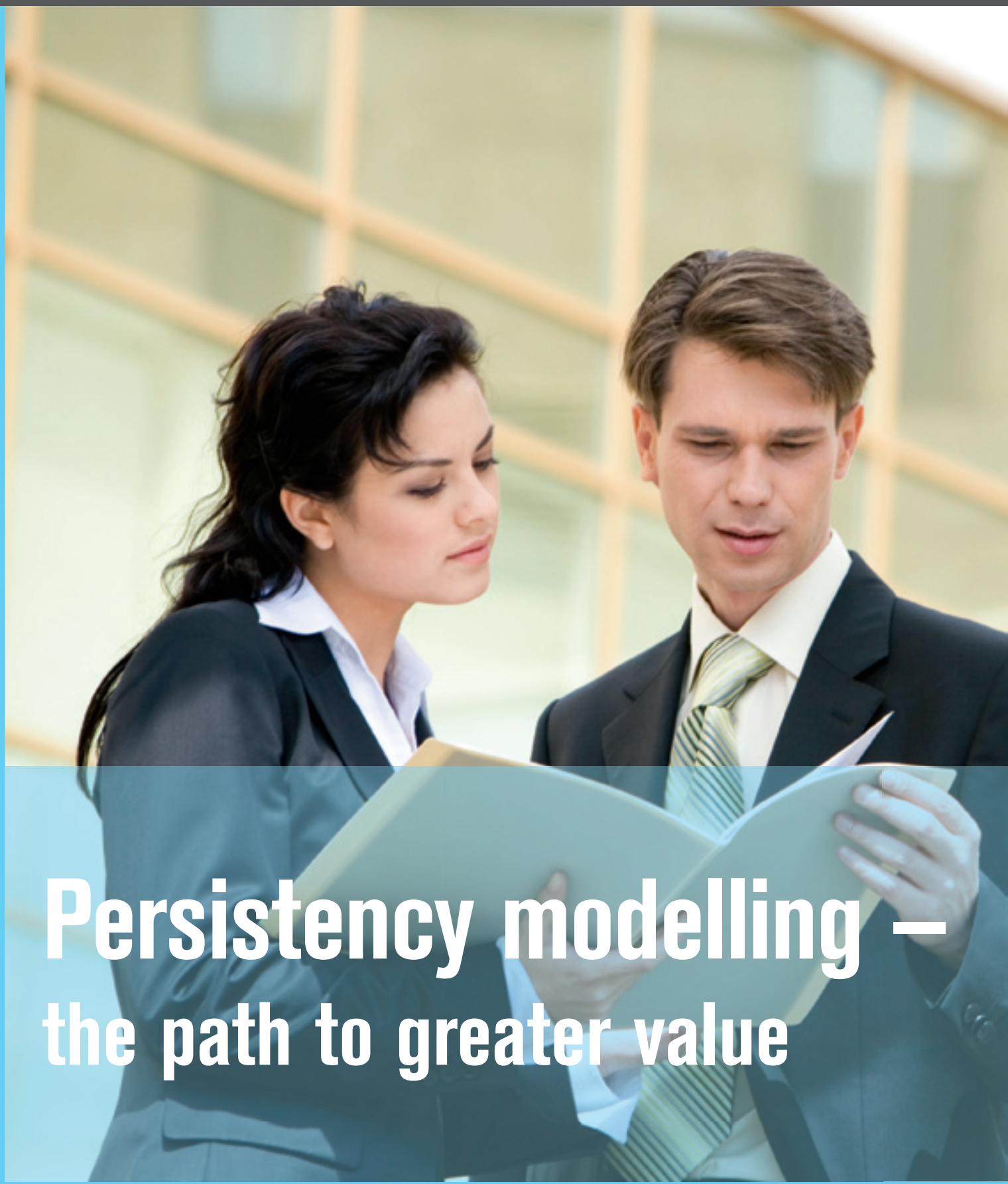


Persistency modelling – the path to greater value

Matthew Edwards shows how multifactor models of life persistency experience can be used to extract greater profitability from your in-force portfolio.

Most life insurers have sought ways to derive increased profits from their existing portfolios. Using the existing policyholder base as a source of further business, whether in the form of life policies or business from other parts of a company's group, has long been an obvious 'easy win' for some companies – although the value of the resulting new business may not be particularly high, and the cross-selling success rate can quickly decline after initial campaigns.

More recently, many insurers have been investigating the potential to increase the profitability of the in-force portfolio by acting to improve policyholder persistency. There are various ways of achieving this aim, but a critical first step is understanding the drivers of policyholder behaviour.

Multi-factor models

Firms need to identify what these drivers are, and quantify their observed influence on withdrawal rates. Historically, companies have analysed their policy surrender experience using simple one-dimensional techniques whereby only the effect of policy duration would be seen, generally conducting separate analyses by product class. Such techniques provide little useful information.

A better approach is to use multi-factor techniques such as generalised linear models (GLMs) – the method that has been used for many years to identify and quantify the many drivers of non-life personal lines claims experience. Allowing automatically for data correlations, taking account of non-linearity and making full use of the underlying probabilistic processes, GLMs allow users to identify and quantify all available factors quickly and robustly. The inset box (opposite) provides a simplified overview of the underlying mathematics, and also notes the advantages GLMs possess compared with other analytic methods that have proliferated in the financial services sector.

With such models, analysts can quickly derive highly predictive multifactor models which allow for the influence on lapse and surrender behaviour of such factors as:

- product class
- duration
- benefit amount
- distribution channel (or specific source of business)
- policyholder age/sex
- socioeconomic characteristics.

Applications

It should be apparent that having an accurate quantification of precisely how such factors affect surrender rates will be of immediate use in financial reporting and product pricing (especially as many products now involve long payback periods, making product profitability particularly sensitive to persistency). More technical readers may also be

GLMs: the broad mechanics

For the purposes of this article, it is enough to be aware that a GLM can model any observed event (for instance, motor third-party claim frequency or annuitant deaths) as a function of various factors (for instance, age, sex and occupation), typically in the following way:

A multiplicative relationship is modelled: for example, we can have:

Modelled quantity = Base level for observed population × Factor 1 (based on duration) × Factor 2 (based on benefit amount) × Factor 3 (based on occupation)

where the modelled quantity would be, in this context, the probability of policyholder surrender.

It is important to note that, despite the word 'linear' in the model, there is no linearity (or other parametric) constraint within any factor, and that non-linearity effects between factors can be allowed for easily via 'interactions'.

For further information, we refer readers to our publication 'A practitioners guide to GLMs', available at watsonwyatt.com/glm

GLMs: advantages over other methods

Compared with the current library of available analytic methods, including one and two-way analysis, linear regression, principal components analysis, survival models, data mining and clustering, GLMs offer a number of advantages. These advantages vary according to the alternative being considered, but generally involve greater predictive power deriving from:

- making full use of knowledge of the underlying probabilistic process (for example, binomial)
- use of logit and other transformations as required
- ability to allow interactions between factors
- transparency – it is simple to see what is going on
- robustness – small changes in the data lead to correspondingly small changes in the results.

interested in how it is possible to compare the historical trends revealed in such analyses with relevant investment market or fund yield trends, potentially revealing relationships which can be used to improve the company's understanding of dynamic policyholder behaviour.

However, the most direct way of using such analyses to improve profitability is likely to be from 'scoring' the in-force portfolio. This refers to a mechanism whereby the results from the multifactor persistency analysis are aggregated into some form of single-figure score – essentially, a statistic which will represent, for every policy or policyholder, their propensity to lapse or surrender. These scores can then be used as the basis for intelligent, targeted customer retention campaigns, whereby the insurer seeks ways to communicate to high-score policyholders (in other words, those most at risk of surrendering) the merits of their policies with a view to pre-empting surrender decisions.

Having set out the basic idea, we illustrate some of the concepts by way of a recent case study.

Case study

The portfolio in question was a portfolio of single premium with-profits contracts of a large European bancassurer. The investigation covered a seventeen-year period, with around 280,000 lapses/surrenders and risk factor information on:

- product
- year of event (equal to year of exposure)
- duration
- age
- sex.

After conducting data cleaning and checks and running simple one-way and two-way analyses of the data to provide some basic familiarity with the portfolio and its history, we were ready to set up initial GLMs combining all available factors. Figures 1 and 2 opposite show the results derived for the two factors of policyholder age and policy duration. Both of these show interesting effects, especially when we bear in mind that the effects shown are for those factors, but taking account simultaneously of the other factor effects in the model (calendar year of event and product group).

Policyholder age

The policyholder age result, as seen in Figure 1, shows the effect of age analysed with age split into four broad levels and a level for corporate clients (in the market in question, many policies are held by small firms as part of their employee retirement benefits planning).

The age 60–79 group was the reference point for this factor, and compared with them, it was clear that younger policyholders had a significantly higher probability of surrender. This probability increased as age decreased: all other things being equal, a policyholder aged below 40 would be 16 per cent more likely to surrender than a 60–79 year old policyholder. The 80+ group also seemed to show a high surrender propensity, but this was a misleading result relating to a data problem (with deaths being classified in the data as surrenders). This age result was intuitively very plausible: we might reasonably expect younger policyholders to be less 'sticky' than older policyholders, whether through greater openness to alternative investments or a greater chance of suddenly needing to realise money.

Particularly striking here was the result for corporate policyholders, the group showing the highest surrender behaviour. This was to be expected: most of these policyholders would dedicate some time every year to assessing their most recent bonus and comparing it against competitor products, quickly surrendering policies which disappointed.

> … the GLM was telling us the 'real effect', to the extent that is ever possible to know the underlying reality behind policy holders' decisions.

Figure 1 | Effect of policyholder age on lapse/surrender

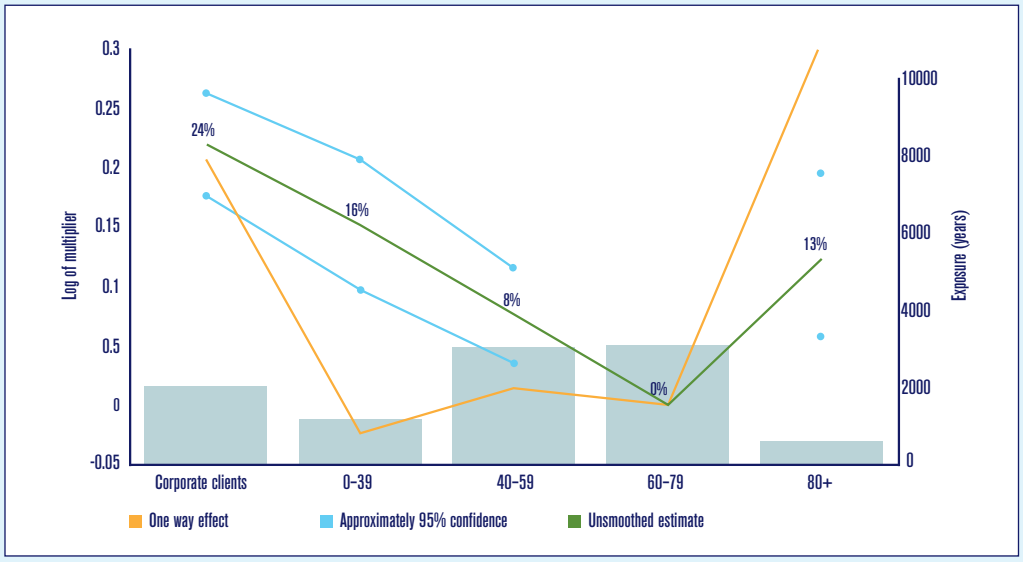


Figure 2 | Effect of policyholder duration on lapse/surrender

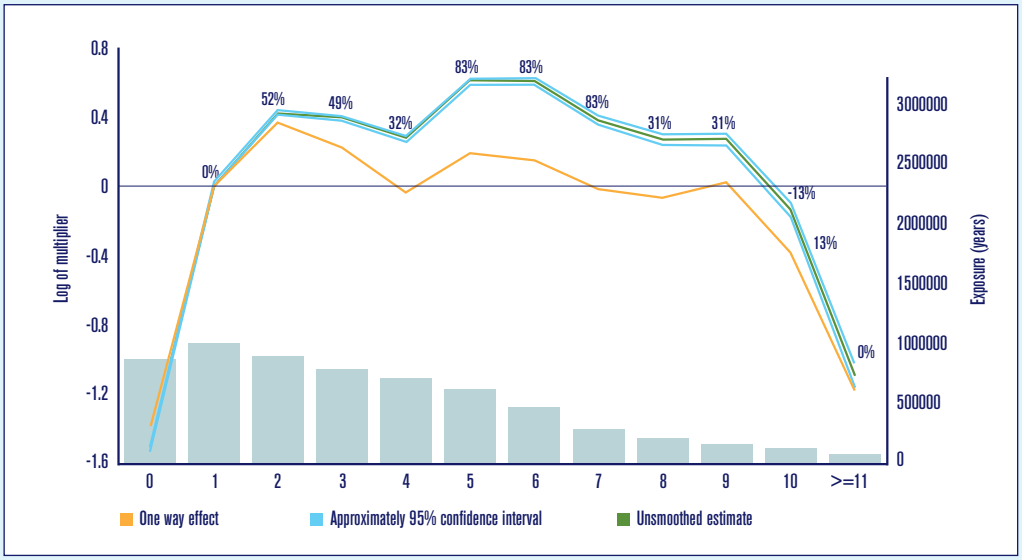
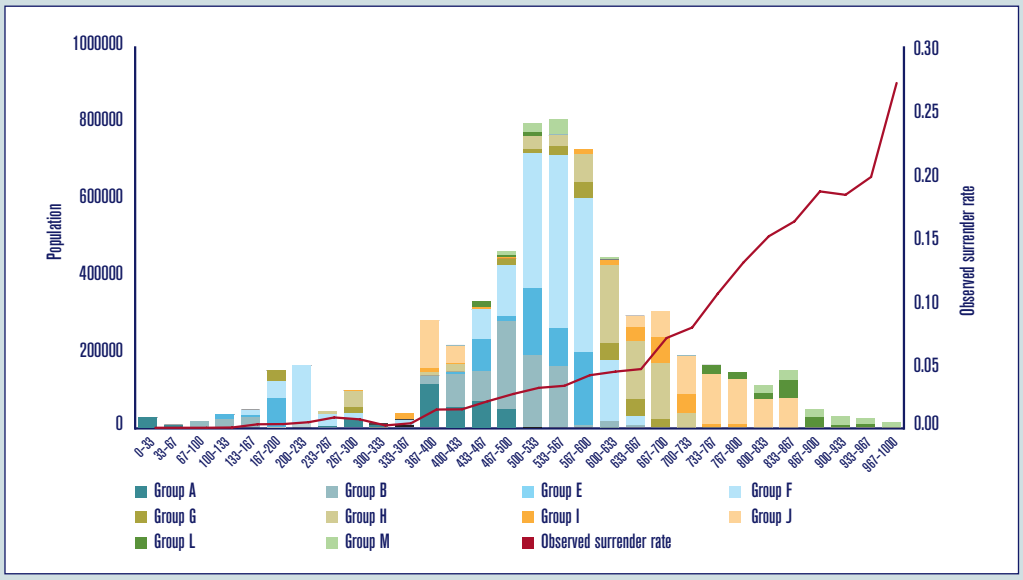


Figure 3 | Persistency scoring – broken down by product subgroup



Policy duration

As would be expected, policy duration shown in Figure 2, had a very strong effect, where we quantified the factor effect as the total multiplicative movement from one extreme to another: in this case the range was 7.9, from the lowest at year zero to the highest at year five. The peaks at year two followed by year five were consistent with the policy conditions applying to most of the products, with substantial surrender penalties applying in the first two years followed by some continuing penalty until the fifth policy anniversary. So what value is the GLM adding, if it tells us what we might already know?

What was happening here was that the GLM was telling us the 'real effect', to the extent that is ever possible to know the underlying reality behind policyholders' decisions. If we compare the green GLM line with the yellow line representing what is seen from the traditional one-way analysis, it is clear that there was a substantial difference in shape after the second year. An insurer pricing long payback period products using a lapse assumption based on the results of a traditional analysis is likely to be significantly mis-pricing.

Other factor effects

Space constraints preclude discussing other factor effects in similar detail. It is worth noting, however, that every factor analysed can lead to some interesting insights. The sex factor, for instance, told us that there was no significant difference in lapse/surrender behaviour between male and female policyholders – and so this factor was removed and the model was re-run.

The time period factor (in these models, calendar year as the models were constructed around yearly rather than quarterly or monthly steps) was also very interesting. Although not directly predictive in the same way that other factors are, it is very useful in two respects:

- First, by absorbing what historical trends or blips there may have been, its presence in the analysis increased the predictiveness of the other factors.
- Secondly, it may be possible to establish relationships between the observed time period lapse/surrender effect and relevant investment market movements. These relationships can then be used predictively: in effect they give us a model of dynamic policyholder behaviour, allowing us to infer a robust surrender assumption for any future economic scenario. This is particularly useful for capital modelling.

Scoring

The results in respect of the many different factors may be useful to actuaries and pricing analysts but will be of little interest to other parts of the company. This is the point where 'scoring', already outlined above, can be done. Essentially we take the model we regard as most predictive after perhaps several iterations of modelling, and then for every policyholder, convert the total lapse/surrender effect implied by the model for that policyholder's combination of circumstances into a one-figure score.

For our case study, a graph of the scores is shown in Figure 3.

The height of each bar shows the number of policyholders belonging in each score band. Colouring the graph according to any known factor (in this example, product group) then helps to see why the 'high-scoring' policyholders (that is those most at risk of lapse or surrender) were

> **An insurer pricing long payback period products using a lapse assumption based on the results of a traditional analysis is likely to be significantly mis-pricing.**

high-scoring when all factor effects were taken into account. The green line superimposed on the same graph shows the observed surrender experience of policyholders in the scoring bands.

These scores are interesting to look at graphically, since they combine all factor effects while the previous factor effect graphs by definition examined individual factor effects. However, their greatest use is in giving the insurer a dataset of useful information. The insurer can take (say) the 10 per cent of policyholders with the highest scores, perhaps cross-indexing with some value measure (since the concern is to identify those policies which are both a potential source of sizeable future income and at high risk of withdrawal), and then set up a campaign to communicate with these policyholders via targeted letters or telephone scripts.

At this stage, quantitative analysis needs to give way to qualitative analysis. How best to communicate with at-risk policyholders to improve their persistency depends on what the reasons are for the past surrenders of similar policyholders. Finding out those reasons is likely to require some form of customer research survey. The multi-factor model (perhaps via the scoring mechanism) can assist here in indicating the policyholder groups who should be 'interrogated' the most.

Conclusion

A number of insurers have set up retention management responsibilities and started to conduct analyses and retention management exercises of these types. It would therefore seem that there is growing awareness of the potential here for greater profit. Although analyses and initiatives of the type outlined above are neither trivial nor cheap, they can be a profitable way to make the most of your customers – who are, after all, your best 'natural resource'.

For more information contact:

Matthew Edwards
+44 (0) 1737 274300
matthew.edwards@watsonwyatt.com

locations

ASIA-PACIFIC ▪ Bangkok ▪ Beijing ▪ Bengaluru ▪ Delhi
Guangzhou ▪ Hanoi ▪ Ho Chi Minh City ▪ Hong Kong ▪ Jakarta
Kolkata ▪ Kuala Lumpur ▪ Manila ▪ Melbourne ▪ Mumbai
Seoul ▪ Shanghai ▪ Shenzhen ▪ Singapore ▪ Sydney ▪ Taipei
Tokyo ▪ Wuhan

EUROPE ▪ Amsterdam ▪ Apeldoorn ▪ Birmingham ▪ Bristol
Brussels ▪ Dublin ▪ Düsseldorf ▪ Edinburgh ▪ Eindhoven
Frankfurt ▪ Lausanne ▪ Leeds ▪ Lisbon ▪ London ▪ Madrid
Manchester ▪ Milan ▪ Moscow ▪ Munich ▪ Nieuwegein
Paris ▪ Purmerend ▪ Redhill ▪ Reigate ▪ Rome ▪ Rotterdam
Stockholm ▪ Vienna ▪ Welwyn ▪ Wiesbaden ▪ Zürich

LATIN AMERICA ▪ Bogotá ▪ Buenos Aires ▪ Mexico City
Montevideo ▪ San Juan ▪ Santiago ▪ São Paulo

MIDDLE EAST ▪ Dubai

NORTH AMERICA ▪ Atlanta ▪ Berwyn, PA ▪ Boston
Calgary ▪ Charlotte ▪ Chicago ▪ Cincinnati ▪ Cleveland
Columbus ▪ Dallas ▪ Denver ▪ Detroit ▪ Grand Rapids
Hartford, CT ▪ Herndon, VA ▪ Honolulu ▪ Houston ▪ Irvine
Kitchener-Waterloo ▪ Los Angeles ▪ Madison, WI ▪ Memphis
Miami ▪ Minneapolis ▪ Montréal ▪ New York ▪ Paramus, NJ
Philadelphia ▪ Phoenix ▪ Portland ▪ Rochelle Park, NJ ▪ St Louis
San Diego ▪ San Francisco ▪ Santa Clara ▪ Seattle ▪ Stamford
Tampa ▪ Toronto ▪ Vancouver ▪ Washington, DC

watsonwyatt.com

21 Tothill Street, Westminster, London, SW1H 9LL UK
Telephone +44 (0) 20 7222 8033
Fax +44 (0) 20 7222 9182

 August 2009 Ref: EU-2009-12568