East of England Forecasting Model Technical report: Model description and data sources

cambridge econometrics
clarity from complexity

Technical Report

September 2016

Cambridge Econometrics
Cambridge, UK

sp@camecon.com
www.camecon.com

Cambridge Econometrics' mission is to provide rigorous, accessible and relevant independent economic analysis to support strategic planners and policy-makers in business and government, doing work that we are interested in and can be proud of.

Cambridge Econometrics Limited is owned by a charitable body,
the Cambridge Trust for New Thinking in Economics.
www.neweconomicthinking.org

Authorisation and Version History

Version	Date	Authorised for release by	Description
1.0	15/12/2016	Mike May-Gillings (Associate Director, CE)	Technical report

Contents

Tables

Figures

1 Introduction

The East of England Forecasting Model (EEFM) was developed to project economic, demographic and housing trends in a consistent fashion and in a way that would help inform spatial economic planning in the East of England. The Model is programmed in Excel spreadsheets, allowing users to produce scenarios under which the impacts of a given scenario can be monitored.

This report provides technical information on the EEFM's coverage, methodology and data sources. (The latest forecast results are presented separately, on the Cambridgeshire Insight website.)

The Model's outputs are just one piece of evidence to assist in making strategic decisions. As in all models, forecasts are subject to margins of error which increase at more detailed geographical levels. In addition, the EEFM relies heavily on published data, with BRES/ABI employment data in particular containing multiple errors at local sector level (though the Model does attempt to correct for these.)

The EEFM is currently maintained and developed by Cambridge Econometrics (CE). CE has a long track record in the development of economic models for strategic planning and policy analysis, at global, national and sub-national level.

The outputs and associated documentation of the EEFM are available on the Cambridgeshire Insight website.

The purpose of this document is to provide a description of the Model's methodology and the data sources used, and act as a companion reference guide to the published results. It will be updated as the Model itself is developed, improved and updated. The report is structured as follows:

- Chapter 2: Description of the Model – This chapter summarises the EEFM coverage with respect to geography, time periods and linkages with other models produced by Cambridge Econometrics.
- Chapter 3: Model Overview – This chapter summarises the structure of the EEFM, and the linkages and relationships between variables.
- Chapter 4: Data Used – This chapter lists the variables in the Model, and indicates the latest data used.
- Chapter 5: Outliers and Data Validity – This chapter summarises Cambridge Econometrics' approach to anomalous data (so-called "outliers") and the methods used to check that the EEFM is internally consistent.

This report does not provide EEFM forecast results. These can be found on the Cambridgeshire Insight website www.cambridgeshireinsight.org.uk/EEFM. The detailed forecasts are set out there in Excel spreadsheets.

2 Description of the model

This chapter provides an overview of the East of England Forecasting Model (EEFM) and summarises its coverage and links to other Cambridge Econometrics models and assumptions. It also contains a list of the variables and geographies used. The forecasting methods and data sources are described in subsequent chapters.

2.1 Structure of the EEFM

The East of England Forecasting Model is a spreadsheet-based model originally designed to help inform and monitor the development and review of the East of England Regional Economic Strategy and Regional Spatial Strategy. It covers a wide range of variables, and is designed to be flexible so that alternative scenarios can be run and the impacts of different assumptions can be measured.

Key features of the Model are:

- A full database including 151 separate variables for each of the East of England's 48 pre-April 2009 local authorities, as well as for historic counties, strategic authorities, selected other local authority groupings, the East as a whole, 10 local authorities in the East Midlands and the region as a whole, 21 local authorities in the South East and the region itself, and the UK;
- Functionality to allow users to develop their own scenarios;
- A comprehensive set of tables allowing users to select and assemble data on the variables, localities, scenarios and results they want; and
- A spreadsheet system containing:
  - Linked worksheets, to facilitate faster updating;
  - Worksheets structured to generate forecasts and scenarios;
  - Worksheets designed to produce tables.

The overall Model structure captures the interdependence of the economy, demographic change and housing at a local level, as well as reflecting the impact of broader economic trends on the East of England. The employment forecasts take account of the supply and demand for labour, the demographic forecasts reflect labour market trends as they are reflected in migration (and natural change indirectly), and the housing forecasts take account of both economic and demographic factors. This structure allows scenarios which test the impact of variables upon each other – for example, the impact of housing supply on economic variables.

2.2 Geography

The Model produces forecasts for each local authority district and unitary authority in the East of England, and selected local authorities in the East Midlands and South East region to allow for LEP aggregation. For the EEFM 2016 forecasts, that equates to 79 local authorities, including the former Mid Bedfordshire and South Bedfordshire districts which have been retained at the request of regional partners. (The new Central Bedfordshire unitary authority is one of the strategic groupings for which forecasts are also provided.)

Forecasts are also available for selected groupings of local authority districts and unitaries. These were decided in consultation with regional partners through the EEFM Model Steering Group, and also include Local Enterprise Partnerships (LEPs). For a full list of the groupings available, refer to the EEFM section of the Cambridgeshire Insight website.

In addition to these geographies, forecasts for the East of England, East Midlands and South East regions, and for the UK, are available.

2.3 Time periods

The EEFM is constructed on an annual basis. Historic data for most variables has been collected over 20 years to provide a basis for estimating the relationships between variables and for forecasting future trends. Forecasts are currently made up to 2045, reflecting the available global, national and regional forecasts. But the longer-term forecasts should be treated with some caution, as unforeseen - but inevitable - future change in the underlying drivers will affect forecast accuracy. Medium-term forecasts are actually more likely to be better approximations than shorter-term ones, as we can usually be more confident about medium-term trends than about short-term random fluctuations around the trend.

2.4 Things to remember when using the model

EEFM forecasts are based on observed past trends only

Past trends reflect past infrastructure and policy environments. Even where major new investments or policy changes are known and have actually started, they can only affect EEFM forecasts to the extent that they are reflected in the currently available data. If they have not yet impacted on the available data, they will not be reflected in the forecasts.

There are two sets of exceptional circumstances in which the currently available data need to be supplemented by other information. The first is where there are concerns about data quality. This issue is explored in Chapter 5. The second is where the Model produces unrealistic forecasts - for example, continuing an employment decline in a particular sector in a particular area until it reaches zero or even negative values. Manual adjustments to the Model are necessary in these situations, and here professional judgement inevitably comes into play. This is discussed further below.

The forecasts are unconstrained

The EEFM forecasts are unconstrained, which means that the forecast numbers do not take into account any policy or other constraints that might prevent their actual realisation on the ground. Forecasts of the demand for dwellings, for example, are the outcome of projected changes in employment, population, etc. If, in reality, planning constraints were to prevent this demand being satisfied, the associated forecast levels of GVA, employment, population, etc. would be less likely to occur.

The forecasts are subject to margins of error

As with all kinds of forecasting, there are margins of error associated with the results which tend to widen over time. Furthermore, the quality and reliability of data decreases at more detailed levels of geography. Under current data-quality

conditions, models are most helpful for identifying trends, average growth rates and broad differentials between areas, sectors, etc. Accordingly, users are encouraged to focus on the patterns over time, not figures for individual years.

Reality is more complex than any model

Several of the modelled relationships are complicated and their treatment in the EEFM is necessarily simplified, despite its large size. In particular, the demand for housing is complex and not all the factors may be fully captured. Questions such as whether migrants' apparent willingness to live at higher densities than the existing population is merely a temporary state which requires much more investigation.

Forecasting models will not all agree

The EEFM's baseline forecasts can be compared with other published forecasts, but close agreement should not be expected and sometimes there can be wide divergences. These can arise from even small differences in underlying assumptions and in the timing and definitions of the data used. But with an awareness of these factors, the EEFM forecasts provide a useful starting point for an understanding of regional and local economic trends in the East of England, particularly when the baseline is accompanied by alternative scenario forecasts with which it can be compared.

2.5 Coverage

Later chapters provide more detailed information on the data used in the EEFM and how the linkages in the Model are used for the forecasting and scenario work. The list below gives an overview of the variables covered by the Model:

Demography

- Population
  - Total
  - Working age (defined as all people aged 16-64)
  - Young (defined as all persons aged 0-15)
  - Elderly (all people aged 65+)
- Migration (Note: domestic and international migration are not differentiated in the EEFM at either the regional or the local level.
- Natural increase

Labour market

- Employee jobs by 31 sectors (workplace-based, SIC 2007 based)
  - Agriculture & fishing **(SIC 01-03)**
  - Mining & quarrying **(SIC 05-09)**
  - Food manufacturing **(SIC 10-12)**
  - General manufacturing **(SIC 13-18, 31-33)**
  - Chemicals excl. pharmaceuticals **(SIC 19-23, excluding 21)**
  - Pharmaceuticals **(SIC 21)**
  - Metals manufacturing **(SIC 24-25)**
  - Transport equipment, machinery & equipment, etc **(SIC 28-30)**
  - Electronics **(SIC 26-27)**
  - Utilities **(SIC 35-37)**
  - Waste & remediation **(SIC 38-39)**
  - Construction **(SIC 41-43)**

  - Wholesale **(SIC 45-46)**
  - Retail **(SIC 47)**
  - Land transport **(SIC 49, 52-53)**
  - Water & air transport **(SIC 50-51)**
  - Hotels & restaurants **(SIC 55-56)**
  - Publishing & broadcasting **(SIC 58-60)**
  - Telecoms **(SIC 61)**
  - Computer related activities **(SIC 62-63)**
  - Finance **(SIC 64-66)**
  - Real estate **(SIC 68)**
  - Professional services excl. R&D activities **(SIC 69-75 excluding 72)**
  - Research & development **(SIC 72)**
  - Business services excl. employment activities **(SIC 77-82 excluding 78)**
  - Employment activities **(SIC 78)**
  - Public administration **(SIC 84)**
  - Education **(SIC 85)**
  - Health & care **(SIC 86-88)**
  - Arts & entertainment **(SIC 90-93)**
  - Other services **(SIC 94-99)**
- Employee jobs – full time and part time by 31 sectors (workplace-based)
- Self-employed jobs by 31 sectors (workplace-based)
- Total employment (employee jobs plus self-employed jobs) by 31 sectors (workplace-based)
- Total number of people employed in an area (consistent with 2001 and 2011 Census points)
- Total number of an area's residents who are employed (consistent with 2001 and 2011 Census points)
- Employment rate of an area's residents (aged 16-74, consistent with 2001 and 2011 Census points)
- Net commuting (number of people employed in an area, minus the number of that area's residents who are employed)
- Unemployed (claimant and ILO)

Output

- GVA by 31 sectors (£m, workplace-based, 2011 prices for the EEFM 2016 forecasts). Note that ownership of dwellings (imputed rents as defined in the Blue Book) is now included within real estate sector.
- Productivity by 31 sectors (per job, including both employee and self-employed jobs)

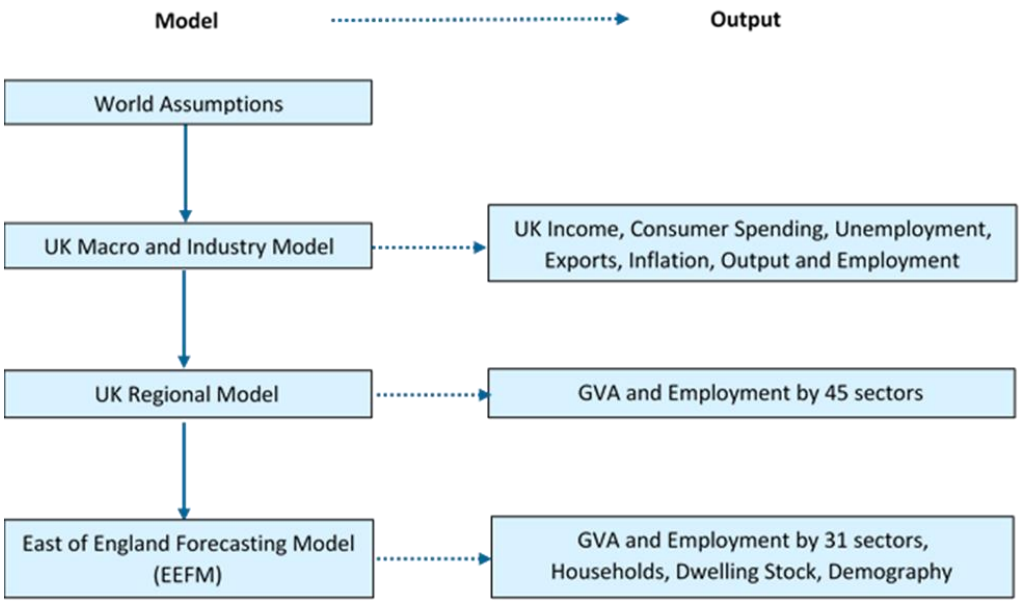
Housing

- Households
- Demand for dwellings

2.6 Links with other models

An important feature of the EEFM is its links to other Cambridge Econometrics forecasting models, ensuring that all EEFM forecasts are consistent with Cambridge Econometrics' world, UK national and UK regional forecasts. The links are summarised in Figure 2.1.

Figure 2.1 Links with Cambridge Econometrics' suite of models



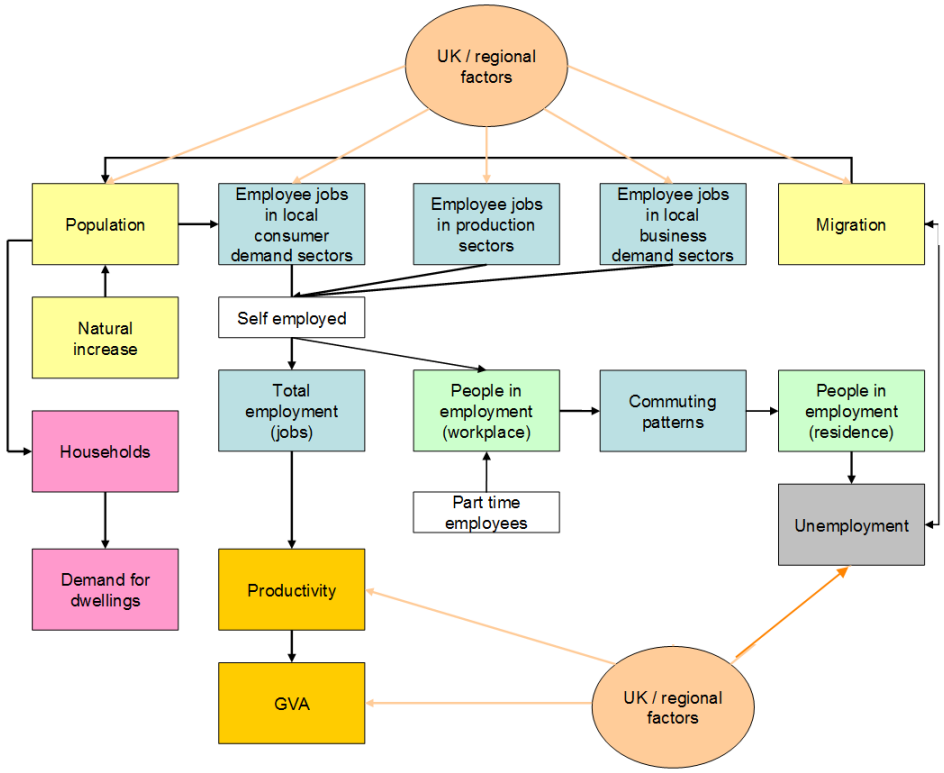
3 Model overview

The structure and data inputs of Cambridge Econometrics' UK Regional Model, which underpins the EEFM, is not set out here. But it can be obtained from Cambridge Econometrics on request.

3.1 Variables in the EEFM

The EEFM is very large, with numerous economic, demographic and housing indicators. Each of these variables is linked to others within the Model, and many key variables are also linked to others in the wider Cambridge Econometrics suite of models. The main internal relationships between variables are encapsulated in Figure 3.1, and the forecasting methodology for each element in the Model is then summarised.

Figure 3.1 Main relationships between variables in the EEFM



3.2 Economic variables

Workplace employees (jobs)

The total number of employee jobs in an area, whether full- or part-time. These can be taken by residents or by commuters from outside. Note that this is a measure of jobs, not workers, so if one person has two part-time jobs, for example, they are counted twice.

This is forecast separately in every area for each of the 31 sectors listed on pages 9 and 10. The forecasts begin with something called a "location quotient" (LQ). This is a ratio which summarises the concentration of a particular sector in a particular area, relative to the regional average. So an LQ of 0.8 (or 80%) for a given sector and area means that that sector is under-represented in the

area. An LQ of 1.25 (or 125%) means that the sector is overrepresented in the area.

The EEFM contains location quotients for every local authority in the East region including the additional local authorities in the East Midlands and South East region required to construct LEP aggregates, for each of the 31 sectors, and for every year since 1991. Forecast trends in the LQs are based on how they have changed over time. So if the LQ for a given sector in a given area has been rising in recent years, the forecasts will project this to continue, and vice versa. LQs which have been stable for a long time (including at zero) will be forecast to remain so.

Three forms of location quotient are used in the EEFM. In the first, the LQ is based on *an area's share of the region's employees in a particular sector*. This is most appropriate for sectors which are essentially independent of the local economy (e.g., manufacturing). Their activities are largely driven by regional, national or international suppliers and customers, and the goods and services they produce are typically traded over long distances. The EEFM treats the following sectors in this way:

- Agriculture
- Mining & quarrying
- Food manufacturing
- General manufacturing
- Chemicals excluding pharmaceuticals
- Pharmaceuticals
- Metals manufacturing
- Transport equipment, machinery & equipment, etc.
- Electronics
- Utilities
- Waste & remediation
- Water & air transport
- Publishing & broadcasting
- Telecoms
- Computer related activity
- Research & development
- Other services

For this group, the local employee growth forecasts in the EEFM come from the interaction of the relevant LQ forecasts with the regional sector employee forecasts from Cambridge Econometrics' UK Regional Model. To take a hypothetical example, if the UK Regional Model forecasts a 5% increase in air transport employees in the East of England, this filters down to the local area forecasts in the EEFM. If the LQ for air transport in a given area is forecast to remain stable, the employee forecasts for air transport in that area will tend to show a 5% increase. (In absolute terms, this means many new jobs in areas

with high LQs and relatively few in areas with low LQs.) If the LQ is forecast to increase (or decrease) in an area, the local employee growth forecasts for air transport will tend to be more than (or less than) 5%.

The LQ in an area can also be based on the number of employees in a given sector *per head of the local population*, relative to the regional average. This is most appropriate for sectors in which employment change is primarily (but rarely exclusively) driven by changes in the local population (e.g., health and education). In the EEFM, this group includes:

- Wholesale
- Retailing
- Hotels & restaurants
- Public administration
- Education
- Heath & care
- Arts & entertainment

For this group, the local employee growth forecasts in the EEFM come from the interaction of the relevant LQ forecasts with the demographic forecasts for the area (which are also in the EEFM) and for the region as a whole (from the Regional Model). To take the example of education, consider an area which has an education LQ of 1.3 (or 130%) - perhaps because it has a university. Suppose that that LQ has been unchanged for a long time and is forecast to stay the same. And suppose that the area's population is also forecast to remain stable. But if the region's population is forecast to increase, education employees in this area will have to increase as well to keep the equation in balance (all other things being equal). This makes sense inasmuch as the area's education institutions clearly serve a market wider than the local area.

Finally, a sector's LQ can be based on the number of its employees *relative to all jobs in the area*, relative to the regional average. This is most appropriate for sectors where changes in employment arise primarily from changes in *total* employment locally - where the latter is effectively a proxy for business activity. (As might be expected, business services sectors tend to be in this group.) In the EEFM, the following are included:

- Construction
- Land transport
- Finance
- Real estate
- Professional services
- Business services
- Employment activities

In this group, the local employee growth forecasts in the EEFM come from the interaction of the relevant LQ forecasts with the regional sector employment forecasts from the Regional Model.

Cambridgeshire County Council and Cambridge Econometrics encourage Local Authorities to view and give feedback on the forecast trends for their areas. We regard such feedback as essential to ensure the EEFM is as credible and as accurate as possible. To that end, a consultation with Local Authorities is carried out before publication of each new forecast. Chapter 5 (Table 5.1) records the instances where local intelligence on employment trends has been used to modify initial EEFM assumptions.

Full-time and part-time employment

The total number of jobs in an area, broken down into full- and part-time jobs.

East of England shares of part-time employees among all employees in the 31 EEFM sectors (which are trend forecasts linked to regional and national projections) are applied to the workplace employee estimates described above. Full-time employees are simply the total of employees minus the part-time employees for each of the 31 sectors.

Workplace self-employment (jobs)

The total number of self-employed jobs in an area.

Self-employment data for the East of England in Cambridge Econometrics' UK Regional Model comes from ONS's Quarterly Workforce jobs.

Self-employment data for local authorities is Census-based, and scaled to the East of England self-employed jobs estimates from the UK Regional Model. It is broken down by the 31 EEFM sectors. The sectors are forecast using the growth in the sectoral employees in employment data and the estimates are scaled to the UK Regional Model's estimate of self-employment by sector for the East of England.

Total workplace employment (people)

This is the total number of people in employment in an area, including both residents and commuters. A person who has more than one job is only counted once, so total workplace employed people is smaller than total workplace employment.

The employment data from the Business Register and Employment Survey (BRES) over the years 2008-14 (and the Annual Business Inquiry (ABI) for earlier years) which is used in the Model measures jobs rather than workers. Because a model aiming to simulate housing demand needs to focus on people, we have to convert the total number of jobs in an area into numbers of employed people.

The 2001 and 2011 Census' give the number of people in employment in an area. For other years, we use BRES/ABI data to estimate residents in employment using the full-time and part-time projections (see above). Individuals are assumed to hold only one full-time job each. Part-time jobs are assumed to account for 0.75 of a full-time job. A simple adjustment is made to scale the indicator so it is consistent with the Census.

This measure is not forecast directly, but is derived from the forecasts of jobs discussed above.

Total workplace employment (jobs)

The total number of employee jobs and self-employed jobs in an area. These can be taken by residents or commuters from outside. Note that this includes all full- and part-time jobs, so if someone has two part-time jobs, they are counted twice.

This is not forecast separately in the EEFM, but derived by summing the workplace-based employee jobs and self-employed jobs forecasts described above, and then adding in a constant for the Armed Forces (see below). (Note: Armed Forces data are added to the public administration & defence sector.)

Residence employment

The total number of employed people living in an area. This includes residents who commute elsewhere to work.

Residence employment is based on a commuting matrix taken from the 2011 Census. This matrix tells us, for any given area, where its residents work. Using this information, each available job (see workplace employment (people) above) is allocated to a resident of one of the authorities with which the area has commuting links, in proportion to the strength of that link. This method assumes that commuting patterns do not change over time.

Net Commuting

The number of people commuting into an area for work, less the number of residents commuting out.

Net commuting requires no specific forecasting method. It is the residual between an area's residence-based and workplace-based estimates of numbers of people in employment. (These variables are used to check the realism of the EEFM's workplace- and residence-based employment forecasts, and can occasionally lead to manual adjustments to the Model.)

Our broad assumption is that commuting flows over the forecast period are in line with past trends. Major changes in transport infrastructure, or significant new housebuilding in an area, may bring about changes in commuting patterns, but as indicated in Chapter 2, the EEFM can only take account of such changes if they are reflected in the available data.

Claimant unemployed

The total number of people in an area without a job and claiming unemployment benefits.

The number of unemployed people in area *i* is projected as:

- the previous year's value
- **plus** β_{1i} X (projected change in working-age population)
- **minus** β_{2i}X (projected change in resident employment)

The two coefficients for each local area (β_{1i} and β_{2i}) were estimated based on unemployment, working age population and resident employment data over the period 1992-2014. All coefficients are less than one, reflecting the fact that many people adding to the local working age population go into education (e.g., students) or directly into employment (e.g., by moving to the area specifically to

take up a new job), and the fact that many new job vacancies in the area will not necessarily be filled by the local unemployed (e.g. migrants, commuters).

ILO unemployment (a wider measure of unemployment that includes those who are actively seeking employment but are not claiming unemployment benefit) is also included in the Model and comes from the Annual Population Survey. This data is available from 2004 and is both back-cast and forecast, using growth rates in the claimant series.

Gross Value Added (GVA)

The total sum of income generated in an area over a specified period, usually a year. It is the sum of wages, profits and rents. An alternative and equivalent definition is the value of gross output less purchases of intermediate goods and services.

GVA forecasts are available for 31 sectors. Previously, a sector entitled 'ownership of dwellings' (imputed rents in the ONS National Accounts) was excluded from the overall business services sector and published as its own sector. In Summer 2011, the ONS changed its methodology to publish data which included imputed rents within the business services sector. To remain consistent with National data, the EEFM now includes this measure of GVA within the real estate sector.

Sub-regionally, limited sector GVA data is available at NUTS 3 level (i.e. for unitaries and shire counties) but not for local authorities. Our initial forecasts at this level are obtained by multiplying forecast regional GVA per job in a sector (from the UK Regional Model) by forecast total workplace employment (jobs) in that sector (from the EEFM) for each local authority.

These initial forecasts are then subject to two adjustments. The first is for wage differentials (from ONS's Annual Survey of Hours and Earnings), which has the effect of increasing GVA disproportionately in areas where wages are higher. The second scales local sector GVA to the most recent published NUTS 3 level GVA estimates for the relevant base year (2011).

Productivity

GVA divided by total workplace employment (jobs). It measures the average amount of income generated in each area by every person working there.

Productivity estimates do not require specific forecasting. They are simply forecast sector GVA divided by forecast total jobs (both employee and self-employed) in that sector.

Relative productivity is simply productivity in a specified area, divided by productivity in the region. A relative productivity value greater than 1.0 implies that productivity in that area (and sector) is higher than the regional average, and vice versa.

3.3 Demographic variables

Total population

The total number of people living in an area.

All population data is taken from ONS's mid-year estimates (MYE) to 2014, and the ONS 2012-based population projections are used thereafter. At local level,

total population is forecast as last year's population plus natural increase plus net migration (domestic and international).

Working age population

The total number of people in an area that are aged 16-64.

Working age population for the region is based on the 2012-based ONS subnational population projections (SNPP).

For local areas, forecast working age population is forecast total population multiplied by a ratio of working age to total population. This ratio is forecast for each year of the forecast period, and calculated as the *previous year's* ratio multiplied by the growth in the ratio regionally according to the ONS (2012-based) projections.

Young population

The total number of children in an area (defined as all people aged 0-15).

The population aged under 16 years is forecast at local authority level using an annual ratio of children to working age people. This ratio is forecast for each year of the forecast period, and calculated as the *previous year's* ratio multiplied by the growth in the ratio regionally according to the GAD (2012-based) projections. The regional forecast for this variable is simply the sum of these local area forecasts.

Elderly population

The total number of elderly people in a given area (defined as all people aged 65+).

The local elderly population forecasts are simply the residual of the total population when the young and working age populations are subtracted. The regional forecast for this variable is simply the sum of these local area forecasts.

Migration

The net flow of people moving into and out of an area, whether this be to/from other parts of the region, the UK or the world. A negative number signifies a net outflow of people from an area, a positive number a net inflow.

- Regional migration:

  This comes from Cambridge Econometrics' assumptions for net migration into the East of England.

  Total net migration into the region in any given year is based on 2012-based ONS subnational population projections.

- Local migration:

  Migration data is sourced from ONS's population mid-year estimates 'Components of Change' data. At local authority level, the number of migrants is the sum of two components: *economic migrants* and *non-economic migrants*.

The number of *economic migrants* into each area in any given year equals:

- previous year's working age population
- **multiplied by** [0.01 + 0.29 (if the area's by the coast, otherwise 0) - (0.16 X previous year's relative unemployment rate differential from the region unemployment rate)] where the unemployment rate has working age population as the denominator)
- This formula implies that the number of migrants into a district will equate to 1% of last year's working age population if the area is not by the coast and the difference between local and regional unemployment rate then was zero. To illustrate with a worked example, in an area not by the coast with 100,000 working age people and a 0.1pp positive difference in relative unemployment rate, net migration the following year will be 100,000 X [0.01 - (0.16 X 0.1)], or 100,000 X [0.01 – 0.016], or 100,000 X -0.006, or -600.

So any change in employment or population in the EEFM which affects unemployment - whether the change is externally-sourced or internally generated within the Model - will affect net migration.

Non-economic migrants are set as a constant - unique to every area - for all future years. The constant for a given local authority is selected on the basis that it both reflects the actual population trend for the area over 1991-2011 (from ONS) and implies a local employment rate trend consistent with that for the region as a whole.

3.4 Housing variables

Households

The total number of households (as defined in official statistics) in an area.

Demand for dwellings

The total number of dwellings (as defined in official statistics) in an area.

The initial household data are as presented in the official DCLG series. The initial dwellings data are the stock data presented in the official DCLG series (table 125 provides total dwelling stock, whilst table 615 provides vacant stock, the residual between these series therefore represents occupied dwelling stock).

The method for forecasting the dwelling stock and the number of households is a three stage process. To produce *household* forecasts, we divide the household population (which is calculated as the total population minus institutional population) by household size, which is in turn based on official DCLG projections.

We then forecast the number of *occupied* dwellings by applying the growth rate in households in a particular year to occupied dwellings in the previous year.

Having calculated occupied dwellings, we use the previous year's ratio of total to occupied dwellings in order to project *total* dwelling stock in a particular year. We call this *"demand for dwellings."* It is intended to proxy dwelling stock, but it is not a conventional stock or supply figure. Rather it tries to estimate what stock might be needed to maintain current occupation ratios in the context of a higher population.

3.5 Carbon emissions

Industry, commercial & energy emissions

The amount of CO2 emissions produced by the industrial, commercial & energy sector in an area in any given year.

Data for the amount of CO2 emissions produced by the industry, commercial & energy sector is published by the Department of Energy and Climate Change (DECC) by local authority.

Local authority CO2 emissions forecasts within the industry, commercial & energy sectors were produced by first creating UK carbon weights by industrial sector. This was done using sectoral employment and carbon emissions forecasts from the DECC projections to 2035, and projections from the Energy-Economy-Environment Modelling Laboratory's Price-Induced Market Equilibrium System (PRIMES) energy system model thereafter. By dividing the emissions in a sector by the number of people in employment in that sector, then dividing this by the emissions for the average UK worker (total UK emissions divided by total UK employment), we are able to get weights showing how carbon intensive specific sectors are.

For each local authority, we then calculate a carbon weighted employment figure based on what the employment breakdown in that area is. So a district which employs significantly more of their workforce in the emissions intensive chemicals and processing industries sector would be forecast to have a higher carbon weighted employment figure than a district which had a large agricultural sector.

This carbon weighted figure is then multiplied by the average emissions per UK employee, to give a pre-adjusted industrial & commercial emissions forecast. The pre-adjusted forecast also takes into account emissions from the energy sector. These emissions are based on the DECC and PRIMES projections, and we have modelled the energy sector as having no employees as such. Otherwise, we could have a problem where a district with a high number of energy sector employees could be a head office and not really emitting much carbon. So we share the energy sector emissions across districts by multiplying UK energy sector emissions by each district's share of total UK employment.

Finally, we adjust our forecasts based on scaling factors capturing the differences between our calculations for 2005-13 and the 2005-13 DECC data.

Domestic emissions

The total number of emissions produced by households in an area in any given year.

Data for the amount of CO2 emissions produced by the domestic sector is published by the Department of Energy and Climate Change (DECC) by local authority.

Local authority CO2 emissions forecasts within the domestic sector are assumed to be a function of population i.e. more people means more households and therefore more domestic energy use. We have calculated the UK average level of domestic emissions per person by taking the total UK household emissions from the DECC and PRIMES projections and dividing by UK total population. Then we applied this UK domestic emissions per person

ratio to the local authority population forecasts in the EEFM to estimate a pre-adjusted domestic emissions forecast by local authority. Then we adjusted the forecasts based on scaling factors capturing the differences between our calculations for 2005-13 and the DECC data during the same years.

Transport emissions

The total number of emissions produced by the transport sector in an area in any given year.

Data for the amount of CO2 emissions produced by the transport sector is published by the Department of Energy and Climate Change (DECC) by local authority.

Local authority CO2 emissions forecasts within the transport sector are assumed to be a function of GVA (for example, more output means more transport use and therefore more emissions from transport). We have calculated the UK average level of transport emissions per unit of GVA by taking the total UK transport emissions from the DECC and PRIMES projections and dividing by UK total GVA from Cambridge Econometrics' UK Regional Model. Then we applied this UK transport emissions per person ratio to the local authority GVA forecasts in the EEFM to estimate a pre-adjusted transport emissions forecast by local authority. Then we adjusted the forecasts based on scaling factors capturing the differences between our calculations for 2005-13 and the DECC data during the same years.

Land use, land use change and forestry (LULUCF) emissions

The total number of emissions produced via land use (e.g. deforestation, emissions from soils, etc.) in an area in any given year.

Data for the amount of CO2 emissions produced by the LULUCF sector is published by the Department of Energy and Climate Change (DECC) by local authority.

Local authority CO2 emissions forecasts within the LULUCF sector are assumed to be a function of land area i.e. more land gives more potential for deforestation, emissions from soils, etc. We have taken land area data, measured in hectares, from the UK Standard Area Measurements for 2007, and assumed that these values have not changed over time. Then we took UK LULUCF emissions data from DECC for 2005-13, and DEFRA forecasts for 2015 and 2020. For the years in between, we assumed a straight line and extrapolated annual data points and beyond 2020 we assumed a continuation of the trend. Then, using data from DECC for 2005-13, we projected the local authority LULUCF emissions by taking the previous year's emissions, and adding the local authority share (calculated by taking each area's share of total UK land area) of the net change in UK LULUCF emissions in each year.

Total emissions

The total number of CO2 emissions produced in an area in any given year.

This is calculated as an aggregate of industry, commercial & energy emissions, domestic emissions, transport emissions and LULUCF emissions.

4 Data sources

4.1 Labour market

Employees in employment

Description: Annual average employee job estimates

Data: 1991 – 1995 Annual Employment Survey (AES)
1995 – 1997 Annual Employment Survey rescaled to ABI
1998 – 2008 Annual Business Inquiry (ABI)
2008 – 2014 Business Register and Employment Survey (BRES)

Full-time/part-time split

Description: Annual average full-time and part-time employee job estimates consistent with the employee job estimates above.

Data: 1991 - 1995 Annual Employment Survey (AES)
1995 - 1997 Annual Employment Survey rescaled to ABI
1998 - 2008 Annual Business Inquiry (ABI)
2008 – 2014 Business Register and Employment Survey (BRES)

Self-employment

Description: Annual average self-employment job estimates

Data: ONS Workforce Jobs (WFJ)
Census 2001 and 2011 for local area estimates

Employees in Armed Forces

Description: Annual average estimate of employees in UK regular Armed Forces stationed in the UK

Data: DASA, ONS Workforce Jobs

Unemployment

Description: Annual average claimant count unemployment – seasonally adjusted

Data: Local authorities: Nomis – Claimant count with rates and proportions
Region: Nomis – Claimant count with rates and proportions

Residence-based employment

Description: Number of people resident in an area who are in employment (irrespective of where they work)

Data:	Local authorities:	Census of Population (2001 and 2011)
		Annual Population Survey (APS)
	Region:	Census of Population (2001 and 2011)
		Annual Population Survey (APS)

Total workplace employment (people)

Description: the number of people who work in an area (irrespective of where they live)

Data: Local authorities: Census of Population
Region: Census of Population

4.2 Commuting

Description: The number of people that travel into, and out of, an area for work

Data: Local authorities: Constructed by Cambridge Econometrics, Census of Population
Region: Constructed by Cambridge Econometrics, Census of Population

4.3 Demography

Population – total

Description: total population, all ages

Data: Local authorities: National Statistics, mid-year population estimates
Region: National Statistics, mid-year population estimates

Working age population

Description: defined as all people aged 16-64

Data: Local authorities: National Statistics, mid-year population estimates
Region: National Statistics, mid-year population estimates

Young population

Description: population aged 0-15

Data: Local authorities: National Statistics, mid-year population estimates
Region: National Statistics, mid-year population estimates

Elderly population

Description: defined as all people aged 65+

Data: Local authorities: National Statistics, mid-year population estimates
Region: National Statistics, mid-year population estimates

Net migration and other changes

Description: net migration flows to/from an area, including other changes (e.g. boundary adjustments, prisoner movements, boarding school pupils, etc)

Data: Local authorities: National Statistics, components of change
Region: National Statistics, components of change

Natural increase

Description: the numbers of births minus deaths

Data: Local authorities: National Statistics, components of change

Region: National Statistics, components of change

4.4 Output

GVA

Description: Gross Value Added in real 2011 prices

Data: Local authorities: Constructed by Cambridge Econometrics, Regional Accounts
Region: National Statistics, Regional Accounts

4.5 Housing

Demand for dwellings

Description: Stock of dwellings.

Data: Local authorities: DCLG – dwelling stock estimates

Number of households

Description: Households

Data: Estimated by Cambridge Econometrics

4.6 Carbon emissions

Industry, commercial & energy emissions

Description: CO2 emissions from the industry, commercial & energy sectors

Data: Local authorities: DECC – Full local CO2 emissions estimates

Domestic emissions

Description: CO2 emissions from the domestic sector

Data: Local authorities: DECC – Full local CO2 emissions estimates

Transport emissions

Description: CO2 emissions from the transport sector

Data: Local authorities: DECC – Full local CO2 emissions estimates

LULUCF emissions

Description: CO2 emissions from the land use, land use change and forestry (LULUCF) sector

Data: Local authorities: DECC – Full local CO2 emissions estimates

Total emissions

Description: Total CO2 emissions

Data: Local authorities: DECC – Full local CO2 emissions estimates

5 Outliers and data validity

Official data (e.g. BRES employment data) are incorporated unchanged into the EEFM, as the crucial starting point upon which local economic data are founded. Data is then adjusted to be consistent with key regional and national series which offer more timely information around recent economic trends. This process allows Model users to reference key variables at the published source, however as data are adjusted this means that users cannot reference data directly, although the broad levels will remain consistent with the published source.

However, in some cases the data can be anomalous - so-called "outliers." This could be because of errors in measuring or recording it. Or perhaps the data is "true" but reflects an unusual circumstance and so does not accurately represent the local situation or local trends. Because of the smaller numbers of observations, data-reporting errors or unusual "outlier" values can be a particular problem at more detailed levels of analysis - for example, when looking at individual sectors in individual local authorities.

This section explores these issues in respect of the BRES (note: prior to 2008, ABI data is used and subject to similar levels of volatility), and outlines Cambridge Econometrics' approach to BRES data outliers. In summary, this is to keep them unchanged within the EEFM spreadsheets, but to adjust them when making forecasts such that the first year of a forecast would incorporate a correction for an outlier value in the BRES data in a previous year.

5.1 BRES outliers

The latest published BRES data is for 2014 and was released in September 2015. Since BRES data is collected by survey whereby individuals/firms complete the questionnaires, there can sometimes be significant discontinuities in the sector data at local level from year to year. Such discontinuities may - or may not - reflect real events. Consider the effects on the data series of an incomplete return from a firm - or an error interpreting or recording it - in one year preceded (or followed) by a complete or correct return in the previous (or subsequent) year. Any recorded change in employees associated with this would be fictitious, and any trend extrapolated from it into the future would be misleading. But equally, a dramatic change could reflect the opening, expansion, contraction or closure of a major business in an area (with potential longer-term effects on other local businesses).

If a discontinuity occurred in say 2008, but was corrected in 2009, producing a "spike" in the time-series data, it can essentially be ignored as it will not affect the forecasting process. Equally, if it were confirmed the following year, it would suggest a 'real' change in the local economy has indeed taken place. In the meantime, local authorities' input is vital to identify whether discontinuities in the data reflect 'real' events or not.

Focussing on the 2 digit SIC 2007 sectors for employee jobs at local authority level, we identified discontinuities showing **more than a 10% change in number of employees in a single year where this change involved more**

than 1,000 employees. These outliers were sent to appropriate local authority representatives for their reaction and input.

Cambridge Econometrics' response to this consultation was as follows: where we were satisfied that a discontinuity genuinely reflected the opening or closure of a firm, or major expansion or contraction, we accepted the change as the correct starting point for the EEFM forecasts. But if we were given evidence by consultees that there was an error in the BRES data or that an outlier gave a misleading picture of the local situation in some way, we corrected for the discontinuity in the first year of the forecast. (In the absence of any information about a discontinuity, we accepted it, in line with our working principle outlined above.)

Table 5.1 sets out those local authorities and sectors where adjustments were made to 2014 BRES data, showing the size and direction of the correction.

Table 5.1 Adjustments made to 2014 BRES data used in setting forecasts

Local authority	Sector	Correction
Braintree	Education	Up by approximately 900 employee jobs
Breckland	Employment activities	Down by approximately 1,200 employee jobs
Broadland	Finance	Up by approximately 1,300 employee jobs
Harlow	Business services	Down by approximately 1,900 employee jobs
Ipswich	Education	Up by approximately 1,100 employee jobs
Norwich	Employment activities	Down by approximately 400 employee jobs
Norwich	Business services	Up by approximately 2,900 employee jobs
St Albans	Business services	Down by approximately 1,200 employee jobs
St Edmundsbury	Business services	Down by approximately 5,800 employee jobs

Notes: The amount of jobs by which a sector has been adjusted does not necessarily reflect the size of the observed anomaly in the BRES data, as the 2014 adjusted value also includes an element of the trend employee growth that would have occurred if the correction had not been made.

5.2 Data checking and validity procedures

A vital foundation of any economic modelling and forecasting work is ensuring that data is correctly sourced and accurately fed into the model. Cambridge Econometrics has a policy of meticulously summing checking variables and carrying out visual checks throughout the process of updating the EEFM to ensure that the data is fully internally consistent.

Data is entered electronically from original official sources and is checked automatically to make sure identities are maintained. It is also checked visually to assess whether trends look plausible and magnitudes are correct.

There are a number of key identities in the EEFM which must hold for the Model to be fully realised, and we have a spreadsheet within it designed specifically to check that this is the case. These identities are:

- Employee jobs by sector = total employee jobs
- Self-employed jobs by sector = total self-employed jobs
- Employment by sector = total employment
- All indicators in each local authority = Eastern totals (note that this does not apply to productivity and unemployment/resident employment rates)

- Total employment = employee jobs + self-employed jobs + HM Armed Forces
- Total population = working age population + young population + elderly population
- Change in population = net migration + natural increase
- People-based employment = net commuting + resident-based employment
- Labour force = employment + unemployment

There are two principal methods that we apply to our models to ensure variables add up correctly over the forecast period:

1 **Scaling:** it is often the case that model input or output variables which are theoretically identical actually have different values. This is usually due to errors or incompleteness in the underlying data or methodological differences in gathering them. Scaling is the process by which two such variables are made equal by raising one to the value of the other, and the procedure can either be multiplicative or additive. Additive scaling takes the difference between the variables and adds it pro rata to the components of the lower of the two (for example, to local authority values when the total of these is less than a regional value to which it should theoretically be equal). Multiplicative scaling takes the ratio of the "target" total to the actual total, and multiplies each component of the actual total by that ratio. In this way, the actual total is shifted upwards (or downwards) to meet a target total which it should theoretically equal.

2 **Residual:** this procedure is used when the value of one component (or a small number of them) can be approximately deduced from the known values of other components and a known total. For example, estimating full time jobs as the residual between total jobs and part time jobs.

6 Employment land use methodology

This chapter outlines our methodology for calculating employment land use forecasts under the 2016 update of the East of England Forecasting Model (EEFM).

6.1 Key outputs

The summary outputs under the employment land module for EEFM 2016 for the East of England and each district include:

- Industrial floorspace (B1c/B2), thousands m²
- Warehouse floorspace (B8), thousands m²
- Office floorspace (B1a/b), thousands m²

Detailed outputs including the variables above split by sector are available on the website.

6.2 Measure of employment

The employment forecasts used in the calculation to estimate employment land requirements are:

- Jobs-based
- Workplace-based
- Full-time equivalents (estimated as the number of full-time employed, plus 75% of the number of part-time employed)

6.3 Employment densities

The employment densities used within the EEFM are based on the Employment Densities Guide, published in 2010[1], which provides guidelines on employment densities by use class. The guide presents densities on a range of different floorspace measures: gross external area (GEA), gross internal area (GIA) or net internal area (NIA). Therefore, it has been necessary to convert all employment densities to the same measure - GIA.

Table 6.1 Employment densities by use, 2010 guide

Use	Use class	Use Type	Area per FTE (m²)	Floor Area Basis	Comment on potential variation
Industrial	B2	General	36	GIA	Range of 18-60 m²
Industrial	B1 (c)	Light Industry (Business Park)	47	NIA	
Warehouse & Distribution	B8	General	70	GEA	Range of 25-115 m²
Warehouse & Distribution	B8	Large Scale and High Bay Warehousing	80	GEA	

[1] Employment Densities Guide, Homes & Communities Agency, 2010

Office	B1 (a)	General Office	12	NIA
Office	B1 (a)	Call Centres	8	NIA
Office	B1 (a)	IT/Data Centres	47	NIA
Office	B1 (a)	Business Park	10	NIA
Office	B1 (a)	Serviced Office	10	NIA

The following employment densities have been adopted for Industry and Warehousing, based on the general use types. The GEA for warehousing has been converted to GIA by using the CLG's Regional Spatial Strategy and Local Development Framework Core Output Indicators – Update 2/2008 guidance[2] which assumes a 3.75% difference.

For office use, the HCA guidance states that the GIA is typically 15-20% higher than net internal space. Using this figure this provides an employment density range for general office of 13.8 m² - 14.4 m².

Table 6.2 Employment densities – industry, warehousing and office (GIA)

Use	Use type	Density: Area per FTE (m²)	Notes:
Industry	B1c/B2	36	Uses General Industry
Warehousing	B8	67	Uses General Warehousing
Offices	B1	14 (based on the average of the 13.8-14.4 range)	Uses General Office

For detailed office uses the same process has been followed for call centres, business parks and serviced office whilst office headquarters are assumed to follow the general employment land density. As the guidance does not provide densities for R&D, science parks and small businesses uses these are assumed to follow the original densities from the 2001 guide. An alternative could be to use the B1c density, given the earlier employment land density guide showed densities for these uses similar to light industry. However, this would result in an overall density of around 60m², which seems very high when compared to the 2001 densities and is very close to the warehousing density.

Overall the following employment densities for detailed office use are used.

Table 6.3 Employment densities detailed office use

Use	Sub-use	Density: Area per FTE (m²)	Notes:
Office	B1b use split:		Based on 2001 density guide
	Science park & Small business units	32	
	High tech R&D	29	
	B1a split:		Based on NIA densities adjusted to GIA (average range of 15-20%)
	General office	14	
	Serviced business centre & Business park	13	
	Call centre	10	

[2] https://www.gov.uk/government/publications/employment-densities-guide

6.4 Allocating employment sectors to use classes

In order to forecast employment land it is necessary to convert the employment sector forecasts into office, warehousing and industrial uses. As the model provides employment sector forecasts by 31 sectors in total (comprising one or several 2 digit SIC codes) we have allocated each sector across the use classes in differing proportions. This analysis has been largely based on reviewing each SIC code in detail and judging the overall proportion that could be expected to be in industry, warehousing or office uses based on our knowledge of the East of England's economy. This is not an exact science as the classification of economic activities does not always lend itself to a straightforward allocation.

The EEFM sectors are mapped to use classes in differing proportions, as outlined in Table 6.4. Those sectors marked with a * need careful consideration given the nature of the activities undertaken, namely:

- Waste and remediation - we have allocated 97% of these activities to industry use to capture waste treatment activities (based on employee share in BRES by detailed SIC codes).
- Construction - we have not included construction in B-use, however, we are aware that often this is classified as industry use.
- Wholesale trade and repair of motor vehicles and motorcycles - we have allocated 75% of this sector to warehousing based on the share of wholesale warehousing activities in the BRES numbers. The remaining 25% associated with the repair of motor vehicles has been allocated to industry.
- Land transport - we have allocated 39% of this sector to warehousing based on the share of warehousing and support activities for transportation in the employee BRES numbers.
- Professional services - we have allocated 96% of this sector to offices. We have excluded veterinary activities based on the share of employees in the BRES numbers.
- Business services - we have allocated 93% of this sector to offices. We have excluded travel agency, tour operator and other reservation services based on the share of employees in the BRES numbers.
- Employment activities - given that this sector includes temporary workers that may work in any industry we have allocated employment based on the weighted shares of all the other sectors' allocations to industry, warehousing and offices.
- Publishing & broadcasting activities - we have allocated all publishing activity to industry. For motion picture, video and television programme production, sound recording and music publishing activities which captures the production side of film and TV we have assigned 80% to warehousing given the large scale production sets often required and 20% to office use. For programming and broadcasting activities which incorporates broadcasting activities which are most likely to be studio based we have assigned 80% of these activities to office use and 20% to warehousing use. The proportions are then scaled depending on the relative employment shares in the BRES data.

- Telecoms - we have allocated 80% of telecoms to warehousing and the remaining 20% to offices.
- Public administration - we have allocated 61% of this sector to offices to take account of the share of general public administration activities; regulation of the activities of providing health care, education, cultural services and other social services, excluding social security; regulation of and contribution to more efficient operation of businesses; and foreign affairs. We have excluded defence activities; justice and judicial activities; public order and safety activities; fire service activities; and compulsory social security activities. The shares are based on the BRES data.

We would appreciate feedback on these sectors or any others, bearing in mind that a simple calculation is applied across the East of England. Densities and allocations are static across the decades in the spreadsheets, as we have made no assumptions about the impacts of changing working practices. We have applied assumptions across the whole region, rather than reflecting any local circumstances. An interactive version of the spreadsheets is available so that users can apply their own assumptions to reflect any specific local circumstances. Please see the Cambridgeshire Insight website for more information.

Table 6.4 Allocation of employment sectors by use class, SIC07

SIC code	SIC description	Industry	Warehousing	Office
		B1c/B2	B8	B1
01-03	Agriculture			
05-09	Mining and quarrying			
10-12	Food manufacturing	100%		
13-18, 31-33	General manufacturing	100%		
19-23 excl 21	Chemicals excl. pharmaceuticals	100%		
21	Pharmaceuticals	100%		
24-25	Metals manufacturing	100%		
28-30	Transport equipment, machinery & equipment	100%		
26-27	Electronics	100%		
35-37	Utilities			
38-39*	Waste and remediation	97%		
41-43*	Construction			
45-46*	Wholesale	25%	75%	
47	Retail			
49, 52-53*	Land transport		39%	
50-51	Water and air transport			
55-56	Hotels and restaurants			
58-60*	Publishing and broadcasting	66%	23%	11%
61*	Telecoms		80%	20%
62-63	Computer related activity			100%
64-66	Finance			100%
68	Real estate			100%

69-75 excl 72*	Professional services			96%
72	Research & development			100%
77-82 excl 78*	Business services			93%
78*	Employment activities	12%	8%	22%
84*	Public administration			61%
85	Education			
86-88	Health and care			
90-93	Arts and entertainment			
94-99	Other services			

6.5 Detailed office uses

The sectors with some element of office use have also been assigned into the more detailed breakdown of office uses as shown in Table 6.5 below. Again, we would appreciate any feedback on these allocations.

Table 6.5 Allocation of office employment sectors by detailed office use classes, SIC07

SIC code	SIC description	Offices	Split by:				
		B1	B1b Science Park & Small business units	B1b Tech/R&D	B1a General Office	B1a Serviced Business Park Centre & Business	B1a Call Centre
58-60	Publishing and broadcasting	11%	0%	0%	11%	0%	0%
61	Telecoms	20%	0%	0%	20%	0%	0%
62-63	Computer related activity	100%	0%	0%	30%	60%	10%
64-66	Finance	100%	0%	0%	100%	0%	0%
68	Real estate	100%	0%	0%	90%	10%	0%
69-75 excl 72	Professional services	96%	7%	7%	79%	2%	1%
72	Research & development	100%	20%	60%	10%	10%	0%
77-82 excl 78	Business services	93%	71%	1%	9%	4%	9%
78	Employment activities	22%	5%	1%	13%	2%	1%
84	Public administration	61%	0%	0%	61%	0%	0%

cambridge econometrics
clarity from complexity

Technical Report | September 2016 | Cambridge Econometrics, Cambridge, UK | sp@camecon.com, www.camecon.com