7 EVALUATION AND CONCLUSIONS

In the previous chapters various regional and spatial models are discussed which are quite different from each other. Chapter 2 deals with short-term interregional input-output forecasting in a small economically developed country, characterized by relatively little regional differences, short distances between regions and strong interregional interdependencies. Chapter 3 is about long-term input-output forecasting and policy simulation for a very large developing country, modelling many regions in very different stages of development and sometimes at long distances from one another with weak interregional ties. The hybrid simulation model in chapter 4 still has the interregional input-output model of chapter 2 as a core element, but is extended with model blocks for land use, traffic congestion, environmental variables and cost-benefit analysis. Finally, chapter 5 and 6 change the modelling perspective from interregional analysis to spatial agglomeration modelling, from the meso-level of sectoral differentiation to the micro level of individual agents, and from the national perspective to the larger level of Europe as a whole.

Theoretically, the main difference between chapter 2-4 on the one hand and chapter 5-6 on the other is about the Leontief model versus the Dixit-Stiglitz approach. These are two different tools of analysis that are aimed at answering different questions. The implementations of both types of models presented in this book do not allow for a general conclusion about which one should be considered to be superior to the other. Therefore, in this chapter both approaches will first be evaluated separately. There are, however, also some more general modelling issues which are relevant for a joint discussion of both approaches. In addition, the recent literature seems to indicate that both types of modelling are converging towards one another. This will be discussed at the end of this chapter.

7.1 Input-output analysis

First, it should be pointed out that the interregional i-o models used in chapter 2-4 are atypical in an international context. Apart from some exceptions in Japan and the Netherlands, the vast majority of currently operational regional i-o models in most countries are not interregional but deal with one region only (Schaffer, 1999). An important advantage of interregional i-o models is that they capture not only intersectoral interdependence within the region, but also trade interdependencies between regions. In addition, an interregional model that covers all regions of a country always has an implicit national result for all endogenous variables. As discussed in detail in chapter 3, this is very useful. When exogenous national forecasts are used as model input, the endogenous national results can be checked and corrected with their exogenous national counterparts putting an error-reduction constraint on the regional and sectoral results.
Of course, when interregional i-o tables are available, no one will disagree that an interregional model is preferable to a single-region model. Even if we are interested in one single region only, the interregional model is still relevant because then we simply add up all other regions into one "rest of the country" region. The resulting biregional model then still has all the advantages of interregional analysis just mentioned.

Apart from the number of regions modelled, an important question is why we should use i-o analysis when we want to make regional economic forecasts. The theoretical limitations are obvious: demand is the only driving force, i-o coefficients are mostly assumed constant, supply is mostly assumed perfectly elastic and therefore prices do not play a role. Nevertheless, there are two good reasons. Firstly, sectoral differentiation is one of the main causes of regional differences at the macro level. Whether a region is dominated by agriculture or manufacturing or services makes all the difference when it comes to modelling its economic growth. Secondly, no matter how supply and price formation on markets of intermediate goods is handled, there is an undeniable "technical" component that will always play its role: car manufacturing needs steel, dairy manufacturing needs milk etc. and when these inputs are simply not produced locally, they will have to come from other regions. In other words, output growth of industry \( i \) in region \( r \) will affect output growth of industry \( j \) in region \( s \), and it is this interdependency that can only be captured by an interregional i-o model.

It is also this lack of interdependency that is one of the major drawbacks of shift-and-share analysis (SAS), which today still is the most commonly used technique and main competitor of i-o analysis when it comes to regional forecasting\(^{60}\). When two industries \( i \) and \( j \), whether in the same region or in different regions, have strong intermediate relations, their growth rates will be correlated, and therefore an independent forecast for each of them, that does not take this interdependency into account, can be expected to produce erroneous results. Compared with SAS, interregional i-o analysis also has a second advantage in that it also enables top-down regional forecasts when there are no national forecasts by industry available. As is particularly illustrated in the SSDM model for Indonesia in chapter 3, even a very simple macro projection of different growth rates for national exports and private consumption enables endogenous estimates of regional output growth by sector due to different export/consumption ratios by industry and region, which are available in the interregional i-o model.

What one needs to account for is the trade-off between theoretical elegance on the one hand and empirical applicability and problem solving capability on the other. Chapter 2-4 have clearly shown that interregional i-o analysis has a high score on the latter and as far as the theory is concerned, the capturing of interregional and intersectoral

---

\(^{60}\) In SAS, exogenous national forecasts of economic growth by industry are applied to the regional structure by industry (the "share" factor), sometimes extended with an estimate of the regional deviation from national growth by industry (the "shift" factor). This is still the basic technique underlying the new multiregional forecasting model for the Netherlands covering all 12 provinces, that is currently being developed by the Dutch Central Planning Bureau (Verkade, 2002). Other examples of SAS analysis and forecasting can be found in Anirikopoulos, Brox & Carvalho (1990) and Gazel & Schwer (1998).
interdependence should also be judged as positive. The fundamental question then remains: how bad is it that i-o coefficients are mostly held constant and that the supply side is very weak, if not totally absent? We will deal with both issues below.

**Constant i-o coefficients**

Firstly, there is evidence from the literature that assuming constant i-o coefficients may not be such a serious drawback as one would expect, in particular for short-term forecasting. When long series of (annual) i-o tables are available it is found that the i-o structure gradually changes over a longer period of time (Saywer, 1992). Hewings & Jensen (1988) and Sonis & Hewings (1992) suggest that major differences in the i-o structure between countries are related to their stage of economic development and can serve as taxonomy of different types of economies. The long-term transformation from a dominant agricultural society to an industrialized country with a mature manufacturing sector, and from a manufacturing economy to a service economy, are the main transitions that lead to major shifts in the sectoral interdependencies and consequently in the i-o structure of the economy.

Secondly and more importantly, even if there is coefficient change in the forecasting period, it should be stressed again that it is not regional growth itself that is predicted in chapter 2-4, but only its deviation from the national growth path. If an i-o forecasting error of sectoral growth is related to any kind of technology change, for example a change of energy inputs in a specific industry, this error is made both at the regional and the national level, and thus only appears as an error in the predicted regional deviation when this technology change takes place in one region and not in the other. In a mature and integrated economy like the Netherlands this is not a very likely assumption. Many firms in the Northern Netherlands are local branches of larger (inter)national companies and their technology level will therefore not be different from any other branch of the same company in an other region. In addition, also for independent firms, that do not belong to any larger conglomerate, for the Netherlands there is no evidence that two firms producing the same good in two different regions use a substantially different technology, in any case different enough to assume a different input structure.

Therefore, as far as the ISAM model is concerned, treating i-o coefficients as fixed for short-term forecasting in the Netherlands is defendable. Moreover, the error-evaluation discussion in chapter 2 has shown that the forecasting results are not bad. In the words of Friedman (1955): “the crucial question is not whether coefficients of production are [...] fixed - quite obvious they are not - but whether treating them as if they were yields good predictions”.

The long-term simulation model SSDM in Indonesia (chapter 3) is a lot more questionable in this respect. As mentioned, this model was to some extent born out of necessity in the context of limited staff and budget, and under heavy time pressure. Fortunately, the exogenous national information did not only comprise projected growth paths of macro expenditures, but also of national sector totals. The projected regional
deviations from national/sectoral growth rates, therefore, did not contain forecasting errors at the national sector level due to the assumption of constant coefficients. However, with so many heterogeneous regions in very different stages of economic development, it is very likely that over a forecasting period of 5 years or more, most changes in the economic structure and intersectoral interdependence will be different between regions. In addition, at that time Indonesia was - and in fact still is - going through a major transition from an agricultural to a manufacturing society. This can be a very discrete process, in which some regions with no manufacturing at all see large scale manufacturing complexes emerge in a short period of time\textsuperscript{61}.

It is also needless to say that when regions develop differently this will also affect their interregional trade patterns. The emergence of new industries in a region reduces the need to import their products from elsewhere, but, in its turn, may induce intermediate import flows of inputs needed for those new industries. This effect has been taken into account in SSDM, by endogenizing trade coefficients for industries that already existed in a region according to the original i-o table. To recall briefly: when a regional industry grows faster than its national counterpart, its products are less imported by other industries in the same region. This mechanism, however, does not account for "zero-one-jumps", e.g. the emergence of an industry in a region which was absent at the start of the forecasting period.

Despite all these shortcomings, one important realistic contribution of the SSDM studies has been that the existing interregional and intersectoral interdependencies are shown to substantially moderate the intended effects of regional policy. In particular the dominant position of the industries in the Jawa regions as suppliers of intermediate products for all other regions implies that any (government) expenditure push in the periphery has substantial indirect demand pull effects in the centre as well. The order of magnitude of these effects may have been modelled in a questionable way, but it was an eye-opener for Indonesian policy makers.

Finally, the long-term simulation of the Dutch deconcentration scenario in chapter 4 takes a middle position. Long-term shifts in interregional trade patterns were accounted for, just as in SSDM, and major transitions or zero-one-jumps were not very likely to occur. One component of the deconcentration scenario, however, was the regional shift of supposed footloose and congestion-sensitive industries from the Rim City to the Northern Netherlands. This should lead to a change of the regional i-o coefficients over the forecasting period other than changes according to the national trend, and the model did not take this into account.

*The treatment of the supply side*

In ISAM supply is supposed to be perfectly elastic. As far as labour is concerned this cannot be a very serious drawback when forecasting short-term deviations from a

\textsuperscript{61} A good example is the fast development of the fish canning industry in North Sulawesi. See Steider (1995a).
national growth path of 2% to maximal 4% per year for a region with relatively high unemployment. Apart from the demand for labour of specific skills and education, during the years that annual forecasts were made with ISAM, the North had no labour supply constraints at the macro level. It is, therefore, only the supply of capital and intermediate products that may serve as a regional supply constraint on the modelled demand-driven regional growth.

For the small open regions modelled in ISAM the assumption of perfectly elastic supply of intermediate inputs can also be defended. The province of Groningen, for example, has a share of 3% of total national GDP and around 70% of its intermediate inputs is imported, either from other Dutch regions or from abroad. The same holds for the supply of investment goods of which 65% is imported into the region. What may be a more serious deficiency is the neglect of production capacity utilization rates. Although there are no data on this subject, it is likely that the utilization rate of production capacity is lower in regions that are lagging behind. The resulting forecasting errors will thus be different by region.

For long-term forecasting, like in SSDM, the situation is the other way around. Capital utilization rates at the start of the forecasting period are not very relevant for 5-25 year projections. The supply of intermediate and investment goods, however, definitely can not be assumed perfectly elastic because, contrary to the biregional case of one small region versus the rest of the country in ISAM, SSDM covers all Indonesian regions.

A positive feature of SSDM, which is also used in the deconcentration study in chapter 4, is that the demand effects of investment itself is taken into account and - more important - the results of the demand-driven model are constrained by the growth of the capital stock. The application of this mechanism in chapter 3 and 4, however, does not allow for an empirical check on its accuracy. In a new future version of ISAM it will be used for short- and medium term forecasting, which will enable a test on whether this limited supply modelling does in fact reduce forecasting errors.

The CGE alternative

In theory, all model deficiencies of interregional i-o analysis just mentioned are "solved" in a CGE model. A fully fledged CGE model has production functions, variable coefficients, substitution elasticities between products and thus a more realistic interaction between supply and demand. The rise of CGE modelling in the 1980s has been followed by a steady growing literature on many different types of implementation, including regional and interregional ones (Partridge & Rickman, 1998). Regional macro forecasting, however, is a very rare CGE application (McGregor, Swales & Yin, 1998). The major part of CGE literature deals with policy analysis where the equilibrium in a base year is "disturbed" by a single isolated variable, such as a tax reduction.

In macro economic forecasting not one but many exogenous variables change at the same time and this makes the change from one equilibrium into the next intractable. In
addition, apart from the i-o tables, much more empirical information is needed which is simply not there, in particular the intersectoral substitution elasticities. Many CGE models in fact keep i-o coefficients constant using a two-step procedure in which a Cobb Douglas production function determines the input of capital and labour in a first step, followed by a distribution of capital input over supplying industries according to the i-o table in the second step (Bröcker, 1995). Finally, one may seriously question the empirical relevance of the CGE assumption of regional price clearing for the Netherlands. Most consumer products are marketed and distributed by nationwide operating companies using national advertising media without regional price differentiation. Also wages hardly have any regional component due to national labour union bargaining by industry.

Despite these considerations a first version of an interregional CGE model is recently developed covering all Dutch regions, but it is aimed at policy simulation applications like regional impact of infrastructure investments (Thissen, 2004).

7.2 New economic geography

As Krugman and Fujita themselves have repeatedly admitted, compared with the existing literature on economic geography there is nothing new in the idea of a joint model of economies of scale and transportation costs. Their main contribution is the rigid mathematical model formulation and the use of the monopolistic competition framework as its "workhorse" (Fujita, 2000).

The NEG approach was originally introduced for an abstract two-region case, but it is actually agglomerations and cities that are the object of study in the NEG literature, although this is not always stated explicitly (Krugman, 1993; Brakman e.a. 2001). Henderson even considers the two-region variant as the regional core-periphery model and the multiregional variants as urban models that "try to better capture interactions across cities, as opposed to regions" (Henderson, 2004, p285). It focuses, therefore, on a lower spatial scale then interregional analysis, and distance and transportation costs play a greater role.

The simulation results with operational NEG models, however, are still not very convincing. In spatially neutral models, there is a strong tendency of all cities collapsing into one agglomeration (the "black hole" result), even when the parameter set is only slightly changed. The introduction of congestion costs can serve as a countervailing force (Brakman et al., 1996; Kubo, 1995) and there are other non-geographical model extensions that lead to more dispersed agglomeration results, like the introduction of a housing market (Helpman, 1998; Suedekum & Pflueger, 2004) or restrictions on the interregional mobility of labour (Puga, 1999). In addition, as was demonstrated in chapter 6, the non-mobile agricultural sector remains a counterforce for the model's centripetal tendencies and when larger parts of the economy are treated as non-mobile the results will also become more realistic. The deconcentration scenario study in
chapter 4 has given an example of how in a multisector specification one can think of each industry being footloose to a certain degree. The aggregate total of all non-mobile fractions per sector can then be treated as the "agricultural" sector of the NEG model (Stelder, 1995c). Finally, when more sectors are introduced, the NEG model can incorporate backward i-o linkages and intermediate flows that also lead to fine-tuning of the agglomeration results (Stelder, 1995c; Venables, 1996; de Vaal & van den Berg, 1999; Knaap, 2004).

In this book, however, only the simple base model of Krugman (1993) is examined on its behaviour in a non-neutral spatial context. The most important result of this exercise is that, just like the non-geographic model extensions just mentioned, the mere introduction of geography itself leads to very differentiated partial agglomeration. When geography is introduced, in other words, when the locations are not spatially neutral, the black hole problem becomes much less dominant and many "black hole" parameter configurations in fact produce only partial agglomeration. The geographical NEG-model presented in chapter 5 and 6 should be seen as the mirror prototype of the geographically neutral models. In stead of using a historical initial agglomeration distribution in a spatially neutral model in order to prove that "history matters", we now have a geographical model with no agglomeration as the initial distribution showing that "geography matters".

The combination of less centripetal variants of the NEG model with the geographical grids used in this book will of course lead to other, and more spreaded city configurations then the ones presented in chapter 6, but there are still important limits to the types of partial agglomeration that a geographical NEG model can handle. What remains impossible, for example, is a model result with more large cities very close to each other, like the Midlands in the UK, the Rim City in the Netherlands and the Ruhrgebiet in Germany. On the larger European scale of the model in chapter 6 these multicity structures always collapse into one agglomeration. However, as an anonymous referee has stated, this may not be considered a major problem because such a multicity structure in fact is one large agglomeration in the sense that the assumed external economies exist on this larger scale.

As is shown in the simulations, large cities appear roughly in a central place pattern more or less equidistant from each other unless there is a major geographical barrier between them with high transportation costs, which is precisely what one would like to see as a model result. This model property is clearly very useful for policy analysis, such as the simulation of the spatial effects of new infrastructure. In order to do that, however, we would need to introduce "history" again: like in traditional CGE analysis, the geographical NEG-model must then be able to reproduce the existing agglomeration structure as a starting "benchmark" equilibrium, which is then to be compared with the new equilibrium with the new infrastructure. Another model option would be to calibrate the model with a benchmark equilibrium earlier in history (e.g. 1500) with sea and inland waterways as the main transportation mode. It will then be possible to examine whether
the model can correctly predict the historical agglomeration shifts as a result of the introduction of rail and road transportation. These types of model calibration are not carried out in this book because the main purpose of the simulations was to isolate the effects of the geographical dimension. The evaluation of the simulations has shown that the effect of geography on the agglomeration structure is clearly there, although there still is no definite and final assessment of the extent to which geographical factors determine the existing agglomeration structure. As is discussed in detail in chapter 6, using the same model grid for reality and prediction is a first priority for future research and other geographical factors like climate, soil fertility and a more realistic treatment of the effect of sea transport and inlands waterways are necessary to reach a more firm conclusion.

7.3 Some concluding remarks

From the policy point of view one of the main contributions of the NEG approach is that self-reinforcing processes can be very important. Large economic centres have the tendency to continue to grow at the cost of the periphery, and infrastructure policy aimed at connecting this periphery to the main centres can lead to the unintended effect of further concentration and agglomeration. Modern models of regions, cities and their interaction should take account of these effects but it is not yet clear whether this is possible in an i-o/CGE framework.

The common denominator of both interregional i-o analysis and agglomeration modelling is the breakdown of an economy at the macro level into components across two dimensions: industries and space. The spatial dimension in agglomeration modelling can be at the detailed level of cities, but information about intersectoral interdependence is usually only available at more aggregated regional levels. There are efforts to achieve convergence of both approaches, from i-o towards CGE with economies of scale (Bröcker, 1995), and from simple NEG towards more differentiated multisector models (Venables, 1996). Recent models that try to actually integrate both methods and their different data levels, for instance by applying the i-o structure of a region to all its cities, are becoming operational, but their calibration is difficult (Bröcker e.a. 2002; Knaap, 2004; Thissse, 2004).

It is a simple law of statistics that at each lower spatial level the spatial units become more heterogeneous and less stable in time. Applying regional i-o structures to cities is therefore problematic, in particular when it comes to endogenized inter-city trade flows for which no data are available at all. The reverse alternative of extending the agglomeration approach to larger regions, however, is also questionable because the major effects of external economies of scale will most likely be at the lower level of actual cities.

Given the experience with both model types, a better alternative could be to use interregional i-o analysis for simulation and forecasting of the region at the macro and
meso/sectoral level without any top-down assumptions about (inter)urban i-o interdependencies. The regional results could then be used as aggregated constraints on the outcome of an agglomeration model that tries to predict the growth of the main urban centres. The aggregate employment growth by city can then be estimated using the recent and detailed available data on the employment structure for each city, which in fact boils down to predicting the "share" factor in the SAS approach mentioned above. It is then the remaining "shift" factor for each city that could be projected by a separate agglomeration model. Some inputs that seem to be implementable for that model in the nearby future are projected infrastructure investments and the development of the housing market following the approach of Helpman (1998).

The need for introducing lower spatial levels in interregional models is clear, also from the policy point of view (Oosterhaven, 1996), and integrated i-o/NEG modelling can be a valuable tool of analysis in this respect. One of its main challenges in the nearby future will be to find a way for consistently using both approaches for the spatial level they are designed for: the region and the city.