assumptions about the stochastic structure of aggregate supply and demand. The MCPML method competes most directly with the variant type 3, which is also based on latent variables. It is, however, more flexible because it does not have the limitations mentioned earlier.

The method is particularly useful for the estimation of any identifiable micro-markets-based disequilibrium model in which the supply and demand variables are latent variables and stochastic functions of other variables. In Chapter 6 an econometric model is formulated that describes two macro markets (the labour and commodity market) specified at the micro markets level. The model allows for different standard errors at each side of the market and for spillover effects between markets. For this type of model, standard estimation techniques based on FIML are out of reach since the analytical computation of the likelihood function is intractable. Another feature of the method is that it does not rely on the “good quality of data assumption” implicit in the variants type 1 and type 2.

Pseudo maximum likelihood techniques are based upon the analytical expressions of the first two moments of the endogenous variables, $E(L)$ and $V(L)$, where E is the expectation and V the variance operator. They are applied to models for which the likelihood function cannot be computed. The Monte Carlo Pseudo

³⁰ The method is called pseudo maximum likelihood to indicate that the likelihood function associated with a family of probability distributions does not necessarily contain the true distribution.

maximum likelihood techniques are used when analytical relations of even the first two moments do not have analytical expressions. The MC technique provides numerical approximations of these moments. To explain the MCPML method, it suffices to start from the simple canonical model, that is without micro markets. Under these circumstances analytical expressions for the two moments exist. In the application of the method to the far more complicated model of Chapter 6 with micro markets and spillovers, analytical expressions for the moments are not available.

In the canonical single market model, labour supply is determined by $L_t^S + \sigma_s \eta_{at}^s$, labour demand by $L_t^D + \sigma_d \varepsilon_{at}^d$, and L_t is defined as:

$$L_t = \text{Min}\left(L_t^S + \sigma_s \eta_{at}^s, L_t^D + \sigma_d \varepsilon_{at}^d\right) \tag{2.3.1}$$

where t is the time subscript. The subscript a stands for “aggregate” indicating the aggregate character of the disturbance term. L_t^S and L_t^D are latent variables which are functions of the observed variables x_t and parameters θ , $L_t^S(x_t, \theta)$, $L_t^D(x_t, \theta)$. L_t is also observed. For all t $(\eta_{at}^s, \varepsilon_{at}^d)$ is a normal disturbance with zero mean and co-variance matrix equal to the identity matrix and independent of x_t . The vector of parameters to be estimated is $(\theta, \sigma_s, \sigma_d)$.

The first two moments of L , conditional on x are given by:

$$E(L) = \Phi^{,S} L^S + \Phi^{,D} L^D - \sigma^* \psi \tag{2.3.2}$$

$$E(L^2) = \left(L^{S^2} + \sigma_s^2\right) \Phi^{,S} + \left(L^{D^2} + \sigma_d^2\right) \Phi^{,D} - \sigma^* (L^S + L^D) \psi \tag{2.3.3}$$

with: $\Phi^{,S} := \Phi(y)$ Φ (is the standard cumulative normal density)
 $\Phi^{,D} := \Phi(-y)$
 $\psi := \psi(y)$ ψ (is the standard normal distribution function)
 $y = \frac{L^D - L^S}{\sigma^*}$

where $\sigma^{*2} = \sigma_s^2 + \sigma_d^2$. Gouriéroux, Monfort and Trognon (1984) were the first to propose three consistent PML procedures. All variants yield consistent estimators of the model parameters, so that convergence of the PML estimators to the true value of the estimator is established.

First-order based pseudo maximum likelihood

The PML₁ technique consists in maximising the likelihood function associated with the pseudo-model:

$$L_t = E(L(x_t, \gamma)) + \epsilon_{1t} \quad (2.3.4)$$

where the pseudo-residuals ϵ_{1t} are assumed to be independent centred normal unit variance variables; γ reflects the identifiable parameters in the model (some combination of θ and σ and τ). The pseudo likelihood function associated with an observation at t and parameter values γ equals:

$$\psi_{1T} = \frac{1}{2} \sum_{t=1}^T \left(L_t - E(L(x_t, \gamma)) \right)^2 \quad (2.3.5)$$

A PML₁ estimator $\hat{\gamma}_{1T}$ of γ minimises ψ_{1T} . Given (2.3.2) it follows that only (θ, σ^*) is identified at the first order in the canonical model. Thus, only (θ, σ^*) can be estimated from the PML₁ technique.

Second-order based pseudo maximum likelihood

The PML₂ technique consists in maximising the likelihood function associated with the pseudo model:

$$L_t = E(L(x_t, \gamma)) + \epsilon_{2t} \quad (2.3.6)$$

where ϵ_{2t} are independent centred normal variables with variance $V(L(x_t, \gamma))$. The corresponding pseudo likelihood function shows:

$$\psi_{2T} = \frac{1}{2} \sum_{t=1}^T \left(\frac{\left(L_t - E(L(x_t, \gamma)) \right)^2}{V(L(x_t, \gamma))} + \log V(L(x_t, \gamma)) \right) \quad (2.3.7)$$

A PML₂ estimator $\hat{\gamma}_{2T}$ of γ minimises ψ_{2T} . Given (2.3.2)-(2.3.3) it follows that all parameters σ , σ_a and θ are identified. From PML₂, σ , σ_a and θ can be estimated in the canonical model (Laroque and Salanié, 1989, p. 834).

Quasi-generalised pseudo maximum likelihood

The last variant of PML estimators follows from a two-step procedure. In the first step, through (for example) PML₂, a consistent estimator $\hat{\gamma}_{2T}$ is obtained. Then, in the second step the likelihood function of the pseudo model

$$L_t = E(L(x_t, \gamma)) + \epsilon_{3t} \quad (2.3.8)$$

where ϵ_{3t} are independent centred normal variables with variance $V\left(L(x_t, \hat{\gamma}_{2T})\right)$ is

maximised. The pseudo likelihood function shows:

$$\psi_{3T} = \frac{1}{2} \sum_{t=1}^T \left(\frac{L_t - E(L(x_t, \gamma))^2}{V(L(x_t, \hat{\gamma}_{2T}))} \right) \tag{2.3.9}$$

The PML₃ estimator $\hat{\gamma}_{3T}$ of γ minimises ψ_{3T}

Simulated (or Monte Carlo) pseudo maximum likelihood methods are used to approximate the first two moments numerically by simulation. The method is especially useful when analytical expressions for the moments do not exist. This is the case when the model is extended to account for two markets (including the commodity market) with micro markets and spillovers. The MCPML estimators are obtained by drawing randomly and independently at each period t , G values of the pair of disturbances $(\eta_{at}^s, \epsilon_{at}^d)$ in their distribution, and by approximating the first two moments of L with their numerical counterparts, $E_{MC}(L)$ and $V_{MC}(L)$, i.e.:

$$E_{MC}(L_t) = \frac{1}{G} \sum_{g=1}^G L_{gt} \tag{2.3.10}$$

$$V_{MC}(L_t) = \frac{1}{G-1} \left(\sum_{g=1}^G L_{gt}^2 - GL_{gt}^2 \right) \tag{2.3.11}$$

Thus, a MCPML₁ estimator of (θ, σ^*) will be a minimiser $\hat{\gamma}_{1T}^{MC}$ of:

$$\psi_{1T}^{MC} = \frac{1}{2T} \sum_{t=1}^T \left(L_t - E_{MC}(L(x_t, \gamma)) \right)^2 \tag{2.3.12}$$

For MCPML₂ $(\theta, \sigma_s, \sigma_d)$ is the minimiser $\hat{\gamma}_{2T}^{MC}$ of:

$$\psi_{2T}^{MC} = \frac{1}{2T} \sum_{t=1}^T \left(\frac{(L_t - E_{MC}(L(x_t, \gamma)))^2}{V_{MC}(L(x_t, \gamma))} + \log V_{MC}(L(x_t, \gamma)) \right) \tag{2.3.13}$$

The MCPML₃ estimator follows by implication.

Laroque and Salanié (1989) prove the convergence of the Monte Carlo approximates of the moments to the exact moments, when the number of drawings G approaches infinity, under some regularity assumption (compare their Theorem 1). The results in Gouriéroux, Monfort and Trognon (1984) then imply:

$$\hat{\gamma}^{MC} \rightarrow \gamma \quad \text{almost surely}$$

when the number of drawings G approaches infinity with sample size. The almost sure convergence here is the convergence of the Monte Carlo approximants to the pseudo likelihood estimator as the amount of Monte Carlo effort increases.

In empirical applications the latter theoretical conditions cannot be met: the number of observations is limited and there are restrictions on the number of Monte Carlo draws to keep the estimation tractable. Since experience with MCPML estimators of “disequilibrium” models is relatively scarce, it is necessary to investigate their performance with respect to the number of Monte Carlo draws under finite (small) sample properties. In Chapter 5, the overall performance of the MCPML method is investigated in which the two aspects mentioned play a role. Performance testing is restricted to MCPML₁ and MCPML₂. Given the results of the two variants, MCPML₃ is expected to have less additional information.

2.4 Identification

In the canonical single market model $\sigma^* = \sqrt{\sigma_s^2 + \sigma_d^2}$ is only identified at the first order, while σ_s and σ_d are separately identified at the second order. But, how are the parameters identified in the model with micro markets? In the single market model with micro markets aggregate transactions, L , are determined by:

$$L = E\left(\text{Min}\left(L^S(x, \theta) + \sigma_s \eta_a^s + \tau_s \eta_i^s, L^D(x, \theta) + \sigma_d \varepsilon_a^d + \tau_d \varepsilon_i^d\right)\right) \quad (2.4.1)$$

where I dropped the time subscript, and $E(L)$ is replaced with L . The expectation is taken with respect to (η_i, ε_i) . The subscript i stands for idiosyncratic to indicate a micro market. The aggregate and idiosyncratic errors are normal disturbances with zero mean and unit variance. Moreover, (ε_a, η_a) are independent from (ε_i, η_i) .

It can be shown that θ , σ_s , σ_d and $\sqrt{\tau_s^2 + \tau_d^2}$ are identified in (2.4.1). The likelihood, however, is intractable in such a model. To allow the analytical computation of the likelihood function, Sneessens and Drèze (1986) use the C.E.S. approximation to (2.4.1) with $L^S(x, \theta) + \sigma_s \eta_a^s$ and $L^D(x, \theta) + \sigma_d \varepsilon_a^d$ as arguments (compare eq. 2.2.17), and the assumptions $\sigma_s = \sigma_d$ and $\varepsilon_a = \eta_a$. This is the earlier mentioned variant type 3 method. Hence, in the empirical application of the model the aggregate shocks of supply and demand are not separately identified for this method. Pseudo maximum likelihood techniques only make use of the first two moments of the endogenous variable. From (2.4.1), The expression for $E(L)$ can be computed. It can be shown that is identical to equation (2.3.2) with the exception that the variance parameter, σ^* , here is defined as $\sqrt{\sigma_s^2 + \sigma_d^2 + \tau_s^2 + \tau_d^2}$. So θ and $\sqrt{\sigma_s^2 + \sigma_d^2 + \tau_s^2 + \tau_d^2}$ are first-order identifiable. This leads to the observation

that the canonical model cannot be identified from the model with micro markets at the first order. Moreover, only an upper bound estimate of the labour market mismatch can be obtained in the PML, application of the micro markets model. In the PML₂ application to the micro markets model the micro supply and demand errors are identified from the aggregate ones (Laroque and Salanié (1989)). It follows that θ , σ_s , σ_d and $\sqrt{\tau_s^2 + \tau_d^2}$ are second-order identifiable.

In the variant of (2.4.1) with observed supply and demand (type 1 method), the micro mismatches are identified from the aggregate errors at the first order. In such a model θ and $\sqrt{\tau_s^2 + \tau_d^2}$ are identified. The aggregate errors σ_s and σ_d are not separately identified, as in the variant type 3. In the empirical application it is, however, possible that, significant shifts in location may entail a temporary increase in micro mismatches in some rare cases (Kooiman and Kloeck, 1979). Then identification will be weak between the micro mismatches and aggregate errors. In case of the variant type 2 method, using data on business survey results to observe latent variables, θ , $\sqrt{\tau_s^2 + \tau_d^2}$ and $\sqrt{\sigma_s^2 + \sigma_d^2}$ are identified, as in the variant type 1. (Laroque and Salanié, 1989, p. 837).

2.5 Review of empirical applications

To complete the description of the four broad macro methods of markets in disequilibrium used in the literature to measure mismatches and the corresponding structural unemployment, I took a large representative sample of recent empirical contributions. For classification purposes, I confine myself to the subdivision in the four categories of Sections 2.2 and 2.3. I explained in Chapter 1 that a fifth category may be added. It covers those studies which have something to say about structural unemployment or structural problems on the labour market while not belonging to the method of markets in disequilibrium. Among them are the traditional *U-V* analysis and the variance method of relative unemployment rates of Layard, Nickell and Jackman (1991). It should, however, be clear that due to differences in definition, the results from these studies are not directly comparable with those of the disequilibrium studies. In reviewing the various contributions, I restrict myself to studies with applications to European economies.³¹

Before going into the specificity of the studies, it is again worth pointing out that the way mismatches are endogenised is not very well developed (see Chapter 1). To allow for technical innovation, most studies define the mismatch parameter as an *ad hoc* function of time. Technical innovation in this interpretation

³¹ Applications to countries outside Europe are scarce, to my knowledge.

is regarded as a factor which widens the imbalance between labour qualifications supplied and demanded. Examples of such applications are Kooiman and Kloek (1979) Sneessens and Drèze (1986) and Stalder (1989). Based on the same motivation used to incorporate a time-trend, labour productivity has been used as an explanatory variable (for instance, by Kooiman and Kloek, 1979 and Lambert, 1988). A third type of application defines the mismatch parameter as a function of aggregate unemployment. In this sense, a high unemployment rate is assumed to impair mobility on both sides of the market (see Lambert, 1988). Chapter 4 addresses the question of endogenising more thoroughly.

Table 2.2 provides a schematic overview of the work done in the field (analogous tables summarising earlier work can be found in Lambert, 1988 and Laffont, 1983). It contains information about the use of regime classification, the functional form of transactions (analytically-based methods only), the stochastic specification, the estimation procedure, correction on vacancy data, and the development of structural unemployment. The table also presents the results of my own research results. They are discussed in Chapters 4 and 6.

Variant type 1

The variant type 1 is the first analytically-based method. It takes supply and demand as observed using a single equation estimation technique. Most applications fall into this category. The applications have two things in common. First, this is the stochastic specification of transactions which simply adds a disturbance term to the deterministic part. The disturbance term includes aggregation errors, errors due to the assumption of log-normality, *etc.* Supply and demand errors are not accounted for. A consequence of this assumption, which may be expected to be unrealistic, is that these errors are not adequately reflected in the disturbance term. It can be shown that allowing for these type of errors would result in non-linear and heteroscedastic disturbances. The effects on the estimation results are not very well established. They may, however, be considered less important if the supply and demand data are observed reasonably accurate from the data. In any case, theory does not disallow the single equation specification in favour of a model formulation with interactions. It is not impossible that the model is badly specified. Second, the estimation technique adopted in these studies is non-linear least squares. The variant type 1 has been particularly popular in the Netherlands. Below I consider three studies carried out for the Netherlands and one for Germany.

In *Kooiman and Kloek (1979)* the labour market is considered irrespective of goods market influences. Equation (2.2.8) is estimated with an additive disturbance term.³² Data on supply and demand are taken from the labour exchange

³² Kooiman and Kloek (1979) also estimated a labour market model with supply and demand unobserved and described by stochastic functions. The results, however, were disappointing. Therefore, I decided not to discuss them further.

registers. This is the limitation of the study. By now it has become clear that vacancy data substantially overestimate the “true levels”. As a result, the structural unemployment rate is estimated too high. The study covers the period 1948 to 1975. The results imply a structural unemployment rate just below 2% in the first half of the sample period and just over 2% in the second half. In *Driehuis* (1987) goods market considerations are partially incorporated in the employment function. At the micro level, transactions are the minimum of labour supply, Keynesian labour demand (or the amount of labour needed to produce the demand for goods) and classical labour demand (the amount of labour necessary to fully utilise the capital stock). The aggregate functional form is of the C.E.S. type with three variables:³³ aggregate labour supply, Keynesian, and classical demand for labour. To measure the two demand variables Driehuis chooses a rather peculiar method. For instance, data on Keynesian demand are obtained from an assumed identity relationship between the concerning variable and goods demand (= model outcome), goods transactions and employment. Ideally, the Keynesian demand for labour figures are an integrated part of the model estimation. In the application of Driehuis, the approximation of Keynesian demand is probably very poor! The estimation period is from 1960 to 1985. The estimated structural unemployment rate is almost constant at about 2.2%. The model of *Muysken, Bierings and De Regt* (1994)³⁴ is dealt with more extensively in Section 4.4. In the empirical application only the labour market is considered and is abstracted from goods market spillovers. The new element in the analysis allows the decomposition of structural unemployment into the familiar mismatch between micro markets and search unemployment due to matching inefficiencies. The estimation covers the period 1960 to 1990 and is based on annual data. The data on vacancies are corrected for systematic biases mentioned earlier. The results indicate that structural unemployment due to mismatches is increasing from slightly over 1% in the mid-sixties to over 3% in the late eighties.

For Germany *Franz and Smolny* (1994) estimated a C.E.S. transactions function for the labour market only, without spillovers from the goods market. An innovative element is the incorporation of the fraction of lagged employment as the third variable in the C.E.S. specification. This specification is derived from a micro model where employment is the minimum of supply, demand and a fraction of lagged employment.³⁵ The estimation period is from 1960 to 1990 and the data are on a quarterly basis. Structural unemployment is estimated at 1.3% between the first

³³ The specification is borrowed from Snessens and Drèze (1986, see also Section 4.2).

³⁴ Further elaborations on this can be found in Muysken and De Regt (1994).

³⁵ Given this micro model, the aggregate result is intuitively difficult to acknowledge given the complex time correlations in micro market employment. An exact proof, however, is not given. I conjecture that the results can only be obtained when the dynamics at the micro level are abstracted from at some stage in the aggregation procedure.

quarter of 1960 (1960I) and the fourth quarter of 1973 (1973IV), and 2.1% between the first quarter of 1974 (1974I) and the fourth quarter of 1990 (1990IV).

Variant type 2

The variant type 2 is the second analytically-based method. The applications consider two macro markets, labour market and goods market, with spillovers and make use of business survey results to measure regime proportions. The applications of Lambert (1988), Stalder (1989) and Gagey, Lambert and Ottenwaelter (1990) have some aspects in common. The equations for supply and demand on the goods and labour market are modelled stochastically with (additive) log-normal disturbances. From possible errors in the regime proportions is abstracted. This concerns the non-linearity of these errors, the errors due to the aggregation procedure and the log-normality approximation. These type of errors, which one would reasonably expect, are therefore not adequately represented in the disturbance terms. The models are estimated by ML taking into account the simultaneity of the endogenous variables and cross-equation restrictions.

Lambert (1988) describes the labour and the goods market at the micro level independently. In the aggregation, first aggregates over labour markets and goods markets are determined independently. Then spillovers are introduced at the aggregate level in an *ad hoc* fashion. The Lambert model is estimated for Belgium period 1965-1980) using annual data. The results imply an average structural unemployment rate of 1.4%.

Stalder (1989) estimated the model for Switzerland with two macro markets. The specification corresponds largely to the one suggested by Lambert (1988). On a more basic level, the aggregation procedure differs from Lambert in that spillovers between the labour market and the goods market are introduced at the micro level and not on the aggregate level as in Lambert (1988). The Stalder method, however, looks superior. Through the use of approximation formulae for the cumulative densities, the difference between the methods fade away and comparable closed forms are obtained. The estimation is carried out on quarterly data for the period 1967 to 1985. A growing labour market mismatch is found as would be implied by the rise of the structural unemployment rate from 1.2% in 1967II to 1.8% in 1985IV.

The *Gagey, Lambert and Ottenwaelter* (1990)³⁶ method is described extensively in Subsection 4.3.2. At the micro level, a hierarchic decision procedure in two steps is suggested. First, the firm decides on its demand, which equals the minimum of Keynesian demand, and classical demand for labour. Then in the second stage, the level of transactions is determined as the minimum of labour demand and labour supply. As a consequence of this procedure, the aggregate expression on transactions discerns two types of mismatches: a labour market

³⁶ Lambert (1990) contains a summary of these results.

mismatch and a goods market mismatch. The model is estimated for France (period 1968-1986) on an annual basis. Structural unemployment rises from about zero in 1968 to 3.3% in 1986.

Variant type 3

The variant type 3 is the third analytically-based variant. In this variant, supply and demand are treated as unobserved latent variables, which are again deterministic functions of well-defined variables. *Sneessens and Drèze* (1986) is one of the early works which set the pace for other studies. A complete model is specified. The labour market is modelled as well as the goods market. The specification of aggregate employment is the same as in *Driehuis* (1987) described above. The estimation is on annual data for Belgium (period: 1955-1982). Structural unemployment is found to be rising from 1.4% in 1955 to 4.5% in 1982. *Mehta and Sneessens* (1990) applied the same procedure as in *Sneessens and Drèze* (1986). The main difference lies in the specification of the model. *Mehta and Sneessens* treat investments and exports as endogenous, whereas in *Sneessens and Drèze* they are exogenous. In the empirical application, the export, import and employment equations are estimated jointly. In *Sneessens and Drèze* this is only the case for the import and the employment equation. Structural unemployment is estimated to rise from 3.3% in 1955 to 4.7% in 1985. Remarkably, the structural unemployment rate is significantly higher in 1955 than in *Sneessens and Drèze*, which is probably the result of endogenising exports and imports. The model in *Sneessens and Mehta* (1993) generalises the employment function of *Lambert* (1988) and *Sneessens and Drèze* (1986) to two labour inputs, skilled and unskilled labour. However, they abstract from spillovers from the goods market. Further, the procedure is largely comparable. As to the results: the structural unemployment rate on the skilled labour market increased from about 0.7% in 1962 to 1.3% in 1989. The aggregate structural unemployment rate shows a sharp rise from 1% in 1962 to more than 6%. These results seem to diverge somewhat from those of *Gagey, Lambert and Ottenwaelter* (1990) for France.

Variant type 4

This is the simulation-based variant. *Laroque and Salanié*, (1989) developed the method. In the macroeconomic model specification the labour and the goods market are described. The formulation is at micro market level, including the spillover terms. Structural unemployment is not a direct estimation result. The estimated variances of micro demands and supplies, however, indicate a labour market that is homogenous and efficient (and a goods market that is rather heterogeneous). It is noteworthy that the other studies for France indicate a significant mismatch on the labour market.

Variant type 5

To test for the presence of structural effects, *Layard, Nickell and Jackman* (1991) formulated a model in which price and wage decisions are taken separately on different segments of the economy. The appropriate mismatch indicator in this set-up is the variance of relative unemployment rates. Using several distinctions (by profession, by region, by industry), the authors find that the value of this mismatch indicator generally has not increased in European countries over the last twenty years.³⁷ Similar results are obtained when the model is extended with vacancies. The authors still point out that mismatches as such may be relevant in Europe, and are a serious problem. The method *Layard et al* (1991) propose is simple. However, it relies on some restrictive model assumptions with flexible wage formation. Besides, the unemployment data used by region, profession, and industry may be subject to large definition and measurement problems.

Traditional *U-V* analysis is another example of studies that fall in the fifth variant. Although, for reasons explained above, traditional *U-V* analysis is inferior to the markets in disequilibrium method, the method should be mentioned to make the list complete. Although the number of contributions in this field is large, I will mention only the recent work of *Cristl* (1994) that differs from the others by its use of corrected vacancy data. The application is to Austria and covers the period 1966 to 1991. The estimation results indicate a rate of structural unemployment of on average 3.1% in the period 1966 to 1983 and 6.1% in the period 1984 to 1991.

³⁷ These results are not incorporated in Table 2.2. For details I refer to the relevant publication.

Table 2.2 Main features and results of recent empirical disequilibrium models with micro markets measuring (mismatch implied) structural unemployment

| | <i>Kooiman and Kloek (1979)</i> | <i>Driehuis (1987)</i> | <i>Muytken, Bierings, De Regt (1994)</i> | <i>Franz and Smolny (1994)</i> | <i>Bierings (Chapter 4 of this book)</i> |
|--|--|---|--|--|--|
| Data | Annual | Annual | Annual | Quarterly | Annual |
| Country | the Netherlands | the Netherlands | the Netherlands | Germany | the Netherlands |
| Period | 1948-1975 | 1960-1985 | 1960-1990 | 1960I-1990IV | 1960-1989 |
| Regime classification information | Supply and demand observed | Supply and demand observed | Supply and demand observed | Supply and demand observed | Supply and demand observed |
| Functional form | Eq. (2.2.8) | C.E.S. | C.E.S. | Dynamic C.E.S. | C.E.S. |
| Specification | Labour market | Labour market, and goods market influence | Labour market | Labour market; exogenous supply/ demand | Labour market |
| Estimation | Non-linear least squares/ single equation | Non-linear least squares/ single equation | Non-linear least squares/ single equation | Least squares/ single equation | Non-linear least squares/ single equation |
| Stochastic or deterministic | Supply/demand deterministic | Supply/ demand deterministic | Supply/ demand deterministic | Supply/ demand deterministic | Supply/ demand deterministic |
| Vacancy data corrected? | No | Not relevant | Yes | Yes | Yes |
| Structural unemployment rate | Just below 2% between 1948 and 1961; just above 2% between 1962 and 1975 | Roughly constant at 2.2% | About 1% in the mid-sixties to just over 3% in the late eighties | Rising from 1.3% (1960-1973) to 2.1% (1974-1990) | Rising from 2.3% in 1960 to 4.9% in 1989 (between 1962-1975) |

Table 2.2 Main features and results of recent empirical disequilibrium models with micro markets measuring (mismatch implied) structural unemployment (continued)

| | <i>Lambert (1988)</i> | <i>Stalder (1989)</i> | <i>Gagey, Lambert and Ottenwaelder (1990)</i> | <i>Sneessens and Drèze (1990)</i> | <i>Mehta and Sneessens (1990)</i> |
|--|--|---|--|--|--|
| Data | Annual | Quarterly | Annual | Annual | Annual |
| Country | Belgium | Switzerland | France | Belgium | Belgium |
| Period | 1965-1980 | 1967II- 1985IV | 1968-1986 | 1955-1982 | 1953-1985 |
| Regime classification information | Business survey data | Business survey data | Business survey data | None | None |
| Functional form | System com- parable to eq. (2.2.22)- (2.2.23) | C.E.S. like closed form | C.E.S. like closed form | C.E.S. | C.E.S. |
| Specification | Labour market and goods market | Labour and goods market with spillovers | Labour and goods market with spillovers (labour and capacity mismatch) | Labour and goods market with spillovers | Labour and goods market with spillovers |
| Estimation | FIML | FIML | FIML | Joint (limited) ML estimation | Joint (limited) ML estimation |
| Stochastic or deterministic | Supply/de- mand stochastic | Supply/de- mand stochastic, regime proportions deterministic | Supply/de- mand stochastic, regime proportions deterministic | Supply/de- mand deterministic | Supply/demand deterministic |
| Vacancy data corrected? | Not relevant | Not relevant | Not relevant | Not relevant | Not relevant |
| Structural unemployment rate | Just over 1.4% on average | Rising from 1% in 1967II to 2% in 1985IV | About zero in 1968; 3.3% in 1986 | Rising from 1.4% in 1955 to 4.5% in 1982 | Rising from 3.3% in 1955 to 4.7% in 1982 |

Table 2.2 Main features and results of recent empirical disequilibrium models with micro markets measuring (mismatch implied) structural unemployment (continued)

| | <i>Sneessens and Mehta (1993)</i> | <i>Laroque and Salanié (1989)</i> | <i>Bierings (Chapter 6 of this book)</i> | <i>Cristi (1994)</i> |
|--|---|---|--|--|
| Data | Annual | Quarterly | Annual | Quarterly |
| Country | France | France | the Netherlands | Austria |
| Period | 1962-1989 | 1963-1984 | 1950-1991 | 1966-1982 |
| Regime classification information | None | Endogenous/ 4 regimes | Endogenous/ 4 regimes | Unemployment/ vacancies observed |
| Functional form | C.E.S. | Analytical form cannot be derived | Analytical form cannot be derived | Hyperbolic <i>U-V</i> |
| Specification | Two types of labour: skilled and unskilled labour/goods market (and spillovers) | Labour and goods market/ multi-market setting (goods market and labour market spillovers at the micro market) | Labour and goods market/ multi-market setting (goods market and labour market spillovers and dynamics at the micro market) | Labour market |
| Estimation | Joint ML estimation with instrumental variables | Monte Carlo pseudo maximum likelihood | Monte Carlo pseudo maximum likelihood | Ordinary least squares/single equation |
| Stochastic or deterministic | Stochastic specifications | Stochastic specifications | Stochastic specifications | Unemployment/ vacancies |
| Vacancy data corrected? | Not relevant | Not relevant | Not relevant | Yes |
| Structural unemployment rate | 1% in 1962 and 6% in 1989 | Not significantly different from zero on average | Around 1% on average | Rising from 3.1% in 1960-1973 to 6.1% in 1974-1990 |

3 The validity of the (log-)normal density to describe micro market supply and demand

3.1 Introduction

Applications of the analytically-based model frequently use the *log-normal* distribution to describe supplies and demands of micro markets (see the references put forward in Section 2.5). Its legitimacy is usually advocated because of its “reasonable” properties (Lambert, 1988, p. 17). An “analogy argument” is asserting the wide applicability of the log-normal distribution to economic variables such as personal incomes, individual wealth, market shares, output and manpower of firms. The frequency distributions of these variables are fairly well described by the log-normal distribution. Evidence from frequency distributions of micro labour supply and demand, however, is not forwarded. Another type of argument emphasises the attractive functional form for aggregate transactions (C.E.S.) that may be derived from the log-normal distribution. Although this is true, the argument is rather artificial. In the simulation-based model (the variant type 4) where micro supply and demand are treated as *normal* variates, no arguments of whatsoever have been put forward to sustain the use of the normal distribution.³⁸

The basis for the log-normality (and the normality) assumption is therefore rather volatile. As a matter of fact, it has not been scrutinised thoroughly in an empirical and/or theoretical sense. The fundamental question left unaddressed is: How fundamental is the choice of the (log)normal distribution in the field of aggregate transactions functions based on stochastic micro supply and demand? And related to this: How sensitive is this choice for the measurement of mismatches and the implied structural unemployment rate?

In this chapter I will explore on the (log-)normality postulate, and investigate the sensitivity of the aggregate functional forms (and of structural unemployment) to these types of distributions in particular. The most obvious way

³⁸ This leaves aside that the normal distribution is in no way crucial to the simulation-based method (see Section 1.1).

to deal with the questions above is to perform a disaggregated analysis of labour supply and demand. In the present context this would require the identification of micro markets more or less operating as independent entities. Unfortunately, such identification is complicated by the lack of sufficient disaggregated and useful data.³⁹ As a consequence, I will not attempt to provide direct evidence about the actual frequency distribution of micro market supplies and demands. Instead, I propose another approach.

This chapter starts with interpreting the validity of the log-normal postulate from the perspective of the theory of stochastic processes.⁴⁰ I will discuss the theoretical validity of the log-normal distribution in relation to real labour market phenomena. This is done by considering the typical conditions for a specific stochastic process (Gibrat's Law) to end up with the log-normal distribution as limiting (stable) distribution. In Subsection 3.2.1 the log-normal distribution is considered with its main contender, Pareto's Law. The reasoning leads to an empirically refutable hypothesis regarding the validity of the log-normal distribution with respect to Pareto's Law. As it turns out, the aggregate transactions (= employment) function based on Pareto's Law is nested in the function based on the log-normal distribution, so that statistical discrimination on nested testing is possible.

In Subsection 3.2.2 I investigate the sensitivity of the aggregate employment function with respect to other distributional assumptions while also taking the normal, Weibull, logistic and exponential distributions into account. Apart from the aggregate employment function based on the Weibull distribution, the employment functions of the other distributions can not be integrated in the nested testing mentioned above. In Section 3.3 the results from nested testing are presented.

As stipulated before, not only the log-normal, but also the normal distribution is used in the literature on markets in disequilibrium. The log-normal distribution is mainly applied in analytically-based models, whereas the normal distribution seems to belong to the domain of the simulation-based model. To settle the "controversy" between the two distributional assumptions I will develop a technique to test normality and log-normality, independently against some general form. This is subject of Section 3.4. The results are used for defining the stochastic specifications of the models in Chapter 4 (which is in the analytically-based tradition) and of Chapter 6 (applying the simulation method). In all this it should be noted that the techniques developed to test the assumptions are first attempts. Further investigations remain necessary to arrive more definite results. Section 3.5 contains the conclusions.

³⁹ See also footnote 4.

⁴⁰ A discussion of the normal postulate from this theoretical notion is excluded since it does not lead to new insights.

3.2 Sensitivity of the aggregate employment function with respect to the stochastic distribution

3.2.1 Evidence from the theory of stochastic processes: *Gibrat's Law against Pareto's Law*

In the literature ample empirical support can be found for the use of the log-normal distribution to describe the size distributions of economic variables (personal income, wealth *etc.*). The rationale for generating the log-normal distribution is "Gibrat's Law of proportionate effect" (see Lambert, 1988, p. 17). It assumes that the values taken by these variables are affected by a great many independent random factors of finite variance that operate in a multiplicative fashion. When this is adopted, a straightforward application of the Central Limit Theorem to the logarithms of the random elements involved leads to the log-normal as the limiting distribution. In contrast, the normal distribution is obtained whenever the value of a variable is determined by a large number of additive independent shocks.

The dispute to describe actual frequencies of economic variables like incomes and prices by either the log-normal or the normal distribution is mostly settled by matter of fact: the log-normal distribution describes the observations but the normal distribution does not. Another argument in support of the log-normal density is that it is by its very nature restricted to the positive range, as are all economic variables.

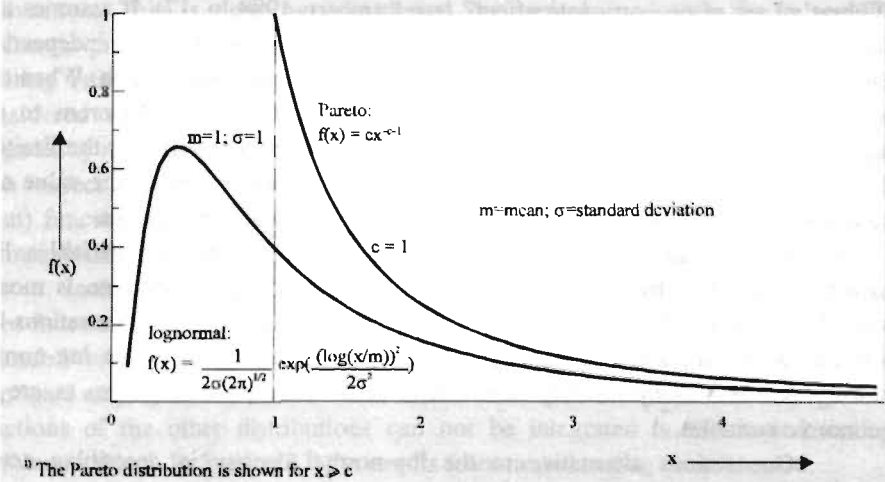
One serious alternative to the log-normal density in describing actual frequencies of economic variables is Pareto's Law. Differences between the distributions occur in the two tails (see Figure 3.1). Pareto's Law is capable only of describing frequencies that decrease towards the right. Evidence from income distributions suggests that the latter is more suited to describe high income ranges, whereas the log-normal does better in the middle income brackets. No decisive argument for or against either alternative can be given. In the case of labour markets the log-normal distribution has more intuitive appeal: reality is probably characterised by a relatively great number of medium sized markets due to specialisation.

From these facts, it is reasonable to infer for the situation at hand, that Pareto's Law fits the larger markets (*i.e.* markets with a relative large supply and demand) more adequately and the log-normal is more suited to describe markets in the medium range. However, as there is no prior information on the actual size distribution it is impossible to test the validity of both distributions on a micro level.

To arrive at some conclusions about the validity of the log-normal distribution in relation to Pareto's Law, I will look into the stochastic processes which govern these distributions and investigate whether the conditions imposed on the process could describe real phenomena reasonably well. A comprehensive evaluation of the log-normal distribution and Pareto's Law from the theory of

stochastic processes is outside the scope of this book. But it is worthwhile to compare some specific features of the two distributions in attempting to answer the “feasibility” question. For a more detailed description of stochastic processes, I refer to Ijiri and Simon (1977).

Figure 3.1 The log-normal and the Pareto distribution (univariate cases)¹⁾



The Gibrat stochastic process is usually adduced to justify the log-normality assumption. In its basic form Gibrat’s Law states that a relative change in a variable x_i is governed by a Markov process in which the probability of a relative change in x_i is independent of the magnitude of x_i . It can be shown that from a Gibrat process the log-normal distribution results as the limiting distribution (Gibrat, 1931). The following proposition applies (Ijiri and Simon, 1977):

PROPOSITION 3.1: *In order to generate asymptotic log-normality, the Gibrat process should satisfy three conditions. These are:*

1. *all elements in the population must start the random walk at the same time,*
2. *there may be no entries in or withdrawals from the population and*
3. *there may not be serial correlation in the rate of growth of x_i .*

Let us develop some intuition as to how far these conditions apply to actual situations. Condition 2 states that entries (in the present context: new markets) and

withdrawals (here: disappearance of markets) may not be the case. For the choice of a model it may make a considerable difference whether the population concerned has a rapid turnover (as in the case of individuals) or not (as with firms) (see: Cramer, 1973). Of course, this condition is not met in real life. But what matters is its relative importance. It is perceivable that for the Dutch labour market turnover is concentrated within labour markets so that the birth and death of markets is a relatively slow process which can be legitimately ignored. Scarce counterexamples could not invalidate the log-normal distribution. In Steindl (1965) it was shown that under certain conditions a stochastic process allowing for entry and exit (of firms) leads to a distribution that has a Pareto tail. In Steindl's model the occurrence of the Pareto tail depends on the numerical values of three parameters, implying the expected growth of individual firms must not be zero, and the rate of entry must ensure a net increase of their number.

Condition 3 implying the absence of autocorrelated growth is hardly perceivable. There are severe problems of autocorrelation in the error terms, particularly in the estimation of analytically-based models of markets in disequilibrium. Several authors suggested that this type of misspecification is likely to come from an inadequate dynamic specification (Lambert, 1988 and Muysken and De Regt, 1992 and 1994). To cope with it, Lambert (1988) extended the static micro market model with log-normally distributed supply and demand by allowing for dynamic interactions between markets, and, as Bierings and Muysken (1987) have shown, also for autocorrelated growth of markets. Although this extension has yielded some interesting insights in the interpretation of the dispersion parameter, within such a dynamic context, it proved an unsuccessful tool for getting rid off the autocorrelation problem.⁴¹ Moreover, the paradoxical situation arises that the micro model allowing for autocorrelated growth is incompatible with the assumption of log-normality (see Bierings and Muysken, 1987). Muysken and De Regt (1994) extended the basic micro model dynamically, allowing for search within a micro market (see also Chapter 4). Despite its theoretical elegance, their model could not solve the autocorrelation problem either: it is still there in their estimation results.⁴²

Ijiri and Simon (1977, Chapter 8) show that the distribution generated in a process governed by autocorrelated growth gives a good approximation of the Pareto or the Yule distribution. For that reason one should derive the aggregate

⁴¹ One could argue whether the Lambert model (Lambert, 1988, Appendix B, p. 126-30) is a dynamic extension at the micro level, since the aggregate specification does not contain lagged terms. Actually, one can show that the stationarity conditions his model assumes, defines away the intertemporal dynamics between markets, initially allowed for (Bierings and Muysken, 1987). This is not the case in Muysken and De Regt (1992 and 1994) in which intertemporal effects are more adequately accounted for, and where finally the absolute change in aggregate supply and demand are variables in the aggregate employment function (see also Chapter 4).

⁴² It is, however, encouraging that in my application the Muysken and De Regt model performs quite well, *without* having the serial correlation problem (Chapter 4, Table 4.1).

employment function from a Pareto distribution and **not** from a log-normal distribution. The Yule distribution is analytically much more difficult to handle, but has a similar form as the Pareto distribution and approaches Pareto's Law asymptotically.

The aggregate function based on the Pareto distribution naturally follows by evaluating the integrals in L (c.f. 2.2.5):

$$L = \int_0^{\infty} \int_{L^s}^{\infty} L^s h(L^s + \theta_s, L^d + \theta_d) dL^s dL^d + \int_0^{\infty} \int_{L^d}^{\infty} L^d h(L^s + \theta_s, L^d + \theta_d) dL^d dL^s \quad (3.2.1)$$

where $h(.,.)$ is the bivariate Pareto distribution (Johnson and Kotz, 1972) transformed to the origin, i.e.:

$$h(L^s + \theta_s, L^d + \theta_d) = \frac{\alpha(\alpha+1)(\theta_d \theta_s)^{\alpha+1}}{(\theta_d L^s + \theta_s L^d + \theta_d \theta_s)^{\alpha+2}} \quad (3.2.2)$$

with $L^s > 0; L^d > 0; \theta_s > 0; \theta_d > 0; \alpha > 0$.

It can then be shown that the following specification can be obtained:⁴³

$$L = \frac{L^S L^D}{L^S + L^D} \quad (3.2.3)$$

This is a rather surprising and at the same time disappointing result. Because of its simple form and its lack of parameters one cannot expect this aggregate employment function to provide a satisfactory description of any actual situation. Note that it implies a structural unemployment rate of 50 percent. Thus, finally, the paradoxical result is obtained that relaxation of the assumption of absence of autocorrelated growth and/or entry/exit in the stochastic process underlying the log-normal distribution does not disqualify the log-normal distribution in describing the stochastic distribution of supply and demand over micro markets (see Bierings and Muysken, 1987). On the contrary, it should be preferred. Evidently, allowing for autocorrelated growth and entry/exit does more harm than good. Therefore, one has to conclude that the stochastic process allowing for autocorrelated growth and entry/exit must restrict the process in another essential aspect than the Gibrat process. Or, in other words, a process monitored by autocorrelated growth and entry/exit is rigidly forced into the wrong direction, whereas the Gibrat process is

⁴³ See Annex 3A for a derivation.

more flexible despite of its more restrictive assumptions.⁴⁴ The above outcomes should be checked by more rigorous analysis. Preferably, the robustness of the distributions under different assumptions on entry and exit should be substantiated by simulation experiments.

Remarkably, the employment function based on Pareto's Law is nested in the employment function based on the log-normal distribution, although the underlying distributions are disjoint. This also allows the use of nested testing, where the nesting refers to aggregate supply, demand and employment (Section 3.3).

3.2.2 Other distributions

In this subsection the sensitivity of the aggregate employment function with respect to some other alternative specifications of $h(L^s, L^d)$ is considered. Aggregate employment functions have been derived for the normal (Kooiman and Kloek, 1979 and Andrews and Nickell, 1984) and the Weibull distributions (Den Broeder, 1983 and Gouriéroux, Laffont and Monfort, 1984). Table 3.1 offers a complete picture of all the experiments with different distributional assumptions. I have added the specifications for the logistic distribution and the exponential distribution.⁴⁵

Note from the table that the employment function based on the (mutually independent) Weibull distributions is a general form of the employment functions based on the transformed Pareto and the log-normal distribution in spite of the fact that all three distributions are disjoint.⁴⁶ There is, however, the possibility that the Weibull distribution on the one hand, and the log-normal and transformed Pareto

⁴⁴ In a recent article of Cowell and Victoria-Veser (1996) about inequality measures (among others the Theil and the Gini index), the values of the inequality measures were less sensitive for contamination of the data in the simulation with log-normal (UK-income) data than with Pareto data. I can not explicitly link these results to mine, it is remarkable, however, that the log-normal distribution is apparently robust under different circumstances.

⁴⁵ See Annex 3B for the derivation on the basis of the logistic distribution. The derivation on the basis of the exponential distribution is straightforward and is left to the reader.

⁴⁶ At first glance it may be surprising that the result for the Weibull distribution is the most general, although contrary to the log-normal (and Pareto) case, independence between micro supply and demand is assumed. However, in case of the log-normal distribution, the covariances between micro supply and demand are not identified in the aggregate specification (see eq. 2.2.8, in which only the mismatch, σ_{ϵ} , is identified). Theoretically, however, the zero-covariance assumption is less plausible: Labour supply and demand in a certain market will not be totally independent.

distribution on the other hand are nested because the domains between the distributions differ (See Table 3.1).

The aggregate employment function based on the Weibull distribution shows:

$$L = \left(a + (L^S - a)^{-\beta_1} + (L^D - a)^{-\beta_1} \right)^{-\frac{1}{\beta_1}} \quad (3.2.4)$$

When parameter a is set to zero, the result is consistent with the case of the log-normal distribution. When next β_1 is set to one, the result is consistent with the case of the Pareto distribution. Hence, the test which of these models provides an adequate explanation of a given situation is rather straightforward, when the differences in domains between the distributions are ignored. It is possible that the Weibull distribution with domain $0 < a < L$ closely corresponds to the lognormal with domain $L > 0$. This could imply that a is almost zero, and needs not be estimated.

The aggregate employment functions based on the normal, logistic and exponential distributions are not nested. As a result, there is no simple way to investigate their empirical validity against the employment functions based on the other distributions. Yet I developed some kind of plausibility testing for the normal distribution in Section 3.4. The employment function based on the logistic or exponential distribution cannot be tested for its empirical plausibility by use of nested testing.

Figure 3.2 provides a graphical representation of the nested aggregate employment functions including the limited nested cases for $\beta \rightarrow 0$ and $\beta \rightarrow \infty$. The case for $\beta \rightarrow \infty$ represents the so-called Leontieff employment function".⁴⁷ In Annex 3C, I prove the Leontieff result, showing it may be derived for any distribution function when the mismatch of micro demands and supplies goes to zero.

⁴⁷ The Leontieff concept is familiar from production function theory. Moreover, the "smoothing by aggregation" is not unknown in production function theory and goes back to Houthakker (1955). It is interesting to note that the Leontieff structure, evidently, may be justified from a probability distribution of corresponding micro variables.

Table 3.1 Contributions to aggregate employment specifications using different assumptions about micro supply and demand ¹

| | | | |
|----------------------|---|--|---|
| Authors: | <i>Kooiman and Kloek (1979)/ Andrews and Nickell (1984)</i> | <i>Kooiman and Kloek (1979)/ Lambert (1988)</i> | <i>Den Broeder (1983)/ Gouriéroux Laffont and Monfort (1984)</i> |
| Density: | bivariate normal | bivariate log-normal | product of two univariate Weibull densities |
| Specific. L: | $\Phi^S_L S + \Phi^D_L D$ | $\Phi^S_L S + \Phi^D_L D - \sigma^2 \psi$ | $\left((L^S - a)^{-\beta_1} + (L^D - a)^{-\beta_1} \right)^{-\frac{1}{\beta_1}}$ |
| | ($L > 0$; compare eq. 2.3.2) | ($L > 0$; compare eq. 2.2.8) | ($\beta_1 > 0$; $0 < a < L$; compare eq. 3.2.4) |
| Approx. of L: | <i>not available</i> | $\left((L^S)^{-\beta_1} + (L^D)^{-\beta_1} \right)^{-\frac{1}{\beta_1}}$ ($\beta_1 > 0$) | <i>not relevant</i> |
| u*: | $\frac{\sigma^2}{\sqrt{2\pi L^S}}$ | $1 - 2\Phi\left(-\frac{1}{2}\sigma\right)$ or: $-\frac{1}{1-2\beta_1}$ | $1 - 2\frac{1}{\beta_1} - \frac{a \left(1 - 2\frac{1}{\beta_1} \right)}{L^S}$ |

¹ For an explanation of the symbols, see the main text.

Table 3.1 Contributions to aggregate employment specifications using different assumptions about micro supply and demand (continued)¹

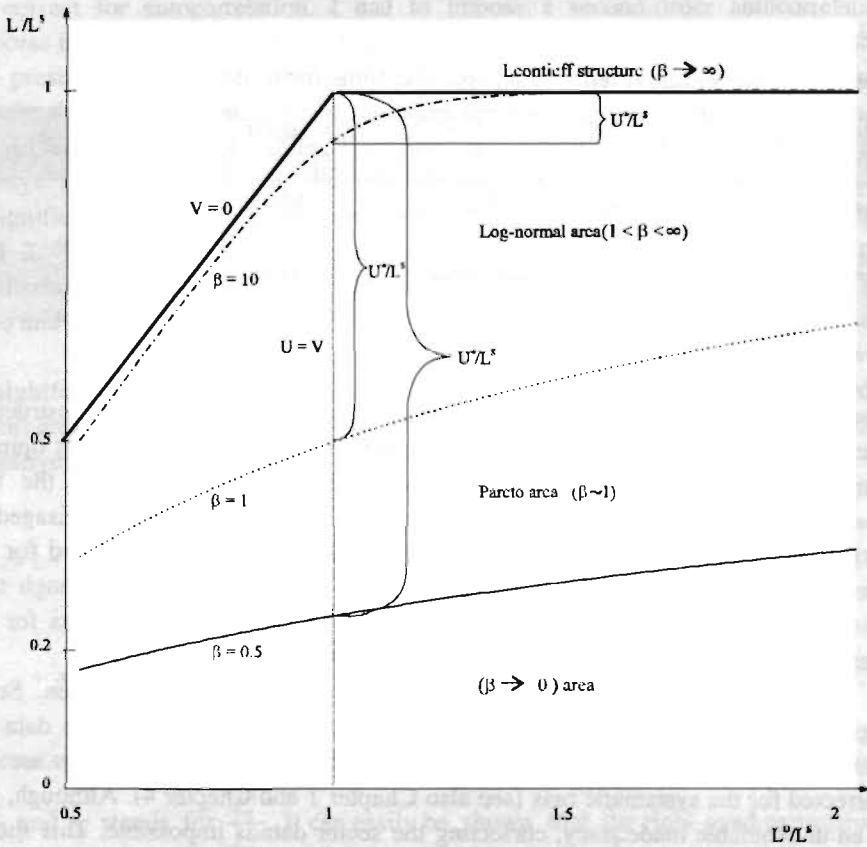
| Author: Bierings (this chapter) | | | |
|---------------------------------|--------------------------------|--|---|
| Density: | bivariate Pareto | bivariate logistic | product of two univariate exponential densities |
| Specific. L | $\frac{L^S L^D}{L^S + L^D}$ | $\frac{L^S + L^D - \sigma \ln \left(e^{\frac{L^S}{\sigma}} + e^{\frac{L^D}{\sigma}} \right)}{\sigma^2}$ | $\sigma^4 \frac{L^S L^D}{L^S + L^D}$ |
| | ($L > 0$; compare eq. 3.2.3) | ($L > 0$; compare eq. 3B.9) | ($L > 0$; proof to the reader) |
| u^* : | $\frac{1}{2}$ | $1 - \frac{1}{\sigma^2} + \frac{\ln 2}{\sigma L^S}$ | $1 - \frac{1}{2} \sigma^4$ |

1) For an explanation of the symbols, see the main text.

3.3 Empirical results of nested testing

In this section, I consider the relative performance of the nested employment functions as presented in Table 3.1. To this purpose, I use the familiar apparatus for nested testing (the likelihood ratio test) on the four types of model specifications. The model specifications are numbered from 1 to 4 depending on the type of employment function (see Table 3.2). For estimation purposes, I use two sets of data. The first set refers to the Netherlands as a whole. The second set relates to the two Dutch construction sectors: the civil and utility sector (c&u) and the roads and waterworks sector (r&w). The decision to include the construction sectors in the empirical analysis rests on several considerations. I figured it would be wise not to rely solely on the estimation results for the Netherlands as a whole to draw inferences about the feasibility of the distributional assumptions, but to include more disaggregated situations. To choose the sectoral dimension seemed quite natural, but then the question is which sectors of economic activity should be considered? It turned out that the two construction sectors are particularly adequate. There are two reasons. The first is that labour supply and demand in the construction industry are quite well-defined and as such readily interpretable concepts.

Figure 3.2 The aggregate employment function $\frac{L}{L^s} = \left[\left[\frac{L^v}{L^s} \right]^{-\beta} + 1 \right]^{-1/\beta}$ for different values of β



This is due to the fact that the unemployment and vacancy data, which are by occupation, can be transformed adequately to the sectoral data, because of the close and rather exclusive link between wanted occupation and sector of activity (Romme and Wolfs, 1986). In turn this property indirectly follows from the relative independence of the labour market in the construction sectors compared to other sectors, indicating a large “*between-sector mismatch*”. Of course one should differentiate by type of occupation. The large *between-mismatch* applies much more to construction workers than administrative or staff (about 20 per cent in both sectors). There is no other sector of economic activity, where the sectoral ties of labour supply and demand by occupation are so close for such a substantial number, and where sectoral supply and demand are so well-defined.

Table 3.2 Four employment model specifications for nested testing

| | $\beta_i \neq 1$ | $\beta_i = 1$ |
|------------|--------------------|-----------------|
| $a \neq 0$ | Model 1 =Weibull | Model 3 |
| $a = 0$ | Model 2 =Lognormal | Model 4 =Pareto |

The second reason pertains to the availability of expert opinion on the construction sectors allowing an additional check on the estimation results. This expert opinion emphasises the relatively small *within-sector mismatch* for each of the two construction sectors separately. The labour market in r&w sector is envisaged as largely self-contained: either labour supply or labour demand are satisfied for the greater part. The c&u sector is also considered rather homogenous, though to a lesser extent than r&w.⁴⁸ (EIB, 1989a and 1989b). The estimation results for the Netherlands as a whole should provide the relevant reference material.

In Annex 3D a complete description of the data used is given. Some aspects about the quality of the data deserve special attention here. The data are from the registers of the Labour Exchange Offices and as such they are not corrected for the systematic bias (see also Chapter 1 and Chapter 4). Although, this is an unavoidable inadequacy, correcting the sector data is impossible. This should be kept in mind when interpreting the estimation results. Another shortcoming concerns the somewhat outdated data set. The more recent observations could not be reconstructed from the available information. Although there are improved data on the Netherlands as a whole I accepted a data set for the Netherlands comparable to

⁴⁸ On the basis of this expert opinion, only the Weibull or the log-normal employment function are serious candidates, since they only include the Leontieff structure as a special case.

the construction sector data, for reasons of consistency. So the estimation outcomes for the construction sectors and for the Netherlands can be compared more purely. My hypothesis is that the mismatch for the Netherlands as a whole is larger. In the next section, where the consistency argument does not apply, I use the corrected data for the Netherlands and a data set with more recent observations.

The employment functions based on the various distributions are estimated by using non-linear least squares with an additive normal disturbance term. The disturbances reflect errors in the distribution assumption. Estimation of model 1 revealed the occurrence of severe autocorrelation in the disturbances. As mentioned before in Subsection 3.2.1, this is a common result in the estimation of employment functions. However, to legitimately employ the likelihood ratio test, it is necessary to correct for autocorrelation. I had to impose a second-order autocorrelation process in the disturbances in order to yield a satisfactory fit for all equations. Table 3.3 presents the estimation results for the four types of employment equations and for the three observation domains. The parameter \mathbf{a} in the Weibull specifications is insignificant for all applications. The first-order autocorrelation parameters, κ_1 , are always significant, whereas the second-order autocorrelation parameter, κ_2 is insignificant only in the application for the Netherlands and for the c&u model 1 and 2. The mismatch parameter β_1 is significant except for model 1 for the Netherlands. Its estimated value is between 14 and 20 for the Netherlands, 22 for c&u and 34 for r&w.

The implied structural unemployment rates cannot be derived straightforwardly from equations of Table 3.1, because a second order autoregression process is imposed in the estimation. Taking the log-normal employment function (\mathbf{a} is after all insignificant), the equation for u^* shows:

$$\begin{aligned}
 u_t^* = & 1 - \kappa_1 x_{t-1} - \kappa_2 x_{t-2} - 2^{-\frac{1}{\beta_1}} + \kappa_1 \left(1 + (y_t - 1)^{-\beta_1} \right)^{-\frac{1}{\beta_1}} \\
 & + \kappa_2 \left(1 + y_t - 2^{-\beta_1} \right)^{-\frac{1}{\beta_1}} \tag{3.3.1}
 \end{aligned}$$

where κ_1 and κ_2 are the parameters of the autocorrelation process, x_t stands for $\frac{L_t}{L_t^D}$ and y_t stands for $\frac{L_t^S}{L_t^D}$. It can easily be shown that the right-hand expression of converges to $1 - 2^{-\frac{1}{\beta_1}}$ when $L_t^S = L_t^D$ holds for every period. In that sense the aggregate rate of structural unemployment for the Netherlands would converge to 4.2%, whereas the corresponding figures for c&u are 3.1%, and for r&w 2.0%.⁴⁹

⁴⁹ Note that these results are consistent with Table 2.1 for the exact specifications.

Table 3.3 Estimation results of nested model specifications for the Netherlands (1960-1985) and the c&u and r&w construction sectors (1960-1983) ¹

| <i>Model variant</i> | <i>Parameter</i> | <i>estimates²</i> | | | <i>log L</i> | <i>DW</i> |
|----------------------|------------------|---|--------------|------------------|--------------|-----------|
| | a | β_1 | κ_1 | κ_2 | | |
| <i>Dutch economy</i> | | | | | | |
| Model 1 | 1208.7 (0.6) | 14.1 (1.5) | 1.5 (7.7) | -0.5 (-2.7) | 134.5 | 2.3 |
| Model 2 | - | 19.5 (9.8) | 1.4 (7.7) | -0.5 (-2.6) | 134.4 | 2.1 |
| Model 3 | | Results are not available ³⁾ | | | | |
| Model 4 | - | - | 1.7 (9.7) | -0.7 (-3.9) | 88.1 | 1.4 |
| <i>c&u</i> | | | | | | |
| Model 1 | 10.0 (0.01) | 22.1 (7.3) | 0.9 (3.8) | -0.03 (-0.15) | 94.6 | 2.0 |
| Model 2 | - | 22.0 (7.3) | 0.9 (3.8) | -0.03 (-0.15) | 94.6 | 2.0 |
| Model 3 | 30049 (0.6) | - | 1.8 (9.0) | -0.8 (-4.0) | 59.7 | 1.9 |
| Model 4 | - | - | 1.8 (9.8) | -0.8 (-4.7) | 59.5 | 1.9 |
| <i>r&w</i> | | | | | | |
| Model 1 | 32.1 (0.002) | 34.3 (5.1) | 1.0 (4.2) | -0.1 (-0.6) | 93.4 | 2.1 |
| Model 2 | - | 34.0 (5.3) | 1.0 (4.3) | -0.1 (-0.6) | 93.4 | 2.1 |
| Model 3 | 9945.1 (1.1) | - | 1.7 (8.1) | -0.7 (-3.3) | 59.8 | 2.1 |
| Model 4 | - | - | 1.7 (8.3) | -0.8 (-3.5) | 58.8 | 2.0 |

1. t-values in parentheses.

2. All models are estimated by dividing the left and right hand side of the equation by L^s . Additive disturbances were assumed to follow an AR(2) process in order to remove serial correlation. To be more specific, the imposed specification shows:

$$\xi_t = \kappa_1 \xi_{t-1} + \kappa_2 \xi_{t-2} + u_t, \text{ Models 1 to 4 correspond to the model specifications of Table 3.2.}$$

3.2. Restricted optimisation with respect to a was carried through in case of *model 1* (c&u and r&w).

3. Optimisation runs into trouble.

However, since aggregate supply and demand were unequal in the past year, these figures cannot be used as proper measures of the current structural unemployment rate. The figures that can be calculated from (3.3.1) are probably more accurate. For 1983, for instance, the structural unemployment rate is 4.4% for the Netherlands, 3.5% for c&u and 2.3% for r&w. These figures are higher because in the early eighties labour supply considerably exceeded labour demand.

The estimates for u^* are consistent with expert opinion. They indicate a small *within-sector mismatch* for each of the two construction sectors. The labour market of the r&w sector approaches the Leontieff structure the most. The estimates of u^* for the construction sectors are significantly lower than for the Netherlands as a whole.

I now come to the results of nested testing. The model selection test is based on the values of twice the log likelihood ratio, which is asymptotically chi-squared distributed. More formally, let $L(u)$ be the value of the likelihood function for the unrestricted model [*model 1*, extended with the AR(2) equation described in footnote 2 of Table 3.3)] and $L(r)$ the value of the likelihood function of one of the restricted models (*model 2* to *4*, also extended with the AR(2) equation). Then asymptotically holds:

$$2\left(\log(L(u) - \log L(r))\right) \approx X_{(j)}^2$$

where $X_{(j)}^2$ is a chi-squared distribution with j degrees of freedom; j is the number of restrictions on the parameters. The likelihood ratio tests for the four nested models are presented in Table 3.4 for the Netherlands, and in Table 3.5 for the c&u and r&w sectors. Not surprisingly, given the estimates for u^* , *model 4* is rejected against the other models in all cases. For the Netherlands as a whole, the evidence points towards not rejecting the employment equation based on the log-normal distribution (*model 2*) against the employment equation based on the Weibull distribution (*model 1*) at a significance level of 2.5%. *Model 2* is far superior to *model 4*. Since the estimation of *model 3* runs into optimisation problems, the relative performance of *models 1, 2* and *4* with respect to *model 3* could not be determined.

The two construction sectors yield similar results. *Model 2* does not perform worse than *model 1*. Moreover, *model 3* does not show a better performance than *model 4*. Hence, the parameter α could rightfully be left out in the regressions. However, *model 3* is rejected against *model 1*, and *model 4* is rejected against both *model 1* and *model 2*, which implies that the restriction of $\beta_1=1$ is rejected.

In conclusion, I can say that both for the data of the Netherlands as a whole and for the two construction sectors, the employment specification based on the Weibull or Pareto distribution should be rejected in favour of the employment function based on the log-normal distribution. There is, however, one reservation to this conclusion, which has to do with the differences in domains between the

Table 3.4 Empirical discrimination between employment function based on Weibull, log-normal and Pareto distribution by means of a likelihood ratio test, 1960-1985 (the Netherlands)

| | | employment function based on | | |
|-------------------------------|---|---|----------------------------|---|
| | | Weibull p.d.f. | | log-normal p.d.f. |
| e.q. tested | against | | | |
| | Pareto p.d.f. | $\left[a + (L^b - a)^e + (L^s - a)^e \right]^{-\frac{1}{e}}$ | | $\left[L^{b^*} + L^{s^*} \right]^{-\frac{1}{e}}$ |
| employment function based on | $\frac{L^{b^*}}{L^b + L^{s^*}}$ | $\chi^2 = 92.73$ | | $\chi^2 = 92.55$ |
| | $\left[\frac{L^{b^*}}{L^b + L^{s^*}} + 1 \right]^{-\frac{1}{e}}$ | | $\chi^2 = 0.1821$ | |
| Conclusion | | Pareto rejected | log-normality not rejected | Pareto rejected |
| degrees of freedom (1% level) | | 2 | 1 | 1 |

Table 3.5 Empirical discrimination between employment function based on Weibull, log-normal and Pareto distribution by means of a likelihood ratio test for the c&u and r&w sector, 1960-1985 (r&w sector between brackets)

| e.g. tested against | | employment function based on | | | | |
|---------------------------------|---------------------|-------------------------------|--|------------------------------|--|---------------------|
| | | Weibull p.d.f. | | log-normal p.d.f. | $\beta = 1$ in Weibull empl. f. | |
| employment function based on | Pareto p.d.f. | $\frac{L^a + L^b}{L^a + L^b}$ | $[a+(L^b-a)^a+(L^b-a)^a]^{-\frac{1}{a}}$ | $[L^b + L^a]^{-\frac{1}{a}}$ | $a + \frac{(L^b-a)(L^b-a)}{(L^b-a)+(L^b-a)}$ | |
| | χ^2 | -70.22 (69.26) | | -70.22 (69.26) | -0.37 (1.36) | |
| | in Weibull empl. f. | | $\chi^2 = 70.22$ (67.90) | | | |
| | log-normal p.d.f. | | | $[L^b + L^a]^{-\frac{1}{a}}$ | | |
| | χ^2 | | -0.0008 (0.0016) | | | |
| Conclusion | | Pareto rejected | Weibull empl. f. rejected | log-normality not rejected | Pareto rejected | Pareto not rejected |
| degrees of freedom (2.5% level) | | 2 | 1 | 1 | 1 | 1 |

distributions. The Weibull and the log-normal distribution could be very much alike when the differences in domains are accounted for. If this is true, the identification of the parameter α is weak, as is the test to discriminate between the Weibull and the log-normal employment function. Within the present context it would, however, go too far to investigate this matter any further. Simulation experiments could be helpful to arrive at more conclusive results.

3.4 Normal and log-normal distribution against general form⁵⁰

3.4.1 Introduction

In Section 3.3, I considered the appropriateness of nested aggregate employment functions. Since the aggregate employment function based on the normal distribution did not fit the nested testing procedure, I could not test for the appropriateness of the normal distribution. However, in addition to the log-normal distribution, the normal distribution has been applied in empirical models of markets in disequilibrium, in particular within the simulation variant. It also has a theoretical underpinning in the theory of stochastic processes. Therefore, it cannot simply be bypassed.

In this section, I will develop a procedure to test the feasibility of the aggregate employment function based on the normal distribution assumption. (*c.f.* Subsection 3.4.2). To that purpose a certain type of distribution function is introduced, which contains the normal distribution as a special case. This makes a specific form of testing for normality possible. The same type of procedure is also carried through for the log-normal distribution (*c.f.* Subsection 3.4.3). All this allows a more thorough judgement about the feasibility of this distribution. Other than in the previous section, the employment functions are “exact” and not approximate functions.

By comparing the results of feasibility testing for the two distributional assumptions, it is possible to give some order of preference. Both equations have the convex shape in common and satisfy equal continuity properties (see Chapter 2 for the log-normal function). They differ in the degree of homogeneity: the log-normal employment function (*c.f.* eq. 2.2.8) is homogeneous of degree one in L^S and L^D , whereas the one based on the normal distribution (*c.f.* eq. 2.3.2) is not. Testing the normality and the log-normality hypothesis as proposed in this section may give an indication of which homogeneity concept serves the data best.

⁵⁰ This section is a revision of Bierings and Muysken (1989).

3.4.2 Weighted normal employment function

In Section 2.3 (eq. 2.3.2), the employment function when $h(L^L, L^d)$ is normal, has been shown to be equal to:

$$L = \Phi\left(\frac{L^D - L^S}{\sigma^*}\right)L^S + \Phi\left(\frac{L^S - L^D}{\sigma^*}\right)L^D - \frac{\sigma^*}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{L^D - L^S}{\sigma^*}\right)^2} \tag{3.4.1}$$

with the well-known symbols.

To test the normality hypothesis, consider the weighted sum of two normal distribution functions, $h(.,.)$:

$$h(.,.) = \Gamma h_1(L^S, L^d) + (1 - \Gamma) h_2(L^S, L^d) \quad 0 \leq \Gamma \leq 1 \tag{3.4.2}$$

where $h_1(.,.)$ and $h_2(.,.)$ have identical means, L^S and L^D , and a diagonal covariance matrix with variances $\sigma_{1s}^{*2}, \sigma_{1d}^{*2}, \sigma_{2s}^{*2}, \sigma_{2d}^{*2}$. This implies that $h_1(.,.)$ and $h_2(.,.)$ only differ with respect to their variances. To facilitate further elaborations, I additionally restrict the variances to be: $\sigma_{2s}^* = c \sigma_{1s}^*$ and $\sigma_{2d}^* = c \sigma_{1d}^*$ with $c \geq 1$.

It may be verified that as a result the distribution $h(.,.)$ lies somewhere between the normal and the t-distribution function.⁵¹ Note that the kurtosis of this distribution for values of Γ between 0 and 1 indicates a longer-tailed density function than the normal distribution function. Symmetry, however, is guaranteed. Evaluating the integrals of (2.2.5) with $h(.,.)$ specified according to (3.4.2), yields the following expression

for aggregate employment $\left(\text{where } \frac{L^D - L^S}{\sigma^*} \text{ is substituted by } y \right)$:

$$L = \Gamma \Phi(y) L^S + (1 - \Gamma) \Phi\left(\frac{y}{c}\right) L^S + \Gamma \Phi(-y) L^D + (1 - \Gamma) \Phi\left(\frac{-y}{c}\right) L^D - \Gamma \sigma^* \psi(y) - (1 - \Gamma) \sigma^* \psi\left(\frac{y}{c}\right) \quad 0 \leq \Gamma \leq 1 \tag{3.4.3}$$

where: σ^* now equals: $\sigma\left(L_1^d - L_1^s\right) = c \sigma\left(L_2^d - L_2^s\right)$.⁵²

⁵¹ Note that for j between one and zero and $\sum_{j=1}^n \Gamma_j = 1$, the expressions $\sum_{j=1}^n \Gamma_j \psi(x_j)$

with ψ is the normal density - is distributed according to the t-distribution when n approaches infinity.

⁵² The subscripts of L^L and L^d refer to the corresponding distributions of eq. 3.4.2.

Since (3.4.1) is nested in (3.4.3) it is possible to test for normality using the familiar likelihood ratio technique for nested models. The test decides on the significance of c and Γ in (3.4.3): when $\Gamma=0$ or 1, or $c=1$, normality is accepted.

3.4.3 Weighted log-normal employment function

When $h(L^S, L^D)$ is log-normal, the aggregate employment function shows (c.f. eq. 2.2.8):

$$L = \Phi\left(\frac{\ln L^D - \ln L^S}{\sigma^*} - \frac{1}{2} \sigma^{*2}\right) L^S + \Phi\left(\frac{\ln L^S - \ln L^D}{\sigma^*} - \frac{1}{2} \sigma^{*2}\right) L^D \quad (3.4.4)$$

Repeating the above procedure for the log-normal case - i.e. with the same assumptions with respect to variances and covariances and using instead of the weighted normal, the weighted log-normal⁵³ - the aggregate employment function

then shows (where $\frac{\ln L^D - \ln L^S}{\sigma^*}$ is substituted by x):

$$\begin{aligned} L \dots & \Gamma \Phi\left(x - \frac{1}{2} \sigma^{*2}\right) L^S + (1 - \Gamma) \Phi\left(\frac{x}{c} - \frac{1}{2} c \sigma^{*2}\right) L^S + \Gamma \Phi\left(-x - \frac{1}{2} \sigma^{*2}\right) L^D \\ & + (1 - \Gamma) \Phi\left(\frac{-x}{c} - \frac{1}{2} c \sigma^{*2}\right) L^D \quad 0 \leq \Gamma \leq 1 \quad (3.4.5) \end{aligned}$$

where: σ^* now equals : $\sigma\left(\ln L_1^D - \ln L_1^S\right) = c \sigma\left(\ln L_2^D - \ln L_2^S\right)$.⁵⁴

Now, one can test for log-normality since (3.4.4) is nested in (3.4.5): When $\Gamma=0$ or 1, or $c=1$ log-normality is accepted.

3.4.4 Empirical results

This subsection presents the empirical results. These results are subsequently used to test the normality and the log-normality hypothesis. I used data on labour supply, labour demand and employment for the Dutch economy covering the period 1960-

⁵³ For convenience, I use the same symbols for the parameters as in the normal case. Other than in the "normal case", the weighted log-normal as well as the log-normal are not symmetrical.

⁵⁴ The subscripts of L' and L^d refer to the corresponding distributions of eq. 3.4.2, now assuming $h_1(\dots)$ and $h_2(\dots)$ to be the log-normal distributions.

1991. The argument to maintain consistency with the construction industry observations is no longer valid in this section. As a result, the data set contains more recent observations and is of higher quality because of the correction for systematic biases of the register data. The data are obtained from Muysken *et al* (1994).⁵⁵ The estimations were carried out using non-linear full information maximum likelihood techniques. Additive (normally distributed) disturbance terms were assumed to account for aggregation or misspecification errors of the structural equations (3.4.1), (3.4.3), (3.4.4) and (3.4.5). Table 3.6 shows the results of the estimations. The results indicated by number 1 stem from the estimation of (3.4.1) for the normal case and (3.4.4) for the log-normal case. As the table suggests, these specifications suffer from severe first order autocorrelation. This situation hardly improves when estimating (3.4.3) and (3.4.5) (not reported in the table). As in Section 3.3 it is necessary to respecify the equations in order to dispose of the autocorrelated errors. The table shows the various alternatives I experimented with.

The first alternative specification indicated by number 2, defines the parameter σ^* of (3.4.1) and (3.4.4) to be a linear function of a constant, and of time. Hence:

$$\sigma^* = \sigma_0 + \sigma_1 t \quad (3.4.6)$$

(+)

The trend-term is *ad hoc*. It is usually motivated by referring to influences of technological innovation (see Section 2.5). Its expected sign is positive. Because (3.4.6) underlies some theoretical notion, it should be preferred above other experiments changing the dynamic structure of the equation by imposing some arbitrary autoregressive structure. These dynamic adaptations are not theoretically founded.⁵⁶ As the table shows, the inclusion of the trend-term still is not sufficient to remove the serial correlation. In addition, I therefore experimented with an AR(1) structure on the disturbance terms (ξ):

$$\xi_t = \kappa \xi_{t-1} + u_t \quad (3.4.7)$$

with $u_t \sim N(0,1)$

When (3.4.6) and (3.4.7) are used the autocorrelation completely disappears. (see the results in the table indicated by the variants 3 and 4). The results presented under 3 refer to (3.4.1) for the normal case and to (3.4.3) for the log-normal case.

⁵⁵ See Annex 3D for a justification.

⁵⁶ In Section 3.3, I did not experiment with β_1 (the equivalent of σ^*) as a function of time, since this would have disturbed the nested testing.

Table 3.6 Estimation results for equations (3.4.1), (3.4.3), (3.4.4) and (3.4.5), the Netherlands, 1960-1991

| Model variant | Parameter estimates ¹ | | | | | log L | DW |
|--|----------------------------------|-------------------------------|--------------|---------------|---------------|--------|-----|
| | σ_0 | σ_1 | κ | c | Γ | | |
| Normal case | | | | | | | |
| 1. σ_0 | 249.0 (28.2) | - | - | - | - | -134.0 | 0.3 |
| 2. σ_0, σ_1 | 183.7 (12.5) | 13.9 (16.2) | - | - | - | -127.8 | 0.4 |
| 3. $\sigma_0, \sigma_1, \kappa$ | 189.6 (5.4) | 15.0 (8.0) | 0.7 (7.8) | - | - | -111.6 | 1.9 |
| 4. $\sigma_0, \sigma_1, \kappa, c, \Gamma$ | 181.3 (5.1) | 14.8 (7.3) | 0.7 (7.7) | 48.1 (0.1) | 1.0 (15.3) | -109.7 | 1.9 |
| Log-normal case | | | | | | | |
| 1. σ_0 | 0.08 (20.4) | - | - | - | - | -156.2 | 0.1 |
| 2. σ_0, σ_1 | 0.05 (15.7) | $2.3 \cdot 10^{-3}$ (14.1) | - | - | - | -125.6 | 0.5 |
| 3. $\sigma_0, \sigma_1, \kappa$ | 0.05 (6.7) | $2.2 \cdot 10^{-3}$ (15.8) | 0.7 (6.2) | - | - | -112.1 | 1.9 |
| 4. $\sigma_0, \sigma_1, \kappa, c, \Gamma$ | 0.05 (6.8) | $2.3 \cdot 10^{-3}$ (8.9) | 0.7 (6.5) | 21.8 (0.1) | | 110.8 | 2.0 |

1. t-values in parentheses.

The results indicated by 4 represent those for (3.4.3) (the normal case) and for (3.4.5) (the log-normal case). The overall picture emerging from the table is that the estimated parameter values are roughly of the same order of magnitude over the alternative specifications tested. The relatively great jump in the value of σ_0 from 1 to 2 in both cases, is mainly the result of the significant influence of the trend-term in σ^* . Moreover, the results on σ^* are consistent with those of Kooiman and Kloek (1979) using data of the Dutch economy for the period 1948-1973. Their employment function, not corrected for autocorrelation, yields a value of σ_0 equal to 181.7 in the normal case and 0.048 in the log-normal case (compare Table 2.1, this corresponds to a structural unemployment rate of 2% on average for the log-normal case). In their specification with autocorrelation, the estimated value of the parameter κ of the AR(1) specification of the disturbances equals 0.54.

Now that the serial correlation problem has been settled, I am able to test the normality and log-normality hypothesis. The nested test is applied to specifications (3.4.1) and (3.4.3) for the normal, and to (3.4.4) and (3.4.5) for the log-normal case. Again as in Section 3.3, a likelihood ratio test is used, now to value the **overall** significance of the normality and log-normality hypothesis. The unrestricted model is described by the equations (3.4.3) or (3.4.5), both extended with (3.4.6) and (3.4.7); the restricted model corresponds to the equations (3.4.1) and (3.4.4) (both with extensions). In Table 3.7 the values of X^2 are presented together with the resulting conclusions about acceptance or rejection of the normality or log-normality hypothesis.

The results indicate that normality as well as log-normality is accepted at the 1% level. They, however, tend to be slightly more in favour of the log-normal density due to the relatively low value of the log-likelihood ratio.

3.5 Concluding remarks

This chapter demonstrated that it is legitimate to use the log-normal distribution to describe the distribution of micro supply and demand in the context of aggregate employment functions. Moreover, the procedure used to test the normality hypothesis suggests that the same is true for the normal distribution. The employment functions based on the Weibull and Pareto distributions do not provide serious alternatives.

The chapter consists of a theoretical part and an empirical part in which the log-normality and normality hypothesis is tested. To validate the use of the log-normal distribution, I inquired into the theory of stochastic processes. Paradoxically, the assumption of autocorrelated growth or entry/exit in the stochastic process - which was more plausible given the high autocorrelation experienced in most empirical studies estimating aggregate employment functions - disqualified the log-normal distribution in favour of Pareto's Law as limiting distribution, whereas at the same time the employment function arising from Pareto's Law was much too simple to be realistic. I concluded that the process monitored by autocorrelated growth or entry/exit is rigidly forced into the wrong direction, whereas the Gibrat process underlying the log-normal distribution) is more flexible despite of its more restrictive assumptions.

I have also shown that several disjoint distributions, gave rise to nested employment functions. This opened up the possibility to test the log-normality assumption by way of the likelihood ratio test. To investigate the feasibility of the log-normal distribution, two different approaches were presented. The first method integrates the employment functions based on the Weibull, log-normal and Pareto distributions in the nested testing. The employment functions based on the normal, exponential or logistic distribution could not be integrated in the nested testing.

Table 3.7 Empirical discrimination between employment function based on weighted log-normal and normal distribution by means of a likelihood ratio test, 1960-1991 (the Netherlands)

| | | employment function based on | |
|-------------------------------|------------------|------------------------------|----------------------------|
| | | weighted normal p.d.f. | weighted lognormal p.d.f. |
| c.q. tested | against | | |
| | normal p.d.f. | $\chi^2 = 3.8$ | |
| employment function based on | normal p.d.f. | | |
| | lognormal p.d.f. | | $\chi^2 = 2.6$ |
| Conclusion | | normality not rejected | log-normality not rejected |
| degrees of freedom (1% level) | | 2 | 2 |

The employment function based on the Weibull distribution is the most general, and of the C.E.S. type, as is the employment function based on the log-normal. The most restrictive is the employment function based on Pareto's Law. The empirical application is for the labour market of the Netherlands as a whole and for two sectors in the construction industry. Including the construction sectors as observation domains allows me to have a consistency check on the estimation results with the available expert opinion. Besides, the feasibility test is not restricted to the situation of the Netherlands as a whole, but also applies to more disaggregated situations. An overall conclusion is that the employment function based on the log-normal is preferred in all cases. There is, however, one reservation to this conclusion which could make the tests to discriminate between the Weibull and the log-normal distribution weak. Namely, both distributions may be very much alike when the differences in domains are accounted for. If this is indeed the case then the identification of the parameter α is weak. I did not go any further into this matter, but simulation experiments could, be helpful to arrive at more conclusive results.

Within the second method, a distribution function is introduced that contains the log-normal as a special case. This makes a specific form of testing for log-normality possible by confronting the corresponding employment functions. Using data for the labour market of the Netherlands as a whole, the nested testing substantiated the use of the log-normal distribution. The second method is also applied for the normal distribution. Because the normal postulate is adopted in empirical models of markets in disequilibrium (in particular the simulation-based variant), its feasibility is worth testing. Using the same data set as for the application based on the log-normal distribution, the nested testing validated the use of the normal distribution too. Perhaps, the results are slightly in favour of the log-normal distribution. As a result, the degree of homogeneity is not sensitive to differences in the degree of homogeneity between the two types of employment functions. It is evident that the techniques to test the distributional assumptions are quite specific, and provide only first results on the subject. Future research could be directed towards finding a super distribution that embeds a large family of common distributions, and derive the corresponding aggregate employment function. The nested tests can be used to test the aggregate employment function based on the super distribution against the other employment functions. Then, more general results are obtained. Another possibility is to consider the use of non-nested tests.

Though, the results in this chapter lend support to the conclusion that there is no need to correct the usual assumptions imposed in the empirical applications of markets in disequilibrium. As a matter of fact, in my applications of the analytically and simulation based variants (Chapter 4 and Chapter 6, respectively), I preserve the usual practise.

ANNEX 3A Derivation of the aggregate employment function for the bivariate Pareto distribution transformed to the origin

The bivariate Pareto distribution $h(L^s, L^d)$ is defined for $L^s > \theta_s$ and $L^d > \theta_d$. I start from the transformed $h(L^s + \theta_s, L^d + \theta_d)$ for $L^s > 0$ and $L^d > 0$. The relationship for aggregate employment is equal to:

$$L = \int_0^{\infty} \int_0^{\infty} L^s h(L^s + \theta_s, L^d + \theta_d) dL^d dL^s + \int_0^{\infty} \int_0^{\infty} L^d h(L^s + \theta_s, L^d + \theta_d) dL^s dL^d$$

$$= I_1 + I_2 \tag{3A.1}$$

where $h(\dots)$ is the bivariate Pareto distribution transformed to the origin. It is defined as:

$$h(L^s + \theta_s, L^d + \theta_d) = \alpha(\alpha+1)(\theta_d \theta_s)^{\alpha+1} (\theta_d L^s + \theta_s L^d + \theta_s \theta_d)^{-(\alpha+2)} \tag{3A.2}$$

with means, $E(L^s)$ and $E(L^d)$, $\frac{\theta_s}{\alpha-1}$ and $\frac{\theta_d}{\alpha-1}$, respectively and covariance matrix:

$$\begin{pmatrix} \frac{\alpha \theta_s^2}{(\alpha-1)^2 (\alpha-2)} & \frac{\theta_s \theta_d}{(\alpha-1)^2 (\alpha-2)} \\ \frac{\theta_s \theta_d}{(\alpha-1)^2 (\alpha-2)} & \frac{\alpha \theta_d^2}{(\alpha-1)^2 (\alpha-2)} \end{pmatrix}$$

I first evaluate I_1 :

$$I_1 = \alpha(\alpha+1)(\theta_d \theta_s)^{\alpha+1} \int_0^{\infty} L^s \left(\int_{L^s}^{\infty} \theta_d L^s + \theta_s L^d + \theta_s \theta_d \right)^{-(\alpha+2)} dL^d dL^s$$

$$= C \int_0^{\infty} L^s I_{1,1} dL^s =$$

$$\begin{aligned}
&= \frac{C}{(\alpha+1)\theta_s} \int_0^{\infty} L^s (\theta_d L^s + \theta_s L^s + \theta_d \theta_s)^{-(\alpha+1)} dL^s \quad (\alpha > 1) \\
&= \alpha \theta_s^\alpha \theta_d^{\alpha+1} \int_0^{\infty} L^s ((\theta_d + \theta_s)L^s + \theta_d \theta_s)^{-(\alpha+1)} dL^s \\
&= C' I_{1,2}
\end{aligned} \tag{3A.3}$$

$I_{1,2}$ is evaluated by means of integration by parts in order to obtain:

$$I_1 = \frac{\theta_s \theta_d^2}{(\theta_d + \theta_s)^2 (\alpha-1)} \tag{3A.4}$$

I_2 is evaluated similarly as I_1 and shows:

$$I_2 = \frac{\theta_s^2 \theta_d}{(\theta_d + \theta_s)^2 (\alpha-1)} \tag{3A.5}$$

$$E(L^s) = \frac{\theta^s}{\alpha-1} \text{ and } E(L^d) = \frac{\theta^d}{\alpha-1}$$

Finally, substituting the equations for $E(L^s)$ and $E(L^d)$ yields the following expression for aggregate employment:

$$L = I_1 + I_2 = \frac{L^s L^d}{L^s + L^d} \tag{3A.6}$$

ANNEX 3B Derivation of the aggregate employment function for the bivariate logistic distribution

Here aggregate employment is defined as:

$$L = \int_{-\infty}^{\infty} \int_{L^s}^{\infty} L^s h(L^s, L^d) dL^d dL^s + \int_{-\infty}^{\infty} \int_{-\infty}^{L^d} L^d h(L^s, L^d) dL^s dL^d = I_1 + I_2 \quad (3B.1)$$

where distribution $h(L^s, L^d)$ is the bivariate logistic density. For reasons of simplicity and efficient exposition, I define the following functions of L^s and L^d , together with $\sigma_s = \sigma_d = \sigma$ (with σ_s and σ_d representing the standard deviation of L^s and L^d , respectively):

$$f(L^s) = -\frac{L^s - \theta_s}{\sigma} \quad (3B.2)$$

$$g(L^d) = -\frac{L^d - \theta_d}{\sigma} \quad (3B.3)$$

$h(L^s, L^d)$ is defined as:

$$h(L^s, L^d) = \frac{2e^{f(L^s)+g(L^d)}}{\sigma^2 (1 + e^{f(L^s)} + e^{g(L^d)})^3} \quad (3B.4)$$

It has means θ_s and θ_d and, variance $\frac{\sigma^2 \pi^2}{3}$ for both supply and demand.

The definition of employment according to (3B.1) is based on the assumption that the probability mass of $h(.,.)$ is negligible for negative variables of L^s and L^d (see for an identical assumption Kooiman and Kloek, 1979). First, I evaluate I_1 :

$$\begin{aligned}
 I_1 &= \frac{2}{\sigma^2} \int_{-\infty}^{\infty} L^s e^{f(L^s)} \int_{L^s}^{\infty} \left(\frac{e^{g(L^d)}}{(1 + e^{f(L^s)} + e^{g(L^d)})^3} dL^d \right) dL^s \\
 &= \frac{1}{\sigma^3} \left(\int_{-\infty}^{\infty} \frac{L^s e^{f(L^s)}}{(1 + e^{f(L^s)})^2} - \frac{1}{(1 + e^{f(L^s)} + e^{g(L^d)})^2} dL^s \right) \\
 &= \frac{1}{\sigma^3} \left(\int_{-\infty}^{\infty} \frac{L^s e^{f(L^s)}}{(1 + e^{f(L^s)})^2} dL^s - \int_{-\infty}^{\infty} \frac{L^s e^{f(L^s)}}{(1 + e^{f(L^s)} + e^{g(L^d)})^2} dL^s \right) \\
 &= \frac{1}{\sigma^3} (I_{1,1} - I_{1,2}) \tag{3B.5}
 \end{aligned}$$

Subsequently, I evaluate $I_{1,1}$ and $I_{1,2}$. $I_{1,1}$ can be worked out using integration by parts to show:

$$I_{1,1} = \lim_{a \rightarrow \infty} \sigma \left(\left[\frac{L^s}{1 + e^{f(L^s)}} \right]_{-a}^a - \int_{-a}^a \frac{1}{1 + e^{f(L^s)}} dL^s \right)$$

The second integral between brackets is equal to (Spiegel 1968, equation 14.515):

$$e^K \int_{-a}^a \frac{1}{e^K + e^{-\frac{L^s}{\sigma}}} dL^s \quad (K = -\frac{\theta_s}{\sigma})$$

$$= e^K \left[e^K L^s + \sigma e^K \ln \left(e^K + e^{-\frac{L^s}{\sigma}} \right) \right]_{-a}^a$$

Substituting this result in $I_{1,1}$ and writing $f(L')$ explicitly yields:

$$\begin{aligned}
 I_{1,1} = \sigma \lim_{a \rightarrow \infty} & \left(\underbrace{\left[\frac{a}{1 + e^{\frac{-a}{\sigma} + \frac{\theta_s}{\sigma}}} \right]}_{T_1} - a + \underbrace{\left[\frac{a}{1 + e^{\frac{a}{\sigma} + \frac{\theta_s}{\sigma}}} \right]}_{T_1} - \right. \\
 & \left. \underbrace{-\sigma \ln \left[e^{\frac{-\theta_s}{\sigma} + e^{\frac{-a}{\sigma}}} \right]}_{T_1} - a + \underbrace{\sigma \ln \left[e^{\frac{-\theta_s}{\sigma} + e^{\frac{a}{\sigma}}} \right]}_{T_1} \right) = \sigma \theta_s
 \end{aligned} \tag{3B.6}$$

One immediately sees that when a approaches infinity, T_1 and T_2 approach zero and T_3 approaches $-\theta_s$. A perhaps less straightforward result is that T_4 approaches zero too, when a goes toward infinity. Next, I evaluate $I_{1,2}$.

$$\begin{aligned}
 I_{1,2} &= \int_{-\infty}^{\infty} \frac{L^s e^{f(L^s)}}{\left(1 + e^{f(L^s) + g(L^s)} \right)^2} dL^s \\
 &= e^{-f(2\theta_s)} \int_{-\infty}^{\infty} \frac{L^s f(L^s + \theta_s)}{\left(1 + e^{f(L^s + \theta_s)} \left(e^{-f(2\theta_s)} + e^{-f(2\theta_d)} \right) \right)^2} dL^s
 \end{aligned}$$

Using integration by parts and substituting $b = e^{\frac{\theta_s}{\sigma}} + e^{\frac{\theta_d}{\sigma}}$ and $c = e^{\frac{\theta_s}{\sigma}}$, $I_{1,2}$ can be written as:

$$I_{1,2} = \frac{c}{b} \sigma \left(\lim_{p \rightarrow \infty} \left[\frac{L^s}{1 + b e^{-\frac{L^s}{\sigma}}} \right]^p - \int_{-p}^p \frac{1}{1 + b e^{-\frac{L^s}{\sigma}}} dL^s \right)$$

The second integral between brackets is equal to:

$$\left[L^s + \sigma \ln \left(1 + be^{-\frac{L^s}{\sigma}} \right) \right]_{-p}^p$$

After substitution of this result in $I_{1,2}$, working out $I_{1,2}$ leads to:

$$I_{1,2} = \frac{c}{b} \sigma \lim_{p \rightarrow \infty} \left(\underbrace{\frac{p}{1+be^{-\frac{p}{\sigma}}} - p}_{T_1} + \underbrace{\frac{p}{1+be^{\frac{p}{\sigma}}} - p}_{T_2} - \underbrace{\sigma \ln \left(1 + be^{-\frac{p}{\sigma}} \right)}_{T_3} + \underbrace{\sigma \ln \left(1 + be^{\frac{p}{\sigma}} \right)}_{T_4} \right)$$

$$= \frac{c}{b} \sigma \ln b \quad (3B.7)$$

When p approaches infinity, T_5 , T_6 and T_7 go toward zero, and T_8 goes toward $\sigma \ln b$. Substituting the results for $I_{1,1}$ and $I_{1,2}$ in I_1 gives:

$$I_1 = \frac{\theta_s}{\sigma^2} - \frac{c}{b\sigma} \ln b \quad (3B.8)$$

In a similar way I_2 can be determined. Then after some manipulation the aggregate employment function can be shown to be:

$$L = \frac{L^s + L^D - \sigma \ln \left(e^{\frac{L^s}{\sigma}} + e^{\frac{L^D}{\sigma}} \right)}{\sigma^2} \quad (3B.9)$$

One sees that the earlier assumption about the probability mass of the density $h(\cdot, \cdot)$ lays restrictions on the values of the parameter σ for specific values of L^s and L^D , since L cannot be negative. One can easily verify that the first order conditions, *i.e.*

$$\frac{\partial L}{\partial L^s}, \frac{\partial L}{\partial L^D} > 0 \text{ are satisfied.}$$

$$\frac{\partial L}{\partial L^S} = -\frac{1}{\sigma^2} \left(1 - \frac{e^{\frac{L^S}{\sigma}}}{e^{\frac{L^S}{\sigma}} + e^{\frac{L^D}{\sigma}}} \right) \quad (3B.10)$$

$$\frac{\partial L}{\partial L^D} = -\frac{1}{\sigma^2} \left(1 - \frac{e^{\frac{L^D}{\sigma}}}{e^{\frac{L^S}{\sigma}} + e^{\frac{L^D}{\sigma}}} \right) \quad (3B.11)$$

Since both L^S and L^D are positive, both derivatives are positive. It follows that the structural unemployment rate is equal to:

$$u^* = 1 - \frac{1}{\sigma^2} + \frac{\ln 2}{\sigma L^S} \quad (3B.12)$$

From this equation it is clear that the structural unemployment rate is positively related to σ , as was to be expected. For large values of L^S , u^* can be approximated by:

$$u^* = 1 - \frac{1}{\sigma^2} \quad (3B.13)$$

ANNEX 3C The aggregate minimum condition

In this annex, I prove that when the mismatch, σ^* [or: $\text{var}(L^d - L^s)$] is approximately zero (but non-zero), the aggregate employment function:

$$L = \Phi^S L^S + \Phi^D L^D \quad (3C.1)$$

with

$$\Phi^S = \frac{\int_0^\infty \int_0^\infty L^s h(L^s, L^d) dL^d dL^s}{\int_0^\infty \int_0^\infty L^s h(L^s, L^d) dL^d dL^s} \quad \text{and} \quad \Phi^D = \frac{\int_0^\infty \int_0^\infty L^d h(L^s, L^d) dL^s dL^d}{\int_0^\infty \int_0^\infty L^d h(L^s, L^d) dL^s dL^d}$$

independent of the type of distribution used, reduces to:⁵⁷

$$L = \text{Min} \left(L^S, L^D \right) \quad (3C.2)$$

From (3C.1) it follows that:

$$\Phi^S = \frac{\int_0^\infty L^s h(L^s) \left(\int_{L^s}^\infty h(L^d | L^s) dL^d \right) dL^s}{\int_0^\infty L^s h(L^s) dL^s} \quad (3C.3)$$

$$\Phi^D = \frac{\int_0^\infty L^d h(L^d) \left(\int_{L^d}^\infty h(L^s | L^d) dL^s \right) dL^d}{\int_0^\infty L^d h(L^d) dL^d} \quad (3C.4)$$

⁵⁷ The expression for Φ^S and Φ^D follow directly from equation (2.2.5).

From (3C.3)-(3C.4) it then follows that, if $\text{var}(L^d - L^s) \approx 0$ for some value of \mathbf{a} , $P(L^d - L^s = \mathbf{a}) = 1$ holds, hence $L^s \approx L^d + \mathbf{a}$. Then the conditional densities $h(L^d | L^s)$ and $h(L^s | L^d)$ in (3C.3) and (3C.4) concentrate around $L^s - \mathbf{a}$ and $L^d + \mathbf{a}$, respectively. Dependent on the value of \mathbf{a} ($\mathbf{a} = 0, \mathbf{a} > 0, \mathbf{a} < 0$), the following combinations of Φ^S and Φ^D apply:

$$(1) \text{ if } \mathbf{a} > 0 \text{ (or } L^s > L^d \text{):} \left[\begin{array}{l} (i) \int_{L^d}^{\infty} h(L^s | L^d) dL^s \approx 1 \rightarrow \Phi^D \approx 1 \\ (ii) \int_{L^s}^{\infty} h(L^d | L^s) dL^d \approx 0 \rightarrow \Phi^S \approx 0 \end{array} \right. \quad (3C.5)$$

$$(2) \text{ if } \mathbf{a} < 0 \text{ (or } L^d > L^s \text{):} \left[\begin{array}{l} (i) \int_{L^d}^{\infty} h(L^s | L^d) dL^s \approx 0 \rightarrow \Phi^D \approx 0 \\ (ii) \int_{L^s}^{\infty} h(L^d | L^s) dL^d \approx 1 \rightarrow \Phi^S \approx 1 \end{array} \right. \quad (3C.6)$$

$$(3) \text{ if } \mathbf{a} = 0 \text{ (or } L^d = L^s \text{)} \int_{L^d}^{\infty} h(L^s | L^d) dL^s \text{ and } \int_{L^s}^{\infty} h(L^d | L^s) dL^d \text{ do not exist.}$$

It is evident that in case 1, $\text{Min}(L^s, L^D) = L^D = L$ and in case 2, $\text{Min}(L^s, L^D) = L^S = L$. In case 3, $\text{Min}(L^s, L^D)$ is indeterminate.

ANNEX 3D Statistical series, definitions and sources

In Section 3.3, I used data for the Netherlands as a whole and for the two construction sectors, civil and utility, and roads and waterworks. The data for the construction industry were not available directly, but had to be constructed from data on occupations. This limits the sample period from 1960 to 1983. From the occupational data on job vacancies, only registered job vacancies are available. I am aware of the fact that these data may have a high measurement error (see also Chapter 4). However, for testing the distribution assumptions in this chapter, I do not consider this a problem. To maintain the greatest possible consistency in definition and measurement with the construction industry data, I used the national data from Gelauff, Wenekers and de Jong (1985). As a consequence, I could not include more recent observations (see also main text, Chapter 3).⁵⁸

Sources and construction

Dutch economy

Source of data on L , L^S and L^D for the period 1960-1982 see Gelauff *et al*, 1985. For the year 1983, I used the Central Economic Plans of 1986 and 1987. L^S is defined as the sum of employment and unemployment, U . L^D is defined as the sum of employment and registered job vacancies, V .

Construction sectors

The employment data for both construction sectors are available from the Economic Institute of Medium and Small Firms. Vacancy and unemployment data are available only by occupational category. To be able to determine the total number of vacancies and unemployed workers per sector, an occupations-sectors matrix was

⁵⁸ In Section 3.4, where I tested the feasibility of the employment function based on the normal and the log-normal, I used data from Muysken and De Regt (1994). These data are not subject to the systematic measurement bias in the register data. The consistency issue with respect to the construction industry data is not relevant in this section. For a description of the data see Annex 4A.

needed. Since the construction occupations are closely linked to one of the two sectors, it was indeed possible to develop such a matrix. Table 3D.1 was constructed using expert opinion.

In constructing the unemployment and vacancy data according to sector, some points need to be kept in mind. The number of vacant jobs for women (especially administrative functions) has been assigned to the respective sectors on the basis of sector's share of vacant jobs for men for the period 1960-1981. For the period 1982-1983, the number of vacant jobs for women is not reported separately but included in the "other occupational categories". The total number of unemployed workers per sector was determined in a similar way. However, the correction for unemployed female workers had to be carried through for the whole estimation period.

Employment data are defined in terms of total hours actually worked divided by the total number of standard hours in one manyear.

Source: Statistics Netherlands (CBS) and the Economic Institute of Medium and Small Firms (data from Romme and Wolfs, 1986).

Unemployment data are defined in terms of total numbers of persons corrected for actual working hours. Formally, $U^* = z U^p$, where U^* is the definition of unemployment in "volume terms", I used in the estimations, U^p which is unemployment in persons and z is the ratio of actual working hours with respect to standard working hours.

Source: Sociale Maandstatieken, Tables 1.5.7 and 1.5.8 (since 1978) and CBS archives; reported by the District Labour Offices on 31 May.

Vacancy data are defined in terms of total numbers of persons corrected for actual working hours. The correction factor is z .

Source: Sociale Maandstatieken, Tables 1.5.7 and 1.5.8 (since 1978) and C.B.S. archives; reported by the District Labour Offices on 31 May.

Labour supply, is the sum of employment and (corrected) unemployment; *Labour demand* is the sum of employment and (corrected) vacancies.

Table 3D.2 contains the data for the construction sectors.

Table 3D.1 Occupational structure of the civil and utility sector and the roads and waterworks sector

| Occupational Category | Civil and Utility sector | Roads and Waterworks sector |
|---------------------------------|--------------------------|-----------------------------|
| Constructional engineers | ■ | |
| Carpenters (c&u) | ■ | |
| Concrete carpenters | ■ | |
| Bricklayers | ■ | |
| Barbenders (concrete work) | ■ | |
| Concrete workers | ■ | |
| Stucco workers | ■ | |
| Reed thatchers and whitewashers | ■ | |
| Hod-carriers | ■ | |
| Masons | ■ | |
| Housepainters | ■ | |
| Pointers | ■ | |
| Other professions (c&u) | ■ | |
| Workers for groundwork | | ■ |
| Other carpenters (r&w) | | ■ |
| Hydraulic engineers | | ■ |
| Workers for dredging work | | ■ |
| Dike-workmen | | ■ |
| Roadmakers | | ■ |
| Diggers | | ■ |
| Rammers | | ■ |
| Crane drivers | | ■ |
| Other professions (r&w) | | ■ |
| Pipefitters | ■ | |
| Metal workers | ■ | |
| Electrical workers | ■ | |

Table 3D.2 Data series on employment, labour supply and labour demand for the construction sectors (civil and utility, and roads and waterworks) ¹⁾

| | L_{cu} | L_{cu}^s | L_{cu}^D | L_{rw} | L_{rw}^s | L_{rw}^D |
|------|----------|------------|------------|----------|------------|------------|
| 1960 | 333661 | 335236 | 343539 | 75209 | 77791.6 | 76166.5 |
| 1961 | 325359 | 326255 | 339178 | 76358 | 77387.2 | 78417.1 |
| 1962 | 328453 | 329034 | 345189 | 83654 | 84388.3 | 85853.3 |
| 1963 | 321006 | 321622 | 337906 | 87045 | 87657.8 | 89167.9 |
| 1964 | 356052 | 356843 | 372898 | 90739 | 91432.6 | 92526.3 |
| 1965 | 362815 | 364257 | 378446 | 87867 | 89175.8 | 89269.0 |
| 1966 | 361234 | 363122 | 376978 | 90115 | 91105.3 | 91605.2 |
| 1967 | 378276 | 392141 | 386712 | 89403 | 93773.0 | 90197.1 |
| 1968 | 370731 | 378117 | 379945 | 84597 | 88377.3 | 85510.5 |
| 1969 | 363820 | 367032 | 377233 | 82265 | 84484.6 | 84032.7 |
| 1970 | 362801 | 365063 | 379761 | 82451 | 83782.7 | 84362.3 |
| 1971 | 355179 | 359294 | 369881 | 82186 | 83875.5 | 83602.2 |
| 1972 | 335592 | 348704 | 342452 | 74321 | 79851.4 | 75056.0 |
| 1973 | 324637 | 336154 | 331754 | 69798 | 74108.1 | 70535.7 |
| 1974 | 302469 | 320773 | 309120 | 62793 | 67360.1 | 63528.6 |
| 1975 | 281642 | 311492 | 286190 | 64709 | 70782.9 | 65424.8 |
| 1976 | 280646 | 303260 | 287545 | 64162 | 70900.9 | 64939.9 |
| 1977 | 283327 | 295361 | 293787 | 61064 | 65568.8 | 61959.6 |
| 1978 | 277741 | 286541 | 291303 | 58819 | 61563.2 | 60362.0 |
| 1979 | 265189 | 273160 | 278229 | 60567 | 62677.8 | 62791.7 |
| 1980 | 277175 | 291488 | 285706 | 57323 | 61003.2 | 58544.8 |
| 1981 | 254781 | 294691 | 257226 | 51699 | 58465.7 | 52011.7 |
| 1982 | 229057 | 293795 | 230062 | 44704 | 55686.2 | 44844.9 |
| 1983 | 207151 | 297999 | 207925 | 44402 | 62200.6 | 44477.3 |

1) Explanation of symbols: L = employment (man-years); L^s =labour supply; L^D =labour demand; cu:=c&u sector; rw=r&w sector.

4 The C.E.S. approach to measuring mismatches on the labour market and structural unemployment

4.1 Introduction

In the previous chapter some theoretical and empirical investigations substantiated the use of the log-normal density to describe the distribution of labour supply and demand on micro markets, thereby establishing the effectiveness of a specific class employment functions within the analytically-based model. Further elaborations within the analytically-based model are now possible. This chapter extends the one level C.E.S. approach of Section 2.2 to measure structural mismatches on the labour market and the corresponding structural unemployment. It has the specific goal of further exploiting the strength of the C.E.S. method, which is its ease of interpretability, *viz.* the interpretation of the C.E.S. parameters (and of structural unemployment) in terms of micro mismatches.

The chapter contains a theoretical part and an empirical part. The empirical method is in the spirit of the variant type 1 and treats supply and demand exogenous and correctly measured by the improved vacancy and unemployment data of Muysken *et al* (1994). As far as the theoretical part is concerned, I suggest to relax some of the restrictive assumptions underlying the (one-level) C.E.S. approach of Chapter 2. This theorising yields some important insights for the empirical investigation of the causes of mismatches and structural unemployment. To value the contributions of this chapter within the proper perspective, I will briefly summarise the main developments in the literature, where they have stopped and where the present chapter continues. For the sake of clarity, I accepted some overlap with Chapter 1 and Chapter 2 in describing the “state of the art”.

One important drawback of the two-variate C.E.S. structure of Chapter 2 is that it at least explicitly ignores a key element of Keynesian analysis: spillover effects from the goods market on to the labour market and *vice versa*.⁵⁹ In the

⁵⁹ Implicitly, this spillover effect is accounted for in the “generalisation” that I propose here. I return to this later.

literature (within the analytically-based method) this limitation has been recognised and dealt with differently.

In Lambert (1988) goods market influences are incorporated in the analysis albeit only at the aggregate level. The real problem, namely to integrate the spillover effect in the micro analysis, is circumvented as such as spillover terms are simply postulated at the aggregate (see also Section 2.4). Instead, Kooiman (1984) derives aggregate expressions allowing for spillover effects at the micro level. Mainly as a result of his comprehensive description of simultaneity of the goods and labour markets, these expressions have a complex form with multiple integrals and therefore lack intuitive appeal. At least in this chapter, I do not wish to compromise on the ease of interpretability - which is exactly the strength of the method. This result is however, not satisfactory.

Another line of research is suggested by Sneessens (1983), who integrated the spillover effect in the C.E.S. methodology by distinguishing “Keynesian” demand for labour, capacity demand for labour, and labour supply at the micro level. The resulting (three-variate) C.E.S. aggregate employment function has three arguments, that is aggregate Keynesian demand, capacity demand, and labour supply. A drawback of this structure, however, is that it is unable to distinguish between “capacity mismatch” (*i.e.* the inadequacy between the available installed capacities and the composition of the demand for goods) and “labour mismatch” (*i.e.* the inadequacy between labour demanded and supplied, in terms of qualification, location *etc.*).

Researchers aware of this shortcoming reacted by developing a two-level C.E.S. employment function: at the first level aggregate Keynesian demand and capacity demand, and, at the second level aggregate Keynesian demand and capacity demand on the one hand and aggregate labour supply on the other. The idea of the two-level C.E.S. within the markets of disequilibrium setting has been put forward independently by Bierings and Muysken (1988) and Gagey, Lambert and Ottenwaeler (1988).

The work of Gagey *et al* (1988) provides the starting point for further elaborations in this chapter. The strength of their approach is that goods market influences are brought into the analysis, while at the same time the simple C.E.S. structure is kept upright. Its weakness, however, lies in the undefined relationship between structural mismatches and the C.E.S. parameters, which must be considered a very undesirable property, especially if one wishes to endogenise the mismatches in the empirical model specification. Contrary to the one level specifications, the derivation of such a relationship is not straightforward [see eq. (2.2.19) for the two-variate case, and eq. (4.2.12) for the three-variate case discussed hereafter]. My contribution is to fill the “interpretation gap” and to derive a more explicit relationship between the C.E.S. parameters (and thus structural un-employment) and micro mismatches. The evidence points towards factors that could explain capacity and labour mismatches. Important insights are obtained for the empirical implementation aiming at discovering the causes of structural unemployment.

In an empirical sense, the innovation of this chapter lies in the econometric model specification which improves over prior studies in two important respects. First, the model takes due account of the theoretical insights above. The second distinctive feature represents the integration of the approach developed by Muysken and De Regt (1994) which allows testing the importance of search relative to structural unemployment. This issue touches on the debate mentioned in Chapter 1 about the relative importance of the rise in mismatch on the labour market since the 1970s (see also Layard *et al*, 1991). The empirical results presented in this volume for the Netherlands establish the quantitative importance of search duration and mismatch in a strict sense (that is the qualification dimension).

The remainder of this chapter proceeds as follows. To provide an equally exhaustive **formal** description of the theoretical developments within the C.E.S. approach, I consider the aggregate specifications for the one-level three-variate case developed Sneessens (1983) in Section 4.2, and the two-level three-variate case based on Gagey *et al* (1988) in the first part of Section 4.3. The second part of Section 4.3 elaborates on the relationship between the C.E.S. parameters and micro mismatches. In Section 4.4. the estimation results for structural unemployment are presented. Section 4.5 contains the conclusions.

4.2 The one-level three-variate C.E.S.

Chapter 2 dealt with the one-level two-variate C.E.S. employment function. For purposes of comparison with the one-level and two-level, three-variate cases, two points with respect to the two-variate C.E.S. function are worth recapitulating:

- The covariance matrix of micro supplies and demands, Σ^* , has not been simplified to make the derivation of the C.E.S. function possible [compare eq. (2.2.6)].
- The C.E.S. parameter $\frac{1}{\beta_1}$ is almost perfectly proportional to the underlying structural mismatch σ^* , for any reasonable value of σ^* ; thus one can simply endogenise mismatches by replacing $\frac{1}{\beta_1}$ by a function of the variables which seem relevant in explaining the mismatches.

In this chapter, the derivation and description of the one-level three-variate C.E.S. rests on a hand-written manuscript of Sneessens (1983). To introduce spillover effects in the analysis, Sneessens assumes that, on each micro labour market, employment is determined by the minimum of Keynesian demand or demand determined employment (*i.e.* the demand for goods translated in terms of

employment), L_i^k , capacity demand for labour (*i.e.* the amount of labour required to run the available production capacity), L_i^c , and labour supply, L_i^s . By analogy of (2.2.7), the micro labour market model with spillover terms then shows (where the supply and demand quantities are already rescaled to the aggregate and i is the index for a micro market):⁶⁰

$$\left. \begin{aligned} \ln L_i^k &= \Lambda^k + \varepsilon_i^k & \ln L_i^c &= \Lambda^c + \varepsilon_i^c & \ln L_i^s &= \Lambda^s + \varepsilon_i^s \\ \ln L_i &= \text{Min}(\ln L_i^k, \ln L_i^c, \ln L_i^s) \end{aligned} \right\} \quad (4.2.1)$$

with

$$E(L^k) = e^{\Lambda^k + \frac{1}{2}\sigma_{\varepsilon^k}^2} = L^K \quad E(L^c) = e^{\Lambda^c + \frac{1}{2}\sigma_{\varepsilon^c}^2} = L^C \quad E(L^s) = e^{\Lambda^s + \frac{1}{2}\sigma_{\varepsilon^s}^2} = L^S$$

and covariance matrix:

$$\Sigma^* = \begin{bmatrix} \sigma_k^2 & \sigma_{kc} & \sigma_{ks} \\ \sigma_{kc} & \sigma_c^2 & \sigma_{cs} \\ \sigma_{ks} & \sigma_{cs} & \sigma_s^2 \end{bmatrix}$$

where I dropped ε and η in the σ 's (see also footnote 21). The specifications for the "average" supply, Keynesian demand and capacity demand are similar to those in (2.2.6).

The intuitive idea behind model (4.2.1) is that for a job to be created, three conditions have to be met: there must be demand for the output, equipment to produce the output and a worker to run the equipment; the actual number of jobs created in a firm is then determined by the minimum of these three quantities. Evaluating the expectation of L with respect to the micro disturbances [*c.f.* eq. (4.2.1)], the following three-variate analogue to (2.2.8) for aggregate employment results:

⁶⁰ In a strict sense the way of reasoning in this section should begin with the micro model in the first step and then change to the macro analogue model in the second step (this is the procedure of Section 2.2). At the cost of rigour in presentation, I choose, however, to skip the first step. This step is not essential, for leaving it out does not influence the final results. I do the same in Section 4.3. There too, the rescaled quantities can be safely interpreted as **micro level** quantities.

$$L = \Phi^K L^K + \Phi^C L^C + \Phi^S L^S \tag{4.2.2}$$

with $\Phi^K = \Phi\left(x - \frac{1}{2} \sigma_{c-k}, -y - \frac{1}{2} \sigma_{k-s} \mid \rho_{k-c, k-s}\right)$

$$\Phi^C = \Phi\left(-x - \frac{1}{2} \sigma_{c-k}, -z - \frac{1}{2} \sigma_{c-s} \mid \rho_{c-k, c-s}\right)$$

$$\Phi^S = \Phi\left(y - \frac{1}{2} \sigma_{k-s}, z - \frac{1}{2} \sigma_{c-s} \mid \rho_{s-k, s-c}\right)$$

Φ denotes the bivariate normal integral and:

$$x = \frac{\ln \frac{L^C}{L^K}}{\sigma_{c-k}} \quad y = \frac{\ln \frac{L^K}{L^S}}{\sigma_{k-s}} \quad z = \frac{\ln \frac{L^C}{L^S}}{\sigma_{c-s}}$$

The variances in the above specifications are again the mismatch-variances between any pair of variables in the above specifications, that is for instance [compare $\sigma_{a,s}$ in (2.2.8)]:

$$\sigma_{c-k}^2 = \text{var}(\varepsilon_c - \varepsilon_k) = \sigma_c^2 + \sigma_k^2 - 2\sigma_{ck}$$

Besides $\sigma_{c,k}$, there are the two other mismatches $\sigma_{k,s}$ and $\sigma_{c,s}$. The parameters appearing in the bivariate integrals, Φ^j (with $j=K, C$ or S) denote the correlation between any two pairs of mismatches. It can be shown that they can all be written in terms of the mismatch variances, *i.e.*:

$$\rho_{k-c, k-s} = \frac{\text{cov}((\varepsilon^k - \varepsilon^c)(\varepsilon^k - \eta^s))}{\sigma_{c-k} \sigma_{k-s}} = \frac{1}{2} \frac{(\sigma_{k-s}^2 + \sigma_{c-k}^2 - \sigma_{c-s}^2)}{\sigma_{c-k} \sigma_{k-s}}$$

Analogous to the two-variate case, transactions, L , is a function of L^K , L^C and L^S and the three mismatches. To make the transformation of (4.2.2) into a C.E.S. structure possible, the same approximation formulae (2.2.15) based on the weighted regime proportions, P^j is used, but now for $j=K, C$ and S .⁶¹ The properties of P^j are the same as for the two-variate case described in Section 2.2. It can be shown that the weighted proportions P^j are defined as [compare (2.2.13)-(2.2.14)]:

⁶¹ The weighting scheme is the same as in Section 2.2.

$$P^K = P(L_i^k < \text{Min}(L_i^c, L_i^s)) = L^K \frac{\Phi^K}{L} \quad (4.2.3)$$

$$P^C = P(L_i^c < \text{Min}(L_i^k, L_i^s)) = L^C \frac{\Phi^C}{L} \quad (4.2.4)$$

$$P^S = P(L_i^s < \text{Min}(L_i^c, L_i^k)) = L^S \frac{\Phi^S}{L} \quad (4.2.5)$$

$$\text{with } P^K + P^C + P^S = 1 \quad (4.2.6)$$

Together the employment function and the weighted proportions P^j (with $j=K, C$ or S) define three functions in terms of $L^K, L^C, L^S, \sigma_{k,s}, \sigma_{c,k}$ and $\sigma_{c,s}$. Hence, in terms of the covariance matrix, Σ^* , only three of the total of six independent arguments, are identified.

Requiring exactness in the point $L^K = L^C = L^S$ as in the two-variate case, implies the following restriction on β_2 :

$$\beta_2^j = \left(\frac{\ln \Phi^j}{\ln(\Phi^K + \Phi^C + \Phi^S)} \right) \Big|_{L^K = L^C = L^S} - 1 \quad j = K, C, S; \beta_2^j > 0 \quad (4.2.7)$$

where the subscript 2 in β indicates the second C.E.S. type as meant in footnote 23.

It is easily verified that for non-identical β_2^j 's it is not possible to solve the system (4.2.1)-(4.2.7) for an explicit employment function.⁶² To be able to still derive a closed form approximation, Sneessens makes the highly restrictive assumption of identical mismatches:

$$\sigma_{c-s}^2 = \sigma_{k-s}^2 = \sigma_{c-k}^2 = \sigma^{*2} \quad (4.2.8)$$

⁶² The following implicit form may, however, be derived:

$$1 = \left(\frac{L^K}{L} \right)^{-\beta_2^K} + \left(\frac{L^C}{L} \right)^{-\beta_2^C} + \left(\frac{L^S}{L} \right)^{-\beta_2^S}$$

The interpretation of the β_2 parameters in this specification is unclear, since it cannot be explicitly related to the mismatches (Kerckhoffs, 1992). This is very undesirable in endogenising the mismatches.

of which

$$\Sigma^* = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix} \tag{4.2.9}$$

is a specific example. Then (4.2.2) is easily solved for a specific C.E.S. employment function. Given the mismatch structure (4.2.8) it is clear that $P^j=1/3$ and all Φ^j 's are the same when the exact system (4.2.2) is evaluated in $L^K=L^C=L^S$. Exactness in this point immediately implies $\beta_2^j = \beta_2$. As in the two-variate case, then for L , the following explicit form results:

$$L \approx \left(L^{K^{-\beta_2}} + L^{C^{-\beta_2}} + L^{S^{-\beta_2}} \right)^{-\frac{1}{\beta_2}} \tag{4.2.10}$$

Equally as (2.2.16) for the two-variate case, there exists a relationship between β_2 and the mismatch parameter σ^* which follows directly from (4.2.7):

$$\beta_2 = \frac{-\ln 3}{\ln 3 + \ln \Phi \left(-\frac{1}{2} \sigma^*, -\frac{1}{2} \sigma^* \middle| \frac{1}{2} \right)} \quad \text{in } L^K=L^C=L^S \tag{4.2.11}$$

By using a Taylor expansion around $\sigma^*=0$, this expression can be simplified as:

$$\frac{1}{\beta_2} \approx \frac{3\sigma^*}{2\sqrt{2}\pi \ln 3} \approx 0.5447 \sigma^* \tag{4.2.12}$$

which is almost the same expression as (2.2.19) for the two-variate case, though, of course, with a different interpretation. Equation (4.2.12) implies that β_2 is inversely related to σ^* as in the two-variate case. For any reasonable value of σ^* , β_2^{-1} and σ^* are almost completely proportional: a 10 percent point rise in σ^* implies a 10 percent point rise in β_2^{-1} (for the relevant ranges of σ^* between, say 0.010 and 0.150, compare Table 2.1). Note that the multiplication factor is 0.5447, which is marginally larger than in the two-variate case. In conclusion, as shown by (4.2.12), as in the two-variate case, mismatches in the (one-level) three-variate C.E.S. case may finally be adequately interpreted in terms of the C.E.S. parameters, thereby allowing a proper endogenisation of σ^* in the empirical analysis. This is, however, done at the expense of a very restrictive covariance matrix.

The implications for the measurement of the structural unemployment rate are straightforward and follow from (4.2.10):

$$u_2^* = 1 - 3 \frac{1}{\beta_2} \quad (4.2.13)$$

So, through β_2 [c.f. (4.2.12)], σ^* determines the structural unemployment rate. For any reasonable value of σ^* , a 10 per cent rise in σ^* results in an almost 10 per cent rise in u_2^* .

Clearly the problem with the approximation (4.2.16) remains that it is based on the assumption of identical mismatches. The result is that σ^* is a composite measure of the dispersion of Keynesian demand for labour, capacity demand for labour and labour supply and as such still difficult to interpret. In the next sections a solution to this problem is suggested by way of a more convenient approximation which is consistent with a two-level C.E.S. structure.

4.3 Unrestricted mismatches and the two-level C.E.S. employment function

4.3.1 Introduction

In Bierings and Muysken (1988) and Gagey *et al* (1988) it was shown that the assumption of identical mismatches is unnecessary for obtaining an explicit employment function. Both studies use more flexible approximation formulae to obtain a C.E.S. employment function with two levels. The first level consists of aggregate Keynesian labour demand and capacity labour demand. The second level consists of aggregate Keynesian labour demand and capacity labour demand on the one hand, and aggregate labour supply on the other.

A major shortcoming of these studies is that the question of how the parameters of the C.E.S. and the mismatches relate is not adequately solved. There is some intuitive reasoning, saying that the C.E.S. parameter at the first level measures the mismatch between Keynesian demand and capacity demand, and the C.E.S. parameter at the second level measures the mismatch between supply and demand. However, structural unemployment implied by the C.E.S. is not rigorously derived from the mismatches and therefore cannot be interpreted correctly. This is a very undesirable property if one wants to identify the variables influencing it.

The aim of this section is to make this relationship explicit. I take the approach of Gagey *et al* (1988) as a benchmark. By calculating partial derivatives I am able to derive explicit expressions between the parameters in the nested C.E.S.

employment function and the mismatch parameters. The exact relationship between the C.E.S. parameters and the mismatches provides the theoretical basis for endogenising the mismatches in the empirical analysis.

4.3.2 Relaxation of mismatch structure

I start by repeating two points relevant to the analysis below. First, the formal presentation is in terms of micro market variables that are rescaled to the aggregate. These rescaled variables may safely be interpreted as micro level variables (see footnote 60). Second, the model presented rests on Gagey *et al* (1988).

The approach of Gagey *et al* begins by recognising that at the micro market, or better the firm level, employment (in the short run) is the result of a two stage process. Imagine that individual firms determine their labour demand, L^d , first by comparing the Keynesian demand for labour L^k to the capacity demand L^c and then confronting this demand with the actual availability of labour, L^s . Given this decision process the *min-condition* may be hierarchically decomposed by introducing explicitly the intermediate demand variable L^d and its macro analogue L^D :

$$\left. \begin{aligned}
 \ln L_i^d &= \text{Min}(\ln L_i^k, \ln L_i^c) & L^D &= E(\text{Min}(\ln L^k, \ln L^c)) \\
 \ln L_i &= \text{Min}(\ln L_i^d, \ln L_i^s) & L &= E(\text{Min}(\ln L^d, \ln L^s))
 \end{aligned} \right\} \quad (4.3.1)$$

where the expectation is again with respect to the micro disturbances as in (4.2.1). With reference to (4.3.1) the composite labour demand part, one can define the following weighted proportions:

$$P^{K|D} = P(L_i^k < L_i^c) \quad (4.3.2)$$

$$P^{C|D} = P(L_i^c < L_i^k) \quad (4.3.3)$$

$$P^{K|D} + P^{C|D} = 1 \quad (4.3.4)$$

where “|” is the conditional sign. The weighting applies to composite labour demand. Likewise, with reference to (4.3.1) the transactions part, the weighted proportions P^s and P^D can be defined:

$$P^S = P(L_i^s < L_i^d) = P(L_i^s < \text{Min}(L_i^k, L_i^c)) \quad (4.3.5)$$

$$P^D = P(L_i^d < L_i^s) = P(L_i^s > \text{Min}(L_i^k, L_i^c)) \quad (4.3.6)$$

$$P^S + P^D = 1 \quad (4.3.7)$$

where the weighting is with respect to employment. Note that this P^S is identical to the one defined previously in (4.2.5), whereas P^D corresponds to $P^k + P^c$ as defined by (4.2.3) and (4.2.4).

Equation (4.3.1) determines demand L_i^d as the minimum of the two log-normally distributed variables L_i^k and L_i^c . Applying the Sneessens approximation as in (2.2.15), L^D can be approximated by:

$$L^D \approx \left(L^{K-} \beta_3 + L^{C-} \beta_3 \right)^{-\frac{1}{\beta_3}} \quad (4.3.8)$$

$$\text{with } \beta_3 = \frac{-\ln 2}{\ln 2 + \ln \Phi\left(-\frac{1}{2} \sigma_{c-k}\right)} \quad \text{and} \quad (4.3.9)$$

$$\sigma_{c-k}^2 = \sigma_c^2 = \sigma_k^2 - 2 \sigma_{ck} \quad (4.3.10)$$

Once again this provides an exact approximation for L^D in $L^k = L^c$.

Equation (4.3.1) also determines micro level employment, L_i as the minimum of demand, L_i^d and the log-normally distributed variable L_i^s . Under the assumption that L_i^d , being the minimum of two log-normally distributed variables, is itself log-normally distributed, the procedure above can be repeated. This approximation works very well for the situation where $\ln L_i^k$ and $\ln L_i^c$ are highly correlated. Only if $\ln L^k$ and $\ln L^c$ have comparable means but substantial unequal variances, or if their mutual correlation is negative, normality of $\ln L_i^d = \min(\ln L_i^k, \ln L_i^c)$ breaks down (see Kerckhoffs, 1992, who in turn refers to Daganzo, 1979, p. 51-58). Given the interpretation of L^k and L^c above, this seems very unlikely. Hence, I regard L_i as the minimum of two log-normally distributed variables. Applying again the Sneessens approximation yields the following expression for L :

$$L \approx \left(L^{D-} \beta_4 + L^{S-} \beta_4 \right)^{-\frac{1}{\beta_4}} \quad (4.3.11)$$

$$\text{with } \beta_4 = \frac{-\ln 2}{\ln 2 + \ln \Phi\left(-\frac{1}{2} \sigma_{d-s}\right)} \quad \text{and} \quad (4.3.12)$$

$$\sigma_{d-s}^2 = \sigma_d^2 = \sigma_s^2 - 2 \sigma_{ds} \tag{4.3.13}$$

where $\sigma_{d,s}$ represents the mismatch between $\ln L^d$ and $\ln L^s$. This provides a very good approximation for L^D , in the point $L^S = L^D$ - not an exact approximation because the minimum of two log-normal variables is only approximately log-normally distributed. By combining (4.3.8) and (4.3.11), the two-level C.E.S. evolves:

$$L \approx \left(\left(L C^{-\beta_3} + L K^{-\beta_3} \right)^{\frac{\beta_4}{\beta_3}} + L S^{-\beta_4} \right)^{-\frac{1}{\beta_4}} \tag{4.3.14}$$

Equation (4.3.14) allows for a convenient interpretation of the parameters involved. The parameter β_3 may be informally interpreted as a measure for the mismatch between the two micro demand concepts and β_4 as a measure for the mismatch between micro supply and micro composite demand. Indeed, according to (4.3.9), β_3 directly reflects the mismatch σ_{c-k}^2 . Analogous to (4.2.12) it can be approximated by:

$$\frac{1}{\beta_3} \approx 0.57556 \sigma_{c-k} \tag{4.3.15}$$

The parameter β_4 reflects the mismatch $\sigma_{d,s}$. By approximation applies:

$$\frac{1}{\beta_4} \approx 0.57556 \sigma_{d-s} \tag{4.3.16}$$

The interpretation of β_4 in terms of the mismatch $\sigma_{d,s}$ is not very revealing, since $\sigma_{d,s}$ is itself related to the three mismatches σ_{c-k}^2 , σ_{k-s}^2 and σ_{c-s}^2 . To get a better understanding, $\sigma_{d,s}$ should be worked out further in terms of these mismatches.

To work out σ_{d-s}^2 according to its definition (4.3.12), the second order moments of $\ln L^d = \text{Min}(\ln L^k, \ln L^s)$ (with respect to the micro disturbances) should be evaluated. Manipulating the expressions in Quandt (1988, p. 57 and p. 222), the following applies (the “ ’ ” sign represents the first derivative⁶³):

$$\sigma_d^2 = \sigma_c^2 \Phi(v) + \sigma_k^2 (1 - \Phi(v)) + \sigma_{c-k}^2 \left(v - (\Phi'(v) + v\Phi(v)) \right) (\Phi'(v) + v\Phi(v)) \tag{4.3.17}$$

$$\sigma_{ds} = \sigma_{cs} \Phi(v) + \sigma_{ks} (1 - \Phi(v)) \tag{4.3.18}$$

⁶³ The “ ’ ” as superscript of Φ here, should be distinguished from the “ ‘ ” in, for example, eq. (2.3.2) or in equations of Chapter 2.

$$\text{with } v = \frac{\Lambda^k - \Lambda^c}{\sigma_{c-k}} \text{ [compare (4.2.1) for the understanding of } \Lambda^k \text{ and } \Lambda^c \text{]} \quad (4.3.19)$$

From (4.3.18) it is clear that the covariance, σ_d , between $\ln L^d$ and $\ln L^s$ is a weighted average between $\ln L^s$ and $\ln L^c$ on the one hand, and between $\ln L^s$ and $\ln L^k$ on the other hand. Their weighting, involving $\Phi(v)$, is obvious. If these two covariances are equal, say $\sigma_{cs} = \sigma_{ks} = \sigma$, it follows that $\sigma_d = \sigma$. If $\sigma_{c-k} = 0$, *i.e.* there is no mismatch between $\ln L^k$ and $\ln L^c$, then $\ln L^d$ is either systematically equal to $\ln L^k$, or to $\ln L^c$. Hence, σ_d^2 is equal to either σ_k^2 or σ_c^2 .

Substituting (4.3.17) - (4.3.18) into (4.3.13), the following expression for σ_{d-s}^2 can be obtained:

$$\sigma_{d-s}^2 = \sigma_{c-s}^2 \Phi(v) + \sigma_{k-s}^2 (1 - \Phi(v)) + \sigma_{c-k}^2 (v - (\Phi'(v) + v\Phi(v))) (\Phi'(v) + v\Phi(v)) \quad (4.3.20)$$

Hence, the mismatch between supply and demand is a function of all three structural mismatches, which can be decomposed in two parts:

- The first part, consisting of the first two terms of (4.3.20), is a weighted average of the structural mismatches σ_{c-s}^2 and σ_{k-s}^2 , with obvious weighting factors. Consequently, if the mismatch between $\ln L^s$ and either component of $\ln L^d$ increases, then the mismatch between $\ln L^s$ and $\ln L^d$ itself increases. This is intuitively obvious.
- The second part concerns the third term of (4.3.20). This term relates the demand-supply mismatch to the inner- L^d mismatch, *i.e.* the mismatch between the two components of demand, $\ln L^k$ and $\ln L^c$. Using the well-known expressions for the conditional expectation of a standard normal variate, it is easily shown that the coefficient on σ_{c-k}^2 is negative for any value of v . Hence, if the mismatch between $\ln L^c$ and $\ln L^k$ **increases**, the mismatch between $\ln L^d$ and $\ln L^s$ **decreases!** At first sight, this result may be counterintuitive. However, it is essentially due to the fact that the min-condition (4.3.1) (for the composite labour demand) tends to smooth the variances of $\ln L^k$ and $\ln L^c$. The conclusion is if σ_{c-k}^2 increases, then β_3 decreases, whereas β_4 increases.

This establishes the relationships between the β parameters in the C.E.S. function and the mismatches which was the prime objective of this chapter.

Yet, one can still be more specific by evaluating σ_{d-s}^2 in $L^k=L^c$, or equivalently $E(L^k) = E(L^c)$. There is no specific reason why $E(\ln L^k)$ and $E(\ln L^c)$ should differ in this situation, so $\Lambda^k = \Lambda^c$. (silently assuming $\sigma_c^2 = \sigma_k^2$).⁶⁴ This implies $v=0$,

⁶⁴ See equation (4.2.1) for Keynesian and capacity demand and the formulated equations for $E(L^k)$ and $E(L^c)$ after that.

$\Phi=1/2$ and $\Phi' = \frac{1}{\sqrt{2\pi}}$, and hence:

$$\sigma_{d-s}^2 = \frac{1}{2} \left(\sigma_{c-s}^2 + \sigma_{k-s}^2 - \frac{\sigma_{c-k}^2}{\pi} \right) \tag{4.3.21}$$

This succinctly shows how the mismatch between supply and demand is related to the underlying three structural mismatches approximately. The expression for σ_{d-s}^2 and hence β_4 , involving a dominant role for σ_{c-s}^2 and σ_{k-s}^2 and a lesser importance of σ_{c-k}^2 (being divided by π) is consistent with the common intuitive interpretation of β_4 , mentioned above. Note that for $\sigma_{c-s} = \sigma_{k-s}$, (4.3.21) is reduced to:

$$\sigma_{d-s}^2 = \sigma_{ck-s}^2 - \frac{\sigma_{c-k}^2}{2\pi} \tag{4.3.22}$$

One should read for the subscript “ck” in the above variance *k or c*.

The mismatch parameter σ_{d-s} can be worked out still further when it is written as the following linear combination of σ_{ck-s} and σ_{c-k} . Define to this purpose:

$$\sigma_{c-k} = \mu \sigma_{ck-s} \tag{4.3.23}$$

By substitution in (4.3.22) this yields:

$$\sigma_{d-s}^2 = \sigma_{ck-s}^2 \left(1 - \frac{\mu^2}{2\pi} \right) \tag{4.3.24}$$

or:

$$\sigma_{d-s} = \sigma_{ck-s} \left(1 - \frac{\mu^2}{2\pi} \right)^{\frac{1}{2}} \tag{4.3.25}$$

Expanding (4.3.25) around $\mu=1$ leads to:

$$\sigma_{d-s} \approx \left(\frac{1}{\left(1 - \frac{1}{2\pi} \right)^{\frac{1}{2}}} \sigma_{ck-s} - \frac{1}{2\pi \left(1 - \frac{1}{2\pi} \right)^{\frac{1}{2}}} \sigma_{c-k} \right) \tag{4.3.26}$$

Substitution of (4.3.26) in (4.3.16), finally results in:

$$\beta_4 \approx \frac{1}{0.628 \sigma_{ck-s} - 0.100 \sigma_{c-k}} \tag{4.3.27}$$

Thus, through (4.3.15) and (4.3.27), the relationship between $\sigma_{ck,s}$ and $\sigma_{c,k}$ and u_3^* is defined. From (4.3.14) it follows:

$$u_3^* = 1 - \left(\frac{\beta_4}{2\beta_3 + 1} \right)^{-\frac{1}{\beta_4}} \quad (4.3.28)$$

Once again one can show that a 10 percent rise in $\sigma_{ck,s}$ and $\sigma_{c,k}$ gives rise to an almost 10 percent rise in u_3^* , when $\sigma_{ck,s}$ and $\sigma_{c,k}$ lie within reasonable ranges. Note that for $\sigma_{ck,s} = \sigma_{c,k} = \sigma^*$ (4.3.27) can be written as:

$$\frac{1}{\beta_4} \approx 0.5283 \sigma^* \quad (4.3.29)$$

This relationship closely resembles the relationship between β_2 and σ^* in the restricted (one-level) three-variate case [see eq. (4.2.12)], and between β_1 and σ^* in the (one-level) two-variate case [see eq. (2.2.19)], with only slight differences in the coefficient values.

4.4 Empirical formulation, data and results

This section deals with the estimation of mismatches and structural unemployment by way of the C.E.S. employment function. In the previous sections three alternative C.E.S. specifications have been put forward [compare equations (2.2.17), (4.2.10) and (4.3.14)]. Evidently, (4.3.14) is to be preferred, because it has the widest application. Contrary to the other specifications it allows for two types of mismatches: the mismatch within the labour demand and capacity demand for labour, and the mismatch between labour demand and labour supply.

Data as to the origin of vacancies, *i.e.* stemming from insufficient demand or capacity constraints on production, are not available. This obstructs the application of a direct and single equation estimation method. Instead, it seems as if the only alternative to overcome this problem is to estimate a complete endogenous model. In this respect the two alternatives of Chapter 2 emerge:

- A complete model estimation with endogenous L^K , L^C and L^S (the variant type 3 of the analytically-based method). **or**
- A complete model estimation in line with the variant type 2 of the analytically-based method. In this variant the labour market equations have been transformed into log-linear specifications with the regime proportions as dependent variables. Besides, this variant uses regime classification information from business surveys.

To my knowledge, there are no applications of the variant type 3. The variant type 2 is adopted by Gagey *et al* (1988) (see also Section 2.5).

In this chapter, however, I am able to employ direct observations when the two-level (nested) C.E.S. function is simplified by capturing the C.E.S. structure of Keynesian demand and capacity demand in a single composite demand term [*c.f.* eq. (4.3.11) or (2.2.17)], and hence to present straightforward estimates of the mismatch (corroborating the statement of footnote 59). This, of course, can only be done when the theoretical lessons of the previous section are kept in mind. At this stage, it is worth repeating one of the prime objectives for undertaking this study was to compare the results for a simple (and readily) interpretable C.E.S. form with those of an intricate endogenous model with four regimes and aggregation over micro markets (the simulation-based method). The latter is the subject of Chapter 6. So, the empirical application stays within the variant type 1 of the analytically-based method.

To justify direct estimation of (2.2.17), two conditions should be met. First, reliable data on labour supply (employed plus unemployed) and demand (employed plus job vacancies) must be available. Second, the lessons of Section 4.3 should be drawn to the empirical specification. With respect to the latter, I suggest to endogenise $\sigma_{d,s}$ (or β_1) taking due notice of the fact that the mismatch between supply and demand, $\sigma_{d,s}$ is a composite measure of $\sigma_{c,s}$, $\sigma_{k,s}$ and $\sigma_{e,k}$ [see eq. (4.3.20)].

Data on vacancies and unemployment

Initially register data on vacancies failed because firms only registered part of their job vacancies and often failed to report back when vacancies were filled or cancelled. Only one in three job vacancies turned out to be registered at the job centres in the Netherlands, Germany and the UK (van Bastelaer and Laan, 1994).⁶⁵ The consequence is a quite a large underestimation of structural unemployment. A similar story applies to the initial unemployment data which were also based on the registers of the job centres. By now matters have changed. Since Statistics Netherlands uses surveys to measure job vacancies and unemployed workers, the reliability of the figures has much improved. The earlier figures could also be enhanced through extrapolation techniques taking account of survey information and data on advertisements (Muysken, Bierings and De Regt, 1994).

⁶⁵ The implication of using poor register data for the estimation of structural unemployment is probably most evident for the situation of Switzerland. Stalder (1989), estimating a disequilibrium model with supply and demand as unobserved endogenous variables, finds a substantial increase in structural unemployment. This clearly contrasts with the very low official unemployment levels (see also Chapter 2, Table 2.2).

I will use the corrected job vacancy series of Muysken, Bierings and De Regt (1994). Their correction procedure uses observed job vacancy data available from the vacancy surveys of Statistics Netherlands for the period 1981-1990 to obtain estimates for the period 1960-1980. I confine myself to a short description. To determine the number of job vacancies, the ratio of personnel advertisements and observed job vacancies was estimated as a function of the duration of vacancies and the capacity utilisation rate -both demand pressure indicators- for the period 1981-1990. Since data on advertisements are only available from 1968 on, this equation can only be used to obtain estimates of observed job vacancies for the period 1968-1980. Estimates for the period 1960-1967 follow from an equation for the vacancy registration rate (*i.e.* the ratio of registered to observed job vacancies) which is estimated as a function of the duration of vacancies, the capacity utilisation rate and registered vacancies for the period 1981-1990. Figure 4.1 displays the estimated series for "observed" vacancies and the series for the registered number of vacancies for the period 1960-1990. Note that the correction is mainly a correction in levels.

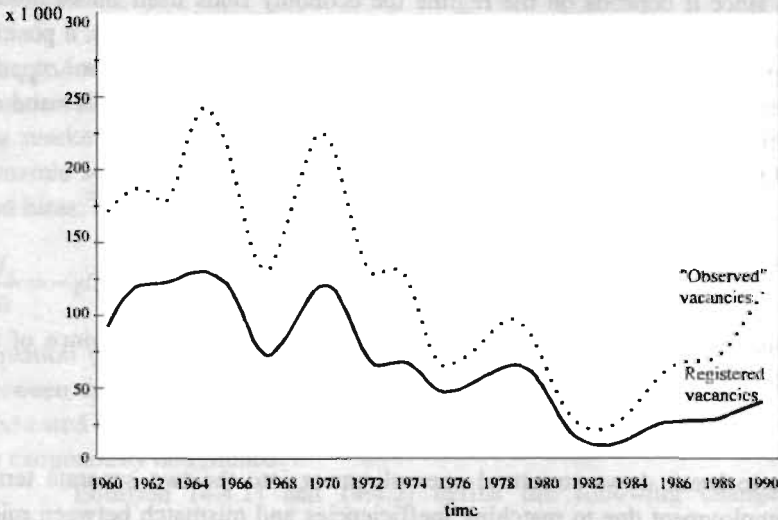
Time-consistent unemployment series are provided by the Central Planning Bureau. They used the survey information from the Labour Force Survey to correct the early register data. A final point of caution should be made, namely, despite the fact that the data are upgraded, the problem remains that vacancies are hard to measure (van Bastelaer and Laan, 1994, p. 14-15). As a consequence the "good quality of data" assumption remains the weakest part of the variant type 1 method (see Chapter 1).

Empirical specification

The empirical specification of the model recognises two important aspects. First, I assume that the coefficient β_1 is not constant and depends on variables influencing $\sigma_{c,s}$, $\sigma_{k,s}$ and $\sigma_{c,k}$. Second, it allows for a decomposition of structural unemployment in (i) search unemployment due to matching inefficiencies and (ii) mismatch between micro markets. This feature allows the importance of search relative to structural unemployment to be tested. As such it relates to a hot debate.

The empirical results presented for the Netherlands establish the quantitative importance of search duration and mismatch in a strict sense (that is the qualification dimension).

Figure 4.1 "Observed" and registered vacancies, 1960-1990



ad. i

Originally, I used four variables to endogenise $\sigma_{d,s}$:

- a) Capacity utilisation rate (*CAPUTIL*);
- b) Competitiveness situation [=price-index of competitive OECD exports (*COMP*)];
- c) Replacement ratio (*REPLRAT*); and
- d) Regional mismatch indicator (*REGMIS*).

CAPUTIL and *COMP* are typical Keynesian indicators⁶⁶ that work on $\sigma_{k,s}$. *REPLRAT* is a typical supply side indicator, which I assume to have a direct effect on $\sigma_{d,s}$.⁶⁷ *REGMIS* is a typical indicator that is usually perceived to be of importance in explaining mismatches. I could not find a suitable (and observable) measure that works on $\sigma_{c,s}$.

As for the effect of *CAPUTIL* on $\sigma_{d,s}$, I hypothesise that a rise in product demand induces *CAPUTIL* to rise, in turn increasing the demand for more qualified

⁶⁶ Why this is, is explained below.

⁶⁷ At this point it should be stressed that the effect of *REPLRAT* on mismatches is *ad hoc*, as is the effect of *REGMIS*, and not founded on the present model. The empirical results, however, indicate a clear influence on mismatches. The result demands further investigations.

personnel which is relatively hard to get ($\sigma_{k,s}$ increases).⁶⁸ With capital fixed in the short run the rise in product demand has no effect on $\sigma_{c,s}$. The effect on $\sigma_{c,k}$ is ambiguous since it depends on the regime the economy finds itself in. Besides, I consider this effect to be relatively small [compare eq. (4.3.20)]. Hence, a positive relationship is expected. *COMP* works in a similar way as *CAPUTIL* on $\sigma_{d,s}$. An improvement in *COMP* increases the product demand and raises the demand for more qualified personnel.⁶⁹

REGMIS is measured as:

$$REGMIS = \sum_r \left| \frac{U_r / U}{V_r / V} \right|$$

with U = unemployment, V = job vacancies and r = region (here province of the Netherlands).⁷⁰ There is a positive influence of *REGMIS* on $\sigma_{d,s}$.

ad. ii

To be able to break down structural unemployment into the two separate terms, search unemployment due to matching inefficiencies and mismatch between micro markets (mismatch in a strict sense: the qualification dimension, see Chapter 1), I use the model of Muysken and De Regt (1994).⁷¹ The empirical results establish the quantitative importance of search duration and mismatch in a strict sense.

To discuss their generalisation, the usual minimum rule is replaced by the matching (or hiring) equation:

⁶⁸ The effect is zero when on all micro markets demand is exerted as capacity demand; a highly unrealistic situation.

⁶⁹ Why do *COMP*, *CAPUTIL* and *REPLRAT* influence variances? The argument rests on the micro model extended with intertemporal spillovers of Lambert (1988, Appendix B). Lambert shows that the mobility parameters are negatively correlated with the variance of micro excess demands. These mobility parameters are in turn dependent on aggregate unemployment. It reflects that the mobility of workers is discouraged because they think their chance to get a job has declined, and employers have fewer incentives to content oneself with less appropriate qualifications with rising aggregate unemployment. The variables *COMP*, *CAPUTIL* and *REPLRAT* are used here in the same way as aggregate unemployment.

⁷⁰ For practical reasons, I use a somewhat different regional mismatch indicator and a different regional classification as in Chapter 7. This poses no extra problems.

⁷¹ The model is advocated by Muysken and De Regt not only for its ability to test for the relevance of search unemployment in structural unemployment, but also as a "dynamic alternative" to the very badly specified static employment functions in prior studies. The static estimations suffer from high autocorrelation and need an *ad hoc* trend term to find a satisfactory fit.

$$h = \frac{1}{\lambda} \text{Min}(u, v) \tag{4.4.1}$$

where λ measures the inefficiency of instantaneous matching or put differently the duration of search unemployment or search vacancies, depending on which side of the market is the shortest; h is the number of hires. To extend the model in a dynamic sense, the change in employment is assumed to be a function of quits (q) and hires:⁷²

$$\frac{dL_i}{dt} = -qL_i + h \tag{4.4.2}$$

Equation (4.4.2) states that the change in employment is equal to the difference between the number of new matches h on the one hand, and separations plus quits (indicated by q) on the other hand. The rate of separations and quits are assumed to be exogenously determined.

Equation (4.4.1) and (4.4.2) define the following change in micro employment:

$$\frac{dL_i}{dt} = -\left(q + \frac{1}{\lambda}\right)L_i + \frac{1}{\lambda} \text{Min}(L_i^d, L_i^s) \tag{4.4.3}$$

For a better understanding of (4.4.3), consider the steady state growth results, where supply and demand grow at some common constant rate κ , even though their levels may differ. Obviously, the growth rate of steady state growth employment L_i^* also equals κ . Hence solving (4.4.3), steady state growth employment L_i^* is equal to:

$$L_i^* = (1 - \varphi) \text{Min}(L_i^d, L_i^s) \tag{4.4.4a}$$

where

$$0 < \varphi = \frac{\lambda(q + \kappa)}{1 + \lambda(q + \kappa)} < 1 \tag{4.4.4b}$$

In Muysken and De Regt (1994) it is shown that aggregation over micro markets then yields the following employment function:

$$L = f\left(e^{-\Psi^S} L^S, e^{-\Psi^D} L^D\right) \tag{4.4.5a}$$

with

$$\Psi^S = \lambda(q + \Delta \ln L^S) \qquad 0 \leq \lambda \tag{4.4.5b}$$

⁷² I use the index i for a micro market only for employment, labour supply and demand, to have consistency in notation with the previous sections.

where $f(\cdot)$ is the C.E.S. function defined in (2.2.17) and $\Delta \ln L^S$ is the growth rate of labour supply. The sum of q and $\Delta \ln L^S$ corresponds to the inflow in unemployment, and as λ is the search duration of the short side of the market, an unemployment rate of at least ψ^S will be observed. A similar specification as (4.4.5b) applies to job search vacancies, ψ^D .

The new element in (4.4.5), compared to (2.2.17), is that supply and demand are weighted with the factors ψ^S and ψ^D , which can be shown to be equal to aggregate search unemployment and vacancies. The "correction" for search unemployment and job vacancies does not influence the mismatch (in the strict sense) represented by β_1 . Note further that λ essentially makes up the difference between the "static" and the dynamic specification. For λ , or search duration, being zero the static version emerges. This enables testing the "contribution" of search to structural unemployment.

The dynamic specification leads, however, to a different interpretation of the mismatch parameter. It can be shown (Muysken and De Regt, 1994) that:

$$\frac{1}{\beta_1^*} = \frac{1}{\beta_1} (1 + 2(1 - \tau) \lambda (\lambda - 1))^{\frac{1}{2}} \quad 0 < \tau < 1 \quad (4.4.6)$$

where β_1 is defined in (2.2.17) and τ represents the strength of intertemporal spillover effects. Thus, the mismatch parameter, $\frac{1}{\beta_1^*}$ is not only positively related to the standard deviation of excess supplies over micro markets - as is the case in the "static" specification above - but it also varies negatively with τ and λ : The higher the intertemporal spillover effect and the higher the search duration on the short side of the micro market will be, the lower the mismatch. Note that search duration and mismatch operate in a capillary way: a higher search duration **within** a market is exchanged by a lower mismatch **between** markets.

The coefficient τ is not separately identified in the model, but is set fixed at 0.8, as in Muysken and De Regt.

The accompanying expression of structural unemployment (with equal growth rates of labour supply and demand) shows:

$$u_4^* = \psi^S + \frac{1}{\beta_1^*} \ln 2 \quad (4.4.7)$$

Equation (4.4.7) states that u_4^* consists of two elements, which may be labelled search and mismatch unemployment. The first part, ψ^S , corresponds to aggregate search unemployment. The second part, $\frac{1}{\beta_1^*} \ln 2$, essentially corresponds to the mismatch unemployment, $\frac{1}{\beta_1} \ln 2$, already identified in the static approach (*c.f.* eq. 2.2.20).

In the estimation of the employment equation, I apply the correction of labour supply for the long-term unemployed from Muysken and De Regt (1994). In the literature, the long-term unemployed are usually said to have a relatively small matching efficiency. Describing the matching efficiency of the long-term unemployed as a fraction $(1 - \alpha)$ of the short-term unemployed, the correction on labour supply shows:

$$L^{SE} = L^S (1 - \alpha u^L) \quad 0 \leq \alpha \leq 1 \quad (4.4.8)$$

where L^{SE} is effective labour supply and u^L is the rate of long-term unemployed. To explain the border cases: When α equals zero, the long-term unemployed fully contribute to labour supply and when α equals one, the long-term unemployed effectively are no part in labour supply. Allowing for different matching efficiencies between the long-term and the short-term unemployed influences the measurement of structural unemployment according to (4.4.7). The new specification may be shown to be:

$$u_4^* = \Psi^S + \frac{1}{\beta_1^*} \ln 2 + \frac{1}{2} \alpha u^L \quad (4.4.9)$$

As before the first term represents search unemployment. Mismatch unemployment constitutes two parts: The first part, $\frac{1}{\beta_1^*} \ln 2$, corresponds to the mismatch unemployment in the static approach. The second part, $\frac{1}{2} \alpha u^L$, arises because of the lower matching efficiency of the long-term unemployed. It is evident that the lower the efficiency, the higher the share of long-term unemployed in total unemployment, the higher the rate of structural unemployment.

Estimation results⁷³

The employment equation is estimated by using additive disturbance terms. The sample period is 1960-1989. In the initial estimation of the employment equation, I included *CAPUTIL*, *COMP*, *REPLRAT* and *REGMIS* as variables operating on $\frac{1}{\beta_1^*}$

Since *COMP* was insignificant (though with the expected sign), I deleted this variable in the estimation. The final results are presented in Table 4.1. Two variants are shown. The first is the "static" version of the model ($\lambda = 0$); the second is the dynamic version ($\lambda \neq 0$). The latter allows the contribution of search to structural unemployment to be tested.

⁷³ A complete description of the data used is given in Annex 4A. In Muysken and De Regt (1994) a justification of the data can be found.

Table 4.1 Estimates of the employment function ("static" and dynamic variant), 1960-1989¹

| Variable | Static model ($\lambda=0$) | Dynamic model ($0 < \lambda < 1$) |
|--|------------------------------|-------------------------------------|
| | Parameter estimates | |
| Constant | - 0.0982** (-4.8) | - 0.0873** (-3.5) |
| Labour supply correction for long-term unemployed (α) | 0.9972 (29.3) | 0.9959** (28.1) |
| CAPUTIL | 0.0008** (5.1) | 0.0007** (3.0) |
| REPLRAT | 0.0005** (10.3) | 0.0005** (10.0) |
| REGMIS ² | 0.0233** (4.0) | 0.0211** (3.2) |
| SEARCH DURATION (λ) | 0 | 0.0098 |
| loglikelihood | 173.035 | 173.466 |
| 100 $\hat{\sigma}$ | 0.068 | 0.068 |
| R ² | 0.9997 | 0.9997 |
| D.W. | 2.0 | 1.9 |

1. t-values in parentheses. ** shows a 1% significance level.

$$2. \text{REGMIS} = \sum_r \left| \frac{U_r / U}{V_r / V} \right| \text{ with } r = \text{province.}$$

The "static" version performs very well. The $\frac{1}{\beta_1}$ variables have the expected signs and are highly significant. This implies that the capacity utilisation rate, the replacement ratio, and the regional mismatch indicator positively influence the mismatch parameter $\frac{1}{\beta_1}$. The coefficient α , measuring the labour supply correction for the long-term unemployed is estimated free and very highly significant. It takes a value that is almost equal to its maximum value one, indicating that the long-term unemployed hardly contribute to effective labour supply. It is worth noting that the estimation results do not exhibit dynamic misspecification! This is a rather spectacular finding in the face of prior studies revealing very low Durbin

Watson statistics. The overall explanatory power remains the same. Also, in Muysken and De Regt (1994) the autocorrelation problem (even in the dynamic version of the model) still overshadows their results. The results presented here indicate that their autocorrelation problem stems from a misspecification of the mismatch which they take a function of time (in the familiar *ad hoc* fashion). It is not following from a dynamic misspecification.

The estimation of the dynamic version of the model shows a value of λ which is not significantly different from zero. This implies that search unemployment must be very low and not contributing to structural unemployment. Evidently, structural unemployment is determined solely by the mismatch between micro markets. This finding for the Netherlands contradicts that of Layard *et al* (1991), who attribute a rather small influence to mismatches between micro markets for most European countries (Layard *et al*, p.38).

Figure 4.2 portrays the estimated structural unemployment rate along with the unemployment and vacancy rate (rates: all in terms of labour supply) for the period 1960-1989 (static estimation). Clearly, mismatch unemployment is rising over time; from 2.3% in 1960 to 4.9% in 1989.⁷⁴ The results on the dynamic version of the model are used to determine the quantitative importance of:⁷⁵

- mismatch unemployment, distinguished in the static mismatch (in a strict sense) and the matching efficiency of the long-term unemployed, and
- search unemployment

for the rate of structural unemployment. The various contributions are shown in Figure 4.3. The negligible influence of search is corroborated. From the figure it further follows that mismatches between markets are a substantial part of structural unemployment, and remarkably, the explanation of the development in the structural unemployment rate lies mainly with the long-term unemployed.

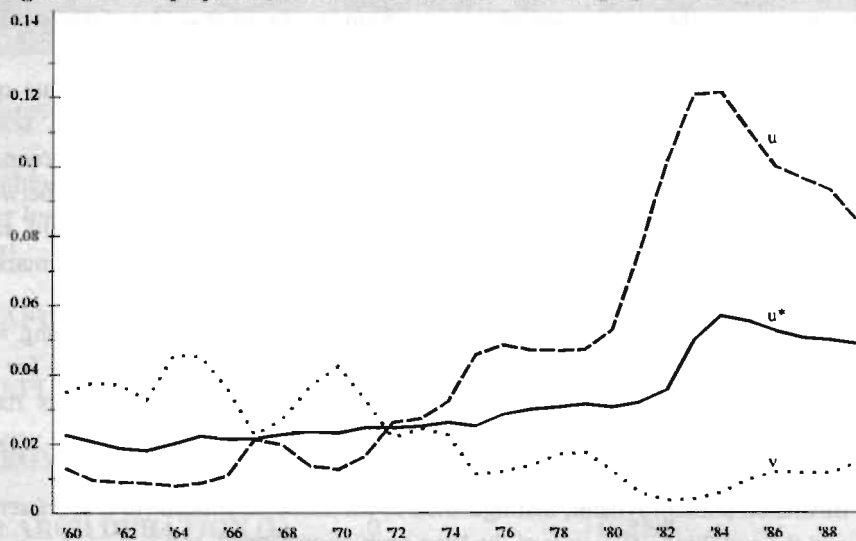
4.5 Concluding remarks

For theoretical reasons, I advocated the use of the two-level employment function - instead of the one-level alternatives - to measure structural unemployment. The two-level C.E.S.-function has at the first level aggregate Keynesian demand and capacity demand, and at the second level aggregate Keynesian demand and capacity demand on the one hand and aggregate labour supply on the other.

⁷⁴ For 1983, u^* is 4.8%. This is remarkably close to the percentage of 4.4%, obtained from the simple model in Section 3.3.

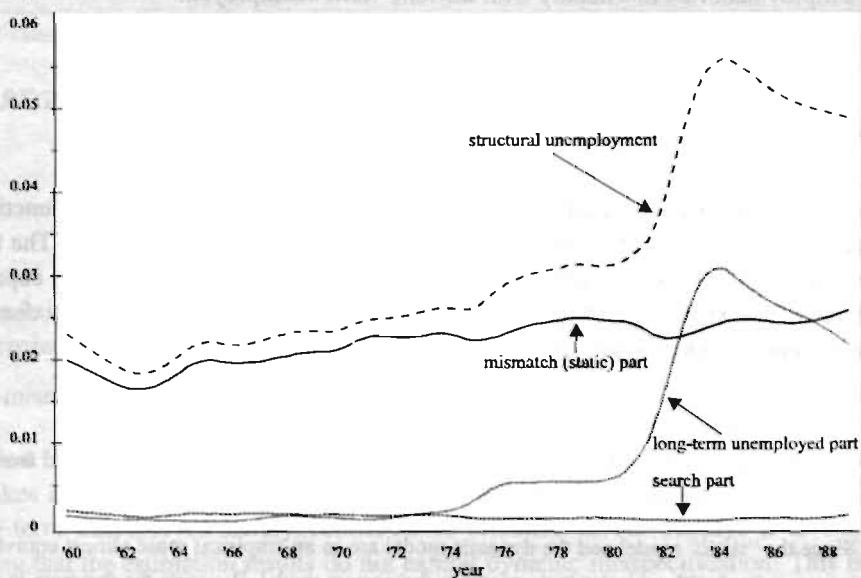
⁷⁵ Since the "static" model and the dynamic model are in an empirical sense almost equivalent, this is permitted, although the static model is the correct model. By using the dynamic model I am able to show the development of search unemployment.

Figure 4.2 Unemployment, vacancies and structural unemployment, 1960-1989¹⁾



¹⁾ Figures are in terms of labour supply; the corrected data for u and v are used.

Figure 4.3 Decomposition of structural unemployment, 1960-1989¹⁾



I elaborated on the use of this C.E.S.-function in the literature by deriving explicit relationships between the parameters of the C.E.S. and the micro mismatches. This in turn made it possible to derive explicit relationships between the rate of structural unemployment and the micro mismatches. Quite a remarkable outcome was the close correspondence of the relationships between the C.E.S.-parameters and the micro mismatches: The C.E.S.-parameter, $\frac{1}{\beta}$ is close to 0.5 times the micro mismatch in all cases. It should be stressed at this point that the theoretical analysis assumes capacity and labour mismatches as distinct phenomena, while of course there could be close links between both. Surely more work is needed to understand the interactions. In the model of Chapter 6 the complex links between the supply and demand determinants are modelled endogenously.

When it comes to the estimation, the two-level C.E.S. is not applicable due to absence of relevant vacancy data. I have to fall back on the one-level C.E.S.-function with observed labour supply and labour demand (the variant type 1 of the analytically-based method). Labour supply is then measured as the sum of employment and unemployment, and labour demand as the sum of employment and job vacancies. Such a procedure can of course only be implemented if the measurement error in the unemployment and vacancy data is relatively small. I used the corrected vacancy data of Muysken and De Regt (1994), and the corrected unemployment data of the Central Planning Bureau. These data are based on survey information from Statistics Netherlands and are far more reliable than the initial available register data. However, part of the measurement problem remains, because vacancies are hard to measure. Consequently, the "good quality" of data assumption is probably the weak spot of the analysis.

In the empirical specification of the model, two important aspects were recognised. First, I used the theoretical insights mentioned earlier to endogenise the C.E.S. mismatch parameter. In the final version of the model, the capacity utilisation rate, the replacement ratio and a regional mismatch indicator were incorporated. They were all hypothesised to have a positive influence on the mismatch parameter. Second, I allowed for dynamics and search duration in the sense of Muysken and De Regt (1994). The main motivation for this was that their introduction of dynamics and search enabled me to test for the contribution of search unemployment to structural unemployment.

The results may be summarised as follows:

- ⇒ The predicted effects of the three variables mentioned on the mismatch parameter are corroborated.
- ⇒ The long-term unemployed fully contribute to structural unemployment.
- ⇒ Structural unemployment rises from 2.3% in 1960 to 4.9% in 1989.
- ⇒ Search unemployment is only a negligible part of structural unemployment.
- ⇒ Structural unemployment is predominantly a phenomenon of mismatches between micro markets.

⇒ The explanation of the development in the structural unemployment rate lies mainly with the long-term unemployed.

The dominant role of the long-term unemployed has a clear implication for future research. It evidently shows how important it is not to treat the effect of the long-term unemployed on effective labour supply exogenously, but instead considering why workers become long-term unemployed (because they lack adequate skills *e.g.*), to explain the mismatches. This is probably one of the key questions.

ANNEX 4A Data

- L : Employment in manyears from the Central Economic Plan (1994).
- U : Unemployed workers (“Werkzoekenden zonder baan) from the Central Economic Plan (1994), Annex A6. The “werkzoekenden zonder baan” comprise those who are out of work, available for work within two weeks, and actively seeking for work.
- U^L : Long-term unemployed, computed as $\frac{LTU(GAB)}{UT(GAB)} * U$, where $LTU(GAB)$ is the number of long-term registered unemployed and $UT(GAB)$ registered unemployed, both obtained from the Sociaal Economische Maandstatistiek (several editions) of Statistics Netherlands.
- V : Corrected vacancies from Muysken *et al* (1994, p. 283).
- VIN : Corrected vacancies inflow, computed as $VINR * VREG$.
- $VINR$: Vacancy inflow from the Sociaal Economische Maandstatistiek (several editions).
- $VREG$: Vacancy registration rate computed as $\frac{VNOT}{V} * 100$.
- $VNOT$: Notified vacancies from the Sociaal Economische Maandstatistiek (several editions) of Statistics Netherlands.
- M : Number of matches computed as $VIN - \Delta V$.
- Q : Quit rate computed as $\frac{M - \Delta L}{L}$.
- CAPUTIL:** Capacity utilisation rate, obtained from the Central Planning Bureau, The Hague.
- REPLRAT:** Replacement ratio (benefits/minimum wage), from the Central Economic Plan (1994), Annex A6.
- REGMIS:** Regional mismatch indicator, computed as $\sum_r \left| \frac{U_r / U}{V_r / V} \right|$, with r =province; time series data on unemployment and vacancies are obtained from the Ministry of Social Affairs and Employment (register data).

5 The performance of the monte carlo pseudo maximum likelihood method

5.1 Introduction

Chapter 4 closed the discussion of the analytically-based method to analyse mismatches and structural unemployment. This chapter and Chapter 6 focus on the simulation-based method, as a means to measure mismatches on the labour market. Compared to the analytically-based methods, the estimation of the macro model using the simulation technique leads to less intuitive and not so readily interpretable results, simply because analytics stops when a number of restrictive assumptions implicit in the analytically-based method are relaxed. This is the price of a more general specification of the macro model in which, in addition to supply and demand, spillovers between the labour market and the goods market are modelled endogenously and unobserved at the micro level (see Chapter 1 and Chapter 2).

The estimation of such a model is theoretically only feasible when a simulation-based technique, or a Monte Carlo Pseudo Maximum Likelihood (MCPML) is used. Whether it is feasible in practice is of another matter. I only have to refer to analytical and statistical problems met in recent studies estimating “only” the “simple” canonical aggregate minimum model using Maximum Likelihood (ML) estimation (see Kooiman and Klok, 1985 and Sneessens, 1983). The likelihood function for these types of models has the remarkable property that it can become unbounded in the parameter space, *i.e.* in the space of potential parameters values there are some isolated points at which the likelihood becomes arbitrarily large. Although this is diagnosed in most empirical studies, since it is invariably accompanied by one or the other residual variance becoming zero (a spurious maximum), this property may sometimes impede the computation of reasonable estimates. The successful implementation of maximum likelihood techniques in the present disequilibrium setting requires that the likelihood function is twice differentiable with respect to the parameters, a property that is not always satisfied.

Although based on ML estimation, these findings indicate the importance of judging the practical feasibility of MCPML estimation before the method is applied to the extended model of Chapter 6. The feasibility of MCPML estimation is investigated by using a simple model of the labour market that allows comparison

with analytical results. I am looking, in particular, at the performance of the technique on a finite number of replications in the Monte Carlo study. To this I consider both the first and the second-order based MCPML method. The feasibility study should provide an answer about:

- a. the method to be used for the estimation of the extended model; and
- b. the weaknesses of the method that can be expected in the extended model estimation.

In all this, one should not overlook the fact that the first-order method is less suited to measure mismatches (see Section 2.4 about the “observational equivalence”). Under the best circumstances the first order method allows an upper bound estimate. This limitation should also be weighted in the final choice of the method.

The rest of this chapter is organised in four sections.⁷⁶ Section 5.2 specifies the model and briefly describes the estimation procedure which has already be discussed in Section 2.3. Section 5.3 reports the results of performance testing with observed data on labour supply and demand. Section 5.4 presents the results from a limited Monte Carlo study with simulated data. Section 5.5 concludes the chapter.

5.2 Model specification and estimation

To explore the feasibility of the MCPML methods, the model to which it is applied should be simple, enabling the comparison with the exact PML results. The model with micro markets does not satisfy this condition. Therefore, I will use the canonical single market disequilibrium model, on which additionally assumptions about the observability of the expected quantities of labour supply and demand are imposed. I consider the following specification:

$$L_t^S = E(L_t^S) + \sigma_s \eta_{at}^s \quad (5.2.1)$$

$$L_t^D = E(L_t^D) + \sigma_d \varepsilon_{at}^d \quad (5.2.2)$$

$$L_t = \text{Min}(L_t^S, L_t^D) \quad (5.2.3)$$

where $E(L_t^S)$ and $E(L_t^D)$ are the expected levels of aggregate labour supply and labour demand. In the testing of the model, these usually latent variables and L_t are observed for $t = 1, \dots, T$ from:

⁷⁶ I refer to Chapter 2 for the principles of the MCPML technology of estimating the model and the different variants. Recent findings on (MC)PML estimation can be found in Gouriéroux, Monfort and Trognon (1984) and Laroque and Salanié (1989).

$$E(L_t^S) = L_t + U_t \tag{5.2.4}$$

$$E(L_t^D) = L_t + V_t \tag{5.2.5}$$

with $U =$ unemployment, and $V =$ vacancies. For all t , the aggregate shocks, η_a^s and ε_a^d are normal disturbances with zero mean and covariance matrix equal to identity.⁷⁷ The standard errors σ_s and σ_d are strictly positive.

The vector of parameters to be estimated then is $\sigma=(\sigma_s, \sigma_d)$. To estimate the canonical model, I only use PML₁ and PML₂ (see Chapter 2). The PML₃ method, which is also a second order method is excluded from the analysis as it is not expected to yield extra information. For the model (5.2.1-5.2.5) the PML estimators are compared with the MCPML estimators. I distinguish an exact and a Monte Carlo (MC) variant of PML estimation. This distinction is made to monitor the extra variability in the parameter estimates induced by the simulation of the objective function. In the exact variants, explicit expressions for $E(L_t)$ and $V(L_t)$ are used, whereas in the MC variants these are assumed to be untractable.

In the MC variants of PML estimation, the procedure comprises the minimisation of approximate PML functions as defined in (2.3.5) and (2.3.7), except that the computation of $E(L_t)$ and $V(L_t)$ is performed by using Monte Carlo methods. As to the working of the MCPML methods in relation to the above labour market model, it suffices to refer to Section 2.3, where I already explained the method for the canonical disequilibrium model. The only difference is that in Section 2.3 aggregate labour supply and demand were treated as latent variables. To gather up the threads it is, however, useful to repeat some aspects.

The exact version of pseudo maximum likelihood estimation is based upon the exact expressions of the first two moments of the endogenous variable L , which show (dropping the time subscript):

$$E(L) = L^S \Phi\left(\frac{L^D - L^S}{\sigma^*}\right) + L^D \Phi\left(\frac{L^S - L^D}{\sigma^*}\right) - \sigma^* \phi\left(\frac{L^D - L^S}{\sigma^*}\right) \tag{5.2.6}$$

⁷⁷ One could also impose a log-normal distribution. This, however, would somewhat complicate the calculations and does not seem necessary in the face of the results of Chapter 3, evidencing the indifference between the normal and the log-normal distribution in the present context.

$$E(L^2) = (L^S + \sigma_s^2)\Phi\left(\frac{L^D - L^S}{\sigma^*}\right) + (L^D + \sigma_d^2)\Phi\left(\frac{L^S - L^D}{\sigma^*}\right) - \sigma^*(L^D + L^S)\varphi\left(\frac{L^S - L^D}{\sigma^*}\right) \tag{5.2.7}$$

The expression for $V(L)$ follows from: $V(L) = E(L^2) - \{E(L)\}^2$. $V(L)$ is not monotonous in any of the parameters. The PML₁ estimator consists in maximising the likelihood function of the pseudo model $L_t = E(L_t(\sigma^*)) + \epsilon_t$, where the ϵ_t 's are independent centred unit variance normal variables, i.e. $\hat{\sigma}^*$ minimises the function $\frac{1}{2} \sum_{t=1}^T (L_t - E(L_t(\sigma^*)))^2$. Under PML₂ the ϵ_t 's are independent centred unit variance normal variables with variance $V(L(\sigma_s, \sigma_d, \sigma^*))$. The likelihood to be minimised with respect to σ_s , σ_d and σ^* shows:

$$\frac{1}{2} \sum_{t=1}^T \left(\frac{(L_t - E(L_t(\sigma_s, \sigma_d, \sigma^*)))^2}{V(L_t(\sigma_s, \sigma_d, \sigma^*))} + \log V(L_t(\sigma_s, \sigma_d, \sigma^*)) \right).$$

It is clear from the specifications that only the estimator $\sigma^* = \sqrt{\sigma_s^2 + \sigma_d^2}$ is identified under PML₁. Under PML₂, σ_s and σ_d are both identified. The Monte Carlo pseudo maximum likelihood estimators make use of the empirical approximations of the first two moments. These are obtained by drawing randomly and independently at each period, G values of the pair of aggregate shocks, η_a^s and ϵ_a^d and then computing:

$$E(L_t) = \frac{1}{G} \sum_{g=1}^G L_g = \bar{L}_t \quad \forall t, \dots, T \tag{5.2.8}$$

$$V(L_t) = \frac{1}{G-1} \left(\sum_{g=1}^G L_{tg}^2 - G\bar{L}_t^2 \right) \quad \forall t, \dots, T \tag{5.2.9}$$

where G is the number of replications in the Monte Carlo part. For each t these specifications are substituted in the optimising function, which is minimised.⁷⁸ From Theorem I in Laroque and Salanié (1989) it is known that for any T , any converging

⁷⁸ Note that numerical errors in the use of explicit expressions for $E(L)$ and $V(L)$ can be seen as special cases of the procedure if G is large, apart perhaps from the normal distribution used.

sequence of approximate estimators converges to the corresponding PML estimators when G goes to infinity.

5.3 Performance testing

Before discussing the empirical results, four points need to be stressed in advance:

1. The computation of $E(L)$ and $V(L)$ in the MC variants is done by using different numbers of replications of η_a^s and ε_a^d with G varying from 25 to 400 and sample of $T = 28$.
2. In generating the random functions, in the MCPML₁ case, I imposed $\sigma_s = \sigma_d$, for convenience.⁷⁹
3. The differences in identifiability of the parameters between the two PML cases complicate the comparison of the parameter estimates. The result from PML₁ estimation is an estimate of the standard error of the function $h(\sigma) = \sqrt{\sigma_s^2 + \sigma_d^2}$, whereas PML₂ gives the standard error of both σ_s and σ_d . To facilitate comparison, an estimate of the standard error of $h(\sigma)$ under PML₂ estimation is needed as well. To that purpose a linearisation of the function $h(\sigma)$ with respect to σ_0 is suggested, i.e.:

$$h(\sigma) = h(\sigma_0) + \frac{\partial h}{\partial \sigma} (\sigma - \sigma_0) \tag{5.3.1}$$

Thus the variance of $h(\sigma)$ can easily be calculated as:

$$Var(h(\sigma)) = \frac{\partial h}{\partial \sigma} Var(\sigma) \frac{\partial h}{\partial \sigma} \tag{5.3.2}$$

The optimising routine produces an estimate of $Var(\sigma)$ and the first derivatives of h with respect to σ can be calculated straightforwardly.

4. Kooiman and Kloek (1979) estimated equation (5.2.7) with an additive disturbance term taking $E(L)$ observed by L . Consequently, their results may be compared with my results of exact PML₁ estimation. I used their data on L^s, L^d to measure the corresponding expected values, and L to have this check on the estimation results (sample period 1948-1975).

To estimate the model the Davidon-Fletcher-Powell (DFP) optimising algorithm was applied using a unimodal line search routine with analytical first derivatives. The first derivatives for the exact variants and the second derivatives for the MC

⁷⁹ The fact that σ_s and σ_d are not individually identified should justify this “restriction”. Note that this restriction does not imply restricted optimisation. The only property used is that $E(L)$ is estimated consistently if $G \rightarrow \infty$.

variants are presented in Annex 5A. Annex 5B derives the covariance matrix in a more general setting.

Table 5.1 presents the results from the experiments. As part of the performance testing, the analytical second derivatives were compared with the (inverse) updated Hessian from the optimising routine. Reported values on H^{-1} correspond to the updated (inverse) Hessian matrix from the DFP algorithm; H_{an}^{-1} is the PML (inverse) Hessian computed from equation (5A.11) of Annex 5A. The reported statistics are based on 25 independent runs for each situation. G , the number of replications required for the computation of the random function is reported in the first column; $G = \infty$ represents the results from PML estimation using the exact expression for $E(L_i)$. It is obvious from the table that with the number of replications increasing, the estimate of σ^* converges to its value obtained from the exact variant. It is also clear that the standard error of the estimates as computed from the Hessian is in all cases very close to the one obtained from the exact version.

Table 5.1. Monte Carlo PML₁ estimates of the canonical labour market model

| MCPML₁ | G | $\hat{\sigma}^*$ | $SD(\hat{\sigma}^*)$ | H' | S.D. | \hat{H}_{an}^{-1} | S.D. | $SD(\hat{\sigma}^*)$ | S.D. |
|--------------------------|----------|------------------|----------------------|--------|-------------|---------------------|-------------|----------------------|-------------|
| SE(MEAN) | 25 | 171.5 | 9.7 | 0.1330 | 0.0128 | 0.1170 | 0.0117 | 2.557 | 0.126 |
| | | | 1.9 | | | | | | |
| SE(MEAN) | 50 | 178.2 | 8.6 | 0.1223 | 0.0101 | 0.1215 | 0.0091 | 2.607 | 0.098 |
| | | | 1.7 | | | | | | |
| SE(MEAN) | 100 | 181.3 | 5.2 | 0.1244 | 0.0161 | 0.1234 | 0.0056 | 2.628 | 0.059 |
| | | | 1.0 | | | | | | |
| SE(MEAN) | 200 | 183.4 | 3.8 | 0.1262 | 0.0048 | 0.1248 | 0.0044 | 2.644 | 0.046 |
| | | | 0.8 | | | | | | |
| SE(MEAN) | 400 | 184.2 | 2.0 | 0.1268 | 0.0031 | 0.1254 | 0.0028 | 2.650 | 0.029 |
| | | | 0.4 | | | | | | |
| Exact | ∞ | 185.6 | | 0.1330 | | | | 2.730 | |

- 1) The reported figures are means computed for 25 independent runs and associated standard deviations can be computed by dividing the reported *S.D.* values by $\sqrt{25}$. In the table only the standard errors of the means of $\hat{\sigma}^*$ are explicitly shown, to facilitate interpretation of the results.
- 2) Estimated from H^{-1} in each run.

It is worth noting that despite the Monte Carlo bias in the simulations, the estimate of the variance is hardly influenced. From one single MCPML₁ estimation it is not possible to obtain an estimate of the variance induced by the simulation, and from several estimations for a single value of G it is not possible to estimate the extra bias. It is encouraging that the values for H^{-1} and \hat{H}_{an}^{-1} are roughly alike; this implies that second derivatives as approximated by the optimisation routine are reasonable (at least, in the one parameter case). The PML results compare relatively well with the estimates found in Kooiman and Kloek (1979). Their estimate of σ^* is 181.7 (my estimate is 185.6). They do not provide estimates of the standard error. Therefore, I estimated the equation again with their data. It turned out that σ^* is estimated to be 181.4 and the standard error 5.7 which should be compared to the square of my estimate $[(2.72)^2=7.5]$.

A very uncomfortable fact is the high number of Monte Carlo replications required to obtain estimates that are acceptably close to the estimates of the exact PML variant. This not only applies to the parameter values but also to the standard deviations. Note, for instance that in the present example, a standard deviation of 2.0 for $G = 400$ in the Monte Carlo simulation has to be compared with the standard error 2.65 as estimated from the Hessian. So to require a minimum of 400 Monte Carlo replications does not seem to be exaggerated. Such large number of replications could make the implementation of MCPML methods in larger and more complex models difficult (for instance, the model of Chapter 6) because of the heavy computational burden.⁸⁰

The results from PML₂ estimation are troublesome (not reported for brevity). Negative $L_t - E(L_t)$ values were found for each t both for the exact variant and the MC variant. This is unacceptable. The estimate of σ^* is much lower than in the PML₁,

⁸⁰ Therefore I attempted to develop a bias-corrected variant of MCPML estimation in order to examine whether consistency at a far smaller number of replications, G , could be obtained. As a theoretical result: a bias-corrected variant for MCPML₁ estimation could be established. For MCPML₂ this is still troublesome (see Annex 5C). This is an improvement to the work of Laroque and Salanié (1989) who proved the consistency only for an infinite number of G . Experiments with the bias-corrected version of MCPML₁ on the above model specification corroborated the theoretical speculations and efficiency gains can indeed be realised. In fact it turns out that relatively low values for G are sufficient to obtain convergence to the exact parameter estimates compared to the situation where no bias correction is implemented (see also Annex 5C).

Despite the fact that the results on the bias-corrected version are interesting and could be used in a number of applications (see: Laffont, Ossard and Vuong, 1995), I decided not to present them in the main text but in an annex. The reason is that an empirically implementable bias-corrected version of MCPML₂ is not available yet. And since on the results of this chapter I decided to implement the MCPML₂ to the larger model of Chapter 6, presenting the results of bias-corrected MCPML₁ version would be a bit much.

case (compare 98.1 with 185.6). But that the corresponding standard error of $h(\sigma)$ in the PML₂ case, which equals 44.9, is substantially larger than the PML₁ value of 2.7 (see Table 5.1).

From asymptotic theory, PML₂ should be more efficient than PML₁, though in this case the sample size is only 28 and under PML₂ one more parameter is estimated. Besides efficiency considerations the explanation may also come from the fact that the model is inadequate because no allowance has been made for micro markets. From Bierings and Sneek (1989) it follows that $E(L_t)$ goes down if the number of micro markets, N , gets larger in the model with micro markets. Hence, if $N=1$ is imposed as a restriction, then under PML₂ indeed all values, $L_t - E(L_t)$, could indeed be expected to be negative. From this it may be concluded that there exists a trade-off between efficiency and robustness: PML₁ is less efficient, but robust with respect to the existence of micro markets, whereas PML₂ is more efficient but overly sensitive to the existence of micro markets. Future research should elaborate on this.

5.4 Monte Carlo evidence

To substantiate the claim made above, I undertook a limited Monte Carlo study by generating artificial data according to the aggregate model implied by equations (5.2.1-5.2.5). I tried four different pairs of pre-set values of σ_s and σ_d . For each of the four sets of parameter values, the following was done 50 times. Two vectors of 28 independent $N(0, 1)$ variables were drawn.⁸¹ These are the values of

η_a^s and ε_a^d . L_t^S , L_t^D and L_t were generated using (5.2.1 - 5.2.5) for the given values of expected supply and demand which is the observed supply and demand in Kooiman and Kloek. Subsequently, PML₁ and PML₂ were used to obtain the parameter estimates.

In Table 5.2 the results from the Monte Carlo simulation are presented. Note that I reported the standard deviations as emerging from 50 estimates (the line entries) and the estimated standard deviations as computed from H^{-1} , denoted by $\hat{SD}(\cdot)$, where the dot indicates the parameter (the column entries). The table shows that both the PML₁ and PML₂ estimates are biased downwards in all cases. The conclusion from the table is that the performance of PML₂ is superior to PML₁, as the downward bias is considerably less (compare in the first line entry 26.6 and 22.0 with 28.3) and the standard deviations of $\hat{\sigma}^*$ are much smaller (compare in the second line entry 3.82 and 15.5). Examining a couple of sets of residuals for both models revealed "random" sequences. However, $\hat{SD}(\cdot)$ under PML₁ severely

⁸¹ $T=28$ represents the actual sample size for which observations on labour supply and demand are available from Kooiman and Kloek.

underestimates the true standard deviation (compare for $\hat{\sigma}^* = 28.3$, the value 6.00 in the first line entry with 15.5 in the second line entry). For PML_2 the bias is apparently upwards. (compare, for $\hat{\sigma}^* = 28.3$, the value 20.30 in the first line entry with 3.82 in the second line entry). Test statistics based on H^{-1} are therefore likely

Table 5.2 Estimates for σ^* under PML_1 and PML_2 (artificial data for different pairs of σ_d and σ_s)¹

| | | PML_1 | | | PML_2 | | | | | |
|------------------------|------|------------|------------------|----------------------|------------------|------------------|------------------|----------------------|----------------------|----------------------|
| (σ_d, σ_s) | | σ^* | $\hat{\sigma}^*$ | $SD(\hat{\sigma}^*)$ | $\hat{\sigma}_d$ | $\hat{\sigma}_s$ | $\hat{\sigma}^*$ | $SD(\hat{\sigma}_d)$ | $SD(\hat{\sigma}_s)$ | $SD(\hat{\sigma}^*)$ |
| (20, 20) | | 28.3 | 22.0 | 6.00 | 18.6 | 18.5 | 26.6 | 20.6 | 19.6 | 20.30 |
| | S.D. | | 15.5 | 2.955 | 3.86 | 4.10 | 3.82 | 4.53 | 4.51 | 3.34 |
| | S.E. | | 2.19 | 0.418 | 0.546 | 0.579 | 0.540 | 0.641 | 0.638 | 0.472 |
| (50, 100) | | 111.8 | 99.5 | 3.27 | 48.0 | 88.3 | 103.5 | 112.3 | 99.5 | 72.6 |
| | S.D. | | 32.9 | 0.649 | 22.1 | 19.5 | 16.3 | 51.6 | 33.7 | 11.0 |
| | S.E. | | 4.65 | 0.091 | 3.124 | 2.754 | 2.307 | 7.305 | 4.765 | 1.557 |
| (100, 50) | | 111.8 | 94.3 | 3.17 | 91.7 | 40.4 | 103.9 | 106.4 | 194.4 | 76.3 |
| | S.D. | | 37.4 | 0.520 | 21.9 | 25.8 | 19.5 | 44.3 | 239.6 | 17.1 |
| | S.E. | | 5.29 | 0.074 | 3.104 | 3.651 | 2.753 | 6.259 | 33.889 | 2.423 |
| (100, 100) | | 141.8 | 122.1 | 2.96 | 95.3 | 87.7 | 133.2 | 160.8 | 163.3 | 90.8 |
| | S.D. | | 44.2 | 0.419 | 24.8 | 27.1 | 19.4 | 62.0 | 70.7 | 16.9 |
| | S.E. | | 6.25 | 0.059 | 3.508 | 3.828 | 2.749 | 8.765 | 9.997 | 2.385 |

1. Means and standard deviations obtained from 50 independent runs. Standard errors of the means are computed by dividing the S.D. by $\sqrt{50}$.

to yield incorrect significance levels; under PML_1 parameter values would be severely underestimated, under PML_2 they would be overestimated. The standard deviations for $\hat{\sigma}^*$ were usually smaller than those of $\hat{\sigma}_s$ and $\hat{\sigma}_d$, because of positive correlation between $\hat{\sigma}_s$ and $\hat{\sigma}_d$. Note that under PML_2 usually none of the estimated parameters would appear to be statistically different from zero. This finding is indeed a problem. It is probably due to incorrectly estimated standard deviations of the parameters. At a late stage of the study I became aware of an

article of Laroque and Salanié (1990), who addressed the problem of spurious maxima in their evaluation of the (MC)PML methods applied to the canonical disequilibrium model. Since their conclusions could also be relevant, I mention them here. They discuss two types of spurious maxima:

1. The zero-variance case characterised by Goldfeld and Quandt (1975) for FIML estimation.
2. The so-called “one-sided” case, where all observations are estimated to belong to the same regime.

Both types of spurious maxima go together with large standard deviations of the estimated parameters. I conjecture that explanation 1 could be true. However, 2 is very unlikely since the optimising function tends to move away from the zero-variance boundary as a result of the imposed penalty function due to a small variability of the supply and demand data. The findings of Laroque and Salanié (1990) argue against this interpretation. They indicate a rather small number of spurious (one-sided) maxima in their simulations compared to, for instance MCPML₁. Consequently, the outcome is indecisive and further investigations into the occurrence of possible spurious maxima in the present application is required. However, it can hardly be contested that the standard errors of the parameters should be used with caution.

Despite the previous point, I believe the following conclusions may be drawn from the (limited) Monte Carlo study. MCPML₂ is more efficient than MCPML₁. Moreover, I conjectured that the aggregate model might not be an adequate model for the data because PML₂ performs reasonably well on the simulated data sets. This is contrary to what I expected, given the experience of the previous section.

When I compare my findings with those of Laroque and Salanié (1990), their results are *grosso modo* in accordance. As far as PML₂ is concerned, I find a small slightly downward bias, for PML₁ the size of the bias is comparable. However, they found a large upward bias where I find a large downward bias. The results of Laroque and Salanié are very much in favour of MCPML₂. MCPML₁ is regarded as unsuitable. They find the asymptotic biases for MCPML₂ to be small with even relatively small values of G ; MCPML₁ is not regarded as an alternative. It requires too high a G to be manageable. In this respect Laroque and Salanié (1990) find for MCPML₂ that even though the estimator is biased, the asymptotic bias appears to be small even for a rather small number, G , of simulations: it is about 2% for $G=20$. These aspects should also be envisaged in the decision on which method should be used for estimating the model of Chapter 6 and contribute to the use of MCPML₂.

Given the results and the experiences of Laroque and Salanié (1990), I do not consider the MCPML₁ estimation and use the MCPML₂ instead, for the larger model of Chapter 6. This offers the additional opportunity to test the presence of

micro markets, since the mismatch variances across markets are identified only under PML_2 .

5.5 Concluding remarks

The results of this chapter lend support to the following conclusions. MCPML₁ estimation of a simple aggregate min model for the labour market with observed supply and demand requires a large number of replications in the Monte Carlo part for given T . MCPML₂ estimation of the model ends up with somewhat senseless results, perhaps because the model is inadequate by not taking into account micro markets. I concluded that there is a trade-off between efficiency and robustness. To further investigate the efficiency question, I undertook a Monte Carlo study with simulated data sets. It follows that (MC)PML₂ is more efficient than (MC)PML₁. Besides, both PML₁ and PML₂ estimates are biased downwards in all cases. The performance of PML₂, however, is superior to PML₁ as the downward bias is considerably less and the standard deviations of $\hat{\sigma}^*$ are far less. In all this two points should be noted. First, the aggregate model is probably not an adequate model for the data under PML₂, because PML₂ performs reasonably well on the simulated data (for which the “true” model is indeed as the aggregate variant). Second, large standard deviations of the parameter estimates were found under MCPML₂. It was conjectured from the results of Laroque and Salanié (1990) that this could be due to a small variability of the supply and demand data, so that the region of a “one-sided” spurious maximum is approached. Given the findings of Laroque and Salanié, indicating that under MCPML₂ spurious maxima of this kind are hardly a problem, this remains an unsolved issue.

The findings of this study and of Laroque and Salanié (1990) added are in favour of MCPML₂ estimation. Hence, the model of Chapter 6 is estimated on the basis of this method. This has the additional advantage that the presence of micro markets can be tested, and the mismatches between them. The problematic points of MCPML₂ estimation may not be forgotten in the application to the larger model. Finally, I stress the problem of the required number of replications in the Monte Carlo part of the estimation. It was found that a minimum of 400 replications is not exaggerated. In Annex 5C, a bias-corrected version of MCPML estimation was developed, which demands far a fewer replications. However, the bias-corrected MCPML₁ can only be implemented.⁸² The uncorrected version must be used for the estimation of the model of Chapter 6 with MCPML₂. To impose a number of 400 replications would make the estimation in this case intractable. So, as a second-best

⁸² This leaves apart that the results of Annex 5C, which established consistency of the estimators for a finite number of replications under MCPML₁, provides interesting applications to other models (see Laffont, Ossard and Vuong, 1995).

solution, I took a number of 50, both for the aggregate as the spatial disturbances. This is in line with Laroque and Salanié (1989), who did not report any problems in this respect. In the practice of optimising, I will, however, critically follow the behaviour of the optimising function for a different number of replications.

ANNEX 5A Computation of derivatives

The first and second derivatives of the likelihood function with respect to the parameters are necessary for the search routine of the DFP optimising routine and for computing the covariance matrix. The expressions used in the estimations are presented below. They apply to the canonical model (5.2.1-5.2.3).

PML₂, exact variant

I present the results for PML₂ using analytical expressions for the moments of the endogenous variable, L . The results for PML₁ simply follow as a special case.

I use E , $E2$ and V to abbreviate $E(L)$, $E(L^2)$, $V(L)$. Let $\sigma^* = \sqrt{\sigma_s^2 + \sigma_d^2}$ and $f = f(\sigma^*, \sigma_s, \sigma_d)$ be differentiable, then:

$$\frac{\partial f}{\partial \sigma_k} = \frac{\partial f}{\partial \sigma_k} + \frac{\partial f}{\partial \sigma^*} \frac{\partial \sigma^*}{\partial \sigma_k} \quad k=d,s \quad (5A.1)$$

where $\frac{\partial f}{\partial \sigma_k}$ on the left hand side means the "total" derivative with respect to σ_k and on the right side it is strictly the partial derivative with respect to σ_k . From this chain rule one obtains for $k = d, s$:

$$\frac{\partial E}{\partial \sigma_k} = \left(L^S - L^D \right) \left(-\frac{Z}{\sigma^{*2}} \right) \varphi \left(\frac{Z}{\sigma^*} \right) - \left(1 + \frac{Z^2}{\sigma^{*2}} \right) \varphi \left(\frac{Z}{\sigma^*} \right) \quad (5A.2)$$

$$\begin{aligned} \frac{\partial E2}{\partial \sigma_k} = & \left\{ \left(L^{S^2} - L^{D^2} + \sigma_s^2 + \sigma_d^2 \right) \left(-\frac{Z}{\sigma^{*2}} \right) \varphi \left(\frac{Z}{\sigma^*} \right) - \right. \\ & \left. \left(L^S + L^D \right) \left(1 + \frac{Z^2}{\sigma^{*2}} \right) \varphi \left(\frac{Z}{\sigma^*} \right) \right\} \frac{\sigma_k}{\sigma^*} - 2 \sigma_k \Phi \left(\frac{Z}{\sigma^*} \right) + 2 \sigma_k \quad (5A.3) \end{aligned}$$

with: $Z = L^D - L^S$; φ = the standard normal density function and Φ = cumulative normal density. From $V = E2 - E^2$ it follows that:

$$\frac{\partial V}{\partial \alpha_k} = \frac{\partial E^2}{\partial \alpha_k} - 2E \frac{\partial E}{\partial \alpha_k} \quad (5A.4)$$

Computation of the derivatives of the PML functions (see main text, this chapter) is straightforward using:

$$\frac{\partial \psi}{\partial \alpha_k} = \frac{\partial \psi}{\partial E} \frac{\partial E}{\partial \alpha_k} + \frac{\partial \psi}{\partial V} \frac{\partial V}{\partial \alpha_k} \quad (5A.5)$$

where $\psi = \psi(E, V)$ for PML₂ and $\psi = \psi(E)$ for PML₁.

PML₁ Monte Carlo variant

$E(L_t)$ is replaced by $\bar{L}_t = \frac{1}{G} \sum_{g=1}^G L_{tg}$ ($\forall t = 1, \dots, T$), where:

$$L_{tg} = \text{Min}\left(L_t^S + \sigma_s \eta_{tg}^s, L_t^D + \sigma_d \varepsilon_{tg}^d\right)$$

Because only the function $\sigma^* = \sqrt{\sigma_s^2 + \sigma_d^2}$ is identified, $\sigma_s = \sigma_d = \frac{\sigma^*}{\sqrt{2}}$ is taken,

i.e.:

$$L_{tg} = \text{Min}\left(L_t^S + \frac{\sigma^*}{\sqrt{2}} \eta_{tg}^s, L_t^D + \frac{\sigma^*}{\sqrt{2}} \varepsilon_{tg}^d\right) \quad (5A.6)$$

One clearly has:

$$\frac{dL_{tg}}{d\sigma^*} = \begin{cases} \frac{1}{\sqrt{2}} \eta_{tg}^s & \text{if } L_t^S + \frac{\sigma^*}{\sqrt{2}} \eta_{tg}^s \leq L_t^D + \frac{\sigma^*}{\sqrt{2}} \varepsilon_{tg}^d \end{cases} \quad (5A.7)$$

$$\frac{dL_{tg}}{d\sigma^*} = \begin{cases} \frac{1}{\sqrt{2}} \varepsilon_{tg}^d & \text{elsewhere} \end{cases} \quad (5A.8)$$

Note that the derivatives are discontinuous.

Define $\delta_{tg}^s = 1$ if the condition in (5A.7) is satisfied and $\delta_{tg}^s = 0$ in all other cases, then it follows that:

$$\frac{d\bar{L}_t}{d\sigma^*} = \frac{1}{G\sqrt{2}} \sum_{g=1}^G \delta_{tg}^s (\varepsilon_{tg}^d - \eta_{tg}^s) + \eta_{tg}^s \quad (5A.9)$$

$$\frac{d\psi_{1T}}{d\sigma^*} = -2 \sum_{t=1}^T (L_t - \bar{L}_t) \frac{d\bar{L}_t}{d\sigma^*} \quad (5A.10)$$

$$\frac{d^2\psi_{1T}}{d\sigma^*} = 2 \sum_{t=1}^T \frac{d\bar{L}_t}{d\sigma^*} \quad (5A.11)$$

Note that in (5A.11) the identity $\frac{d\bar{L}_t}{d\sigma^*} = 0$ is used. In Table 5.1 (main text) the inverse of the Hessian matrix using (5A.11) is compared with the one obtained from the DFP algorithm.

PML, Monte Carlo variant

In this case $E(L_t)$ is replaced by $\bar{L}_t = \frac{1}{G} \sum_{g=1}^G L_{tg}$ ($\forall t = 1, \dots, T$) and $V(L_t)$ by

$$\text{Var}(L_t) = \frac{1}{G-1} \left(\sum_{g=1}^G L_{tg}^2 - G\bar{L}_t^2 \right) \quad (\forall t = 1, \dots, T). L_{tg} \text{ is defined as in (5A.6) but}$$

without the restriction $\sigma_s = \sigma_d = \frac{\sigma^*}{\sqrt{2}}$. Then it follows that:

$$\left. \begin{array}{l} \frac{dL_{tg}}{d\sigma_s} = 0 \\ \frac{dL_{tg}}{d\sigma_d} = \epsilon_{tg}^d \end{array} \right\} \begin{array}{l} \text{if } E(L_t^S) + \sigma_s \eta_{tg}^s > E(L_t^D) + \sigma_d \epsilon_{tg}^d \end{array} \quad (5A.12)$$

$$\left. \begin{array}{l} \frac{dL_{tg}}{d\sigma_s} = \eta_{tg}^s \\ \frac{dL_{tg}}{d\sigma_d} = 0 \end{array} \right\} \begin{array}{l} \text{if } E(L_t^S) + \sigma_s \eta_{tg}^s < E(L_t^D) + \sigma_d \epsilon_{tg}^d \end{array}$$

All derivatives are now obtained through simple substitution.

ANNEX 5B Computation of covariance matrix for PML₁ and PML₂ (general results)

The asymptotic covariance matrix of is $J^{-1} I J^{-1}$ (Gouriéroux *et al*,1984), where

$$J_{ij} = E_x E \left(-\frac{\partial^2 \psi(x, \theta)}{\partial \theta_i \partial \theta_j} \right) \text{ and } J_{ij} = E_x E \left(-\frac{\partial \psi(x, \theta)}{\partial \theta_i} \frac{\partial \psi(x, \theta)}{\partial \theta_j} \right) \quad (5B.1)$$

and x is a vector of exogenous variables; θ is a vector of parameters.

PML₁

The pseudo function according to PML₁ is given by:

$$\Psi(x, \theta) = \frac{1}{2} (y - f(x, \theta))' \Sigma^{-1} (y - f(x, \theta)) = \frac{1}{2 \sigma^2} \sum_{i=1}^T \psi_i(x, \theta) \quad (5B.2)$$

where Σ is the diagonal covariance matrix of the endogenous variable y ; $f(\cdot, \cdot)$ represents the expectation of y as a function of x and θ . Note that for expository reasons y and f are taken to be vectors instead of matrices, *i.e.* for each t , a scalar is dealt with. From standard differentiation rules one has:

$$E \left(-\frac{\partial^2 \psi(x, \theta)}{\partial \theta_i \partial \theta_j} \right) = \frac{\partial f'}{\partial \theta_i} \Sigma^{-1} \frac{\partial f}{\partial \theta_j} \quad (5B.3)$$

$$E \left(\frac{\partial \psi(x, \theta)}{\partial \theta_i} \frac{\partial \psi(x, \theta)}{\partial \theta_j} \right) = E \left(\frac{\partial f'}{\partial \theta_i} \Sigma^{-1} (y - f)(y - f)' \Sigma^{-1} \frac{\partial f}{\partial \theta_j} \right) = \frac{\partial f'}{\partial \theta_i} \Sigma^{-1} \frac{\partial f}{\partial \theta_j} \quad (5B.4)$$

$$E_x \left(\frac{\partial \psi(x, \theta)}{\partial \theta_i} \frac{\partial \psi(x, \theta)}{\partial \theta_j} \right) = \frac{1}{T} \sum_{i=1}^T \frac{\partial \psi_i(x, \theta)}{\partial \theta_i} \frac{\partial \psi_i(x, \theta)}{\partial \theta_j} \quad (5B.5)$$

PML₂

The pseudo function according to PML₂ is given by:

$$\Psi(x, \theta) = \frac{1}{2} (y - f(x, \theta))' g^{-1}(x, \theta) (y - f(x, \theta)) + \frac{1}{2} \log \det g(x, \theta) \quad (5B.6)$$

where g is the diagonal covariance matrix of y as a function of x and θ . From Laroque and Salanié (1989) it is known that:

$$E \left(\frac{\partial^2 \psi(x, \theta)}{\partial \theta_i \partial \theta_j} \right) = \frac{\partial f'}{\partial \theta_i} g^{-1} \frac{\partial f}{\partial \theta_j} + \frac{1}{2} \text{trace} \left(g^{-1} \frac{\partial g}{\partial \theta_i} g^{-1} \frac{\partial g}{\partial \theta_j} \right) \quad (5B.7)$$

$$E \left(\frac{\partial \psi(x, \theta)}{\partial \theta_i} \frac{\partial \psi(x, \theta)}{\partial \theta_j} \right) = \frac{\partial f'}{\partial \theta_i} g^{-1} \frac{\partial f}{\partial \theta_j} + \frac{1}{2} \text{trace} \left(g^{-1} \frac{\partial g}{\partial \theta_i} g^{-1} \frac{\partial g}{\partial \theta_j} \right) \quad (5B.8)$$

ANNEX 5C The bias-corrected MCPML variant

5C.1 Theoretical results

In this annex, a modification to the procedure in Laroque and Salanié (1989) is suggested. They proved consistency of the simulated PML estimators if T and G go to infinity, with G is the number of Monte Carlo replications. For the MCPML₁ case a bias correction is introduced in this section such that consistency is achieved even for finite G . Let:

$$\psi_T = \frac{1}{T} \sum_{t=1}^T (L_t - E(L_t))' (L_t - E(L_t)) \quad (5C.1)$$

where $E(L_t)$ is a function of the parameter θ of interest and where L_t may be a vector (in this section only). The MCPML method is based on the fact that ψ_T converges to some function $\psi(\theta)$; if ψ has a unique minimum at θ_0 and θ_0 is the “true” parameter I then hope that $\hat{\theta}^T$ converges to θ_0 , where $\hat{\theta}^T$ minimises ψ_T . In the MCPML₁ method, $E(L_t)$ is replaced by a random variable that converges to $E(L_t)$ if some index G goes to infinity. The corresponding function shows:

$$\psi_{2T} = \frac{1}{T} \sum_{t=1}^T (L_t - E(L_t) + B_{tG})' (L_t - E(L_t) + B_{tG}) \quad (5C.2)$$

where B_{tG} is a random variable that is independent of L_t ; B_{tG} is chosen to have zero mean and finite second order moments. It should be noted that the variables are independent for different values of t . Then using the results of Gouriéroux *et al* (1984) and Laroque and Salanié (1989) the following theorem applies:

THEOREM 1: $\psi_{2T} - \left(\psi_T + \frac{1}{T} \sum_{t=1}^T B_{tG}' B_{tG} \right)$ converges almost surely to zero if

$$T \rightarrow \infty \therefore \quad (5C.3)$$

Note that the expression in (5C.3) equals $\frac{2}{T} \sum_{t=1}^T B_{tG}' (L_t - E(L_t))$.

B_{tG} is a simple average of simulated $(L_t - E(L_t))$ values according to so-called “crude” Monte Carlo (Hammersley and Handscomb, 1964). Of course variance reducing techniques such as importance sampling, control variates or antithetic variables may alleviate the computational burden of the simulated PML method substantially in general, but is not applied here as it is rather specific to each application. For the same reason I did not consider replacing simple averages of simulated values by more robust and under certain conditions more efficient functions of simulated values.

From the theorem, I conclude that there are two possibilities with respect to consistency:

1. if $Var(B_{tG})$ does not depend on θ , then $\frac{1}{T} \sum_{t=1}^T B_{tG}' B_{tG}$ will become a ‘flat’ function if $T \rightarrow \infty$. In this case the sequence $\{B_{tG}\}$ will not have any influence on the location of the minimum of $\psi(\theta)$ and if T is large enough, then $Var(B_{tG})$ does not necessarily have to be very small for all t . Consistency, in this case is approved for finite G .
2. if $Var(B_{tG})$ is a function of θ then there is a problem, because the shape of $\psi(\theta)$ will be different from $\psi(\theta) + \frac{1}{T} \sum_{t=1}^T B_{tG}' B_{tG}$. In this case the number of replications G must be large enough to have a small $\frac{1}{T} \sum_{t=1}^T B_{tG}' B_{tG}$ relative to $\psi(\theta)$. Consistency is only obtained if $G \rightarrow \infty$.

It is precisely point 2 of the conclusion above which applies to the MCPML₁ method: $Var(B_{tG})$ depends on θ , since $Var(L_t)$ depends on σ_r , σ_d and G . Thus, the method is not consistent for any finite G . To see how consistency can be obtained for finite G , replace ψ_T in (5C.2) by:

$$\psi_{3T} = \frac{1}{T} \sum_{t=1}^T \left((L_t - E(L_t) + B_{tG})' (L_t - E(L_t) + B_{tG}) - trace(Var(B_{tG})) \right) \quad (5C.4)$$

Then the following theorem applies:

THEOREM 2: Let the bias-corrected MCPML₁ estimator be given by the parameter value $\hat{\theta}^T$ that minimises ψ_{3T} , then this estimator is consistent if $T \rightarrow \infty$, even for finite G ⁸³ ∴.

⁸³ This result which dates back to 1989 (Bierings and Sneek, 1989) was found independently by Laffont, Ossard and Vuong (1995).

Note from the definitions of ψ_T and ψ_T that:

$$\psi_{3T} - \psi_T = \frac{1}{T} \sum_{t=1}^T \left(B_{tG}' B_{tG} - \text{trace}(\text{Var}(B_{tG})) + 2B_{tG}'(L_t - E(L_t)) \right) \quad (5C.5)$$

Under appropriate moment conditions this expression converges almost surely to zero if $T \rightarrow \infty$ for any G . In the present set-up convergence is uniform over θ

(note that $\text{Var}(B_{tG}) = \frac{1}{G} \text{Var}(L_t)$).

It is unlikely that the asymptotic distribution of the bias-corrected MCPML₁ estimator coincides with that of the PML₁ estimator itself for any finite G . Developing ψ_T and ψ_T into the familiar Taylor expansions yields:

$$\sqrt{T} \frac{\partial \psi_3(\hat{\theta}_3^T)}{\partial \theta} = 0 = \sqrt{T} \frac{\partial \psi_3(\theta_0)}{\partial \theta} + \sqrt{T} (\hat{\theta}_3^T - \theta_0) \frac{\partial^2 \psi_3(\theta_0)}{\partial \theta \partial \theta'} \quad (5C.6)$$

$$\sqrt{T} \frac{\partial \psi_1(\hat{\theta}^T)}{\partial \theta} = 0 = \sqrt{T} \frac{\partial \psi_1(\theta_0)}{\partial \theta} + \sqrt{T} (\hat{\theta}^T - \theta_0) \frac{\partial^2 \psi_1(\theta_0)}{\partial \theta \partial \theta'} \quad (5C.7)$$

The two matrices of second derivatives will converge to the same (constant) matrix if $T \rightarrow \infty$. (compare Laroque and Salanié, 1989). What is needed is the same

asymptotic distribution for $\sqrt{T} \frac{\partial \psi_3(\theta_0)}{\partial \theta}$ and $\sqrt{T} \frac{\partial \psi_1(\theta_0)}{\partial \theta}$ but from (5C.5) would

result:

$$\sqrt{T} \frac{\partial \psi_3(\theta_0)}{\partial \theta} = \sqrt{T} \frac{\partial \psi_1(\theta_0)}{\partial \theta} + K \quad (5C.8)$$

where K is a non-degenerate finite stochastic variable. The variance of K would approach zero if $G \rightarrow \infty$.

The bias correction can have considerable practical advantages. Assume that the PML₁ estimator should be approximated as closely as possible by the bias-corrected MCPML₁ estimator. The first option obviously is to increase G , but computationally this is not very attractive. Moreover, one would still have to guess about the variability introduced by simulating the objective function for that particular value of G (just one realisation of the variable K is obtained). Another option is to do several independent estimations for fixed G ; the estimates can be averaged, the variability introduced by simulation goes down and one will have an idea about the magnitude of this variability (note the sample is over K). It should be stressed, however, that for fixed G and T one may still have a small sample bias.

The results for bias-corrected version for MCPML₂ could not be completely established. I present some preliminary findings. For expository reasons, it is again assumed that L_t is a scalar. Suppose $E(L_t)$ and $V(L_t)$ are obtained through simulation as discussed before and consider:

$$\frac{1}{T} \sum_{t=1}^T \left(\frac{(L_t - E(L_t) + C_{tG})^2}{V(L_t) + D_{tG}} + \log(V(L_t) + D_{tG}) \right) \quad (5C.9)$$

where for simplicity sake, C_{tG} and D_{tG} are assumed to be independent (one can make them independent!), and $Var(C_{tG})$ and $Var(D_{tG})$ approach zero if $G \rightarrow \infty$. Through expansion of denominator and logarithm it follows that:

$$\frac{1}{T} \sum_{t=1}^T \left\{ \left(\frac{(L_t - E(L_t))^2}{V(L_t)} + \frac{C_{tG}^2}{V(L_t)} + \frac{2C_{tG}(L_t - E(L_t))}{V(L_t)} \right) \left(1 - \frac{D_{tG}}{V(L_t)} + \frac{D_{tG}^2}{(V(L_t))^2} \dots \right) + \log V(L_t) + \left(\frac{D_{tG}}{V(L_t)} - \frac{D_{tG}^2}{(V(L_t))^2} \dots \right) \right\} \quad (5C.10)$$

Taking the expectation of this expression one can use the independence between C_{tG} and D_{tG} . If D_{tG} is constructed from an unbiased estimator of $V(L_t)$, then

$$\frac{1}{T} \sum_{t=1}^T \frac{D_{tG}}{V(L_t)}$$

will not cause the problem (apart from the existence of sufficiently

high order moments and also $\frac{1}{T} \sum_{t=1}^T \frac{C_{tG}^2}{V(L_t)}$ may be dealt with. Both series within

the curly brackets, however, are in general non-converging asymptotic series in G (see Sneek, 1983 for a definition of asymptotic series), so the optimal cut-off point depends on G . Although for finite, but large enough G it is possible to reduce the asymptotic bias if $T \rightarrow \infty$ of the MCPML₂ method, I was unsuccessful in completely reducing the bias proved and decided that it can not be implemented in the empirical analysis.

5C.2 Empirical results

In Table 5C.1, I present some results for the bias-corrected MCPML₁ estimation method together with its uncorrected counterpart. It is very obvious that at a sample size of $T = 28$ the bias is reduced considerably for all reported values of G ; in all cases the averages of the parameter estimates are within one standard error of the PML₁ estimate with an exact expression for $E(L_t)$, *i.e.* the bias is essentially removed. In Table 5C.1 on the other hand, even for $G = 400$ a bias was reported that

is statistically different from zero. Some experiments were undertaken with a larger number of runs. Using 200 independent runs for $G = 25$, a smaller value of $\hat{\sigma}^*$ resulted (187.93), with associated value $SD(\hat{\sigma}^*) = 12.4$ and thus with standard error 0.875.

Figure 5C.1 is presented to illustrate the importance of the bias correction graphically. The left side shows the estimation results for the ordinary MCPML, estimations based on the 25 independent runs. The right hand side shows the bias-corrected results. Note that in the figure 1.96 times standard errors are displayed and furthermore that the standard errors were subtracted from the heights of the bars for estimation. It is clear that the bias correction is at the cost of only a slight increase in the variation due to the Monte Carlo part.

Table 5C.1 Bias-corrected Monte Carlo PML₁ estimates of the canonical labour market model¹

| | G | $\hat{\sigma}^*$ | $SD(\hat{\sigma}^*)$ | H^{-1} | <i>S.D.</i> | <i>S.E.</i> ²⁾ | <i>S.D.</i> |
|-----------------|----------|------------------|----------------------|----------|-------------|---------------------------|-------------|
| SE(MEAN) | 25 | 188.7 | 12.0 | 0.1787 | 0.0342 | 3.150 | 0.287 |
| | | | 2.4 | | | | |
| SE(MEAN) | 50 | 187.0 | 9.5 | 0.1787 | 0.0138 | 3.161 | 0.123 |
| | | | 1.9 | | | | |
| SE(MEAN) | 100 | 185.7 | 5.5 | 0.1785 | 0.0136 | 3.160 | 0.120 |
| | | | 1.1 | | | | |
| SE(MEAN) | 200 | 185.7 | 3.1 | 0.1787 | 0.0063 | 3.163 | 0.056 |
| | | | 0.8 | | | | |
| SE(MEAN) | 400 | 185.4 | 2.0 | 0.1789 | 0.0045 | 3.165 | 0.039 |
| | | | 0.4 | | | | |
| | ∞ | 185.6 | | 0.1330 | | 2.730 | |

1) The reported figures are means computed for 25 independent runs and associated standard deviations.

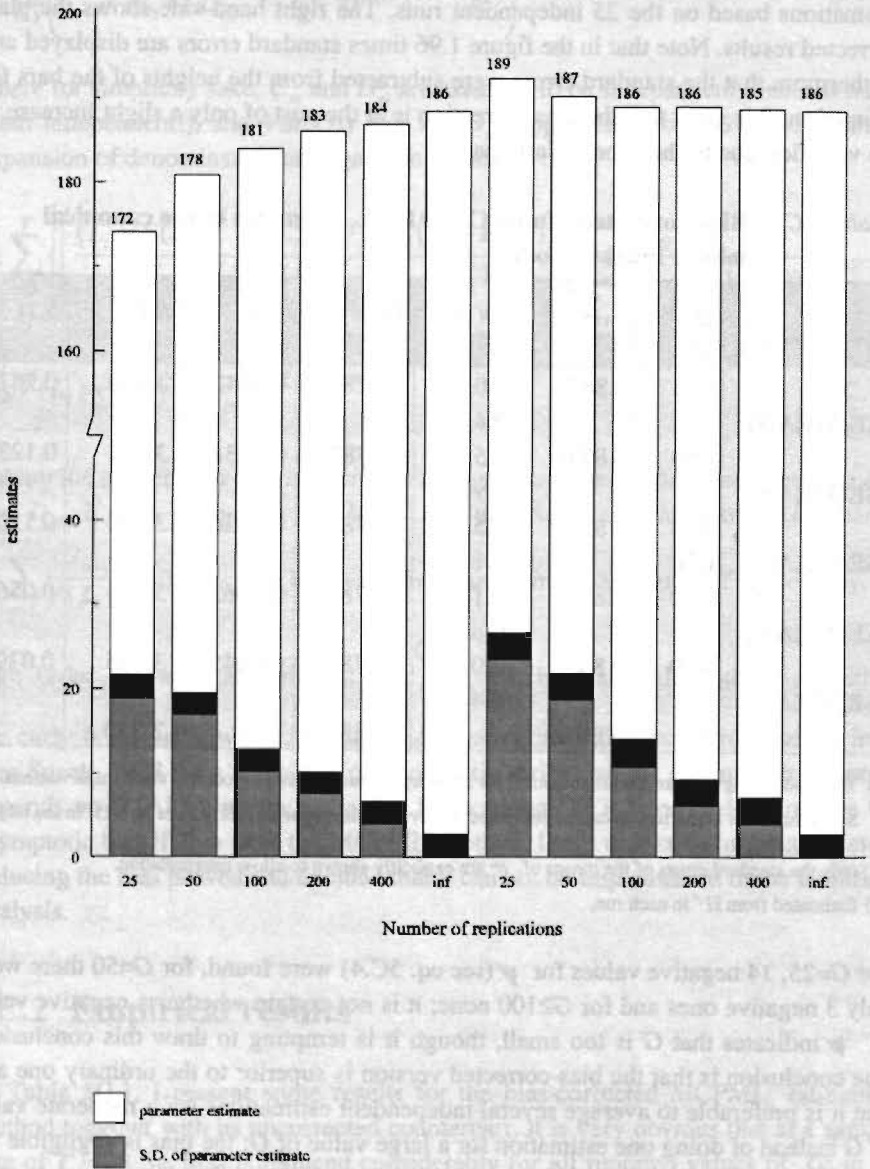
Standard errors of the means can be computed by dividing the reported *S.D.* values by $\sqrt{25}$. In the table

only the standard errors of the means of $\hat{\sigma}^*$ are explicitly shown to allow interpretation.

2) Estimated from H^{-1} in each run.

For $G=25$, 14 negative values for ψ (see eq. 5C.4) were found, for $G=50$ there were only 3 negative ones and for $G \geq 100$ none; it is not certain whether a negative value of ψ indicates that G is too small, though it is tempting to draw this conclusion. The conclusion is that the bias-corrected version is superior to the ordinary one and that it is preferable to average several independent estimations for a moderate value of G instead of doing one estimation for a large value of G ; the bias is negligible for moderate values of G , the standard deviation due to the simulation approaches zero anyway through averaging and at the same time one can estimate the latter standard deviation.

Figure 5C.1 Monte Carlo PML estimation and the bias-corrected Monte Carlo estimation



6 An econometric disequilibrium model for the Dutch economy estimated by Monte Carlo pseudo maximum likelihood

6.1 Introduction

This chapter presents an econometric macro model taking account of micro markets. In deviation from the approach in Chapter 4, the model fully accounts for endogeneity of incomplete market clearing at the micro level. The (effective) demand and supply variables are unobserved latent variables. In the model structural mismatches on the labour market (and the goods market) follow from a complex system of interrelated forces. As a result the causes of mismatches can not be strictly identified.⁸⁴ It is, however, clear that the effects on structural mismatches are regime specific, i.e depending on the typical commodity demand/supply and labour demand supply configuration. This fits the approach of Malinvaud (1977), who distinguished unemployment into a classical and a Keynesian component. Keynesian unemployment results from a lack of demand with respect to the usable productive capacity. The classical part is directly linked to the lack of employment posts, i.e. ultimately to a lack of profitable-usable capacity.

The focus of the chapter is on drawing inferences from the model estimates of τ_s and τ_d which are the mismatches of supply and demand, respectively. Related to the fact that τ_s and τ_d are not separately identified (see also Chapter 2), only an indication of the presence of structural mismatches can be obtained. To get an idea of the importance (or the size) of structural problems on the labour market, I compute the rate of structural unemployment from the estimations of unobserved supply and demand. These results can then be compared with those from the analytically-based method of Chapter 4.

To estimate the model, I apply the MCPML₂ simulation technique. I use annual data for the Netherlands covering the period 1950-1991.

⁸⁴ By contrast, in the approach of Chapter 4 structural mismatches were no intrinsic part of the model, but were envisaged by distinguishing a Keynesian (or a commodity demand) and a capacity component in the mismatch parameter.

The remainder of the chapter is organised as follows. I start with the description of the model in Section 6.2. This section also addresses the identification question of τ , and τ_u . The practice of optimising and the empirical results are presented in Section 6.3. Section 6.4 contains the conclusions.

6.2 The model

The model proposed here builds on the accumulated experience in estimating disequilibrium models since Kooiman and Kloek (1985), Sneessens (1983), Artus, Avouyi-Dovi and Laroque (1985), Laroque and Salanié (1989). The specification draws heavily on Laroque and Salanié (1989), which in turn relies on Artus, Avouyi-Dovi and Laroque (1985). This has the advantage of having a workable base-line model, thus allowing comparison of the results.

6.2.1 Structure of the model

Let me explain the basic features of the model. The model distinguishes two macroeconomic markets, the commodity market and the labour market. They are considered simultaneously, distinguishing firms and households. There can be excess demand (ED_Q) or excess supply (ES_Q) on the commodity market, excess demand (ED_L) or excess supply (ES_L) on the labour market. As is standard in this type of model four regimes (at given price and wage rate) may appear (Malinvaud, 1977). These are: repressed inflation (ED_Q, ED_L), underconsumption (ES_Q, ED_L), classical unemployment (ED_Q, ES_L) and Keynesian unemployment (ES_Q, ES_L).

Firms and households adapt their notional quantities on the one market in the face of rationing on the other (spillovers). In the econometric formulation of the model, three types of spillovers are allowed for. Firms may correct notional supply when rationed on the labour market, or notional labour demand when rationed on the commodity market. Households can only change their commodity demand (consumption and imports goods) when rationed on the labour market. Spillovers on labour supply from an excess commodity demand are considered unrealistic in European economies (Kooiman and Kloek, 1985).

The building blocks of the model are the supply and demand functions for labour and commodities. The model consists of eleven econometric equations. Five equations determine five (reduced-form) endogenous variables, *viz.* employment, L , gross domestic product, Q , household consumption, C , imports, M and exports, X . Investment demand, the demand for inventories and government demand are assumed to be exogenous, and captured by \bar{D} . The other six equations define the latent supply and demand variables, Q_e^s , Q_e^d , C_e^d , M_e^d , X_e^d , L_e^d and L_e^s , where the

subscript e denotes effective quantities. These are to be distinguished from notional quantities which will be indicated by “~”.

All equations are regime specific. They define four types of equilibria, which are shown in Scheme 6.1. As an example, consider the repressed inflation regime. In this regime, firms are rationed on the labour market and households on the commodity market. Hence, firms’ notional demand for labour equals its effective demand for labour. Firms’ effective goods supply is below its notional level as firms anticipate the rationing on the labour market. The notional quantities for households are equal to the effective quantities. On the labour market, transactions (L) equals labour supply, on the commodity market the transacted level (Q) equals effective goods supply. The transacted quantities of consumption (C), exports (X) and imports (M) follow from the respective notional quantities corrected by a fraction ρ of the unsatisfied goods demand. In all regimes, except the Keynesian unemployment regime, the quantities Q , C , M and X follow easily by substitution. In the Keynesian unemployment regime a system of two simultaneous equations has to be solved, since L_e^a is a function of itself through Q_e^s and Q_e^a .

6.2.2 Specification of the model

The specification of the model is at the micro market level. In line with Chapter 2, the stochastic specification of the model distinguishes an idiosyncratic error term and an aggregate error. The idiosyncratic error refers to the micro market, and the aggregate error to the aggregate supply and demand shocks. The error η is used for all supply variables and the error ε to the demand variables, The index i stands for idiosyncratic and the index a for aggregate. The vectors η and ε are independent centred normal variables with unit variance (see also Section 2.4). The vectors τ and σ represent the corresponding standard errors, where the σ is related to the aggregate shocks, and τ to the micro markets. The parameter τ reflects the mismatch parameters, which are of particular importance in this chapter.

Firms

I will now consider the precise specification of the model. Let firms maximise their monetary profit at current prices given a production function with an implicit fixed capital stock. The **production function** is of the Cobb-Douglas type with constant returns to scale:

$$\log \tilde{Q}^s(L) = \alpha_0 \log L + (1 - \alpha_0) \log K + \alpha_1 t + \alpha_2 \quad 1 \geq \alpha_0 > 0 \quad (6.2.1)$$

where $\tilde{Q}^s(L)$ denotes notional goods (only) when firms are not rationed on the labour market ($L^d < L^s$) supply, K is the capital stock; t is a time trend

Scheme 6.1 Equilibrium regimes^{1,2}

| | Repressed Inflation (RI) | Underconsumption (UC) | Classical Unemployment (CU) | Keynesian Unemployment (KU) |
|---|---|--|-----------------------------------|---|
| Commodity market | | | | |
| Effective supply: | $Q_e^s = Q^s(L^s)$ | $Q^s(L^s)$ | \bar{Q}^s | $Q^s(L_e^d)$ |
| Effective demand: | $Q_e^d = \bar{Q}^d$ | \bar{Q}^d | $Q^d(L_e^d)$ | $Q^d(L_e^d)$ |
| Consumption: | $C_e^d = \bar{C}^d$ | \bar{C}^d | $\bar{C}^d(\bar{L}^d)$ | $C^d(L_e^d)$ |
| Imports: | $M_e^d = \bar{M}^d$ | \bar{M}^d | $\bar{M}^d(\bar{L}^d)$ | $M^d(L_e^d)$ |
| Exports: | $X^d = \bar{X}$ | \bar{X} | \bar{X} | \bar{X} |
| Transactions: | $Q = Q_e^s$ | \bar{Q}^d | $Q^d(L_e^d)$ | $Q^d(L_e^d)$ |
| Consumption: | $C = \bar{C}^d - \rho_c(\bar{Q}^d - Q)$ | $\bar{C}^d - \rho_c(\bar{Q}^d - Q)$ | $C^d(L_e^d)$ | $C^d(L_e^d)$ |
| Imports: | $M = \bar{M}^d + \rho_m(\bar{Q}^d - Q)$ | $\bar{M}^d + \rho_m(\bar{Q}^d - Q)$ | $M^d(L_e^d)$ | $M^d(L_e^d)$ |
| Exports: | $X = Q - C + M$ | $Q - C + M$ | $Q - C + M$ | $Q - C + M$ |
| Labour market | | | | |
| Effective demand: | $L_e^d = \bar{L}^d$ | $L^d(\bar{L}^d, Q^s(L^s) - \bar{Q}^d)$ | \bar{L}^d | $L^d(\bar{L}^d, Q^d(L_e^d) - Q_e^s(L_e^d))$ |
| Effective supply: | $L_e^s = \bar{L}^s$ | \bar{L}^s | \bar{L}^s | \bar{L}^s |
| Transactions: | $L = \bar{L}^s$ | \bar{L}^s | \bar{L}^d | L_e^d |
| <p>1) A "-" refers to notional quantities and a "~" to given quantities. 2) RI: firms constrained on labour market, households on the commodity market; UC: only firms constrained; CU: workers constrained on labour market, firms on the commodity market; KU: only workers constrained.</p> | | | | |

incorporating technological change. Whenever firms are rationed on the labour market (labour shortages), equation (6.2.1) represents effective goods supply with L equal to L^s (compare the underconsumption regime). The supply of output is determined by the production function, once the level of employment is given. Effective supply, $Q^s(L)$, is assumed to adjust slowly towards the long-term target, \bar{Q}^s , by expressing $Q^s(L)$ as a geometric mean of the target and output in the previous period with respective weights λ_0 and $(1-\lambda_0)$:

$$\log Q^s(L) = \lambda_0 \log \bar{Q}^s + (1-\lambda_0) \log \bar{Q}_{-1} + \sigma_Q^s \eta_{Qa}^s + \tau_Q^s \eta_{Qa}^s \quad 1 > \lambda_0 > 0 \quad (6.2.2)$$

The parameter λ_0 measures the speed of adjustment. It contains information on the mean lag of adjustment towards the long run target⁸⁵, which is defined as $\frac{1-\lambda_0}{\lambda_0}$.

\bar{Q}_{-1} is the level of lagged (predetermined) output.

The determination of firm's **labour demand** may be described as follows. First a notional target (\bar{L}^d) is defined which follows from maximisation of short-run profit given predetermined price, wage and capital stock:

$$\log \bar{L}^d = \frac{1}{1-\alpha_0} \left(\log \alpha_0 + \log \frac{p}{w(1+t')} + (1-\alpha_0) \log K + \alpha_1 t + \alpha_2 \right) \quad (6.2.3)$$

where p is the product price, w is the nominal wage rate and t' is the rate of firms' tax and social contribution. As with output, it is assumed that firms slowly adapt their demand for labour to the long-term target. This gives for the demand for labour when firms do not experience rationing constraints:

$$\log \bar{L}^d = \lambda_1 \log \bar{L}^d + (1-\lambda_1) \log L_{-1} + \sigma_L^d \varepsilon_{La}^d + \tau_L^d \varepsilon_{Li}^d \quad 1 > \lambda_1 > 0 \quad (6.2.4)$$

The parameter λ_1 is a measure for the speed of adjustment of labour demand, \bar{L}^d , to the target; $\frac{1-\lambda_1}{\lambda_1}$ measures the mean lag. It can be given a theoretical justification from the rational expectations adjustment costs models for labour demand (Pfann, 1989) and explains why labour hoarding may occur.

Finally, firms also allow for rationing on the commodity market, which yields for effective labour demand:⁸⁶

⁸⁵ I have chosen to treat the dynamics in labour demand and good supply independently and thus *ad hoc*. Preferably, they should follow from the same program of the firm maximising profits under uncertainty conditions.

⁸⁶ This spillover effect allows the model to cope with the productivity cycle (Neftçi, 1984).

$$\log L^d = \log \bar{L}^d + \varsigma(\log Q - \log Q^s) \quad 1 > \varsigma \geq 0 \quad (6.2.5)$$

The second term of equation (6.2.5), *i.e.* $+\varsigma(\log Q - \log Q^s)$, represents the spillover effect from unsatisfied goods supply to the demand for labour. It is only effective in the computation of the equilibrium allocation in the Keynesian unemployment regime. In the excess demand regimes (repressed inflation and underconsumption) labour supply rules the game. In the classical unemployment regime, firms are not constrained on their sales implying, $L^d = \bar{L}^d$.

Households

Three commodity demand categories are distinguished: consumption demand (C^d), the demand for exports (X^d), demand for imports (M^d) and the residual demand (\bar{D}) which is exogenous:

$$Q^d(L) = C^d(L) + X^d(L) - M^d(L) + \bar{D} \quad (6.2.6)$$

Consumption demand depends on disposable wage income, R^d , and lagged consumption:

$$C^d(L) = \mu_c \chi_1 \left(\frac{R^d(L)}{P_c} \right) + (1 - \mu_c) C_{-1} + \sigma_c^d \varepsilon_{ca}^d + \tau_c^d \varepsilon_{ci}^d \quad 1 \geq \mu_c > 0 \quad (6.2.7)$$

The lagged consumption term is exogenous and should account for the smoothness in consumption over time which is explained for example by the permanent income theory.⁸⁷ The spillover from the labour market on consumption demand is realised through real disposable wage income which is partly endogenous through L . The specification for real disposable wage income shows:

$$R^d(L) = \bar{R}_{tot}^d - w(1 - t'')(\bar{L} - L) \quad (6.2.8)$$

where \bar{R}_{tot}^d is exogenous households disposable wage income, w is the gross wage rate per head; t'' is the rate of social security contributions and direct taxes on wage income paid by the employees, and \bar{L} is the observed level of employment. In

⁸⁷ The "permanent income" notion has often been used to justify the presence of the lagged dependent variable in a regression equation, even though the full permanent income theory (see Friedman, 1957) is not adopted. To fix ideas, consider the model: $C_t = k Y_t^P$ (with k is a positive constant and Y^P is permanent income) and $Y_t^P = (1 - \gamma) \sum_{j=0}^{\infty} \gamma^j Y_{t-j}$ (with γ a constant between zero and one). Then, on subtracting γ times the lagged value of each side of the equation and rearranging, yields: $C_t = \gamma C_{t-1} + k(1 - \gamma) Y_t$. This transformation is introduced by Koyck (1954).

(6.2.8) disposable income, $R^D(L)$, consists of an exogenous part, $\bar{R}_{tot}^D - w(1-t)L$ and an endogenous part, $w(1-t)L$.⁸⁸

Exports are determined by the traditional demand factors, world trade and the competitive position:

$$X^d = WD e^{-\varepsilon_2 \left(\frac{p_x}{p_{cx}} \right) + \varepsilon_0} + \sigma_x^d \varepsilon_{xa}^d + \tau_x^d \varepsilon_{xi}^d \quad (6.2.9)$$

where WD is world trade represented by the quantity of competing (OECD)

exports; $\frac{p_x}{p_{cx}}$ is a measure for exports competitiveness, or the exports price

relative to the price of competing exports. The quantity as well as the price of competing exports are reweighted according to the shares of the various commodities in Dutch exports and the shares of the various countries in Dutch trade.

Imports are modelled as a function of final demand and the imports price relative to the national product price. The spillover from the labour market works through via consumption demand.⁸⁹

$$M^d(L) = (C^d(L) + X^d + \bar{D}) e^{\varepsilon_2 \left(\frac{p_m}{p} \right) + \varepsilon_3 t + \varepsilon_0} + \sigma_m^d \varepsilon_{ma}^d + \tau_m^d \varepsilon_{mi}^d \quad (6.2.10)$$

The **supply of labour** has not received much attention in macroeconomic studies, since it does not appear in Keynesian models. This explains its rather poor treatment in empirical models (see also Kooiman and Kloek, 1985 and Sneessens, 1981). Ideally, one would want to derive labour supply from a formal treatment of optimal household behaviour with incomplete information and uncertainty. This is beyond

⁸⁸ Note that equation (6.2.8) in way represents a "trick" to make households disposable **wage** income endogenous. First, the observed disposable wage income is subtracted from the observed total disposable income. Then the unobserved disposable wage income which equals $w(1-t)L$ is added again.

⁸⁹ As well in the exports equation as in the imports equation, I considered to include the rate of capacity utilisation as explanatory variable. This would reflect the so called "Zijlstra effect". This effect implies that in a situation of depressed demand, firms may decide to sell their redundant goods at the foreign market. It further predicts that when domestic demand rises again, firms chose to sell their products in their own country again, due to the high costs of exports. A similar effect arises for imports. I did not incorporate these interesting spillover effects, since it would complicate the estimation of the model significantly. This comes mainly from the fact that the capacity utilisation rate should be modelled endogenously. In future applications of the model, it is worth experimenting with the "Zijlstra effect".

the scope of the present study and should be topic of a specific study on labour supply. Besides, one must be careful not to try to solve all the problems at once, so that the, still intricate, estimation of the model is not unnecessarily increased. Therefore, it seems wise to join the "bad" habit of earlier studies which modelled the unobserved effective labour supply as a linear and *ad hoc* function of the labour force, P (observed by employment plus unemployment):

$$L^s = (1 - \beta_i) P + \sigma_L^s \eta_{L_a}^s + \tau_L^s \eta_{L_i}^s \quad 1 > \beta_i \geq 0 \quad (6.2.11)$$

with:

$$\beta_i = \beta_0 \quad \forall 1950 \leq t < 1967$$

$$\beta_i = \beta_0 + \beta, \text{Max}(t-1968, 0) \quad \forall 1968 \leq t \leq 1991$$

The term $(1 - \beta_i)$ is a correction factor on the active population, P , for "natural" frictions (see Kooiman and Kloek, 1985, who adopt the same specification). It should reflect the fraction of unemployment which is unavoidable while people have to spend time to find a job (search unemployment). Compared to Chapter 4, search unemployment is modelled rather simply in this equation and should ideally be modelled endogenously. Still, it is interesting to compare the estimation outcomes about the significance of search unemployment over time.

Determination of endogenous variables (transactions)

The last five equations are definitional identities. Endogenous transactions, *i.e.* employment, L , gross domestic product, Q , are determined by the minimum rule:

$$L = \text{Min}(L^s, L^d(Q, L)) \quad (6.2.12)$$

$$Q = \text{Min}(Q^s(L), Q^d(L)) \quad (6.2.13)$$

The effective demand for goods is a linear function of L . It consists of four components: $C^d(L)$, $M^d(L)$, X^d and an exogenous residual, \bar{D} . A linear rationing scheme is imposed, so that when demand is rationed:

$$C = C^d(L) - \rho_c(Q^d(L) - Q) \quad (6.2.14)$$

$$X = X^d(L) - \rho_x(Q^d(L) - Q) \quad (6.2.15)$$

$$M = M^d(L) + \rho_m(Q^d(L) - Q) \quad (6.2.16)$$

with: $\rho_c + \rho_x + \rho_m = 1$

For given values of the parameters and disturbances Q, L, C, M, X can be easily determined. In the Keynesian unemployment regime $L (= L_e^d)$ must be solved from a system of two equations:

$$a) \log L_e^d = \log \tilde{L}^d + \zeta (\log Q^d(L_e^d) - \log Q^s(L_e^d))$$

$$b) \log Q^s(L_e^d) = \log Q^s + \alpha_0 \lambda_0 (\log(L_e^d) - \log(\tilde{L}^d))$$

Substitution of b) in a) yields:

$$c) (1 + \zeta \alpha_0 \lambda_0) \log L_e^d - (1 + \zeta \alpha_0 \lambda_0) \log \tilde{L}^d - \zeta \log Q^d(L_e^d) + \zeta \log Q^s(L_e^d) = 0$$

This equation is solved for L_e^d numerically using the Gauss-Newton algorithm (a maximum step length at each iteration is imposed to ensure positive values of the parameter Taylor approximation. Write c) as $f(x) = 0$ with $x = L_e^d$ then solving for x in $f(x)$:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

Starting with a given value of x_0 - I use the observed employment level - x is determined, iteratively. Convergence of x_0 to x is very quickly.

For the present model, I did not address the question of uniqueness of equilibrium. From Laroque (1991) and others it is, however, clear that the model under consideration has a unique equilibrium when the parameter restrictions are satisfied.

6.2.3 Identification

In Chapter 2 it was said that PML₂ estimation for the aggregate-min (or canonical) model allows the identification of the standard deviations of the common aggregate shocks of supply and demand, separately. From Laroque and Salanié (1989) it is known that PML₂ estimation for the model with micro markets additionally identifies a function of τ_d and τ_s (the micro variances), namely $\sqrt{\tau_d^2 + \tau_s^2}$. They are, however, not separately identified (in the theoretical model with only normal disturbances). I agree with Laroque and Salanié (1989, p. 846) that for the complex model under consideration, with some disturbances normal and other log-normal and interactions between the two markets, it is possible that all parameters are

identified. However, identification is possibly weak. Like Laroque and Salanié, I therefore impose *ad hoc* identification restrictions (Scheme 6.2).

Scheme 6.2 Identification restrictions

| | Labour market | Goods market |
|----------------------------|-----------------------------|--|
| <i>supply disturbances</i> | $\tau_L^s = k_1 \sigma_L^s$ | $\tau_Q^s = k_2 \sigma_Q^s$ |
| <i>demand disturbances</i> | $\tau_L^d = k_1 \sigma_L^d$ | $\tau_c^d = k_2 \sigma_c^d$ $\tau_x^d = k_2 \sigma_x^d$ $\tau_m^d = k_2 \sigma_m^d$ |

where k_1 and k_2 are non-negative constants estimated simultaneously with σ . Through the parameters k_1 and k_2 information is obtained about the size of the spatial disturbances. From this, in turn, an indication of the presence of structural mismatches on the labour market (and the goods market) may be derived. A k parameter of zero (implying τ_d and τ_s to be zero) indicates that the labour market or goods market functions as efficient aggregate markets and structural mismatches do not occur. A value of $k \geq 1$ should be interpreted as large and indicates a heterogeneous market (see Laroque and Salanié, 1989, p. 851).

It is clear, however, that through the set of identification restrictions σ_s and σ_d (the aggregate variances) are no longer separately identified as in the theoretical model in Section 2.4. So, with identification restrictions the MCPML₂ method compares to the analytically-based C.E.S method, which also does not identify the aggregate variances separately.

6.3 Estimating the model

6.3.1 Data

The model is estimated for the Dutch economy as a whole using annual data. The sample period is 1950-1991. The employment and capital stock series are for the enterprise sector only. Most data used in this study are obtained from the files of the Central Planning Office. They are in turn constructions of the data of Statistics Netherlands. Table 6.1 reports the details on measurement of the model variables.

Table 6.1 Glossary of variables¹*Endogenous variables*

- Q : Gross national product at market prices, billions of guilders at 1980 prices.
 C : Private consumption, billions of guilders at 1980 prices.
 X : Exports of commodities excluding energy, billions of guilders at 1980 prices.
 M : Total imports, billions of guilders at 1980 prices.
 L : Employment enterprise sector (millions of workers).
 \bar{R}^D : Households disposable income, billions of guilders.

Exogenous (predetermined) variables

- C_{-1} : Lagged private consumption, billions of guilders at 1980 prices.
 L_{-1} : Lagged employment enterprise sector (persons).
 K : Gross capital stock (structures, equipment and transport) enterprise sector at 1980 prices.
 p : Gross national product deflator.
 p_c : Household consumption deflator.
 p_x : price of exports exclusive energy.
 p_{ex} : price of competing OECD exports exclusive energy, doubly reweighted.
 w : Gross wage rate per head in the enterprise sector at current prices.
 t' : Average rate of social contributions supported by employers as a fraction of gross wages in enterprise sector.
 t'' : Average rate of social security contributions paid by the employees and direct taxes on wage income (actual receipts); fraction of gross wage income.
 WD : World trade (= quantity of competing exports excluding energy, doubly reweighted).
 \bar{D} : Exogenous part of demand for goods at 1980 prices (calculated as a rest category from the income identity).
 P : Labour force (in persons), defined as the sum of the level of employment in the enterprise sector and the level of unemployment.
 t : time, 1 in 1950 etc.

1. Here, as anywhere else in the book, one billion means one thousand million.

As the overall data period, 1950 to 1991, contains two changes of base, the data are converted to a common base. For this a correction factor is computed from the two observations available in the year of base year change. The capital stock series are obtained directly from Statistics Netherlands. The capital stock per sector data were easily matched with the enterprise classification of the Central Planning Office (CPB, 1992: Appendix V).

6.3.2 *The practice of optimising*

The optimisation was carried out on a Pentium 100 Mhz machine. The estimation of the model using the MCPML₂ method turned out to be very computationally demanding. The optimisation algorithm uses the Davidon-Fletcher-Powell (DFP) algorithm with analytical first derivatives. The analytical first derivatives were checked numerically, which by the way, did prove to be a necessary exercise. The DFP algorithm employs the identity matrix as an initial Hessian or an initial matrix set by the user. At each iteration, a new approximation to the inverse Hessian is computed by the DFP updating process. From Belsley (1980), it is known that the process requires a minimum of p iterations to obtain a reliable approximation of the Hessian, with p is largely equal to the number of parameters (=26 in the present context). Given the experiences of Laroque and Salanié (1989, 1990), I considered a number of 50 drawings (G) of each of the disturbances vectors ε and η in the Monte Carlo part a necessary minimum. The computer burden is already considerable at a number of $G = 50$, corroborating my observation in the previous chapter. Starting values were obtained using 20 drawings. I accepted the starting values of the estimation run with the largest function value and when:

- a) the function value did not change anymore, and
- b) the parameter estimates were almost insensitive to another set of disturbance terms and to a somewhat larger number of drawings.

During this process of different estimation runs, a number of local maxima occurred. In the ultimate estimation of the model it turned out that at relatively good starting values, each iteration took about 5 minutes (real time, not computer time!). As convergence required 240 iterations the estimation of the model took about 20 hours.

Preceding the final run in the practical chore of optimising the non-linear pseudo function, I encountered several problems. As multiple local maxima were found it became clear very quickly that the routine demands quite good starting values of the parameters. In a first experiment I used the parameter estimates of other studies (Laroque and Salanié, 1989 and Kooiman and Kloek, 1985) as starting values. This proved to be unsuccessful: the optimising routine was very sensitive to the starting values, so local maxima could easily occur. Having set G at 20 the value

of the pseudo likelihood function became too large and the routine immediately stopped as division by zero occurred.

Obviously, a finely tuned procedure was necessary to find an acceptable set of starting values. First, I did OLS estimation on each of the equations using observed values for the endogenous variables and some dummy for the latent variables. These parameter estimates subsequently served as input for the model estimation with G set at a manageable level of 20. There were several runs with different sets of disturbances ε and η , and the parameter estimates of these runs combined yielded the set of good starting values for the runs with $G = 50$.

I got an impression about the behaviour of the optimising function from the various runs with $G = 20$, an impression was obtained. In the course of the process I found several problems and some practical solutions were made. Still it was not at all clear what caused the problems. Finally it became evident that they were the result of relatively bad starting values. I will discuss two such problems. First, in the early experiments negative values of the latent variables (which by definition should be positive) occurred. After some iterations all variables were negative for one specific t or more, so that the pseudo likelihood function value cannot be computed. The solution consisted of neglecting negative values in calculating the function value. This procedure worked quite well, as all variables were positive from a certain number of iterations on.⁹⁰ Second, the routine could force the parameters to move outside its theoretical boundaries. When this happened a penalty function became effective, stopping the routine from searching further outside these boundaries: in effect the function value was artificially set at a very large value to discourage the search direction. Whenever one or more parameter values came very close to the boundary value the search direction changed conditionally: if the search was directed outside the parameter space, then the search continued parallel to the boundary. The final parameter estimates are no boundary solutions.

Once good starting values were found, I started the runs with $G = 50$. To detect the possible influence of the random disturbances on the estimates, I did experiments with two different sets of random terms. Moreover it was necessary to investigate the sensitivity for outliers on the parameter estimates. The model performance for the year 1982 was relatively weak. Therefore, the model was also estimated without this outlier, *i.e.* the pseudo likelihood function value for this year takes a value zero. It turned out that the results for the different runs are not fundamentally different. Therefore, I will present the estimation results for the complete sample 1950-1991. In the estimation runs with $G = 50$, there were no serious problems of spurious maxima (a result that is in accordance with the MCPML₂ experiences of Laroque and Salanié, 1990). As a matter of fact, I only found one local maximum which furthermore had very unrealistic properties.

⁹⁰ In the final runs with $G = 50$, this problem did not occur even a single time. Evidently, the approach to start with a smaller number of $G = 20$, has proved its value.

A final point to be made is that the parameter estimates of the run with $G = 20$ used for the starting values for the run with $G = 50$ and the final parameter estimates (for $G = 50$) are quite close in most of the cases. From this one could conjecture that a higher number of replications for most of the parameters is probably not necessary. Given, the findings of Chapter 5, where a number of 400 seemed no exaggeration, this is a quite relieve.

6.3.3 *Estimation results*

This section summarises the results from the estimation of the model. First, I will discuss the parameter estimates. Subsequently, I will consider the calculation of the standard errors of the parameter estimates, model performance and the implications for the measurement of structural mismatches and structural unemployment.

Parameter estimates

The results are presented in the Tables 6.2A and 6.2B. In the Tables I included the results of Laroque and Salanié (1989) for France to serve as reference. The models are not completely identical, so for some parameters no reference point is available.

Production function and labour demand

The exponent of labour, α_0 , should reflect the share of wages in total output. The estimate of 0.58 comes quite close to the actual figure. As such, it quite favourably contrasts with that of others who predicted too low estimates (Mehta and Sneessens, 1990 and Laroque and Salanié, 1989). The parameter estimate of α_1 reflects a rather small rate of exogenous technological change of almost 0.3% a year. Other studies also indicate small values. They are, however, somewhat larger: the relevant percentages are 1 for France, for Belgium it is between 0,7% and 2% (see Laroque, 1991 and Sneessens, 1981 respectively). Table 6.2A shows a very low value for λ_1 of 0.03, indicating a mean lag of adjustment towards the long run target of almost 31 years, which is unrealistically long. Laroque and Salanié (1989) find similar results for France; Laroque (1991) finds a value for λ_1 which is not significantly different from zero for the Unites States. It is conjectured that profit maximisation with a Cobb-Douglas production function is not the proper specification for the data, besides a underutilisation of the capital stock is not accounted for. In this respect, I noted that from iteration 210 on, the dependence between α_0 and λ_1 was strongly negative, which could indicate a weak identification. Laroque (1991), using identical technology equations, comes to a similar interpretation (p.24). The spillover effect from the goods market on labour demand as reflected in ζ is substantial and estimated higher than for France. This result corroborates the importance of modelling spillovers at the micro level.

From Table 6.2B it follows that the standard deviation of the common aggregate shock, σ_Q^s is small and resembles the value found in Laroque and Salanić (1989).

Table 6.2A MCPML₂ estimation results for the Netherlands (sample period 1950-1991) and France^{1,2}

| Parameters | Estimates | |
|-----------------|---------------|----------------|
| | Dutch economy | French economy |
| α_0 | 0.58142 | 0.473 |
| α_1 | 0.00026 | . |
| α_2 | 1.99350 | - 3.117 |
| λ_0 | 0.41891 | . |
| λ_1 | 0.03142 | 0.023 |
| ζ | 0.29929 | 0.161 |
| μ_c | 0.50246 | . |
| χ_1 | 0.91001 | 0.832 |
| ε_1 | 1.04780 | 1.089 |
| ε_2 | 0.58116 | 0.640 |
| ε_0 | 0.50634 | 3.128 |
| l_2 | - 0.33419 | - 0.037 |
| l_3 | 0.01120 | 0.010 |
| l_0 | - 1.07860 | - 2.488 |
| β_0 | 0.00970 | 0.101 |
| β_1 | 0.00066 | . |
| ρ_c | 0.29804 | 0.389 |
| ρ_x | 0.12565 | - 0.236 |

1. As estimated by Laroque and Salanić (1989) using quarterly data for the period 63II-81IV.
2. Value of pseudo likelihood function: 50.553, D.W. with respect to: $Q = 0.82$; $C = 1.09$; $M=1.08$; $L = 0.62$.

Table 6.2B MCPML₂ estimates of the standard deviations of the disturbances

| Parameters | Estimates | |
|-----------------------|---------------|----------------|
| | Dutch economy | French economy |
| σ_Q^s | 0.00858 | 0.0104 |
| σ_L^d | 0.01237 | 0.0027 |
| σ_m^d | 0.80139 | . |
| σ_x^d | 1.72170 | 1100.3 |
| σ_c^d | 1.16300 | 1729.8 |
| σ_L^s | 0.00783 | 28.0 |
| k_1 (labour market) | 1.06030 | 0.071 |
| k_2 (goods market) | 1.31440 | 1.249 |

Labour supply

The β_0 estimate equals 0.0097. This is higher than the value found in Kooiman and Kloek (1985) for the Netherlands, 1951-1979. As β_1 is not identified in their model, they fixed it at 0.00045. Remarkably, this value deviates only little from my estimate, which is 0.00066. The estimate of β_1 implies a rise of just over 1.5% from 1968 to 1991. This result differs from that in Chapter 4, where search unemployment remained relatively stable at a very low level. As far as this level of search unemployment is concerned the estimates differ: the value of β_0 reflects a rate of search unemployment of about 2%, whereas in Chapter 4, the rate of search unemployment remained clearly under 1%. In Chapter 8, I will develop some reasoning explaining this result. Compared to France, my estimate of β_0 is even modest. The value found by Laroque and Salanié (1989) is about 0.1. This result could reflect the problems they met in estimating the labour supply function. They reported, for instance, a higher number of spurious maxima in this respect. For Germany and the United Kingdom the estimates of β_0 are consistent with mine (see Laroque, 1991).

Finally, I note that, as in Kooiman and Kloek (1985), I find a low value for σ_L^s , indicating an almost perfect fit. It is not strange that the figure for France is much higher. This is because it should be related to a far more higher level of labour supply.

Household consumption and external trade

The consumption function estimates point to a propensity to consume real disposable income of 0.91, which is almost similar to the French estimate. The parameter of the autoregressive term μ_c is about 0.5 reflecting a certain smoothness in consumption (compare footnote 87). The distribution parameter ρ_c is almost 0.3 (for France 0.4), implying that in situations of excess goods demand, consumption bears 30% of the adjustment.

The estimation results for the export and import equations are consistent with those of other studies for the Netherlands. The elasticity of exports demand, ϵ_1 , with respect to world trade is about one. The price elasticity with respect to exports demand is -0.58 and almost equal for France. The distribution parameter ρ_x is 0.13. Import demand reacts in a similar way on price changes as do export demand on exports prices ($\iota_2 = -0.33$). The distribution parameter ρ_m (which follows from the identity $\rho_c + \rho_x + \rho_m = 1$) is quite large (about 0.5) which means that it bears the largest part of the adjustment in situations of excess goods demand. For France a value of 0.8 can be computed; this, however, exaggerates the “true” value since ρ_x is not significantly different from zero. The comparable figures are 0.4 and 0.6 for Germany and for the United Kingdom are 0.4 and 0.6 (Laroque, 1991).

Estimating the standard errors of the estimates

The computation of standard errors of the estimates under MCPML₂ is unsettled. When the likelihood is exact, the asymptotic covariance matrix of say a parameter θ is equal to I^{-1} , with:

$$I_{i,j} = E_x E_0 \left(\frac{\partial \psi(x, \theta_0)}{\partial \theta_i} \frac{\partial \psi(x, \theta_0)}{\partial \theta_j} \right)$$

where x denotes the vector of endogenous variables and exogenous variables, θ is the vector of parameters and ψ is the likelihood function evaluated at (x, θ) .

The matrix I is not exact under PML₂ estimation and the second order derivatives are also needed. The asymptotic covariance matrix of θ then equals $J^{-1}IJ^{-1}$ (Gouriéroux *et al*, 1984, and Annex 5B), with:

$$J_{i,j} = E_x E \left(- \frac{\partial^2 \psi(x, \theta^*)}{\partial \theta_i \partial \theta_j} \right)$$

Then the following convergence in distribution applies:

$$\lim_{T \rightarrow \infty} \left(\lim_{G \rightarrow \infty} \sqrt{T} \left(\hat{\theta}_{GT} - \theta^* \right) \right) = N(0, J^{-1} I J^{-1})$$

for any sequence of approximate estimators $\hat{\theta}_{GT}$, with θ^* is the optimum. This result provides a consistent asymptotically normal estimator (see Laroque and Salanié, 1989, p. 839)

For the model at hand, MCPML₂ estimation makes use of analytical first derivatives. As a result the matrix I can be estimated quite accurately, if one assumes a sufficient number of Monte Carlo replications G . This is not so for the matrix J . As Laroque and Salanié (1991) put it:

“To compute a consistent estimator of J would have required the consistent evaluation of second derivatives which is a difficult task for these non-differentiable methods.”

In the practical work they suggest an approximate formula (p. 21). From simulation experiments with the canonical disequilibrium model they have to conclude from a Monte Carlo study that the estimated standard errors, however, underestimate the empirical standard errors (see also Chapter 5). In particular they find a large bias for MCPML₂. On page 22, they state:

“Thus it appears that the application of the standard formulae for estimating standard errors, ..., particularly to MCPML₂, tends to produce estimates of the dispersion that are somewhat optimistic.”

Instead of using the “standard” formula, I use three empirical methods to calculate 95% confidence intervals. In combination they hopefully contain information about the significance of the parameter estimates. I find some confidence for this in Artus, Laroque and Michel (1984, their footnote 6) who conjecture that most results in asymptotic theory apply to a more general class of models.

Still, the results here should be used with caution. More satisfactory estimates would come from a Monte Carlo study on the present model (similar to Chapter 5 for the simple labour market model). This should be the subject of future research.

Theory says that when the likelihood is exact (and some other requirements such as exact normality of the parameter estimates are satisfied), the inverse of the Hessian as computed from the DFP optimising routine equals the asymptotic covariance matrix (see Gouriéroux, Monfort and Trognon, 1984). Then $J^{-1} I J^{-1} = I^{-1} = J^{-1}$. Assuming exact normality of the parameter estimates, a 95% confidence interval can be constructed as the set of parameter values that would not be rejected if tested as the null hypothesis. This interval can be computed from:

$$\psi(x, \theta^* + A) = \psi(x, \theta^*) - 3.84$$

where θ^* is in the optimum and A is a constant which can be computed numerically. The value of 3.84 is equal to $(1.96)^2$. As with the non exact pseudo function, it is unknown how useful these estimates are. I only hope that the combined results give some information on the significance of the parameter estimates. Table 6.3 presents the results based on J^{-1} (which is the inverse of the Hessian from the DFP routine), $J^{-1} I J^{-1}$ according to the standard formula above and on the “numerical” method.

Ideally the confidence intervals should be conservative: it is better to wrongly accept than reject significance from zero. When for the three methods in the table the most conservative 95% confidence intervals are constructed, it then still follows that all parameters are different from zero, except for α_1 . The estimates based on J^{-1} and $J^{-1} I J^{-1}$ are quite the same and differ somewhat from the numerical method. The estimates for the latter method indicate quite small intervals for the first 13 parameters in the table (until parameter β_1). For the other parameters, in particular the dispersions, the confidence intervals are larger and less symmetrical.⁹¹

Model Performance

The availability of statistical tests for the type of model under consideration here has not been developed very much, if at all. Therefore, I will present some intuitive results to develop some more faith in the model’s performance. In Figure 6.1 the actual and estimated series of the four endogenous variables, Q , C , M and L are displayed. From the figure, the overall model fit seems satisfactory. For the year 1982 the fit is relatively poor for output and consumption. This year defines the turning point in economic activity. Employment is predicted relatively poorly in the year 1983. The estimated imports series seem to follow the data quite good over the entire sample. Figure 6.2. plots the disturbances of the endogenous variables. A casual glance leads one to suspect the randomness of the disturbances. Knowing that adequate test statistics have not been developed for the type of model, I computed the usual Durbin Watson Statistic (see Table 6.2A). In the absence of lagged endogenous variables, it can be shown that this statistic follows asymptotically a normal distribution with mean 2 and standard error $2/\sqrt{T}$. Of course, there are lagged endogenous variables and the significance of the test is therefore doubtful. Nevertheless, I reported the results as an illustration. If the test would be applicable its values would point to indeterminate outcomes, implying that there is no evidence of first order autocorrelation.

⁹¹ Symmetry should follow from the normal distribution.

Table 6.3 Parameter estimates and 95% confidence intervals of the parameters (3 methods)

| Parameters | Parameter estimates | 95% confidence-intervals estimated from: | | |
|--------------|---------------------|--|---------------------|-------------------------------------|
| | | Function increase of 3.84 | \hat{J}^{-1} | $\hat{J}^{-1} \hat{I} \hat{J}^{-1}$ |
| α_0 | 0.58142 | 0.57821 - 0.58471 | 0.52455 - 0.63829 | 0.53589 - 0.62695 |
| α_1 | 0.00026 | -0.00036 - 0.00080 | 0.00026 - 0.00026 | 0.00026 - 0.00026 |
| α_2 | 1.99350 | 1.97420 - 2.01140 | 1.64522 - 2.34177 | 1.71488 - 2.27212 |
| χ_1 | 0.91001 | 0.90112 - 0.91911 | 0.89031 - 0.92972 | 0.89102 - 0.92900 |
| μ_c | 0.50246 | 0.44500 - 0.56346 | 0.38724 - 0.61768 | 0.39153 - 0.61339 |
| ϵ_1 | 1.04780 | 1.04521 - 1.05054 | 1.03189 - 1.06371 | 1.03273 - 1.06288 |
| ϵ_2 | 0.58116 | 0.56849 - 0.59381 | 0.52227 - 0.64005 | 0.52619 - 0.63613 |
| ϵ_0 | 0.50634 | 0.49395 - 0.51873 | 0.45181 - 0.56087 | 0.46138 - 0.55130 |
| ι_3 | 0.01120 | 0.01092 - 0.01144 | 0.01026 - 0.01214 | 0.01032 - 0.01209 |
| ι_2 | -0.33419 | -0.34219 - -0.32681 | -0.37882 - -0.28967 | -0.37610 - -0.29228 |
| ι_0 | -1.07860 | -1.08682 - -1.07125 | -1.14621 - -1.01102 | -1.14154 - -1.01566 |
| β_0 | 0.00970 | 0.00238 - 0.01411 | 0.00404 - 0.01536 | 0.00514 - 0.01426 |
| β_1 | 0.00066 | 0.00015 - 0.00069 | 0.00024 - 0.00082 | 0.00026 - 0.00080 |
| ρ_c | 0.29804 | 0.19743 - 0.38412 | 0.25925 - 0.33683 | 0.26578 - 0.33030 |
| ρ_x | 0.12565 | -0.03265 - 0.25189 | 0.12263 - 0.12867 | 0.12312 - 0.12818 |
| ζ | 0.29929 | 0.20322 - 0.43655 | 0.25971 - 0.33887 | 0.26799 - 0.33059 |
| λ_1 | 0.03142 | 0.02433 - 0.03744 | 0.01924 - 0.04359 | 0.02221 - 0.04063 |
| λ_0 | 0.41891 | 0.34878 - 0.49172 | 0.28993 - 0.54789 | 0.33198 - 0.50584 |
| k_1 | 1.06030 | 0.60113 - 1.76412 | 0.97587 - 1.14473 | 0.99134 - 1.12926 |
| k_2 | 1.31440 | 0.98967 - 1.73422 | 1.13803 - 1.49078 | 1.18237 - 1.44643 |
| σ_Q^s | 0.00858 | 0.00590 - 0.01422 | 0.00609 - 0.01108 | 0.00754 - 0.00963 |
| σ_L^d | 0.01237 | 0.00937 - 0.01709 | 0.00937 - 0.01536 | 0.01082 - 0.01391 |
| σ_m^d | 0.80139 | 0.55510 - 1.34781 | 0.77461 - 0.82817 | 0.77800 - 0.82478 |
| σ_x^d | 1.72170 | 1.30635 - 2.43204 | 1.71141 - 1.72935 | 1.71469 - 1.72871 |
| σ_c^d | 1.16300 | 0.87591 - 1.70461 | 1.14928 - 1.17672 | 1.15015 - 1.17585 |
| σ_L^s | 0.00783 | 0.00012 - 0.03298 | -0.00747 - 0.02314 | 0.00193 - 0.01374 |

The outcomes should, however, be used with caution, as long as the worth of this statistic for the present situation has not been established.

A by-product of the Monte Carlo simulation is the number of times the computation leads to each of the four regimes. At each date there are 2500 (50x50) observations distributed over the four regimes, each observation corresponding to a specific set of disturbances ε and η . From this, I computed the regime proportions. They are displayed in Figure 6.3. In the fifties the Keynesian unemployment regime dominates, its importance is declining from the early sixties to mid seventies, then its influence again increases. The significance of classical unemployment is increasing from the late sixties/early seventies on. According to these results, one may thus say that rationing of goods demand has been an almost permanent feature of the Dutch economy between 1973 and 1986. A few periods between 1970 and 1982 were characterised by inflationary pressures on both markets.

I find the large underconsumption regime in the fifties and sixties rather suspect. It is probably the result of the very small estimate for λ_1 ; the large excess demand regime in this period, however, is not strange. A possible explanation for the low λ_1 estimate could be the result of not taking into account the underutilisation of the capital stock in the model.

Roughly, the estimated regime proportions are in line with my expectations. It would be nice to compare these findings with those of other studies to get an idea of reference. This is rather difficult, particularly because the estimated regime proportions are very country-specific and very sensitive to political measures (see Artus, Laroque and Michel, 1987, their remarks on the "Plan Giscard" of 1964). Some rough comparisons, however, can be made. The development of the Keynesian unemployment regime fits in nicely with the findings of Kooiman and Kloek (1985) for the Netherlands (period 1952-1979), although they estimated the canonical model. It also compares very well with Mehta and Sneessens (1990) for Belgium (period 1955-1985). The differences for the classical unemployment regimes are greater, although there are similarities: Kooiman and Kloek predict a substantial classical unemployment for the period 1976-1978, which is in accordance with my results. Moreover, the excess goods demand regimes are comparable. Compared to Belgium and France there are differences and similarities for the repressed inflation, underconsumption and classical unemployment regimes.

Mismatches and structural unemployment

The model estimates cannot be directly translated in a measure of u^* as defined in Chapter 4. Information about the heterogeneity of the labour market and goods market supply and demand is here reflected in the parameters k_1 and k_2 respectively. Table 6.2B shows that k_1 and k_2 are convincingly above one, which clearly indicates the presence of structural mismatches on both the labour market and the goods market. Besides, the confidence intervals as shown in Table 6.3 indicate that the

estimated values are significantly different from zero (the homogenous market case).

For France Laroque and Salanié (1989) find corresponding values of 0.071 for the labour market, and 1.249 for the goods market. From this, they conclude that the labour market in France behaves as if it were homogenous and the goods market heterogeneous.

The estimate of k points to the existence of structural problems on the Dutch labour market. However, it does not indicate how important the problem is. Therefore, I computed the rate of structural unemployment from the estimates of unobserved supply and demand. More specifically, the following formulae is used:

$$u^* = \frac{\text{Min}(L^S - L, L^D - L)}{L^S} * 100\% \quad (6.3.1)$$

When L^S equals L^D , this measure gives an estimate of the rate of structural unemployment. To reduce the sensitivity of this measure for cyclical effects, I only computed u^* for a period for which supply and demand are almost equal (see also Muysken, Bierings and De Regt, 1990). From the estimates I could distinguish two such periods: 1959-1962 and 1976-1979. For these two periods, the rate of structural unemployment may be estimated as roughly 0.8% and 0.9%, respectively. These estimates are lower than those obtained from the C.E.S. method of Chapter 4 with 1.7% and just under 3%, respectively.

6.4 Concluding remarks

In this chapter I considered the type four variant of estimating structural mismatches and structural unemployment. To this purpose, a model was formulated that took full account of endogeneity of incomplete market clearing at the micro level. Spillovers between the goods market and the labour market were explicitly modelled at the micro level. Besides, lagged variables of the endogenous variables were included in the micro specifications. The estimation method is MCPML₂. Only the second order technique allows testing for the presence of structural mismatches. The estimation of the model turned out to be very demanding, using a great deal of computer time. Moreover, it demanded some creative solutions to problems met in the course of the optimising process. The final results, however, seem to be satisfactory, although the relevance of an excess demand situation on the goods market is perhaps overestimated. There were no serious problems with respect to spurious maxima, and the estimated regime proportions seem to have reasonable values. For the computation of the standard errors of the parameters, I suggested some practical solutions, since the matter does not yet have an adequate theoretical solution. As a consequence, the presented significance levels should be used with caution. Further theoretical investigations remain necessary.

Figure 6.1 Estimated and actual values of the endogenous variables, 1961-1991

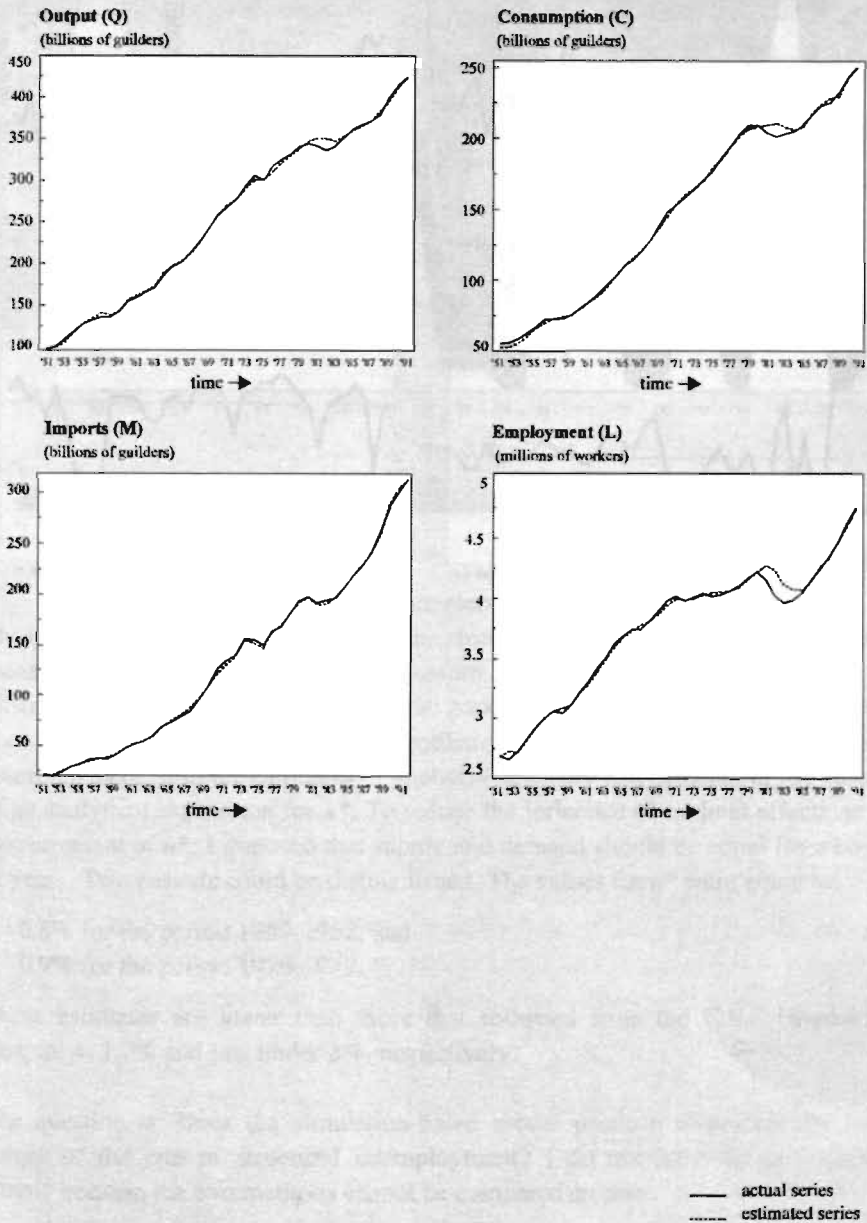


Figure 6.2 Estimated disturbances relative to the corresponding values of the endogenous variables, 1951-1991

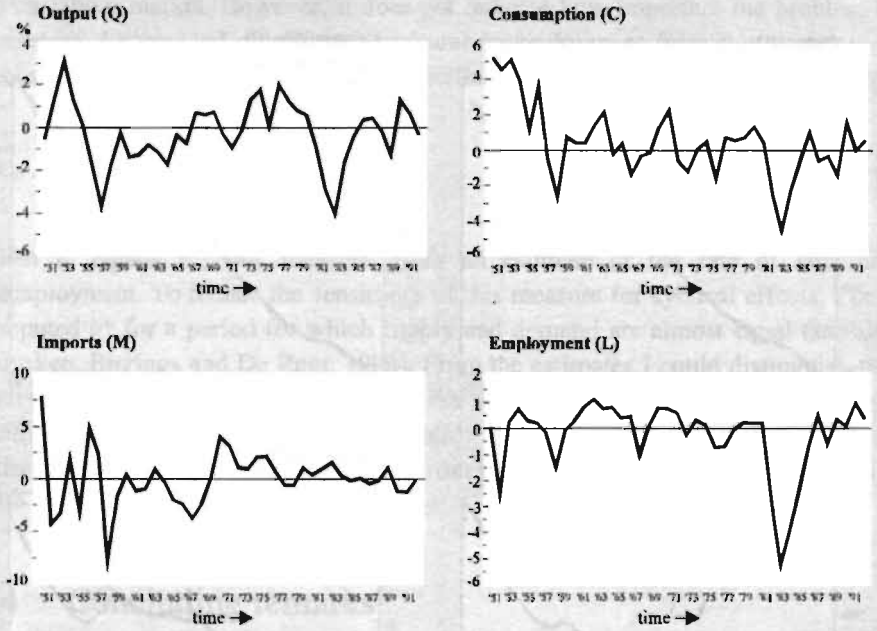
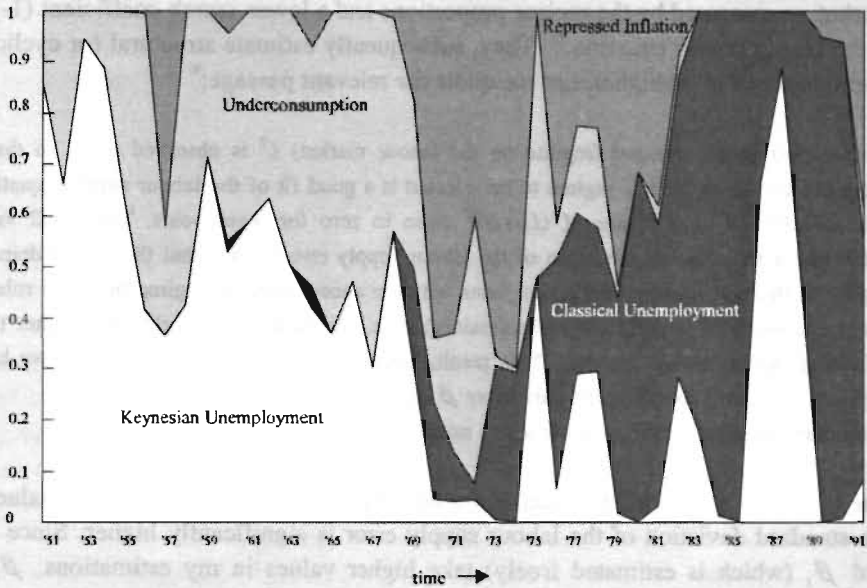


Figure 6.3 Regime proportions (cumulative), 1961-1991



The relevant mismatch parameters are identified through a set of identification restrictions. As a result, the structural mismatch indicator is a derived measure. The estimations of this measure indicate the presence of structural mismatches on the labour market (and the goods market). To develop some intuition about the importance of structural problems, I computed the rate of structural unemployment from the estimates of unobserved supply and demand, in the absence of an analytical expression for u^* . To reduce the influence of cyclical effects on the measurement of u^* , I imposed that supply and demand should be equal for a couple of years. Two periods could be distinguished. The values for u^* were equal to:

- 0.8% for the period 1959-1962; and
- 0.9% for the period 1976-1979.

These estimates are lower than those that followed from the C.E.S. method of Chapter 4: 1.7% and just under 3%, respectively.

The question is: Does the simulation-based model produce systematically lower values of the rate of structural unemployment? I do not have definite answers simply because the two methods cannot be compared on this.

However, another interesting comparison is to confront the outcomes of Chapter 6 with those of Kooiman and Kloek (1985) who estimated the canonical disequilibrium model for the Netherlands. Compared to my estimation outcomes, they estimate a very minor relevance of an excess demand regime on the labour market, as measured by the regime proportions and a lower search coefficient ($1-\beta_1$) in the labour supply equation.⁹² They, subsequently estimate structural (or cyclical) unemployment to be higher. Let me quote the relevant passage:⁹³

“Under (an) excess demand (regime on the labour market) L^s is observed as L , so that a necessary condition for this regime to be selected is a good fit of the labour supply equation. This is reflected in estimates of $(L^s-L)/L^s$ close to zero for these years. The small value obtained for the standard deviation of the labour supply error entails that the excess demand regime is strongly rejected ... This explains why the excess demand regime plays no role so that unemployment is predicted to be (mainly) of a structural or cyclical nature rather than reflecting labour market frictions. This result, however, depends on the value that we have imposed (= a priori fixed) for the parameter β_1 reflecting the increase in frictions”....
(Kooiman and Kloek, 1985, p.60, brackets mine).

The fit of the labour supply equation in the present study is also good. The value of the standard deviation of the labour supply error is significantly higher. Since β_0 and β_1 (which is estimated freely) take higher values in my estimations, β_1 is always estimated higher. As a further result, the actual level of unemployment is more easily matched by search, so that the excess demand regime on the labour market is more likely, and structural (or cyclical) unemployment plays a smaller role. This is also evidenced by Kooiman and Kloek (1985), when they fix β_1 at a higher level. By way of reference: For the relevant period in the seventies (1976-1979) for which I computed the rate of structural unemployment, about 50 percent of the markets are in the excess demand for labour regime, whereas Kooiman and Kloek reach a figure of 0%.

For the fifties and sixties, it is possible that my estimations overestimate the excess demand for labour regimes. This may be concluded from the very small estimate for λ_1 (the speed of adjustment of labour demand to the target), implying that it takes very long to reach the target. This in turn is perhaps the result of not taking into account the underutilisation of the capital stock in the production function. The lesson to be learned is that when it is indeed true that my estimations overestimate the excess labour demand regimes, structural unemployment could be underestimated. Then the estimates of structural unemployment in Chapter 4 and 6 probably come closer.

There is also another possibility that could explain the different outcomes, namely that the C.E.S. method overestimates the structural unemployment rate. This

⁹² Where $\beta_1 = \beta_0 + \beta_1 t$.

⁹³ They exclude the repressed inflation regime in the estimation.

would be the case when the aggregate labour supply and demand shocks affect the dispersion of supply and demand at the micro level. In Section 2.4 I conjectured, however, that this dependency is not very likely to occur.

7. Mismatches on the regional labour market (Micro evidence)

7.1 Introduction

In Chapter 4 and Chapter 6 the phenomenon of mismatches or structural unemployment was approached entirely from a high aggregate level. The procedure was to impose a stochastic structure on the unobserved micro level supply and demand. At the macro level, estimates of mismatches (and thus of structural unemployment) could be obtained by using time-series data. The results clearly showed the importance of mismatches on the labour market.

The strength of the approach is that in the absence of data to describe a micro market, quantitative estimates of mismatches can be obtained. A major drawback of the use of aggregation to test workers behaviour is that testing does not take place at the (individual) level at which it is formulated. There are some other reasons why micro studies are important. Mismatches reflect imbalances between the supply and demand for labour across occupational groups, regions, age groups, ethnic groups *etc.*, and combinations thereof. Therefore, micro studies could reveal important complementary information on the causes and nature of mismatches. Or as Sneessens and Mehta (1993) put it:

“The huge shifts observed in the Beveridge curve together with the concentration of unemployment in certain areas or work groups have motivated more disaggregated studies where the causes of structural problems could be better accounted for.” (Sneessens and Mehta, 1993, p. 2).

This chapter recognises the importance of studying structural problems from a more disaggregated level as it attempts to establish a direct link between mismatches and individual behaviour. To implement this idea, I had to make some practical choices. The first concerns the choice of the data source. I took the *Labour Force Survey*, LFS) of Statistics Netherlands. It is particularly suited because it provides the most comprehensive and integrated supply side information of the Dutch labour market. Because it is a household survey, however, it does not contain any information on vacancies.

Taking the LFS as the data source, the next question was how to operationalise the micro market concept empirically. In Chapter 4 and Chapter 6 a micro market is defined as a typical and *exclusive* configuration of supply and demand characteristics. In this sense, supply and demand are assigned to one (perfectly self-contained) market, thereby disallowing mobility between markets. In theory, such a classification is possible, even if the empirical implementation, however, is virtually or totally impossible. The present micro study, therefore has to permit some degree of mobility. Therefore, in this chapter a micro market may be defined loosely and indirectly as a market demanding a more than average effort to leave.

The next question was how many dimensions to consider; it is conceivable that region, skill, experience, sex, age, occupation and class of industry play a role. Since the LFS is a survey, each cell must represent a sufficiently large number of individuals to obtain statistically reliable conclusions. In a multidimensional approach the limits of what is statistically acceptable are reached very quickly, not to mention the heavy burden such an approach places on the analytical part. Therefore, I opted for a unidimensional approach, the *regional dimension*. There are three reasons for this. First, it allows statistically reliable results; ⁹⁴ second its empirical relevance: recall the results of Chapter 4 which established the correspondence between regional mismatches and structural unemployment. The third is a statistical argument. It states that the implementation of the relevant alternative dimensions, *i.e.* the occupational or industrial dimension would demand too many cells to distinguish, with too small a number of persons within a cell as a consequence.

⁹⁴ A certain degree of disaggregation into classes may influence the measurement of mismatch, a highly undesirable property. This is related to the modifiable areal unit problem described in Openshaw and Taylor (1979). Integrity requires choosing the degree of disaggregation on *a priori* grounds that justify the assumption of more or less self-contained classes. Distinguishing a too small number of, say, occupations would leave the inevitable frictions within these occupational classes unobserved, while too large a number of classes would indicate mobility where actually no above average effort is undertaken. In the case of occupations, at least 87 classes Occupational Classification 1984 of Statistics Netherlands, 1985) should be distinguished (see also Loozen and Dekker, 1993). This would leave hardly any room for distinguishing factors that explain occupational mobility. For classes of industries similar arguments apply (see Corpeleijn, 1980). Also note that the group of school-leavers (about 200 thousand) must be excluded because - per definition - they had no earlier occupation or class of industry. This further decreases the number of persons within a cell. One could counter the above argument by saying that several survey years should be taken together to increase the number of persons within a cell. In this chapter about regional mobility it is necessary to keep the one year condition, to get a conceptually acceptable measure of regional labour market mobility (see Section 7.3).

The regional classification I use here does not have this objection. Moreover it satisfies the requirement of footnote 94 that the classification should represent more or less self-contained units. The final choice to be made in order to allow empirical implementation was a redefinition of the problem. Instead of “mismatches and individual behaviour”, I will consider the mirror image of mobility behaviour of the unemployed.⁹⁵ In particular it proved necessary to delineate the problem further, by studying the mobility behaviour of unemployed workers who have recently found a job.⁹⁶ The concept of a vacancy then naturally comes in through the concept of a job entrant.

So finally, I come to the purpose of this chapter, which is to explore the factors hampering regional labour mobility of the unemployed workers in particular. Among these factors, I distinguish job characteristics, but also economic factors such as search time and relative vacancy situation. It is especially interesting to compare the influence of the search variable on mobility with the results of Chapter 4.

The rest of this chapter is organised as follows. Section 7.2 presents some facts about the structure of unemployment by region. In Section 7.3 the concepts of job entrant and regional mobility are defined and operationalised on the basis of the Labour Force Survey. Section 7.4 describes the data and the estimation method. The estimation results explaining the probability a job entrant has in being mobile across RBA-regions are presented in Section 7.5. In Section 7.6 a summary of the main findings is given, together with some concluding comments. Annex 7A presents the results of a reliability test of the data used. In Annex 7B, I present the results of an alternative model, in which mobility is treated as a function of the commuting distance.

7.2 Structure of unemployment by region: some facts

I used the construct of RBA-region to delineate the regional labour market. RBA-regions are the working areas of the employment offices. Their introduction was

⁹⁵ So, in this chapter mismatches are used as an implicit concept and are not explicitly measured. This contrasts with Chapter 4, where mismatches were measured by the variances of supply and demand across markets.

⁹⁶ In a strict sense, the analysis requires data on the regional search behaviour of the unemployed. These data are not available. This motivates my choice for the proxy category of previously unemployed workers who have recently found a job. I assume this category was looking for a job before they were accepted in the job. A statistical advantage of studying the group of persons who have recently found a job is that one does not have to rely on subjective questions, but on realised actions.

part of the decentralisation of labour market policy. The regions were composed in such a way that they represent more or less self-contained labour market areas (CBA, 1990).

Over time the pattern of unemployment rates by RBA-region is relatively stable (Table 7.1). In Layard, Nickell and Jackman (1991), this persistence of unemployment is identified as a "mismatch". To measure inadequacy of supply they use the variance of relative unemployment rates $\left(\text{var} \frac{u_r}{u} \right)$.

The unemployment rate is the number of unemployed in terms of the labour force.⁹⁷

These data yield some evidence about the significance of regional mismatches. As such they provide the necessary background for the regional mobility behaviour of previously unemployed job entrants, which is considered as indicative evidence for regional search behaviour of the unemployed. The next section is concerned with how regional labour market mobility is measured. It nicely combines the concepts of job entrant and regional mobility to define regional labour market mobility.

7.3 Job entrants and regional labour mobility: operational definitions and measurement

7.3.1 Job entrants

A job entrant is defined as a person who, at the date of interview, has been on the current job for less than 12 months and who is receiving a wage-income. Those in military service are excluded. This implies that persons who started a job in the previous year, but who are out of the job again at the date of interview, are not counted as job entrants. The self-employed and family workers are not considered job entrants in my set-up. Their reasons to accept a job in another region are quite different from those with a wage-income job. Within the framework of regional

⁹⁷ In Chapter 4, I used a somewhat different measure of dispersion, including also vacancies (see footnote 70). Moreover the measure was computed for the provinces of the Netherlands. The reason for this was only a practical one. For provinces time-series data from 1960 on were available. For RBA-regions this is not the case. This leaves unchanged the view that the RBA classification should be preferred: provinces are not composed initially to describe more or less self-contained labour market areas, as RBA-regions are (see footnote 94). The fact that vacancies are not incorporated in the mismatch indicator here fits the dominantly supply side treatment of mismatches in this chapter.

Table 7.1 Unemployment rate by RBA-region, 1981-1991 (persons aged 15-64) ¹

| RBA-region | 1981 | 1983 | 1985 | 1988 | 1989 | 1990 | 1991 |
|--|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | % | | | | | | |
| Groningen | 8.3 | 13.1 | 12.8 | 12.2 | 11.1 | 11.4 | 8.8 |
| Friesland | 6.6 | 11.8 | 11.8 | 10.0 | 10.0 | 8.0 | 8.0 |
| Drenthe | 6.5 | 12.6 | 11.9 | 9.4 | 7.7 | 7.1 | 7.2 |
| Ijssel-Vecht | 5.6 | 9.9 | 9.2 | 6.4 | 5.8 | 5.4 | 5.5 |
| Twente | 7.3 | 14.5 | 12.1 | 10.1 | 9.1 | 8.7 | 8.4 |
| Mid-Ijssel | 4.6 | 10.3 | 9.0 | 9.5 | 6.5 | 6.9 | 5.6 |
| Veluwe | 4.3 | 7.9 | 7.0 | 6.7 | 6.3 | 5.2 | 5.9 |
| Arnhem/East-Gelderland | 6.4 | 11.3 | 10.4 | 9.1 | 8.1 | 7.0 | 6.5 |
| Nijmegen Rivierenland | 7.9 | 14.1 | 10.5 | 9.5 | 9.9 | 9.3 | 7.4 |
| Flevoland | 7.1 | 10.6 | 10.7 | 7.6 | 8.1 | 7.7 | 6.4 |
| East Utrecht | 4.4 | 7.3 | 6.2 | 6.1 | 7.0 | 4.7 | 4.8 |
| West Utrecht | 5.3 | 7.0 | 7.4 | 8.6 | 6.9 | 6.9 | 5.8 |
| Het Gooi en Vechtstreek | 3.8 | 6.6 | 6.7 | 4.2 | 6.3 | 5.5 | 5.3 |
| North-Holland North | 5.6 | 8.6 | 7.5 | 7.2 | 6.4 | 6.8 | 5.9 |
| Amsterdam/Zaanstr./Waterl. | 6.5 | 12.1 | 11.5 | 11.6 | 10.1 | 9.2 | 8.8 |
| Kennemer-, Amstelland en Meerl. | 3.4 | 6.3 | 5.5 | 5.9 | 4.5 | 4.1 | 4.4 |
| Rijnstreek | 3.7 | 6.9 | 6.0 | 4.7 | 4.5 | 4.9 | 3.7 |
| The Hague/Delft | 4.8 | 7.9 | 7.3 | 7.5 | 6.9 | 7.1 | 6.3 |
| Drechtsteden | 4.3 | 8.2 | 7.2 | 6.0 | 6.5 | 4.4 | 4.9 |
| Rijnmond | 5.8 | 9.8 | 8.7 | 9.8 | 9.9 | 8.3 | 8.0 |
| Zeeland | 4.8 | 9.3 | 6.9 | 6.7 | 6.8 | 5.0 | 4.8 |
| West North-Brabant | 8.0 | 12.3 | 9.9 | 7.3 | 6.3 | 6.7 | 6.5 |
| Breda | 6.3 | 11.0 | 8.8 | 7.0 | 5.9 | 4.8 | 5.1 |
| Mid Brabant | 6.8 | 11.2 | 10.4 | 9.8 | 7.8 | 6.2 | 6.3 |
| North/East Brabant | 7.2 | 12.1 | 9.9 | 7.9 | 7.5 | 6.6 | 5.5 |
| South/East Brabant | 7.4 | 12.0 | 9.3 | 8.4 | 7.4 | 6.7 | 7.0 |
| North and Mid Limburg | 6.4 | 11.1 | 9.7 | 7.1 | 5.5 | 5.3 | 5.4 |
| South Limburg | 8.3 | 12.2 | 11.7 | 9.8 | 8.4 | 7.8 | 6.9 |
| All | 5.9 | 10.2 | 9.1 | 8.3 | 7.6 | 6.9 | 6.5 |
| var(u_t/u) | 5.6 | 6.1 | 5.8 | 5.4 | 4.9 | 5.8 | 4.2 |

1. The labour force, using the 1992 CCS definition, consists of the employed and the unemployed labour force. The unemployed labour force comprises all persons who do not work, or are working fewer than 12 hours a week and who: (i) have found work bringing the total number of weekly working hours to at least 12 hours, or (ii) declare they want to work for at least 12 hours a week, are available for at least 12 hours a week and have undertaken specific steps to find work for at least 12 hours a week. The employed labour force comprises all persons who work 12 hours or more a week.

See: Bierings, Imbens and van Bochove (1991) for the conceptual background of this definition.

Source: LFS 1981-1991 (Statistics Netherlands).

mobility this would require a separate treatment of both categories of workers. I have chosen to concentrate on those receiving a wage-income, which concerns the majority of people.

Figure 7.1 indicates how a job entrant is identified in the labour force questionnaire. Note that, the starting-date of the job is given priority in setting the definition of a job entrant. Hence, a respondent indicating that the period between the starting-date and date of interview is less than a year is always considered a job entrant, irrespective of the answer to the first question. Below I will only present the results for job entrants who accepted a job of 12 hours or more a week. This is in line with the new definition of the labour force Statistics Netherlands introduced in 1992 (see Bierings, Imbens and van Bochove, 1991). According to this definition the employed labour force consists only of people with a major labour market involvement (operationalised by the 12 hours threshold). To impose the threshold of 12 hours is also necessary to accurately measure the category of previously unemployed job entrants.

To identify the *previously unemployed* job entrants, some compromise is necessary. I would prefer to adopt the objective requirements as in the definition of unemployment (see footnote 1, Table 7.1). However, in the LFS questionnaire these requirements are not determined retrospectively (for one year ago). Instead the LFS has some subjective questions about the main labour market status of one year prior to the interview. People are asked about their main activity one year earlier (employed, school, housekeeping *etc.*). Within the category of job entrants, I defined the previously unemployed as those people who indicated they had no work before they got hired.⁹⁸ Most of them are likely to have searched for a job. For a small category of school-leavers this is not the case. They found work directly after leaving school. The inclusion of this category, however, will not significantly influence the results.

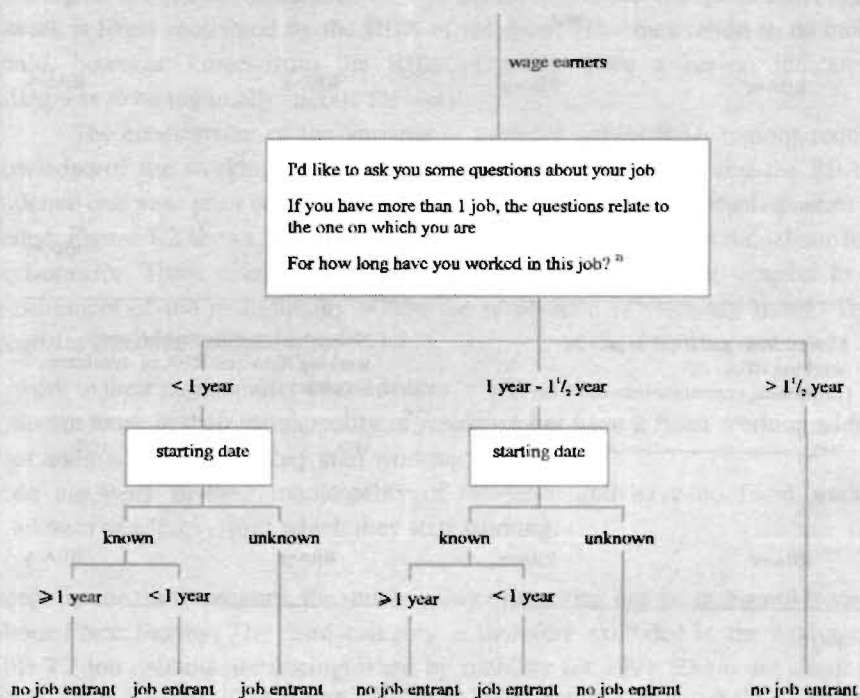
7.3.2 Regional mobility⁹⁹

As regional classification I use the classification in RBA-regions (see Figure 7.3, for a chart). Job entrants are mobile across RBA-regions if the RBA-region where they lived one year prior to the interview differs from the RBA-region where they worked at date of interview. Those who lived in a foreign country are thus excluded. Mobile job entrants encompass commuters as well as migrants. Scheme 7.1 contains the classification of mobile/immobile people.

⁹⁸ This excludes the category of persons who were employed a year ago, became unemployed and then found a job again. This category is not observed by the LFS, but is expected to be small.

⁹⁹ The definitions on job entrants and regional mobility used in this section, compare to those used in Teulings (1990).

Figure 7.1 Measurement of job entrants in the labour force questionnaire¹⁾



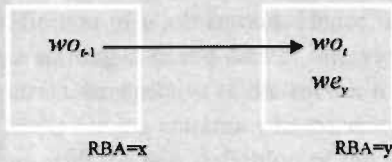
1) LFS 1991.

2) The respondent can choose between the following five categories:

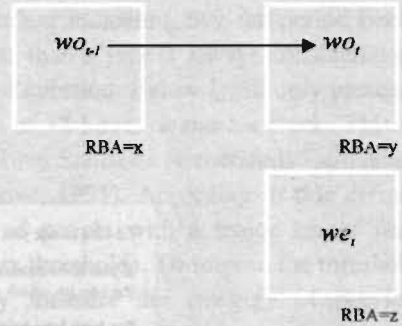
- a. less than 4 months
- b. 4 months up to 1 year
- c. 1 year
- d. somewhat more than a year up to 1 1/2 year
- e. more than 1 1/2 year

Scheme 7.1 Classification of mobile persons across RBA-regions

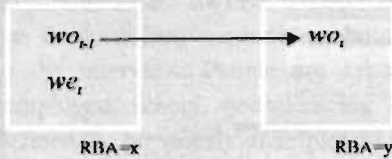
1.a Moved to present working RBA
(Mobile)



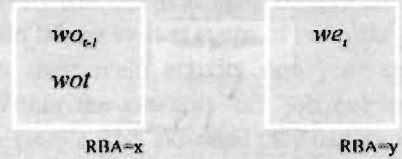
1.b Moved but not to present working RBA
(Mobile, commuter between RBA's)



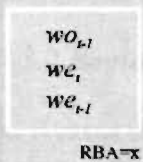
1.c Moved from and lived in present working RBA
(Immobile, commuter between RBA's)



1.d Not moved, commuting between working RBA and RBA of residence
(Mobile)



1.e Stayer



Symbols:

wo : RBA of residence
 we : working RBA
 t : year of interview
 x, y, z : arbitrary RBA

The definition contains an arbitrary element. It concerns the delineation of persons under 1.c as regionally immobile. Consider the following: a couple migrates to another region because she accepted a new job there, he accepted a new job in the region where they originally lived. In this example she is correctly counted as regionally mobile, whereas he is perhaps incorrectly counted as not mobile. Ideally, one would like to have the reason for mobility, to get a more precise measure of regional mobility. Such a question, however, is not included in the labour force

questionnaire. Therefore, I had to accept that in some - probably rare - cases, the measurement of regional mobility contains an arbitrary element. A related point is that the definition cannot fully exclude the possibility that regional mobility (read: migration) is not associated with the new job accepted (contrary to what I wanted). In the definition of mobility, I deliberately do not use the previous working RBA. This implies, for instance, that a *stayer* (case 1.e) who previously worked in another RBA-region is not considered mobile. The reason is that the change in RBA region of work is likely motivated by the RBA of residence. The motivation to be mobile should, however, come from the RBA of work. Then a person indicates a willingness to be regionally mobile for work.

The construction of the variable of mobility across RBA-regions requires knowledge of the working RBA-region at the date of interview and the RBA of residence one year prior to that date. For this the corresponding municipalities are needed. Figure 7.2 shows how mobile job entrants are identified in the labour force questionnaire. Three categories of persons are distinguished with respect to the measurement of the municipality where the respondent is currently living. These categories consist of people who:

- work in their municipality of residence;
- do not work in their municipality of residence but have a fixed working address or address from which they start working;
- do not work in their municipality of residence and have no fixed working address or address from which they start working.

Except for the third category, the municipality of working can be measured from the Labour Force Survey. The third category is therefore excluded in the analysis. In Table 7.2 job entrants are distinguished by mobility for 1991. There are about 682 thousand job entrants of whom 153 thousand were mobile across RBA-regions. This implies that just over 3 out of 4 job entrants were immobile. Among the previously unemployed (303 thousand persons) this was 4 out of 5 persons. Evidently, most jobs are found within the region of residence.¹⁰⁰ This is more pronounced for the previously unemployed than for the previously employed job entrants.

The definition of mobility across RBA-regions treats all job entrants the same way, irrespective of the commuting distance or time. This can be undesirable when the commuting distance or time hampers mobility. In such situations the asymmetric treatment of mobile persons just crossing the border to start work and immobile persons with a long distance to commute is controversial, and weighting by distance or time would be appropriate. Whether distance or time act as barriers

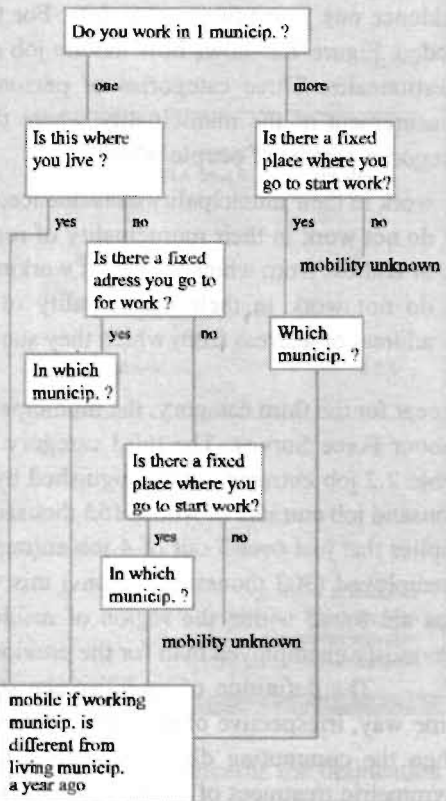
¹⁰⁰ This fits the definition of an “independent” labour market in Smart (1974).

Figure 7.2 Measurement of previous and present working municipality and regional mobility (wage earners) in the labour force questionnaire¹⁾

Previous municipality of residence



Present working municipality



1) LFS 1991

for mobility is unclear.¹⁰¹ This is not observed in the LFS. Most likely there are adverse effects: allowances for moving or travelling expenses could decrease barriers, whereas traffic-jams increase them. Table 7.3 shows the distribution of previously unemployed mobile and immobile job entrants by commuting distance. The 15 kilometre boundary is meant to distinguish people with a short distance to commute from those with a long distance to commute. A further distinction is made by area of RBA-regions. It follows that most of the previously unemployed mobile job entrants travelled more than 15 kilometres to their work (about 88%). In small RBA-regions this percentage is somewhat lower. Most immobile job entrants travelled less than 15 kilometres to their work (81%). In small RBA-regions this percentage is higher.

Table 7.2 Job entrants (15-64 year, working 12 hours a week or more) by mobility across RBA-regions (commuters and migrants), 1991

| | <i>Total</i> | <i>of whom were previously without work</i> |
|----------------------------------|--------------|---|
| | x 1 000 | |
| Job entrants ¹ | 682 | 303 |
| of whom were | | |
| • mobile across RBA's | 153 | 60 |
| • immobile | 529 | 245 |

1. Only job entrants for which the working municipality is known from the LFS are counted. When the components of a total do not add to that total this is due to rounding errors.

Source: LFS 1991 (Statistics Netherlands)

These facts indicate that the majority of mobile people travelled longer distances and the majority of immobile people shorter distances. So, at first, distance does not seem to cause a problem with the concept of mobility across RBA-regions. Or in other words, it seems unnecessary to include a distance criterion in the definition of regional mobility. This conclusion is substantiated in our logistic regressions. In Annex 7B, I present the logistic regression variant that takes a correction of mobility for commuting distance into account. The correction works as follows: mobile people with a short commuting distance have a smaller weight than those with a

¹⁰¹ There is some indicative evidence in Vos (1980) who studied the willingness of unemployed workers to accept a job outside the region where they live. He found that 70% of the unemployed who applied for a job would accept a job at a distance of more than 200 kilometres. To my opinion this (large) percentage, should be used with caution, because "willingness" is a subjective concept and does not reflect actual behaviour. This is exactly the reason why I in this chapter have chosen for an approach on actual (mobility) behaviour.

long distance and, immobile persons with a long distance to commute have a smaller weight than those with a short distance. The results for the weighted logistic regression indicate that weighting with distance hardly influences the parameter estimates.

Table 7.3 (Previously unemployed) job entrants (aged 15-64) mobile across RBA- regions by area-class of RBA-regions by commuting distance, 1991¹

| Area of RBA-regions ² | Mobile across RBA regions | | | Immobile | | |
|----------------------------------|---------------------------|--------------------------|---------|----------|--------------------------|---------|
| | Total | with commuting distance: | | Total | with commuting distance: | |
| | | <15km | ≥15 km | | <15km | ≥15 km |
| | % | | | % | | |
| Small | 100 | 14 (16) | 86 (84) | 100 | 85 (87) | 15 (13) |
| Middle | 100 | 11 (12) | 89 (88) | 100 | 81 (84) | 19 (16) |
| Large | 100 | 10 (11) | 90 (89) | 100 | 73 (77) | 27 (23) |
| Total | 100 | 12 (14) | 88 (86) | 100 | 81 (84) | 19 (16) |

1) Between brackets are the results for all job entrants.

2) The area class "small", "middle", "large" contain all RBA-regions with area <750 km, 750-1500 km, >1500 km respectively.

Source: LFS, 1991 (Statistics Netherlands)

7.4 Data and estimation method

Data

The data set used to examine barriers of regional labour mobility is the continuous Dutch LFS of 1991 (conducted by Statistics Netherlands). It is a stratified random sample of about 12,000 addresses per month. The stratification is by labour market position, region and municipality. The LFS sample contains an annual of 2.1% of all addresses in the Netherlands. The response is about 60%. All members of a household are covered and their family ties, sex, age, marital status, and nationality are obtained. Questions about the labour market position are submitted to all persons over 15.¹⁰²

¹⁰² For more details on data collection and processing see van Bastelaer (1988).

Information on vacancies is obtained from the job vacancy survey of Statistics Netherlands (and not! from the registers of the job centres, compare Chapter 4). All information is combined in one data file and per individual; the matching criterion for the vacancy information which is per RBA-region is the RBA-region.

The following variables are introduced to control for worker characteristics. There are dummies that capture gender (*MALE*) and ethnic origin (*NATIVE*). Their expected sign is not obvious, but their inclusion will answer interesting questions. Does mobility differ between men and women or between workers from different ethnic origins? I also included the household composition (*SINGLE*, *HEAD WITHOUT CHILD*, *HEAD WITH CHILD*, *SPOUSE WITHOUT CHILD*, *SPOUSE WITH CHILD* and *CHILD*). The likely outcome is that having dependants, especially children, makes it more important to find a job at home. Four dummies are included to capture differences in educational attainments: workers with primary education (*PRIM*), workers with lower secondary general and vocational education (*LOWSEC*), workers with higher secondary general and vocational education (*HIGHSEC*), workers with higher vocational education (*HIGHVOC*), and workers with university education (*UNIV*). It is expected that educational level is positively related to mobility. I also use an age variable. Intuitively, the age profile with respect to mobility is concave - mobility increases with age over some interval, but starts to decline at the older ages. In this respect I introduce *AGE* and *AGE*².

I will also use information on job characteristics, such as working hours (*HOURS*) and flexibility of job relation (*FLEX*). The effects of these variables on mobility are not obvious, but are likely to be determined by the same type of arguments, Why be mobile for a minor job?; or is it the only way to get a job. Flexible labour comprises such categories as temporary employees, seasonal labour, homeworkers, etc. There are seven occupational dummies: Professional, Technical and Related Workers (*PROF*), Administrative and Managerial Workers (*ADMIN*), Clerical and Related Workers (*CLERICAL*), Sales Workers (*SALES*), Service Workers (*SERVICE*), Agriculture, Animal Husbandry and Forestry Workers, Fishermen and Hunters (*AGRIC*) Production and Related Workers, Transport Equipment Operators and Labourers (*PRODUCTION*). As regional dummy variable I use *REGION*. It contains the RBA-regions of residence 12 months before date of interview.

Finally, two economic indicators are included. First, there is a dummy that measures the **period without work before acceptance of the current job**; it proxies the search time before being accepted in the current job. When it is less than 12 months it is indicated by *PERIOD*. The effect of *PERIOD* on mobility is not obvious. Here, I recall the finding in Chapter 4 on the long-term unemployed who are not part of effective labour supply, thereby contributing maximally to structural unemployment (see eq. 4.4.9). This result is consistent with other findings indicating that the proportion of unemployed people who leave unemployment within a given time period is much lower for those who have been unemployed for

long periods (see Jackman and Layard, 1990). Whether this is the case in a regional context is not certain. Gorter, Nijkamp and Rietveld (1993) find that the willingness to move decreases unemployment duration, although the effect is not significant. This could mean that people aware of this effect, start searching in other regions as well just after becoming unemployed. Vissers, de Vries and Schepens (1986), however, find that people's willingness to move (or commute) for a job does not increase when they are offered (a fictitious) higher wage or unemployment benefit in the new region. I did not include a variable measuring relative wages in my regressions. The differences are too small across regions for a job searcher to base his or her mobility decision on. The second economic indicator measures the relative labour market situation between domestic and working region before job acceptance. It is the relative vacancy situation, indicated by *VACINDEX*. It takes into account that most people are unwilling to move or travel over a long distance to their work and would prefer to take a job close to home. So, the likelihood of job seekers in region r seeking work outside the region will be primarily influenced by the availability of jobs in region r rather than by vacancies in the country as a whole. I allow for the discouraging effect of a distance on job search by the discount factor δ_{rs} . The vacancy index then shows:

$$VACINDEX = \frac{\sum_{r=1}^R V_r}{\sum_{r \neq s}^R \delta_{rs} V_s} \quad (7.1)$$

$$\text{with: } \delta_{rs} = \frac{1}{1 + d_{rs}^2}$$

d = distance between the centre of RBA-region r and RBA-region s .

V_r = vacancies in RBA-region r .

The parameter δ_{rs} indicates that the vacancy-index decreases quadratically if the distance of a RBA-region from the RBA-region of residence increases.

The exact dating of the vacancy information used for the decision to become mobile in the period under consideration is unknown. I took a weighted average of the vacancy situation in 1989/1990 (factors are 9/32 and 23/32, respectively) somewhat arbitrarily. I checked the robustness by comparing the results based on other weighting schemes. It turned out that the results are not significantly influenced by them.

There is some indicative evidence of the influence of vacancies on mobility in Gorter (1991). He finds that the relative number of vacancies decreases the "regional inflow intensity".

Estimation technique

I applied logistic regression to explain regional mobility. The functional form of this probability, $P(Y=1)$ is:

$$p_i = P(Y_i = 1 | x_i) = \frac{e^{x_i' \omega}}{1 + e^{x_i' \omega}} \quad i=1, \dots, n \quad (7.2)$$

where $\{Y_i\}$ is a sequence of independent random variables taking the values 0 or 1; x_i is a k-vector of known constants; ω is a k-vector. Logistic regression is applied on bias-corrected sample outcomes. In fact each individual in the LFS is weighted such that the sample distributions of sex, age, and marital status correspond to the population distributions. In this way the bias resulting from selective non-response is corrected for. I used the software program SUDAAN to estimate the parameters and variances of the model. This program can handle the complex nature of the design of the LFS (stratification *etc.*) by using the Taylor series linearisation method of variance estimation. Standard statistical packages as SPSS and STATA are not suited.

7.5 Results ¹⁰³

Here I can present the results of four logit regressions variants. The logit variants are summarised in Scheme 7.2. The four variants are combinations of two different samples on the one hand and two different regional indicators on the other hand. The two sample categories are the previously unemployed job entrants and all job entrants (reference group). Comparison of the results with the group of all job entrants may reveal some interesting information about how a previously unemployed job entrant estimated his/her chances of a job offer outside the resident region. A lower probability (relative to the probability of previously employed workers) may indicate that an unemployed worker expects his chances of getting a job in another region to be small and will therefore search for a job mainly within the resident region (all else being equal). There may be good reasons to expect that employers will prefer to offer jobs to a previously employed worker in their own region and not so easily to a previously unemployed worker from another region.

The regional indicators are the vacancy index or the typical RBA-region. I could not include both regional indicators in one regression, because of their strong dependency. To be able to classify a typical RBA-region on the mobility scale and to determine the effect of vacancies, I did separate regressions.

¹⁰³ Only, the results for 1991 are presented. I also estimated the models for 1989 and 1990. There were no significant differences with 1991. It is just to say that the results I present for 1991 have a wider applicability to the years 1989 and 1990.

Scheme 7.2 Logit variants

| <i>Sample</i> | <i>Regional indicator</i> | |
|---|---------------------------|------------|
| | Vacancy index | RBA-region |
| Job entrants previously unemployed | 1 | 3 |
| All job entrants | 2 | 4 |

Variant 1 and 2 (with vacancy index)

Table 7.4 presents the results of the logit regression with the vacancy index. The equation appears satisfactory in terms of overall performance and in terms of t-statistics on individual explanatory variables. Besides, the parameters have the expected signs. Significant signs are the same for the previously unemployed and the reference group of all job entrants. Remarkably, the direction of the effects hardly differs between the previously employed and the previously unemployed job entrants. There is, however, a difference in overall mobility performance. The mobility probability of the previously unemployed job entrants and of all job entrants can be computed from the estimation results. It follows that the previously unemployed job entrants are less mobile across regions than previously employed job entrants. This in turn seems to indicate that the previously unemployed may perceive their chances of getting a job elsewhere as relatively small (one could speak of regional discouragement or self-selection).

The age effects are not significant for the previously unemployed job entrants. The results show a concave age-mobility pattern, with a peak at age 36 for those who had a job already. The hours variable is significant. The smaller the number of hours worked a week, the smaller the probability of being mobile. In this respect, persons who changed their job did not behave differently from those who were previously unemployed. It seems as if people are mobile when it pays. There is quite a strong and significant vacancy effect. Unemployed job entrants as well as employed workers living in regions with a low number of vacancies are more mobile. There is no significant gender effect for the unemployed job entrants. Among those who changed job, men were more mobile across RBA-regions than women. There are no significant effects from ethnicity, type of job relation, (flexible

Table 7.4 Logit results of the probability of job entrants (aged 15-64, working 12 hours) to become mobile across RBA-regions, 1991 ¹

| Variable | Parameter estimates ² | |
|----------------------------|-----------------------------------|------------------------------|
| | Previously unemployed (variant 1) | All job entrants (variant 2) |
| CONSTANT | -1.739 (-1.6) | -3.887 ** (-5.2) |
| Continues variables | | |
| AGE | 0.058 (0.1) | 0.108 ** (2.5) |
| AGE ² | -0.001 (-0.1) | -0.002 ** (-2.4) |
| HOURS | 0.015 * (1.8) | 0.020 ** (3.1) |
| VACINDEX | -0.490 ** (-4.4) | -0.541 ** (-8.0) |
| Dummy variables | | |
| MALE | 0.156 (0.9) | 0.244 * (2.1) |
| NATIVE | 0.076 (0.3) | -0.058 (-0.4) |
| FLEX | -0.146 (-0.9) | -0.121 (-1.1) |
| PERIOD | -0.012 (-0.0) | -0.058 (-0.6) |
| PRIM | - | - |
| LOWSEC | -0.061 (-0.2) | 0.178 (1.0) |
| HIGHSEC | 0.648 ** (2.4) | 0.523 ** (3.1) |
| HIGHVOC | 1.126 ** (3.6) | 0.949 ** (4.9) |
| UNIV | 1.269 ** (3.6) | 1.299 ** (5.8) |
| SINGLE | - | - |
| HEAD WITHOUT CHILD | 0.015 (0.0) | 0.104 (0.7) |
| HEAD WITH CHILD | -0.199 (-0.6) | -0.146 (-0.9) |
| SPOUSE WITHOUT CHILD | -0.070 (-0.3) | 0.205 (1.4) |
| SPOUSE WITH CHILD | -0.573 * (-1.8) | -0.423 * (-2.0) |
| CHILD | -0.387 * (-1.7) | -0.082 (-0.6) |
| PROF | 0.864 ** (3.6) | 0.903 ** (6.6) |
| ADMIN | 1.416 * (2.1) | 0.922 ** (3.8) |
| CLERICAL | 0.202 (0.9) | 0.142 (1.1) |
| SALE | 0.031 (0.1) | 0.097 (0.6) |
| SERVICE | 0.081 (0.3) | 0.187 (1.2) |
| AGRIC | 0.686 * (2.0) | 0.497 * (2.2) |
| PRODUCTION | - | - |
| Sample size | 2598 | 6152 |
| Loglikelihood | -1162 | -2981 |
| Multiple-R ² | 0.12 | 0.13 |

1. See Scheme 7.1.

2. t-values in parentheses. *, ** show a 5% and 1% significance level.

Source: LFS, 1991 (Statistics Netherlands).

or not) and the search period on mobility behaviour. This is rather revealing. Among the (unemployed) job entrants, those with a long search time were not less inclined to accept a job elsewhere than those with a short search time. One issue should not be overlooked. The variable *PERIOD* could be a bad proxy for search time, especially for women. Women change labour market position more frequently than men during a life-time. Many women interrupt their careers, for instance, to take care of the children or to run the household, and are outside the labour force for a time, so for them *PERIOD* may be a bad proxy for search time. For men *PERIOD* is likely to be a good proxy. I estimated the model for men and women separately to find out what the gender effect was, if any. There was no such effect, so I decided to present the results for men and women together. There may, however, be some compensating effects. Among the previously unemployment job entrants those who were at school one year earlier were relatively more mobile than the other group. This category of school-leavers had relative short search periods - on average about 4 months (Berkhout, Loozen and van der Valk, 1993). For the "other" category the rule applies: the longer the search time the less mobility across regions. It is, clear, however, that the evidence of Chapter 4 indicating a maximum contribution of the long-term unemployed to structural unemployment, is not corroborated in a regional context.

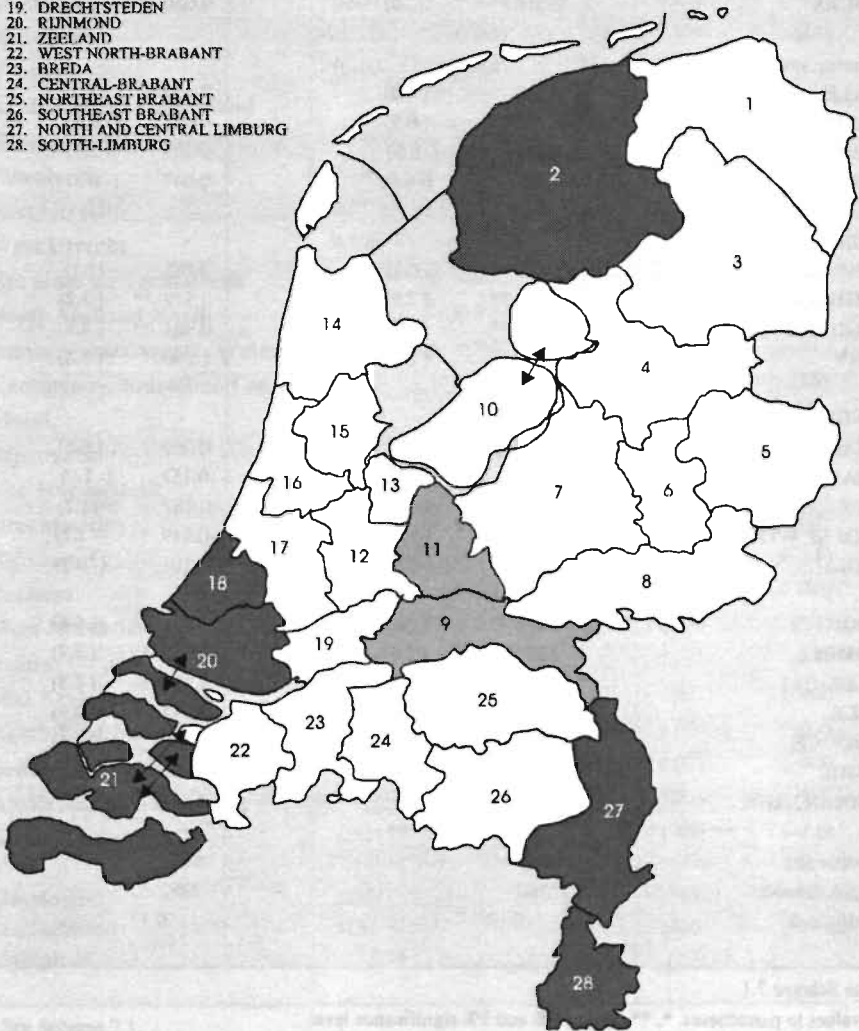
There is no ethnicity effect. Native job entrants are not more or less mobile across regions than foreign job entrants. A flexible (temporary) job in another region is not more or less attractive than a regular job. The professional and the educational variables created some large and highly significant coefficients. A low educational level is associated with low mobility and a "highly-valued" profession with high mobility. As expected, children in a household negatively influenced mobility, though the level of significance differs. Especially spouses with children compare relatively badly on the mobility-scale. They are significantly more likely to work in the region of residence.

Variant 3 and 4 (with RBA-region)

Second, I considered the results based on logit model with RBA-region. I introduced this variant to find out which RBA-regions can be characterised as typically mobile/immobile. To this purpose, I excluded the vacancy index from the regressions, since it reflects the same information contained in the RBA-regions. Table 7.5 presents the results of the logit estimations. The non-RBA-parameter estimates compare fairly well to those of Table 7.4, which is a prerequisite. On the scale of mobility Amsterdam *etc.*, The Hague, Rijnmond, Zeeland and South-Limburg score low and Nijmegen *etc.*, East-Utrecht high (see also Figure 7.3). They constitute a large part of the regional mismatches.

Figure 7.3 The 28 RBA-regions

1. GRONINGEN
2. FRIESLAND
3. DRENTHE
4. IJSSEL-VECHT
5. TWENTE
6. CENTRAL-IJSSEL
7. VELUWE
8. ARNHEM/EAST-GELDERLAND
9. NIJMEGEN/RIVIERENLAND
10. FLAVOLAND
11. EAST-UTRECHT
12. WEST-UTRECHT
13. THE GOOI AND VECHTSTREEK
14. NORTH-HOLLAND NORTH
15. AMSTERDAM/ZAANSTREEK/WATERLAND
16. KENNEMER AMSTEL-LAND EN MEERLANDEN
17. RIJNSTREEK
18. THE HAGUE DELFT
19. DRECHTSTEDEN
20. RIJNMOND
21. ZEEI AND
22. WEST-NORTH-BRABANT
23. BREDA
24. CENTRAL-BRABANT
25. NORTHEAST BRABANT
26. SOUTHEAST BRABANT
27. NORTH AND CENTRAL LIMBURG
28. SOUTH-LIMBURG



high mobility regions
 low mobility regions

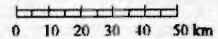


Table 7.5 Logit results of the probability of job entrants (aged 15-64, working 12 hours) to become mobile across RBA-regions, 1991¹

| Variable | Parameter estimates ² | |
|----------------------------|-----------------------------------|------------------------------|
| | Previously unemployed (variant 1) | All job entrants (variant 2) |
| CONSTANT | - 2.201 * (- 2.0) | -4.722 ** (- 6.0) |
| Continues variables | | |
| AGE | 0.057 (0.1) | 0.116 ** (2.7) |
| AGE ² | - 0.001 (- 0.2) | -0.002 ** (- 2.6) |
| HOURS | 0.016 * (2.0) | 0.020 ** (3.2) |
| Dummy variables | | |
| MALE | 0.169 (1.0) | 0.264 * (2.2) |
| NATIVE | 0.164 (0.7) | -0.012 (- 0.1) |
| FLEX | - 0.153 (- 0.9) | -0.123 (- 1.1) |
| PERIOD | - 0.004 (- 0.0) | -0.047 (- 0.5) |
| PRIM | - | - |
| LOWSEC | - 0.019 (- 0.1) | 0.183 (1.1) |
| HIGHSEC | 0.705 ** (2.6) | 0.529 ** (3.2) |
| HIGHVOC | 1.191 ** (3.7) | 0.961 ** (4.9) |
| UNIV | 1.269 ** (3.5) | 1.290 ** (5.8) |
| SINGLE | - | - |
| HEAD WITHOUT CHILD | - 0.003 (0.0) | 0.069 (0.5) |
| HEAD WITH CHILD | - 0.170 (- 0.5) | -0.153 (- 1.0) |
| SPOUSE WITHOUT CHILD | - 0.092 (- 0.4) | 0.187 (1.2) |
| SPOUSE WITH CHILD | - 0.517 * (-1.6) | -0.449 * (- 2.1) |
| CHILD | - 0.379 * (-1.6) | -0.101 (- 0.7) |
| PROF | 0.836 ** (3.4) | 0.924 ** (6.6) |
| ADMIN | 1.335 * (1.8) | 0.963 ** (3.8) |
| CLERICAL | 0.217 (0.9) | 0.176 (1.3) |
| SALE | 0.048 (0.2) | 0.144 (0.9) |
| SERVICE | 0.080 (0.3) | 0.216 (1.4) |
| AGRIC | 0.698 * (1.9) | 0.562 * (2.4) |
| PRODUCTION | - | - |
| Sample size | 2598 | 6152 |
| Loglikelihood | - 1162 | - 2981 |
| Multiple-R ² | 0.12 | 0.13 |

1. See Scheme 7.1.

2. t-values in parentheses. *, ** show a 5% and 1% significance level.

Source: LFS, 1991 (Statistics Netherlands).

Table 7.5 Logit results of the probability of job entrants (aged 15-64, working 12 hours) to become mobile across RBA-regions, 1991¹ (continued)

| Variable | Parameter estimates ² | | | |
|---------------------------------|-----------------------------------|---------|------------------------------|---------|
| | Previously unemployed (variant 1) | | All job entrants (variant 2) | |
| Groningen | - 0.706 | (- 1.5) | - 0.267 * | (- 1.7) |
| Friesland | - 1.016 * | (- 2.3) | - 0.148 | (- 0.5) |
| Drenthe | - 0.625 | (- 1.3) | 0.045 | (0.2) |
| Ijssel-Vecht | - 0.215 | (- 0.5) | 0.330 | (1.1) |
| Twente | - 0.657 | (- 1.1) | - 0.282 | (- 0.9) |
| Mid-Ijssel | -0.053 | (- 0.1) | 0.309 | (1.0) |
| Veluwe | - 0.160 | (- 0.4) | 0.322 | (1.3) |
| Arnhem/East-Gelderland | - | - | - | - |
| Nijmegen Rivierenland | 0.548 * | (1.6) | 0.547 * | (2.1) |
| Flevoland | 0.366 | (0.8) | 0.622 * | (2.0) |
| East Utrecht | 0.533 | (1.2) | 0.701 ** | (2.6) |
| West Utrecht | 0.156 | (0.4) | 0.405 | (1.5) |
| Het Gooi en Vechtstreek | 0.234 | (0.4) | 0.638 * | (1.7) |
| North-Holland North | - 0.560 | (- 1.3) | 0.030 | (0.1) |
| Amsterdam/Zaanstr./Waterl. | - 0.231 | (- 0.6) | - 0.192 | (- 0.8) |
| Kennemer-, Amstelland en Meerl. | -0.352 | (- 0.8) | 0.427 | (1.6) |
| Rijnstreek | 0.075 | (0.2) | 0.584 * | (2.4) |
| The Hague/Delft | - 0.884 * | (- 2.1) | - 0.727 * | (- 2.5) |
| Drechtsteden | - 0.342 | (- 0.7) | 0.518 * | (1.8) |
| Rijnmond | - 1.217 ** | (- 3.0) | - 0.703 ** | (- 2.6) |
| Zeeland | - 1.238 * | (- 2.1) | - 0.874 * | (- 2.3) |
| West North-Brabant | - 0.275 | (- 0.6) | - 0.434 | (1.4) |
| Breda | 0.034 | (0.1) | 0.181 | (0.6) |
| Mid Brabant | - 0.477 * | (- 1.0) | - 0.071 | (- 0.2) |
| North/East Brabant | - 0.393 | (- 0.8) | - 0.056 | (- 0.2) |
| South/East Brabant | - 0.346 | (- 0.8) | - 0.071 | (- 0.3) |
| North and Mid Limburg | - 0.704 | (- 1.5) | - 0.703 * | (- 2.4) |
| South Limburg | - 1.624 ** | (- 3.1) | - 1.402 ** | (- 4.0) |
| Sample size | 2598 | | 6152 | |
| Loglikelihood | - 1134 | | - 2925 | |
| Multiple-R ² | 0.14 | | 0.15 | |

1. See Scheme 7.1.

2. t-values in parentheses. *, ** show a 5% and 1% significance level.

Source: LFS, 1991 (Statistics Netherlands).

7.6 Concluding remarks

Contrary to the previous chapters, the phenomenon of mismatch was studied from a micro perspective. In particular, I considered mismatches at the regional level. I started with the observation that regional differences in unemployment are quite persistent and that the regional contribution to mismatches already evidenced in Chapter 4 is corroborated when some kind of variance indicator is used to describe mismatches. Then the question was what caused these mismatches. For reasons of empirical implementation, I had to approach the phenomenon of mismatches from its mirror image, that is from the perspective of mobility behaviour. Factors decreasing mobility have an increasing effect on mismatches.

I used cross-section data from the Labour Force Survey of Statistics Netherlands for 1991 to this end. It was noted that information on regional search behaviour of the unemployed is necessary to study the causes of regional mismatches. Data on this, however, fail. Therefore, I used the category of job entrants previously without work as a proxy category. I assumed that the regional mobility behaviour of this category would reflect the regional search behaviour of the unemployed.

One remarkable finding is that for job entrants there is no conclusive relationship between a weak labour market position and mobility. On the one hand, job entrants with highly valued professions or higher education are more mobile across regions. On the other hand, increasing age, a flexible job, being a foreigner, a long search period, or the characteristics commonly associated with a weak labour market position, had no influence on mobility and thus on mismatch.¹⁰⁴

The effect of search period, strictly cannot be compared with that of Chapter 4, which indicated a great contribution of the long-term unemployed to structural unemployment. This is because in my sample the long-term unemployed who did not find a job at all are not included. Since, I do not know the size of this group, it is unwise to compare the outcomes.¹⁰⁵

The strong influence of the relative vacancy position on regional labour mobility is striking. Hence, a high number of vacancies in a region decreases regional mismatches. On average people move (or commute) from a region with a small number of vacancies to a region with a large number of vacancies. The ranking by region showed that regions such as Amsterdam, The Hague, Rijnmond, Zeeland and South Limburg score low and Nijmegen, East Utrecht high.

¹⁰⁴ Of course, people who accepted a flexible (temporary) job were not always forced to do so because of a weak labour market position. Most people, however, prefer a regular job.

¹⁰⁵ In a very specific study of Kloosterman (1987) it was found that about 80% of the long-term unemployed are **not** prepared to migrate for a job. This percentage is almost the same as the percentage of previously unemployed immobile job entrants in this study. The results in Kloosterman, however, apply to a very specific situation and do not have a wider applicability.

The micro approach presented here established the effect on regional mobility behaviour in the face of a relatively bad excess demand situation. These types of effects cannot be examined within the analytically-based macro approach, and they cannot be made visible in the simulation-based approach. They are simply not identified. Consider in this respect the micro model with intertemporal spillovers in Lambert (1988, Appendix B) which I already mentioned in Chapter 4:

$$\ln L_t^d = \Lambda_t^D + \varepsilon_t^d - \gamma \left((\ln L_{t-1}^d - \ln L_{t-1}^s) - (\Lambda_{t-1}^D - \Lambda_{t-1}^S) \right)$$

$$\ln L_t^s = \Lambda_t^S + \eta_t^s + \theta \left((\ln L_{t-1}^d - \ln L_{t-1}^s) - (\Lambda_{t-1}^D - \Lambda_{t-1}^S) \right)$$

$$\ln L = \text{Min}(\ln L_t^d, \ln L_t^s)$$

where symbols have the familiar interpretation. Λ^S and Λ^D represent the “average” values of $\ln L^s$ and $\ln L^d$, and are positive constants. On the labour market $\gamma > 0$ implies that employers are tempted to recruit people who are less qualified, instead of engaging in the long and costly search for the best qualified personnel; $\theta > 0$ represents the mobility of workers moving to more prosperous micro markets. Imposing $\gamma + \theta < 1$, the aggregate transactions specification which can be derived on this micro model does not identify the parameters γ and θ in the empirical analysis. As such it is observationally equivalent to the model without intertemporal spillovers. In the micro study the relative vacancy index is incorporated as an explanatory variable for regional mobility and had a high and significant influence. Thus a high vacancy level in a certain region decreases regional mismatches. This points to a $\theta > 0$. In this respect the micro study is complementary to the macro approach in which θ (and γ) are not identified.

ANNEX 7A Plausibility of mobility and job entrants data (Labour force survey)

7A.1 Introduction

The population statistics of Statistics Netherlands observe total internal migration. Total internal migration is defined as the total number of migration moves within the Netherlands across municipal borders. The population statistics are an important source for determining the quality of the LFS data on internal migration and in turn of the quality of the labour force data on regional mobility and job entrants. In this annex the differences between the two statistics are indicated, and the causes of measurement error are analysed (Section 7A.2). To this purpose, I use data from the LFS 1990 and the population statistics 1989/1990. The outcomes of this comparison are representative for 1991. The implication for the quality of regional mobility of persons who have accepted a new job is considered in Section 7A.3. Section 7A.4 compares the number of persons who accepted a new job according to the LFS 1990 (the “new” LFS) with those of the “Arbeidskrachtentelling” (AKT= the “old” LFS) of 1985 and with vacancy data from the vacancy survey of Statistics Netherlands.

7A.2 Quality of data on internal migration and causes of measurement error

In the LFS everyone is asked if they have moved in the previous year. If so, people are asked whether they stayed in the same municipality or moved to another (see Figure 7.2). These answers are used to measure regional mobility of job entrants: if a person moved in the previous year and his/her current working region differs from the one he/she lived one year prior to the interview, then this person is classified as regionally mobile.

Total internal migration, *i.e.* the total of moves across municipal borders in the Netherlands, was 373 thousands persons (aged 15-64) according to the LFS of 1990. According to the populations statistics, 471 thousands persons (aged 15-64) moved (two-yearly average 1989/1990). This implies that the LFS underestimated the number of people who moved by 98 thousand.

From table 7A.1 it follows that 21% of all persons who moved in 1990 was not observed in the LFS. Given the age category, the percentage underestimation is almost the same for men and women. In absolute terms it is the largest in the age categories 15-19 and 20-24 years, *i.e.* 15 and 33 thousand, respectively. In a relative sense, this is also the case: in the age category 15-24 year 25% is missing, whereas in the other age categories this is 18%. Percentages hardly differ by sex.

Table 7A.1 Internal migration by age and sex according the population statistics 1990 and the LFS 1990 (age: 15-64)

| Age | Population statistics 1989/1990 | | | LFS % of population statistics | | |
|--------------|------------------------------------|---------|-------|-----------------------------------|---------|-------|
| | Men x 1 000 | Females | Total | Men % | Females | Total |
| 15-19 | 24 | 31 | 55 | 73 | 71 | 72 |
| 20-24 | 59 | 75 | 134 | 77 | 75 | 76 |
| 25-29 | 56 | 48 | 105 | 91 | 84 | 88 |
| 30-34 | 33 | 27 | 60 | 84 | 83 | 83 |
| 35-39 | 21 | 16 | 38 | 81 | 89 | 84 |
| 40-44 | 16 | 12 | 28 | 67 | 66 | 67 |
| 45-49 | 9 | 8 | 17 | 80 | 76 | 78 |
| 50-54 | 7 | 6 | 13 | 78 | 94 | 85 |
| 55-59 | 6 | 5 | 11 | 72 | 68 | 70 |
| 60-64 | 5 | 5 | 10 | 69 | 74 | 71 |
| Total | 237 | 234 | 471 | 80 | 78 | 79 |

Source: LFS, 1991 and Population Statistics (Statistics Netherlands).

There are three possible causes of underestimation:

- a) Underrepresentation of newly built dwellings in the labour force sample;
- b) Multiple migration moves;
- c) Underestimation of non-resident students.

Let me go through these points.

The **underrepresentation** of newly built dwellings in the labour force **sample** arises from the fact that the sample has been drawn some months before it is actually set out in the field. For example, the sample for the first half of 1990 was drawn from the address data source of August 1989. For the second half of 1990 the sample was drawn from the March 1990 data source. This explains the underestimation of newly built dwellings. Using data on dwellings completed in

1989 and 1990, an estimate of the magnitude of underestimation on average of 64 thousand dwellings can be calculated. Because not all newly built dwellings are immediately occupied the number of 64 thousand is the maximum estimate.

If I assume that a migration move to a newly built dwelling involves two persons in the age category 15-64 year, the underestimation in persons amounts to 130 thousand. From the population statistics it follows that from the 1.65 million persons who moved in 1989, 596 thousand crossed municipal borders. The comparable figure in 1990 is 566 thousand out of 1.55 million persons. On average 36.3% of all persons who moved crossed the municipal border in 1989/1990. When it is additionally assumed that of all persons who moved and were not captured by the LFS, 36.3% crossed municipality borders as well, implies that the LFS underestimates the number of domestic migrants (aged 15-64) by about 47 thousand.

The number of **multiple moves** can be estimated from the monthly pattern of the number of moves one year prior to the interview. It turns out that the number of persons who moved 1 month before they were interviewed is larger than 12 months earlier. The reason is that persons who moved several times in the year before they were interviewed only indicate the most recent move, since the questionnaire only allows them to mention one move. Assuming that the number of moves 1, 2, 3, .. 12 months before the date of interview is uniformly distributed, the number of multiple moves not measured by the LFS can be estimated at about 8 thousand persons at most. From these 8 thousand about 3 thousand crossed a municipal border.

The **underrepresentation of non-resident students** in the labour force can also be roughly estimated. As indicated earlier, 25% of domestic migrants in the age category 15-24 year are not observed, whereas it is 18% for the other age categories. This indicates that a specific group is not observed, probably students living in lodgings. Experiences in observing such students are negative (compare the "Woningbehoefteonderzoek" 1985/1986). If I assume 18% to be the "normal" underestimation, then the number of unobserved moves in the LFS of students (aged 15-24) living in lodgings can be estimated as $25\% - 18\% = 7\%$ or about 15 thousand persons.

From Table 7A.2 it follows that the extent of underestimation unexplained comprises 33 thousand domestic migrants. One can, however, imagine a selectivity in the non-response of people who have moved shortly before interviewing should have taken place since they are probably less inclined to participate.

Table 7A.2 Overview of the differences in internal migration moves between the LFS 1990 and the population statistics 1989/1990 (age: 15-64)

| | x 1 000 |
|--|---------|
| Internal migration moves population statistics | 471 |
| Internal migration LFS | 373 |
| Difference | 98 |
| originating from: | |
| • Underrepresentation of newly built dwellings in the labour force sample | 47 |
| • Multiple moves | 3 |
| • Underrepresentation of non-resident students | 15 |
| • Unexplained | 33 |

Source: LFS, 1990 (Statistics Netherlands)

7A.3 Consequences of underestimation for the measurement of mobile job entrants

In the LFS 63% of all persons who moved had a paid job. The percentage of workers among persons who did not move was also 63. This implies that the underestimation of moves in the LFS does not influence the measurement of the number of workers (and wage earners).

There is, however, some influence on the measurement of the number of job entrants. In the LFS 1990 the percentage of job entrants among the migrants was about 15 percent points higher than among the non-migrants. Assuming that the percentage of job entrants who moved to newly built dwellings and to already existed dwellings is the same, the number of unobserved job entrants in the LFS can be estimated at 15 thousand persons ($0.15 \cdot 98\ 000$) (= 1.5% of all job entrants).

The number of job entrants mobile across RBA regions is also underestimated. The number of migrant-job entrants who were mobile across RBA regions was 37 thousand. About 21% of all moves were unobserved (see above). Assuming that this percentage is the same for all persons who moved across RBA regions, this would imply that actually 47 thousand job entrants moved across RBA regions. This again implies that the number of mobile job entrants measured in the LFS is underestimated by 10 thousand persons (6% of all mobile job entrants).

7A.4 Reliability of job entrants data

In this section, the reliability of job entrants data (LFS 1990) is considered by comparison with other sources. I use the LFS 1985 and the job vacancy survey of the Statistics Netherlands in this respect. In Teulings (1990) the number of job entrants per province was published. A job entrant was defined as a wage earner of 15-64 years old whose main activity was paid work, and who worked at least 15 hours a week at the interview date, and who did not work or worked in another profession a year prior to the interview. In order to compare the LFS 1985 and the LFS 1990 data I had to count the number of job entrants in the LFS 1990 using this definition. The results are presented in Table 7A.3. In the table, job entrants are distinguished by province, and shows the close correspondence of the percentage distribution of job entrants by province. The number of job entrants however, differs substantially between both sources: in the LFS 1990 the number of people who find a job is 290 thousand higher. Part of the difference can be explained from the rise in the number of wage earners in the period 1984-1985 as compared to the period 1989-1990. For the period 1984-1985 the rise of the number of wage earners (who worked at least 15 hours a week) is estimated at 63 thousand persons. For the period 1989-1990 the corresponding estimate is 135 thousand persons a difference of 72 thousand persons. The other part can be explained from the fact that people have become more mobile. A higher employment rate implies the creation of new jobs. These jobs are partly filled by persons who were previously unemployed and by persons who changed job. So, a higher employment rate implies more mobility. The number of job entrants in 1990 was about 70 % points higher than in 1985.

The number of filled vacancies in 1989/1990 is another source of comparison. From the **vacancy survey** data are available on the number of filled vacancies. The sample excludes public services, schools, temporary employment agencies and supervised work for the disabled. To allow comparison I excluded job entrants in economic activity class 9 (*i.e.* "other services"). I specifically compared the number of job entrants not in class 9 working at least 20 hours a week (average 1989/1990) with the number of filled vacancies for at least 20 hours a week (average 1989/1990).

The number of filled vacancies is estimated as follows. From the job vacancy survey it follows that the number of filled vacancies for wage earners, not in class 9 amounts to 535 thousand. Moreover, about 83% of all vacancies refers to jobs for wage earners working 20 or more hours a week. From this I estimate the number of filled vacancies (of 20 hours and more a week, not in class 9) at 444 thousand ($0.83 \cdot 535$ thousand), and the number of job entrants 539 thousand. So there is a difference of 95 thousand. There are three qualitative explanations. First, vacancies for temporary employment agencies are not included. This, however, can only explain a very small part of the difference. Second, persons are offered jobs for which no vacancy has been announced. Third, it is possible that the LFS 1990 also counts persons who only change position within the same firm as job entrants. To

conclude, on the basis of the vacancy data, the number does not pose questions with respect to the reliability of job entrants data in the LFS 1990.

Table 7A.3 Job entrants (aged 15-64) by province of residence one year prior to the interview ¹

| <i>Province of residence 1 year prior to the date of interview</i> | <i>LFS 1985</i> | <i>LFS 1990</i> | <i>LFS 1985</i> | <i>LFS 1990</i> | <i>percentage of growth LFS 1990 in relation to LFS 1985</i> |
|--|-----------------|-----------------|-----------------|-----------------|--|
| | x 1 000 | | % | | |
| Groningen | 14 | 22 | 3 | 3 | 57 |
| Friesland | 16 | 24 | 4 | 4 | 71 |
| Drenthe | 10 | 16 | 2 | 2 | 60 |
| Overijssel | 29 | 42 | 7 | 6 | 45 |
| Flevoland | . | 11 | . | 2 | . |
| Gelderland | 44 | 83 | 11 | 12 | 89 |
| Utrecht | 31 | 58 | 7 | 8 | 87 |
| North-Holland | 74 | 123 | 18 | 17 | 58 |
| South-Holland | 92 | 149 | 22 | 21 | 62 |
| Zeeland | 8 | 15 | 2 | 2 | 88 |
| North-Brabant | 62 | 105 | 15 | 15 | 69 |
| Limburg | 29 | 46 | 7 | 7 | 59 |
| Total | 413 | 704 | 100 | 100 | 70 |

1. A '.' indicates the absence of sufficient reliable data (it concerns figures below 5 000).

Source: LFS, 1990 (Statistics Netherlands)

ANNEX 7B Job entrants and mobility across RBA-regions by commuting distance and other characteristics

In this annex I will present the results of the logistic regression variant that uses a mobility measure corrected for commuting distance. The weighted case estimation of the logit model uses sample data which are weighted by commuting distance. Weighting is such that a mobile job entrant with a commuting distance of 15-30 kilometres, 30-45 kilometres, or 45 kilometres, respectively becomes a weighting factor that is three times, four times, or five times as large as of a mobile job entrant with a commuting distance of < 15 kilometres. For immobile job entrants I used 1/3, 1/4 and 1/5. Additionally imposing that weighting may not influence the total number of mobile and immobile job entrants yields the factors indicated in Table 7B.1.

Table 7B.1 Weighting scheme of sample data by commuting distance of job entrants (15-64 year), 1991

| <i>Commuting distance</i> | <i>Job entrants mobile across RBA-regions</i> | <i>Job entrants immobile</i> |
|---------------------------|---|------------------------------|
| < 15 kilometres | 0.27 | 1.10 |
| 15-30 kilometres | 1.06 | 0.36 |
| 30-45 kilometres | 1.42 | 0.27 |
| ≥ 45 kilometres | 1.77 | 0.22 |

I obtained the information on the commuting distance (not available in the LFS) between municipalities from the private company ORTEC. The commuting distance between municipality of residence and municipality of work is measured by the postal codes of the centre of each municipality. The centre of a municipality is determined on the basis of the population density. The distance data are coupled to our data file on an individual basis. The matching criterion is the municipality of work and residence.

The results for the weighted logistic regression estimation are presented in Table 7B.2. There is hardly a weighting effect on the parameter estimates as can be easily verified.

Table 7B.2 Logit results of the probability of job entrants (aged 15-64, working 12≥ hours) to become mobile across RBA-regions, weighting by commuting distance, 1991^{1,2,3}

| Variable | Parameter estimates ² | | |
|-----------------------------|-----------------------------------|---------|------------------------------|
| | Previously unemployed (variant 1) | | All job entrants (variant 2) |
| CONSTANT | -2.105 | (- 1.9) | -4.586 ** (- 5.8) |
| Continuous variables | | | |
| AGE | 0.049 | (0.2) | 0.124 ** (2.7) |
| AGE ² | -0.001 | (- 0.1) | - 0.002 ** (- 2.7) |
| HOURS | 0.016 * | (1.8) | 0.023 ** (3.5) |
| VACINDEX | -0.371 ** | (- 3.2) | - 0.409 ** (- 5.9) |
| Dummy variables | | | |
| MALE | 0.301 * | (1.7) | 0.365 * (3.0) |
| NATIVE | 0.055 | (0.2) | - 0.100 (- 0.6) |
| FLEX | -0.196 | (- 1.2) | - 0.125 (- 1.1) |
| PERIOD | -0.037 | (- 0.2) | - 0.058 (- 0.5) |
| PRIM | - | - | - |
| LOWSEC | -0.136 | (- 0.4) | 0.189 (1.1) |
| HIGHSEC | 0.947 ** | (3.3) | 0.690 ** (4.0) |
| HIGHVOC | 1.495 ** | (4.5) | 1.248 ** (6.1) |
| UNIV | 1.569 ** | (4.1) | 1.523 ** (6.5) |
| SINGLE | - | - | - |
| HEAD WITHOUT CHILD | 0.017 | (0.1) | 0.064 (0.4) |
| HEAD WITH CHILD | -0.149 | (- 0.4) | - 0.076 (- 0.5) |
| SPOUSE WITHOUT CHILD | 0.049 | (0.2) | 0.228 (1.5) |
| SPOUSE WITH CHILD | -0.866 ** | (- 2.5) | - 0.584 ** (- 2.6) |
| CHILD | -0.509 * | (- 2.1) | - 0.154 (- 1.0) |
| PROF | 0.977 ** | (4.1) | 0.956 ** (6.7) |
| ADMIN | 1.457 * | (2.1) | 1.085 ** (4.0) |
| CLERICAL | 0.246 | (1.0) | 0.053 (0.4) |
| SALE | 0.123 | (0.4) | 0.081 (0.5) |
| SERVICE | 0.178 | (0.7) | 0.255 (1.6) |
| AGRIC | 0.790 * | (2.2) | 0.518 * (2.2) |
| PRODUCTION | - | - | - |
| Sample size | 2598 | | 6152 |
| Loglikelihood | - 1100 | | - 2870 |
| Multiple-R ² | 0.17 | | 0.17 |

1. For the mobile persons who migrated to another region, the migration distance has been taken.

2. For an explanation of the variants see Scheme 7.1 (main text).

3. t-values in parentheses. *, ** show a 5% and 1% significance level.

Source: LFS, 1991 (Statistics Netherlands).

8 Summary and conclusions

This thesis examines the problem of structural mismatches on the labour market in relation to the measurement of structural unemployment. With the empirical contributions presented in this study, I seek to answer two fundamental questions:

1. What is the significance of mismatches and structural unemployment on the Dutch labour market, and
2. What is the nature of these mismatches or what causes them?

To this purpose, I concentrated on two macro approaches of markets in disequilibrium and a micro approach.

Why these three methods? First of all, to get a complete picture of the problem of mismatches in the Netherlands. Inevitably, this is achieved as a result of the complementarity of the methods: restrictive assumptions applying to the one method do not apply to the other, and *vice versa*. As a consequence, the combined outcomes are less dependent on the arbitrariness in the assumptions of the individual methods. Secondly, in as far the three methods share the measurement of some common effects, these can be compared.

The Chapters 1 and 2 put the present study in a proper theoretical perspective. Objectives are set out and the state of the art studies are discussed. The two macro approaches used in this study originate from the theoretical works of Barro and Grossman (1976), Muellbauer (1978) and Malinvaud (1980). Their models are in turn inspired by Keynes (1936). The common message is that prices and wages adjust too slowly to clear markets permanently in the short run, so that transactions are determined by supply or demand, while the other side of the market gets rationed. Another key feature of these disequilibrium models is that the effects of incomplete market clearing are not exclusively confined to one single market. If workers or firms can not realise their plans in, say, the goods market, they will change behaviour on the labour market and *vice versa*. Both the assumptions of fixed prices and spillover effects are at the core of this thesis. In the application of the two macro methods in this study, prices are assumed to be exogenously fixed. As such, the recent theoretical contributions which succeeded in bringing in endogenous prices in the disequilibrium analysis are not taken into account. I accept this as a shortcoming of the present study.

The two macro approaches which are used in this study, essentially are empirical model applications of the theoretical Keynesian disequilibrium model. Central is the assumption that aggregate markets such as the labour and goods market of an economy consist of a large number of micro markets on which supply and demand ratios differ. The minimum condition is applied to these micro markets, that is transactions are determined by the minimum of supply and demand. Consequently, either unemployment or vacancies exist in a specific market. Aggregation smoothes the “sharp edges” off the micro markets min-condition to get macro outcomes which are continuous in aggregate supply and demand. In this type of model, aggregate transactions are always smaller than the minimum of aggregate supply and demand, implying that unemployment (U) and job vacancies (V) are positive simultaneously. This feature implicitly defines the level of structural (or equilibrium) unemployment. It is equal to the unemployment level at $U = V$.

The first macro approach used in this study is an analytically-based variant. From this variant analytical expressions for aggregate transactions are obtained through integration over density functions that describe the distribution of supply and demand ratios across markets. The resulting aggregate specification is a continuous non-linear relationship which maps aggregate supply and demand onto aggregate transactions (employment or output). An approximate form of this aggregate specification is of the C.E.S. type. In the empirical application this variant takes it that supply and demand can be observed from the data. A clear advantage of the method is that it leads to an aggregate transactions specification which is readily interpretable and intuitively reasonable: the mismatch indicator is simply a parameter in the aggregate transactions function and the link with structural unemployment is easily established. Endogenising the mismatches is straightforward.

There are also problematic parts which underlie the method. Here I will mention two of them. The first relates to the assumed observability of labour supply and demand. Even when vacancy (and unemployment) data is improved from the early official register data and considered for empirical application, the “good quality of data” assumption still constitutes a critical element. An extreme example of the negative consequences of using official data applies to the situation of Switzerland, where the official figures indicate that unemployment hardly ever exceeded 1%, despite very volatile output and employment. It is not surprisingly that econometric models have difficulties in explaining developments over a longer period of time for Switzerland. From this Stalder (1989) concludes that structural modelling with unobserved and endogenous supply and demand should be preferred to a data-based approach as a device to explain these developments. When improved data on supply and demand are used, the method is probably less vulnerable, but the results should still be interpreted carefully, especially while the measurement of job vacancies is complicated, and the improved data are obtained from a **statistical** correction procedure applied to the official series. The second problematic part of

the method relates to the assumptions imposed to maintain the simple C.E.S. structure. These assumptions for instance lead to a less rigorous treatment of goods market spillovers. Full account of spillovers at the micro level would imply that analytical relationships cannot be derived (compare Kooiman, 1984). It is unclear how these assumptions influence the measurement of mismatches and structural unemployment. In the simulation-based method spillovers can be modelled rigorously at the micro level.

The second approach used in this study is based on simulation. Essentially this is an estimation technique which allows a full formalisation of a macro model with endogenous unobserved supply and demand at the micro market level. The key feature here is that spillovers are fully taken into account at the micro level. Inevitably, the method is very general as it hardly requires simplifying restrictions (contrary to the analytically - based method). It can *a priori* fit any distribution (log-normal or others) of micro supply and demand, which makes it powerful. Moreover, the simulation-based method allows the identification of the variances of the aggregate supply and demand variables in theory. These aggregate quantities operate as common factors on micro market supply and demand. In the analytically-based methods, these variances are not separately identified. On the other hand, however, the method requires a great deal of performance testing. Furthermore, the technique is neither user-friendly, nor easily adapted to new model specifications, while the numerical optimisation is a time consuming exercise. Moreover, it is impossible to test the assumptions formally for these type of methods. The simulation-based method is not sensitive to measurement errors in supply and demand as the data-based method.

The third approach used in this study entails a micro study that focuses on the factors hampering (regional) mobility, the reverse side of (regional) mismatch. It seeks to explain which personal characteristics and economic variables influence individual mobility behaviour. Since the micro study makes it possible to examine the mobility behaviour of workers looking for a job in the face of a relative excess demand situation on their market, it is complementary to the analytically-based macro approach. The latter method cannot measure such an effect. I will come back on this below.

The empirical implementation of the analytically-based and simulation-based models are described in Chapter 4 and 6. Chapter 3 concentrated on the feasibility of the distribution assumptions underlying the models. The results corroborated the legitimate use of the log-normal and normal distributions to describe markets in disequilibrium. From this I decided to stay with common practice by adopting the log-normal distribution for the analytically-based variant and the normal for the simulation-based variant (the normal distribution is, however, not a necessary condition, since the simulation method can fit any distribution; in the larger model of Chapter 6 some disturbances are also log-normal). The choice of the log-normal distribution has been in favour of the Weibull and Pareto distributions, although the latter was initially regarded as superior in a dynamic

sense from the viewpoint of the theory of stochastic processes. I have shown that the three transactions functions from the Weibull, lognormal and Pareto distributions give rise to aggregate transactions functions which are nested, thereby allowing nested testing. In all this it should be kept in mind that the tests are specific. To obtain more conclusive results more work is needed.

In Chapter 4, an extension of the C.E.S. transactions function is suggested. The generalisation which is of the two-level form, accounts for goods market spillovers on the labour market. Under some simplifying assumptions a C.E.S.-type transactions function holds, which describes aggregate employment in terms of two labour demand variables (Keynesian and capacity demand) and labour supply, and two mismatch parameters. The mismatch parameters reflect the so-called mismatch on the labour market” and “mismatch on the goods market” (*i.e.* between the composition of the demand for goods and the composition of the available production capacity). The chapter provides new theoretical and empirical insights. As a theoretical result, an explicit relationship between micro dispersions and the C.E.S. parameters is made. This paves the way for more adequately endogenising the mismatch parameters for the empirical analysis. A point to note is that the theoretical analysis treats the labour and the goods mismatch as independent phenomena, while of course there may be close links between both. Surely more work is needed to understand the complex interactions between the supply and demand determinants.

Since aggregate Keynesian demand and capacity demand are not separately observed from the data, I had to fall back to the one level C.E.S. function with labour supply and one composite demand term when it comes to estimation. The spillover effects are indirectly accounted for through the endogenisation of the mismatch parameter. The results of the theoretical analysis which established the relationship between the mismatch indicator and the variances of Keynesian demand, capacity demand and labour supply across markets are used for this purpose. The empirical model specification incorporates search elements into the framework of micro markets in disequilibrium. Including search creates a matching inefficiency within a micro market, which has to be distinguished from the mismatch between markets. The resulting aggregate C.E.S. transactions structure is a function of aggregate supply and demand, a mismatch parameter, and of search unemployment. Search unemployment in turn is related to job search duration. The influence of the long-term unemployed on structural unemployment is accounted for through the modelling of aggregate effective labour supply. This impact depends on the ratio of effective labour supply to labour supply. It should be stressed that I assumed search and mismatch as distinct phenomena within the disequilibrium setting. This is probably a simplifying assumption. Is it imaginable that for example, search intensity is a function of the degree of labour mismatch perceived by labour and firms (Lambert, 1989). Future work should elaborate on this.

The estimation results indicate that structural unemployment is predominantly determined by mismatches between micro markets and not by

search. Furthermore, the long-term unemployed are not effective on the labour market, thereby maximally contributing to structural unemployment. Over time, the part of structural unemployment involving the long-term unemployed closely follows total structural unemployment.

The dominant role of the long-term unemployed has a clear implication for future research. It shows how important it is to consider why workers become long-term unemployed (because they lack adequate skills *e.g.*), to explain the mismatches. This is probably one of the key questions for future research! The estimated rate of structural unemployment rises from 2.3% in 1960 to 4.9% in 1989 and is positively influenced by the capacity utilisation rate, replacement ratio and the regional mismatch indicator. The two last mentioned variables are used *ad hoc*, not founded on the theoretical micro model. Nevertheless, the influence of these variables on labour market mismatch is convincing.

Before implementing the simulation-based, or Monte Carlo Pseudo Maximum Likelihood to the complete model of Chapter 6, the first and second order variants of MCPML estimation are evaluated in Chapter 5. It is however, on beforehand, that using the MCPML₁ variant only makes it possible to estimate an upper-bound for the level of mismatches. The application is to a very simple model of the labour market so that the simulation results can be compared with the analytical outcomes. In the actual testing, I distinguish a variant with observed data of supply and demand, and a Monte Carlo study with supply and demand generated from the model. The findings of this study are in favour of MCPML₂ estimation because (MC)PML₂ is more efficient than (MC)PML₁.

Performance testing reveals three specific problems. First, both PML₁ and PML₂ estimates are biased downwards in all cases. PML₂ performs superior to PML₁, as the downward bias and the standard deviations are considerably less. Second, large standard deviations of the parameter estimates were found under MCPML₂. It was conjectured from the results of Laroque and Salanié (1990) that this could be due to a small variability of the supply and demand data, so that the region of a "one-sided" spurious maximum is approached. Third, the required number of replications in the Monte Carlo part of the estimation should not be smaller than 400. Unfortunately only for the MCPML₁ variant I could establish a bias-corrected variant of MCPML estimation which requires a far fewer replications. Since I decided to use the MCPML₂ method for estimation of the model in Chapter 6, this was not much help, especially because the estimation of the larger model is only feasible for a maximum of 50 replications. Given the positive findings of Laroque and Salanié (1989), I decided not to be discouraged by this and estimate the larger model of Chapter 6 using MCPML₂.

The model of Chapter 6 takes full account of the endogeneity of incomplete market clearing at the micro level, including spillovers, to estimate mismatches and structural unemployment. Additionally, lagged variables are included in the micro specifications. The estimation of the model turned out to be very demanding. Moreover, some creative solutions to problems met in the course of the optimising

process were necessary. The final results, however, seem to be satisfactory. There were no serious problems with respect to spurious maxima, and the estimated regime proportions seem to have reasonable values. I suggested some practical solutions for the computation of the standard errors of the parameters, since the matter has not been solved theoretically for these type of models.

The relevant mismatch parameters are identified through a set of identification restrictions. As a result, the structural mismatch indicator is a derived measure. The estimations of this measure indicate the presence of structural mismatches on the labour market (and the goods market). In the absence of an analytical expression for structural unemployment I computed the structural unemployment rate from the estimates of unobserved supply and demand to develop some intuition about the importance of structural problems. For the two periods 1959-1962 and 1976-1979 for which supply almost equalled demand, the structural unemployment rate computed is 0.8% and 0.9%, respectively. These estimates are lower than those that followed from the C.E.S. method of Chapter 4. I conjectured that this may be the result of a relatively high search coefficient in the labour supply equation. This result is consistent with the relatively great importance that my estimates of the regime proportions attribute to the excess demand for labour regime. The study I used for reference in this respect is Kooiman and Kloek (1985). Laroque and Salanié (1989) did not find evidence for labour market mismatches using the simulation-based model for France.

Chapter 7 presents a micro study to examine mismatches on the spot. The mere fact that a micro market supply and demand is considered as statistical constructs in the macro approaches, justifies such a micro approach. Opting for the regional dimension of mismatch in this chapter was motivated by the result of Chapter 4 which established the significant influence of a regional mismatch indicator on structural unemployment. I had to approach the phenomenon of mismatches from its reverse side, that is from the perspective of mobility behaviour of job entrants for reasons of empirical implementation. Estimation results are obtained from a cross sectional logit model.

A remarkable finding is that there is no conclusive relationship for (previously unemployed) job entrants between a weak labour market position and mobility. On the one hand, job entrants with highly valued professions or higher education are more mobile across regions. On the other hand, increasing age, a flexible job, being a foreigner, a long search period, or specific personal characteristics commonly associated with a weak labour market position, had no influence on regional mobility and thus on regional mismatch. The effect of search period cannot be compared with that of Chapter 4, which indicated a great contribution of the long-term unemployed to structural unemployment. This is because in my sample the long-term unemployed who did not find a job at all are not included. Although my sample includes job entrants who had been out of work for a long time, I do not regard these results representative for all long-term unemployed. The strong influence of the relative vacancy position on regional

labour mobility is striking. Hence, there is a declining effect on regional mismatches from high (relative) vacancies. On average people move (or commute) from a region with a small number of vacancies to a region with a large number of vacancies. This result is particularly interesting since it provides information that cannot be obtained explicitly from the macro analyses.

To measure mismatches on the labour market this book brings together two, in the literature opposing, empirical macro methods of markets in disequilibrium. I called them the analytically-based method and the simulation-based method. Both methods have their advantages and disadvantages. The simulation method is more general since less restrictive assumptions are required on the micro model specification. Still, the analytically-based method is not simply a special case. It allows a direct interpretation of mismatches in terms of causes. This is not possible with the simulation method. Therefore, it seems wise not to put the emphasis on what separates the methods but what brings them together, and so to benefit the most. I suggest the following.

The intriguing point that follows from this thesis is that both methods lead to different estimates of the extent and development of labour market mismatches. For this, I could not give a well-founded explanation. Therefore I opt for a research project to analyse the topic more rigorously. This would undoubtedly increase our understanding of mismatches on the labour market and structural unemployment. As a result, I hope, Keynesian disequilibrium analysis could regain the place in the literature that it deserves so much.

References

- Andrews, M. and S. Nickell (1984), "A Disaggregated Disequilibrium Model of the Labour Market", Working Paper, No. 669, London: Centre of Labour Economics.
- Artus, P., S. Avouyi-Dovi and G. Laroque (1985), "Estimation d'une maquette macroéconomique trimestrielle avec rationnements quantitatifs", *Annales de l'INSEE* 57: 3-25.
- Artus, P., G. Laroque and G. Michel (1987), "Estimation of a Quarterly Macroeconomic Model with Quantity Rationing", *Econometrica* 52: 1387-1414.
- Azariadis, C. (1975), "Implicit Contracts and Unemployment Equilibria", *Journal of Political Economy* 83: 1183-1202.
- Azariadis, C.C. and J.E. Stiglitz (1983), "Implicit Contracts and Fix-Price Equilibria", *Quarterly Journal of Economics* XCVIII Supplement: 1-22.
- Bailey, M.H. (1974), "Wages and Employment under Uncertain Demand", *Review of Economic Studies* 41: 1-22.
- Barro, R.J. and H.I. Grossman (1976) - *Money Employment and Inflation*, Cambridge and New York: Cambridge University Press.
- Bastelaer, A., van (1988), "The continuous Labour Force Survey in the Netherlands: data collection and processing", Unpublished paper, Statistics Netherlands.
- Bastelaer, A., van and J. Laan (1994), "The Current Job Vacancy Survey in the Netherlands", Chapter 2 in J. Muysken (ed.): *Measurement and Analysis of Job Vacancies: An International Comparison*, Newcastle upon Tyne: Athenaeum Press Ltd.
- Bean, Ch. and A. Gavosto (1989), "Outsiders, Capacity Shortages and Unemployment in the United Kingdom, in: J. Drèze *et al* (eds.), *Europe's Unemployment Problem*, Cambridge, Mass: MIT Press.

Becker, G.S. (1964) - *Human Capital: A Theoretical and Empirical Analysis*, 2nd edition, New York.

Belsley, S. (1980), "On the Efficient Computation of the Nonlinear Full-Information Maximum-Likelihood estimator", *Journal of Econometrics* 14: 203-25.

Benassy, J.-P. (1975), "Neo-Keynesian Disequilibrium Theory in a Monetary Economy", *Review of Economic Studies* 42: 503-24.

Benassy, J.-P. (1977), "On Quantity Signals and the Foundation of Effective Demand Theory", *Scandinavian Journal of Economics* 79: 147-68.

Ben-Horim, M. and D. Zuckerman (1987), "The Effect of Unemployment Insurance on Unemployment Duration", *Journal of Labour Economics* 5: 386-90.

Berg, G., van den (1990), "Search Behaviour, Transitions to Non-participation and the Duration of Unemployment", *Economic Journal* 100: 842-65.

Berkhout, P., S.W.H.C. Loozen, J. van der Valk (1993), "De positie van schoolverlaters op de arbeidsmarkt: Uitkomsten van de Enquête beroepsbevolking, 1992", *Supplement bij de Sociaal-Economische Maandstatistiek* 5.

Bierings, H.B.A., J.C.M. Imbens and C.A. van Bochove (1991), "The labour force definition", *CBS select statistical essays* 7: 41-62.

Bierings, H.B.A. and J. Muysken (1987), "The Distribution Approach to Employment Functions: Some Theoretical Reconsiderations and an Empirical Application to the Dutch Building and Construction Sector", Paper presented at the European Meeting of the Econometric Society, Copenhagen, August 1987.

Bierings, H.B.A. and J. Muysken (1988) "Nested Employment Functions and Structural Unemployment: With an Empirical Application to the Dutch Economy 1960-1985", Research Memorandum, RM 88-005 (March), Maastricht: University of Limburg.

Bierings, H.B.A. and J. Muysken (1989) "A Note on the Robustness of the Normal and Lognormal Density Function in Quantity Rationing Models", *Kwantitatieve Methoden* 31: 55-65.

Bierings, H.B.A. and J.M. Sneek (1989), "Pseudo Maximum Likelihood Techniques in a Simple Rationing Model of the Dutch Labour Market", Research Memorandum, Nr. 1989-63, Amsterdam: Free University.

- Bierings, H.B.A and A. van Witteloostuyn (1989), "Werkloosheid en Vacatures: Inzichten vanuit de Economische Theorie", *Maandschrift Economie* 53: 382-400.
- Blanchard, O. and L. Summers (1986), "Hysteresis and the European Unemployment Problem," in: S. Fischer (ed.), *NBER Macroeconomics Annual 1986*, Cambridge, Mass.: MIT Press.
- Blaug, M. (1976), "The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey", *Journal of Economic Literature* 14: 827-55.
- Bluestone, B., Murphy, W.M. and M. Stevenson (1976) - *Low Wages and the Working Poor*, Ann Arbor: Institute of Labour and Industrial Relations.
- Broeder, G. den (1983) "A Family of Market Transaction Functions", Series Foundations of Empirical Economic Research, Rotterdam: Dutch Economic Institute.
- Carlton, D.W. (1989), "The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics", in: R. Schmalensee and R.D. Willig (eds.), *Handbook of Industrial Organisation*, Amsterdam: North-Holland.
- Centraal Bureau voor de Arbeidsvoorziening (CBA, 1990), "Informatie aan Regionale Besturen voor de Arbeidsvoorziening", Rijswijk: CCBA.
- Centraal Plan Bureau, Centraal Economisch Plan (1986, 1987, 1992, 1994), Staatsuitgeverij, Den Haag.
- Clower, R. (1965), "The Keynesian Counterrevolution," in: F. Hahn and F. Brechling (eds.), *The Theory of Interest Rates*, 103-25, London: McMillan.
- Corpeleijn, A.W.F. (1980), Arbeidsmobiliteit en werkgelegenheidsstructuur: Oorzaken van sectorale verschillen in personeelsverloop", *Economisch Statistische Berichten*, januari: 88-92.
- Cowell, F.A. and M.P. Victoria-Feser (1996), Robustness Properties of Inequality Measures", *Econometrica*, 64: 77-101.
- Cramer, J.S. (1973) - *Empirical Econometrics*, Amsterdam: North Holland Publishing Company.

Cristl, J. (1994), "Mismatch Unemployment in Austria", Chapter 11 in J. Muysken (ed.): *Measurement and Analysis of Job Vacancies: An International Comparison*, Newcastle upon Tyne: Atheneum Press Ltd.

Dixon, H.D. and N. Rankin (eds.), (1995) - *The New Macroeconomics: Imperfect Markets and Policy Effectiveness*, Cambridge: Cambridge University Press.

Daganzo, C. (1979), "Multinomial probit: The Theory and Application to Demand Forecasting", Series Economic Theory, *Econometrics and Mathematical Economics*, New York: Academic Press.

Döringer, P. and M. Piore (1971) - *International Labour Markets and Manpower Analysis*, Lexington: Lexington Books.

Draper, N. and P. Kooiman (1991), "Een Klein Macroeconomisch Niet-Evenwichts-Model", Research Memorandum, No. 84, The Hague: Central Planning Bureau.

Drèze, J. (1975), "Existence of an Exchange Equilibrium under Price Rigidities", *International Economic Review* 16: 301-20.

Driehuis, W. (1987), "A Disequilibrium Analysis of Unemployment in the Netherlands, 1961-1985", Research Memorandum, No. 8722, Amsterdam: University of Amsterdam.

Economisch Instituut voor de Bouwnijverheid (1989a), "De samenstelling van de vraag naar werken in de grond-, water- en wegenbouw", Amsterdam.

Economisch Instituut voor de Bouwnijverheid (1989b), "De bouwarbeidsmarkt in het najaar van 1988", Amsterdam.

Evers, G.H.M. and A. van der Veen (1985), "A Simultaneous Non-linear Model for Labour Migration and Commuting", *Regional Studies* 19: 444-53.

Franz, W. and H. Koenig (1990), "A Disequilibrium Approach to Unemployment in the Federal Republic of Germany", *European Economic Review* 34: 413-22.

Franz, W. and W. Smolny (1994), "The Measurement and Interpretation of Vacancy Data and the Dynamics of the Beveridge Curve: The German Case", Chapter 10 in: J. Muysken (ed.): *Measurement and Analysis of Job Vacancies: An International Comparison*, Newcastle upon Tyne: Atheneum Press Ltd.

Friedman, M. (1957) - *A Theory of the Consumption Function*, Princeton, N.J.: University Press for National Bureau of Economic Research.

Gagey, F., J.-P. Lambert and B. Ottenwaelter (1988), "A Disequilibrium Estimation of the French Labor Market Using Business Survey Information", Paper presented at the May 1988 Meeting of the European Unemployment Program, Chelwood Gate.

Galbraith, J.K. (1967) - *The New Industrial State*, Boston: Houghton Mifflin.

Gelauff, G.M.M., A.R.M. Wennekers and A.H.M. de Jong (1985) "A Putty-Clay Model with Three Factors of Production and Partly Endogeneous Technical Progress", *De Economist* 133: 327-51.

Gibrat, R. (1931) - *Les Inégalités Économiques*, Paris: Sirey.

Goldfeld, S.M. and R.E. Quandt (1975), "Estimation in a Disequilibrium Model and the Value of Information", *Journal of Econometrics*, 3: 325-48.

Gorter, C. (1991) - *The Dynamics of Unemployment and Vacancies on Regional Labour Markets*, Ph.D. Thesis, Amsterdam: Tinbergen Institute.

Gorter, C. (1993), *Regionaal Arbeidsmarktonderzoek in Nederland*", Research Memorandum, RM 1993-53, Amsterdam: Free University.

Gorter, C., P. Nijkamp and P. Rietveld (1993), "Barriers to employment: Entry and Re-Entry Possibilities of Unemployed Job Seekers in the Netherlands", *De Economist* 1: 71-95.

Gouriéroux, C., J.J. Laffont and A. Monfort (1984), "Économétrie des Modèles d'Équilibre avec Rationnement: une Mise à Jour", *Annales de l'INSEE* 55-56: 5-38.

Gouriéroux, C., A. Monfort and A. Trognon (1984), "Pseudo Maximum Likelihood Methods: Theory", *Econometrica* 52: 681-700.

Gouriéroux, C. and G. Laroque (1985), "The Aggregation of Commodities in a Quantity Rationing Model", *International Economic Review* 26: 681-99.

Grandmont, J.-M. and G. Laroque (1976), "On Keynesian Temporary Equilibria", *Review of Economic Studies* 43: 327-34.

Grip, A. de (1987), "Causes of Labor Supply and Demand Mismatches in the Dutch Building Trades", *De Economist* 135: 182-200.

- Hamersley, J.M. and D.C. Handscomb (1964) - *Monte Carlo Methods*, London: Chapman and Hall.
- Hart, O.D. (1982), "A Model of Imperfect Competition with Keynesian Features", *Quarterly Journal of Economics* 97: 109-38.
- Hartley, M.J. and P. Mallela (1977), "The Asymptotic Properties of a Maximum Likelihood Estimator for a Model of Markets in Disequilibrium", *Econometrica*, 5: 1205-20.
- Hartog, J. (1980) - *Tussen vraag en aanbod*, Leiden: Stenfert Kroese.
- Hartog, J., E. Mekkelholt and H. van Ophem (1987), "Een Empirische Studie naar de Arbeidsmobiliteit in Nederland, *OSA werkdocument* W32, Den Haag.
- Hicks, J.R. (1965) - *Capital and Growth*, Oxford: Oxford University Press.
- Hirsch, B.T. and J.T. Addison (1986) - *The Economic Analysis of Unions: New Approaches and Evidence*, Boston: Allen and Unwin.
- Holden, K. and D.A. Peel (1975), "The Determinants of Unemployment and the U-V relationship", *Applied Economics* 7.
- Holt, C.C. (1970), "Job Search, 'Phillips' Wage Relation, and Union Influence: Theory and Evidence", in: E.S. Phelps *et al* (eds.), *Micro Foundations of Employment and Inflation Theory*, 53-123, London.
- Holzer, H. (1988), "Search Method Use by Unemployed Youth", *Journal of Labour Economics* 6: 1-20.
- Houthakker, H.S. (1955) "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis", *Review of Economic Studies* 23: 27-31.
- Ijiri, Y. and H.A. Simon (1977) - *Skew Distributions and the Sizes of Business Firms*, Amsterdam: North Holland Publishing Company.
- Jackman, R. and R. Layard (1990), "Does Long-Term Unemployment Reduce a Person's Chance of a Job? A Time-Series Test", *Economica* 58: 93-106.
- Johnson, N.L. and S. Kotz (1972) - *Distributions in Statistics: Continuous Multivariate Distributions*, New York.

- Jones, S. (1989), "Job Search Methods, Intensity and Effects", *Oxford Bulletin of Economics and Statistics* 51: 277-96.
- Jovanovic, (1979), "Firm Specific Capital and Turnover", *Journal of Political Economy* 87: 1246-60.
- Kerckhoffs, C. (1992), "Some Notes on Transaction Functions and Regime Proportions in Quantity Rationing Model, Unpublished Manuscript", Maastricht: University of Limburg.
- Kerr, C. (1954), "The Balkanization of Labour Markets", in: E. Wight Bakke *et al* (eds.), *Labor Mobility and Economic Opportunity* 92-110, Cambridge, Mass: MIT Press.
- Keynes, J.M.(1936) - *The General Theory of Employment, Interest and Money*, New York: Harcourt Brace.
- Killingsworth, M.R. (1983) - *Labour Supply*, New York: Cambridge University Press.
- Kloosterman, R.C. (1987), "Achteraan in de rij: Een onderzoek naar de factoren die (her)intreding van langdurig werklozen belemmeren", *OSA werkdocument* W38.
- Kooiman, P. and T. Kloek (1979), "Aggregation of Micro Markets in Disequilibrium: Theory and Application to the Dutch Labour Market, 1948-1975", Working Paper, Rotterdam: Erasmus University.
- Kooiman, P. and T. Kloek (1985), "An Empirical Two Market Disequilibrium Model for Dutch Manufacturing", *European Economic Review* 29: 323-54.
- Kooiman, P. (1984), "Smoothing the Aggregate Fix-Price Model and the Use of Business Survey Data", *The Economic Journal* 94: 899-913.
- Koyck, L.M. (1954) - *Distributed Lags and Investment Analysis*, Amsterdam: North-Holland.
- Laffont, J.J. (1983), "Fix-Price Models: A Survey of Recent Empirical Work", Working Paper, No. 8305, GREMAQ, University of Toulouse.
- Laffont, J-J, H. Ossard and Q. Vuong (1995), "Econometrics of First Price Auctions", *Econometrica* 63: 962-82

Lambert, J.-P. (1988) - *Disequilibrium Macroeconomic Models*, Cambridge: Cambridge University Press.

Lambert, J.-P. (1990), "The French Unemployment Problem: Lessons from a Rationing Model Relying on Business Survey Information", *European Economic Review* 34: 423-33.

Laroque, G. (1991), "Comparative Estimates of a Macroeconomic Disequilibrium Model: France, Germany, the U.K. and the U.S.A.", Preliminary first draft, Paris: INSEE.

Laroque, G. and B. Salanié (1989), "Estimation of Multi-Market Fix-Price Models: An Application of Pseudo Maximum Likelihood Methods", *Econometrica* 57: 831-60.

Laroque, G. and B. Salanié (1990), "Estimating the Canonical Disequilibrium Model: Asymptotic Theory and Finite Sample Properties, Working Paper, Paris: INSEE.

Laroque, G. and B. Salanié (1992), "Macroeconomic Disequilibrium Models", in: H. Pesaran and M. Wickens (eds.), *Handbook of Applied Econometrics*, First Draft of a Chapter prepared for the book.

Layard, R.S., Nickell, S. and R. Jackman (1991) - *Unemployment, Macroeconomic Performance and the Labour Market*, Oxford: Oxford University Press.

Leijonhufvud, A. (1968) - *On Keynesian Economics and the Economics of Keynes*, London: Oxford University Press.

Lindbeck, A. and D. Snower (1985), "Explanations of Unemployment", *Oxford Review of Economic Policy* 1: 34-59.

Lippman, S.A. and J.J. McCall (1976), "The Economics of Job Search: A Survey", Parts I and II, *Economic Inquiry* 14: 155-89 and 347-68.

Loozen S.W.H.C. and R.J.P. Dekker (1993), "De relatie tussen beroep en opleiding: Uitkomsten van de Arbeidskrachtentelling en de Enquête beroepsbevolking", *Supplement bij de Sociaal Economische Maandstatistiek* 4: 4-10.

Lucas, R. and L. Rapping (1970), "Real Wages, Employment and Inflation", in: E.S. Phelps (ed.): *Microeconomic Foundations of Employment and Inflation Theory*, New York: W.W. Norton, 257-305.

- McFadden, D. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration", *Econometrica* 57: 995-1026.
- Malinvaud, E. (1977) - *The Theory of Unemployment Reconsidered*, Oxford: Basil Blackwell.
- Malinvaud, E. (1980) "Macroeconomic Rationing of Employment," in: E. Malinvaud and J.P. Fitoussi (eds.), *Unemployment in Western Countries*, London: Macmillan.
- Meersman, H. (1987) - *A Disequilibrium Macroeconomic Model with Endogenous Prices*, Ph.D. Thesis, Antwerpen: Universiteit Antwerpen.
- Mehta, F and H.R. Sneessens (1990), "Belgian Unemployment or the Story of a Small Open Economy Caught in a Worldwide Recession", Research Memorandum, Université Catholique de Louvain and Faculté Libre de Sciences Économiques de Lille.
- Mortenson, D.T. (1970), "A Theory of Wage and Employment Dynamics", in: E.S. Phelps *et al* (eds.), *Microeconomic Foundations of Employment and Inflation Theory*, New York: Norton.
- Muellbauer, J. (1978), "Macrotheory vs. Macroeconometrics: the Treatment of Disequilibrium in Macro Models", Birbeck Discussion Paper 59, London: Birbeck College.
- Muysken, J., H.B.A. Bierings and E. de Regt (1990), "Structurele werkloosheid in Nederland, 1960-1988", *OSA werkdocument*, W82, januari 1991, Den-Haag.
- Muysken, J., C. de Neubourg and H. van der Burg (1982), "Regional and Occupational Labor Market Imperfections, the Netherlands, 1955-1980", Research Memorandum 105, Groningen: University of Groningen.
- Muysken, J. and H. Meijers (1988), "A Post-Mortem on U-V Analysis", Research Memorandum, No. 88-031, Maastricht: University of Limburg.
- Muysken, J. (ed.), (1994) - *Measurement and Analysis of Job Vacancies: An International Comparison*, Newcastle upon Tyne: Atheneum Press Ltd.
- Muysken, J., H.B.A. Bierings and E. de Regt (1994), "Vacancy Data, the Employment Function and Structural Unemployment in the Netherlands", Chapter 12 in: J. Muysken (ed.): *Measurement and Analysis of Job Vacancies: An International Comparison*, Newcastle upon Tyne: Atheneum Press Ltd.

- Muysken, J. and E. de Regt (1992), "Imperfect Matching and the Dynamics of the Employment Function", Research Memorandum (March), Maastricht: University of Limburg.
- Muysken, J. and E. de Regt (1994), "A Dynamic Analysis of Structural Unemployment in 5 European Countries: Austria, Germany, the Netherlands, Sweden and the United Kingdom", Research Memorandum (March), Maastricht: University of Limburg.
- Neftçi, S.N. (1984), "Are Economic Time Series Assymmetric over the Business Cycle", *Journal of Political Economy* 92, p. 307.
- Negishi, T. (1978), "Existence of an Underunemployment Equilibrium", in: G. Schwodiauer (ed.), *Equilibrium and Disequilibrium in Economic Theory*, Dordrecht: Reidel.
- Neubourg, C. de, (1987) - *Unemployment, Labour Slack and Labour Market Accounting, Dissertation*, Groningen.
- Oi, W. (1971), "Labor as a Quasi-fixed Factor", *Journal of Political Economy* 70: 538-55.
- Openshaw, S. and P.J. Taylor (1979), "A Million or so Correlation Coefficients: Three experiences on the Modifiable Areal Unit Problem", in: N. Wrigley (ed.), *Statistical Applications in the Spatial Sciences*, London.
- Padoa Schioppa, F. (1990), "Classical, Keynesian and Mismatch Unemployment in Italy", *European Economic Review* 34: 434-42.
- Patinkin, D. (1956) - *Money, Interest and Prices*, New York: Harper and Row.
- Pissarides, C.A. (1985), "Job Search and the Functioning of Labour Markets", in: Carline *et al* (eds.), *Labour Economics*, 159-85, New York: Longman.
- Pissarides, C.A. (1992), "Loss of Skill during Unemployment and the Persistence of Employment Shocks", *Quarterly Journal of Economics*, November: 1371-91.
- Pfann, G.A. (1989), - *Stochastic Adjustment Models of Labour Demand*, Ph.D. Thesis, Maastricht, Universiteit van Limburg.
- Quandt, R.E. (1988) - *The Econometrics of Disequilibrium*, New York: Basil Blackwell.

- Romme, S. and G. Wolfs (1986) "Loonkosten per Gewerkt Uur; een Inventarisatie van Bronnen en Tijdreksconstructie", Research Memorandum, No. 86-010, Maastricht: University of Limburg.
- Salop, S.C. (1979), "A Model of the Natural Rate of Unemployment", *American Economic Review* 69: 117-25.
- Shapiro, C. and J. Stiglitz (1984), "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, 74: 433-44.
- Siebert, W.S. (1985), "Development in the Economics of Human Capital", in: S. Carline *et al* (eds.), *Labour Economics*, 5-77, New York: Longman.
- Smart, M.W. (1974), "Labour Market Areas: Uses and Definition, *Progress in Planning* 2-4: 239-353.
- Sneek, J.M. (1983), "Some Approximations to the Exact Distribution of Sample Autocorrelations for Autoregressive Moving Average Models," in: O.D. Andersen (ed.): *Time Series Analysis: Theory and Practise*, Amsterdam: North Holland.
- Sneessens, H.R. (1981) - *Theory and Estimation of Macroeconomic Rationing Models*, Berlin: Springer Verlag.
- Sneessens, H. (1983), "Aggregation in a Quantity Rationing Model", Unpublished Manuscript, Département des Sciences Économiques, Louvain: Université Catholique de Louvain.
- Sneessens, H.R. and J.H. Drèze (1986), "A Discussion of Belgian Unemployment, Combining Traditional Concepts and Disequilibrium Econometrics, *Economica* 53: S89-S119.
- Sneessens H.R. and F.S. Mehta (1993), "Real Wages, Skill Mismatch and Unemployment Persistence, France 1962-1989", Paper presented at the Applied Econometrics Association, April 28-29, 1994, Aix-en-Provence.
- Spiegel, M.R. (1968) - *Mathematical Handbook of Formulas and Tables*, Schaum's outline series, McGraw-Hill Book Company.
- Stalder, P. (1989), "A Disequilibrium Model With Smooth Regime Transitions and a Keynesian Spillover for Switzerland's Labor Market", *European Economic Review* 33: 863-93.

Statistics Netherlands, *Sociaal Economische Maandstatistiek*, several editions, The Hague.

Steindl, J. (1965) - *Random Processes and the Growth of Firms*, London: Griffin.

Teulings, C.N. (1990) - *Conjunctuur en Kwalificatie*, Amsterdam: Stichting voor Economisch Onderzoek der Universiteit van Amsterdam.

Vissers, A.C.M. de Vries and Th. Schepens (1986), "Arbeidsmarktgedrag ten Tijde van Massale Werkloosheid: De eerste uitkomsten van een landelijk onderzoek", *OSA Voorstudie 12*, Staatsuitgeverij, 's-Gravenhage.

Walras, L. (1874) - *Eléments d'économie politique pure*, 1974, Lausanne: Corbaz, English translation, *Elements of Pure Economics*, 1954, Homewood, Ill.:Irwin.

Wright, E.O. (1979) - *Class Structure and Income Determination*, New York: Academic Press.

Yellen, J.L. (1984), "Efficiency Wage Models of Unemployment", *American Economic Review*, Papers and Proceedings 74: 200-5.

Nederlandse samenvatting

Dit proefschrift gaat over de gebrekkige aansluiting van vraag en aanbod op de arbeidsmarkt en structurele werkloosheid. In de ruimste betekenis is structurele werkloosheid het gevolg van (i) kwalitatieve discrepanties op de arbeidsmarkt ('mismatches') en (ii) zoekwerkloosheid die zich voordoet bij het normale verkeer op de arbeidsmarkt. Tot uitdrukking komt deze vorm van werkloosheid in het gelijktijdig voorkomen van werkloosheid en vacatures op de arbeidsmarkt. In het empirische gedeelte van het proefschrift staat de vraag centraal naar de omvang en de oorzaak van 'mismatches' en van structurele werkloosheid op de Nederlandse arbeidsmarkt. Om een zo'n compleet mogelijk beeld te verkrijgen is voor het beantwoorden van deze vraag gebruik gemaakt van macro-economische tijdreeks-analyse en micro-economisch dwarsdoorsnede onderzoek.

De macro-economische methoden die in dit boek een rol spelen, zijn empirische toepassingen van de onevenwichtigheids- of 'fix-price' theorie. Deze theorie heeft een micro-economische grondslag in de zin dat zij is opgezet vanuit het bedrijf en het huishouden. Keynes en later onder andere Barro en Grossman en Malinvaud zijn belangrijke namen die aan deze theoretische stroming verbonden zijn. Een belangrijk uitgangspunt van de 'fix-price' theorie is de veronderstelling van starre lonen en prijzen. Zo zorgt een starre loonvoet ervoor dat een onevenwichtige situatie van werkloosheid kan blijven voortbestaan. Daarbij spelen zogenaamde 'spillover' effecten tussen de goederenmarkt en de arbeidsmarkt een grote rol. Huishoudens en bedrijven zullen hun vraag of aanbod op de ene markt aanpassen als zij op de andere markt beperkt worden, bijvoorbeeld in hun afzet of in consumptieve vraag. Zo ontstaan vier mogelijke configuraties van vraag- of aanbod overschotten, ook wel regimes genoemd. Deze zijn door Malinvaud systematisch in kaart gebracht. In het zogenaamde regime van klassieke werkloosheid zijn ondernemers bij de gegeven loonvoet niet bereid zoveel arbeiders in dienst te nemen dat aan de goederenvraag kan worden voldaan.

Het begrip 'mismatches' wordt geïntroduceerd door een expliciete formulering van deelmarkten in niet-evenwicht situaties. De uitwerking van de idee van deelmarkten in deze context is afkomstig van Muellbauer. Bij arbeidsdeelmarkten moet men vooral denken aan markten voor specifieke soorten arbeid. Daarbij kunnen ook (persoons)kenmerken als leeftijd, geslacht, nationaliteit, ervaring, opleiding *etc.* te onderscheiden dimensies zijn. Het uitgangspunt van de methode is een economie bestaande uit een groot aantal deelmarkten, waarbij op

iedere deelmarkt de werkgelegenheid (of afzet) wordt bepaald door het minimum van vraag en aanbod (de zogenaamde *minimum conditie*). Op een deelmarkt is dus alleen sprake van werkloosheid of vacatures. Vraag en aanbod zijn op al deze markten verschillend, maar kunnen niet worden gekwantificeerd door het meerdimensionale karakter van deelmarkten en het ontbreken van datamateriaal. Daarom wordt verondersteld dat vraag en aanbod over deelmarkten een specifieke verdeling volgen. Na aggregatie van de werkgelegenheid over deelmarkten ontstaat een 'glad' verband tussen geaggregeerde werkgelegenheid enerzijds en de geaggregeerde vraag- en aanbodhoeveelheden anderzijds. Dit verband impliceert het gelijktijdig voorkomen van werkloosheid en vacatures. De door aggregatie verkregen maat voor 'mismatches' weerspiegelt de variantie van vraagoverschotten. Deze maat staat in directe relatie tot de structurele werkloosheid.

In het proefschrift worden twee verschillende varianten van de macro-economische tijdreeksanalyse toegepast. De eerste is de analytische variant, de tweede de simulatie variant. In de analytische variant wordt een geaggregeerde werkgelegenheidsfunctie afgeleid waarbij een specifieke verdeling van micro vraag en aanbod is verondersteld. De functie is niet-lineair in de geaggregeerde vraag en aanbodhoeveelheden en bevat de 'mismatch' als parameter. Een benadering van deze functie heeft een eenvoudige C.E.S. structuur. In de empirische toepassing van de analytische variant is verondersteld dat de geaggregeerde vraag en het aanbod waarneembare grootheden zijn en overeenkomen met respectievelijk de som van werkgelegenheid en vacatures en werkgelegenheid en werkloosheid. Voor deze variant is derhalve wezenlijk dat goede data voorhanden zijn. Nog niet zo lang terug waren de data van vacatures en werkloosheid behept met grote meetfouten. Dit kwam doordat de registratie van de arbeidsbureaus niet up-to-date was. Inmiddels zijn verbeterde cijfers van het CBS beschikbaar waarbij gebruik is gemaakt van enquêtes. De oudere gegevens zijn door Muysken en anderen middels statistische correctie opgewaardeerd, waarbij onder andere gebruik is gemaakt van de CBS enquête-informatie. Door (ondermeer) de statistische correctie blijft de veronderstelling van 'goede' data voor een langere tijdreeks een mogelijk kwetsbaar punt van deze variant. Daarom blijft de vergelijking van de uitkomsten over 'mismatches' en structurele werkloosheid op basis van andere methoden die deze veronderstelling niet hanteren, geboden. Een belangrijk voordeel van betreffende variant is dat zij leidt tot intuïtief en economisch gemakkelijk interpreteerbare uitkomsten en eenvoudige schattingsmethoden toelaat. Bovendien leent de methode zich om oorzaken van 'mismatches' en structurele werkloosheid op te sporen. Daar staat weer tegenover dat vereenvoudigende veronderstellingen nodig zijn om überhaupt een analytisch verband te kunnen afleiden. Wanneer bijvoorbeeld 'spillover' effecten van de goederenmarkt naar de arbeidsmarkt en *vice versa* op micro niveau worden gemodelleerd is geen analytische uitdrukking voor het aggregaat meer te bepalen.

De simulatievariant is in essentie een schattingstechniek die de specificering van een zeer algemeen model op microniveau toestaat. Het opnemen van vraag en aanbod relaties van arbeid als ook van goederen en 'spillover' effecten is mogelijk. Aggregatie over deelmarkten wordt uitgevoerd door simulatie en is geïntegreerd in de schattingsmethode. Anders dan de analytische methode vereist de simulatiemethode geen specifieke verdeling van micro vraag en aanbod. De schattingsmethode, wel bekend onder de naam Monte Carlo Pseudo Maximum Aannemelijkheid (MCPMA), maakt gebruik van de eerste twee momenten (verwachting en variantie) van de endogene variabelen (bijvoorbeeld werkgelegenheid, afzet, consumptie *etc.*). Laroque en Salanié hebben deze methode voor de toepassing van onevenwichtige markten ontwikkeld. Analytische uitdrukkingen zijn niet meer beschikbaar, hetgeen de interpretatie van de uitkomsten bemoeilijkt en de methode ongeschikt maakt de oorzaken van 'mismatches' en structurele werkloosheid expliciet aan te geven. Een ander bezwaar dat aan de simulatiemethode kleeft, is dat het zeer computer intensief is hetgeen beperkingen oplegt aan 'spelen' met het model. Bovendien is toetsing van de methode vereist om achter de eigenschappen van de schatters te komen. Voor de berekening van standaardfouten van de parameters, het vaststellen van autocorrelatiepatronen zijn geen theoretische resultaten beschikbaar en moet teruggegrepen worden op toetsen waarvan vooralsnog de bruikbaarheid niet is aangetoond. In dit proefschrift worden beide macro methoden toegepast om de 'mismatch' problematiek voor Nederland te analyseren.

De micro-economische dwarsdoorsnede studie completeert de uitkomsten van de macro-economische tijdreeksanalyses. Ondermeer omdat de data beschikbaarheid een meerdimensionale benadering niet toelaat, is een unidimensionale aanpak gekozen en wel die van de regionale arbeidsmarkten. De analyse beschouwt de andere kant van 'mismatches' door zich te richten op het vaststellen van regionale arbeidsmobiliteit in Nederland en de factoren die daarop van invloed zijn.

De inhoud van de acht hoofdstukken is als volgt. **Hoofdstukken 1 en 2** zijn introductiehoofdstukken. Het probleem wordt gedefinieerd en de plaats in de literatuur aangegeven. Als voorbereiding op de schatting van de modellen volgens de analytische en de simulatievariant wordt in de hoofdstukken 3 en 5 op de methodologische aspecten van de beide varianten ingegaan. **Hoofdstuk 3** beziet zowel vanuit theoretisch als empirisch gezichtspunt de bruikbaarheid van de log-normale en de normale verdeling voor het beschrijven van micro vraag en aanbod. Deze verdelingen zijn gangbaar in de literatuur. In de theorie van stochastische processen worden aanwijzingen gevonden dat de Pareto verdeling een beter alternatief is voor het beschrijven van (dynamische) arbeidsmarkten dan de log-normale verdeling. Een verrassend resultaat is dat de werkgelegenheidsfunctie op basis van de Pareto verdeling een zeer vereenvoudigde afspiegeling is van de werkgelegenheidsfunctie gebaseerd op de log-normale verdeling. Zij impliceert een

uiterst irrealistische structurele werkloosheidsvoet van 50%! Werkgelegenheidsfuncties afgeleid op basis van andere verdelingen laten zich niet gemakkelijk beoordelen ten opzichte van de functie gebaseerd op de (log-)normale verdeling. Gebruikmakend van geneste toetsing wordt de log-normale verdeling niet verworpen ten opzichte van de algemenere werkgelegenheidsfunctie volgens de Weibull verdeling. In hetzelfde hoofdstuk is nog een andere toets ontwikkeld om na te gaan of the (log-)normale verdeling bruikbaar is. Deze toets berust op de eigenschap dat de som van twee (log-) normale verdelingen niet (log-)normaal is verdeeld. De vorm van de werkgelegenheidsfunctie die op basis van deze nieuwe verdeling kan worden afgeleid fungeert als algemene vorm bij de geneste toetsing. Toegepast voor meerdere situaties volgt ook hier dat de (log-)normale verdeling niet wordt verworpen. Dit onderbouwt de toepassing van deze verdelingen voor de hoofdstukken 4 en 6. In **hoofdstuk 5** wordt de MCPMA methode geëvalueerd voor het geaggregeerde minimum model (geaggregeerde werkgelegenheid is het minimum van geaggregeerde vraag en aanbod). Het model is bewust zo eenvoudig gekozen, omdat hiervoor analytische resultaten beschikbaar zijn die vergeleken kunnen worden met die van de simulatie. Op grond van deze resultaten ondersteund met die van een extra simulatie is besloten de MCPMA₂ techniek (en niet de MCPMA₁ techniek) te gebruiken voor de schatting van het algemene model van hoofdstuk 6. De MCPMA₂ techniek maakt gebruik van de eerste twee momenten, de MCPMA₁ techniek alleen van de verwachting. Bij de keuze voor de MCPMA₂ techniek was ook van betekenis dat de MCPMA₂ schatter meer efficiënt is dan de MCPMA₁ schatter. Een belangrijk voordeel van de MCPMA₂ boven de MCPMA₁ schatter is dat met deze schatter de omvang van de 'mismatches' is vast te stellen, terwijl op basis van de MCPMA₁ schatter alleen een bovengrens kan worden bepaald. In dit hoofdstuk is eveneens de consistentie van de schatters aangetoond voor een eindig aantal replicaties in het Monte Carlo gedeelte wanneer een 'bias' correctie wordt toegepast. Voor MCPMA₂ is de 'bias' gecorrigeerde variant nog te weinig ontwikkeld om toegepast te kunnen worden. Dit betekent dat vooralsnog geen efficiency winsten kunnen worden bereikt bij de schatting van het algemene model van hoofdstuk 6. De analyse van hoofdstuk 5 geeft evenwel aan dat het aantal vereiste replicaties in de niet bias gecorrigeerde variant beduidend hoger ligt, en hoger dan voor het algemene model kan worden gerealiseerd. De schatting zou dan vanwege het enorme tijdsbeslag nauwelijks meer uitvoerbaar zijn.

In **hoofdstuk 4** wordt een generalisatie van de C.E.S. werkgelegenheidsfunctie gesuggereerd. Daarbij gold als uitgangspunt dat de situatie op de arbeidsmarkt alleen zinvol kan worden beoordeeld in samenhang met de goederenmarkt (zie hierboven). De werkgelegenheid is een functie van de Keynesiaanse vraag naar arbeid, de capaciteitsvraag naar arbeid en het arbeidsaanbod. De twee C.E.S. parameters staan voor de goederenmarkt 'mismatch' (d.w.z. tussen de compositie van de vraag naar goederen en de aanwezige productiecapaciteit) en de arbeidsmarkt 'mismatch'. In het theoretische deel van het hoofdstuk wordt de relatie tussen micro 'mismatches' en de C.E.S.

parameters expliciet gemaakt. Dit is nodig om tot een adequate endogenisering van de C.E.S. parameters en dus van 'mismatches' en structurele werkloosheid te komen. In eerdere toepassingen is dit aspect veronachtzaamd. Uit de schatting van de weliswaar afgeslankte C.E.S. functie met arbeidsvraag en aanbod als waargenomen variabelen, blijkt de verhouding uitkeringsniveau/minimum loon, een regionale 'mismatch' indicator en de bezettingsgraad de 'mismatch' tussen arbeidsmarkten positief te beïnvloeden. Het empirische gedeelte geeft ook inzicht in het belang van drie componenten voor het verklaren van structurele werkloosheid:

- (i) Mismatches verklaren een substantieel deel van de structurele werkloosheid.
- (ii) Zoekwerkloosheid beïnvloedt nauwelijks het niveau van de structurele werkloosheid.
- (iii) De ontwikkeling van structurele werkloosheid wordt bijna volledig bepaald door het niet effectief zijn van langdurig werklozen op de arbeidsmarkt.

Volgens de schattingen steeg de structurele werkloosheidsvoet van 2,3% in 1960 tot en met 4,9% in 1989.

In het model van **hoofdstuk 6** zijn de goederen vraag en aanbod relaties, en de arbeidsvraag en aanbod relaties gemodelleerd op micro niveau. In deze relaties spelen 'spillover' effecten een rol. Tevens zijn vertraagde variabelen in de micro relaties opgenomen. De vraag en aanbod variabelen zijn niet waargenomen grootheden. De relevante 'mismatch' parameters zijn geïdentificeerd via een set van identificatierestricties. De schatting van het model is erg computer intensief, waarbij veel tijd nodig is om acceptabele startwaarden te vinden. Bovendien waren aanpassingen in de optimalisatieroutine nodig, bijvoorbeeld wanneer het zoekgebied naar de optimale parameterwaarden buiten de parameterruimte dreigde te komen.

De structurele werkloosheid kan worden berekend uit de geschatte waarden voor de geaggregeerde vraag en het aanbod. Dit is slechts mogelijk voor de subperioden 1959-1962 en 1976-1979. De structurele werkloosheidsvoet bedroeg respectievelijk 0,8% en 0,9%. Deze schattingen liggen beduidend lager dan die volgens de analytische (C.E.S.) variant. Waarom dit het geval is, is moeilijk na te gaan. Waarschijnlijk speelt de hoge zoekcoëfficiënt in de arbeidsaanbodrelatie hier een belangrijke rol. Duidelijk is dat op dit punt meer onderzoek gewenst is. De overige modeluitkomsten zijn acceptabel. De parameters hebben de juiste tekens, en de waardes zijn vergelijkbaar met andere studies. De berekende regime proporties (deze geven het belang aan van een bepaalde configuratie van goederenvraag en aanbod overschotten of tekorten in een bepaald jaar) zijn over het algemeen niet strijdig met intuïtie en andere studies.

Hoofdstuk 7 presenteert de resultaten van een micro (dwarsdoorsnede) studie naar regionale arbeidsmobiliteit. Factoren die mobiliteit beperken zijn geïnterpreteerd als factoren die 'mismatches' bevorderen. De studie completeert de macro-analyses doordat uitgegaan wordt van individueel gedrag en deelmarkten niet beschouwd zijn als theoretische constructies. Een andere aanleiding voor de

micro studie is dat in het micro model met intertemporele “spillover” effecten, de mobiliteitsparameters niet zijn geïdentificeerd.

In de logit model specificatie die is gebruikt om regionale arbeidsmarkt mobiliteit te verklaren, is de kans dat een baanvinder regionaal mobiel is afhankelijk gesteld van een aantal factoren zoals leeftijd, geslacht, opleiding, beroep, het soort regio, de relatieve vacaturepositie van de regio. De uitkomsten van de schatting tonen aan dat voor baanvinders die voorheen werkloos waren, factoren die normaal gesproken op een zwakke arbeidsmarktpositie duiden, regionale mobiliteit niet negatief beïnvloeden. Daarbij moet wel bedacht worden dat in de steekproef uitsluitend baanvinders zijn opgenomen. Langdurig werklozen die nauwelijks of geen kans hebben op een baan, zijn ondervertegenwoordigd. Opmerkelijk is dat de relatieve vacaturepositie van de regio de kans op regionale mobiliteit zeer sterk beïnvloedt. Kent de eigen regio naar verhouding veel vacatures trekt zij mensen aan, in het omgekeerde geval stoot zij af.

Hoofdstuk 8 bevat de conclusies en een samenvatting.

Curriculum vitae

Harry Bierings is op 20 november 1959 geboren in Den Haag . In 1978 behaalde hij zijn Atheneum *B* diploma aan het Christelijk Lyceum te Rijswijk. Van 1978 tot 1984 studeerde hij Economie aan de Erasmus Universiteit Rotterdam. Na het behalen van zijn doctoraal diploma (richting ontwikkelingsprogrammering) is hij in januari 1985 als Wetenschappelijk Assistent in dienst getreden bij de vakgroep Algemene Economie der Rijksuniversiteit Limburg te Maastricht. Vanaf juni 1989 is hij, aanvankelijk als statistisch onderzoeker en later ook als projectleider, werkzaam bij het Centraal Bureau voor de Statistiek te Heerlen.

Mismatches on the Labour Market and Structural Unemployment

Unemployment rates differ widely between groups. They tend to be much higher among unskilled people than among those with skills. There are differences between occupations, regions, age groups and people from various ethnic backgrounds. It is remarkable how persistent these differences are. This book examines the imbalances between supply and demand, or mismatches, in relation to structural unemployment. Two empirical questions underlie the study. The first question is: How important are mismatches and structural unemployment on the Dutch labour market and how did they develop over time. The issue of mismatch development touches on a debate between researchers from the equilibrium oriented school and exponents of the disequilibrium framework. The second question is: What causes labour market mismatches.

A macroeconomic approach and a microeconomic approach are adopted to study the subject. The reason for considering the two approaches simultaneously is that the combined outcomes are less dependent on the arbitrariness of the individual methods. The macro method plays a key role in this book. It fits the familiar analysis of micro markets in disequilibrium to describe the necessary heterogeneity of aggregate markets. Two alternative and in part competing variants are discussed. The disequilibrium line of thinking assumes that wages and prices adjust too slowly to clear markets permanently and restore the imbalance between demand and supply. In the short run employment may be determined one-sidedly by demand or supply, while the other side of the market gets rationed.

The book also has a strong methodological character in a sense that the macro methods are evaluated, and the assumptions scrutinised, theoretically as well as empirically.

Harry Bierings was born on November 20, 1959 in The Hague, the Netherlands. He studied economics at the Erasmus University in Rotterdam from 1978 until 1984. Between 1985 and 1989 he worked as a research fellow at the department of economics of the University of Limburg at Maastricht. Since 1989 he has been statistician at Statistics Netherlands. His scientific interest is mainly focused on labour market models and survey design and sampling.