ERC
Enterprise Research Centre

The UK's high growth firms and their resilience over the Great Recession

ERC Research Paper 62

September 2017

www.enterpriseresearch.ac.uk

The UK's high growth firms and their resilience over the Great Recession

Not for quotation without permission

Michael Anyadike-Danes
Enterprise Research Centre and Aston Business School
m.anyadike-danes@aston.ac.uk

Mark Hart
Enterprise Research Centre and Aston Business School
Mark.hart@aston.ac.uk

The Enterprise Research Centre is an independent research centre which focusses on SME growth and productivity. ERC is a partnership between Warwick Business School, Aston Business School, Imperial College Business School, Strathclyde Business School, Birmingham Business School and Queen's University School of Management. The Centre is funded by the Economic and Social Research Council (ESRC); Department for Business, Energy & Industrial Strategy (BEIS); Innovate UK and the British Business Bank. The support of the funders is acknowledged. The views expressed in this report are those of the authors and do not necessarily represent those of the funders.

ABSTRACT

During the last decade High-Growth Firms (HGFs) – sometimes referred to as 'Scale-Ups' – have increasingly become an established feature of the UK business policy landscape. Indeed, HGFs are mentioned in the government's recently published policy document "Building our Industrial Strategy", and are now considered sufficiently important that the Minister for Small Business has taken on the role of "Scale-Up Champion". Whilst we know something of the characteristics of these firms – about their age, size, sector and location – we know relatively little about the dynamics of the HGF population as it evolves over time. For the most part attention is focused simply on the annual count which, as we shall see, is not an entirely appropriate measure of HGF activity.

Our point of departure is to develop a measurement framework which distinguishes the first High-Growth Episode (HGE) and marks the 'birth' of an HGF (i.e., the first time it can be categorised as a HGF) from its subsequent repeat HGEs. Tracking HGFs over their lifetime, we show that almost two-thirds of HGEs recorded during a growth period, and conventionally referred to as HGFs, are actually repeat episodes being recorded by HGFs 'born' some years previously. By making the firm/episode distinction we are able to calculate the age of HGFs more appropriately (average HGF age at birth not average age at which an HGE is recorded). Finally, we are able to show that because the population of HGFs is relatively long-lived, by about age 15 HGFs make up around 40% of all firms with more than 10 employees, or to put it slightly differently, by age 15 40% of all firms with more than 10 employees have experienced at least one HGE.

With the onset of the Great Recession the number of HGEs declined quite markedly. Initially, there was a fall in HGF births, which was (necessarily) followed by a reduction in repeat episodes (induced by the fall in births). However, in the upswing repeat episodes recovered first: existing HGFs recorded an unusually large number of repeat episodes; and this was followed by a return of HGF birth numbers to their pre-Recession level. What our analysis of the Great Recession period shows is that there is, apparently, a relatively stable mechanism producing HGEs period by period.

Keywords: high-growth firms; birth cohorts; firm demography; firm growth
JEL codes: D22; E24; L11; L25; M13

CONTENTS

1. CONTEXT AND MOTIVATION

During the last decade High-Growth Firms (HGFs) – sometimes referred to as 'Scale-Ups' – have increasingly become an established feature of the UK business policy landscape. Indeed, HGFs are mentioned in the government's recently published policy document "Building our Industrial Strategy", and are now considered sufficiently important that the Minister for Small Business has taken on the role of "Scale-Up Champion". Whilst we know something of the characteristics of these firms – about their age, size, sector and location – we know relatively little about the dynamics of the HGF population as it evolves over time. For the most part attention is focused simply on the annual count which, as we shall see, is not an entirely appropriate measure of HGF activity.

We develop a measurement framework designed to track the population of high-growth firms (HGFs) between 1998 and 2015: the innovation here is the explicit allowance for firms which experience 'repeat' episodes of high growth. The relationships between the categories in the measurement framework are used to build a simple 'model' of the evolution of the HGF population. We use ONS data from the 14 birth cohorts of UK firms born between 1998 and 2011, covering the three year growth periods from 1999/02 to 2012/15 to parameterise the model.

We provide a plausibility check of the models properties by generating a large set HGF birth cohorts which are then aged to match the pre-1998 cohorts of HGFs which are not distinguishable in our data, these simulated old cohorts are aggregated and compared to the observed pre-1998 cohort data over the 1998/01 to 2013/16 growth periods. Analysis of the synthetic cohorts is also used to compute the share of repeats in the annual count of HGEs and the average age of HGFs at birth. Finally, the model is used to generate a counterfactual path for the UK HGF population over the Great Recession (GR) period, the difference between the observed and the average path provides an estimate of how the HGF population was affected by the downturn and its re-bound during the recovery phase.

2. MEASUREMENT MATTERS: TAKING THE MANUAL OF BUSINESS DEMOGRAPHY SERIOUSLY

Here we investigate HGFs using the recommended three year growth period. The Eurostat-OECD metric for identifying an HGF (see EUROSTAT-OECD [2007, Chapter 8]) requires that we count firms which,

- are born before the beginning of the period
- are alive at the end of the period[1]
- have at least 10 employees at the beginning of the period
- **record an annual average growth of 20% in employment[2] over the period**

A little later on, and in a rarely noticed section, **The Manual of Business Demography** continues,

> "The identification of high-growth enterprises on an annual basis may lead to the inclusion of an enterprise in the population of high-growth enterprises in several years. The question arises whether a high-growth enterprise ... should be counted in more than one reference year if it fulfils the given definition. The recommendation is to do so." EUROSTAT-OECD [2007, p.63]

So we have here is an explicit recognition that when HGFs are to be counted over successive annual – and therefore necessarily overlapping – growth periods a firm may be counted more than once. And this recognition is the starting point for our discussion here.

Specifically we introduce the distinction between HGFs and high growth episodes (HGEs). Adapting the terminology of **The Manual of Business Demography**, we define a High Growth Episode (HGE) as a three year period in the life of a firm over

[1] Taken together these first two conditions imply that in each period we will have a 'balanced panel' of firms – the same firms are always present throughout the period.
[2] Alternatively, an annual average growth of 20% in turnover over the period can be used as the criterion, but only employment is used here.

which its job numbers grow by 72.8% from a base of 10 or more. This, in turn, allows us to distinguish the date of a firm's first episode from any subsequent episode. In other words the first HGE is, essentially, the period in which a firm is first categorised (or, 're-born') as an HGF.

Obviously distinguishing between episodes and firms gives some structure to the relationship between HGE numbers and HGF numbers: having experienced one HGE an HGF may (possibly) never experience another, or (more likely) it may experience many more. Consequently there will no longer be (necessarily) a one-to-one relationship between the numbers in the HGF population and the number of HGEs. So, a key purpose of the measurement framework set out here is to track the evolution of the population of HGFs and alongside it the numbers of HGEs they record. However it should be borne in mind that the number of HGEs is, in fact, the conventional measure of the number of HGFs.

Before going any further it is worth anchoring our discussion of the HGE/HGF distinction by having a brief look at some aggregate count data[3] for HGEs, labelled HGFs in the conventional treatment. Figure 1 displays the HGEs for 15 growth periods from 1998/01 to 2012/15. Two features of the data stand out. First, the striking 'bulge' in HGEs between 1999/02 and 2001/04. This seems to have been a by-product of the 'high tech boom' (it was concentrated in just a few sectors) around the turn of the century (see Anyadike-Danes et al. [2012, p. 12 and Figure 5.6, p.52]). The second is the 'bite' taken out of the series in 2008/11 and 2009/12 and associated with the Great Recession (GR) period.

[3] The data used here is from the Jobs version of the Longitudinal Business Structure Database which can be accessed through the Secure Lab. A more detailed account of data construction can be obtained from the authors on request. The use of these data does not imply the endorsement of the data owner or the UK Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

With short time series and two major disturbances it is quite hard to determine the underlying trends from these numbers alone. Nonetheless we can see that the number in 2013/16 (10,865) is almost the same as it was a decade earlier in 2003/06 (10,757): typically the overall count is in a relatively narrow band between 9,000 and 11,000. As we shall show almost two-thirds of these episodes in each growth period are 'repeats': HGEs recorded by HGFs born previously.

3. ACCOUNTING FOR HGES

Because the most important aspects of firm performance – survival and growth– are strongly age-dependent, we start by classifying the population of firms into birth cohorts. We can then use the first HGE recorded by a firm within a birth cohort of firms to distinguish 'sub-cohorts' of HGFs. So the 'birth' of an HGF sub-cohort is registered in the year of its first HGE. In other words, HGFs within a birth cohort of firms are recorded as 'sub-cohorts' indexed by firm age at the beginning of a three year growth period. As noted earlier **The Manual of Business Demography** requires (in our terminology) that if a firm is to be counted as an HGF it must be at least one year old when it records an HGE. So, for example, the first opportunity for firms in the 1998 firm birth cohort (cohort98) to record an HGE is 1999/02: that is firms born in 1998 which record an HGE in the period 1999/02 – 10 or more jobs in 1999, and 72.8% (or more) growth in jobs between 1999 and 2002 – and these comprise the first sub-cohort of cohort98.

Beginning, for simplicity, with a single birth cohort of firms, we have in the first block of the HGF/HGE accounts the numbers born into each, successive sub-cohort. Using cohort98 for concreteness (and in obvious notation) we can write the HGF numbers for the first three growth periods (1999/2002, 2000/2003, and 2001/2004) as,

$$HGF_{99/02} \equiv sub_1^1$$
$$HGF_{00/03} \equiv sub_1^2$$
$$HGF_{01/04} \equiv sub_1^3$$

The superscript attached to 'sub' denotes the 'age' at which the sub-cohort is 'born': so it is the birth cohort birth year plus one. For example, sub[1] refers to the sub-cohort of cohort98 HGFs 'born' in 1999[4]. Each of the sub-cohorts also has a subscript which denotes its place in the sub-cohort order. Since all of these three are the first in their respective growth periods they all have a '1' subscript. As we shall see, the subscript allows us to keep track of a sub-cohort as the cohort ages.

We have also a parallel set of accounts for HGEs. Using an 'r' pre-fix to denote a repeat episode in a sub-cohort, 'rsub1' denotes the number of firms in 'sub1' recording a repeat episode. However because, as noted earlier, a sub-cohort is (potentially) long-lived, we also add a subscript to denote at what firm age the repeat episode takes place. By definition the earliest date at which the first sub-cohort can record a repeat is in the second growth period – $rsub_2^1$; and of course the sub-cohort can record a further repeat in the third growth period – $rsub_3^1$.

So we can write the cohort98 HGE accounts as,

$$HGE_{99/02} \equiv HGF_{99/02} \equiv sub_1^1$$
$$HGE_{00/03} \equiv HGF_{00/03} + rsub_2^1 \equiv sub_1^2 + rsub_2^1$$
$$HGE_{01/04} \equiv HGF_{01/04} + rsub_3^2 + rsub_3^1 \equiv sub_1^3 + rsub_3^2 + rsub_3^1$$

As we can see, the terms accumulate: each growth period there is one new sub-cohort born and each earlier sub-cohort adds a further 'batch' of repeats.

Before going any further it is worth illustrating the contribution of sub-cohorts to the HGE total. Figure 2 provides data for a single growth period, 2012/15. Each of the sub-cohort contributions for six of the 15 birth cohorts which make up the growth period total have been plotted to display the relationship between the age and the relative size of the sub-cohort 'strata'. On the left of the plot we have three 'old' cohorts, 1998 to 2000, on the right three 'young' cohorts, 2009 to 2011. In this display (as noted earlier) we have just one sub-cohort per cohort (sub1 for cohort11, sub-cohort 2 for cohort12, down to sub-cohort14 for cohort98), the rest

[4] Since HGFs are defined by a growth period a HGF birth is here being recorded using the initial year of the growth period as the reference date.

of the contributions are repeat episodes from earlier sub-cohorts (repeat of sub1 for cohort10, repeats of sub1 and sub2 for cohort09, down to 13 sub-cohorts for cohort98). Clearly, the age of firms producing HGEs in any period – the age of HGFs themselves in the conventional treatment – is a weighted average of all birth cohorts of firms where the weights are the size of the sub-cohorts. However, as we can now see, a very large proportion of the HGEs being recorded are in fact repeats, contributed by HGFs which recorded their first HGE sometime in the past. Evidently the average age of HGFs will be very different: they are likely to be very much younger than HGEs. We will return to the question of age below.

4. THE BUILDING BLOCKS EMPIRICALLY EXAMINED

Our ultimate purpose here is to investigate the evolution over time of the number of HGFs and the number of HGEs they record. We treat the sub-cohort birth numbers as given and then, as we can see from the accounts, 'explaining' HGE numbers reduces to modelling the number of repeat episodes. So we assume,

- the time path of sub-cohort births in a cohort
- the number of repeat episodes recorded by a sub-cohort in a growth period is proportional to the size of the sub-cohort in the previous growth period
- the numbers of a sub-cohort alive in a growth period depend on the proportion of the sub-cohort which died during the previous period

Now we examine, in turn, data on the three components of the 'model': sub-cohort births; the death ratio; and the repeat ratio.

4.1 sub-cohorts of HGFs

Figure 3 plots the numbers at birth in each sub-cohort (on a log scale) against firm age at the beginning of the growth period for the birth cohorts from 1998 to 2009[5] with a separate panel for each cohort. And, of course, the number in each sub-cohort is, by our definition, the number of HGFs being born at that age. The overall count of HGFs in a particular growth period is then a sum across cohorts. For

[5] The cohort for 2010 has just two data points and 2011 just one, and these are not easily visible on the multi-panel display, so the values are given in the notes to Figure 3.

example, the count for the 2005/08 growth period is the cohort98 count for age 7 (i.e. sub-cohort 7), plus the cohort99 count for age 6 (i.e. sub-cohort 6), and so on up to cohort04 age 1 (i.e. sub-cohort 1).

The fine solid line on each panel has been added to aid visualisation. It is an average over cohorts of births by sub-cohort, and it gives us a simple cross-cohort frame of reference. As we can see, within each birth cohort the number of HGFs in each sub-cohort declines with age, but none of the cohorts appear to depart very far, or permanently, from the cross-cohort average. Most cohorts start with sub-cohort 1 relatively close to the average figure of 677 episodes.

The two striking exceptions are cohort09 and cohort10 (not plotted) which are more than 15% below the average, but as you will appreciate both these sub-cohorts are born in (potentially) GR-affected years (which we will discuss below). The rate of decline in the average from sub-cohort 1 to sub-cohort 2 is very steep – there are around 450 episodes in sub-cohort 2, about one third lower than sub-cohort 1 – the rate of decline then slackens off, and falls quite gradually by between 10% and 15% per period. By age 14 the number of sub-cohort births is down to around 90.[6]

The dotted vertical line on each panel is drawn at the 2008/11 period as a further aid to cross-cohort comparison. It draws attention to one striking feature of the data: the sharp dip in HGF numbers associated with the GR. Since HGE identification is based on growth over a three-year period it is not obvious *a priori* which growth period(s) will be most affected by a sui generis event 'like' the GR, however, the results of our non-parametric approach seem intuitively plausible. Of course, the firm age at which the GR downturn occurs varies systematically across the cohorts, from age 10 for cohort98 to age 1 for cohort07. The size of the impact appears to be inversely related to age – it appears larger in the older cohorts/sub-cohorts – but remember the data is plotted on a log scale, so what we are observing is a larger proportionate effect.[7] The effect of the 2008/11 drop in numbers is made

[6] It should be noted that as age increases the number of observations in the average decrease. By age 14, the only observation is that of cohort98.

[7] The impact the on the 1998 to 2005 birth cohorts is quite similar – between 50 and 60 – when measured in absolute numbers. The effects on the relatively 'new-born', cohort06 and cohort07 are rather larger – about 75 – despite being proportionately smaller.

more striking by the return to trend after the GR which is quite pronounced too, although there is some cross-cohort variation in the timing of the recovery. Moreover, there is also an almost universal 'up-tick' in 2012/15, the last growth period, which may be a delayed recovery effect. We will discuss the GR and the recovery from it in section 6 below.

4.2 the death ratio

So far we have only counted the episodes recorded in the 'birth year' of a sub-cohort. However in every growth period subsequent to that birth year some firms in the sub-cohort will, as we know, record a further HGE, a 'repeat' episode. Rather than simply count these episodes for each of the sub-cohorts within each birth cohort, we compute the 'repeat ratio': the number of repeat episodes in a growth period divided by the number of HGFs in the sub-cohort still alive at the beginning of the growth period. However, in order to compute the repeat ratio we need first to look at its denominator – the number of HGFs alive – which, in turn, requires an examination of the death ratio: the proportion of the sub-cohort which dies from one period to the next.

For each sub-cohort within a birth cohort we can compute a death ratio – the number of HGFs in the sub-cohort which die during a period divided by the number of HGFs alive at the beginning of the period. Now it is well-established that the chance of a firm dying declines with age, and so it is for HGFs too. However, the series of ratios by age for each sub-cohort of each birth cohort are quite 'noisy'.[8] At least in part because, and certainly for older ages, the numbers in a sub-cohort still alive are quite small, and the number of deaths, necessarily even smaller (and the latter seem quite sensitive to 'events'). Nevertheless the overall pattern is quite clear: the HGF death ratio, like the 'all firm' death ratio, declines with age. Figure 4 displays the

[8] Indeed for evidence of 'noisy' year-to-year variation in the UK all-firm death ratio see Anyadike-Danes and Hart [2017b, Figure 5]

ratio, within a cohort averaged over sub-cohorts, cohort-by-cohort from cohort98 to cohort09, in a multi-panel display. Again a cross-cohort average has been add to aid visualisation and comparison.

Whilst declining with age, the pattern across the age range shows considerable variation, and there do appear to be some discrete shifts. In the periods just preceding the onset of the GR cohort04 and cohort05 start well above average (the first deaths are 2006/09 and 2007/10), whilst later on, in periods of GR recovery, cohort09 and cohort10 (deaths in 2012/2012 and 2010/13, see note to Figure 4) were well below the overall average. The sub-cohort ratios[9] within a cohort depend on the age of the firm, not the age of the sub-cohort. So, for example, the death ratio for the firms from sub-cohort 2, that is firms from the sub-cohort which died at firm age 3, is similar to the death ratio for firms age 3 from the first sub-cohort, that is firms which died between the second and third periods for sub-cohort 1.

4.3 the repeat ratio

For each growth period in the life of each sub-cohort we can express the number of HGFs recording a repeat episode as a ratio to sub-cohort size that is the numbers in the sub-cohort still alive. An analysis of this ratio, sub-cohort by sub-cohort, within each cohort reveals a rather surprising finding: the repeat ratio varied very little across sub-cohorts. In other words, it is the years since the birth of the sub-cohort, not years since the birth of the cohort, which is important for the chance of a repeat episode. Clearly, for HGFs age is a more subtle influence on growth than for firms in general for whom (on average) the chance of growth declines with age (see Anyadike-Danes and Hart [2015] and Anyadike-Danes and Hart [2016a]). Figure 5 displays the proportion averaged over sub-cohorts from our 14 different cohorts plotted against sub-cohort 'age'. The proportion displays a striking

[9] The numbers by sub-cohort are not reported here, not only do they show wide variation they are, in many cases, below the threshold for disclosure.

pattern. At age 2, the first year after the birth of a sub-cohort, half of the HGFs record a repeat episode; at age 3 the proportion is one third; but after age 3 the proportion drops steeply to about 10%, and thereafter declines very slowly.

5. PUTTING THE BUILDING BLOCKS TOGETHER

5.1 a simple mechanical 'model'

We can now bring together the simple model built from the empirical relationships. In the first growth period the number of HGEs is just the number of HGFs born in the first period (the first sub-cohort, sub_1^1),

$$HGE_1 = sub_1^1 \qquad (1)$$

In the second growth period the number of HGEs is the new-born period 2 sub-cohort (sub_1^2) plus the repeat episodes from period 1's sub-cohort ($rsub_2^1$). Symbolically,

$$HGE_2 = sub_1^2 + rsub_2^1 \qquad (2)$$

Using δ to represent the death ratio, and β the repeat ratio, we can write,

$$sub_2^1 = (1 - \delta_2) \times sub_1^1 \qquad (3)$$
$$rsub_2^1 = \beta_2 \times sub_2^1 \qquad (4)$$

Substituting (3) and (4), we can re-write (2) as,

$$HGE_2 = sub_1^2 + \beta_2 \times (1 - \delta_2) \times sub_1^1 \qquad (5)$$

Using the 'reduced form', equation (5), as a template for the relationship in the third growth period, we have the new-born third sub-cohort plus the repeat episodes from period 1 and period 2, simply adjusting the time subscripts.

$$HGE_3 = sub_1^3 + \beta_2 \times (1 - \delta_3) \times sub_1^2 + \beta_3 \times (1 - \delta_3) \times (1 - \delta_2) \times sub_1^1$$

There is one minor, but noteworthy, feature of the time subscripts: the repeat coefficient on sub² is β_2 whilst the death ratio is δ_3. This reflects the two different 'calendars' in operation here: the repeat ratio depends on sub-cohort age; whilst the death ratio depends in birth cohort age (that is firm age).

5.2 the pre-1998 cohorts

We use the averages for the sub-cohort numbers (from Figure 3), to compute values for HGFs and HGEs for a single cohort. But before investigating the characteristics of that cohort, we provide some evidence on the plausibility of the model. As mentioned earlier, we have no birth date information for cohorts born before 1998, so here we use our model to provide a projection of the HGE numbers which might be produced for the growth periods from 1998/01 to 2013/16 by this collection of these older cohorts.

We start by computing HGEs for 20 'average' cohorts, using the averages (as set out above) for 14 years and then projected trends in those averages for ages 15 to 20.[10] We then treat one replicate as if it were born in 1977, the next in 1978, and so on, up to and including 1997. The data start with the growth period for 1998/2001, which is the first observed growth period in our data. In that period we have the episodes recorded by: the first sub-cohort of the (hypothetical) cohort97; for the (hypothetical) cohort96 we have the second sub-cohort, plus repeats from its first sub-cohort; and so on, going back to the repeats from the 20th sub-cohort of the (hypothetical) 1977 birth cohort. Using the model parameters the numbers are then projected forward to 2012/15. The results are plotted against a

[10] The age 20 cut-off is arbitrary but since we have no actual data on beyond age 14 it seemed best not to push the projected trends too far into the future.

log scale on Figure 6 together with the observed data on the combined pre-1998 cohort HGEs for 1998/01 to 2012/15.

There is a considerable degree of consistency between the observed data and the projected data produced by accumulating values from the 20 average cohorts (the underlying model components after age 14 are log-linear trends which produce a straight line when the data are plotted on a log scale). There are two large swings in the observed around the projected: 1999/02 to 2001/04; and 2008/11 to 2011/14. We mentioned both of these earlier when we discussed Figure 1. The first is apparently associated with the 'high tech' boom, the effect of the first was largely confined to cohort97 (Figure 3 shows little evidence of it affecting cohorts 1998 to 2001), whilst the second, is the GR which as we saw (again from Figure 3) affected cohorts 1998 to 2006).

5.3 a single hypothetical cohort

Let us now look at the evolution of a single hypothetical cohort. Period by period numbers of HGFs alive ('alive'), numbers of repeat episodes ('repeats') and the total HGEs ('episodes') are plotted on Figure 7, together with the sub-cohort numbers ('births') which drive the model. Results are displayed for 20 years although, as we know, beyond age 14, births, the death ratio, and the repeat ratio are 'projected' trends in the averages since we only have data for the first 14 years.

Starting at the top of the chart we have the size of the HGF population. Initially the numbers rise quite steeply, but as births slow (albeit at a steady rate) the numbers of HGF deaths rise with population size (even though the death ratio is declining) and, from about age nine birth and death numbers are almost equal, so the population size flattens. At age nine it is 2,145, and peaks at 2,198 at age 14, but by age 20 it is only slightly lower at 2,166. The number of HGEs follows a rather different path, from a very young age, after age 2, it begins to fall, and then continues to do so at a fairly steady rate. The divergence between HGF and HGE numbers can be explained quite straightforwardly. As we saw earlier (on Figure 3) the sub-cohort birth average falls, and the detail is easier to see on Figure 6: the number drops from around 700 at birth down to around 50 by age 20. The number of repeats declines too, from very early in the life of the cohort (age 3 in fact) but very much more slowly because it depends on the size of the HGF population.

Repeat numbers continue to accumulate, even though the contributions from the birth of the successively smaller sub-cohorts is limited: by age 20, many sub-cohorts are only contributing 3 or 4 episodes each, but by that stage there are 20 sub-cohorts. Looking back to Figure 2, at the cohort98 bar, and comparing it to our Figure 7 plot at year 14 – around 300 episodes of which 80 are births, so just over 200 repeats –approximately matches the numbers in cohort98. And we can see that for cohort98 in at age 14 much of the 'weight' of repeats in the total comprises very small numbers contributed by relatively elderly sub-cohorts.

Age has been a subject of (continuing) discussion in the HGF literature, it had been expected that HGFs were typically young firms, but this was not confirmed (for the UK see Anyadike-Danes et al. [2012, Table 5.1, p43]). In the conventional treatment age is measured as the age at the time a growth period's HGF count is recorded. 'Age' is of course is the age at which an HGE has been recorded, but unless it is a firm's first HGE it will simply be the age at which an HGF born previously has recorded a repeat episode. Clearly, this affects conclusions about the age distribution of HGFs, which is likely to be quite different from the age distribution at which an HGE is recorded. By age 20 our hypothetical cohort has recorded just over 3,900 births, and half of these HGFs were born before age 5 (as measured from the beginning of the growth period, age 8 at the end). If an HGE is recorded by a firm more than 5 years old it is more likely to be a repeat, and at age 20 it is three times more likely to be a repeat.

In fact, we can compile a more complete picture of the importance of repeats for the growth periods covered here. For the cohorts from 1998 onwards we of course have the data, but for the pre-1998 cohorts we have to rely on our aggregated synthetic cohort data.[11] Figure 8 displays the repeat proportion across the growth periods and we see immediately that with just one (very early) exception it is in the 60% to 70% range. In other words, on average **around two-thirds** of the HGEs

[11] One further adjustment is necessary, though, the data on the pre-1998 cohorts is generated by summing over replicated 'average' cohorts which, as we saw from Figure 6, does not incorporate the effects of either the 'high tech bulge' of the early 2000s or the GR. Here, as a first approximation, we have applied the proportion of births from the replicated aggregate pre-1998 cohorts to the observed total. This is not an ideal procedure since the fluctuations about the average may affect the proportions, but it seems a plausible starting point.

recorded by the UK's HGFs in the last decade and a half are repeat episodes. To be sure the proportion fluctuates, and the widest swings occur in and around the GR period, with the largest shares being recorded in 2008/11 and 2011/14. We will look at these GR related phenomena in more detail in the next section.

Figure 8 also displays, plotted against the right hand axis, the average age of HGFs – that is the average age at which a firm records its first HGE – over the growth periods 1998/01 to 2012/15. The average is computed, for each period, by weighting each birth sub cohort by its age, where we use the age at the initial year of the growth period, so for example the first sub-cohort of cohort98, an HGE recorded in 1999/02 is assigned the age 1. For the pre-1998 cohorts we use the same (simulated) data just described in the discussion of the repeat proportion. The most striking feature of the age series is its remarkable constancy – it hardly varies at all – average age is never greater than 7.0 years, and rarely lower than 6.0 years. It is also possible to compute the average age at which an episode is recorded – it is just episode-weighted age. Of course we expect it to be considerably older since, as we know there is a long tail of relatively old (albeit small) sub-cohorts, and the result is around 11 years. However, because of the relative importance of the tail in the calculation, and the fact that for the pre-1998 (oldest) cohorts we are using simulated data, that estimate should not be regarded as reliable as the estimate of the average age of HGFs.

We saw on Figure 7 that the number of HGFs alive grows considerably over the life of our synthetic cohort. To put that finding in context we can compare HGF numbers alive as a ratio to the total number of firms with 10+ employees, computed growth period by growth period.[12] This ratio has been computed by age for each cohort and is displayed on Figure 9 a multi-panel chart against a log scale. To improve visualisation a vertical dashed line has been added at 2009/12 where the ratio seems to have a distinct, possibly GR related, 'kink'; and as in previous cases the mean has been computed age-by-age and is plotted as a fine solid line. Clearly the HGF share of the 10+ firm population increases very steeply early in the life of the cohort: it doubles between age 1 and age 5, from 0.16 to 0.32. The rate of

[12] To be precise, the denominator is the number of firms with at least ten employees at the beginning of a growth period which survive to the end of the growth period.

change seems then to slow, reaching around 0.4 by age 10, and appears to be approaching an asymptote around 0.45 (although we have only a handful of observations after age 10). Although there is some variation across cohorts the picture looks pretty consistent: by age 10 40% of all 10+ firms have experienced at least one HGE.

6. HGFS AND HGES BEFORE, DURING, AND AFTER, THE GREAT RECESSION

We have already seen evidence of the effect of the GR on sub-cohort births (Figure 3) but here we provide a more systematic account which covers repeat episodes as well, and as we shall see repeat episodes are an important part of the story. The approach is quite straightforward. We use our simple model to project HGEs cohort by cohort (as we did for the pre-98 group) and then compute the difference between observed and projected. As noted earlier the GR effect on the pre-1998 cohorts is clearly visible on Figure 6. In absolute numbers, the HGEs were around 750 below projected in 2008/11 and then by about 1,000 in 2009/12. Although the 'gap' between observed and projected narrowed in 2010/13 the recovery was quite slow and HGE numbers for these older cohorts did not return to the projected trend until 2012/15.

We now present some more detailed results covering the periods 2006/09 to 2012/15 for the post-1997 cohorts which we have divided into two groups: 1998 to 2005; and 2006 to 2011. The solid line on the chart is the value for observed less projected summed over both groups and the GR period stands out very clearly. In 2008/11, the impact was to reduce the number of HGE episodes by about 1,500; the observed less predicted numbers remain about the same amount in 2009/12; but in 2010/13 the negative effect is roughly halved. Although both groups feature substantially in the downswing, they perform rather differently in the recovery phase. The older group – cohorts 1998 to 2004 – recover relatively quickly: they are at or above projected average values from 2010/13 onwards. By contrast, the younger group had a rather slower, recovery, but in 2012/15 they recorded about 500 HGEs above average (about the same as the older group).

Panel (b) covers the same two groups of cohorts but here they are combined, and instead births and repeats are distinguished. There are some striking differences.

First we can see in the GR period that in 2008/11 the impact is comprised entirely by births then, in the following period, 2009/12, it is a lower level of repeats which accounts (almost) exclusively for the reduced HGE numbers. Combining the birth and repeat effects has the effect of producing similar-sized GR effects in 2008/11 and 2009/12 as we saw in panel (a). Since the mechanism driving repeat episodes in the model is births in the preceding periods, the GR effect on repeats in 2009/12 and 2010/13 is hardly surprising.

The recovery phase cannot be so simply interpreted. Although births appeared to pick up slightly in 2010/13, clearly numbers were still well below average in 2011/14. Looking back to Figure 3 we can see the proximate source of this effect: it is almost entirely attributable to two very young cohorts, cohort09 and cohort10. For cohort09 births were below average in 2010/13, its first growth period, as well as in 2011/14; whilst 2011/14 is the first growth period for cohort10 and the number of births (recorded in the note at the foot of the figure) is just over 100 below the first sub-cohort average. The GR may have a longer term depressing effect on these two cohorts.

Even more anomalous is the 2011/14 surge in repeats which is very clearly not being driven by a preceding upturn in births. Whilst we have no 'explanation' for 500 additional repeats, we know (as a matter of accountancy) that it must represent an upward shift in the repeat ratio. To dig a little deeper into the data here we have computed the repeat ratio by sub-cohort against age, cohort-by-cohort and it is displayed on Figure 10. A vertical line has been added at 2011/14 to make it easier to locate the 'up-tick' in the ratio which produced the divergence between the observed figure for episodes and that produced by the model. For ease of reference a pair of horizontal dashed lines have also been added at the mean values of the first and second repeat ratios. You will notice too, that this upward shift is not sustained. In every case it turned down in the following period, in most cases it had fallen back quite close to around the mean. It appears that this aspect of the recovery – a temporary upward shift in the repeat ratio – might be best described by a surge in 'animal spirits': existing HGFs recorded an unusually large number of repeat HGEs.

7. WHAT HAVE WE LEARNED?

Our key departure is to develop a measurement framework which distinguishes the birth of a high growth firm – the growth period in which it records its first high growth episode – from the 'repeat' high growth episodes which it records subsequently.

Since the early 2000s the number of HGEs per three year growth period has fluctuated within a relatively narrow band – between 9,000 and 11,000 – so the average number over that period, 10,000, provides a reasonable estimate of the numbers which might be expected. Of course this degree of predictability of HGEs at the population level – the conventional measure of HGF numbers – does not imply that it is similarly easy to predict which firms might comprise that population total. We have argued elsewhere (see Anyadike-Danes and Hart [2015]) that firm growth by size should be regarded as a non-stationary markov chain – the proportions moving from one size-band to another are relatively predictable, but an individual firm's growth prospects have a chance element attached – and the analysis here suggests that the HGE population in any growth period might be regarded similarly.

Not only is an individual firm's growth prospects – the chance of recording an HGE – difficult to predict, but we now know that the evolution of the HGE time series of counts combines two different dynamics. First the time path of HGF births, what we refer to as sub-cohorts within a birth cohort of firms, and secondly the time path of the proportion of each sub-cohort which records a repeat HGE. Both these time paths depend very strongly on age. The first HGF sub-cohort in any cohort is around 700 firms, but the number falls away very quickly (exponentially) with firm age, and by age 10 the number in the new sub-cohort is less than 150. The value of the repeat proportion decays even more quickly: the chance of a repeat in the period after an HGF birth is around one half; and then a third after two periods; after which it drops to around 10%, where it remains.

Putting these dynamics together in a simple model we find that the average age at which an HGF is born – that is records its first HGE – is about six years. However, because HGFs continue to record repeat episodes for many years after birth, almost two-thirds of the HGEs in any given three year growth period are repeats, that is firms which have previously recorded at least one HGE.

We have shown elsewhere how the 'collapse' in firm births and its subsequent effects on the size of the population of continuing firms played a central role in accounting for the Great Recession (see Anyadike-Danes and Hart [2017b]). We find a similar mechanism at work in the population of HGFs: at the onset of the Great Recession HGF births (that is, firms recording their first HGE) decline, and pull repeat episodes down with them. The take-off into recovery is initiated by a strong (autonomous) increase in repeats, which is followed by a return of HGF 'births' to longer term average levels.

Our analysis provides a clearer understanding of the dynamics of HGFs over time, and although there is a clear age decay effect, it reveals the importance of HGEs in driving growth in these firms. This is an important consideration in any attempt by policy makers to identify the lead indicators for HGFs, a group of firms which plays such an important role in job creation.[13] Currently, as flagged in the Industrial Strategy Green Paper in January 2017, the Scale-up agenda is to have a prominent role in driving local growth, with the focus on the importance of identifying, targeting and supporting more HGFs or scale-ups. In practice though, identifying a firm about to have its first HGE may involve quite different lead indicators than the indicators which might be relevant for firms having a second or third HGE. Moreover, it is important to recognise that the first sub-cohort in a birth cohort of firms grows faster and makes a more substantial contribution to job creation than do subsequent sub-cohorts (see Anyadike-Danes and Hart [2017a, Figure 7] for evidence).

[13] According to Anyadike-Danes and Hart [2016b] HGFs contribute between one third and one half of job creation by job creating firms between 199/02 and 2011/14.

REFERENCES

Anyadike-Danes, Michael, Karen Bonner, and Mark Hart (2012) “Exploring the incidence and spatial distribution of high growth firms in the UK and their contribution to job creation,” Working Paper 13/05, NESTA.

Anyadike-Danes, Michael and Mark Hart (2015) “All grown up? The fate after 15 years of the quarter of a million UK firms born in 1998,” mimeo, Aston Business School and Enterprise Research Centre. To appear: Journal of Evolutionary Economics.

(2016a) “Peeling back the layers: separating the effects of age and size on UK job growth, 2008-2015,” mimeo, Aston Business School and Enterprise Research Centre. Available from: https://www.researchgate.net/profile/Michael Anyadike-Danes.

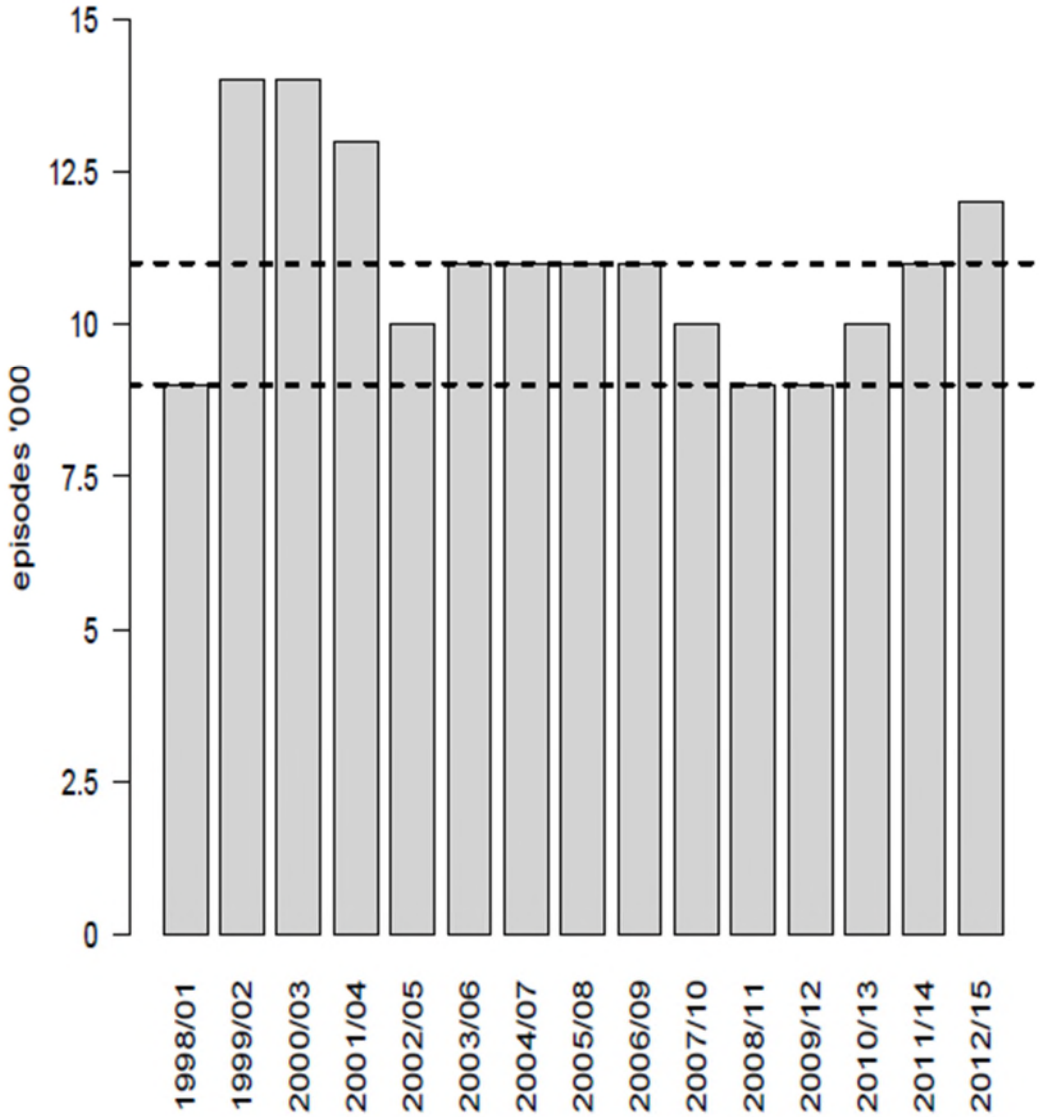
(2016b) “Should we be worrying about high growth firms?: A forensic investigation of job growth in the UK,” mimeo, Aston Business School and Enterprise Research Centre. Available from: https://www.researchgate.net/profile/Michael Anyadike-Danes.

(2017a) “An integrated treatment of firm growth and job creation,” mimeo, Aston Business School and Enterprise Research Centre. Available from: https://www.researchgate.net/profile/Michael Anyadike-Danes.

(2017b) “Firm and job dynamics in the United Kingdom before, during and after the global financial crisis: Getting in under the hood,” in OECD ed. Business Dynamics and Productivity, Paris: OECD, Chap. 4, pp. 87–109.

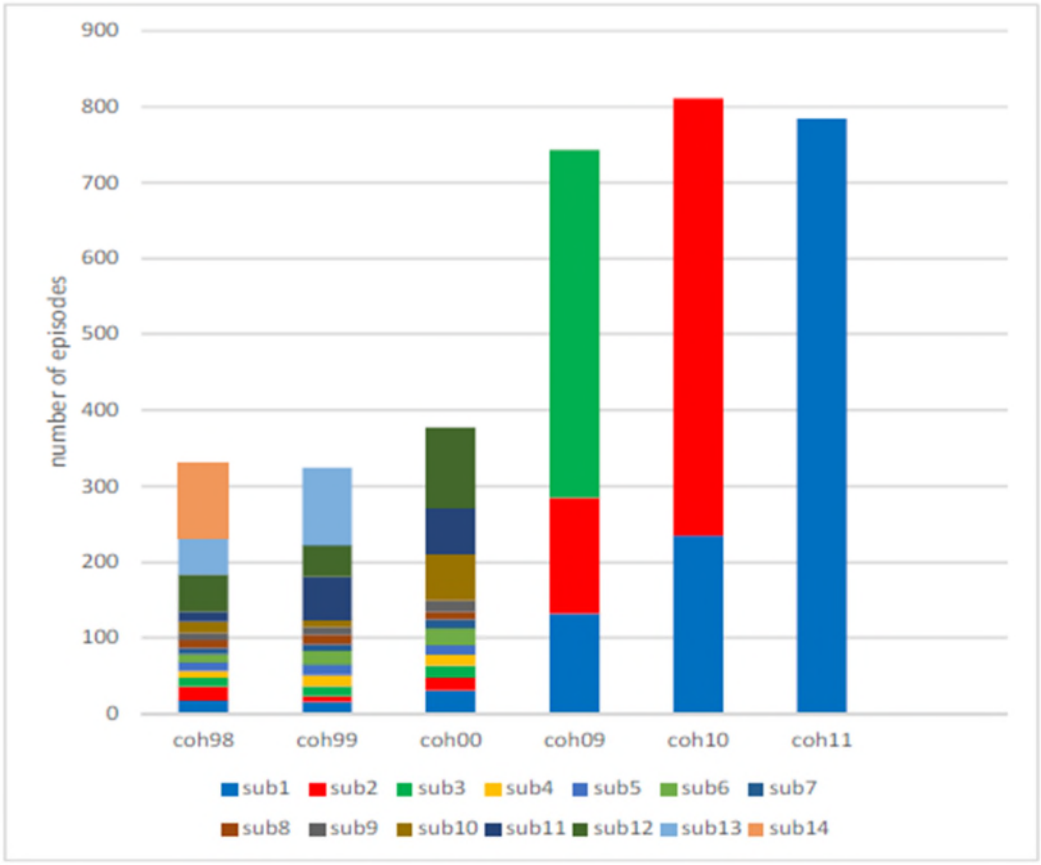
EUROSTAT-OECD (2007) EUROSTAT – OECD Manual on Business Demography Statistics, Luxembourg: EUROSTAT.

Figure 1: HGE episodes, 1998/01 to 2012/15, numbers (’000)



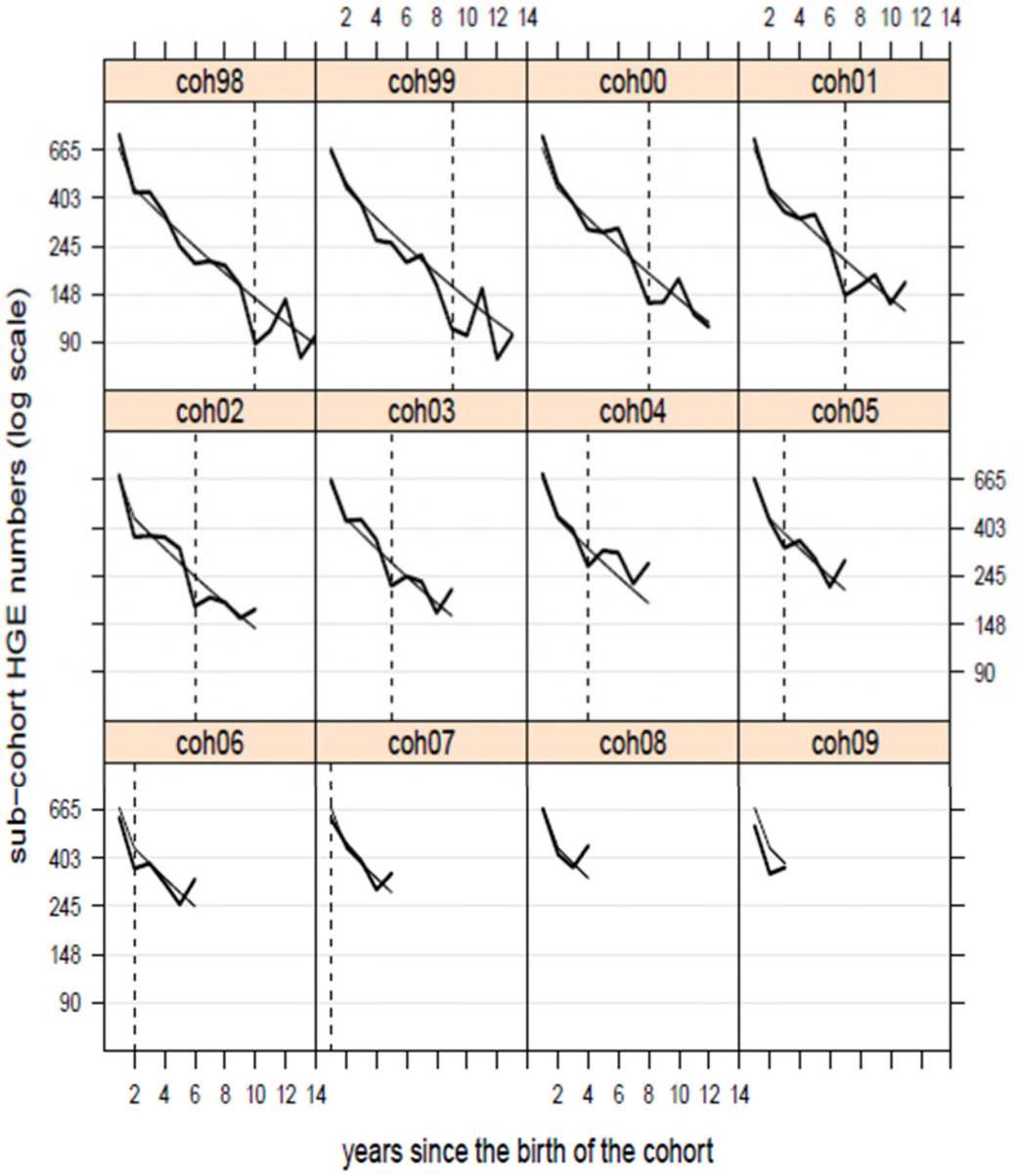
Note: the dashed horizontal lines are at the 9,000 and 11,000

Figure 2: HGE in 2012/15, sub-cohort contribution by selected cohorts, number of episodes



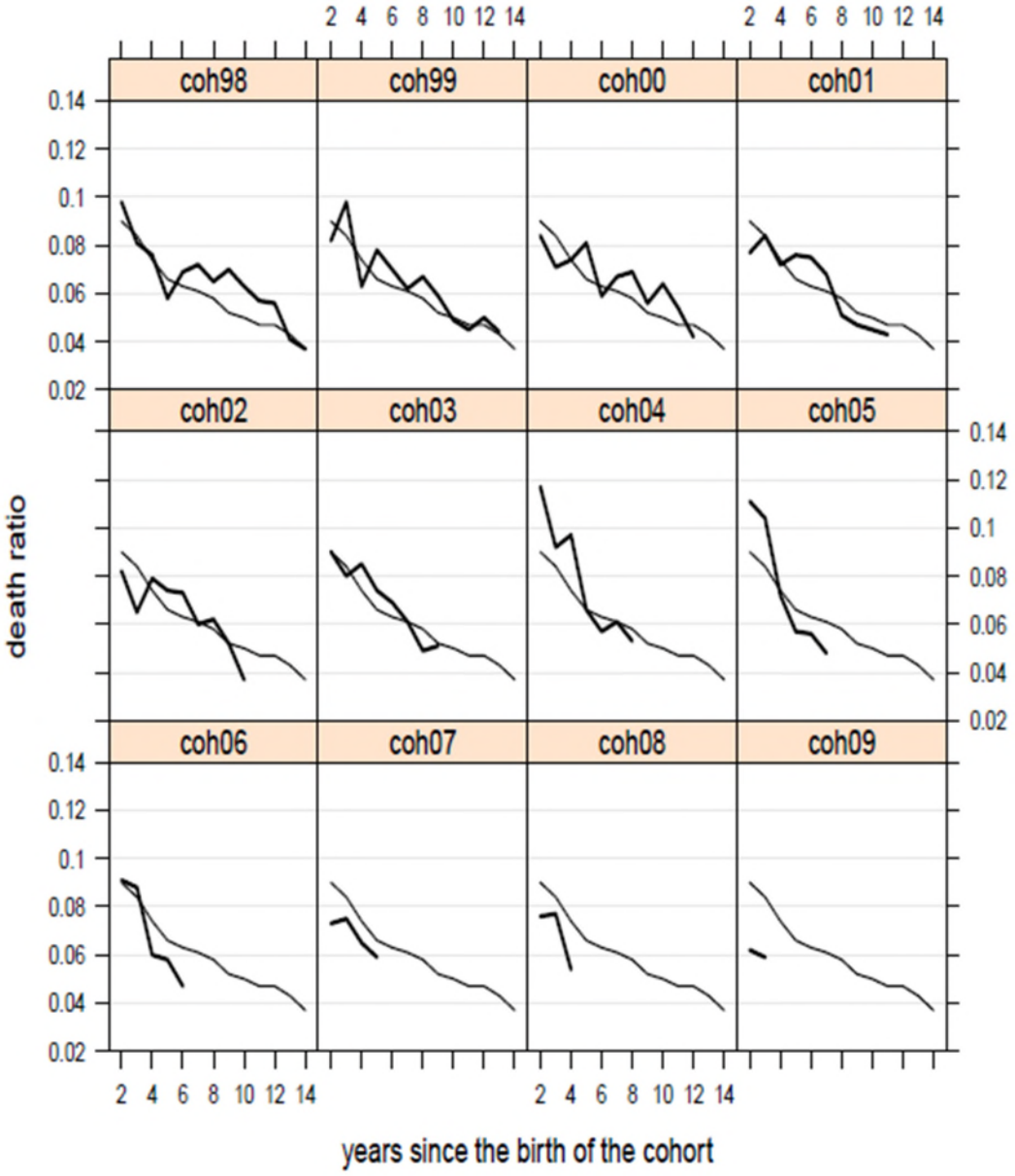
Note: only the sub-cohort ('sub') at the top of each 'stack' is the 'birth' sub-cohort, all the others are 'repeats'.

Figure 3: HGF sub-cohort births by birth cohort, cohort98 to cohort09, numbers (log scale)



Note: the sub-cohort numbers for most recent cohorts are: cohort10, 562, and 576; cohort11, 784.

Figure 4: HGF sub-cohort death rate by birth cohort, cohort98 to cohort09, ratio



Note: the first sub-cohort death ratio for cohort10 is 0.064; by definition there are no deaths for cohort11

Figure 5: HGF sub-cohort repeat ratio by age, averaged over all sub cohorts, ratio

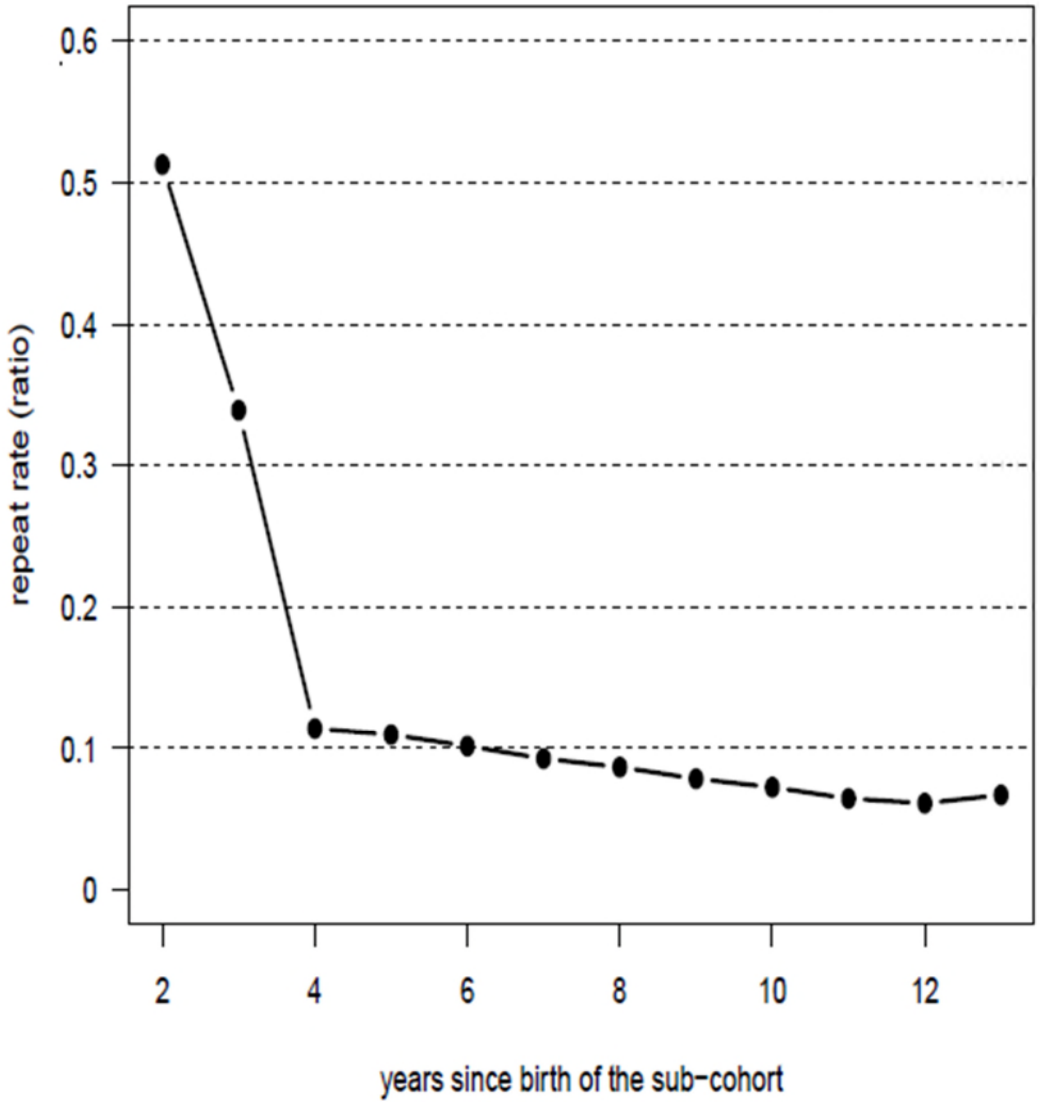


Figure 6: projected HGE numbers pre-1998 cohorts (log scale)

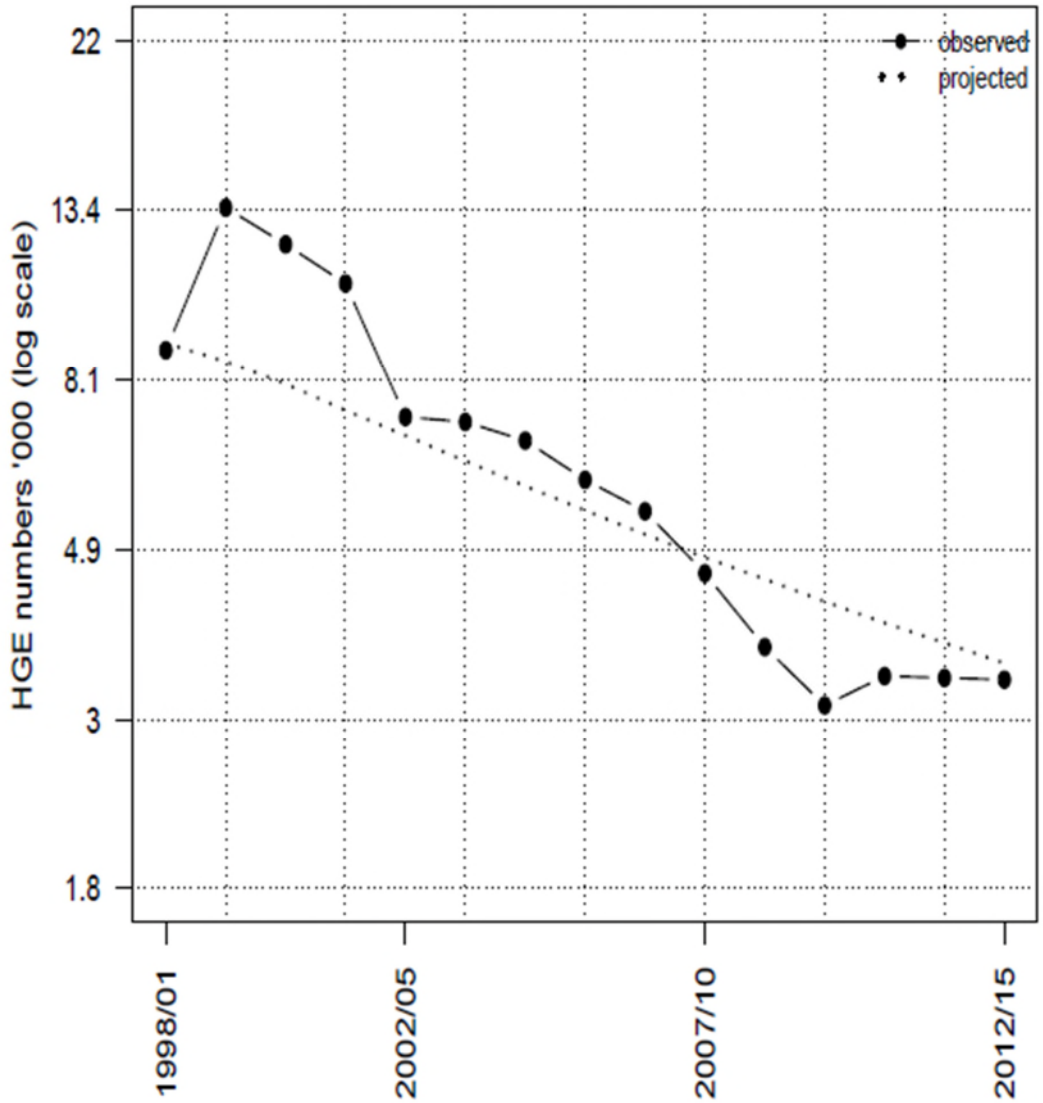


Figure 7: HGFs, HGEs, and repeat episode numbers projected from birth to age 20 (log scale)

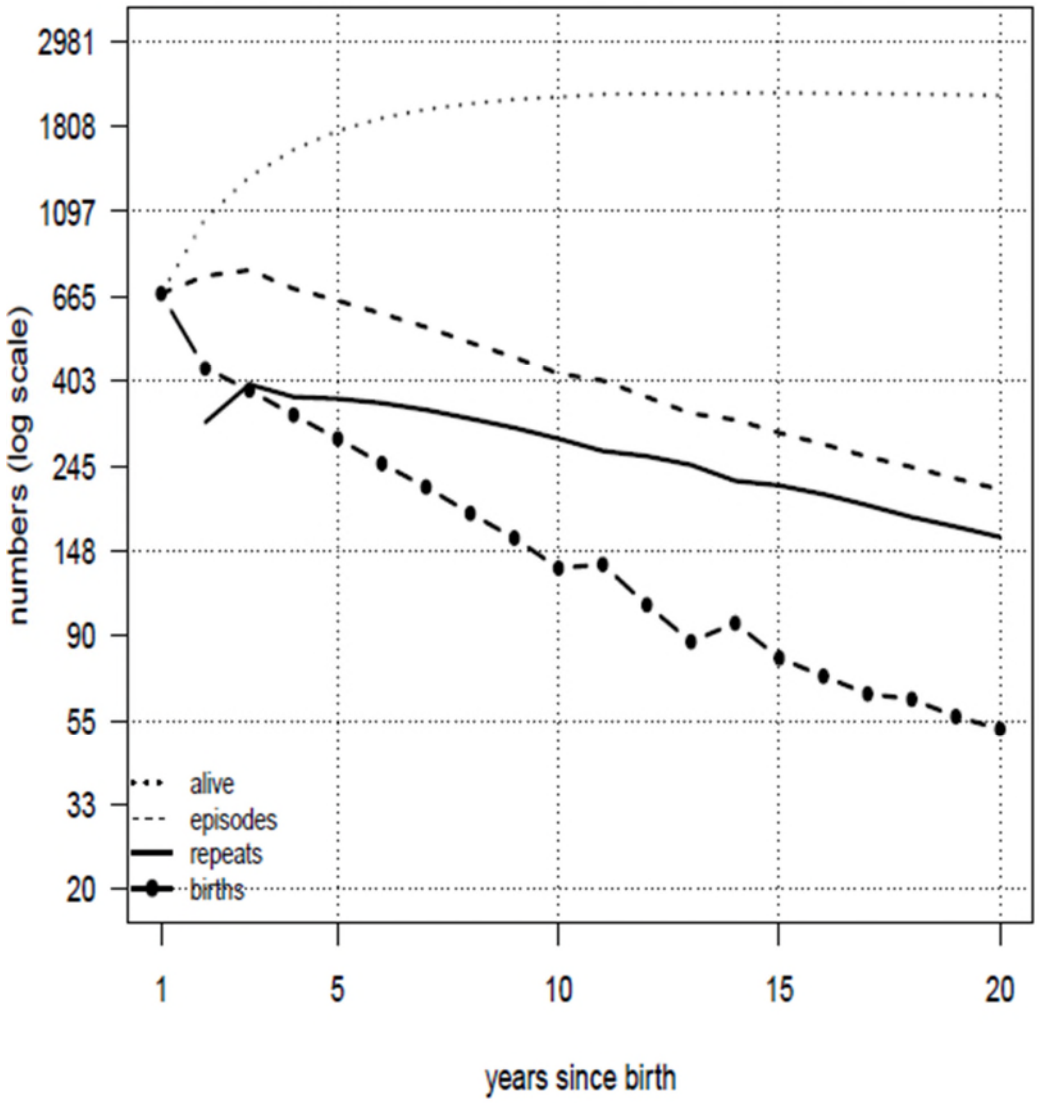


Figure 8: HGE repeat proportion (%) and HGF average age (years), 1998/01 to 2012/15

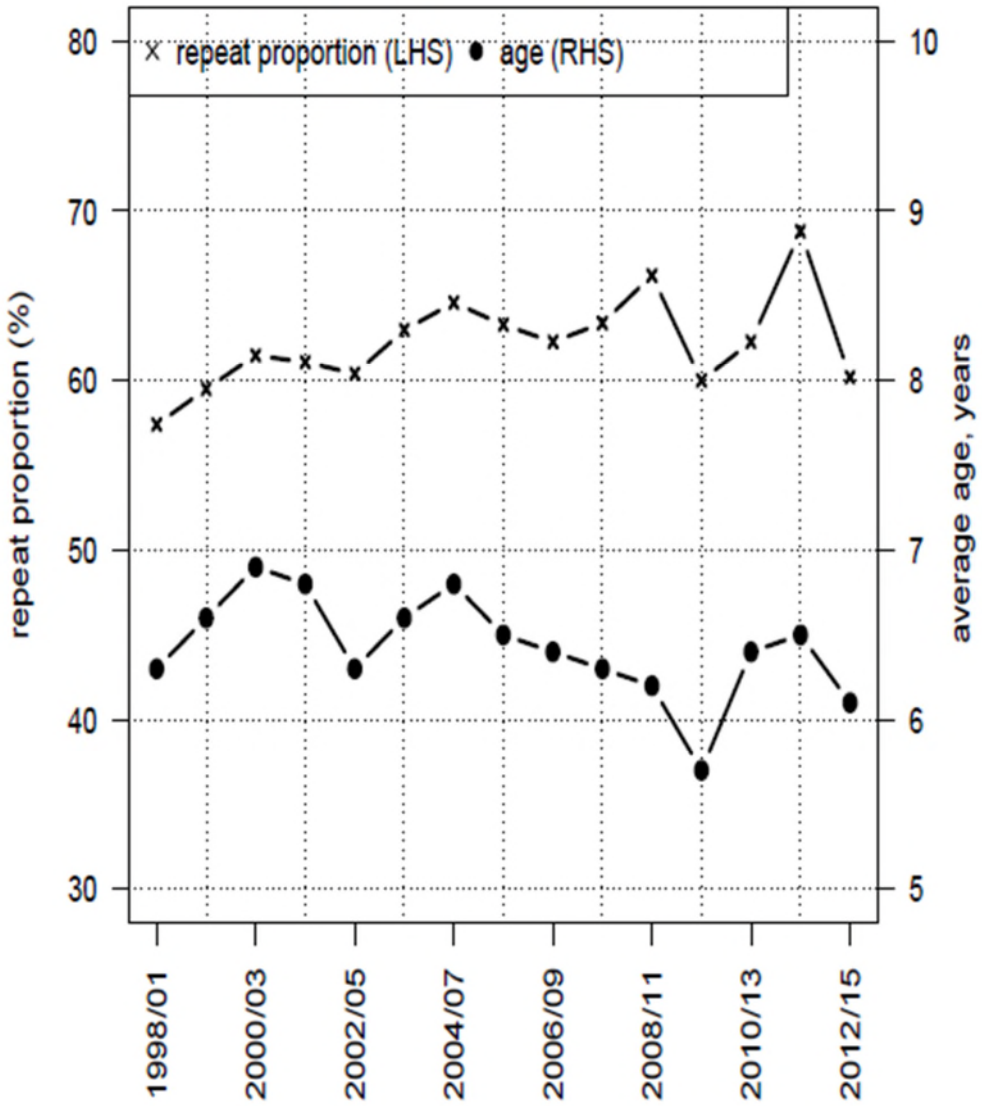
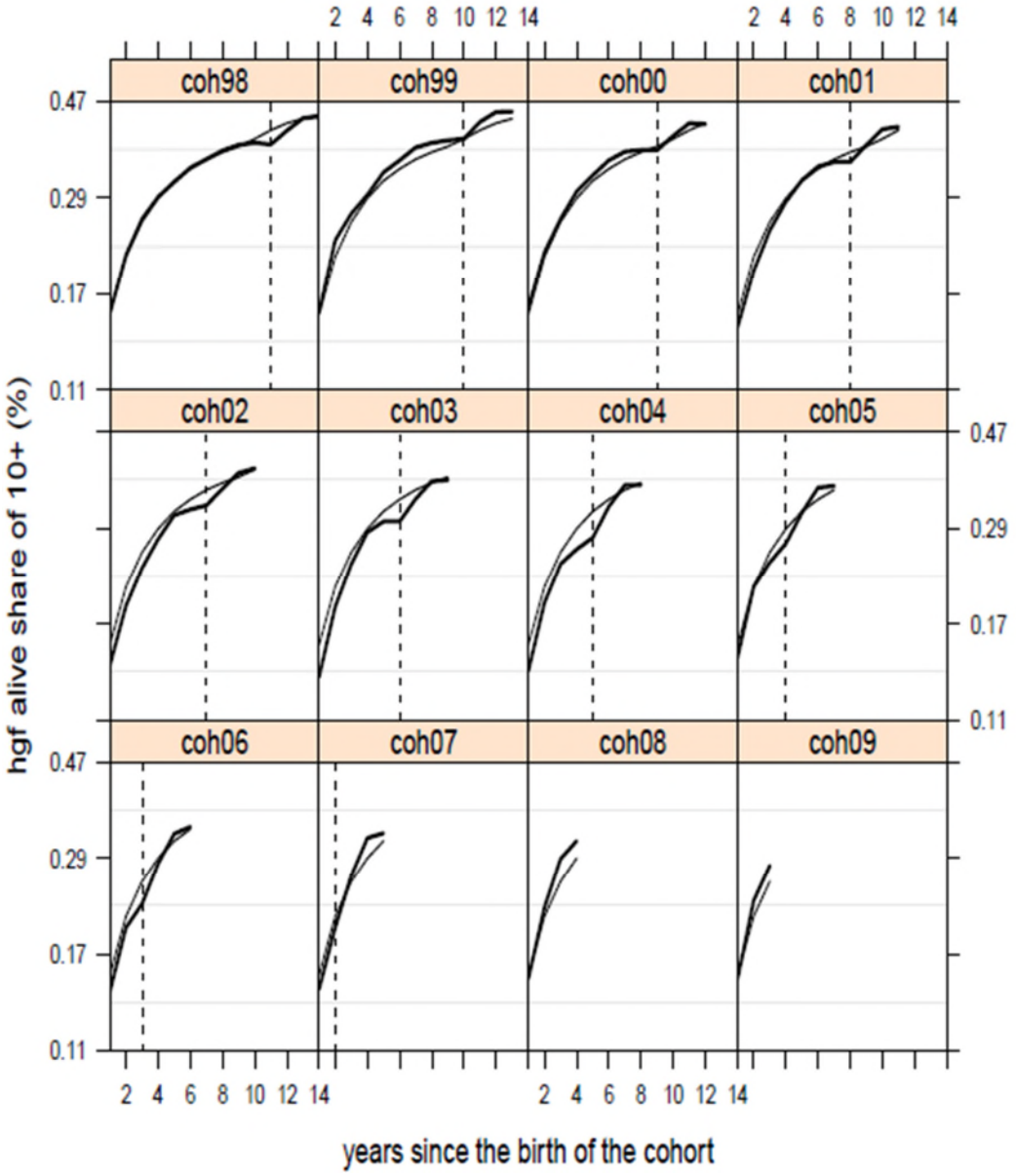


Figure 9: HGF alive share of 10+ alive by age, cohort98 to cohort09, ratio (log scale)



Notes:
1. the ratios for the most recent cohorts are: cohort10, 0.178, and 0.231; cohort11, 0.231.
2. vertical dashed lines at 2009/2012

Figure 10: HGEs actual less projected, births and repeats, 2006/09–2012/15, numbers

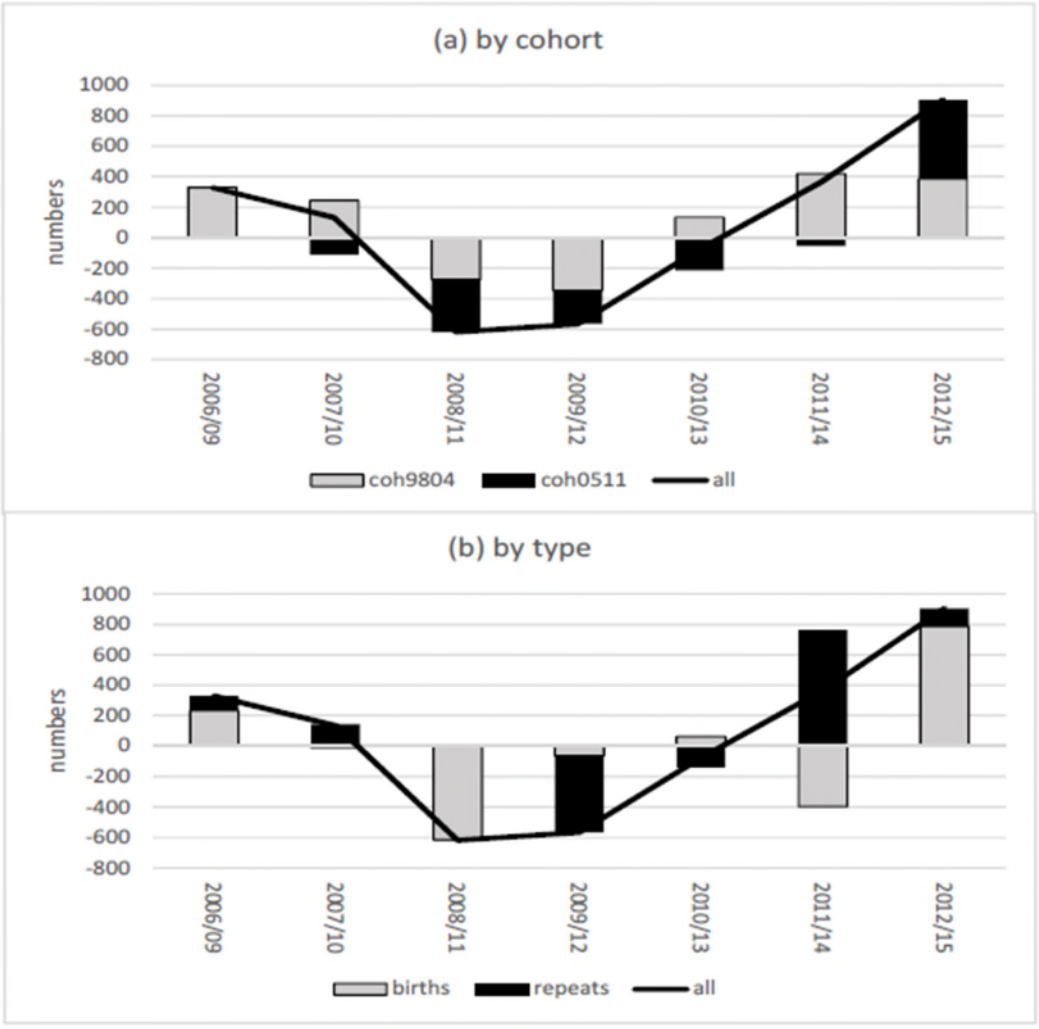
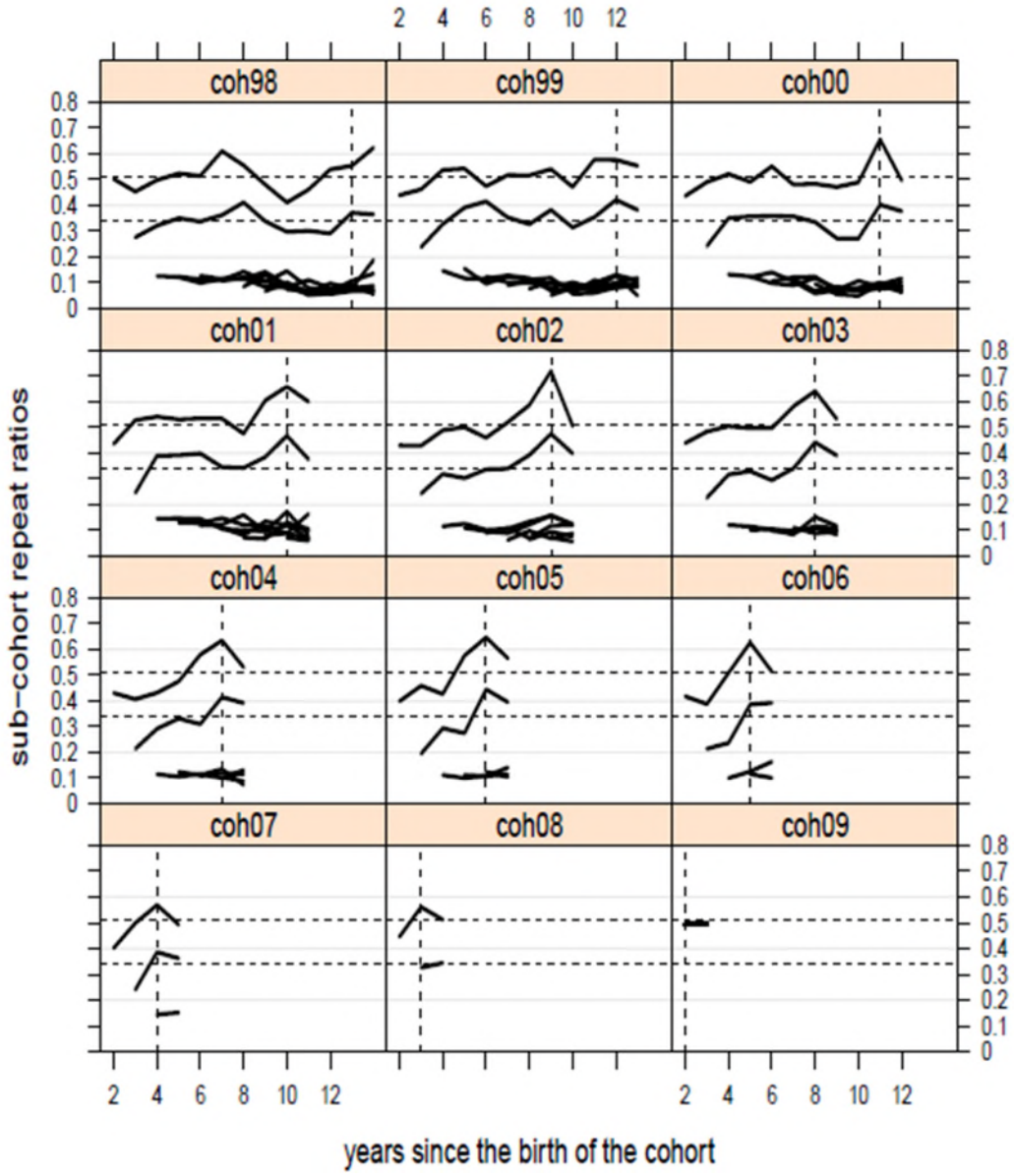


Figure 11: HGF sub-cohort repeat ratios by birth cohort, cohort98 to cohort09, years since the birth of the cohort, ratio



Notes:
1. horizontal dashed line at mean of the first repeat ratio (0.51) and second repeat ratio (0.34), see Figure 5
2. vertical dashed lines at 2011/2014

ERC
Enterprise Research Centre

Centre Manager
Enterprise Research Centre
Warwick Business School
Coventry CV4 7AL
CentreManager@enterpriseresearch.ac.uk

Centre Manager
Enterprise Research Centre
Aston Business School
Birmingham, B4 7ET
CentreManager@enterpriseresearch.ac.uk