Netherlands Official Statistics

Volume 14, Winter 1999

Voorburg

Prinses Beatrixlaan 428
P.O. Box 4000
2270 JM Voorburg (Netherlands)

Telephone : . .31 (070) 337 38 00
Fax : . .31 (070) 387 74 29
E-mail: lhka@cbs.nl
Internet: http://www.cbs.nl

Heerlen

Kloosterweg 1
P.O. Box 4481
6401 CZ Heerlen (Netherlands)

Telephone : . .31 (045) 570 60 00
Fax : . .31 (045) 572 74 40

Key figure A-125/1999

Subscription: Dfl. 42.00 per year
Price per copy: Dfl. 20.00

ISSN 0920-2048

Postage will be charged.

Contents

Structural Business Statistics: a key role in the system of economic statistics, the Dutch case

Symon Algera

1. The Dutch system of economic statistics

1.1 Background

Fifty years ago the various Dutch economic statistics were a rather chaotic collection, each with their own aims, background, population, definition etc. Until the introduction of the National Accounts in the 1940s there was no conceptual framework to judge the consistency, comprehensiveness and co-ordination of the economic statistics.

Two crucial factors helped evolve the various Dutch economic statistics on the basis of the development of the National Accounts. First the fact that most of the statistical work is organised in a single centralised office, and further that the National Accounts are compiled in that office.

Since then the various economic statistics have evolved into a coherent system. This section contains some general remarks about this system.

The system of economic statistics is based on the view that the economic process is a circular flow-of-income system. Income is generated in the production process and distributed among economic actors in the income distribution process. The current spending power of the actors is then increased or reduced by the financing process. The resulting funds are spent in the expenditure process on goods and services produced in the production process, thus closing the cycle. This cycle is, however, not truly closed. There are two types of flows into and out of the system as a whole. The first set of these is international: the flows to and from the rest of the world in each of the four processes. The second type of in - and outflows is intertemporal: the value of additions to and withdrawals from stocks of goods and other assets.

As a consequence of this view, the Dutch basic system describes four processes: production, income distribution, financing and expenditure, as well as their international and intertemporal relations. The production process can be considered as the heart of the economic process, since it generates the income and the goods and services that are distributed and allocated in the other processes. Put differently, without the production process there would be no income distribution, no financing process, nor an expenditure process. For this reason, the description of the production process in the Netherlands traditionally occupies a major position in the collection of economic statistics. The main source for this description is the Structural Business Statistics (SBS). These annual statistics are supplemented with the system of short-term indicators containing quarterly and monthly statistics and some expectation indicators.

1.2 Institutional versus functional approach

In principle the Dutch basic system is an institutional one. This means that within each process the system focuses on the actors rather than on the phenomena of the process. The latter approach is referred to as functional. To clarify the difference between institutional and functional statistics, let us take the simple example of the statistical description of the production of printed products. In a functional approach, a complete coverage of the production process would require all printing output to be covered by a series of production statistics on printing, irrespective of who produces the printed products. In contrast, an institutional approach would yield a series for printing companies, whose main activity is printing. It would provide a breakdown of printed products into a number of distinct products and would also show the printing companies' output of non-printed products.

This example clarifies the different meaning of “completeness” in the two approaches. In functional statistics, completeness means complete coverage of the phenomena of a process (output of printed products), while in institutional statistics it implies complete coverage of all the actors participating in a process.

The institutional approach has two major advantages. First it is more practical. Consider again the printing example: in a functional approach, total production of all printed products has to be covered, irrespective of the producer. This could be achieved by asking every producer in every industry sector how much of each type of printed product he produces. Obviously this would be impractical in many cases. In institutional statistics on the other hand, the method of observation is quite straightforward: the actors are simply asked to specify their major activities.

The second major advantage of institutional statistics is analytical. As institutional statistics are summaries at the meso level of the behaviour of economic actors, these data can be used immediately for the analysis of economic behaviour. This is mostly why the basic system of economic statistics usually employs the institutional approach. This does not mean that it contains no functional statistics at all. Statistics on foreign trade in goods are functional because the focus is on the international flows of (highly disaggregated) groups of commodities and not on the actors that export or import them. Recent developments with respect to Electronic Data Interchange (EDI) may increase the use of functional statistics.

2. Co-ordination

Institutional statistics take the perspective of the actors whose economic transactions they describe as their point of departure. Since perspectives and viewpoints may differ from actor to actor and from industry to industry, a radical institutional approach would lead to statistical anarchy if there were no countervailing force to provide order. This countervailing force is co-ordination: co-ordination is what cements the individual economic statistics into a coherent system.

There are three steps in the co-ordination process for institutional statistics. The first step is to define the actors that are to be used as statistical units. The second step in the process of co-ordination is drawing up classifications of actors and goods and services and the third step is the formulation of definitions for the variables measured (for an elaboration see*: The industrial statistics of the Netherlands, 1992*).

3. Annual statistics – Short-term indicators

Generally, the most detailed and complete observations deal with annual data, which are normally available about one year after the end of the reporting year. In some cases these structural data are collected less frequently. The system of annual statistics is supplemented by short-term statistics: quarterly and monthly statistics, and business sentiment indicators, the latter based on opinions and expectations. The quarterly statistics are less detailed than the annual ones, but are available earlier: one or two quarters

after the end of the quarter under review. The monthly statistics are available even earlier, but again are less detailed.
Thus, in the case of quarterly and monthly data there are two time dimensions: the length of the reporting period and the speed with which the data become available. In the case of expectation indicators, the emphasis is on the latter dimension: they provide approximate indications of current and future trends expected by the actors themselves.
The choice of detail of the short-term statistics partly depends on the potential variability of the phenomena concerned; the purpose of the short-term statistics in the basic system of economic statistics is that they provide, together with annual statistics, a comprehensive and timely picture of the economic process. In manufacturing, rather detailed short-term information is needed for an adequate picture of economic trends. In all cases, the short-term statistics are institutional whenever possible and co-ordinated with respect to both actors and definitions. This way, they are conceptually fully consistent with the annual data.

4. Integration in National Accounts

The National Accounts describe the economic process as a whole, that is: they describe not just all processes of the circular flow-of-value system but also the relations between these processes. The description of the Dutch economy in the National Accounts has to be complete, so the results of the economic behaviour of all the actors must be incorporated. Also all relevant phenomena must be dealt with. It is important in this respect that the National Accounts are the base for the systematic approach of the economic statistics.
To construct the National Accounts from the different economic statistics the results of many separate statistics, each describing only a part of the economic process, must be integrated. This process of integration has – in principle – three components: *standardisation, completion* and *confrontation and balancing*.

5. Matrix presentation

The coherence between statistics which focus on the relationships between SBS and SI on the one hand, and between SBS and NA on the other can be represented in a “statistical matrix” (for elaboration see Algera 1995).

The matrix dimensions are *time* distinguishing between statistics relating to future, monthly, quarterly and annual data, and *degree of integration:* single, combined or integrated.
The column headed “single” contains statistics which are obtained from a single survey of statistical units. The data are simply the survey findings. In the “combined” column the statistics are the outcome of combining different statistics (and different surveys). An example of statistics of this kind is the deflation of turnover (from one source) using prices (from another).

Figure 1
“Statistical matrix” (an example for manufacturing)

	Degree of integration		
Time	**Single**	**Combined**	**Integrated**
Future	Tendency Survey	Producers confidence index	
Month	Turnover	Production index	
Quarter	Prodcom		Quarterly Accounts
Year	S.B.S.	“Indusry Report”	National Accounts

The “integrated” column contains data that are the outcome of an integration process involving detailed checks of all available information, with the original statistics being adjusted as necessary. The National Accounts are an example of integrated statistics.
We can use such a matrix, to help explain some relevant aspects of coherence. Going from the top to the bottom of each column, the report becomes more reliable and detailed, but less timely. Crossing the rows, from left to right, the data become more comprehensive and reliable, but as a rule again less timely.
It is policy at Statistics Netherlands to place emphasis on coherence in statistics. In this respect, it is worth highlighting the specific character of a central statistical office. The main strength of central statistical offices, and major international institutions lies in the fact that statistics can be presented from within a central coherent system. The underlying statistics gain from this in practical value. Interrelationships can be established, provided the various statistics can be linked together in terms of basic data. In this respect central offices differ from other institutions. A matrix of the kind described above can be a valuable instrument in developing a system of statistics. It will be possible, on the basis of specific selection criteria, to assess for each variable whether it is feasible and meaningful to fill in a particular cell in the matrix. The standards for reliability and timeliness are key factors in arriving at this judgement. A policy decision to obtain more statistics in the “year/combined” cell was central to a reorganisation of Statistics Netherlands some years ago.

6. Relationships between Structural Business Statistics and short-term indicators

6.1 Features of short-term indicators

Short-term indicators can be regarded as the predictors of SBS. This means that for the whole system of statistics the definitions of variables and the used classifications need to be co-ordinated. (Which is easier to organise if the production of statistics is centralised in one institution).
Monthly indicators are not combined in a statistical sense and they are not systematically checked for consistency (Algera and Janssen, 1991). Furthermore, they are not balanced, as is the case in the description of the production process in the quarterly and annual accounts, where supply and use tables are compiled. There are both conceptual and practical reasons for not doing this on a monthly basis. An important feature of short-term indicators compared with SBS is that they provide a much earlier indication of the development of a certain variable than annual statistics. On the other hand they are less detailed and less accurate. Short-term indicators are restricted to specific subjects of the economic process. Several aspects are taken into account to determine which subjects merit early economic indicators.

6.2 Selection criteria for short-term indicators

For the question of what should be measured, first of all the variability of a (macro-) economic phenomenon can be pointed out. For example, short-term data on consumption of fixed capital are not very meaningful. If these figures nevertheless are needed (e.g. for completeness), it is clearly relatively easy to estimate them.
A second selection criterion for short-term indicators is the degree of implications for an economy of a change in the variable to be measured: in business terms, a decline in exports of goods and services is much more serious than a decline in exports of financial capital.
The extent to which a short-term indicator is really indicative for a variable to be measured is also important in the selection. An example of this is prices of imports and exports. “Unit values” as a direct result of customs statistics can be very indicative for these

prices. For goods which vary greatly in quality, however, the changes in “unit values” differ strongly from price changes.
A fourth and last (but not least) selection criterion is the significance of a short-term indicator in the whole description of the economic process. In other words, the indicator has to fill a gap in the statistical description of the economic process.
For the question of how to compile short-term indicators, the main requirement is that they have to be available quickly. This means that indicators must be based on information that can easily be provided by respondents. So: relatively simple questions on (for the respondent) simple variables and, importantly, the number of questions has to be kept to a minimum.
The second aspect in this respect is that it must be possible to compile the indicators continuously. There is little use in basing an indicator on a sample from a variable population, if the representativeness of the sample cannot be continued, as the indicator then does not describe the real development. The same holds for indicators based on fixed ratios that cannot be updated in time or indicators that are biased precisely in times of important business cycle changes. Nothing is more frustrating than an indicator that is found wanting at the moment really interesting developments occur.

6.3 Consistency (ex-post)

A specific aspect of reliability concerns the ex-post consistency of the data. If it is accepted that short-term data are less reliable than those that become available at a later stage, a decision must be made as to whether- for reasons of consistency- the short-term figures should be adjusted to these more definite figures. This is done in the Netherlands for macro-economic statistics, but not for most others. There is a dilemma here. If a monthly figure must always be consistent with a “later” figure, which may be independent and superior, then, for example, the monthly household consumption expenditure index should always be adjusted to the publication of a new set of Quarterly National Accounts (QNA). This is what is done in the Netherlands and what some users want (not two different sets of figures on the same subject). On the other hand, for understandable reasons, users also want stability in statistics. It should be explained to these users that “adjustments” are a logical consequence of the choice of a particular package of statistics. If people want rapid statistics, they must also accept adjustments.
The recently developed “one figure” philosophy of Statistics Netherlands argues in favour of adjustment.

7. Relationships between Structural Business Statistics and National Accounts

7.1 Background

The National Accounts describe the economic process as a whole. They are constructed from many different economic statistics, of which the SBS are certainly one of the most important, perhaps *the* most important.
In principle, the process of integration in the National Accounts has three components.

– *Standardisation.*
  While separate statistics may employ different definitions for similar concepts, the National Accounts are based on standardised definitions. An important part of the integration process is the adjustment of the data for differences in definitions. The same applies to the differences in classifications and in the degree of detail employed in the various underlying statistics.
– *Completion.*
  The individual statistics taken together often do not provide a complete description of the economic process as a whole. Statistics on parts of the process are lacking; other statistics may not provide complete coverage of their subject. In the National Accounts these gaps are closed by means of estimates on the basis of all kinds of ad hoc information.
– *Confrontation and balancing.*
  Conceptual relations may exist between many variables in different statistics. Thus total production and imports of electricity equals total national uses plus exports and transport losses. In the National Accounts these conceptual relations are made explicit and the data from different sources are confronted on the basis of these explicit conceptual relations. Generally speaking, this process of confrontation yields data of a better quality than those in the separate underlying statistics.

In the course of time, there have been shifts in the relative importance of these three components of the integration process in the Dutch National Accounts. Originally the emphasis was on standardisation and completion: there was no system of economic statistics yet, so standardisation featured large in the integration process. Similarly, a lot of completion had to be done because of the many gaps in the underlying statistics which were partly caused by the lack of a system. But as indicated earlier, the presence of a system of National Accounts provided the impulse needed for the systematisation and complementing of the body of economic statistics. Eventually this led to the emergence of the basic system, and the subsequent growth of this system led to the availability of many new and better statistical data. It also prompted a change in the nature of the integration process. The basic system is, in principle, fully co-ordinated with respect to actors and phenomena and is, again in principle, complete. Consequently, standardisation and completion have become less important elements of the integration process and confrontation and balancing are now the most important factors. The compilation of supply and use tables play a key role in the integration process.

7.2 Basic features of the integration process

Timing and contents of the Dutch National Accounts
At Statistics Netherlands three estimates of the National Accounts data are compiled every year (on the basis of the Quarterly Accounts three more annual estimates are available for only a limited number of macro-economic variables: T + 0 months, T + 2 months, and T + 4 months). The three annual estimates are scheduled as follows:

- T + 6 months: first estimate (“preliminary”)
- T + 17 months: second estimate (“improved preliminary”)
- T + 27 months: third estimate (“final”)

These estimates contain a complete set of National Accounts data:

- supply and use tables in current prices and in prices of the preceding year;
- I/O tables (industry-by-industry) in current prices and in prices of the preceding year – in producers’ prices and basic value;
- sector accounts for all main sectors including financial accounts.

Social accounting matrices (SAM) and National accounting matrices including environmental accounts (NAMEA) are compiled for the “final” year. Balance sheets are currently only compiled on an experimental basis; the same applies for homogeneous or product-by-product I/O tables.
The final estimate of the supply and use table consists of around 250 industries by 800 product groups. Naturally, because of confidentiality regulations not all data can be published for a broad public. After application of the Dutch confidentiality rules a supply and use table of some 150 activities and 600 product groups is available. The corresponding activity by activity I/O table is available for around 150 activities.
The preliminary estimates are made with 100 industries and 250 product groups. This is very close to the publication level of the supply and use table and I/O table.

Sources and units
At Statistics Netherlands, all source statistics for the supply and use tables are “institutional statistics”. This means basically, that Statistics Netherlands surveys enterprises as they present themselves to for example the tax authorities. The main exceptions to this rule are the larger companies, which are often structured in a complex or diverse way. Such companies are asked to create special statistical units which often correspond to their business units. This implies that an industry classification in S&U and I/O tables in the Netherlands should be interpreted in the institutional way: they describe establishments that are as homogeneous as possible within the limits of what is reasonably possible from a survey point of view. As a result, in Dutch S&U tables output of most industries consists of a main product (or products) and a number of secondary products that are not always directly related to the main product.

Working procedures
The working procedures of the compilation of supply and use tables at Statistics Netherlands can (chronologically) be summed up as a column - row - column scheme.

A. Columns: Input from specialists
The data received from source statistics are made complete and consistent with the level of detail of the reporting year. This work is done by NA experts (referred to as “specialists”) who are each specialised in a group of industries. They are responsible for the necessary adjustments to meet NA definitions and to estimate the “white spots” not covered by source statistics. Furthermore, specialists are responsible for a number of additional estimates, as the source statistics never contain all the necessary details. Outputs and inputs are deflated separately using prices from a central price database in which price data are stored on foreign trade, producers’ prices and consumer prices. For the services part, prices of inputs or other indicators are used for output prices; these prices are compiled by NA specialists.
The input from specialists (on production and uses by industry, on final use components and on foreign trade) in the main automated integration system are the columns of the supply and use tables. After introduction of the data in the system, the data are then checked again by the specialist.

B. Rows: The integration process
At the start of the integration process, the automated integration system contains a full description at product and activity level of the year under compilation in current prices and in prices of the year before (800 product groups and 250 industries in the final estimate). The data set also includes the corresponding set of data of the preceding year in current prices.
The integration process is based on the balancing of the rows of the supply and use tables. During the process, data can only be approached row wise; the columns are “locked”. Product groups are aggregated into about 200 “statistical groups” of related product groups. These statistical groups are attributed to an integrator: the only person allowed to change the data of his/her statistical groups and the underlying product groups.

The integrator now manually balances every product group by making supply and use (or better: sales and purchases, in order to avoid valuation problems with changes of stocks) equal - both in current prices and prices of the previous year. Large discrepancies between supply and demand of a product are analysed and discussed by the integrator and the most relevant specialists.
The programme allows for automatic balancing of a product group or even a statistical group, but these facilities are hardly used. No statistical discrepancy between supply and demand is left after balancing. The integrator normally does not change data on domestic output, but finds solutions by altering data on imports or intermediate and final uses (including stocks). This integration method implies that GDP according to the production method and GDP according to the expenditure method are made equal.

One of the consequences of this method is that value added by industry or total imports/exports or final uses can and will be changed from the data from specialists used as input. Wherever this leads to “unacceptable” changes to the input data, a third step is necessary.

C. Columns: Checks and “repairs”
As value added and the input structure of industries can be changed in the second step, the results are checked by the specialists to see if they are acceptable. If not, data are changed to accommodate the wishes of the specialists. In most cases, these changes are only minor.
Although the description of the integration process may give the impression of being a very lengthy and labour intensive operation, this is not the case. The balancing of the final estimate (S&U and I/O tables) takes about two and a half months and involves six to seven people full-time. The other two estimates are finished within four to six weeks each. The preparation of the inputs by specialists takes about the same time per estimate. At the National Accounts department, around twenty people are involved.

Compiling Input/Output-tables
Balanced S&U tables give information about value added per industry, the input-output structure per industry in terms of products and, of course, the major macro-economic figures like GDP, consumption etc.
S&U tables, however, do not give information on the input-output structure of the economy in terms of industry by industry. One of our main users, the forecasting agency of the Dutch government (Netherlands Bureau for Economic Policy Analysis) uses I/O tables of the industry-by-industry type in their forecasting models.
These industry-by-industry I/O tables are derived from the S&U tables in the following way. To start with, a complete I/O table is compiled for each product group. As only limited information is available on the relationship between producing units/imports and the users (intermediate or final) in many cases a proportional distribution is used. Of course when information is available this is used as a starting point. In general, there is no manual balancing process; the matrix is adjusted by applying a mathematical programme based on a Lagrangian adjustment method.
This procedure results in I/O tables for each product group. Adding them up gives the national industry-by-industry I/O table.

7.3 Quarterly National Accounts

Quarterly National Accounts (QNA) are also compiled by means of supply and use tables. The general method is more or less the same as for the annual accounts, although because of a lack of information the supply and use tables for the quarterly accounts are much more aggregated.
Moreover the compilation process of QNA can be described as an “extrapolation process” (Janssen and Algera, 1988). The basis of the compilation is the supply and use tables of the corresponding quarter of the previous year. These tables contain all kinds of structural information.

The calculation of macro-economic figures of the quarter under review extrapolates from these supply and use tables of the base quarter with volume and price indicators, and supplements these with information on levels from direct sources. Where no trend indicators or independent data are available for a particular field, assumptions are made. The values thus obtained are then all integrated into supply and use tables for the quarter under review. The integration process employs all kinds of plausibility tests to ensure that the original estimates are mutually consistent.
The trend indicators used in this process are often the same short-term indicators discussed in the previous section. Clearly then the quality of the QNA (assessed by a confrontation with the later annual accounts) is determined by the quality of the short-term indicators (assessed by a confrontation with the later SBS).

7.4 The role of Structural Business Statistics in the compilation process

The basic source in the compilation process of the NA, as described in the previous sections, is the SBS (Gorter et al., 1990). The "specialists" treat data drawn from the SBS with great detail. Plausibility checks may involve contacting the primary departments or even the survey respondents. The production statistics generally refer only to companies with more than ten employees (or, more accurately, man-years). Some specific data are requested only from companies with more than fifty employees. For the purposes of the National Accounts these data have to be grossed up to provide figures for the relevant economic class as a whole. Differences arising from such factors as reclassification of units, changes in the question etc. must be offset to ensure that the data are consistent with those from earlier periods. Subsidiary activities, particularly trade, must be separated out. Other adjustments concern the use of producers' prices for intermediate consumption, corrections for fraudulent returns, etc.
The branch specialists compile a number of statements to be used in the subsequent integration process. The statement of output comprises a number of data concerning a given product group such as output value, stock changes, various margins, exports of the product group concerned, and the relevant deflators. The raw and auxiliary materials statement contains similar information about intermediate consumption. Such statements are also compiled for economic classes for which no production statistics are available. Here we have to fall back on other kinds of information.
All this information is the starting point for the integration process. After this integration process it may appear that the initial information has to be changed because it did not fit with other information. Again this may lead to a feedback to the departments responsible for the SBS.

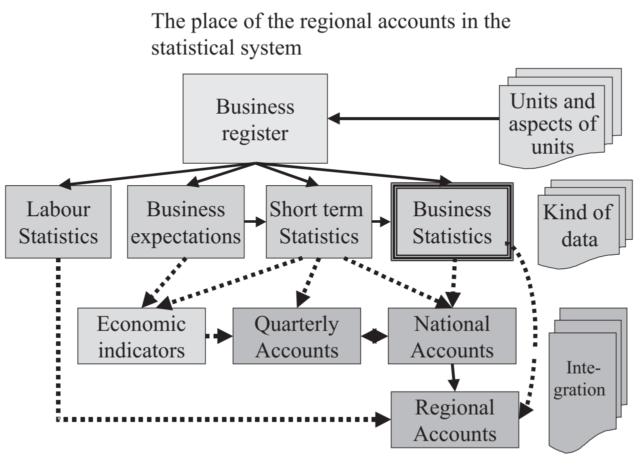
7.5 Organisational issues

The SBS and the NA fall under the responsibility of different divisions within Statistics Netherlands. The three components of the integration - *standardisation*, *completion* and *confrontation and balancing* - all used to be carried out within the NA division. Nowadays there is a trend to move standardisation and completion tasks to the department responsible for the SBS (the subject-matter department), as in recent years most of the information needed for these activities becomes available within these departments. The main task of the NA department is confrontation and balancing and making the necessary adjustments to the various information sources. Nevertheless there are some elements of completion and standardisation which still belong to the responsibilities of the NA department, for example estimations on fraud and illegal activities.
This organisational issue can be illustrated with the help of the matrix presented above. The "single" column used to belong to the subject-matter department, and the other two columns "combined" and "integrated" to the NA department. Nowadays there is a trend to shift parts from the column "combined" to the task field of the subject-matter departments, giving these departments more and more responsibility for the standardisation and completion steps. In the last major reorganisation of Statistics Netherlands it was decided to place more emphasis on making this kind of statistics, which are in a sense a bridge between the SBS and the NA.

7.6 Regional accounts

Like the national and quarterly accounts, the regional accounts are also part of the statistical system (De Vet, 1999). The regional accounts are a regional specification of the comparable national accounts, using the same definitions, classifications and sources as the national accounts. However, there are in principle two extensions of a specific concept of the national accounts. These refer to the allocation of transactions to a regional economy.
Firstly, in the national accounts, economic transactions are allocated to the actors that reside in the economic territory of the country. Using the same concept to allocate the transactions to the territory of a region poses the problem that the total of the regions does not cover the total economic territory of the country: transactions of embassies, military bases and platforms on the continental shelf cannot in principle be allocated to one of the regions that constitute the geographic territory of a country. For this reason, ESA extends the national definitions with the so-called "extra regio". This extension has a limited influence for business statistics. Only for private industries which have a unit on the continental shelf should transactions be allocated to the extra regio; for example, oil extraction. In all other cases, private industries are situated within the geographic territory of the country if they have a residence in that country.
Secondly, practical problems are posed by the type of production units used. The European System of Accounts (ESA) recommends the use of the local kind of activity unit (local KAU). However, because of the response problem Statistics Netherlands uses the KAU as a production unit, which introduces the existence of the so-called multi-regional KAU, i.e. a KAU with two or more regional KAUs in different regions. Whether a KAU is multi-regional depends on the size chosen for the regional classification. The existence of the multi-regional KAU implies some form of regionalisation of the transactions of the KAU.
Depending on the sources, two methods are used as the basis for the regionalisation: the production and the income method. The production method basically requires data from the business statistics, for instance production and intermediate consumption. The income method is based on data on the components of value added: wages, social contributions, production-related taxes and subsidies, consumption of fixed capital and operating surplus. The labour statistics are the main source. At the national level, the latter method is principally used for non-market services, like general government. In the regional production accounts of the Netherlands both methods are practised.
The methods for regionalisation vary, depending on the available sources, from bottom up to top down methods. These methods refer to the regional breakdown of the national data. Briefly speaking, a bottom up method can be used if all available required data enable a regional breakdown, and the data add up to the corresponding data in the national accounts. However, this is rarely the case. The top down method makes use of regional indicators for the regional allocation of the variables in the national accounts. An example of a regional indicator is the number of jobs of a KAU in a local KAU or in a municipality. Other indicators refer for instance to wages paid at the local level. In practice, regionalisation is a mix of the two methods, using both production and income methods.

Figure 2
The place of the regional accounts in the statistical system



The regional accounts in the Netherlands are principally based on linking business and labour statistics at the level of the KAU. Thus elements of production statistics are linked to the regional elements in the labour statistics at the level of the KAU. The reliability of the regionalisation depends on the homogeneity of the production activities in the local units of the KAU in relation to the regional indicator.

From the point of view of the National Accounts, but also from the point of view of the regional accounts, a coherent system of statistics (business register, business statistics and labour statistics) is important. Within this framework, business statistics play a key role.

8. Recent developments

8.1 Electronic Data Interchange (EDI)

The term EDI is normally used for the technical link between two information systems by means of data communication (de Bolster, 1997). At Statistics Netherlands we use the term EDI more for the main link: the link with the source, and we therefore distinguish two types of EDI, based on different types of source:
– primary EDI: direct from the original source.
– secondary EDI: the use of data already collected by another institute (e.g. tax authorities).

The two types of EDI require different approaches. Using primary EDI we have to make connections with a large number of different administrative systems, which requires software that supports a flexible link. On the other hand, secondary EDI usually comes from the administrative system of another data collector, so the link can be tailor-made. The main problem in this case is that the information has already been “translated” for the purpose of that data collector, which often presents problems with the “translation” process at Statistics Netherlands. As these external registrations are usually not set up for statistical purposes, they use their own definitions, units and classifications, and data verification will be focused on the original purpose of the registration. Raw data from these registrations therefore need cleaning up and adjusting to statistical definitions and classifications before they can be used as input for surveys. As it is further not possible to return to the original source to obtain additional information in the case of errors or missing data, it will involve a lot of work at Statistics Netherlands to transform these data into statistical information. But if we succeed in doing so, we shall be able to minimise the administrative burden for the respondents.

Primary EDI does not really affect the system of economic statistics as described in this article.
However, secondary EDI, which has been used with quite some success in the Netherlands, does affect the system, and in two ways. Firstly in some cases the principle of a pure institutional approach is left because a certain *phenomenon* is registered and not the institutional actor.

Secondly, because what we get from registers is often not exactly what we want, we have to make adjustments to the definitions and classifications used at Statistics Netherlands (otherwise we may have overlaps or gaps or we may count apples in with oranges).

The increase of secondary EDI means an increased effort in the standardisation and completion phases in the integration process.

8.2 Quality assessment

As indicated in the previous section there is a clear coherence between statistics. With respect to the SBS there are two types of linkages. One is the linkage *short-term indicators - SBS* and the other *SBS - NA*.
Assuming that this statistical coherence exists, the reliability can be assessed in two ways. First in terms of “predictive power” (SI - SBS) and secondly in terms of “consistency” (SBS - NA). In this way it is possible to get quantitative measures of quality. For example: the monthly turnover index indicates an annual growth of 4.2% and the SBS indicates a growth of turnover for the same year of 3.8%, in this case the turnover index failed by 0.4 of a percent point.
Recent policy at Statistics Netherlands places more emphasis on this kind of quantitative quality measurement and aims to report on this in publications. In this approach emphasis is put on the *output* of the statistical production process, rather than on the process itself. Of course this approach goes hand in hand with quality management of the process of the compiling of statistics.

For further information or comments on this contribution, please contact Symon Algera: sala@cbs.nl.

Notes

– This article was presented at the seminar on “Structural Business Statistics”, Warsaw, 7–9 June 1999.
– The section on basic features of the integration process is taken from Van Nunspeet and Takema, 1998.

References

Algera, S.B. and R.J.A. Janssen, 1991. The development of short-term macro-economic statistics. In: *CBS Select, Statistical Integration* (Statistics Netherlands).

Algera, S.B., 1995. *Reliability and timeliness*. Paper presented at the ad hoc group on main economic indicators OECD, Paris, October 1995.

Bolster, G.W. de, 1997. The impact of EDI on data collection. In: *Statistical Journal of the United Nations* ECE 14 (1997) 335–343, IOS Press.

Gorter, C.N. et al., 1990. *A summary description of sources and methods used in compiling the final estimates of Dutch National Income 1986*. National Accounts Occasional Paper no. 35 (Statistics Netherlands).

Janssen R.J.A. and S.B. Algera, 1988. *The methodology of the Dutch system of Quarterly Accounts*. National Accounts Occasional Paper no. 25 (Statistics Netherlands).

Nunspeet, W. van and T. Takema, 1998. *Ten years of supply and use tables in the Netherlands*. Paper presented at NSSG/Eurostat Seminar on the implementation of ESA ’95, Athens, December 1997.

Statistics Netherlands, 1992. *The Industrial statistics of the Netherlands.* ISBN no: 90.35714717.

Vet, B. de, 1999. Data for regional economic policy: forty years of regional accounts. In: J.G.S.J. van Maarseveen and M.B.G. Gircour (eds.) *A Century of statistics – counting, accounting and recounting in the Netherlands*. ISBN no: 90.6861.168.2.

Labour force scenarios for the European Union

Andries de Jong

1. Introduction

In 1995 the European Commission commissioned Statistics Netherlands to revise and extend labour force scenarios for the countries of the European Union. These scenarios are part of a comprehensive Eurostat research programme. In an earlier phase population scenarios were formulated based on alternative assumptions on fertility, mortality and international migration. The labour force scenarios are a refinement of these population scenarios in that the dimension labour force participation has been added.

The assumptions of the scenarios apply to the period up to 2020 and concern labour force participation rates laid down in three scenarios. The *Baseline Scenario* is supposed to predict the most likely development in the coming decades. The Low and High Scenarios describe alternatives in different economic and cultural contexts. The *Low Scenario* presents a rather gloomy economic development and resentment of cultural change, which have a negative influence on labour participation, and in the *High Scenario* a flourishing economy and a positive attitude toward cultural change boost labour market participation.

The labour force participation rates have been combined with the population scenarios to arrive at labour force scenarios. As the population scenarios at the national level were available until 2050, computations of the labour force have also been made up to this year, by keeping labour force participation rates constant after 2020. Three main scenarios on the labour force were compiled by combining the Low, Baseline and High Scenarios of labour force participation rates with the Low, Baseline and High Population Scenarios.

The labour force comprises both the employed and the unemployed. Employed people either have a paid job or are self-employed, and pursue occupational activity of at least one hour a week. Unemployed people are not in any employment, but are currently available for work, are actively seeking a job, or intend to be self-employed.

The labour force scenarios cover the 15 countries of the European Union and project the labour force on 1 January by sex, age in years (between 15 and 75 years) and three categories of working hours (1-19, 20-31 and 32 or more hours a week). In addition to scenarios compiled at the national level, a further differentiation to the regional level (NUTS II) has also been made.

The present article gives an overview of the main characteristics of these labour force scenarios. More information can be found in De Jong and Broekman (1999) and De Jong (1998).

2. Three long-term labour force scenarios

Three internationally consistent, long-term labour force scenarios were compiled.

The *Baseline Scenario* assumes that recent trends will continue for the most part. The main assumptions are:
– continued growth of the EU economy;
– modest increases in labour demand and employment growth;
– moderate rise in the labour force participation of young people (under 25);
– significant increase in the labour force participation of women;
– slight fall in the participation rate of middle-aged men;
– considerable increase in the participation of women aged 50 to 75;
– slight fall in the participation of elderly men.

The main characteristics of the *Low Scenario* are:
– lower economic growth than in the past;
– lack of jobs for young people leading to a decline of the participation among young people;
– deterioration of general conditions for middle-aged men and women to be active on the labour market;
– continued trend towards early retirement, especially among men.

The main features of the *High Scenario* are:
– higher rates of economic growth than in the past;
– rising demand for labour;

Table 1
Participation rates by age group in the European Union

			Low scenario		High scenario		Baseline scenario	
	1995		2020		2020		2020	
	men	women	men	women	men	women	men	women
15–19	29	24	23	21	44	42	33	31
20–24	70	61	60	58	80	76	71	67
25–29	89	73	81	72	93	87	87	79
30–34	95	71	89	75	97	89	93	81
35–39	96	71	91	75	97	88	93	81
40–44	95	71	90	73	97	87	93	78
45–49	93	66	88	69	97	86	92	76
50–54	87	58	81	60	95	81	89	70
55–59	67	39	58	39	86	66	74	52
60–64	32	14	24	12	56	38	39	24
65–69	10	4	6	3	19	12	13	7
70	5	2	2	1	7	4	5	2

– more young people combining education with paid employment;
– ample incentives for men and women in their prime to be active on the labour market;
– older persons postponing retirement because of more flexible working arrangements.

A comparison of the situation in 2020 with that in1995 according to the three scenarios led to following conclusions. For men, all age-specific activity rates of the Low Scenario are well below those observed in 1995. In contrast, all activity rates of the High Scenario are substantially above the latest observed figures. For the Baseline Scenario a mixed pattern emerges, for young and older men the rates are slightly above the 1995 pattern and slightly below it for men in the middle age group.
The picture is quite different for women. The Low Scenario resembles the 1995 pattern most, especially at younger and older ages. A slight rise is foreseen for the middle working ages. A rise is projected for all ages in the Baseline Scenario and even more so in the High Scenario.

3. National long-term labour force scenarios

Three geographic clusters of countries can be distinguished on the basis of age patterns of participation rates, namely the northern, western and southern part of the EU. In compiling national scenarios on labour force participation it was assumed that future participation trends of countries belonging to a same cluster bear more resemblance than those between countries belonging to different clusters.

In most EU countries the labour participation rate of men has shown a steady downward trend. The economic recession of the early nineties increased the effect in several countries, such as Denmark and Spain, while in other countries, such as the Netherlands, male participation remained constant.
International differences are not very impressive on the whole in 1995, although the male participation rates in Belgium and Italy were over 10 percentage points lower than in the leading country Denmark. Large variability in male participation rates exists predominantly at both young and old ages. Both Denmark and the United Kingdom are characterised by high participation rates among young men. In contrast, in Belgium, France and the countries of the southern cluster youth participation is rather low. When it comes to low participation for older people Belgium and France are in the lead. Both in the northern and southern cluster senior participation is comparatively high.

According to the Baseline Scenario the male participation rate will be somewhat lower in most countries in 2025, primarily caused by falling participation rates at the mid-range working ages. According to the Low Scenario current differences between the countries will largely persist while in the High Scenario they will become much smaller, although they will not disappear completely. In the Baseline Scenario an intermediate situation applies.

The average participation rate of women in the EU member states has generally shown a significant rise over the last decades, although the pace of change has been far from uniform. The Netherlands used to have very low participation rates, especially compared with the surrounding countries, but they have surged ahead recently, putting the country in the middle ranks. Perhaps a parallel can be found in the fertility rates, which used to be quite high, but via a very rapid fall became ‘average’. Combining childcare with paid work used to meet with disapproval, but Dutch women have found a way out of this situation by changing from full-time to part-time work when they have children, enabling them to have the ‘best of both worlds’. Nowadays the Scandinavian countries have the highest female participation rates, closely followed by the United Kingdom. It seems as if the participation rate of women has reached its peak in this part of Europe, for Denmark was confronted with a falling rate for the first time in the early nineties.

The current international differences in female participation rates are considerably larger than those for men. The northern cluster takes the lead at almost all ages while in the south female participation is still uncommon. In the southern cluster the traditional pattern of leaving the labour market after childbirth still prevails. Most countries of the western cluster are in a transition stage, in which having children has a limited effect on the labour participation rate of women. In the northern cluster, family obligations no longer interfere with having a paid job.

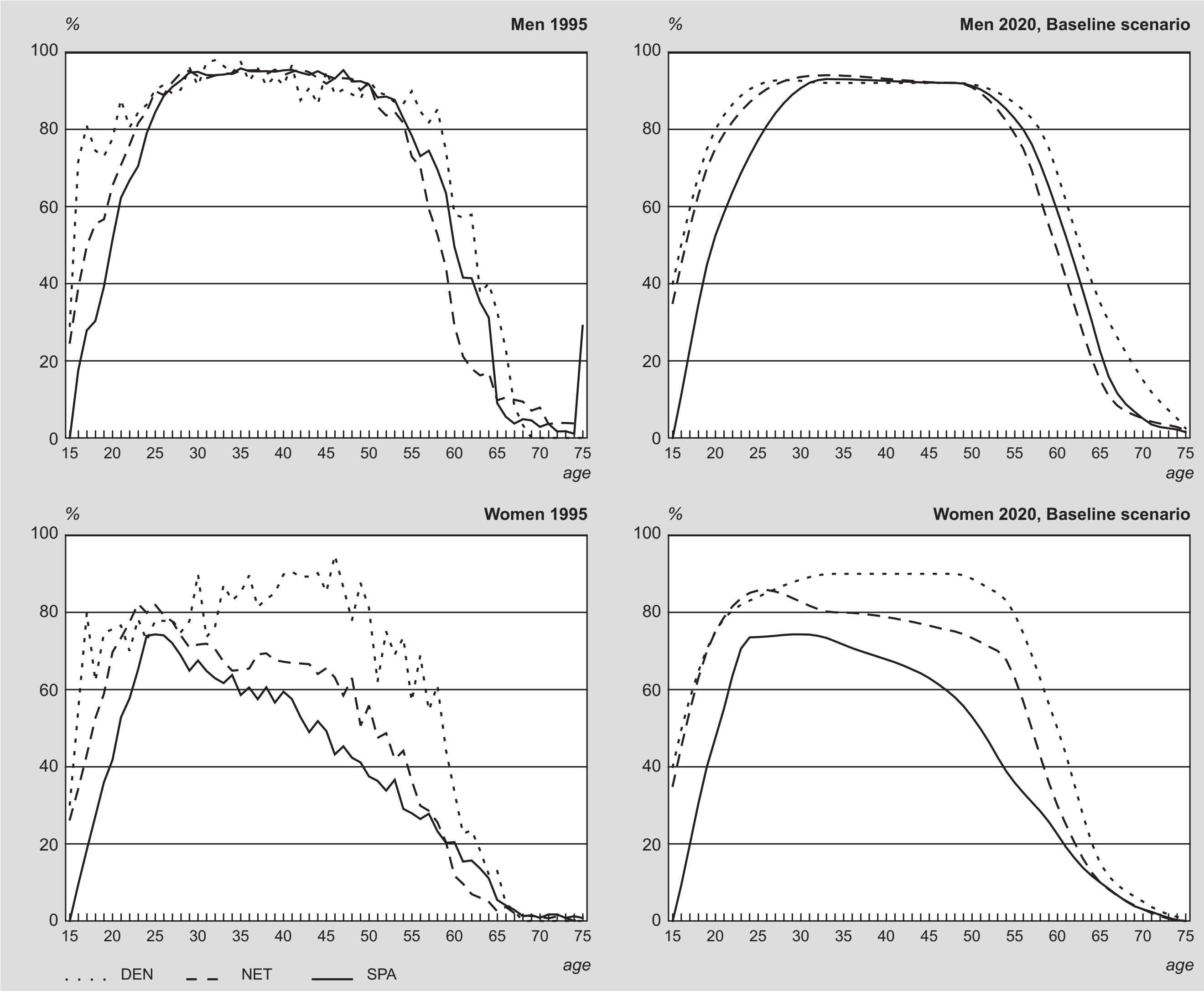
According to the Low Scenario the participation rate for women in the western cluster will move substantially upwards, in the direction of the northern cluster. For the southern cluster only a minor rise is assumed. Both under the Baseline and High Scenario, female

Table 2
Participation rates per country [1)]

			Low scenario		High scenario		Baseline scenario	
	1995		2020		2020		2020	
	men	women	men	women	men	women	men	women
AUS	73.5	53.6	61.4	51.0	77.1	69.5	69.3	60.8
BEL	64.0	44.3	53.6	41.8	66.1	58.4	59.2	48.6
DEN	76.9	63.9	67.2	58.2	80.4	72.2	73.5	65.2
FIN	68.2	60.2	57.1	50.4	67.7	65.3	62.6	57.3
FRA	66.3	52.1	53.6	47.9	71.3	65.1	64.4	56.4
GER	72.5	52.7	64.1	52.1	74.2	66.5	69.8	58.3
GRE	69.9	39.4	63.4	39.0	76.5	56.8	69.5	45.1
IRE	71.0	42.4	63.7	36.6	78.6	59.9	72.4	46.9
ITA	64.8	36.7	56.7	34.2	74.2	53.0	64.7	42.0
LUX	68.9	38.5	56.5	33.9	72.1	54.6	64.3	44.4
NET	73.0	51.5	58.9	48.3	75.6	65.7	68.0	56.3
POR	71.6	53.0	64.8	53.2	78.4	68.4	70.4	59.8
SPA	68.6	40.9	62.3	36.9	75.9	54.7	68.6	44.6
SWE	73.2	67.1	61.1	54.7	74.4	72.1	67.5	62.6
UKI	76.2	58.4	66.5	54.7	79.5	71.3	72.0	63.2

[1)] Number of active persons related to the number of persons aged 15–75.

Figure 1
Participation rates in Denmark, the Netherlands and Spain



participation rates will increase sharply in the western and southern countries. In the High Scenario this will lead to a virtual disappearance of the traditional age pattern in the southern countries. Only marginal differences in participation rates will be left between the western and northern clusters.

4. Labour force scenarios by working hours

In order to gain more insight in the fundamental trends in employment, labour force participation was also distinguished by working hours. The economic recovery of the mid-nineties seemed to be accompanied by an increase in part-time jobs. An ongoing trend towards a more flexible organisation of work makes it easier to employ part-time workers or people on fixed-term contracts. This creates more job opportunities for mothers with young children in particular.

The quantitative assumptions of the labour force scenarios by working hours divided the Low, Baseline and High scenarios into three categories, i.e. 1–19, 20–31 and 32 or more hours a week. Data on working hours at the outset of the scenario period were derived from the Labour Force Survey and are the number of hours people normally work each week in their main job, including any customary overtime or extra hours. In the scenarios the remaining category of the labour force, namely the unemployed, is omitted as this would require a projection of labour demand. In the starting year of the scenarios, the share of the unemployed in the activity rate was distributed proportionally across the three categories of working hours.

The qualitative assumptions of the labour force scenarios on working hours are mainly connected with the effects of economic growth on the division between full-time and part-time jobs. The Baseline Scenario assumes that a limited reorganisation of the labour market towards more flexibility will take place. A satisfactory growth of the economy will lead to a moderate creation of new jobs, and most of the employment growth will go into part-time jobs.

In the Low Scenario a relatively low economic growth rate will lead to a meagre creation of new jobs, and they will be part-time rather than full-time. On the other hand in some countries recession might lead to a decline in employment and if companies have to lay off employees, especially part-timers will have to leave. Job incentives for marginal groups are minimal.

The conditions in the High Scenario are just the opposite. High economic growth will lead to abundant job opportunities. The labour market will become more flexible. Part-time employment will grow rapidly, absorbing groups who previously stood on the sidelines.

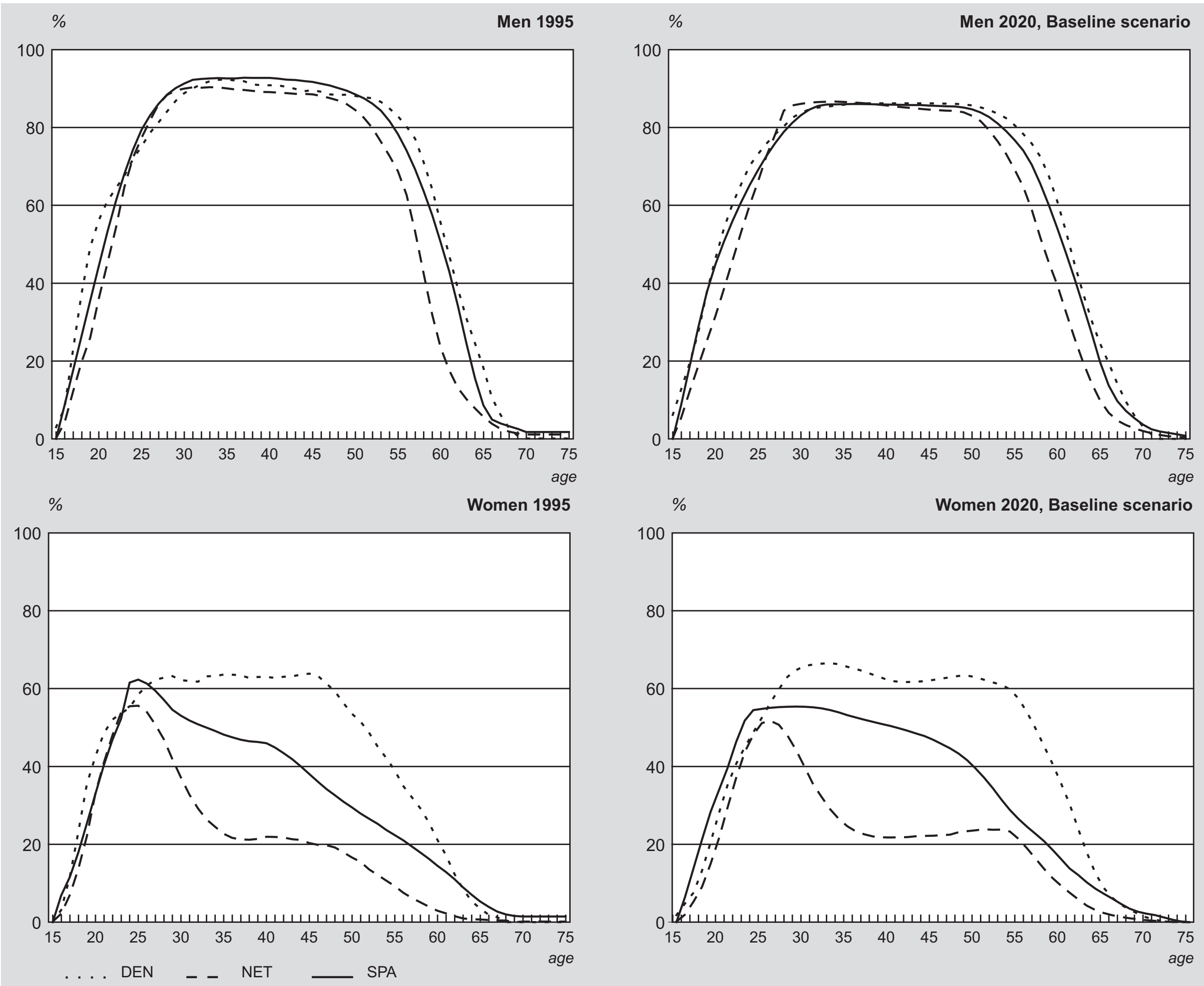
Table 3
Participation rates by working hours per week in the European Union

	Observed	Low scenario			High scenario			Baseline scenario		
	1995	2000	2005	2020	2000	2005	2020	2000	2005	2020
1–19 hours										
Men	1.9	1.9	2.0	2.0	2.6	3.4	5.4	2.4	3.0	3.6
Women	6.2	6.2	6.2	6.2	7.5	8.8	11.3	7.0	7.7	8.5
20–31 hours										
Men	2.4	2.3	2.2	2.2	2.8	3.3	4.6	2.7	3.0	3.4
Women	8.9	8.9	8.9	8.8	10.6	12.1	15.0	9.9	10.7	11.6
32 and over hours										
Men	58.7	55.5	52.6	52.6	59.2	60.4	61.4	57.7	57.2	57.2
Women	30.4	30.5	30.6	30.7	32.6	34.7	35.8	31.2	32.0	33.0

5. National labour force scenarios by working hours

Differences in the participation rates of men in full-time jobs between the EU countries are limited. Portugal's high score is partly based on the relatively large size of the primary sector, where part-time jobs are rare. But the gap between Portugal and countries with low rates such as Finland, Belgium and the Netherlands is only just over ten percentage points. These countries are characterised by early retirement. At young and middle age ranges the differences with other countries of the EU are limited.

Figure 2
Participation rates 32+ working hours in Denmark, the Netherlands and Spain



According to the Baseline Scenario, national differences in activity curves will more or less remain the same in the future.

With respect to the participation of women in full-time jobs the differences between the countries are more impressive. The Scandinavian countries, Austria and Portugal lead the field, and the gap with the Netherlands, which has by far the lowest participation rate, is about twenty percentage points.
In the countries with high activity rates the female activity curve for full-time jobs is similar to the male curve. In the rest of the EU the male and female activity curves hardly resemble each other, because of very different attitudes towards working mothers. In Scandinavia and in several western countries, such as France, bringing up children does not constitute a reason for leaving the labour force or working fewer hours. In the Netherlands in particular, women who have children cut back their hours by to differing extents. This results in a sharp drop in the activity rates for full-time jobs in the 25-34 age bracket, followed by a fairly stable activity rate up to women in their fifties. In several western and southern countries the activity curve has a left-hand peak, reflecting a gradual change from full-time jobs to part-time jobs with relatively many hours, and at higher ages a retreat from the labour market.
According to the Baseline Scenario, the activity rates for full-time jobs in the northern and western part of the EU will only show a significant rise at higher ages, mainly an expression of the general rise in participation at these ages. In the southern countries a rise in full-time participation is expected for both middle and older age groups, again the result of the expected rise in female participation, but this time coupled with less orientation towards part-time working than in the rest of the EU.

In all the countries of the EU it is still quite rare for men to work part-time: activity rates are only a few percentage points above zero for each age. In the Baseline Scenario the age-specific activity rates belonging to part-time jobs with a large number of hours (20–31 hours a week) will slowly rise and fluctuate around 5% in 2020.
Part-time jobs with a relatively large number of hours are popular among women in several countries, especially Sweden, but also Denmark, France, the Netherlands and the United Kingdom. These countries have a bell-shaped activity curve, with a top around the age of 45. The lowest activity rates for 'large' part-time jobs can be found in the southern countries, where the age pattern of the activity rates for women resembles that of men: rather flat and only a few percentage points above zero.
In the Baseline Scenario the female activity curve will rise slightly in the southern countries and continue to reflect the male age pattern. For Denmark and Sweden no important changes in the current activity curve are foreseen. Activity rates for 'large' part-time jobs have fallen in Denmark over the last ten years. It is assumed in the Baseline Scenario that a growing popularity of part-time work might

Figure 3
Participation rates 20–31 working hours in Denmark, the Netherlands and Spain

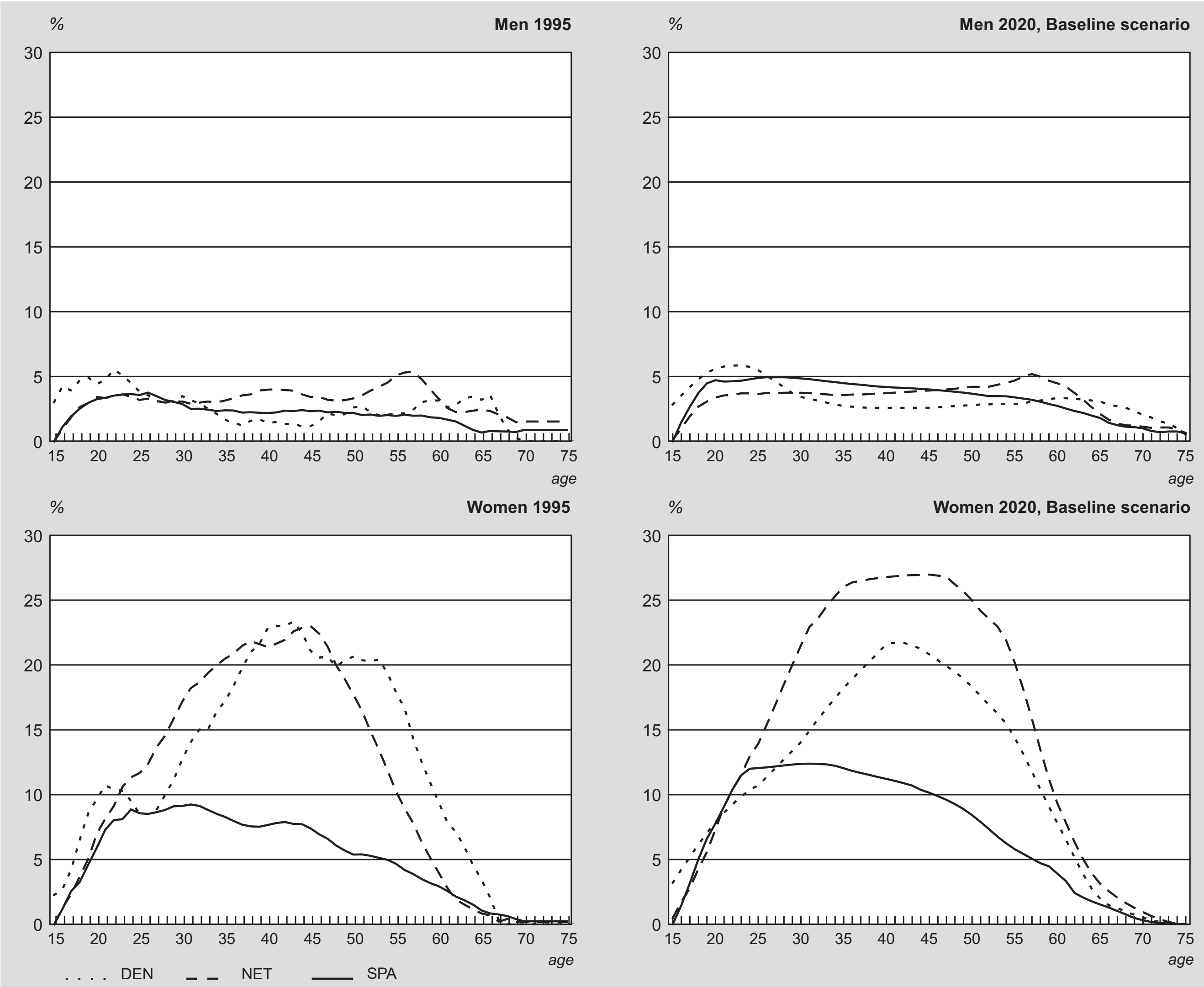


Figure 4
Participation rates 1–19 working hours in Denmark, the Netherlands and Spain

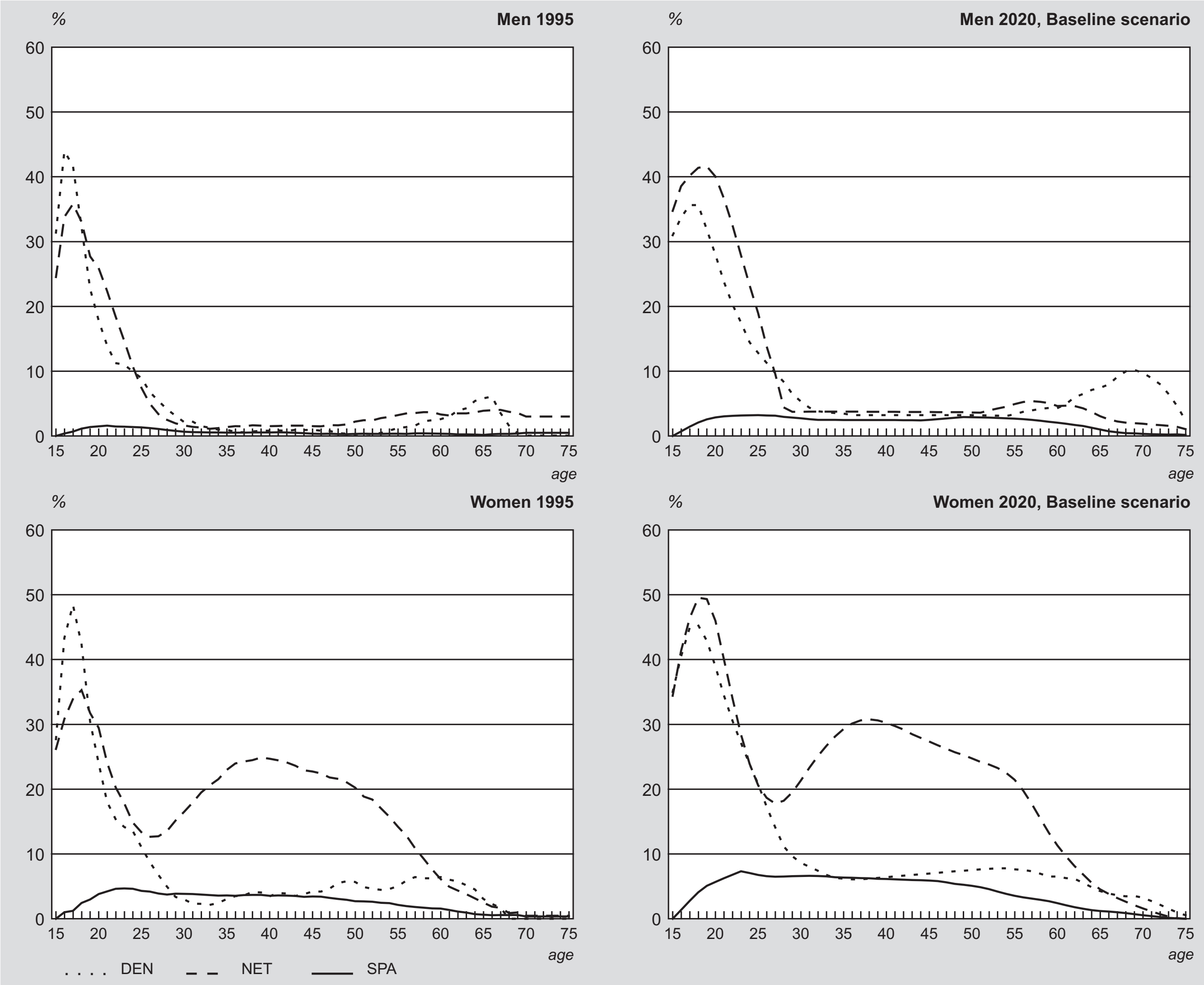


Table 4
Participation rates by working hours per week per country [1)]

	32+ working hours				20–31 working hours				1–19 working hours			
	men		women		men		women		men		women	
	1995	2020	1995	2020	1995	2020	1995	2020	1995	2020	1995	2020
AUS	61.5	60.4	36.1	42.1	1.6	2.2	9.5	12.7	1.1	2.7	3.1	4.1
BEL	53.7	52.0	25.6	28.8	2.6	3.0	9.1	12.2	0.8	2.1	5.2	7.6
DEN	61.2	59.1	38.5	41.3	2.5	3.2	11.9	11.3	5.7	8.8	7.9	11.5
FIN	52.4	51.9	43.7	43.3	3.5	4.6	7.0	8.9	3.6	5.4	4.3	6.2
FRA	55.1	55.2	33.6	36.0	2.9	4.1	10.0	13.7	1.2	2.3	4.1	6.1
GER	61.0	60.6	32.7	35.8	1.4	2.4	9.7	12.3	1.4	2.4	6.7	8.4
GRE	61.5	58.8	30.3	33.3	3.7	5.4	5.4	8.6	0.8	1.9	1.4	2.2
IRE	61.7	59.8	25.8	28.3	4.0	5.9	7.6	10.2	1.3	2.5	4.2	6.0
ITA	55.8	56.1	25.5	29.9	2.2	2.9	6.0	8.8	1.5	2.6	2.6	3.8
LUX	57.2	57.1	24.1	28.2	1.2	2.0	7.0	10.8	0.6	1.4	2.9	4.1
NET	53.8	53.1	18.1	19.6	3.0	3.3	11.0	14.9	6.2	9.4	15.7	21.5
POR	64.3	59.4	40.9	45.3	3.1	4.0	6.5	8.4	1.1	2.8	3.0	4.0
SPA	60.0	57.2	29.1	32.0	2.1	3.3	5.1	7.1	0.6	2.1	2.5	4.2
SWE	58.4	53.7	38.2	35.2	4.5	5.4	18.3	18.4	3.7	5.3	5.8	7.3
UKI	61.5	57.5	29.0	29.2	3.0	3.8	11.0	13.8	4.0	6.6	13.6	17.8

1) 2020 according to Baseline scenario.

prevent a further fall. In contrast, in most western countries, especially in the Netherlands and France, working in part-time jobs has been becoming more common for many years now. In the future this trend will continue, resulting in higher activity rates for 'large' part-time jobs. This also means that the further rise of female activity rates which will take place in most western and southern countries of the EU will predominantly stem from a growing preference for working in part-time jobs with a large number of hours.

Within the EU large differences can be detected between the participation rates for 'small' part-time jobs, i.e. 1–19 working hours a week. The three Scandinavian countries, the Netherlands and the United Kingdom have high activity rates for both men and women for 'small' part-time jobs up to the age of twenty. Then the activity rates go into a deep decline in the early twenties, and slow down in the late twenties. In the Scandinavian countries the activity rates of the two sexes no longer go hand in hand above the age of thirty, when female participation rates become notably higher. In the Netherlands and the United Kingdom the female pattern no longer resembles that of Denmark as age progresses: the bell curve peaks around the age of 40.

In the Scandinavian countries most mothers have 'large' part-time jobs or continue to work full-time in combination with parenthood. In the Netherlands and the United Kingdom the strategy is slightly different: most mothers change from a full-time job to either a 'large' or a 'small' part-time job after childbirth. In the southern countries hardly any men or women have part-time jobs.
The Baseline Scenario assumes that in line with current developments, activity rates of people in their twenties will rise significantly in the Scandinavian countries, and even more so in the Netherlands and the United Kingdom. Above that age a limited rise in activity rates is foreseen in all countries of the EU.

6. Labour force

The labour force scenarios were drawn up by multiplying population numbers with labour force activity rates. The underlying population scenarios were provided by Eurostat and compiled by Statistics Netherlands and the Netherlands Interdisciplinary Demographic Institute (Beer and de Jong, 1996, Gaag et al. 1998). These scenarios are based on various combinations of low, medium and high variants of fertility, life expectancy, and net migration. The Baseline Scenario is the main reference scenario and describes the outcome of a continuation of current trends, while the other two give an indication of the uncertainty of the future population size.

The Low, Baseline and High scenarios of labour force participation rates were combined with the Low, Baseline and High population scenarios respectively. Because the low labour force scenario is based on both low population growth and low participation rates while the high scenario combines high population growth with high participation rates, the two scenarios on the labour force represent quite extreme developments. The assumptions on the activity rates apply to the period until 2020. However, as the population scenarios at the national level were available until 2050, computations of the labour force were also made up to that year, by keeping labour force participation rates constant after 2020.

In the Baseline Scenario the EU population in the age-group 15–75 will grow moderately up to 2020, and subsequently start to decline. In the Low population scenario no population growth is foreseen and the EU will be confronted with a rapid decrease after 2020. In the High scenario a rather fast population growth is expected in the short run followed by stabilisation.
The combination of population trends with the trends in the activity rates according to the three scenarios lead to more or less the same developments in the labour force and the working age population.
In the Baseline Scenario, the labour force will grow from nearly 170 million in 1995 to about 180 million in 2020. Population decline will lead to a reduction to around 155 million in 2050. In the Low Scenario the labour force will decline continuously, leaving around 150 million people in 2020. The High Scenario envisages a rise leading to a workforce of over 210 million in 2020.

7. National labour force trends

According to the Baseline Scenario the labour force will have grown slightly in all member countries of the EU by 2020, in line with the labour force trends in the EU as a whole. As population numbers will fall in the second quarter of the 21st century, the labour force will shrink and will drop just below the current size in most countries by 2050. However, in Germany, Italy and Spain the labour force will be significantly smaller by then, because of sustained low fertility levels. France, on the other hand will have a considerable larger labour force by the middle of the 21st century, partly as a consequence of its rather high fertility rate. In the High Scenario all countries will have a growing labour force in the first half of the 21st century, while in the Low Scenario all countries will face a continuously diminishing labour force.
The trends in the labour force are not alike for men and women. The Baseline Scenario predicts a male labour force of about the same size in 2020 as in 1995 in most member states, the main exceptions

Table 5
Population and labour force in the European Union

	Observed			Low scenario			High scenario			Baseline scenario		
	1985 [1]	1995 [1]	1995	2000	2020	2050	2000	2020	2050	2000	2020	2050
	x mln											
Population	253	268	285	288	285	224	290	308	307	289	296	262
Labour force	148	158	169	167	152	114	183	212	210	176	180	155

[1] EUR 12.

Table 6
Labour force per country

	Observed		Low scenario			High scenario			Baseline scenario		
	1985	1995	2000	2020	2050	2000	2020	2050	2000	2020	2050
	x mln										
AUS		3.9	3.8	3.5	2.6	4.1	5.1	5.2	4.0	4.2	3.7
BEL	4.0	4.2	4.1	3.7	2.9	4.5	5.2	5.3	4.4	4.3	4.0
DEN	2.8	2.8	2.8	2.5	2.0	3.0	3.4	3.6	2.9	2.9	2.8
FIN		2.5	2.4	2.1	1.7	2.6	2.8	3.0	2.5	2.4	2.2
FRA	24.6	25.6	25.4	23.2	18.4	28.1	33.3	34.5	27.6	30.7	30.2
GER	36.1	39.8	39.2	36.2	25.9	42.5	47.9	45.6	40.8	41.6	34.1
GRE	4.0	4.4	4.4	4.1	3.2	4.9	5.8	5.9	4.7	4.8	4.4
IRE	1.3	1.4	1.5	1.4	1.1	1.7	2.1	2.2	1.6	1.9	1.9
ITA	23.0	22.8	22.4	18.8	12.4	24.7	28.5	25.6	23.6	23.0	17.8
LUX	0.2	0.2	0.2	0.2	0.1	0.2	0.3	0.3	0.2	0.2	0.2
NET	5.8	7.4	7.4	6.7	5.4	8.0	9.8	10.3	7.7	8.2	7.8
POR	4.5	4.7	4.7	4.6	3.6	5.2	6.2	6.3	4.9	5.2	4.8
SPA	13.9	16.3	16.3	14.6	9.7	18.1	20.8	19.1	17.2	17.3	13.5
SWE		4.5	4.2	3.9	3.3	4.6	5.4	6.2	4.4	4.5	4.7
UKI	27.6	28.8	28.4	27.0	21.6	30.6	35.8	36.7	29.7	31.0	27.6

being Luxembourg and Ireland, where some growth is still to come. In Ireland net migration will cease be negative in the future and, stimulated by a relative high fertility rate, the labour force will have enough potential to grow.
While the male labour force will stagnate in the coming decades, the female labour force will grow in the first quarter of the 21st century in a substantial part of Europe. However, in the Scandinavian and southern countries growth is expected to be meagre. In Scandinavia this is because the participation rates are already high and approaching those of men, while in the south the cause is a long-term low fertility level. Also, the labour force growth for women will be reversed when the population starts to decrease. By the middle of the 21st century the female labour force will be back at its current size in most countries, whereas Italy and Spain will be confronted with a much smaller female labour force.

8. Labour force by working hours

The effect of the assumptions on the activity rates for full-time jobs according to the three scenarios can be judged by looking at the future trends in the EU labour force. In the Low Scenario the size of the male labour force working full-time will fall in the coming decades. In the Baseline Scenario the current size will be maintained up to about 2020, after which a decline is foreseen. The High Scenario is the only one where a rise is expected in the near future, followed by a nearly constant size from 2020 onwards.
The expected trends for women are slightly different. In the Low Scenario the size of the full-time labour force will first remain constant for about ten years. Both the Baseline and the High scenarios expect a rise in the short run. After 2015 the female labour force will maintain this higher size under the High Scenario, while under the Baseline Scenario a moderate fall is foreseen.

Because part-time work is more popular among women, and because more men work to begin with, the male labour force is nearly twice as large as the female labour force. Under the Baseline Scenario this ratio will be somewhat less biased in the future, as the decline in the number of men working full-time will be greater than that for women.

The number of men working in ‘large’ part-time jobs is still modest in the EU, namely four million. Although the High Scenario predicts this will double, part-time work will continue to be a rather marginal phenomenon for men.

Table 7
Labour force by working hours per week in the European Union

	Observed	Low scenario			High scenario			Baseline scenario		
	1995	2000	2020	2050	2000	2020	2050	2000	2020	2050
	x mln									
1–19 hours										
Men	2.8	2.8	2.8	2.2	3.8	8.4	8.6	3.5	5.3	4.8
Women	9.4	9.6	8.9	6.9	11.7	17.6	17.5	10.8	12.8	11.1
20–31 hours										
Men	3.6	3.5	3.3	2.6	4.3	7.3	7.3	4.1	5.2	4.6
Women	13.8	14.1	13.1	9.9	16.7	23.6	23.2	15.6	17.7	15.2
32 and more hours										
Men	91.9	88.9	79.9	59.4	94.8	99.8	98.6	92.3	89.8	77.1
Women	47.9	48.2	44.4	32.9	51.6	55.6	54.6	49.4	49.5	42.3

The prospects for women working in ‘large’ part-time jobs are better, especially in the High Scenario. Only the Low Scenario foresees a slight fall, while both in the Baseline and the High scenario a notable grow is foreseen up to around 2020. At the moment, nearly four times as many women as men are working in a ‘large’ part-time job and this ratio is expected to remain more or less the same in the future.
Nearly three million men work in ‘small’ part-time jobs in the EU. The High Scenario expects this figure to triple in the future, while no growth will take place in the Low Scenario.
The number of women working in ‘small’ part-time jobs is three times as high as that of men. The High Scenario and the Baseline Scenario foresee a significant growth in the future. The Low scenario on the other hand foresees a moderate decline.

For further information or comments on this contribution, please contact Andries de Jong: ajgg@cbs.nl.

References

Beer, J. and A. de Jong (1996), National population scenarios for the countries of the European Economic Area. In: *Monthly bulletin of population statistics,* 7/1996, pp. 7–19. Statistics Netherlands.

Gaag, N. van der, L. van Wissen, E. van Imhoff and C. Huisman (1998), *National and Regional Population Trends in the European Union*, 1975–2025. NIDI.

Jong, A. de (1998), *Analysis and projection of labour force by gender and age for the countries of the European Economic Area*. Report for the European Commission.

Jong, A. de and R. Broekman (1999), *National and Regional Trends in the Labour Force in the European Union*. Report for the European Commission

Uncertainty of population forecasts: a stochastic approach

Joop de Beer and Maarten Alders

1. Introduction

The Dutch population forecasts published by Statistics Netherlands every other year project the future size and age structure of the population of the Netherlands for the next fifty years (de Beer, 1999). The forecasts are based on assumptions about future changes in fertility, mortality, and international migration. Obviously the validity of assumptions on changes in the long run is uncertain. It is important that users of forecasts are aware of the degree of uncertainty. In order to provide information about the degree of uncertainty of population forecasts it has become common practice among population forecasters to publish several variants. In addition to the medium variant which is assumed to describe the most plausible future development, low and high variants are also usually published. The low variant which is based on the assumption of low fertility, high mortality and low net migration projects low population growth. Accordingly the high variant projects high growth. One disadvantage of these variants is that they tend to overestimate the degree of uncertainty of future population size and to underestimate the degree of uncertainty of the ageing of the population. The reason is that the low and high variants are deterministic variants, which are based on the assumption of perfect correlation between the three components of population growth (positive correlation between fertility and net migration and negative correlation between these components and mortality) and perfect correlation between successive forecast years. Since it is unlikely that fertility and net migration are high (low) in each forecast year, while at the same time mortality is low (high), the low and high variants are rather extreme. The probability that the interval between the low and high variants of population size will cover the true value is much higher than the probabilities that the separate intervals between the low and high values of fertility, mortality and migration will cover the true value. Monte Carlo simulations show that if the intervals between the low and high values of fertility, mortality, and migration each correspond with a probability of two thirds, the resulting interval of population size corresponds with a probability of over 95 per cent in the long run. On the other hand, the low and high variants tend to underestimate the degree of uncertainty of ageing, since in the low (high) variant both the number of elderly people and the size of the population are small (large). Thus the ratio of the number of elderly people to population size hardly differs between the low and high variants.

In order to give more accurate information about the degree of uncertainty of population forecasts, Statistics Netherlands decided to produce stochastic population forecasts. Instead of publishing two alternative deterministic variants in addition to the medium variant, forecast intervals are published. These intervals are calculated by means of Monte Carlo simulations. The simulations are based on assumptions about the probability distributions of future fertility, mortality, and migration. Recently stochastic population forecasts were published for various other countries also: Austria (Hanika et al., 1997; Lutz and Scherbov, 1998b), Finland (Alho, 1998), Germany (Lutz and Scherbov, 1998a), and the United States (Lee and Tuljapurkar, 1994).

This article gives a concise description of the methodology and the main assumptions underlying the Dutch stochastic population forecasts. More details are given in de Beer and Alders (1999). The main focus of this article is on the interpretation of the results of stochastic population forecasts.

2. Deterministic low and high variants

Population forecasts are based on the cohort-component model. This model projects the future population by age and sex on the basis of assumptions about future changes in fertility, mortality, and migration by age and sex. In the Dutch population forecasts assumptions on fertility refer to age-specific rates distinguished by parity, mortality assumptions refer to age and sex-specific mortality rates, assumptions about immigration refer to absolute numbers distinguished by age, sex and country of birth, and assumptions on emigration are based on a distinction of emigration rates by age, sex and country of birth. Uncertainty of population forecasts depends on the uncertainty of the validity of these assumptions. The effect of the uncertainty on future fertility, mortality, and migration on the future size and age structure of the population can be assessed by means of calculating variants based on alternative values of these three components. Usually high and low variants are based on combining high and low values of fertility, life expectancy, and net migration respectively. These deterministic variants imply perfect correlation between fertility, mortality, and migration.

Even if deterministic variants are specified, assumptions about the degree of probability are required, although in practice they are often made only implicitly. In specifying low and high variants one crucial question is how wide the intervals between the low and high values of fertility, life expectancy, and net migration should be. Obviously a wide interval implies a higher level of probability than a small interval. One implication of the fact that assumptions about the degree of probability are usually not stated explicitly, is that there are remarkable differences between the forecasts for different countries. For example, in the Belgian forecast the difference between the low and high variants of average family size equals 0.3 children per woman, while in the French forecast the interval is 0.6 children (Sorvillo, 1999). It is questionable whether the degree of uncertainty of fertility projections differs so much between Western European countries that this can justify these differences. It seems more likely that the intervals underlying the forecasts of different countries are implicitly based on different assumptions about the level of probability. Consistency between forecasts of different countries can be achieved if the level of probability corresponding with the forecast interval is specified explicitly. Moreover, explicit assumptions about the level of probability are useful for achieving consistency between the components. Assumptions about the future trends of fertility, mortality and migration are expressed in different measures: numbers of children per woman, life expectancy in years, numbers of immigrants and age-specific emigration rates. The intervals for these measures cannot be compared directly. Does an interval of 0.6 children per women reflect the same degree of uncertainty as an interval of 6 years for the life expectancy at birth? An explicit assumption about the level of probability provides a criterion for comparing the intervals of the components.

Since deterministic variants imply perfect correlation between components and between forecast years, the low and high variants of population size are more extreme than the underlying low and high values of fertility, life expectancy, and net migration. In other words: the probability that the true population size will be higher than in the high variant or lower than in the low variant is considerably lower than the corresponding probabilities for fertility, life expectancy, and net migration. For example, in the 1996-based Dutch population forecasts it was assumed that the probability is 67% that the average number of children of women born in 2020 will lie between 1.4 and 2.0. Furthermore it was assumed that the probability is 67% that life expectancy at birth in 2050 will lie between 81 and 85 years for women and between 78 and 82 years for men, and that the probability is 67% that net migration will lie between zero and 70 thousand in 2010. Put together, these low and high values forecast a total population in 2050 somewhere between 12.7 and 21.5 million. Clearly such a wide interval is not very informative. However, this interval does not correspond with a probability of 67%, but rather with a probability of over 95%.

Whereas the interval between the low and high variants of population size tends to overestimate the degree of uncertainty, the interval between the two variants for the percentage of the population aged 65 or over seems to be too small. According to the 1996-based population forecasts the interval between the low and high variants of the proportion of people aged 65 or over in 2050 was only slightly more than 4 percentage points. It should be noted that the proportion in the high variant was even lower than in the low variant, in spite of the higher life expectancy in the high variant. The reason is that the interval between the low and high variants of the population size was larger than the interval for the number of elderly people. To reflect the degree of uncertainty of the proportions of elderly people alternative variants can be calculated, such as young and old variants (de Beer and de Jong, 1996). In the young (old) variant high (low) fertility is combined with high (low) net migration and low (high) life expectancy. This results in a wider interval for the proportion of elderly people but in a relatively narrow interval for population size.
Thus for each result different types of variants are needed to reflect the degree of uncertainty of that result. These kinds of problems are characteristic for deterministic variants. Stochastic forecasts do not have this problem. For each result the corresponding forecast interval can be calculated on the basis of the same set of simulations.

3. Stochastic population forecasts: methodology and assumptions

Based on assumptions about the probability distribution of future fertility, mortality, immigration and emigration, Monte Carlo simulations can be used to calculate the probability distribution of the future size and age structure of the population. Assumptions need to be specified about the type of probability distribution for each component (e.g. whether the distribution is symmetric or asymmetric), the parameters of the distributions in each forecast year (the value of the standard deviation is particularly important), and the correlations between successive forecast years, between ages, between men and women, and between the components. It is assumed that the expectations of the distributions are equal to the medium variant which is assumed to describe the most plausible future. Furthermore it is assumed that the forecast errors of fertility, mortality and migration are serially correlated, that for each component there is perfect correlation across ages and between men and women and that the three components are independent.
The main assumption concerns the standard deviation. Three methods can be used to assess the value of the standard deviation: an analysis of forecast errors of previous population forecasts, a model-based estimate of forecast errors, and expert judgement. In making the Dutch population forecasts, Statistics Netherlands applies a combination of these methods (de Beer and Alders, 1999).

The time series of life expectancy at birth can be modelled by a random walk model with drift. The width of the 95% forecast interval for the year 2050 derived from this model equals 12 years. A time series analysis of historic forecast errors, however, suggests that an interval of 8 years may be appropriate. The discrepancy between these two analyses implies that judgement is required. For this reason the processes underlying the changes in life expectancy need to be analysed. One main assumption underlying the forecast of future age-specific mortality rates is that increase in life expectancy at birth is caused by an increase in the percentage of people becoming old rather than an increase in the percentage of old people becoming very old. This assumption implies that the uncertainty of future life expectancy depends more strongly on the uncertainty of the median age of dying than on the uncertainty about changes in the maximum life span. This analysis supports the assumption that the 95% interval will equal 12 years in 2050 rather than 8 years (de Beer and Alders, 1999).
A time series model of total fertility rates produces a 95% forecast interval of the level of fertility in 2050 ranging from 0.6 to 2.8 children per woman. An analysis of fertility rates distinguished by birth order for successive cohorts suggests, however, that this may be too wide. The lower limit would imply that only a minority of women would have children, whereas the upper limit implies that a majority of women would have three children or more. Both assumptions are highly unlikely, even as boundaries of a 95% interval. An analysis of birth-order specific fertility rates suggests that an interval ranging from 1.1 to 2.3 children per woman is more appropriate (de Beer and Alders, 1999).
Assumptions about the uncertainty of future immigration are based on a distinction by country of birth. In contrast to the other components an asymmetric distribution is assumed. The reason is that for some countries the immigration numbers are already very low. Assuming a symmetric forecast interval would result in very narrow intervals. To obtain an interval for total immigration an assumption is needed about the correlation between the immigration flows from individual countries. This correlation will not be perfect because a high number of migrants from one country does not imply high numbers from other countries. For example, in the 1990s the number of immigrants from countries like Turkey and Morocco has declined, whereas the number of asylum migrants has increased. On the other hand, the separate immigration flows are not independent either. For example restrictive immigration policies affect the size of migration from different countries. Assumptions about the uncertainty of emigration are based on emigration rates distinguished by age, sex and country of birth and the relationship between the level of the emigration rates and the duration of stay of immigrants in the Netherlands.

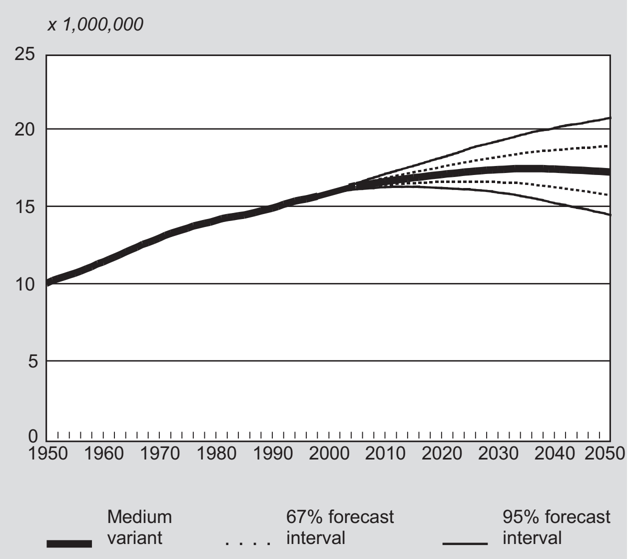
Based on the assumptions about the probability distributions of future fertility, mortality, immigration and emigration, Monte Carlo simulations provide a probability distribution of the size and age structure of the population. One thousand time paths of fertility, mortality, immigration and emigration were simulated and from these simulations one thousand different developments of the future size and age structure of the population were calculated. The results are used to compute forecast intervals. For example, the 67% forecast interval of the total population size is calculated on the basis of the assumption that if two thirds of the simulations of the population size in a certain year lie within a certain interval, the probability is two thirds that the real population size in that year will be inside that interval. The results of the simulations can also be used to calculate forecast intervals for the size of age groups, for example the number of people aged 65 or over, or for ratios such as the old age dependency ratio (the number of people aged 65 or over divided by the number of people aged 20–64 years).

4. Main results

4.1 Population growth

According to the medium variant the Dutch population will continue to increase in the next decades, from 15.9 million in 2000 to 17.4 million in 2035. After 2035 it will decline to just over 17 million people in 2050. On the basis of Monte Carlo simulations it can be calculated that the probability is two thirds that the real population size in 2035 will lie between 16.5 million and 18.5 million people (figure 1). This interval implies that it is fairly certain that the population will continue to grow in the coming years. According to the lower limit of the 67% forecast interval, the population will grow at least until 2025. The

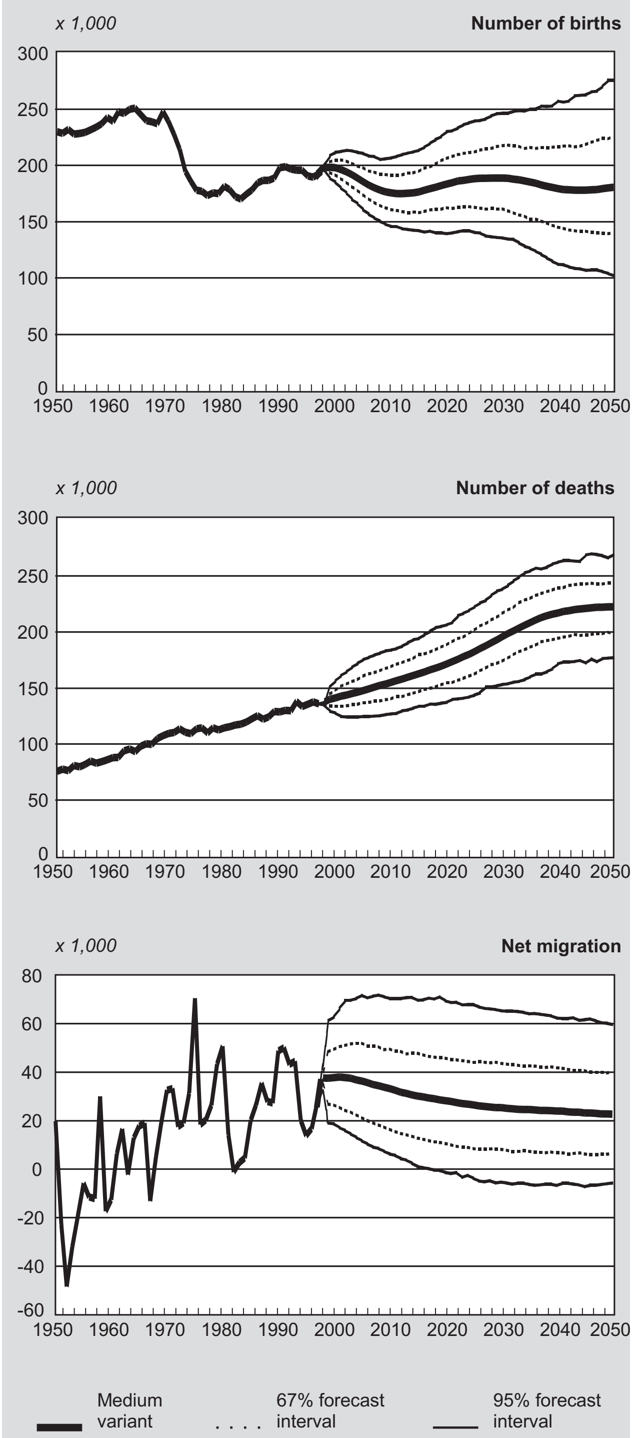
Figure 1
Population size, medium variant, 67% and 95% forecast intervals



lower limit will not fall to the current population size before 2050. Even according to the lower limit of the 95% forecast interval the population will increase in the next 15 years. Thus it is very likely that it will continue to grow. However, just how strong this growth will be remains rather uncertain. The upper limit of the 67% forecast interval continues to increase during the entire forecast period. However, it will remain well below the 20 million people projected in the 1960s. The probability that the Netherlands will count more than 20 million people in the next decades is rather small, as the upper limit of the 95% forecast interval will exceed 20 million not before 2040. This implies that it is very unlikely that the population will grow at the same speed in the next fifty years as in the last fifty years. Since 1950 the population has increased by more than 5.5 million. The probability that it will grow by more than 5 million people in the coming 50 years is assumed to be very small. The reason for this is the ageing of the population, which means that the annual number of deaths will increase strongly in the next decades, while the number of births will decline slightly. In the first decades the number of births will decrease because the number of people in the reproductive ages will decline, as the generations born after 1975 are smaller than those born in the fifties and sixties. In the long run, the number of births will not increase either because the average number of children per woman is relatively low.

Figure 2 shows the separate forecast intervals for the three components of population growth, the annual numbers of births, deaths, and net migration. The figure shows clearly that the degree of uncertainty of the number of births increases more strongly than that of the number of deaths. This is because the uncertainty of fertility in the long run consists of two components. First, the level of the fertility rate is uncertain. Secondly, in the long run, say after 2030, uncertainty increases as children will be born of whom the parents themselves are not yet born at the moment the forecast is made. The degree of uncertainty of the number of deaths increases relatively slowly. People who will die in the next fifty years are to a large extent people who are already living in the Netherlands at the moment. The degree of uncertainty of net migration increases rapidly in the short run compared with the numbers of births and deaths. In the long run it levels off, mainly because the uncertainty of migration is to a large extent determined by short term fluctuations. In the long run the degree of uncertainty is limited due to the negative mutual effects of immigration policies and migration pressure: if migration pressure increases, immigration policies tend to become more restrictive. Moreover, there is a positive correlation between emigration and immigration, limiting the width of the uncertainty interval of net migration.

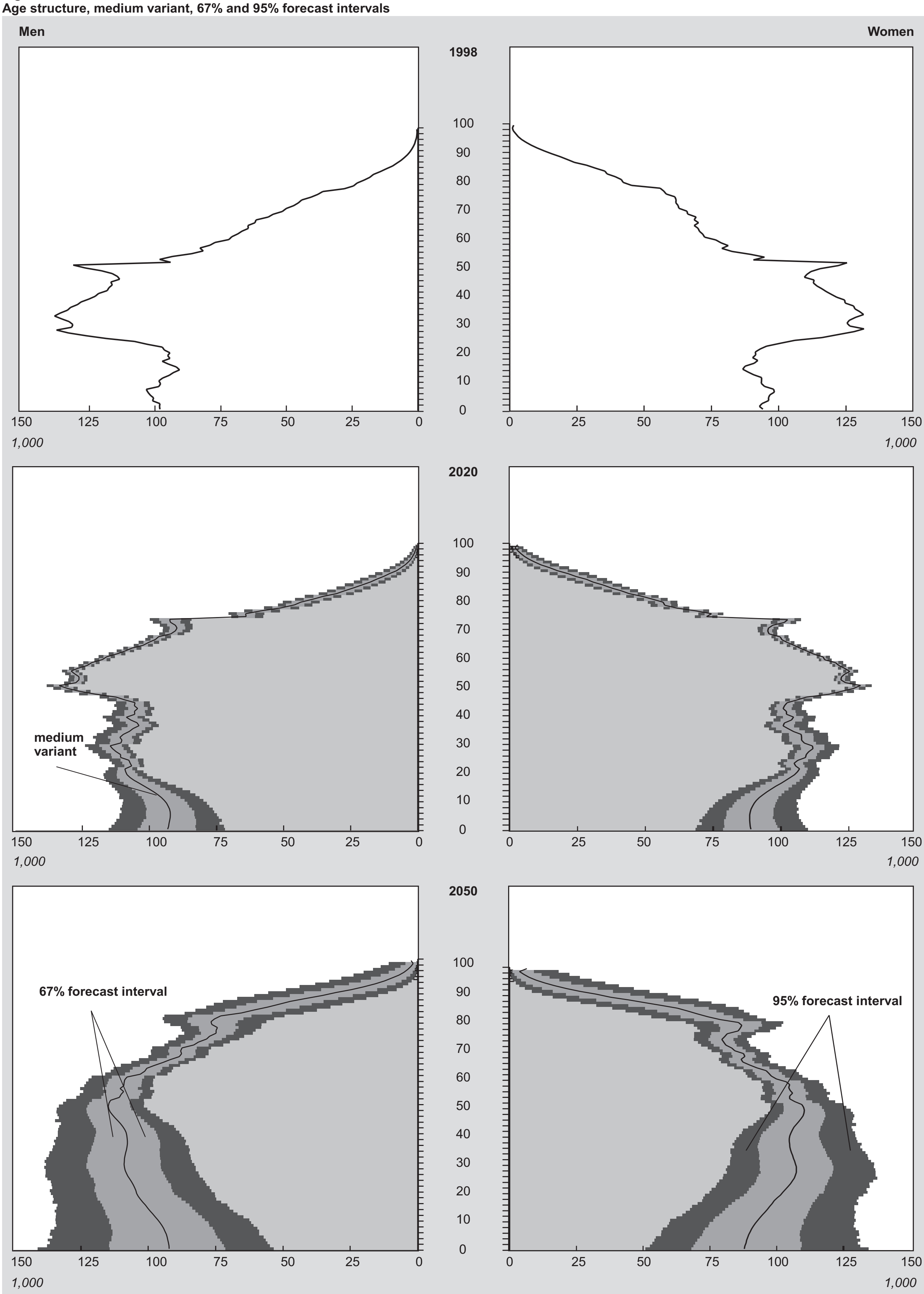
Figure 2
Births, deaths and migration, medium variant, 67% and 95% forecast intervals



4.2 Age structure

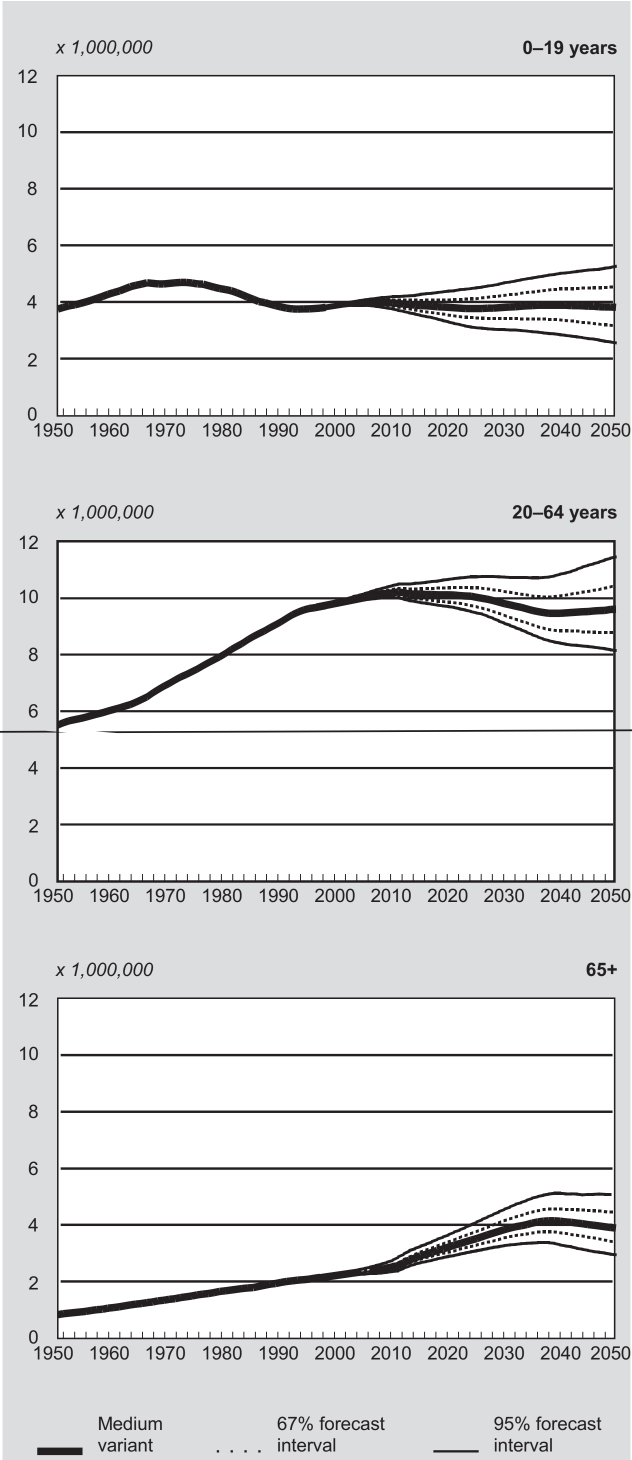
The degree of uncertainty of future developments differs between age groups. The uncertainty of the youngest age group is mainly affected by the uncertainty of fertility, while the uncertainty of the oldest age groups depends to a large extent on the uncertainty of mortality. Uncertainty of migration mainly affects the middle age groups. Figure 3 shows the 67% and 95% forecast intervals for each single year of age.
In 2020 the forecast intervals are relatively wide for those younger than 20 years. These are people who are born during the forecast period. The forecast intervals of the people aged 20 to 50 are

Figure 3
Age structure, medium variant, 67% and 95% forecast intervals



relatively small, as these people are already born at the beginning of the forecast period. Because the chances are very small that they will die before 2020 the uncertainty of their number is to a large extent determined by migration. Most people who migrate are younger than 40. The uncertainty of the number of people aged between 50 and 75 is very low. This number is not affected by fertility and hardly by migration. Furthermore, mortality does not yet play a significant role. Above age 75 the forecast intervals seem to be small, but this group is small. The relative size of the intervals is large, in particular for the oldest group.
The situation in 2050 is rather different from that in 2020. Obviously the forecast intervals are wider than in 2020, because the degree of

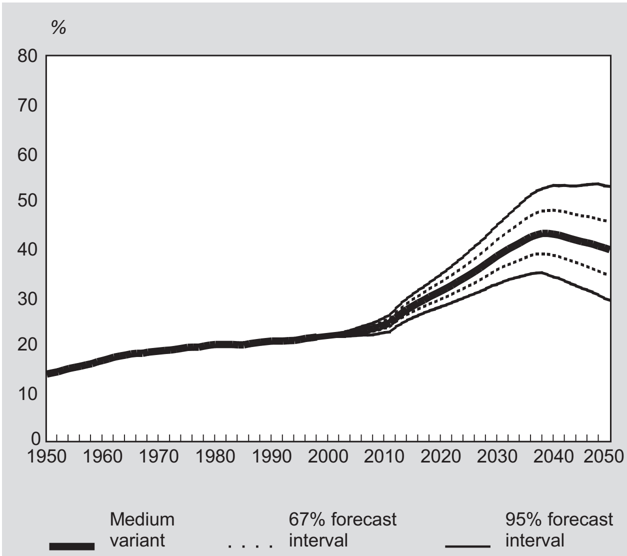
Figure 4
Size of age groups, medium variant, 67% and 95% forecast intervals



uncertainty increases with the length of the forecast horizon. Furthermore, different developments interact in the long run. Everyone under 50 was born during the forecast period. The uncertainty of the number of people younger than 20 is particularly large, since these are the children of people who were themselves born during the forecast period. For the older ages the degree of uncertainty is high due to the uncertainty of mortality and of the number of people who migrate in the decades before 2050.

Figure 4 shows the degree of uncertainty of the number of people aged 65 and over. According to the medium variant this age group will double, from just over 2 million in 1998 to more than 4 million in 2040. After 2040 the number will fall slightly to 3.8 million in 2050. The proportion of elderly people will increase from 13.5% to 23.4% in 2040. The forecast intervals show that it is very likely that this age group will continue to grow. Even the lower limit of the 95% forecast interval shows a strong increase by more than one million in the next 40 years.
In contrast, the size of the working age population will hardly grow at all. The medium variant projects that the number of 20–64 year-olds will rise by 400 thousand to 10.2 million in 2010. After 2010 a decline will set in, to reach 9.6 million in 2050. A decrease in the size of the potential labour force is just as likely as an increase. The 95% forecast interval in 2040, the peak of the ageing, ranges from 8.4 million to 10.8 million people.
The old age dependency ratio is an important factor in discussions about whether pension schemes will remain affordable in the future. This ratio is defined by the number of people aged 65 and over divided by the number of people aged between 20 and 65. Figure 5 shows that it is rather certain that the old age dependency ratio will increase. At present the ratio is about 22%. By 2040 it is quite likely to be over 34%, since the probability is 95% that the old age dependency ratio will range from 34% to 53%. This means that for every person older than 65 there will be at the most three, but perhaps even no more than two people of working age, compared with five people today. After 2040 the old age dependency ratio may start to decrease, but this is by no means certain.

Figure 5
Old age dependency ratio, medium variant, 67% and 95% forecast intervals



4.3 Forecast intervals versus low and high variants

A comparison of the 95% forecast interval of population size of the 1998-based stochastic forecasts with the low and high variants of the 1996-based deterministic forecasts shows that the latter variants are rather extreme. The interval between the variants of the population size in 2050 ranges from 12.7 million to 21.5 million. This interval is even wider than the 95% forecast interval of the stochastic forecasts. The interval between the low and high variants of the old age dependency ratio in 2050 according to the 1996-based forecasts ranges between 41% and 47%. The interval is smaller than the 67% forecast interval of the stochastic forecasts which ranges from 35% to 46%. So, in contrast to the interval between the variants of the population size, the interval between the variants for the old age dependency ratio is relatively small. Thus the low and high variants for population size and the old age dependency ratio correspond with different levels of probability. It is much more likely that the interval of population size in 2050 will cover the true value than the interval of the old age dependency ratio.

4.4 Using stochastic population forecasts

One of the consequences of using a stochastic population forecast is that 95% forecast intervals for different age groups do not sum up to the 95% forecast interval of the total population. While the sum of all age groups in the medium variant equals the medium variant of the total population, this does not hold for the forecast intervals. For example, the addition of the lower limits of the 95% forecast intervals of all individual ages in 2050 equals 13.3 million, whereas the lower limit of the 95% forecast interval of the total population size equals 14.4 million. This is caused by the fact that the numbers of people in the separate age groups are not perfectly correlated. For example, the degree of uncertainty of the number of 10-year-olds mainly depends on the uncertainty of fertility, while the uncertainty of the number of 50-year-olds is to a large extent affected by the uncertainty of migration. Because fertility rates and migration numbers are assumed to be independent, the numbers of 10-year-olds and 50-years olds are not perfectly correlated. Note that even though fertility rates and migration numbers are assumed to be independent, there is some correlation between the number of births and the number of migrants: if the number of immigrants is high, more children will be born in later years because some of the immigrants will have children in the Netherlands. Another reason for the imperfect correlation between age groups is that the correlation across time is not perfect. The numbers of births in successive years are not perfectly correlated and, as a consequence, neither are the different age groups.
Another consequence of using a stochastic population forecast is that for each ratio a separate interval needs to be calculated. The 95% forecast interval of the number of people aged 65 or over in 2050 ranges from 2.9 million to 5.0 million. The 95% interval for the total population ranges from 14.4 million to 20.7 million. The lower (upper) limit of the 95% forecast interval for the proportion of people aged 65 or over cannot be calculated by dividing the lower (upper) limit of the interval of the number of elderly people by the lower (upper) limit of the interval of the total population. This would result in an interval from 19.9% to 24.2% which is smaller than the 95% interval based on the simulations: range 16.9% to 28.1% (see table). The degree of uncertainty of ratios is underestimated by dividing the intervals of the separate age groups by the intervals of the total population. This implies of course that the intervals of the proportions for separate groups do not sum up to 100%.

5. Conclusions

On the basis of assumptions about the probability distributions of future fertility, mortality, and migration, a probability distribution of the future population size and age structure can be calculated. Even though the ‘true’ probability is not known, as a probability distribution of the future is a forecast, a probability distribution seems to be the best way in which forecasters can communicate the uncertainty of forecasts to the users. A probability distribution allows users to make their own choices. One major problem in using deterministic variants is that they reflect different degrees of uncertainty for different outcomes. For example, high and low variants produce intervals for population size that correspond with a higher probability than intervals for the old age dependency ratio. The opposite is true if so-called young and old variants are published. Thus different variants may be useful for different users. One major benefit of stochastic population forecasts is that the intervals for all outcomes are comparable. Thus the outcomes can be used by any type of user.

Stochastic population forecasts are based on assumptions about the degree of uncertainty of future fertility, mortality, and migration. These assumptions can be based on time series models, on an analysis of errors of historic forecasts, and on judgement. Even though time series models and analyses of historic forecast errors play an important role in the specification of the assumptions, judgement based on an analysis of the underlying processes is the determining factor in assessing forecast intervals for the Dutch population forecasts.

In the 1998 Dutch population forecasts it is assumed that the width of the 95% forecast interval for fertility in 2050 equals 1.2 children per woman, the interval for life expectancy at birth equals 12 years and the interval for net migration equals 66 thousand. Furthermore it is assumed that forecast errors of fertility, mortality, and migration are serially correlated, that for each component there is perfect correlation across age and that the three components are independent. Simulations based on these assumptions lead to a 67% forecast interval of total population size of 3.2 million in 2050. This amounts to 19% of the medium variant. This interval is smaller than the interval between the low and high variants of the 1996-based population forecasts. This is explained by the fact that the low and high variants are deterministic variants which imply perfect correlation between the three components of population growth, fertility, mortality, and migration, and perfect correlation between the values of these components across forecast years. Although the intervals between the low and high variants for fertility, life expectancy and net migration were assumed to roughly correspond with 67% forecast intervals, the interval between the low and high variants of total population size was much wider than the 67% interval for population size. According to the 1996-based population forecasts the interval between the low and high variants of population size in 2050 equals 8.8 million in 2050, which is twice the width of the 67% forecast interval according to the 1998-based stochastic population forecasts.

In interpreting the results of a stochastic population forecast it should be noted that the lower and upper limits of the 95% forecast interval of total population size are not equal to the sum of the lower and upper limits of the 95% forecast intervals of the separate age groups. For example, the sum of the lower limits of the intervals of all individual ages for the year 2050 equals 13.3 million persons, whereas the lower limit of the 95% interval of total population size equals 14.4 million. Similarly the sum of the upper limits equals 22.0 million, whereas the upper limit of the 95% interval equals 20.7 million. This is due to the fact that the age groups are not perfectly correlated. One reason for this is that the components of population growth are not perfectly correlated. As the separate components affect different age groups (the uncertainty of fertility influences the uncertainty of numbers of young people, while mortality mainly affects the size of the elderly population), the forecast intervals of separate age categories are not perfectly correlated. Another reason is that there is no perfect correlation across time. The number of births in one year is not perfectly correlated with the number in the previous year. As a result the size of successive age groups is not perfectly correlated.

For further information or comments on this contribution, please contact Joop de Beer: jber@cbs.nl, or Maarten Alders: mals@cbs.nl.

Table 1
Age groups, medium variant and 95% forecast interval

	0–19 years			20–64 years			65+			population size		
	medium variant	lower limit	upper limit	medium variant	lower limit	upper limit	medium variant	lower limit	upper limit	medium variant	lower limit	upper limit
	x million											
1998	3.8			9.7			2.1			15.7		
2000	3.9	3.9	3.9	9.8	9.8	9.8	2.2	2.1	2.2	15.8	15.8	15.9
2010	4.0	3.8	4.2	10.2	10.0	10.4	2.5	2.3	2.6	16.6	16.2	17.1
2020	3.8	3.3	4.4	10.1	9.7	10.7	3.2	2.8	3.5	17.1	16.2	18.2
2030	3.8	3.0	4.7	9.8	9.1	10.7	3.8	3.2	4.4	17.4	15.9	19.2
2040	3.9	2.8	5.0	9.5	8.4	10.8	4.1	3.2	5.0	17.4	15.2	20.1
2050	3.8	2.6	5.3	9.6	8.1	11.5	3.8	2.9	5.0	17.2	14.4	20.7
	as a percentage of total population											
1998	24.3			62.2			13.5					
2000	24.4	24.3	24.4	62.0	62.0	62.1	13.6	13.6	13.7			
2010	23.9	22.9	24.9	61.3	60.4	62.2	14.8	14.0	15.7			
2020	22.3	19.9	24.6	59.2	57.4	61.4	18.5	16.7	20.3			
2030	21.9	18.6	25.4	56.4	53.8	59.3	21.7	18.7	24.8			
2040	22.3	17.9	26.5	54.3	51.1	57.9	23.4	19.2	27.9			
2050	22.1	16.9	27.1	55.7	51.8	59.6	22.2	16.9	28.1			

References

Alho, J.M., 1998. *A stochastic forecast of the population of Finland*. Reviews 1998/4, Statistics Finland, Helsinki.

de Beer, J. and A. de Jong, 1996. National population scenarios for countries of the European Economic Area. In: *Maandstatistiek van de bevolking*, July 1996, 7–19.

de Beer, J. and M. Alders, 1999. *Probabilistic population and household forecasts for the Netherlands*. Paper presented at the Joint ECE/Eurostat Work Session on Demographic Projections, 3–7 May 1999, Perugia.

de Beer, J., 1999. Bevolkingsprognose 1998-2050 (Population Forecasts 1998-2050). In: *Maandstatistiek van de bevolking*, January 1999, 8–19.

Hanika, A., W. Lutz and S. Scherbov, 1997. Ein probabilistischer Ansatz zur Bevölkerungsvorausschätzung für Österreich. In: *Statistische Nachrichten*, 984–988.

Lee, R. and S. Tuljapurkar, 1994. Stochastic population forecasts for the United States: beyond high, medium, and low. In: *Journal of the American Statistical Association* 89, 1175–1189.

Lutz, W. and S. Scherbov, 1998a. Probabilistische Bevölkerungsprognosen für Deutschland. In: *Zeitschrift für Bevölkerungswissenschaft* 23, 83–109.

Lutz, W. and S. Scherbov, 1998b. An expert-based framework for probabilistic national population projections: the example of Austria. In: *European Journal of Population* 14, 1–17.

Sorvillo, M.P., 1999. *Fertility projections: an international comparison*. Paper presented at the Joint ECE/Eurostat Work Session on Demographic Projections, 3–7 May 1999, Perugia.

The importance of innovation for company performance

Luuk Klomp and George van Leeuwen

Using the second Community Innovation Survey (CIS-2) for the Netherlands we analyse the input and output stages of the innovation process and the links between the innovation process and economic performance. We investigate the existence of feedback links running from past economic performance to the input and the output stage of the innovation process and compare the results of a singe-equations approach with the results obtained from a simultaneous-equations model.

Keywords: innovation, simultaneous-equations models, economic performance.

1. Introduction

The perception of knowledge as an important production factor is becoming more and more widespread. The boost of literature from policy makers and scientists alike is a good indicator for the interest in the knowledge-based economy and thus in the innovation process, (e.g. Kleinknecht, 1996, Brouwer, 1997, OECD, 1998 and 1999, The Dutch Ministry of Economic Affairs, 1999, Acs, Carlsson and Karlsson, 1999, and Audretsch and Thurik, 1999). The availability of new and improved indicators collected in the Community Innovation Surveys (CIS) provided an opportunity to study innovation as a separate process, with R&D expenditures as the most important input into innovation and newly created or improved products or process innovation as the output of the innovation process. The data referring to the technological environment of companies and to the organisational aspects of their innovation processes created a major impetus for the explanation of differences in innovation activity as well as an analysis of the importance of company-specific innovation characteristics for the output of the innovation process and the effects of the innovation output on company performance. Recently, the interest in the innovation process has shifted from the input (R&D) to the output stage (realised innovations). Moreover, the focus is now also on the linkages between the three stages of the innovation process: input, throughput and output, with the role of innovation as a driving factor of long-term macro-economic growth taken for granted.

The purpose of this paper is to investigate whether companies that are more active in the innovation process show better economic performance [1]. The novelty of our approach is that we link innovation and company performance using a so-called simultaneous- equations model. In the traditional approach, one stage of the innovation process (for instance R&D expenditures or the realisation of innovations) has been isolated and subsequently linked to economic performance, thereby neglecting the joint dependence of measures of innovativeness on company specific innovation characteristics and the joint dependence of innovative output and the overall economic performance of companies. We analyse the input and the ouput stage of the innovation process simultaneously, including the feedback mechanisms between the stages, and then confront the innovation process with indicators of economic performance, like turnover growth and employment growth.

We matched the CIS-2 data set to data sets from the production surveys to be able to link innovative activities to economic performance at the company level. A quick inspection of the data showed that innovative companies outperformed their non-innovative counterparts, although the differences for turnover growth are more pronounced than those for employment growth. We then tried to model the relationships between the different stages of the innovation process and economic performance. The model of Kline and Rosenberg was used as a basis. This, being a ‘chain-link model’, captures feedback mechanisms within the innovation process. The model is an extension of the traditional linear model where it is assumed that innovation output follows the input and throughput phase without paying attention to feedback loops.

Our empirical model is a so-called simultaneous-equations model of innovation input, innovation output, sales growth and employment growth, in which we include links between the input, throughput and output stage of the innovation process as well as links between the innovation process and the economic activity. The advantage of estimating a simultaneous model over single-equations models for either innovation input or output is that, for instance, the effect of technological opportunities open to the company are disentangled into two parts: an impact on the innovation input and an impact on the innovation output. We observe that the results following both approaches – the simultaneous-equations approach and the single-equations approach – diverged. A notable difference is that the feedback loop that ties together the innovation process and overall economic performance changes from the output stage to the input stage of the innovation process when one takes into account the joint dependence of the different stages. Therefore, we recommend that simultaneous-equations models should be used and developed in future work.

2. Comparing the performance of innovating and non-innovating companies

2.1 Matching CIS-2 and the production survey data

To select the data used in the econometric part of the article, we started out from the 10,664 companies that responded to CIS-2, which covers the period 1994–1996. Most of these companies were also covered in the Statistics Netherlands’ production surveys, which provide data on total sales, employment, value added and profitability. However, a number of responding companies belong to sectors for which no production surveys were available, and for these we had to use the data on total sales and employment in 1994 and 1996 as collected in CIS-2. We used the 1996 employment data in CIS-2 for all companies to check the comparability of the unit of observation in both surveys. On the basis of this consistency check we decided to reject 1,250 companies because of the large discrepancies in the employment figures [2]. In addition we omitted 1,032 companies from the analysis because their data on total sales and employment were missing in the production surveys of 1994 or 1996, and 54 CIS-2 respondents were rejected at this stage because of an implausible score for their innovation intensity [3]. After this preliminary data cleansing 8,328 companies were selected with a complete record of total sales and employment for 1994 and 1996.

Table 1 presents a breakdown of the initially selected companies according to some response characteristics. About 48% of the selected companies reported that they had implemented product or process innovation in 1994–1996. The rate of innovativeness (measured by the number of innovating companies as a percentage of all companies) varies between 66 % for manufacturing and 30 % for other sectors and increases with company size in all sectors. Contrary to rate of innovativeness, which increases with company size, the share of companies reporting innovative output does not differ much between classes of company size. So on the face of it, for companies which have implemented product or process innovation, the success rate of this innovation does not depend on the size of the company. The level of innovative output is measured

Table 1
The selected companies with complete data on sales and employment

Sector of principal activity and classes of company size [1]	Number of companies	Of which:			
		Non-innovating companies	Innovating Companies	Companies performing R&D on a permanent basis	Innovating companies with innovative output [2]
Manufacturing	2,969	1,002	1,967	1,074	1,652
small	1,296	607	689	243	564
medium- sized	1,313	345	968	599	821
large	360	50	310	232	267
Services	4,170	2,496	1,674	571	
small	1,852	1,264	588	144	
medium-sized	1,892	1,076	816	295	
large	426	156	270	132	
Other industries [3]	1,189	835	354	91	276
small	529	402	127	22	96
medium-sized	573	392	181	50	144
large	87	41	46	19	36
All sectors	8,328	4,333	3,995	1,736	1,928
small	3,677	2,273	1,404	409	660
medium-sized	3,778	1,813	1,965	944	965
large	873	247	626	383	303

[1] Small companies: 10-50 employees;
Medium-sized companies: 50-200 employees;
Large companies: 200 or more employees.
[2] 0–100% of products sold in 1996 are 'new to the company'.
[3] Agriculture, forestry and fishing, mineral extraction, public utilities, and construction industry.

by the share in total sales of products which are 'new to the company'. This variable is available for manufacturing companies and companies belonging to other industries but could not be obtained for companies in the service industries. Lastly, Table 1 confirms the well-known empirical fact that formal R&D activities are predominantly concentrated in manufacturing and that the probability of performing R&D on a permanent basis also is size dependent.

2.2 A comparison of the growth rates for total sales and employment

Although the importance of innovation for economic activity is often widely acknowledged, this does not imply that by definition non-innovating companies perform worse than their innovating counterparts. One can not even exclude the possibility that non-innovating companies perform better on average. Matching the CIS-2 data and the production survey data enables us to compare the performance of innovating and non-innovating companies. A clear picture emerges when we look at the distributions of the two performance measures presented in Figures 1a and 1b. Evidently, innovating companies performed better than non-innovating companies on total sales growth but the differences are less pronounced for the growth rates of employment. However, the main message from the distributions presented in these figures is the overwhelming heterogeneity in company performance for both the innovating and the non-innovating companies. Consequently, it is expected that technological innovation will not be able to explain all observable heterogeneity.

Figure 1a
The distribution of total sales growth (n = 8328)

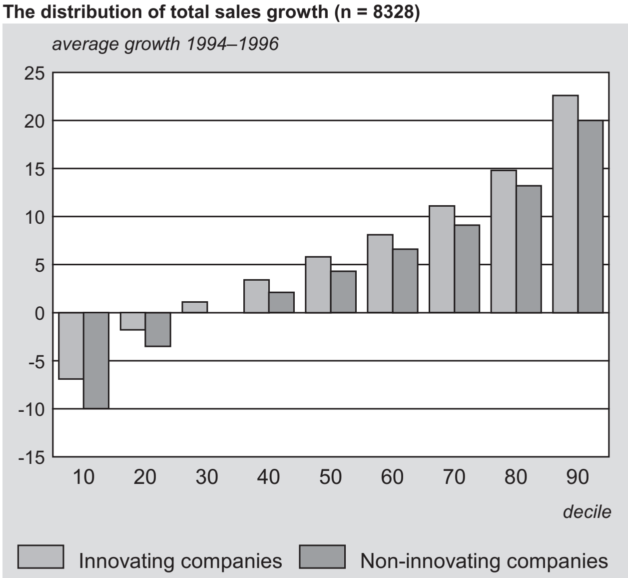


Figure 1b
The distribution of employment growth (n = 8328)

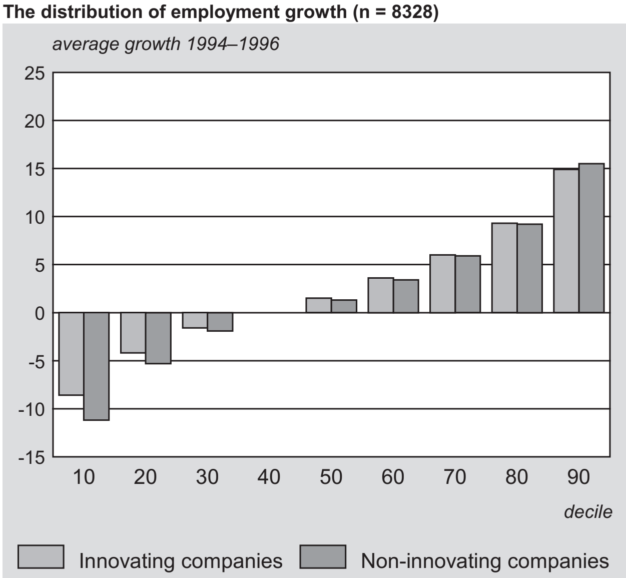
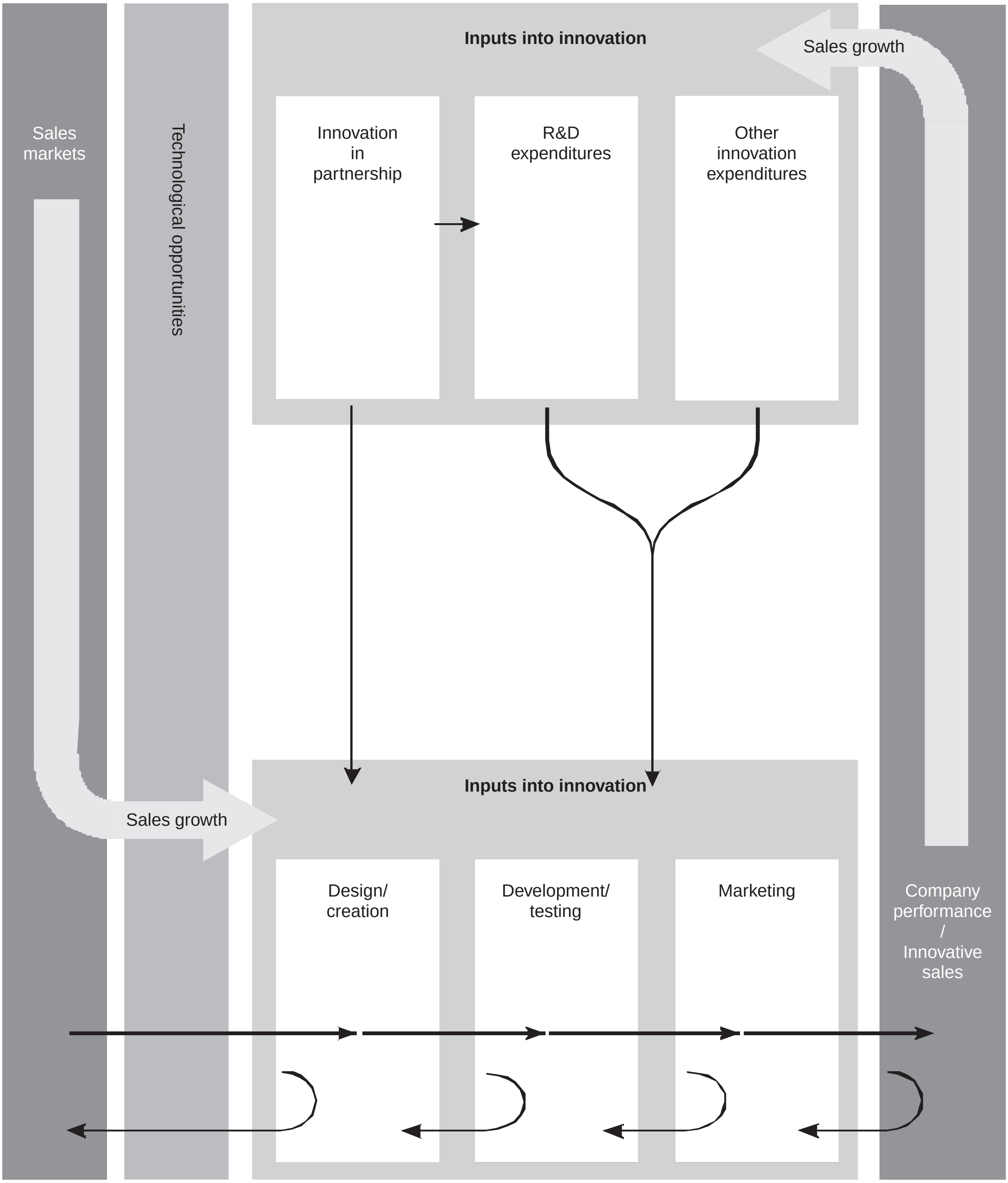


Figure 2
The innovation process and company performance



3. A tentative structural model for the analysis of innovation and economic performance

Using a formal model of the innovation system as a frame of reference and taking into account the data at hand, there are different routes to the empirical testing of the many dimensions underlying the relationship between innovation and overall economic performance. In order to enhance further discussion and as a motivation for the route chosen in this article, we first present a condensed and adapted summary of the innovation model of Kline and Rosenberg (1986). Figure 2 presents some of the main features of their well-known innovation model which we have adopted for the specification of the empirical models. The central part of the figure summarises the innovation process and the surrounding bars indicate its positioning within the technological and economic environment. The two bars at the left side of Figure 2 represent a company's market potential and technological environment respectively. In the empirical application we use sectoral dummy variables to take account of the sector-specific market potentials open to a company. The technological environment of a company is represented in the empirical model with the help of different company-level data. We include two variables which are derived from a factor analysis of the technological opportunities open to the company by applying a principal components analysis to the data collected on the use of information sources. Following Felder et al. (1996), we use two factors to represent the use of technological opportunities: technological information sourced from 'science' and technological information sourced from 'other companies' such as suppliers or customers or competitors.

A company's technological environment may also affect its organisational arrangements. In the empirical model we use proxy variables which refer to organisational aspects in order to take account of the notion that a company may absorb knowledge from the environment via supplier-producer-customer-interactions, or may build up and maintain its own knowledge-base via R&D investment and R&D co-operation [4)]. The latter two important examples are indicated separately in Figure 2. We use two dummy variables to indicate the presence of permanent R&D facilities and the emergence of innovation in partnerships with other companies. The relation between the presence of permanent R&D facilities, 'innovation in partnership' and the two technological opportunity variables (the factors 'science' and 'other companies') can be outlined as follows. One may expect a 'cost-push' effect on innovation expenditure of the technological opportunity factor 'science' because of the absorptive capacity argument (see e.g. Cohen and Levinthal, 1989, and Leiponen, 1999). A co-operation between R&D companies and research institutes and universities requires relatively high internal research skills in order to assimilate the fruits of the co-operation and to internalise and commercialise the knowledge created during the co-operation. Co-operation with suppliers, customers and competitors, by contrast, is expected to have lower research competence requirements, a smaller impact on the organisation of companies, and thus a lower 'cost-push' effect on innovation expenditure than the factor 'science'. On the other hand, one can imagine that non-R&D co-operation affects innovation throughput more directly than R&D co-operation and consequently may have a larger effect on the level of innovation output than the technological opportunity factor 'science'.

The central part of Figure 2 represents the links within the innovation system itself and the links between a company's innovation system and its technological environment. Given the market potentials and the technological opportunities open to the company, the use of technological opportunities may both affect the level of innovation expenditures, as well as the innovation throughput directly. Therefore, the extent of innovation success is expected to depend on investment in innovation, e.g. by building up or maintaining own R&D capital stock or investment in the exploitation of technological opportunities via R&D co-operation, but also on more informal not R&D driven co-operation with other companies. The upper block of the central part of the figure represents the inputs into innovation, as measured by innovation expenditures. Notice that it is assumed that R&D co-operation increases R&D expenditures. The lower block represents the well-known functional breakdown of the innovation production process (a functional split-up of the innovation activities) which is adopted from the traditional linear innovation model. The interactions between the different innovation activities are indicated with dark arrows. Because we only have data pertaining to the final result of innovation production process, we cannot incorporate the underlying activities into the empirical models.

Lastly, the bar on the right-hand side of Figure 2 indicates that innovation output as measured by the share of new or improved products in total sales contributes to a company's total sales growth and thus affects its overall economic performance which in turn is assumed to affect the inputs into innovation as measured by the level of innovation expenditures. Notice further, that in addition to this link it is also assumed that a company's overall sales growth may affect the level of innovation output directly. These feedback links close our tentative model. In a broad view, the model links a company's own innovation performance to the exogenously given market potentials and to the availability of technological opportunities. Moreover, the model also establishes a framework for the empirical testing of the existence of a persistent relation between its own overall economic performance and its level of innovativeness. The latter will be represented by feedback links running from a company's own record of sales performance, either to the input stage or to the output stage of the innovation process (or the throughput stage of the complete system if we take a broader view). These feedback loops explicitly acknowledge the dynamic nature of the system.

Ideally, the proper empirical testing of all lead and lag structures involved would require the availability of longitudinal company level innovation data. Some longitudinal aspects are present in CIS-2, but in a crude way and moreover the longitudinal aspects are restricted to the measures of overall economic performance only. With the available data, the best we can do is to take account of the interdependency of the different stages of the system. Besides establishing a framework for the testing of a persistent relation between innovation and economic performance, the model also allows the empirical testing of other hypotheses. With the different variables at hand we can estimate separately equations for the levels of innovation intensity and innovation output. We can also start from the assumption that the inputs into innovation, innovation output and the overall economic performance are jointly determined, and then test for example whether technological opportunities and past performance have a separate role in the explanation of differences in innovation output, apart from their impact on innovation expenditures.

4. Model specification

4.1 The choice of the endogenous variables

For the econometric estimates presented below we use the data of the innovating companies. Following the exposition given in section 3, our main objective is to estimate the contribution of innovation to growth in sales and employment and to investigate the importance of company-specific innovation characteristics and the existence of a persistent relation between innovation and overall economic performance. The basic assumption is that a company's total sales and employment growth depends on its innovation output as measured by the share of new or improved products in total sales. This measure refers to the level of innovative output in 1996. Consequently, we have to assume that a company's innovation output and its total sales growth and employment growth are jointly endogenous. We also consider the inputs into innovation to be an endogenous variable because the variables which measure the resources devoted to innovation refer to 1996 as well. There are different ways to choose the measure of the inputs into innovation. CIS-2 uses a rather broad definition of the sources devoted to innovation, with R&D expenditures as the most important component. In spite of the well-known deficiencies of R&D as an indicator for the innovation process, it remains one of the most frequently applied measures (see for instance Kleinknecht (1996) and Kleinknecht and Bain (1993) for a discussion of the deficiencies of the R&D indicator) [5)]. We could restrict ourselves to R&D expenditures (both intramural and extramural R&D), but in doing so we would neglect more than half of all innovation expenditures. Therefore, we use the total of all innovation expenditure components. Following general practice, a company's total innovation expenditure is scaled by its total sales and the resulting innovation intensity is used as the measure for the inputs into innovation [6)]. Thus, the jointly endogenous variables in the model are the innovation intensity in 1996, the quantitative innovation-output indicator (the share of new or improved products in total

sales in 1996) and the average growth rate of total sales and employment in 1994–1996. These variables are labelled as *ININT*, *P*, $\Delta logS$ and $\Delta logE$ respectively.
In the empirical application we use the 'log-odds' transformation of the quantitative innovation-output indicator (labelled as *LOP*) [7]. The 'log-odds' ratio will be included as an explanatory variable in the equations for total sales growth and employment growth. The 'log-odds' ratio was chosen so that the predicted value of the throughput measure lies in between 0% and 100%. Then, by using (the endogenous) 'log-odds' ratio as an explanatory variable in the equation for a company's total sales and employment growth we can directly infer from the estimates of the simultaneous model the effects on total sales growth and employment growth of changes in the dichotomous innovation variables. An obvious disadvantage of the 'log-odds' ratio is that companies with a zero or a hundred percent share of innovative sales in principle cannot be taken into account. To remedy this problem, we apply a Tobit analysis to the innovation-output equation and we also re-estimate the simultaneous model using all observations after the imputation of a value of 0.001 or 0.999 for *P* for the companies with a quantitative innovation-output indicator equalling 0% and 100% respectively.

4.2 The choice of the exogenous variables

The joint endogeneity of the inputs into innovation, the level of innovation output and the economic performance in general deliver an interrelated sequence of endogenous variables, each of which may be affected by different exogenous variables. However, it is not immediately clear which variables should be considered exogenous for the explanation of an arbitrarily chosen endogenous variable. Another instance of the specification problem is that our candidate (exogenous) variables for the innovation characteristics of companies are of a qualitative nature and in many cases dichotomous or derived from the response to questions that are not mutually exclusive. In addition, some of the qualitative innovation variables may be more or less closely related to other variables. As an example we mention the variables that refer to the use of technological opportunities derived from information sources and the variables referring to innovation in partnership. A company that innovates in co-operation with universities or research institutes will probably also consider the associated information sources to be (very) important.
The selection of the exogenous variables was guided by the following considerations. We make a distinction between financial resources, variables that reflect some of the organisational aspects of the innovation process, variables that refer to the objectives of innovation, and industry-specific and size characteristics. For many companies the innovation expenditures consist largely of investment components, for example expenditures on in-house R&D, licences and patents and equipment purchased for the implementation of process innovation. We assume that these investment type expenditures are affected by the availability of financial resources and therefore take into account two financial variables: the ratio of cash-flow to total sales for 1994 (CF_{1994}) and a dummy variable referring to the awarding of innovation subsidies (D_{subs}).
Furthermore, and following the reasoning presented in section 3, we use two dummy variables indicating the presence of permanent R&D facilities ($D_{R\&D}$) and innovation in partnerships ($D_{co\text{-}op}$) in order to represent a company's organisational arrangements regarding the innovation process. Similar to CIS-1, the use of technological information is measured on a four-point scale (not used, of little importance, important and very important), and companies were asked to rate the importance of thirteen information sources. Using a principal components analysis we transformed these data into two continuous variables, representing the use of technological opportunities offered by the 'scientific world' and other companies such as customers, suppliers and competitors. These factors will be denoted by '*SCIENCE*' and '*OTHER*' respectively [8].

By collapsing the response to the CIS-2 questions on objectives of innovation processes into four dummy variables, we also constructed four exogenous variables that represent the objectives underlying the implementation of innovations. If the replacement of old products or the improvement of the quality of existing products or the extension of market shares and product ranges are rated as (very) important, the dummy variable D_{dpull} takes on a value of one (and zero otherwise). Similarly, we constructed three 'cost-push' dummy variables for the objectives 'economising on labour cost considered (very) important' (D_{push1}), 'economising on the cost of material inputs and energy considered (very) important' (D_{push2}) and 'legal regulations considered (very) important' (D_{push3}). The list of exogenous variables is completed by including in the models the logarithm of the age of the companies in January 1994 (LA_{1994}), a set of industry dummy variables, the logarithm of total sales in 1994 (LS_{1994}) and a set of dummy variables in order to capture size effects [9].
As a consequence of the choice of the endogenous and exogenous variables not all innovating companies could be used in the estimation procedure. Of the 3,995 innovating companies with a complete record for total sales and employment growth, 936 had to be rejected due to missing or implausible data for the exogenous variables [10]. Data on the inputs into innovation were available for the remaining 3,059 companies and also data on the level of innovation output for 1,977 of these [11].

5. The estimation results

5.1 Single-equations estimation

This section presents the econometric estimates for various specifications. As discussed above our aim was to estimate a simultaneous model for the four jointly endogenous variables. Before discussing the variables that enter the simultaneous model, we look first at the single-equation estimates for the inputs into innovation and for the innovation output as measured by the share of innovative sales. In addition to the exogenous variables discussed in the preceding section we also included the average growth rate for total sales in both equations in order to capture the feedback effect from general company performance.

The determinants of the innovation inputs
The first two columns of Table 2 present the single-equation estimates for the input stage of the innovation process using the whole sample of innovating companies. The most significant explanatory variables are those that refer to the use of technological opportunities and the presence of permanent R&D facilities. It is found that the impetus on the resources devoted to innovation from the relationship with customers, suppliers and competitors is smaller than the 'cost-push' effect associated with the technological opportunities 'science'. In addition, innovating in partnership also has a significantly positive effect on the financial sources devoted to innovation.
These results may be interpreted as a corroboration of the absorptive capacity hypothesis which conjectures that co-operation with 'science' requires higher internal R&D skills and thus higher R&D expenditures [12]. As for the financial variables, we found a significant positive estimate for the effects of internal cash flows, but also that the awarding of innovation subsidies contributes significantly and with the expected sign to the inputs into innovation. Another notable result is that innovation intensities decrease with size (measured by the coefficient of LS_{1994}) as well as with age, indicating that younger companies spend relatively more resources on innovation than older companies. Notice further, that the coefficients of the exogenous variables representing the objectives underlying the implementation of product or process innovation are not significantly different from zero in all cases. Our last comment concerns the estimates of the coefficients of the variable 'sales

Table 2
Single-equation estimates of the determinants of innovation inputs and output

Explanatory variables	Innovation intensity		Innovation output					
			Two-limit TOBIT model [1]		Generalised TOBIT model [2]			
					Probit part		OLS part	
	Est.	t-value	Est.	t-value	Est.	t-value	Est.	t-value
Number of companies	3,059		1,977		3,059		1,662	
Continuous variables:								
CF_{1994}	0.039	5.5			–0.003	–1.9		
ININT			0.007	6.4	0.008	1.4	0.045	6.8
LS_{1994}	–0.650	–6.0	0.001	0.1	–0.125	–5.5	0.003	0.1
ΔlogS	–0.010	–2.1	0.001	2.4	–0.003	–1.9	0.006	2.7
SCIENCE	0.578	5.6	0.017	2.7	–0.012	–0.4	0.079	2.2
OTHER	0.340	3.3	0.034	5.1	0.058	1.9	0.184	4.3
LA_{1994}	–0.437	–4.0	–0.013	–1.8	0.161	5.2	–0.124	–2.4
Dummy variables:								
D_{subs}	0.959	4.1			0.460	6.7		
$D_{R\&D}$	1.087	4.9	0.097	7.1	0.226	3.4	0.399	3.9
$D_{co\text{-}op}$	0.538	2.4	0.038	2.6	–0.022	–0.3	0.103	1.1
D_{dpull}	0.428	1.1	0.157	5.2	0.540	5.0	0.318	1.5
D_{push1}	0.294	1.4	–0.007	–0.5	–0.135	–2.3	0.140	1.6
D_{push2}	–0.171	–0.8	–0.007	–0.4	0.195	3.1	–0.007	–0.1
D_{push3}	–0.002	–0.1	0.004	0.3	0.070	1.1	–0.036	–0.4
Pavitt1	–1.571	–2.3	0.040	1.9	1.999	14.1	–0.415	–1.3
Pavitt2	1.176	1.8	0.090	3.8	1.728	11.6	0.051	0.2
Pavitt3	–2.265	–3.0	–0.021	–1.0	1.417	12.7	–0.267	–0.8
Pavitt4	–2.417	–3.7	–0.004	–0.2	1.289	12.4	–0.156	–0.5
Pavitt5	–1.945	–3.0	0.036	1.6	1.349	12.0	–0.250	–0.9
Pavitt6	–2.004	–3.0						
Pavitt7	–1.443	–2.4	0.117	5.8	1.746	13.4	0.446	1.4
Pavitt8	–2.926	–6.1						
Constant term	12.093	9.8	0.028	0.4	–0.909	–3.1	–1.419	–2.2
Dummy variables for sector-size interactions included	yes		no		no		yes	
σ			0.263	55.9			1.484	54.4
ρ							–0.152	–0.8
Log Likelihood			–443				–4,408	
R^2	0.133							

[1] Dependent variable is the share of innovative sales in total sales (the innovative-output indicator P).
[2] Dependent variable is the 'log-odds' ratio of P.

growth'. In Table 2 these estimates are significantly negative and thus contradict the Schmookler hypothesis of a positive feedback effect from own past performance to innovativeness (see e.g. Schmookler, 1966). This hypothesis has also frequently been tested on the level of innovation output (see e.g. Brouwer and Kleinknecht, 1997a, Brouwer et al, 1997b and Cosh et al., 1999). Therefore, it seems natural to proceed along this line of research by estimating a model for the extent of innovation success, measured by the share of innovative sales in total sales.

The determinants of innovation output
The determinants of the level of innovation output are estimated by applying several variants of Tobit models. In doing so we also investigate the possibility of a selectivity bias caused by the double truncation for the 'log-odds' ratio of the quantitative innovation-output indicator or the missing data for the dependent variable (for companies in business services). Firstly, we use a two-limit Tobit model (see e.g. Maddala, 1994) for the 1977 companies for which we have the quantitative innovation-output indicator (*P*). Secondly, we estimate a generalised Tobit model for the sample of companies with a complete record of the explanatory variables. In this approach we estimate simultaneously the probability of observing a score for the innovative-output indicator that lies between 0 and 1 (the Probit part of the model) and the actual level of this indicator (the OLS part of the model) [13].
In the generalised Tobit model the variables that determine the probability of observing *LOP* may differ from the set of variables that explains its actual level, if observed. Thus, different choices can be made from the list of exogenous variables discussed in Section 4. With the exception of the endogenous variables 'employment growth' we include all variables in the Probit part of the generalised Tobit model, but we assume that financial variables do not affect the level of innovative sales directly and they are therefore excluded from the OLS part of the model [14].
The estimation results for the two Tobit models are presented in the other columns of Table 2. It should be noted that the estimates of the two-limit Tobit model and the OLS-part of the generalised Tobit model are not directly comparable because of the different dependent variables. The two-limit Tobit model has the advantage that the estimates are directly interpretable as the contribution to the share of innovative sales in total sales of the corresponding variables. We take this model as a starting point for commenting on the estimation results. It is found that the level of innovation intensity (*INNINT*), performing R&D on a permanent basis ($D_{R\&D}$), the

objective ‘demand factors considered important’ (D_{dpull}) and the use of technological opportunities offered by other companies (*OTHER*) are the most significant variables. An interesting result is that the (significant) contribution to *P* as well as to *LOP* of *SCIENCE* is half the size of the estimated coefficient of *OTHER,* indicating that the interactions with customers, suppliers and competitors contribute more to innovation throughput than the use of information from the ‘science’ sector. Furthermore the results show that the share of innovative sales in total sales decreases with age, indicating that younger companies have higher innovative output than older ones.

The estimation results for the OLS-part of the generalised Tobit model show that – with the exception of the coefficients of the dummy variables for the ‘demand-pull’ objective and ‘innovation in co-operation’ – the same variables as in the two-limit Tobit model are statistically significant. The insignificance of the ‘demand-pull’ dummy variable can be explained by the fact that only relatively few companies rated demand factors as an unimportant objective for implementing product or process innovation. This result is confirmed by the corresponding estimate of the Probit part of the model, which shows a strong and positive correlation between this objective and the probability of observing innovative sales. The general picture that emerges from the other Probit estimates is also reasonable in light of the underlying selection mechanism. The empirical facts that – on average – ‘business service’ companies are smaller, younger, perform R&D on a permanent basis less frequently and also showed a higher growth rate for their total sales than manufacturing companies [15)] is confirmed in the Probit estimates of the generalised Tobit model. A negative estimate in the Probit part of the generalised Tobit model was found for the coefficients of $\Delta logS$, LS_{1994} and a positive estimate for the coefficient of LA_{1994} and $D_{R\&D}$. Furthermore, and according to the sign of the Probit estimates for D_{push1} and $D_{push2,}$ the objectives ‘economising on labour cost’ and ‘economising on the cost of material inputs’ discriminated in favour of ‘business-service’ and manufacturing companies respectively. This result seems also plausible, if one takes into account the inter-industry differences in the cost shares of the corresponding production factors. Furthermore, the probability of having observed innovative sales is negatively correlated with past internal cash-flows (CF_{1994}) and positively (but not statistically significant) correlated with the inputs into innovation (*ININT*), indicating that – on average – the sample with innovative sales does not consist of companies which were more profitable or which spent more resources on innovation than the ‘business service’ companies.

A simultaneous-equations model for the relation between innovation and company performance

The estimates of the two-limit Tobit model and the OLS part of the generalised Tobit model of Table 2 indicate that – conditional on having innovative sales – a company’s sales growth for 1994–1996 has a positive effect on the level of innovative sales. This estimate of the feedback of past performance to innovation output for the generalised Tobit model is of the same order of magnitude as the corresponding result of Brouwer et al. (1997) [16)]. Recalling that the corresponding estimate for the innovation-intensity equation was significantly negative, this leaves us with contradicting results for the feedback effects from general company performance to the innovation process. This contradictory result may be caused by the joint endogeneity of the growth rate of total sales, the quantitative innovation-output indicator and the inputs into innovation. In this subsection we relax the implicit exogeneity assumptions underlying the single-equations estimates by assuming that the inputs into innovation, the extent of innovation success and the overall company performance are jointly determined.

In setting up our simultaneous model we maintain as a basic assumption that the impact of innovation on company performance is channelled through the sales of innovative products only. According to this reasoning the company-specific innovation characteristics that are assumed to determine the inputs into innovation and the results of the throughput process should not appear in the equations for a company’s total sales and employment growth. Therefore, the equations for the latter two variables are kept very simple. Besides the extent of innovation success as measured by the ‘log-odds’ ratio of the innovative-output indicator we include industry and size dummy variables in the equations for the growth rate of total sales and the growth rate of employment. In addition, we also include a company’s total sales growth as an explanatory variable in the equation for its employment growth. Therefore, the simultaneous system consists of the following four equations:

$$ININT = F_1(LOP, \Delta logS, X_1, IND) \quad (1A)$$

$$LOP = F_2(ININT, \Delta logS, X_2, IND) \quad (1B)$$

$$\Delta logS = F_3(LOP, SIZE, IND, IND*SIZE) \quad (1C)$$

$$\Delta logE = F_4(LOP, \Delta logS, SIZE, IND, IND*SIZE) \quad (1D)$$

where *SIZE* and *IND* denote size and industry dummy variables [17)] and X_1 and X_2 are vectors of exogenous or predetermined explanatory variables. As a base-line model we use the same exogenous and predetermined variables as in the single-equation approach. Thus, X_1 is the set of all exogenous and predetermined variables except the variables that refer to the objectives underlying the implementation of product or process innovation (D_{dpull} – D_{push3}), and X_2 includes all exogenous or predetermined variables with the exception of the financial variables (CF_{1994}, D_{subs}). Then, after adding disturbance terms to the equations, the system (1A) – (1D) can be estimated with the method of Full Information Maximum Likelihood (FIML) [18)].

The FIML estimates are presented in Table 3. We first comment on the estimation results for the base-line model. The estimates for *ININT* and *LOP* can be compared with the single-equations estimates of Table 2. A striking difference for the *ININT*-equation is that many of the variables that were significantly different from zero in the single-equation approach are no longer so in the base-line simultaneous model. Notable examples are the coefficients of the technological opportunities (*SCIENCE* and *OTHER*), the age of the companies (LA_{1994}) and the variables that refer to performing R&D on a permanent basis ($D_{R\&D}$) and to innovation in co-operation ($D_{co\text{-}op}$). Furthermore, we see a substantially smaller size effect on the inputs into innovation level (measured by the coefficient of LS_{1994}) in the simultaneous-equation approach than found in Table 2. The estimates that are comparable in both methods concern the financial variables. Notice further, that the coefficient of a company’s total sales growth turns out to be significantly positive in the equation for *ININT* but becomes not significantly different from zero in the equation for *LOP*. Thus, the feedback loop that ties together the innovation process and overall economic performance changes from the output stage to the input stage of the innovation process when one takes into account the endogeneity of the inputs into innovation, the innovation output and overall economic performance.

Next, we look at the estimates of the base-line model for the quantitative innovative-output indicator and compare these results with the estimates for the OLS-part of the generalised Tobit model presented in Table 2. Contrary to the equation for *ININT*, many of the single-equation estimates and also the ranking of the importance of the various qualitative innovation characteristics for the innovation throughput is preserved in the simultaneous approach. The differences between the two sets of estimates concern the coefficient of (the endogenous) innovation intensity which turns out to be insignificant in the simultaneous model and the feedback effect measured by the coefficient of a company’s total sales growth, already discussed above. Notice further that according to the sign of the estimate of LS_{1994} the share of innovative sales decreases with size. Thus, it is found that large companies do not have a higher share of innovative products in total sales than small ones. Combining the estimation results for the innovation output equation and the equation for total sales growth also enables us to quantify the effects on overall company performance of some important innovation characteristics. For instance, using the estimates for $D_{R\&D}$ and $D_{co\text{-}op}$, it can be found that the effect on a company's total sales growth of performing R&D on a permanent basis amounts to approximately 2% and that innovating in co-operation with other companies or research institutes (ceteris paribus) leads to about 0.5% higher sales growth.

The two other sets of estimates presented in Table 3 concern a modification of the base-line model and the application of this modified model (labelled model A in the table) to the subset of companies that also includes the companies with a score for *P* equal to 0 or 1 (labelled model B in the table). In model A we removed the statistically insignificant variables LA_{1994}, $D_{R\&D}$ and $D_{co\text{-}op}$ from the equation for *ININT*. Another reason for eliminating the first two variables is our belief that they measure similar things to *SCIENCE* and *OTHER*. We prefer to use the latter two 'technological opportunity variables' because these are continuous and thus are expected to have more information content than the dummy variables [19]. It can be seen in Table 3 that the coefficients of the equation for *ININT* of model A are estimated more precisely, whereas – taken on the whole – the estimates for the innovative-output equation and also for the contribution of the innovation throughput to a company's total sales growth do not diverge much from the base-line estimates. Another notable difference between the two models is that the estimate of *SCIENCE* in the equation for *ININT* and the estimate for $D_{co\text{-}op}$ in the equation for *LOP* become statistically significant, whereas the effect of the technological opportunity *OTHER* on the inputs into innovation remains insignificant.

Lastly, we comment on the results of the modified simultaneous model after including the companies which reported zero innovative sales (N = 280) and the companies with sales in 1996 consisting entirely of new or improved products (N = 35) [20]. The differences are as to be expected. Including these companies leads to changes for some of the estimates of the equation for *LOP* in particular, but

Table 3
FIML estimates of the simultaneous model

	Base-line		A [1]		B [2]	
	Est.	T [3]	Est	T [3]	Est.	T [3]
Number of companies	1,662		1,662		1,977	
ININT						
Constant term	8.130	2.5 ***	5.579	1.9 *	5.191	1.7 *
CF_{1994}	0.040	2.9 ***	0.042	3.6 ***	0.023	2.6 ***
LS_{1994}	–0.353	–1.0	–0.454	–2.6 ***	–0.507	–2.9 ***
ΔlogS	0.385	2.7 ***	0.480	3.0 ***	0.674	2.5 ***
LOP	2.504	0.8	0.545	0.5	0.117	0.2
SCIENCE	0.137	0.3	0.346	1.8 *	0.428	2.7 ***
OTHER	–0.564	–0.7	–0.226	–0.9	0.072	0.3
LA_{1994}	0.087	0.2				
D_{subs}	0.906	2.1 **	0.912	2.3 **	0.917	2.5 ***
$D_{R\&D}$	–1.079	–0.7				
$D_{co\text{-}op}$	0.033	0.1				
Industry dummy variables included	yes		yes		yes	
LOP						
Constant term	–0.736	–1.1	–0.488	–0.7	–3,535	–2.9 ***
ININT	0.028	0.5	0.032	1.6 *	0.256	1.6 *
LS_{1994}	–0.082	–1.7 *	–0.093	–1.9 *	0.009	0.1
ΔlogS	–0.006	–0.2	–0.023	–0.6	–0.183	–1.2
SCIENCE	0.117	2.4 ***	0.123	2.4 ***	0.031	0.3
OTHER	0.207	4.3 ***	0.218	4.2 ***	0.338	3.4 ***
LA_{1994}	–0.125	–2.3 ***	–0.152	–3.1 ***	–0.163	–2.4 ***
$D_{R\&D}$	0.464	4.7 ***	0.446	4.4 ***	0.950	4.5 ***
$D_{co\text{-}op}$	0.143	1.5	0.192	2.3 ***	0.348	2.3 ***
D_{dpull}	0.247	1.2	0.305	1.4	1.522	4.5 ***
D_{push1}	0.117	1.4	0.155	1.8 *	–0.109	–0.9
D_{push2}	0.006	0.1	0.018	0.2	–0.128	–1.0
D_{push3}	0.033	0.4	–0.008	–0.1	0.048	0.4
Industry dummy variables included	yes		yes		yes	
ΔlogS						
Constant term	14.013	5.7 ***	13.876	6.0 ***	9.329	6.1 ***
LOP	4.346	3.9 ***	4.401	4.0 ***	1.278	2.8 ***
Industry and size dummy variables included	yes		yes		yes	
ΔlogE						
Constant term	0.735	0.5	0.715	0.5	–0.427	–0.4
LOP	0.377	0.6	0.430	0.7	–0.168	–0.6
ΔlogS	0.029	0.8	0.041	1.1	0.039	0.9
Industry and size dummy variables included	yes		yes		yes	
Log Likelihood	–21,346		–21,347		–26,519	

1) Base-line model, but the variables LA1994, DR&D and Dco-op excluded from the equation for ININT.
2) Model A, including companies with a score for P equal to 0 or 1.
3) *, ** and *** denote significance at the level of 10, 5 and 1 % respectively.

the estimates of the equation for *ININT* are not affected too much (an exception is the stronger significance of the coefficient of *SCIENCE*) [21]. It can be noted that the constant term of the equation for *LOP* decreases substantially [22] and that the effects of performing R&D on a permanent basis and the importance of the 'demand-pull' objective to implement product or process innovation are substantially larger. In addition, the contribution of the innovation throughput to a company's total sales growth is substantially smaller (although still significantly positive) in the final estimates than in the preceding results.

A recurrent feature in the estimates of the simultaneous model is that a company's overall economic performance has a larger influence on its financial sources devoted to innovation than on its innovation output. On the other hand the technological environment of a company seems to contribute more directly to its innovation output. Experiences built up in interactions with customers, suppliers and competitors combined with permanent own R&D efforts appear to be most favourable for the results of the innovation process and thus for the overall economic performance. Their importance can be determined by performing some simulations with the help of the estimates of the model. These simulations were carried out using the final estimates of Table 3 and after using averages of the continuous variables for the definition of the reference company [23]. Then, it can be shown that the combined effect of R&D and 'producer-client' interactions raises the share of innovative products in total sales by 14 percentage points. In turn, this raises total sales growth by 1.9 % (1.2 % for performing R&D on a permanent basis, 0.45 % for innovating in partnership and 0.25 % for the use of information sources provided by other companies).

6. Summary and conclusions

The econometric research reported in this article was inspired by the observable divergence between formal innovation models and the econometric practice. Formal models of innovation stress the importance of links between the input, the throughput and the output stage of the innovation process and the links between the innovation process and economic activity. However, notwithstanding newly available company level innovation data, the partial or reduced-form approach to empirical modelling has dominated the mainstream of micro econometric research of these data. By way of example we refer to the availability of new innovation output indicators. This shifted much of the attention away from the single-equation estimation of the determinants of traditional input measures such as R&D intensities, to the single-equation estimation of the determinants of innovation output, thereby neglecting the possibility of a joint dependence of both measures of innovativeness on company specific innovation characteristics. For instance, the technological opportunities open to the company may affect its level of innovation expenditures (reflecting, for example, the absorptive capacity argument to build up and maintain its own R&D stock of knowledge) but the same opportunities may also affect its level of innovation output directly. Therefore, the role of technological opportunities should preferably be analysed using a systems approach in which the impact of different opportunities on the inputs into innovation and the innovation output are estimated simultaneously. Similar arguments hold for the link between the innovation process and overall economic performance. If one is willing to accept the notion that the impact of innovation on economic activity is mainly channelled through innovative sales, then the level of innovative sales and the overall economic performance measured by the growth rate of total sales are jointly dependent and this should be taken into account in the estimation procedure.

The interdependencies and complementarities referred to above have been investigated in this paper. With the innovation model of Kline and Rosenberg (1986) serving as a frame of reference, an attempt was made to estimate simultaneously the determinants of differences in innovation intensities, the factors determining the share of innovation sales in total sales and the contribution of innovation output to sales and employment growth at the company level. We compared the results of the single-equation estimation of the determinants of the inputs into innovation and share of innovative sales with the results of a simultaneous model, and showed that the results of the two approaches diverge. A notable difference is that the feedback loop that ties together the innovation process and overall economic performance changes from the output stage to the input stage of the innovation process system when one takes into account the endogeneity of the inputs into innovation, the innovation output and overall economic performance.

Other empirical results from the simultaneous model are:

- Financial variables such as past internal cash flows and innovation subsidies contribute significantly to the level of innovation intensity. Furthermore, we observed a significant 'cost-push' effect of the technological opportunity 'science' on the level of innovation intensity but not so for the technological opportunity offered by 'other companies'.
- In addition to the level of the innovation intensity, a separate role was observed for R&D and for innovation in partnership. Companies which perform R&D on a permanent basis or innovate in partnership have a significantly higher level of innovation output than other companies.
- The use of technological opportunities offered by customers, suppliers and competitors has a larger effect on the level of innovation output than the use of these opportunities offered by 'science'.
- A significantly positive effect of the level of innovation output on sales growth was observed, but this is not the case for employment growth.
- Finally, Schmookler's demand-pull hypothesis is confirmed in the significantly positive feedback effect from company level sales growth to the level of innovation intensity.

For further information or comments on this contribution, please contact Luuk Klomp: lklp@cbs.nl, or George van Leeuwen: glwn@cbs.nl.

Notes

1) This paper uses the data of the second Dutch Community Innovation Survey. The first results of this survey were presented in Statistics Netherlands (1998). Klomp and van Leeuwen (1999) also presents a comparison with the macro results of other countries. The econometric analysis presented in this paper is an extended and substantially revised version of section 6.5 of the first mentioned publication

2) Companies were rejected if the employment figure for 1996 reported in CIS-2 was below 66 percent or above 150 percent of the corresponding employment figure reported in the Production Survey.

3) Companies were rejected if the ratio of total innovation expenditure to total sales was higher than 50 percent.

4) Of course we are aware of the possibility that a company's innovative activity may change the technological environment, but the analysis of these interactions is beyond the scope of this paper.

5) The availability of time series data for most Western countries seem to be responsible for the frequent use of R&D as a measure for the innovation process. R&D is often the only indicator that allows for international comparison over a time period.

6) An alternative is to use the value added of companies as the scaling measure, but this would require the imputation of depreciation cost for some of the innovation expenditure components, notably investment in R&D, fixed and intangible assets. Furthermore, using total sales as a scaling factor seems to be more in line with the throughput measure available and value added is not available for every company.

7) The 'log-odds' ratio can be expressed as with *P* the ratio of innovative sales to total sales.

[8] The use of publicly available information sources such as journals, scientific literature, fairs and exhibitions is also included in the principal components analysis. These information sources obtain the highest scores for the factor loadings of 'other companies'. Details on the principal components analysis can be found in appendix I of Klomp and van Leeuwen (1999).

[9] For the construction of industry dummy variables, the companies were classified into nine groups of the Pavitt classification and for the 'size' dummy variables, we classified the companies into three groups (small, medium sized and large companies). The 'supplier dependent' manufacturing industry and small companies are the reference categories.

[10] For 917 companies data on profitability and for 17 companies data on the age in January 1994 were missing and 2 companies were rejected because of an exceptional score of the profitability indicator.

[11] See section 4.3 of Klomp and van Leeuwen (1999) for a comparison of the remaining companies and the whole sample of innovating companies.

[12] We also estimated an innovation intensity equation with the R&D intensity as the dependent variable. The results show a small and insignificant estimate for the technological opportunity 'other companies' and a contribution of 'science' which is of the same order of magnitude, but estimated more precisely.

[13] See section 5.2 of Klomp and van Leeuwen for further details.

[14] For similar reasons we do not include the financial variables in the two-limit Tobit model.

[15] Approximately 90 percent of the companies with data on innovative sales available belong to the manufacturing industry.

[16] Using the first wave of CIS, Brouwer et al. (1997) estimated a threshold model in order to explain the sales of innovative products under the assumption of fixed cost of introduction.

[17] We recall that small companies in the 'supplier dependent' manufacturing industry are the reference group.

[18] We used the FIML estimation procedure implemented in TSP 4.3.

[19] Estimating the model with *SCIENCE* and *OTHER* excluded but LA_{1994}, $D_{R\&D}$ and $D_{co\text{-}op}$ included, yielded results that were not diverging much from the results of the baseline model.

[20] We recall that as a pragmatic solution we imputed a value for *P* of 0.001 respectively 0.999 in order to construct the variable *LOP* for these companies.

[21] Taking the standard errors of the estimates into account, the final estimates of the equation for *ININT* are not statistically different from the estimates for model A.

[22] This forces the constant term for P into the direction of zero.

[23] For the reference company we pinpointed in the equation for *LOP* the dummy variable for the 'demand-pull' objective at one and the other dummy variables at zero.

References

Acs, Z.J., B. Carlsson, and C. Karlsson (eds.), 1999, *Entrepreneurship, Small & Medium-Sized Enterprises and the Macroeconomy*, Cambridge University Press, Cambridge, Mass. USA.

Audretsch, D.B., and R. Thurik (eds.), 1999, *Innovation, Industry Evolution, and Employment*, Cambridge University Press, Cambridge, United Kingdom.

Brouwer, E., 1997, *Into Innovation, Determinants and Indicators*, Academic Thesis, University of Amsterdam, Elinkwijk B.V., Utrecht.

Brouwer, E. and A. Kleinknecht, 1997a, 'Determinants of Innovation. A Microeconometric Analysis of Three Alternative Innovation Output Indicators', in E. Brouwer, *Into Innovation, Determinants and Indicators*, Academic Thesis, University of Amsterdam, Elinkwijk B.V., Utrecht.

Brouwer, E., G. Ridder, K. van Montfort and A. Kleinknecht, 1997b, 'Innovation and fixed costs', in E. Brouwer, *Into Innovation, Determinants and Indicators*, Academic Thesis, University of Amsterdam, Elinkwijk B.V., Utrecht.

Cohen, W.M. and D.A. Levinthal, 1989, 'Innovation and Learning: The Two Faces of R&D', *The Economic Journal*, vol. 99, pp. 569–596.

Cosh, A., A. Hughes and E. Wood, 1998, 'Innovation in UK SME's: Causes and Consequences for Company Failure and Acquisition', in Acs, Z.J., B. Carlsson and C. Karlsson (eds.), *Entrepreneurship, Small and Medium-Sized Enterprises and the Macro-Economy*, Cambridge University Press, Cambridge, Mass., USA.

Felder, J., G. Licht, E. Nerlinger and H. Stahl, 1996, 'Factors Determining R&D and Innovation Expenditure', in Kleinknecht, A. (ed.), *Determinants of Innovation, The Message from New Indicators*, Macmillan Press Ltd., London.

Kleinknecht, Alfred (ed.), 1996, *Determinants of Innovation: The Message from New Indicators*, MacMillan Press Ltd., London.

Kleinknecht, Alfred and Donald Bain (eds.), 1993, *New Concepts in Innovation Output Measurement*, St. Martin's Press, Ipswich.

Kline, S.J. and N. Rosenberg, 1986, 'An overview of innovation', in Landau, R., and N. Rosenberg, (eds), *The Positive Sum Strategy. Harnessing Technology for Economic Growth*, National Academy Press, Washington DC.

Klomp, L. and G. van Leeuwen, 1999, 'The importance of innovation for company performance', working paper Department of Statistics on Short Term Indicators and Technology, Statistics Netherlands, Voorburg/Heerlen.

Leiponen, A., 1999, 'Competencies, Collaboration and Innovation', mimeo.

Maddala, G.S., 1994, '*Limited-Dependent and Qualitative Variables in Econometrics*, Econometric Society Monographs No. 3. Cambridge University Press, Mass., USA.

Ministry of Economic Affairs, 1999, *Room for Industrial Innovation*, (in Dutch: Ruimte voor industriële vernieuwing), The Hague, The Netherlands, June 1999.

OECD, 1998, *Technology, Productivity, and Job Creation*, Paris.

OECD, 1999, *Main Science and Technology Indicators 1999/1*, Paris.

Revermann, Crista and Elke Maria Schmidt, 1999, 'Measuring Research and Development in Service Industries of Germany', paper presented at the Working Party of National Experts on Science and Technology (NESTI), *OECD*, Paris, 1–2 July 1999.

Schmookler, J, 1966, *Invention and Economic Growth*, Harvard University Press, Cambridge, Mass., USA.

Statistics Netherlands, 1998, *The Knowledge Based Economy 1998*, (in Dutch: *Kennis en economie 1998*), Voorburg/Heerlen, The Netherlands, December 1998.

Explaining gender wage differences

Bart F.M. Bakker, Kea G. Tijdens [1] *and Jeroen W. Winkels*

1. Introduction

In industrialised countries women's wages tend to be lower than men's, although by how much differs from country to country. In 1995, the average hourly wage of women in the countries of the European Union (EU) was 78% of that of men. The ratio ranges from 70% in Greece to 90% in Sweden. The Dutch gender difference is one of the highest: on average Dutch women earn 71% of average male wages (Eurostat, 1999). The present study reveals an overall earnings ratio of about 70% in the Netherlands in 1993.
Since the early 1970s research on the gender wage difference has focused primarily on two issues. The first is the extent of the wage difference, adjusted for workers' characteristics, with economists focusing on human capital and sociologists on occupational variables. The second question deals with identifying the impact of labour force interruptions on wages. This study estimates how various factors contribute to the wage difference between men and women in the Netherlands. Three groups of variables are used, describing differences in human capital (including interruptions of the occupational career), differences in jobs and differences in the match between qualifications and jobs.
Traditionally, human capital is measured as educational level and work experience. Apart from the level, the type of education is also important for labour market opportunities. The segmentation of the labour market implies that some segments are overcrowded, and women tend to choose types of education which qualify them for these overcrowded segments only. For this reason, we introduce the type of education as a factor to explain differences in gender wage differences.
Labour market interruption is another explanation of female-male wage differentials. Approximately one third of working women interrupt their occupational careers to have and bring up children. Another reason for career interruption is unemployment. Both sorts of interruptions are supposed to influence the value of the qualifications: educational qualifications become outdated when they are not kept up through work experience.
Another possible explanation for gender wage differences is that men and women are unequally distributed over occupations. This article looks at five aspects of occupation that could explain male-female wage differences: the occupational level, the concentration of women in the occupation, the level of supervision, the number of working hours and the type of industry. The introduction of the gender composition of occupations as a factor determining wage differences needs some further explanation.
Wages in female-dominated occupations are usually low. Most explanations for this focus on skill-related arguments: women more often work in occupations in which skill requirements are lower than occupations dominated by men. Furthermore, predominantly female jobs pay lower wages to women and men because these jobs require relatively many care-giving skills, which are valued less in terms of pay. Other explanations focus on women's progress through occupations over the career. The gender wage gap will widen with age, as women tend to choose occupations that will not penalise them too heavily for spells out of the labour force. In these occupations age-earnings profiles are flat, a fact supported by Mikulic (1998), who finds that wage mobility of women is lower than that of men.
Lastly, overeducation is introduced as one of the indirect factors that may explain gender wage differences. Overeducation influences wages because for a given job the people with "too much" education for that job are supposed to be more productive and earn more. For a given educational level, overeducation reduces the worker's returns to education: the returns are lower for the overeducated than the returns for the adequately educated.
In order to estimate the impact of these indicators properly, we need to control for other relevant variables. Here we control for some important family background characteristics, like having a partner or not, having children to care for, and region and urbanisation of residence.
The analysis is based on the *Netherlands Socio-economic panel survey*. This survey contains detailed information on wages, education, occupation, human capital and family background. Labour force interruptions are incorporated in the study using the panel data.

2. Data and variables

2.1 The Socio-economic panel survey

The data used in this analysis come from the 1993 Socio-economic panel survey (SEP). This is an annual panel of approximately 5,000 households, and contains information on household members over 15 years of age. Statistics Netherlands has been carrying out this survey since 1984, in an effort to describe the main elements contributing to the prosperity of individuals and households. It covers topics such as income, transfer of income, living conditions, hours worked, domestic production and the perception of prosperity.
The initial SEP sample was drawn in two stages. First municipalities were selected and then addresses within these municipalities. A number of fieldwork procedures were designed to decrease attrition from the panel. Because of selectivity of response in the first wave and selectivity of attrition, entry and re-entry into the panel, the sample is weighted according to sex, marital status and urbanisation. In doing so, the 1993 wave becomes a representative sample of the Dutch population. For our analysis we selected respondents who were already in the sample in 1986, and from them we selected employees who were employed for at least twelve hours a week in 1993. The SEP data do not include information on people working for fewer hours. The self-employed were excluded. After this selection, and after excluding individuals with implausible values for hourly wages and missing information, the sample contained 2,760 respondents.

2.2 Variables

This study aims to explain the differences between the wages of men and women. Means and t-values of the variables are presented in Table 1. In this section we discuss the measurement of the variables and the differences between men and women.

Wages
The dependent variable is the natural logarithm of the hourly earnings in Dutch guilders before tax. Measured in Dutch guilders, men earn on average DFL 30.51 per hour, women DFL 21.25, a raw earnings ratio of 70%. The hourly wages were calculated from the labour income variables (including overtime and shift premiums) for the calendar year 1992 and the number of working hours in April 1992 and 1993. Respondents' annual income is thus converted to hourly earnings according to the average number of hours worked in one year. All variables, except age, show a linear relationship with the log hourly earnings.

Measurement and means of occupational variables
Five indicators of occupation were used in the model: occupational level, the proportion of women within the occupation, the level of

supervisory responsibilities of the individual, the number of working hours and the type of industry.

We define an occupation as a category of similar jobs, a job being a set of tasks performed or designed to be performed by one individual. In the SEP survey, respondents are asked for their job title, a description of the main tasks within that job, whether they have supervisory tasks, and for a description of the economic activity of the company or firm for which they work. Those with supervisory tasks are asked how many subordinates they have and to describe their supervisory responsibilities. The occupations are coded according to the 1992 Netherlands' Standard Classification of Occupations (Bakker, 1993). The NSCO'92 is meant to provide a framework for the classification of all jobs in the labour force. It allows the grouping of jobs into successively broader categories, including a ranking in levels. This ranking can be used as an indicator of vertical segregation between occupations.

Five *occupational levels* were assigned according to the required skills for the job defined as the level of training that best prepares the employee for the job's tasks and duties. The following levels were distinguished: (1) no certificate in secondary education, (2) junior general or junior vocational training, (3) senior general or senior vocational training (4) vocational college or first stage of university, and (5) second stage of university. If there is no appropriate training programme for a job, as in the case of management positions, the duration of required job experience is taken into account. The assignments were based on detailed secondary analysis of approximately 3,400 job descriptions. Information from several occupational data systems was used.

As a second indicator of vertical segregation we used *supervision*, consisting of two dummy variables, one indicating 1 to 9 subordinates, the other 10 or more. No subordinates is the reference category. Mean values of the occupational and supervisory variables are higher among men than women. For both male and female employees their respective average occupational levels are higher than their average educational levels.

To measure the *gender composition of the occupation* as an indicator for horizontal segregation, data are taken from the annual Labour Force Survey, for which the NSCO'92 is also applied. For the 121 occupational groups at a three-digit level we calculate the

Table 1
Means and t-values of the variables (Socio-economic panel survey, N=2,760)

	Men	Women	t-test
Wages			
ln hourly wage, before taxes	3.33	2.97	20.85 **
Occupation			
Occupational level (1 to 5)	3.04	2.73	7.70 **
Supervision			
No subordinates (reference)	.64	.86	–13.58 **
1-9 subordinates	.22	.09	9.43 **
10 subordinates or more	.14	.05	8.29 **
Gender concentration of occupation (1 to 5)	2.56	3.89	–38.03 **
Working hours (1 to 3)	2.93	2.33	23.85 **
Type of industry: public sector	.28	.53	–13.10 **
Human capital			
Educational level (1 to 5)	2.39	2.45	–1.20
Type of education			
Language and culture	.01	.01	–1.12
Technical	.26	.01	22.74 **
Home economics and services	.02	.16	–11.42 **
Agriculture	.03	.00	4.94 **
Teacher training programmes	.04	.06	–3.12 *
Social sciences	.02	.04	–2.63 *
General (reference)	.37	.40	–1.33
Law, public administration, law enforcement and security	.04	.02	3.08 *
Economics, clerical and commercial	.13	.15	–1.30
Transport, communications and traffic	.02	.00	4.96 **
Medical and paramedical	.01	.09	–7.97 **
Mathematics and natural science	.01	.01	–.47
Age	39.11	35.50	9.25 **
Firm-specific work experience	11.59	6.84	15.26 **
Interruption because of unemployment	1.12	1.18	–3.73 **
Interruption because of withdrawal	1.01	1.21	–15.53 **
Matching			
Skills mismatch			
Overeducated	.04	.06	–1.97 *
Miseducated	.06	.12	–4.81 **
Not over- or miseducated (reference)	.90	.82	5.24 **
Family background			
No partner	.16	.27	–6.27 **
Children	.85	.65	7.53 **
Urbanisation (1 to 5)	2.79	3.05	–4.88 **
Region			
East	.21	.21	–.29
South	.25	.23	1.15
North	.11	.09	1.49
West (reference)	.43	.46	–1.64

* significant, p<.05
** significant, p<.01

percentage of women in the occupation from within our sample. The distribution is bimodal, because in the majority of the occupational groups the proportion of women is either less than 20% or more than 80%. The percentages of women were classified into five categories. The boundaries for the five categories were chosen in such a way that they gave approximately a discretesized normal distribution. Completely male-dominated occupations are rated 1, while completely female-dominated occupations are rated 5. Not surprisingly, compared with men, women appear more often to be employed in female-dominated occupations.
Working hours are measured as the contractual working time, and coded as (1) 12-19 hours, (2) 20-31 hours, and (3) 32 hours or more. Hours are not coded continuously, because we do not expect a linear relationship between hours and log hourly wages. Lastly, *sector* is specified as either public or private. Compared with men, women appear to be employed in the public sector much more and to be much more likely to have a part-time job.

Measurement and means of human capital variables
Six indicators for human capital were used in the model. *Educational level* is defined as the highest certificate attained according to the five categories already used for occupational level: (1) no certificate in secondary education, (2) junior general or junior vocational training, (3) senior general or senior vocational training (4) vocational college or first stage of university, and (5) second stage of university. This type of classification is quite usual in the Netherlands and provides more information than 'years of education' as an explanatory variable. Working women have a higher educational level than working men (variable means 2.45 and 2.39 respectively). This is mainly because working women are predominantly younger and therefore better educated than the overall population. *Type of education* is measured with eleven dummy variables for the highest certificate attained. Table 1 reveals the largest gender differences in technical education and in education for home economics and services.
Firm-specific work experience is measured as experience with the current employer in years, based on employment history questions in wave 1993. Total work experience in years is not measured as such, nor could it be estimated. It is proxied by *age* in years. This is a better proxy for men than for women, because of women's career breaks. As tests of linearity show that age does not have a linear relationship with log hourly wages, we add *age squared* to the wage equations. Table 1 shows that, although the men in the panel are on average only four years older than the women (39.1 versus 35.5 years), they have nearly twice as many years of firm-specific work experience (11.6 versus 6.8 years). This may be caused by periods out of the labour force. Periods of at least a half-year between 1984 and 1993 are measured using earlier waves of the panel, but job interruptions that began before 1984 could not be investigated because of left-censoring. Two variables were created, indicating *interruption because of unemployment* and *interruption because of withdrawal* from the labour market. As expected, women interrupted their working career for longer periods than men, in particular because they withdrew temporarily from the labour force.

Measurement and means of the matching process
How individuals with a given education are allocated to jobs – the matching process – is measured by indicating the mismatch between education and occupation. Individuals can be said to be overeducated if they have more skills and abilities than their jobs utilise, and if they have received more schooling than is required for their occupation. Halaby (1994) argues that this measure of overeducation ignores the diversity of qualitatively distinct types of skills. Therefore we use two dummy variables for skills mismatch: (1) *overeducation,* if an individual has a higher educational level than required within the type of education typical for the occupation, and (2) *miseducation,* if someone has a higher educational level than required, but in an education not typical for the occupation. The matching process is less perfect for female employees. In the panel, 10% of the men and 18% of the women are overeducated or miseducated.

Measurement and means of family background variables
To measure family background, we use the presence of *children* and a *partner* (married or cohabiting). Table 1 shows that 27% of female employees are single and 65% have children, compared with 16% and 85% respectively in the male work force. Following Van de Stadt and Vliegen (1994), *urbanisation* is measured as the address density of the areas where respondents live. This density varies from (1) fewer than 500, (2) 500–1,000, (3) 1,000–1,500, (4) 1,500–2,500, (5) 2500 or more addresses. Urbanisation levels differ for men and women in our sample, as employed women are significantly more likely to live in towns. *Regional differences* are controlled for using the three dummy variables, east, north and south. West is the reference category. However, there are no significant differences in the distribution of the sexes across regions.

3. The method to decompose group differences

Several steps preceded decomposition of the gender wage difference. First, we looked for outliers in a linear regression analysis without interaction terms, with hourly wage as the dependent variable, and the variables listed in Table 1 as independent variables. We use dummy variables to estimate the effects of region, type of education, supervision and skill mismatch. Forty outliers were removed. We then proceeded with the same analysis, but without the outliers and used the 'new' regression coefficients in an optimal scaling procedure for the four dummy variables. This procedure optimises discrimination in the dependent variable for each dummy variable in a multivariate way, by controlling for the other factors listed in Table 1. The results of the optimum scaling technique are given in the second column of Table 2. These first two steps both employ pooled data on men and women.
Differences in wages between men and women can be decomposed into compositional differences (e.g. differences in the level of education) and differences in returns (e.g. differences in the effects of education on wages). Compositional differences are measured by the means of men and women (Table 1, second and third column). Differences in returns are measured by linear regression coefficients estimated for men and women separately (Table 2, third and fourth column). Following the method of decomposition developed by Oaxaca (1973), the total wage difference between men and women can be decomposed by:

$$\ln w_m - \ln w_f = B_m' (X_m - X_f) + X_f' (B_m - B_f)$$
or
$$\ln w_m - \ln w_f = B_f' (X_m - X_f) + X_m' (B_m - B_f)$$

where ln w = the natural logarithm of the hourly wage, B = the vectors of the regression coefficients including the intercept, X = the vectors of the averages, and the subscripts m = male and f = female. The first component [$B_m' (X_m - X_f)$ or $B_f' (X_m - X_f)$] indicates the wage differences caused by differences in the structure of the labour force. The second component [$X_f' (B_m - B_f)$ or $X_m' (B_m - B_f)$] indicates the gender differences caused by the different effects for men and women of the variables on the hourly wages. Thus, all interaction terms with gender are specified. The sums of the two components are alike, but the components vary, because they have been weighted differently.
Jones and Kelley (1984) established that even minor changes in the model specification, like the location of the zero point of each continuous exogenous variable, the choice of the omitted category of a dummy variable, and the assignment of a 0-1 score to a dichotomous variable, lead to a substantially different estimation of the wage difference due to compositional differences and differences in returns. However, their joint contribution per variable is entirely correct. Because we are not aware of work in which a solution is given for this problem, we apply Oaxaca's (1973) decomposition rules, but we must be careful in our interpretation of the results. The results are shown in Table 2.

Table 2
Scale values, wage equations and decomposition of the male-female wage difference (Socio-economic panel survey, N=2,760)

	Scale values	Wage equations		Contribution to wage difference
		men	wome n	
Occupation				
Occupational level		.167 **	.116 **	.190 **
Supervision		1.137 **	.228	.035 **
No subordinates	0			
1–9 subordinates	.067			
10 subordinates or more	.134			
Gender concentration of occupation		.003	–.035	.141 **
Working hours		–.022	–.026	–.003
Type of industry: public sector		–.151 **	.020	–.052 **
Human capital				
Educational level		.006	.032 *	–.062 **
Type of education		.882 **	1.094 **	–.009 **
Language and culture	–.099			
Technical	–.065			
Home-economics and services	–.050			
Agriculture	–.038			
Teacher training programmes	–.037			
Social sciences	–.036			
General	0			
Law, public administration, law enforcement and security	.016			
Economics, clerical and commercial	.024			
Transport, communications and traffic	.051			
Medical and paramedical	.064			
Mathematics and natural science	.140			
Age		.082 **	.083 **	.270 *
Age squared		–.000 ** 9	–.001 ** 0	–.073 *
Firm-specific work experience		–.000	.008 **	–.056 **
Interruption because of unemployment		–.099 **	–.029	–.075 **
Interruption because of withdrawal		–.127	–.184 **	.096 **
Matching				
Skills mismatch		1.295 *	.192	.003 **
Overeducated	–.013			
Miseducated	.063			
Nnot over- or miseducated	0			
Family background				
No partner		–.092 **	.006	–.015 **
Children		.012	–.069 **	.055 **
Urbanisation		.010	.005	.011 **
Region		.949 *	1.235 *	.005 **
East	–.045			
South	–.035			
North	–.016			
West				
Constant		1.275 **	1.373	–.098
Adjusted R^2		.548	.426	
male-female wage difference				.363

* significant, p<.05
** significant, p<.01

4. Findings

The overall wage difference is .363, which means that women's mean log hourly wage is .363 lower than men's. In general, the gender wage difference is influenced by differences in the composition of the respective work forces as shown by the means in Table 1, by differences in the returns for the variables as shown by the respective wage equations in Table 2, or by combined effects. The last row in table 2 shows that some variables contribute positively to the wage difference, others negatively. A variable with a positive sign would narrow the difference if the two work forces were equal with respect to this variable and if there were similar returns, all other things being equal.

4.1 The impact of occupation

As far as occupational characteristics are concerned, the means in Table 1 show that *occupational level* is significantly higher for male than for female employees. Moreover, the regression results in Table 2 show that the returns on the *occupational level* are positive and significant for both sexes, but that they are higher for men. Therefore, *occupation level* makes a very substantial contribution to the gender wage difference (.190 of .363), due to compositional effects and differences in returns.

Men are more likely to have supervisory jobs, although the vast majority of men and women do not have any supervisory tasks. The returns on the *supervision* variable to men's wages are positive, very high and significant. It does not pay for a woman to be a supervisor. Overall, the *supervision* variable makes a contribution to the gender wage difference of .035, which is significant but not particularly large.

The *gender composition of the occupation* makes no significant contribution to the explanation of either men's or women's earnings. Although the returns for being in a female-dominated job on women's wages are not statistically significant, the coefficient is not really low. This leaves room for the possibility that women profit from being in male dominated occupations. Because women are much more often employed in female dominated occupations, the *gender* *composition of the occupation* contributes substantially to the explanation of the gender wage difference (.141 of .363). Thus the gender wage difference is predominantly caused by wage difference between, and not within male or female dominated occupations.

Men are far more likely than women to have a job of 32 hours or more a week. Neither men's nor women's hourly wages depend on their respective *working hours*. This may be surprising since Rubery (1997) argues that much of the low pay found in the UK is concentrated among part-time workers. This is not the case in the Netherlands, partly because since the mid-eighties government, unions and employers have aimed for equal treatment of part-time and full-time workers. The decomposition in Table 2 shows that working hours do not contribute significantly to the gender wage difference.
Lastly, we study the impact of *sector*. For men returns are less in the public than in the private sector. For women, they are equal in both sectors. As relatively more women than men are employed in the public sector, the overall effect is –.052. This suggests that the wage difference would have been larger if composition and returns had been equal with respect to sector.

4.2 The impact of human capital

We now turn to the human capital characteristics. Table 1 showed that there are no significant sex differences in *educational level*. The regression coefficients show that the returns on education are positive and significant for female and low and insignificant for male employees. Altogether, educational levels contribute -.062 to the male-female wage difference. And although in contrast to this, there are large gender differences in the *type of education*, this variable does not affect men's or women's wages significantly. The overall effect of type of education on the wage difference is slightly negative (-.009 of the .363 wage difference). Thus, if the female and the male work force had been similar with respect to education and type of education and the returns on education had been similar, the gender wage difference would have been substantially larger, all other things being equal.
Age differs significantly between the sexes: the men are older. Although age has comparable effects on the wages of both sexes, it has a large effect on the gender wage difference (.270) because of the large compositional differences. *Age squared* negatively influences the wages of both sexes in a comparable way. It contributes -.073 to the wage difference.
Firm-specific work experience does not influence men's wages, and has only a small but significant and positive effect on women's wages. Men have more years of experience than women. Overall, firm-specific experience contributes -.056 to the wage difference.
Lastly, women interrupt their employment careers more often, either through *unemployment* or *intermittence*. Interruptions have a negative effect on the wages of both sexes, although only men's wages are negatively affected by unemployment, and only women's wages are negatively affected by intermittence. Therefore, the overall -.075 contribution of unemployment to the wage difference almost counterbalances the .096 contribution of intermittence.

4.3 The impact of the matching process

As far as the matching variables are concerned, the large majority of employees are neither overeducated nor miseducated. However, women are more often both overeducated and miseducated than men. Mismatch, whether *overeducation* or *miseducation*, does not contribute to the log hourly wages of women, but it does contribute slightly positively to the log hourly wages of men. This means that only men receive sufficient benefits of having a job for which the required educational level is higher, but the required educational sector is different from their own educational background. This surprising result can be explained by the well-established fact that men find wages more important when changing jobs than women. Married women with a high education in particular, who acquire part of their household income through their partners' work (in most cases also highly educated and earning a relatively high income), choose high status jobs. As income and status are key determinants in the total social rewards, they substitute part of the income rewards by social status rewards (Frank, 1985). As income rewards are more important for men, they are more likely to invest in a more profitable job in a sector outside that of their educational background, while women are more likely to be found in (or forced into) jobs in a different sector that pays less. The overall effect of skill mismatch is slightly advantageous to men (.003).

4.4 The impact of family background

The wage effects of family background differ between the sexes. Being *single* affects men negatively, while having *children* has a similar effect on women. Female employees are less likely to have a partner or children than men. This results in an overall contribution to the gender wage difference for being a single of -.015 and for having children of .055. Thus, if the female and the male work force were similar with respect to having children and if there were no wage penalty for women with children, the gender wage difference would be smaller, ceteris paribus. In contrast, the wage difference would have been larger if things had been equal with respect to having a partner.
Urbanisation does not effect wages within sex, but because its means differ substantially between the sexes, it does influence the gender wage difference. The urbanisation variable can explain .011 of the .363 wage difference. *Region* does not differ between the sexes, but living in the west of the country shows significant positive returns for both. These returns are higher for men than for women. The capital and the large cities are located in the western part of the Netherlands. Region only very slightly contributes to the wage difference (.005 of .363).

5. Conclusion and discussion

This study aimed to explain the gender wage difference by estimating the impact of human capital, occupation and various factors. Three groups of variables are used to describe differences in human capital (including interruptions of the occupational career), differences in jobs and differences in the qualifications-jobs match. The analysis is based on data of 2,760 respondents in the Socio-economic panel survey, which contains detailed information on wages, occupation, human capital variables and family background. The total gender wage difference in log hourly wages is .363. The largest contribution to the wage difference is made by occupation, and to a lesser extent by human capital. Job matching processes hardly contribute to the explanation, and family background is particularly important for women.
Our conclusions are that the largest part of the gender wage difference is explained by the five variables measuring occupation (.311 of the .363 difference). Overall, a major contribution is made by occupational level and gender composition of the occupation, and a minor contribution by supervision. Wages in female dominated occupations are lower than in male dominated occupations, but there are hardly any gender pay differences within these occupations. If the female and the male work force were similar with respect to occupational level, supervisory tasks and 'femaleness' of the occupation, and there were no differences in returns, the gender wage difference would totally disappear. On the other hand, it would be larger if things were equal with respect to sector.
Altogether, human capital variables contribute substantially to the gender wage difference. The major contribution comes from age and age squared (together explaining .197 of the .363 wage difference). The returns on age hardly differ between the sexes,

suggesting that the wage difference can be attributed to compositional effects. If female employees had the same age profile as their male counterparts, the gender wage difference would be more than halved, all other things being equal.

Some other human capital variables are advantageous for women. If the female and the male work forces were similar with respect to educational level, type of education, firm-specific work experience and there were no wage penalties, the gender wage difference would be larger, ceteris paribus. Interruptions of the occupational careers lead to lower wages for men and women when they re-enter the labour market. However, they contribute only slightly to the gender wage difference. The effects of unemployment are negative for men and the effects of time out are disadvantageous for women's wages. These effects counterbalance each other.

The gender wage difference has no basis in any sort of mismatch between education and occupation. This leads to the conclusion that women's occupational choices are made given adequate access to different types of education.

The study shows that the method to decompose group differences in compositional effects and returns as developed by Oaxaca (1973) can be very useful for research on gender differences.

For further information or comments on this contribution, please contact Bart Bakker: bbkr@cbs.nl, Kea Tijdens: kea@butler.fee.uva.nl, or Jeroen Winkels: jwns@cbs.nl.

Note

1) Kea Tijdens is associate professor and research coordinator at the Amsterdam Institute of Advanced Labour Studies at the University of Amsterdam.

The authors would also like to thank Jacques Siegers of Utrecht University and Louise Grogan and other colleagues at the University of Amsterdam for helpful comments.

References

Bakker, Bart F.M., 1993, 'The development of the Standard Classification of Occupations 1992', In: *Netherlands Official Statistics,* vol. 8, pp. 5–22.

Eurostat, 1999, *Living conditions in Europe. Statistical pocketbook 1999 edition* (Luxembourg: Office for Official Publications of the European Communities)

Frank, R.H., 1985, *Choosing the right pond, human behavior and the quest for status* (New York: Oxford University Press)

Halaby, C.N., 1994, Overeducation and skill mismatch, In: *Sociology of education,* vol. 67, pp. 47–59

Jones, F.L. & J. Kelley, 1984, Decomposing differences between groups. A cautionary note on measuring discrimination, In: *Sociological Methods and Research,* vol. 12, pp. 323–343

Mikulic, B., 1998, Loonmobiliteit in de eerste helft van de jaren negentig (Wage mobility in the first half of the nineties), In: *Sociaal-economische dynamiek 1998* (Voorburg/Heerlen: CBS)

Oaxaca, R., 1973, Male-female wage difference in urban labor markets, In: *International Economic Review*, vol. 14, pp. 693–709

Rubery, J., 1997, Wages and the labour Market, In: *British Journal of Industrial Relations*, vol. 35, pp. 337–366

Van de Stadt, H. & M. Vliegen, 1994, A new measure of urbanization, the address density of the surrounding area, In: G. Beets *et al.* (eds.) *Population and family in the low countries 1993, late fertility and other current issues* (Lisse: Swets and Zeitlinger), pp. 151–65