The Trimmed Bootstrap: Evidence from international markets and inflation adjusted returns

Brett Doran and Graham Bornholt

Abstract

The trimmed bootstrap introduced in Doran and Bornholt (2017) provides pension finance researchers inside the US with a tool that was shown to offer heightened accuracy over pre-existing empirical simulation models. While researchers in the US can benefit from this addition to the literature, it leaves a question regarding the model's functionality in other markets. This study examines this question by applying the methodology derived in Doran and Bornholt (2017) on a global scale including both nominal and real returns.

The results suggest the trimmed bootstrap outperforms existing stochastic simulation techniques in a variety of markets. The study corroborates the findings of Doran and Bornholt (2017) and suggests researchers who employ techniques which fail to incorporate time dependency in their methodology do so at the detriment of defensible results. Further, through a novel forward-looking test, we show evidence that the trimmed bootstrap should become the default methodology for researchers regardless of the market in which they reside.

JEL: Simulation, Bootstrap, Pensions, Retirement, Forecasting returns.

Introduction

Pension finance researchers frequently employ stochastic parametric and nonparametric techniques to simulate retirement outcomes. In a previous paper, we introduced the trimmed bootstrap, which was designed to control for unrealistic issues which arise in pension finance research when utilising nonparametric simulation, namely performance risk, and sequencing risk (Doran & Bornholt, 2017). Through a series of out-of-sample tests, we derived a new technique which controls for biases that were argued to be too frequent in techniques that assume no time dependency in returns, such as the Efron (1979) bootstrap. The trimmed bootstrap factors in both the short-to-midterm and long-term return reversals, and through doing so, was shown to be a more-accurate technique over existing methods. This heightened accuracy is an important contribution to the literature, as it enables researchers to have better informed models that optimise variables such as asset allocation and contribution/withdrawal profiles. In turn, such optimisations can play a sizeable role in public policy decisions.

However, our previous work was limited to a solely US analysis. The technique was not validated in alternative markets. The reliability of the model in countries which did not achieve the same level of stock market performance as that of the US over the data period is unknown. This study closes this gap by providing the results through examining the range of economies available in the updated Dimson, Marsh and Staunton (DMS) (2002) database.

The contributions of this study are twofold. Firstly, it is the first international study into nonparametric bootstrap techniques on the 21 countries in the DMS database. Hence, the results offer insight into the validity of using nonparametric testing procedures on each country. Secondly, the analysis shows the reliability of the trimmed bootstrap internationally. The results provide global pension finance researchers with the necessary knowledge to assess which technique is optimal in certain markets.

The out-of-sample results show that for the majority of countries listed in the DMS database, the trimmed bootstrap is the superior nonparametric technique. However, in our previous paper we argued that the trimmed bootstrap becomes better informed with a larger dataset. Hence, looking beyond the metrics we presented in that paper, we present a forward-looking estimator of the trimmed bootstrap in each market. Essentially, the trimmed bootstrap has the ability to learn from additional data, much more than alternative methods which do not account for mid- and long-term return reversals. The forward-looking analysis allows more data into the analysis by utilising the entire dataset. Hence, this forward-looking analysis gives the trimmed bootstrap the ability to produce a more robust estimate. The results of this analysis suggest that the trimmed bootstrap is the superior methodology in all countries in the DMS database, unless future market returns are expected to be larger than those experienced in the historical sample. The analysis concludes that researchers would be wise to cease utilising a simple Efron (1979) bootstrap to simulate long-term returns, over alternatives such as the trimmed bootstrap, regardless of the market in which they reside. Evidence suggests that the trimmed bootstrap should become the default choice for pension finance researchers who currently employ other bootstrap techniques.

Literature review and background

Global pension schemes face a range of contemporary risks. In particular, the majority of countries in the western world are experiencing an aging population (United Nations Department of Economic and Social Affairs, 2015). Compounding this issue is the fact that generations are living longer. Using data from the World Health Organisation, Mathers, Stevens, Boerma, White, and Tobias (2015) show that in just the past two decades, high income countries have experienced an increase in life expectancy at age 60 to the value of 3.2 years for men, and 2.8 years for women. An increase of approximately three years in life expectancy across the population requires significantly more savings to self-fund the extra three years of consumption in retirement. This trend may require public policy changes to ensure that older generations have the capacity to self-fund their retirement.

Countries need adequate retirement policies to ensure that the impacts of ageing populations and greater longevity do not impact the balance sheets of governments. Looking to country-specific examples, the pension system in the United Kingdom has historically been centred around annuities in retirement, a system similar to the US (Sweeting, 2009). This process creates a private-pension scheme whereby public accounts are not as heavily burdened by future retirement liabilities from those who depleted their retirement savings too soon. The system ensures those with greater longevity are partly funded by those with shorter longevity. An annuity system divides the pre- and post-retirement research into two categories. The pre-retirement research looks into factors such as contribution rates and asset allocation designs to maximise the annuity equivalent value at retirement date for the plan member, while the post-retirement research examines the optimal annuitisation strategy for different life expectancies (see, for example, Pfau and Kitces, 2014). These markets produce research which aims to ensure greater wealth accumulation at retirement date, so that the post-retirement annuity strategy can be purchased to self-fund retirement.

In the southern hemisphere, Australia, which has been quoted as a leading player in pension reform with its mandatory defined contribution schemes (Rashbrooke, 2009), is framing the debate for both pre-and-post retirement phases. Australia is continuing to push reform with the implementation of the country's MySuper reforms. These policy changes are aimed at tailoring asset allocation strategies to the individual, which will maximise the longevity of retirement accounts (Bateman & Kingston, 2012; Howard, 2012). Contrary to research in an annuity system, maximising the longevity moves the debate from a 'pre-retirement' followed by a 'post-retirement' discussion, to a 'pre-and-post' retirement discussion.

Following the policy style of Australia, New Zealand implemented the KiwiSaver in 2007, which has since seen contribution rates increase in retirement savings accounts. The policy places a focus on the optimal asset allocation strategy for particular member profiles (MacDonald, Bianachi and Drew, 2012). Furthermore, in the past decade, both New Zealand and Italy have significantly increased policy relating to pension coverage and safety net benefits that aim to decrease poverty in the aging generation (OECD, 2013).

Figure 1 provides a snapshot of the summarised pension policy efficacy by country, as determined by the Organisation for Economic Co-operation and Development (OECD). Several categories are defined by the OECD: adequate coverage of the population; adequacy of the system itself in providing a sustainable retirement income; the sustainability of the system between accumulation and decumulation generationally (that is, an aging population will not create unsustainability in the system); incentives to work; and administrative efficiency. In terms of top-rankings, only the United Kingdom achieves a pass in every category, and Australia is the only country to achieve a pass in six out of the seven categories. In contrast, the US system is poorly ranked, achieving pass marks in only three out of seven. However, these three categories (that is, coverage, adequacy and sustainability) are considered the most important.

As we discussed in Doran and Bornholt (2017), empirical nonparametric stochastic simulation is a powerful tool used by pension finance researchers. Two approaches are ubiquitous in the pension finance literature: utility theory and stochastic estimation.[1] We focus on the latter approach. The original bootstrap methodology, coined by Efron (1979),

[1] Utility theory is a common approach used in US pension finance research. The optimisation of asset allocations using relative risk aversion utility functions provided the basis for contemporary asset allocations. These de-risk as retirement draws near, as the human capital bond-type annuity decreases with less remaining working years (Bodie et al., 1992; Cairns, Blake, & Dowd, 2006; Kingston & Thorp, 2005).

involved the process of randomly sampling with replacement for the purpose of statistical inference. Pension finance literature commonly employs this simulation technique to investigate a range of metrics pertinent to the discipline. With a plethora of empirical nonparametric simulation techniques available in the pension finance literature, researchers lacked consensus on the optimal technique.

In our previous paper, we argued that this lack of consensus was due to the absence of a robust study examining the relevant accuracy of each technique. The study examined 11 existing nonparametric bootstrap techniques to test their individual accuracy in estimating retirement outcomes. These techniques included the Efron (1979) simple bootstrap, a series of varying block sizes from Künsch's (1989) moving block bootstrap (also referred to as the rolling block bootstrap), and the Politis and Romano (1994) stationary bootstrap. We concluded through a series of out-of-sample tests that contemporary techniques can be wildly inaccurate. This finding led to the development of a new technique, the trimmed bootstrap, which offered heightened accuracy over existing stochastic nonparametric simulation techniques.

While the trimmed bootstrap was found to be a superior technique, the analysis to date was limited to US nominal returns data. This naturally raises the question of whether the superiority of the trimmed bootstrap extends to other countries. Can we assure pension finance researchers outside of the US of the appropriateness of the trimmed bootstrap methodology for their markets? Related to this issue is the fact that, over the data period, the US was an extremely prosperous nation and became the largest economy in the world. Other countries' distributions of long-term returns may be quite dissimilar to those of US in ways that affect the relative performance of different bootstrap methods. For example, many other countries' stock markets had less stellar performances than that of the US. Also, perhaps the serial dependencies of the US stock market differ in some materially-important way to those that may exist in other countries. Thus, we need to investigate the effectiveness of trimmed bootstrap in other markets.

Figure 1: Overview of pension reform measures in 34 OECD countries 2009-2013

	Coverage	Adequacy	Sustainability	Work incentives	Administrative efficiency	Diversification/ security	Other
Australia	x	x	x	x	x		x
Austria	x	x	x				x
Belgium				x			
Canada	x		x	x		x	x
Chile	x	x			x	x	x
Czech Republic			x	x		x	
Denmark				x	x		
Estonia		x	x	x	x	x	
Finland	x	x	x	x		x	
France	x	x	x	x			x
Germany		x	x	x			
Greece		x	x	x	x		
Hungary		x	x	x		x	x
Iceland							x
Ireland	x		x	x		x	x
Israel	x	x				x	
Italy		x	x	x	x		
Japan	x	x	x		x		x
Korea	x		x		x		
Luxembourg	x		x	x			
Mexico		x			x	x	
Netherlands						x	
New Zealand		x	x				x
Norway		x	x	x			
Poland	x		x	x		x	
Portugal	x	x	x	x		x	
Slovak Republic			x		x	x	
Slovenia			x	x			
Slovenia	x	x	x	x	x	x	x
Spain		x	x	x			
Sweden		x	x	x	x	x	
Switzerland			x			x	
Turkey				x		x	x
United Kingdom	x	x	x	x	x	x	x
United States	x	x	x				

Source: OECD (2013)

The trimmed bootstrap methodology trims the randomly generated long-term return paths to produce a distribution of simulated return paths loosely calibrated to the historical distribution of long-term paths. These historical distributions can be expected to vary significantly across countries (perhaps because some countries have experienced major shifts in their economies due to inflationary or political issues). Such variations constitute an added reason to check the robustness of the US-based findings.

Whereas our previous study was based solely on nominal returns, both nominal and real returns are employed in pension finance research. The debate between employing nominal or real returns when forecasting (or estimating) outcomes is a ubiquitous debate in finance. Bruno and Chincarini (2010) show that real returns can be a useful tool for markets which have experienced hyperinflationary periods. They also argued that any inflation-adjusted optimisation (i.e. real returns) also produces nominal return optimisation. This finding means that employing real returns allows the researcher to utilise a more stable distribution of returns not impacted by hyper-inflationary periods, with no downside to accuracy or practical application. Many pension researchers prefer using real returns for modelling policy alternatives. Using real returns in cross-country research has intuitive appeal because countries differ significantly in their inflation histories in the sample. Hence, it is hypothesised that understanding the outcomes in a real-return environment is of greater importance than understanding those in a nominal return environment. However, while the nominal versus real return debate is evident in pension finance research, the aim of this study aim is not to investigate which provides the superior research approach. Instead, we replicate the analysis using both nominal and real return data to satisfy both sides of the debate.

Data and methodology

The data is sourced from the updated DMS database and covers 114 years of annual data from 1900 to 2013 (Dimson et al., 2002). The DMS database offers data across 21 countries, which allows us to examine a further 20 countries beyond the US. For robustness, both nominal and real returns are included in the results. Incorporating the real return analysis is important, not only for the reasons given in the previous section, but also because of idiosyncrasies with global data. In particular, it is hypothesised from the summary statistics in Table 1 that for certain countries, such as Germany and Austria, which experienced extreme inflationary periods in the test sample, any simulation methodology conducted on such extreme nominal outcomes will offer little guidance for the expected future outcomes. Using real returns removes the effect of such hyperinflationary periods from the data and should produce more-realistic simulated outcomes.

We use two approaches for comparing the suitability of the trimmed bootstrap method with other bootstrap techniques in individual countries. The first approach utilises the procedures from Doran and Bornholt (2017) and involves splitting the full sample into two parts: an in-sample component and a 40-year out-of-sample component. The second approach, a forward-looking approach, uses the whole sample to simulate future return paths and hence future wealth outcomes. For each country, the analysis in this second approach involves determining how extreme the next 40-year wealth outcome would have to be for the trimmed mean to underperform any of the alternative bootstraps. Armed with this knowledge, the researcher for a particular country can then see that selecting a bootstrap method other than the trimmed bootstrap is an implicit judgement about how extreme the next 40-year outcome is likely to be. These two approaches are described next.

Out-of-sample testing

While in our previous paper we examined four out-of-sample periods, in this study, only the most recent out-of-sample period is used. This change is due to our previous findings which highlighted that the trimmed bootstrap continues to improve the longer the available data upon which the trimming is based. Hence, we only examine the final out-of-sample period to allow the most data into the model. That is, we use the data from 1900 to 1973 to simulate 1974–2013 for the 21 countries in the DMS database.

The tests statistics are the same as Doran and Bornholt (2017). Specifically, both the mean absolute error (MAE) and the root mean squared error (RMSE) of the alternative bootstrap methods in relation to the historical outcome are examined. These variables provide a basis for examining the accuracy of the simulation techniques in the study. The MAE and RMSE statistics for each simulation technique are compared to the statistics derived for the trimmed bootstrap. The significance in the MAE methodology is determined using a paired *t*-test for the differences in the MAEs between the trimmed bootstrap and the alternative techniques. The significance testing in the RMSE methodology employs a nonparametric Fisher randomization test of the RMSE differences for 10,000 permutations (Noreen, 1989).

Forward-looking comparisons

In addition to this significance testing, we employ a forward-looking approach to examining the optimal technique to apply for each market. A fundamental issue with the significance tests is that the trimming is based on in-sample returns which end in 1973. This means, for example, that when we are undertaking the trimming of 40-year returns so as to calibrate to the historical distribution of rolling 40-year returns, we are basing this calibration on just 35 rolling 40-year returns, whereas the full sample has 75 such returns. In other words, when testing, we are excluding 40 years, or 40 data points. This is a sizeable amount to exclude, considering out dataset is 114 data points per country.

The novel forward-looking approach uses the full sample in a forward-looking MAE analysis. By examining the simulated wealth outcomes produced by each bootstrap technique when simulating using the entire dataset, we can calculate the MAE that would be produced by each competing method for a given hypothetical terminal wealth outcome achieved over the next 40 years. For example, if you assume that a sum of 500,000 currency units will be the retirement result in the next 40 years in a particular market, then we can determine which bootstrap technique would have been the best choice, given this hypothetical future outcome (i.e., the bootstrap with the lowest MAE). The analysis derives a ‘range’ of hypothetical outcomes where a particular technique is optimal. We provide this analysis for the case, using real returns for every 100,000 currency unit block between 0 and 2,000,000, for every country in the dataset. The equivalent internal rates of returns for these hypothetical wealth outcomes are also provided to help standardise the values for the reader.

Commentary on the data

The annualised summary statistics for the 21 countries in the DMS database are shown in Table 1. The nominal data is displayed on the left-hand side (LHS) of the table and the real returns are shown on the right-hand side (RHS). This table offers preliminary insight into some issues around nominal returns. Instances of hyperinflation may significantly impact the estimates derived from simulations using historical data. Such examples are seen in Germany and Austria, where the maximum one-year return was 2.6E+9 per cent and over 1,100 per cent, respectively. Such extremes are not apparent in real returns data. This is corroborated by the real returns on the RHS, which reports maximum one-year returns for Germany and Austria at 155 per cent and 127 per cent respectively.

We compare the trimmed bootstrap with the 11 empirical bootstrap techniques studied in Doran and Bornholt (2017) on the 21 countries listed in Table 1. The 11 empirical bootstrap techniques are made up of the Efron (1979) bootstrap, a series of Künsch (1989) moving block bootstraps (specifically block sizes of 2 to 10), and finally the Politis and Romano (1994) stationary bootstrap. The number of simulated paths in each simulation is held consistent at 1,000, to keep the calculations manageable.[2]

For a complete description of the trimmed bootstrap, please refer to equation [1] in Doran and Bornholt (2017). For clarity, the following summarises the output distribution of the technique. The trimmed bootstrap begins with simulated *Z*-year return paths produced by

[2] There are 756 simulation analyses to undertake (21 countries x 12 simulation methods x 3 wage assumptions). Hence total simulations equate to 756 x 1,000 = 756,000.

the Efron (1979) bootstrap procedure. The trimming process excludes paths that are judged unsuitable. Possible mid-term return reversal is incorporated by trimming out 40-year paths with rolling 10-year compounded returns that are more extreme than observed historically. Secondly, potential long-term return reversals are accommodated by aligning the 40-year returns of simulated outcomes (where $Z = 40$) to history. This is enabled by ensuring no simulated 40-year return is beyond historical extremes, while also requiring that the upper and lower quartiles of the distribution of simulated returns depend on the corresponding upper and lower quartiles of the historical distribution of 40-year returns. As noted in our previous paper, the Efron (1979) bootstrap tends to produce too many simulated long-term return paths that are too extreme. The trimmed bootstrap process was argued to ameliorate this problem and so provides a more accurate set of simulated return paths.

A set of summary statistics from the in-sample data (1900–1973), which are the inputs for the trimmed bootstrap procedure, are displayed in Tables 2 and 3 for nominal and real data, respectively. This study holds Z constant, where $Z = 40$. The contrast between the real and nominal amounts in Tables 2 and 3 further emphasises the issues pertinent to nominal data. The maximum 10-year compounded return for any country in nominal returns was for Germany, with a 1,397.1 per cent return per annum, due to extreme inflationary influences. This figure reduces to 37.7% per annum for real (inflation-adjusted) data.

The data required for the trimmed bootstrap are the minimum and maximum 10-year compounded returns, along with the compounded 40-year minimum, maximum, 25th and 75th percentiles and are given in Table 2. The country list provides a diverse range for testing. In the nominal returns case, 12 of the listed 21 countries had a 10-year period with a negative annualised return. In 40-year returns, no country recorded a negative annual return over the 40 years. Sweden recorded the lowest 40-year nominal return at 2.7% per annum.

In the real returns case, only 2 countries escaped recording a negative per annum return over a 10-year period in the sample (Australia and New Zealand). The real data also produced seven countries which experienced a negative annual return over a 40-year period in the sample (Austria, Belgium, Finland, France, Germany, Italy and Japan). Austria also had both a mean and median rolling 40-year annualised return in negative territory at -0.6% and -1.5% respectively.

These historical performances in Tables 2 and 3 provide the foundations for the trimming procedures based on the 10-year and *Z*-year (40 years) returns. As the trimmed bootstrap has more reliance on historical data, we hypothesise that it may underperform in countries which experienced fundamental shifts in market returns between the in-sample and the out-of-sample period. Possible factors which could create such shifts include impacts of war, abnormal or hyper-inflation, and significant political changes which impact their respective economies. Our understanding from our previous study of the benefits of the trimmed bootstrap method leads us to expect that underperformance of the trimmed approach will occur only if the out-of-sample returns are relatively extreme.

Table 1: 1900–2013 Annual summary statistics for the 21 countries, nominal (LHS) and real (RHS)

	Mean	Median	Standard Deviation	Skewness	Kurtosis	Minimum	Maximum
Australia	0.13	0.14	0.18	-0.04	0.59	-0.40	0.67
Austria	0.29	0.05	1.23	7.16	57.98	-0.59	11.09
Belgium	0.10	0.08	0.25	1.17	4.01	-0.48	1.28
Canada	0.10	0.11	0.17	-0.15	0.04	-0.33	0.52
Denmark	0.11	0.08	0.22	1.65	6.48	-0.48	1.20
Finland	0.17	0.11	0.31	1.46	4.39	-0.51	1.67
France	0.13	0.08	0.25	0.69	0.38	-0.41	0.89
Germany	2.2E+7	0.09	2.4E+8	10.68	114.00	-0.88	2.6E+9
Ireland	0.11	0.06	0.23	0.72	2.49	-0.65	0.87
Italy	0.14	0.12	0.33	1.82	5.61	-0.47	1.60
Japan	0.15	0.09	0.29	1.13	2.16	-0.44	1.21
The Netherlands	0.10	0.09	0.23	1.34	6.12	-0.49	1.30
New Zealand	0.12	0.11	0.20	1.81	9.48	-0.49	1.19
Norway	0.11	0.09	0.28	2.37	11.67	-0.53	1.79
Portugal	0.17	0.09	0.39	2.18	6.90	-0.70	1.87
South Africa	0.15	0.12	0.23	1.02	2.22	-0.30	1.08
Spain	0.12	0.08	0.23	0.96	2.77	-0.37	1.16
Sweden	0.12	0.12	0.22	0.25	0.25	-0.38	0.70
Switzerland	0.08	0.07	0.19	0.32	0.50	-0.34	0.61
United Kingdom	0.11	0.11	0.22	2.05	12.88	-0.49	1.46
United States	0.12	0.14	0.20	-0.39	-0.09	-0.44	0.57

	Mean	Median	Standard Deviation	Skewness	Kurtosis	Minimum	Maximum
Australia	0.09	0.12	0.18	-0.38	0.23	-0.43	0.51
Austria	0.05	0.00	0.30	1.12	3.14	-0.60	1.27
Belgium	0.05	0.03	0.24	0.54	1.69	-0.49	1.05
Canada	0.07	0.07	0.17	0.00	-0.08	-0.34	0.55
Denmark	0.07	0.05	0.21	1.39	5.49	-0.49	1.08
Finland	0.09	0.07	0.30	1.24	5.05	-0.61	1.62
France	0.06	0.04	0.23	0.33	-0.26	-0.41	0.66
Germany	0.08	0.07	0.32	1.40	5.64	-0.91	1.55
Ireland	0.07	0.03	0.23	0.26	0.99	-0.65	0.68
Italy	0.06	0.05	0.29	0.72	2.43	-0.73	1.21
Japan	0.09	0.06	0.30	0.49	2.27	-0.86	1.21
The Netherlands	0.07	0.07	0.22	0.83	2.97	-0.50	1.02
New Zealand	0.08	0.08	0.20	1.18	6.87	-0.55	1.05
Norway	0.07	0.05	0.27	2.09	10.55	-0.54	1.67
Portugal	0.09	0.05	0.35	1.61	5.21	-0.77	1.52
South Africa	0.10	0.08	0.22	0.86	2.50	-0.52	1.03
Spain	0.06	0.03	0.22	0.70	1.80	-0.43	0.99
Sweden	0.08	0.08	0.21	0.11	0.18	-0.43	0.68
Switzerland	0.06	0.05	0.20	0.34	0.36	-0.38	0.59
United Kingdom	0.07	0.07	0.20	0.59	3.59	-0.57	0.97
United States	0.08	0.11	0.20	-0.22	-0.28	-0.38	0.56

Table 2: Summary statistics for the trimming process – nominal data 1900–1973

	10-year summary statistics				40-year summary statistics					
	Minimum	Maximum	Mean	Median	Minimum	Maximum	Mean	Median	25th percentile	75th percentile
Australia	0.050	0.174	0.116	0.112	0.101	0.125	0.113	0.114	0.111	0.116
Austria	-0.103	1.132	0.220	0.100	0.086	0.282	0.200	0.237	0.113	0.251
Belgium	-0.083	0.234	0.072	0.064	0.038	0.096	0.073	0.075	0.067	0.084
Canada	-0.002	0.181	0.090	0.074	0.059	0.120	0.087	0.087	0.072	0.098
Denmark	-0.003	0.163	0.074	0.067	0.054	0.099	0.070	0.069	0.064	0.076
Finland	0.053	0.201	0.128	0.131	0.107	0.156	0.132	0.134	0.123	0.141
France	-0.024	0.262	0.111	0.089	0.069	0.159	0.122	0.117	0.110	0.138
Germany	-0.149	13.971	1.969	0.070	0.049	1.041	0.701	0.978	0.086	1.011
Ireland	-0.001	0.213	0.067	0.060	0.032	0.100	0.062	0.054	0.050	0.078
Italy	-0.030	0.370	0.128	0.092	0.076	0.193	0.146	0.154	0.133	0.166
Japan	-0.013	0.449	0.147	0.134	0.062	0.196	0.137	0.140	0.103	0.166
The Netherlands	-0.027	0.183	0.068	0.069	0.033	0.115	0.064	0.057	0.044	0.086
New Zealand	0.056	0.128	0.090	0.091	0.077	0.094	0.085	0.085	0.083	0.088
Norway	-0.055	0.166	0.059	0.066	0.043	0.095	0.061	0.060	0.055	0.066
Portugal	0.011	0.259	0.114	0.100	0.081	0.152	0.117	0.120	0.101	0.130
South Africa	0.007	0.210	0.096	0.090	0.081	0.120	0.099	0.100	0.091	0.105
Spain	-0.018	0.178	0.083	0.077	0.049	0.119	0.078	0.078	0.062	0.088
Sweden	-0.071	0.163	0.064	0.077	0.027	0.110	0.061	0.051	0.038	0.084
Switzerland	-0.042	0.175	0.061	0.060	0.032	0.100	0.058	0.055	0.040	0.074
United Kingdom	0.016	0.184	0.080	0.068	0.046	0.112	0.074	0.078	0.058	0.087
United States	-0.019	0.190	0.091	0.090	0.054	0.125	0.087	0.092	0.069	0.100

Table 3: Summary statistics for the trimming process – real data 1900–1973

	10-year summary statistics				40-year summary statistics					
	Minimum	Maximum	Mean	Median	Minimum	Maximum	Mean	Median	25th percentile	75th percentile
Australia	0.006	0.171	0.085	0.086	0.055	0.103	0.085	0.085	0.078	0.092
Austria	-0.227	0.226	-0.008	-0.009	-0.073	0.056	-0.006	-0.015	-0.036	0.038
Belgium	-0.130	0.204	0.009	0.008	-0.022	0.043	0.005	0.004	-0.011	0.021
Canada	-0.025	0.167	0.066	0.060	0.041	0.091	0.067	0.067	0.054	0.082
Denmark	-0.020	0.093	0.039	0.040	0.025	0.055	0.039	0.037	0.033	0.043
Finland	-0.149	0.185	0.035	0.056	-0.005	0.068	0.036	0.038	0.018	0.050
France	-0.152	0.191	0.016	0.007	-0.024	0.040	0.013	0.015	0.006	0.022
Germany	-0.193	0.377	0.023	0.011	-0.070	0.067	0.015	0.002	-0.009	0.052
Ireland	-0.094	0.174	0.035	0.025	0.011	0.056	0.032	0.029	0.020	0.047
Italy	-0.120	0.213	0.026	0.037	-0.017	0.072	0.025	0.031	0.009	0.037
Japan	-0.309	0.311	0.048	0.085	-0.043	0.079	0.016	0.021	-0.008	0.029
The Netherlands	-0.053	0.159	0.040	0.031	0.011	0.069	0.040	0.037	0.026	0.056
New Zealand	0.013	0.122	0.064	0.064	0.044	0.075	0.064	0.063	0.059	0.071
Norway	-0.104	0.114	0.027	0.028	0.017	0.058	0.033	0.033	0.025	0.038
Portugal	-0.102	0.202	0.045	0.045	0.019	0.099	0.049	0.045	0.031	0.063
South Africa	-0.051	0.184	0.072	0.067	0.064	0.090	0.076	0.076	0.070	0.081
Spain	-0.068	0.143	0.031	0.030	0.005	0.039	0.022	0.022	0.017	0.028
Sweden	-0.106	0.114	0.036	0.054	0.010	0.068	0.036	0.029	0.015	0.060
Switzerland	-0.111	0.162	0.039	0.040	0.011	0.092	0.040	0.038	0.018	0.055
United Kingdom	-0.056	0.144	0.052	0.045	0.025	0.080	0.051	0.056	0.036	0.064
United States	-0.039	0.168	0.067	0.061	0.038	0.095	0.066	0.068	0.050	0.083

Retirement assumptions

As this study will examine both nominal and real returns, a variation between the assumed wage growth profile is required. Our previous paper studied nominal returns employing a 4 per cent wage growth formula. This study holds this assumption constant for the nominal return analysis. However, the real return analysis incorporates two variations (our previous nominal growth assumption included an inflationary factor). In an OECD working paper on pension outcomes, Antolin et al. (2010) factor in a 2 per cent per annum wage productivity growth rate to stochastic inflation. Thus, the real analysis in this study imposes this 2 per cent annual real wage growth assumption. We also offer a secondary assumption of 0 per cent for robustness. This second assumption allows for the possibility that wages will not grow faster than inflation in the future. At the time of writing, the current climate within financial markets is suggesting markedly lower wage growth outcomes than in the proceeding decades.[3]

We construct a straw man for the analysis. Starting at age 25 and earning 40,000 currency units, this plan member contributes 9 per cent of their salary once per annum (at years end) over 41 years, to create 40 years of return experience (the first year is excluded as the first contribution is at the end of year 1). The tag of currency units is used, as the analysis is examining multiple countries, all with their own individual currencies.

We would like to note here that, while 40,000 currency units, along with a 9% contribution profile, may be applicable for some countries for members at the age of 25, other countries may have a smaller, or larger starting salary. Even the contribution rate is likely to differ. On the face of it, this seems to be a limitation of this international study. For example, at the time of writing, 40,000 USD is over 100 times

[3] Commentators in Australia have conjectured that the low wage growth in western countries is due to poor macroeconomic conditions and the importation of labour (Jacobs & Rush, 2015; Scutt, 2016).

greater than 40,000 Yen. However, mathematically the starting salary is not the issue. Recall that the analysis is testing the trimmed bootstrap against other common techniques in the literature. Hence, all such techniques will be subjected to the same currency earning assumptions. This enables us to observe if a particular country's return experience can benefit from the trimmed bootstrap procedure under the constant wage assumption. The currency unit starting value could be assumed to be 'x'. There is no difference in growing 40,000 units against 1 unit, or 1 million units. Algebraically, and like for like currency units, the starting salary, and the contribution rate as a percentage of income, will not impact the standardised scale of the outcomes, as long as they are constant between simulation techniques.

Where the model could be subject to some criticism is in the growth rates of the contributions. The growth assumption can impact the results, as compounded values can impact the weighting of individual contributions, and in turn, place greater importance on returns experienced later in the wealth accumulation, impacting the dollar-weighted return. Relating back to a previous discussion, this further supports the rationale for implementing multiple salary growth rates. Two growth rates are employed in this study: 4% per annum for nominal returns (made up of an assumption of 2% inflation and 2% real productivity growth); and 2% per annum real growth for real returns, along with an alternative assumption of 0% per annum real growth included for robustness purposes. The 0% per annum real growth allows analysis of equally weighted contributions and also supports the forward-looking possibility that wage growth in the future will be driven largely by inflation.

Empirical results

The results are presented in two sections. The out-of-sample tests section reports the out-of-sample results for each simulation technique in nominal returns. After illustrating case specific outcomes, we present the results for the estimates of the real returns for both the 2 per cent and 0 per cent wage growth per annum cases. This analysis leads into deriving the optimal technique for each country as suggested by the out-of-sample period results. The final analysis is a capstone forward-looking examination of each technique that utilises the entire dataset from 1900 to 2013.

Empirical results: out-of-sample tests

Table 4 provides the differences in MAE in millions of currency units, between the trimmed bootstrap and the respective simulation methodology. The difference has been derived as (S – T) where S is the MAE for the simulation methodology in question and T is the trimmed bootstrap MAE. Hence, positive values represent a superior outcome for the trimmed bootstrap, and negative values represent an inferior result for the trimmed bootstrap.

A positive value is classed as being a 'win' for the trimmed bootstrap and is classed as significant at alpha = 0.05 in a paired t-test. Significant values are formatted bold black in the table for clarity, while those that are red and underlined are instances where the trimmed bootstrap has significantly underperformed the respective simulation method. The trimmed bootstrap has significantly outperformed at the 5% level in 11 of the 21 countries tested, and has significantly underperformed in six countries against the Efron (1979) bootstrap. The Künsch (1989) moving block bootstrap displays varying accuracy depending on block length, with the trimmed bootstrap outperforming between 10 to 11 out of 21 countries tested. The Politis and

Romano (1994) stationary bootstrap has accuracy similar to the Efron (1979) bootstrap, with the trimmed bootstrap again outperforming the technique in 11 out of 21 countries, the same as for the Efron (1979) bootstrap. This corroborates with previous our previous study, that the existing methods in the literature provide similar results relative to the trimmed bootstrap, although the block and stationary bootstraps tend to outperform the Efron bootstrap as they were designed to accommodate some short-term serial dependence (Künsch, 1989).

A graphical representation of the data provides further insights into the mechanics at play between the bootstrap methods. Figure 2 shows the difference in MAE between the particular simulation technique and the trimmed bootstrap. When the column is above the line, the trimmed bootstrap has outperformed the given bootstrap methodology. The figure shows that when the trimmed bootstrap is inferior, the severity of the loss is not on par with the instances when it outperforms.

Table 5 displays the RMSE estimates for the nominal data. The RMSE methodology for comparison differs from the MAE comparison in that it penalises extreme errors to a greater extent than does the MAE. Across all countries, the RMSE results are somewhat stronger than the corresponding MAE results for the trimmed bootstrap. For example, the trimmed bootstrap significantly outperforms the Efron (1979) bootstrap in 14 out of the 21 countries using RMSE as opposed to the 11 observed in the MAE framework. This suggests that the trimmed bootstrap is estimating outcomes which do not produce instances of extreme errors, unlike the competing techniques. The RMSE methodology also observes sign changes (that is, underperformance in MAE is now outperformance for the trimmed bootstrap). This occurs for a number of cases, such as Italy, New Zealand, and Spain. Further, in the

United Kingdom, the trimmed bootstrap now records a significant outperformance against the Efron (1979) bootstrap under the RMSE framework.

While the trimmed bootstrap has not displayed stochastic dominance over all countries, it has outperformed all other techniques in 9 of the 21 countries. A comparison of the countries – which fail and which succeed – sheds further light on the trimmed bootstrap methodology. Countries which had produced extreme abnormal events over the in-sample period of 1900–1973, and that were not reasonably expected to eventuate again in the future, contributed to the instances of failure by the trimmed bootstrap, particularly if their out-of-sample stock performance was a big improvement over their in-sample performance. Countries such as Denmark and Norway, which were occupied by the Nazi party in World War II, and Sweden, which was surrounded by Nazi-occupied warzones that starved their economic development, are among the list. The United Kingdom and Ireland had a similar economic experience over the data used to simulate the out-of-sample period and were both impacted by two world wars.[4]

Spain was involved in its own civil war over the data period (1936–1939), once again impacting economic prosperity (and not replicated in the out-of-sample period). The wealthier republican army lost the three-year conflict and were estimated to have wasted large quantities of their wealth via inefficient war efforts (Martin-Acena, Martinez Ruiz, & Pons, 2012). Such events create political unrest, which influences the nominal figures to a larger extent than real data (as inflation captures a large component of political issues).

[4] See http://www.worldwar2history.info/ for a discussion and experience of individual countries during World War II.

Table 4: Nominal data – differences in MAE between given simulation strategy and trimmed bootstrap (millions of currency units), t-values in parentheses (positive values represent trimmed bootstrap winning and are bold, significant negative, loss, values are underlined). Salary growth = 4 per cent per annum.

Country	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	**1.517 (19.40)**	**1.824 (16.94)**	**1.446 (18.03)**	**1.138 (18.49)**	**1.014 (16.95)**	**1.038 (18.50)**	**0.924 (16.98)**	**0.596 (14.91)**	**0.568 (13.20)**	**0.408 (11.23)**	**0.934 (19.45)**
Austria	**1113.351 (2.71)**	**16926.862 (3.89)**	**27782.215 (3.42)**	**188250.565 (3.24)**	**320142.197 (2.20)**	**560049.523 (2.07)**	12907057.22 (1.08)	2455393.638 (1.38)	2581745.728 (1.07)	1200302.533 (1.25)	**184361.662 (2.15)**
Belgium	**0.617 (10.03)**	**1.221 (6.03)**	**1.260 (7.75)**	**0.920 (8.54)**	**0.712 (8.90)**	**0.700 (10.09)**	**0.885 (9.53)**	**0.668 (8.00)**	**0.551 (7.47)**	**0.421 (7.72)**	**0.694 (8.36)**
Canada	**0.723 (11.42)**	**0.885 (12.42)**	**0.746 (10.50)**	**0.745 (11.29)**	**0.570 (11.28)**	**0.584 (11.39)**	**0.507 (10.79)**	**0.488 (9.88)**	**0.509 (10.01)**	**0.398 (8.77)**	**0.479 (9.78)**
Denmark	-0.342 (-9.30)	-0.368 (-10.67)	-0.316 (-9.23)	-0.302 (-9.58)	-0.237 (-7.43)	-0.246 (-8.20)	-0.154 (-5.35)	-0.184 (-6.61)	-0.178 (-5.99)	-0.180 (-6.66)	-0.399 (-11.68)
Finland	**8.067 (11.53)**	**7.559 (11.45)**	**5.942 (10.37)**	**5.134 (12.73)**	**4.181 (13.67)**	**3.347 (9.86)**	**2.024 (13.93)**	**1.222 (11.11)**	**1.063 (10.54)**	**0.861 (9.04)**	**5.895 (13.80)**
France	**0.436 (2.48)**	**1.414 (4.93)**	**2.342 (4.27)**	**2.918 (2.69)**	**2.385 (5.95)**	**2.013 (6.54)**	**2.498 (6.37)**	**2.827 (7.96)**	**2.773 (6.80)**	**2.303 (7.72)**	**1.639 (5.07)**
Germany	1.16E+34 (1.05)	2.04E+37 (1.00)	1.66E+38 (1.11)	1.63E+39 (1.00)	7.57E+36 (1.24)	8.34E+40 (1.70)	3.15E+41 (1.01)	8.59E+40 (1.00)	4.85E+38 (1.04)	1.48E+41 (1.00)	3.98E+39 (1.00)
Ireland	-0.028 (-0.84)	-0.090 (-2.58)	-0.094 (-2.99)	-0.085 (-2.82)	-0.090 (-2.81)	-0.041 (-1.16)	-0.071 (-2.32)	-0.036 (-0.93)	-0.071 (-2.10)	-0.038 (-1.13)	-0.117 (-3.54)
Italy	1.366 (0.99)	-1.726 (-1.56)	-2.231 (-3.03)	-1.317 (-1.61)	1.239 (0.86)	2.273 (1.55)	0.891 (0.96)	0.871 (0.99)	**5.995 (2.96)**	**4.884 (2.94)**	-0.126 (-0.16)
Japan	**13.342 (6.88)**	**18.488 (8.46)**	**13.577 (8.46)**	**12.704 (6.87)**	**20.788 (7.75)**	**18.279 (7.57)**	**10.857 (6.44)**	**16.153 (5.86)**	**15.355 (8.56)**	**25.607 (4.52)**	**18.663 (7.32)**
The Netherlands	**0.209 (4.86)**	**0.236 (3.68)**	**0.131 (2.43)**	**0.100 (2.46)**	**0.158 (3.73)**	0.070 (1.66)	**0.148 (3.57)**	**0.231 (3.81)**	**0.112 (2.69)**	**0.122 (2.90)**	**0.220 (4.39)**
New Zealand	-0.171 (-6.25)	-0.214 (-8.92)	-0.282 (-12.42)	-0.328 (-15.23)	-0.302 (-14.69)	-0.321 (-15.62)	-0.273 (-13.78)	-0.257 (-13.25)	-0.244 (-13.19)	-0.202 (-11.77)	-0.233 (-11.38)
Norway	-0.323 (-8.96)	-0.348 (-9.73)	-0.314 (-9.60)	-0.261 (-8.00)	-0.218 (-6.83)	-0.180 (-6.24)	-0.068 (-2.39)	-0.111 (-3.74)	-0.069 (-2.43)	-0.111 (-3.82)	-0.415 (-8.37)
Portugal	**3.562 (10.61)**	**4.975 (8.53)**	**3.074 (11.57)**	**3.779 (10.72)**	**3.639 (11.23)**	**3.174 (9.84)**	**3.041 (9.33)**	**2.786 (9.60)**	**2.196 (11.08)**	**2.339 (10.30)**	**5.372 (7.78)**
South Africa	-0.902 (-6.93)	-1.148 (-8.36)	-0.576 (-5.05)	-0.550 (-4.88)	-0.097 (-0.96)	-0.383 (-3.67)	-0.411 (-4.42)	-0.349 (-4.19)	-0.189 (-2.40)	-0.148 (-1.97)	-0.421 (-4.24)
Spain	-0.472 (-6.81)	-0.374 (-5.56)	-0.372 (-5.62)	-0.395 (-7.18)	-0.443 (-8.73)	-0.302 (-5.54)	-0.392 (-7.95)	-0.379 (-7.79)	-0.345 (-7.29)	-0.357 (-7.80)	-0.534 (-9.91)
Sweden	-0.168 (-3.74)	-0.292 (-6.01)	-0.321 (-6.45)	-0.233 (-5.19)	-0.196 (-4.65)	-0.219 (-5.29)	-0.149 (-3.63)	-0.132 (-3.47)	-0.181 (-4.45)	-0.158 (-3.95)	-0.200 (-4.51)
Switzerland	**0.130 (2.97)**	**0.206 (5.38)**	**0.133 (4.41)**	**0.129 (4.46)**	**0.090 (3.63)**	**0.075 (2.94)**	**0.084 (3.48)**	**0.056 (2.53)**	0.030 (1.35)	0.022 (1.05)	**0.102 (4.31)**
United Kingdom	0.006 (0.13)	-0.147 (-3.07)	-0.198 (-4.84)	-0.269 (-6.47)	-0.231 (-5.70)	-0.278 (-6.57)	-0.189 (-4.64)	-0.239 (-5.82)	-0.201 (-4.90)	-0.168 (-4.20)	-0.062 (-1.54)
United States	**0.811 (8.61)**	**0.647 (7.63)**	**0.401 (6.03)**	**0.401 (6.40)**	**0.270 (4.98)**	**0.295 (5.54)**	**0.229 (4.73)**	**0.303 (6.22)**	**0.222 (4.19)**	**0.155 (3.29)**	**0.305 (5.80)**
Trimmed Bootstrap wins*	11/21	11/21	11/21	11/21	11/21	10/21	10/21	10/21	10/21	10/21	11/21
Trimmed Bootstrap loses*	6/21	8/21	9/21	8/21	8/21	7/21	8/21	7/21	8/21	7/21	7/21
Insignificant	4/21	2/21	1/21	2/21	2/21	1/21	3/21	4/21	3/21	4/21	3/21

*Significance defined at alpha = 0.05

Figure 2: Nominal data - differences in MAE between given simulation strategy and trimmed bootstrap (millions of currency units).

NB: Figure excludes Germany and Austria due to extreme values.

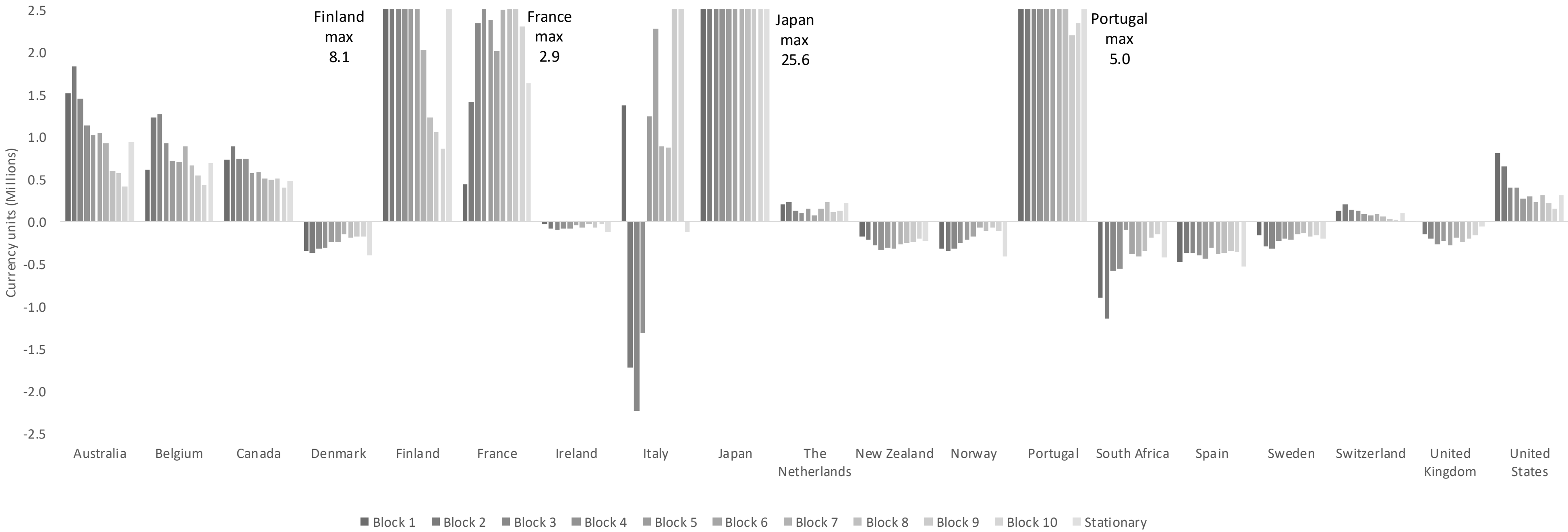


Table 5: Nominal data – differences in RMSE between given simulation strategy and trimmed bootstrap (millions of currency units)

Country	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	2.38***	3.27***	2.38***	1.70***	1.57***	1.50***	1.38***	0.82***	0.86***	0.60***	1.27***
Austria	12779.56***	138319.3***	257951.8***	1847071***	4605356***	8582355***	3.77E+08***	56229887***	76461229***	30382471***	2718679***
Belgium	1.23***	5.42***	4.28***	2.57***	1.69***	1.46***	2.16***	1.77***	1.45***	0.90***	1.80***
Canada	1.52***	1.83***	1.72***	1.61***	1.09***	1.13***	0.96***	0.99***	1.05***	0.84***	1.00***
Denmark	-0.17***	-0.23***	-0.19***	-0.20***	-0.15***	-0.16***	-0.09***	-0.12***	-0.11***	-0.13***	-0.27***
Finland	21.96***	20.53***	17.38***	12.20***	9.05***	9.60***	3.61***	2.16***	1.76***	1.50***	13.23***
France	1.81***	5.82***	13.89***	30.74***	9.50***	6.49***	9.29***	8.31***	9.97***	6.48***	6.85***
Germany	3.31E+29***	6.45E+38***	4.74E+39***	5.14E+40***	1.92E+32***	1.55E+42***	9.87E+42***	2.71E+42***	1.48E+40***	4.69E+42***	1.25E+41***
Ireland	0.04	0.00	-0.03	-0.04*	-0.03	0.04	-0.02	0.11	0.01	0.03	-0.04*
Italy	30.78***	20.80***	8.96***	11.67***	31.63***	32.95***	16.08***	14.39***	51.14***	40.01***	10.85***
Japan	48.81***	58.60***	38.77***	46.13***	73.57***	64.64***	39.80***	74.05***	44.85***	165.72***	68.97***
The Netherlands	0.42***	0.88***	0.59***	0.27***	0.37***	0.30***	0.32***	0.78***	0.29***	0.32***	0.57***
New Zealand	0.03	-0.06***	-0.14***	-0.20***	-0.19***	-0.21***	-0.17***	-0.17***	-0.17***	-0.14***	-0.13***
Norway	-0.19***	-0.22***	-0.21***	-0.16***	-0.13***	-0.11***	-0.01	-0.04	-0.02	-0.05**	-0.12
Portugal	9.33***	17.05***	7.12***	9.84***	9.00***	8.69***	8.76***	7.57***	4.75***	5.61***	20.32***
South Africa	-0.61***	-0.82***	-0.36***	-0.34***	0.06	-0.21***	-0.29***	-0.26***	-0.11*	-0.08	-0.27
Spain	-0.03	0.01	0.00	-0.17***	-0.27***	-0.10*	-0.25***	-0.24***	-0.23***	-0.25***	-0.31***
Sweden	-0.12	-0.23	-0.25	-0.20	-0.17	-0.19	-0.13	-0.12	-0.16	-0.14	-0.17
Switzerland	0.66***	0.60***	0.35***	0.33***	0.20***	0.22***	0.18***	0.11***	0.10***	0.05***	0.19***
United Kingdom	0.16***	0.03	-0.11***	-0.17***	-0.13***	-0.17***	-0.10***	-0.15***	-0.12***	-0.10***	0.01
United States	1.93***	1.63***	1.00***	0.90***	0.63***	0.63***	0.49***	0.55***	0.58***	0.40***	0.63***

Key: Positive values (trimmed bootstrap superior) in black, negative values (trimmed bootstrap inferior) in red.

* significant at alpha = 0.10; ** significant at alpha = 0.05; *** significant at alpha = 0.01.

A common link is present in the majority of countries where the trimmed bootstrap records an inferior result in nominal figures. The formation of the European Economic Community (EEC), which would become the European Union (EU) occurred in 1958, towards the end of the in-sample period. The EEC formed the European Currency Unit in the out-of-sample period (1979), which provided a basket of European currencies and aimed at minimising fluctuations between their values to stimulate trade. This was followed by the formation of a common currency (the Euro) for the majority of European states in 1999. These events represent significant changes in the monetary policy for these countries, attaching a bias not captured in historical data. We do not expect the trimmed bootstrap to outperform in scenarios where the future is dramatically more extreme than history: such political changes fall within this description.

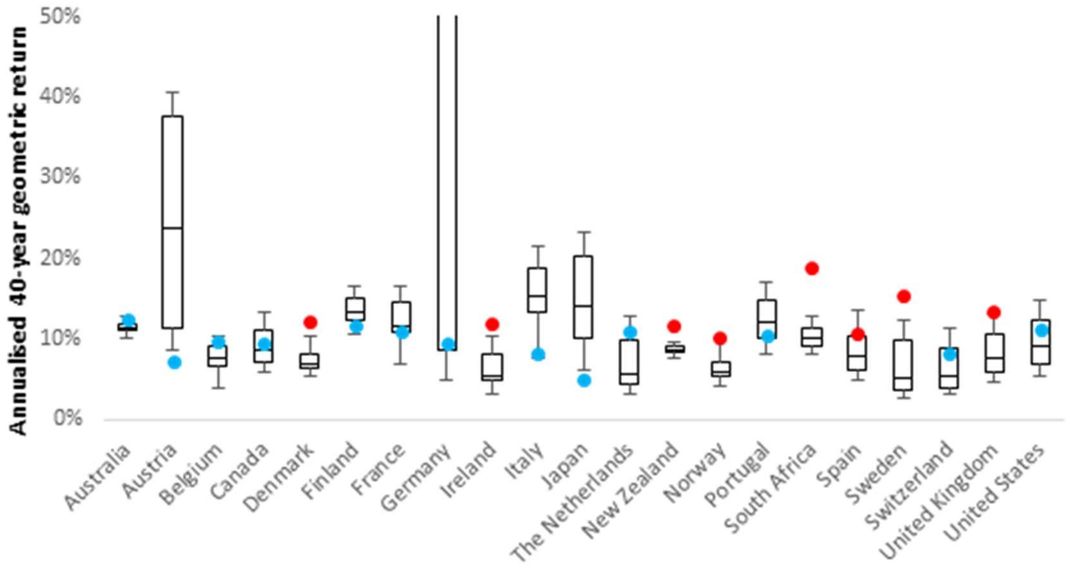
Further instances of underperformance are observed in South Africa. The in-sample period saw South Africa experience the second Boer War (1899–1902) and a reign of apartheid which put in place a range of economic and political sanctions (Bayoumi, 1990; Laverty, 2007). These events fundamentally impacted the economic prosperity of the country with large scale global sanctions on trade (Johnson & Dickinson, 2015). The increased economic activities in South Africa over the out-of-sample period were politically motivated and are hypothesised to have impacted the nominal return results of the trimmed bootstrap significantly as it relies more heavily on future returns relating to historical experience than on other methods.

New Zealand offers the only instance of the failure of the trimmed bootstrap while having a complete detachment from conflict (albeit by geographical location only). While New Zealand was detached geographically, it did experience a fundamental shift in its economy, from the test data to the out-of-sample period. Prior to the 1980s, New Zealand was an extremely closed economy (Evans, Grimes, Wilkinson, & Teece, 1996). Over the period of 1983 to 1993, the OECD (1994) found that the New Zealand economy measure of 'openness' (a ratio of imports plus exports to GDP) increased by 42 per cent. The country was subjected to a highly protectionist political stance which was liberalised during the out-of-sample period. New Zealand experienced a political shift between the periods under measure, which is a factor in the unreliability of the trimmed bootstrap.

However, while these instances of failure can be discussed, it is not a blanket solution. The reader may look to Japan, a country where the trimmed bootstrap has been superior, yet which had an in-sample period which saw two atom bombs decimate the country against a war-free out-of-sample period. This does not seem to resonate with the previous discussion of the trimmed bootstrap failing in countries that had structural differences between the in-sample and out-of-sample periods. Instead, this provides the foundation for examining the data differences between the in-sample and out-of-sample periods in order to understand what drives the failure of the trimmed bootstrap.

Figure 3 depicts the in-sample 40-year geometric mean return box-and-whisker plot for the 21 countries examined. The ends of a 'box' denote the positions of the 25th and 75th percentiles, while the line across the box denotes its median value. The ends of the 'whiskers' denote either the minimum or maximum value. The overlaid dots on the figure are the geometric mean returns from the 40-year out-of-sample period. Those coloured red are instances of significant underperformance of the trimmed bootstrap relative to one or more of the 11 alternative bootstrap methods. The blue dots represent either a significant outperformance of the trimmed bootstrap, or an insignificant outcome to any technique. The figure shows a clear image that, aside from Spain, in instances where the 40-year out-of-sample period was beyond the upper-end of historical observations, the trimmed bootstrap was an inferior estimator. In the case of Spain, the trimmed underperformance was due to an out-of-sample result that was in the upper tail of what had been observed in-sample. While this shows that the trimmed bootstrap may not be the best method if the future outcome is relatively extreme on the upside compared to past history, the same issue is not observed at the lower extreme. Instances where the out-of-sample 40-year return was below the in-sample observations did not record an inferior result for the trimmed bootstrap (Austria and Japan). In every one of the eight cases of trimmed bootstrap failure, the out-of-sample result is well above the median in-sample result. This suggests that the trimmed bootstrap failures are driven by instances where the future is a lot better than for previous norms, something which would be of little concern to fund managers and particularly to plan members.

Figure 3: Box and whisker plot of in-sample 40-year geometric nominal returns with overlaid out-of-sample period observation



Red dots are instances the trimmed bootstrap was statistically insignificant against one or more of the 11 other techniques and those coloured blue, the trimmed bootstrap, significantly outperformed all other techniques.

The results applying nominal returns are consistent with the US analysis in our previous paper. However, the trimmed bootstrap does not display stochastic dominance in all indexes. As discussed previously, the trimmed bootstrap is observed to be the suboptimal simulation technique for indexes that experienced a fundamental change in their return environment between the in-sample and out-of-sample periods. The study continues with analysis into real returns, which are adjusted for inflation, a macroeconomic factor which encapsulates a degree of political instability (Aisen & Veiga, 2006).

Table 6 shows the real returns results when assuming a two per cent wage growth per annum. The number of instances where the trimmed bootstrap has outperformed on a significant basis has increased from a range of 10-11/21 to 14/21 against the range of simulation techniques in the MAE tests. Thus, when the influence of inflation is removed, the trimmed bootstrap outperforms to a greater extent.

When the real returns are contrasted with those of the nominal returns, similarities arise. Firstly, the countries in which the trimmed bootstrap outperforms the majority of the other methodologies in both nominal and real returns are 9 of the 21 countries (Austria, Australia, Canada, Italy, Japan, The Netherlands, Portugal, Switzerland, and the United States). In contrast, there are three countries where the trimmed bootstrap is inferior in both nominal and real returns (Norway, Spain, and Sweden). Geographically speaking, the trimmed bootstrap has performed well across the world; however, there are poor results from countries close to or involved with WWII, the formation of the ECC, or their own civil war.

A number of countries recorded a change in the trimmed bootstrap's performance against existing bootstrap techniques. In two countries, the trimmed bootstrap was significantly superior in nominal returns, though inferior in real returns (Denmark and

France). Contrary to recording worse results in real returns, two countries moved in favour of the trimmed bootstrap, recording inferior results for the trimmed bootstrap in nominal returns, but in favour of the technique in real returns (New Zealand and South Africa).

Table 7 provides the results for the RMSE methodology. Across the 21 countries, the trimmed bootstrap is inferior to the Efron (1979) bootstrap in only three countries (Denmark, Norway and Sweden). Further, Belgium, France and Spain all reduced the prevalence of estimating a failure for the trimmed bootstrap when employing the RMSE methodology. This again supports the nominal estimates and suggests the trimmed bootstrap's accuracy benefits from the removal of extreme results that record very large errors.

The performance of the trimmed bootstrap is more reliant on the consistency between the in-sample and out-of-sample test periods. Previous discussions had suggested that if history was exceeded in the future by a considerable margin, the trimmed bootstrap would be inferior to those methodologies which do not have such reliance on history observations. However, the results from this study suggest that the trimmed bootstrap may be inferior only if history is exceeded on the up-side, not so the down-side.

Table 6: Real data - differences in MAE between given simulation strategy and trimmed bootstrap (millions of currency units), t-values in parentheses (positive values represent trimmed bootstrap winning and are bold, significant negative, loss, values are underlined). Salary growth = 2 per cent per annum

Country	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	0.289 (8.56)	0.589 (10.52)	0.446 (10.83)	0.470 (12.15)	0.435 (10.60)	0.488 (12.06)	0.548 (12.59)	0.390 (10.50)	0.321 (9.53)	0.256 (8.53)	0.215 (6.68)
Austria	0.078 (5.64)	0.084 (5.24)	0.111 (6.03)	0.122 (6.26)	0.146 (5.74)	0.140 (7.64)	0.127 (6.62)	0.229 (7.07)	0.143 (6.63)	0.230 (6.26)	0.158 (6.67)
Belgium	-0.019 (-1.66)	-0.042 (-2.75)	-0.037 (-3.30)	-0.063 (-6.19)	-0.053 (-4.24)	-0.053 (-4.51)	-0.051 (-4.85)	-0.026 (-1.94)	-0.041 (-4.27)	-0.043 (-4.74)	-0.053 (-5.18)
Canada	0.163 (6.83)	0.346 (9.93)	0.283 (7.47)	0.321 (8.74)	0.230 (8.20)	0.260 (8.94)	0.236 (8.26)	0.203 (6.47)	0.200 (7.02)	0.120 (5.65)	0.151 (5.83)
Denmark	-0.075 (-7.64)	-0.110 (-11.16)	-0.082 (-8.69)	-0.076 (-8.35)	-0.051 (-6.00)	-0.063 (-7.45)	-0.038 (-4.47)	-0.032 (-4.14)	-0.031 (-3.99)	-0.032 (-4.61)	-0.077 (-8.54)
Finland	-0.023 (-0.71)	0.036 (0.96)	0.058 (1.68)	-0.034 (-1.10)	0.046 (1.20)	-0.049 (-2.18)	-0.014 (-0.66)	-0.081 (-3.92)	-0.067 (-3.67)	-0.118 (-6.42)	0.025 (0.64)
France	-0.039 (-3.65)	-0.030 (-1.93)	0.007 (0.26)	0.016 (0.48)	-0.006 (-0.30)	-0.031 (-2.15)	-0.017 (-1.17)	-0.029 (-1.94)	-0.028 (-2.17)	-0.028 (-2.31)	-0.013 (-1.12)
Germany	0.536 (5.23)	0.464 (6.95)	1.048 (3.21)	0.590 (5.83)	0.809 (5.48)	0.733 (6.09)	0.648 (5.43)	0.957 (4.51)	1.006 (4.57)	1.131 (4.95)	0.716 (6.33)
Ireland	0.051 (4.90)	0.038 (3.68)	0.044 (3.91)	0.037 (3.84)	0.048 (4.46)	0.055 (4.76)	0.051 (4.46)	0.077 (5.61)	0.047 (3.99)	0.040 (3.48)	0.044 (4.06)
Italy	0.343 (9.35)	0.306 (9.18)	0.313 (8.38)	0.270 (10.06)	0.281 (9.49)	0.297 (10.71)	0.272 (9.74)	0.324 (11.25)	0.277 (9.82)	0.330 (9.11)	0.265 (8.82)
Japan	1.475 (12.52)	2.699 (10.03)	2.724 (11.09)	2.603 (9.49)	3.034 (10.65)	2.400 (14.05)	1.869 (13.53)	1.797 (15.07)	2.017 (11.95)	2.044 (10.45)	2.241 (11.09)
The Netherlands	0.102 (4.81)	0.086 (4.07)	0.066 (3.41)	0.057 (3.86)	0.076 (5.44)	0.042 (2.68)	0.061 (4.15)	0.105 (4.66)	0.044 (2.93)	0.037 (2.54)	0.084 (5.64)
New Zealand	0.218 (16.76)	0.231 (16.46)	0.189 (16.15)	0.198 (17.61)	0.169 (15.16)	0.156 (15.55)	0.183 (15.73)	0.161 (15.33)	0.125 (13.51)	0.101 (12.34)	0.146 (14.13)
Norway	-0.080 (-6.97)	-0.098 (-8.88)	-0.080 (-7.53)	-0.068 (-6.42)	-0.048 (-4.52)	-0.055 (-5.44)	-0.022 (-2.24)	-0.029 (-2.93)	-0.018 (-1.82)	-0.024 (-2.45)	-0.116 (-9.73)
Portugal	0.367 (8.78)	0.404 (9.02)	0.337 (8.85)	0.309 (7.80)	0.261 (6.27)	0.167 (6.49)	0.136 (5.28)	0.118 (5.26)	0.132 (5.12)	0.104 (4.58)	0.463 (10.53)
South Africa	0.686 (14.41)	0.898 (13.33)	0.672 (13.36)	0.807 (11.76)	0.808 (5.87)	0.720 (13.26)	0.730 (9.92)	0.600 (14.43)	0.530 (15.32)	0.484 (15.80)	0.592 (12.94)
Spain	-0.142 (-9.51)	-0.129 (-9.97)	-0.114 (-8.52)	-0.107 (-8.43)	-0.112 (-9.19)	-0.089 (-7.04)	-0.106 (-11.23)	-0.095 (-9.26)	-0.096 (-9.87)	-0.112 (-12.93)	-0.125 (-12.35)
Sweden	-0.054 (-3.71)	-0.115 (-6.95)	-0.103 (-6.40)	-0.098 (-5.98)	-0.083 (-5.33)	-0.111 (-6.98)	-0.091 (-6.04)	-0.066 (-4.79)	-0.076 (-5.25)	-0.052 (-4.09)	-0.062 (-4.17)
Switzerland	0.075 (6.35)	0.113 (5.76)	0.091 (5.80)	0.097 (4.94)	0.085 (6.05)	0.075 (4.56)	0.099 (6.22)	0.095 (5.94)	0.076 (5.11)	0.035 (3.11)	0.081 (6.28)
United Kingdom	0.139 (9.75)	0.193 (9.73)	0.116 (7.86)	0.110 (6.62)	0.120 (8.00)	0.127 (8.10)	0.115 (7.67)	0.111 (6.65)	0.075 (5.49)	0.048 (4.06)	0.117 (9.05)
United States	0.493 (11.55)	0.510 (10.77)	0.360 (9.57)	0.373 (11.30)	0.309 (10.69)	0.332 (11.61)	0.324 (11.39)	0.336 (10.01)	0.294 (9.84)	0.200 (8.60)	0.274 (10.30)
Trimmed Bootstrap wins*	14/21	14/21	14/21	14/21	14/21	14/21	14/21	14/21	14/21	14/21	14/21
Trimmed Bootstrap loses*	5/21	6/21	5/21	5/21	5/21	7/21	5/21	7/21	6/21	7/21	5/21
Insignificant	2/21	1/21	2/21	2/21	2/21	0/21	2/21	0/21	1/21	0/21	2/21

*Significance defined at alpha = 0.05

Figure 4: Real data – differences in MAE between given simulation strategy and trimmed bootstrap (millions of currency units)

Salary growth = 2 per cent per annum.

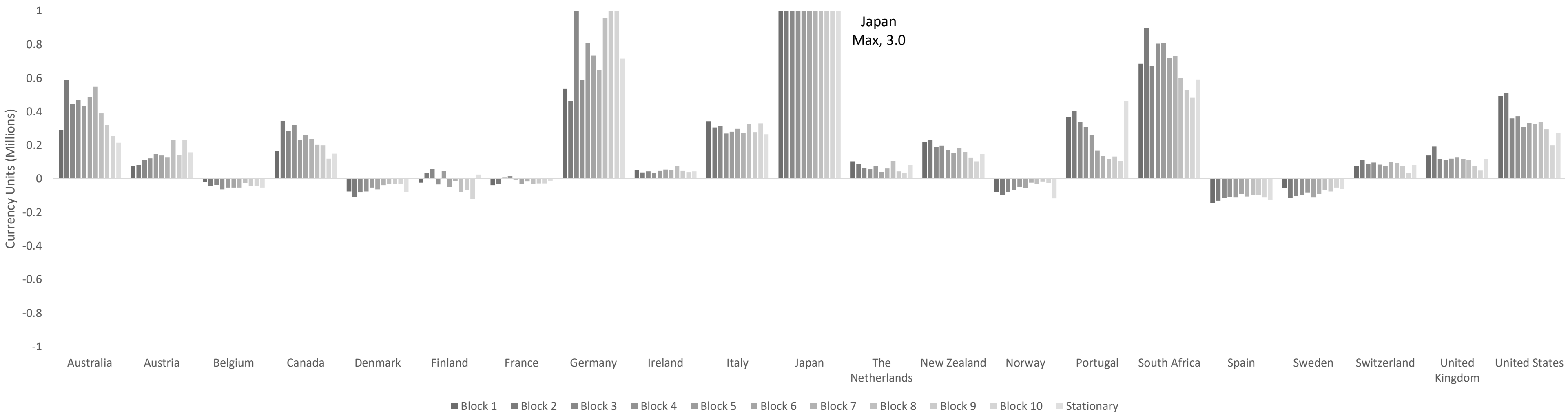


Table 7: Real data 2 per cent salary growth – differences in RMSE between given simulation strategy and trimmed bootstrap (millions of currency units)

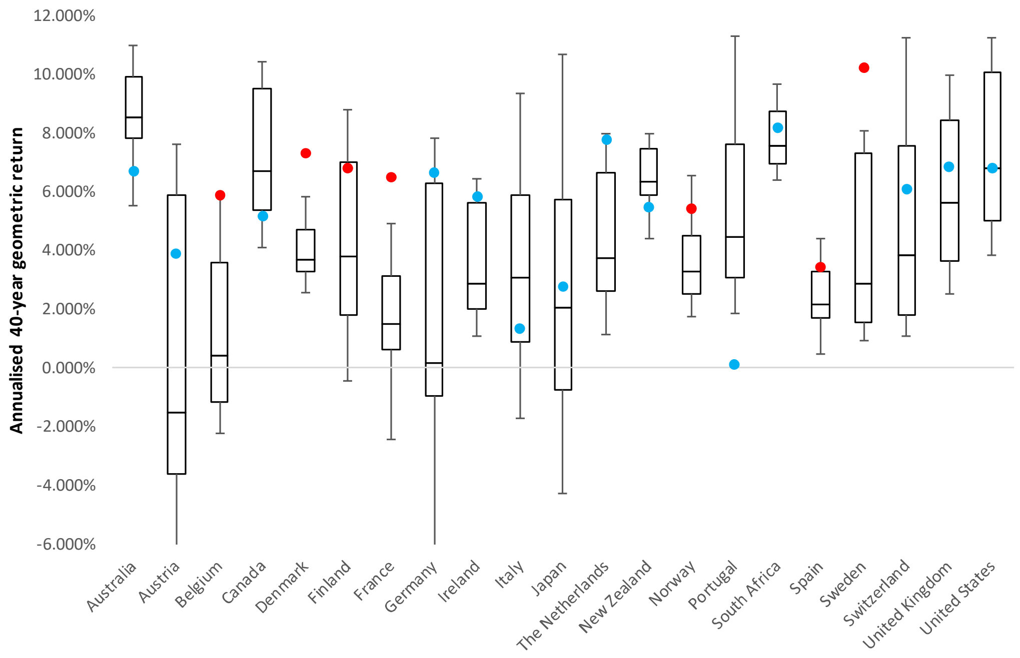
Country	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	0.62***	1.38***	0.89***	0.86***	0.88***	0.91***	1.04***	0.76***	0.62***	0.49***	0.54***
Austria	0.17***	0.23***	0.31***	0.35***	0.52***	0.32***	0.33***	0.75***	0.41***	0.88***	0.47***
Belgium	0.03*	0.07*	0.01	-0.03***	0.01	0.00	-0.01	0.05	-0.02**	-0.03***	-0.02*
Canada	0.36***	0.77***	0.83***	0.79***	0.50***	0.56***	0.52***	0.58***	0.51***	0.26***	0.39***
Denmark	-0.03***	-0.07***	-0.04***	-0.04***	-0.02***	-0.04***	-0.01**	-0.01*	-0.01**	-0.02***	-0.05***
Finland	0.29***	0.42***	0.40***	0.26***	0.47***	0.09***	0.08***	0.02	0.00	-0.05***	0.47***
France	0.00	0.08*	0.33***	0.51***	0.20***	0.06**	0.08***	0.08**	0.05*	0.03*	0.05***
Germany	2.79***	1.71***	9.83***	2.76***	4.23***	3.38***	3.32***	6.23***	6.52***	6.79***	3.16***
Ireland	0.09***	0.08***	0.09***	0.07***	0.09***	0.11***	0.10***	0.18***	0.11***	0.09***	0.09***
Italy	0.95***	0.85***	0.97***	0.64***	0.73***	0.68***	0.67***	0.72***	0.68***	0.94***	0.72***
Japan	3.52***	8.41***	7.72***	8.52***	8.98***	5.41***	4.25***	3.71***	5.20***	5.99***	6.30***
The Netherlands	0.28***	0.26***	0.22***	0.11***	0.13***	0.12***	0.12***	0.31***	0.11***	0.10***	0.14***
New Zealand	0.35***	0.38***	0.30***	0.29***	0.27***	0.23***	0.29***	0.25***	0.19***	0.15***	0.23***
Norway	-0.03**	-0.05***	-0.04***	-0.03***	-0.01*	-0.02***	0.01	0.00	0.01	0.00	-0.06***
Portugal	0.96***	1.04***	0.82***	0.87***	0.90***	0.38***	0.38***	0.25***	0.35***	0.25***	1.04***
South Africa	1.32***	1.98***	1.40***	1.96***	4.00***	1.52***	2.07***	1.13***	0.91***	0.79***	1.24***
Spain	-0.01	-0.04**	-0.02	-0.02	-0.03*	-0.01	-0.06***	-0.04***	-0.05***	-0.08***	-0.07***
Sweden	-0.03***	-0.08***	-0.07***	-0.07***	-0.06***	-0.08***	-0.07***	-0.05***	-0.06***	-0.04***	-0.04***
Switzerland	0.10***	0.29***	0.20***	0.28***	0.15***	0.20***	0.20***	0.20***	0.17***	0.06***	0.13***
United Kingdom	0.21***	0.38***	0.21***	0.24***	0.22***	0.24***	0.21***	0.24***	0.15***	0.09***	0.18***
United States	1.10***	1.25***	0.90***	0.78***	0.64***	0.65***	0.63***	0.78***	0.65***	0.42***	0.56***

Key: Positive values (trimmed bootstrap superior) in black, negative values (trimmed bootstrap inferior) in red.

* significant at alpha = 0.10; ** significant at alpha = 0.05; *** significant at alpha = 0.01.

Figure 5 shows the box-and-whisker plots of the in-sample 40-year geometric real returns overlaid with the 40-year out-of-sample return (the real return analysis of Figure 3). Again, the instances where the trimmed bootstrap underperformed any one of the 11 techniques are coloured red, while the blue out-of-sample points refer to a superior result for the trimmed bootstrap over all other techniques. As with Figure 3, in the instances where the trimmed bootstrap has been inferior, the out-of-sample period recorded a very high return, exceeding the in-sample maximum in four out of six occasions (and above the 75th percentile in the remaining two). However, the most striking point to make from the combination of Figure 3 and Figure 5 is that in a total of three instances between both nominal (Austria & Japan) and real (Portugal), the out-of-sample return was below the minimum in-sample observation, yet the trimmed bootstrap was more-accurate than the Efron (1979) bootstrap. This corroborates with our views in our previous paper, where the trimmed bootstrap was impacting the upper-tail more than the lower tail. This is reflected in the RMSE results, where the trimmed bootstrap performs better as the test methodology places greater weight on larger errors (which will be more prevalent on the up-side of the distribution).

Figure 5: Box and whisker plot of in-sample 40-year geometric real returns with overlaid out-of-sample period observation.



Red dots are instances where the trimmed bootstrap was statistically insignificant against one or more of the 11 other techniques; those coloured blue, the trimmed bootstrap, significantly outperformed all other techniques.

Table 8 provides a summary of the relationships between nominal and real return outcomes, against the results for the Efron bootstrap. As this analysis is a binary determination, it requires disregarding significance. Hence, while Belgium and Finland are in the bottom-left window (inferior-superior), the trimmed bootstrap was not statistically inferior in real returns. Once the countries are grouped, a visual geographical relationship forms. In regard to real returns, all with the exception of Norway (which borders Sweden) are in the European Union, which formed over the out-of-sample period. This represented a significant change in the underlying economic prosperity within these countries and rendered the trimmed bootstrap, which relies on historical data more than other techniques used, an inferior methodology for these locations. Three countries in which the trimmed bootstrap was inferior in the nominal returns data, but was superior in real data: Ireland, New Zealand and South Africa. The impact of political shifts between the in-sample and out-of-sample periods in these countries is reduced by removing inflation.

Table 8: Summary of the relationship regarding country-specific performance between nominal and real returns against the Efron bootstrap*

Nominal Returns (Table 4)

Real Returns (Table 6)	Trimmed superior	Trimmed Inferior
Trimmed Superior	Australia Austria Canada Italy Japan Netherlands Portugal Switzerland United States United Kingdom	Ireland New Zealand South Africa
Trimmed Inferior	Belgium Finland France	Denmark Norway Spain Sweden

*Germany was excluded from this table as it provided too many insignificant results in nominal return analysis due to its hyper-inflationary period.

Tables 9 and 10 are robustness tests that test the sensitivity of the previous real results to the productivity wage growth assumption of two per cent per annum. Therefore, Tables 9 and 10 are based on the alternative assumption of wage growth of zero per cent per annum, thus equally weighting contributions. Across the table, no MAE altered its sign (+/-); however, a total of three changes in *t*-stat significance were observed. There is now a significantly inferior result for the trimmed bootstrap in France against the Block 7, and two now insignificant superior results in The

Netherlands against both the Block 6 and Block 10 bootstraps. The robustness test signifies that the wage growth assumption has little impact on the results. With regard to the comparisons between the RMSE results in the real data, only France recorded a sign change (+/-) against the Efron (1979): Block 1, from 0.00 in two per cent wage growth to -0.01 when using the zero per cent wage growth assumption.

Stochastic Dominance (SD) and Almost Stochastic Dominance (ASD) are two further ways to measure the accuracy of the trimmed bootstrap. Table 11 displays the two measures for the real data analysis with the assumption of a zero per cent salary growth rate. The analysis is conducted with the assumption of no salary growth to offer robustness to forward-looking opinions around lower wage growth. The light green shade in the table signifies the trimmed bootstrap displaying SD in the country against the simulation technique in question. The dark green shading demonstrates that the trimmed bootstrap achieved the criteria for ASD to hold, which is a violation of less than 0.059 (Leshno & Levy, 2002).

The trimmed bootstrap displays either SD or ASD against at least one other technique in 12 of the 21 countries tested. Of the 231 observations in Table 11, 39 are observations where the trimmed bootstrap displayed SD against an alternative bootstrap technique. The trimmed bootstrap displays ASD with an area of violation below 0.059 a total of 73 times across Table 32. Furthermore, there was no situation where another bootstrap technique displayed either SD or ASD over the trimmed bootstrap.

Table 9: Real data – differences in MAE between given simulation strategy and trimmed bootstrap (millions of currency units), t-values in parentheses (positive values represent trimmed bootstrap winning and are bold, significant negative, loss, values are underlined). Salary growth = 0 per cent per annum.

Country	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	**0.253 (8.80)**	**0.505 (10.44)**	**0.387 (11.04)**	**0.406 (12.24)**	**0.378 (10.75)**	**0.423 (12.24)**	**0.475 (12.73)**	**0.339 (10.72)**	**0.279 (9.75)**	**0.223 (8.76)**	**0.189 (6.88)**
Austria	**0.065 (5.88)**	**0.070 (5.48)**	**0.091 (6.07)**	**0.101 (6.53)**	**0.123 (5.79)**	**0.113 (7.60)**	**0.102 (6.49)**	**0.189 (6.95)**	**0.117 (6.66)**	**0.188 (6.00)**	**0.127 (6.56)**
Belgium	-0.021 (-2.28)	-0.038 (-3.02)	-0.035 (-3.88)	-0.055 (-6.88)	-0.046 (-4.59)	-0.046 (-4.93)	-0.046 (-5.35)	-0.024 (-2.30)	-0.036 (-4.82)	-0.038 (-5.38)	-0.046 (-5.85)
Canada	**0.138 (6.87)**	**0.292 (9.89)**	**0.240 (7.49)**	**0.271 (8.69)**	**0.194 (8.26)**	**0.218 (8.95)**	**0.198 (8.19)**	**0.172 (6.58)**	**0.167 (7.01)**	**0.103 (5.76)**	**0.126 (5.82)**
Denmark	-0.064 (-8.10)	-0.092 (-11.63)	-0.068 (-9.02)	-0.063 (-8.64)	-0.044 (-6.38)	-0.052 (-7.84)	-0.032 (-4.77)	-0.028 (-4.49)	-0.026 (-4.30)	-0.027 (-4.87)	-0.065 (-8.99)
Finland	-0.027 (-0.96)	0.024 (0.77)	0.042 (1.40)	-0.037 (-1.38)	0.032 (0.96)	-0.050 (-2.71)	-0.022 (-1.30)	-0.076 (-4.48)	-0.065 (-4.38)	-0.104 (-6.95)	0.011 (0.34)
France	-0.037 (-4.56)	-0.033 (-2.63)	-0.003 (-0.16)	0.006 (0.19)	-0.013 (-0.79)	-0.035 (-3.08)	-0.023 (-2.02)	-0.033 (-2.61)	-0.033 (-3.18)	-0.031 (-3.26)	-0.020 (-2.18)
Germany	**0.448 (5.24)**	**0.388 (6.80)**	**0.906 (3.12)**	**0.487 (5.85)**	**0.680 (5.23)**	**0.608 (6.03)**	**0.541 (5.15)**	**0.816 (4.42)**	**0.858 (4.34)**	**0.942 (4.87)**	**0.599 (6.27)**
Ireland	**0.034 (4.05)**	**0.019 (2.23)**	**0.026 (2.88)**	**0.021 (2.65)**	**0.028 (3.21)**	**0.034 (3.66)**	**0.029 (3.15)**	**0.048 (4.27)**	**0.026 (2.76)**	**0.020 (2.20)**	**0.026 (2.90)**
Italy	**0.286 (9.29)**	**0.250 (9.15)**	**0.262 (8.32)**	**0.223 (10.13)**	**0.233 (9.51)**	**0.243 (10.67)**	**0.224 (9.69)**	**0.268 (11.21)**	**0.228 (10.03)**	**0.274 (9.05)**	**0.221 (8.87)**
Japan	**1.266 (12.34)**	**2.306 (9.89)**	**2.370 (10.82)**	**2.243 (9.31)**	**2.623 (10.49)**	**2.067 (13.73)**	**1.593 (13.47)**	**1.542 (14.90)**	**1.739 (11.71)**	**1.754 (10.21)**	**1.939 (10.79)**
The Netherlands	**0.073 (4.12)**	**0.058 (3.33)**	**0.046 (2.73)**	**0.032 (2.67)**	**0.049 (4.34)**	0.016 (1.28)	**0.040 (3.33)**	**0.073 (3.90)**	0.022 (1.83)	0.018 (1.52)	**0.061 (5.13)**
New Zealand	**0.185 (17.16)**	**0.194 (16.66)**	**0.160 (16.48)**	**0.166 (17.97)**	**0.144 (15.50)**	**0.132 (15.88)**	**0.155 (16.09)**	**0.137 (15.81)**	**0.106 (14.03)**	**0.087 (12.93)**	**0.126 (14.71)**
Norway	-0.068 (-7.60)	-0.082 (-9.30)	-0.068 (-7.90)	-0.058 (-6.81)	-0.042 (-5.04)	-0.048 (-5.91)	-0.021 (-2.75)	-0.026 (-3.34)	-0.018 (-2.27)	-0.021 (-2.76)	-0.099 (-10.26)
Portugal	**0.305 (8.63)**	**0.336 (8.92)**	**0.280 (8.74)**	**0.261 (7.67)**	**0.218 (6.11)**	**0.139 (6.52)**	**0.117 (5.48)**	**0.099 (5.38)**	**0.111 (5.18)**	**0.088 (4.72)**	**0.389 (10.49)**
South Africa	**0.591 (14.26)**	**0.766 (13.18)**	**0.572 (13.22)**	**0.681 (11.65)**	**0.694 (5.58)**	**0.612 (12.96)**	**0.621 (9.83)**	**0.512 (14.49)**	**0.450 (15.25)**	**0.411 (15.67)**	**0.502 (13.00)**
Spain	-0.115 (-9.28)	-0.106 (-9.94)	-0.094 (-8.81)	-0.088 (-8.53)	-0.093 (-9.43)	-0.074 (-7.17)	-0.089 (-11.99)	-0.078 (-9.41)	-0.079 (-10.31)	-0.091 (-13.37)	-0.102 (-12.54)
Sweden	-0.047 (-3.97)	-0.096 (-7.11)	-0.088 (-6.66)	-0.082 (-6.19)	-0.071 (-5.61)	-0.091 (-7.17)	-0.076 (-6.27)	-0.056 (-5.09)	-0.064 (-5.48)	-0.044 (-4.33)	-0.052 (-4.33)
Switzerland	**0.060 (6.24)**	**0.091 (5.57)**	**0.075 (5.71)**	**0.081 (4.90)**	**0.068 (5.91)**	**0.059 (4.43)**	**0.077 (6.00)**	**0.074 (5.72)**	**0.059 (4.88)**	**0.027 (2.90)**	**0.065 (6.11)**
United Kingdom	**0.108 (9.21)**	**0.147 (9.06)**	**0.086 (7.07)**	**0.080 (5.72)**	**0.088 (7.04)**	**0.094 (7.27)**	**0.081 (6.58)**	**0.082 (5.94)**	**0.053 (4.69)**	**0.031 (3.13)**	**0.092 (8.59)**
United States	**0.417 (11.49)**	**0.431 (10.54)**	**0.305 (9.58)**	**0.313 (11.21)**	**0.262 (10.72)**	**0.280 (11.67)**	**0.271 (11.33)**	**0.283 (10.10)**	**0.246 (9.83)**	**0.171 (8.72)**	**0.234 (10.37)**
Trimmed Bootstrap wins*	14/21	14/21	14/21	14/21	14/21	13/21	14/21	14/21	13/21	13/21	14/21
Trimmed Bootstrap loses*	6/21	6/21	5/21	5/21	5/21	7/21	6/21	7/21	7/21	7/21	6/21
Insignificant	1/21	1/21	2/21	2/21	2/21	1/21	1/21	0/21	1/21	1/21	1/21

*Significance define at alpha = 0.05

Figure 6: Real data – differences in MAE between given simulation strategy and trimmed bootstrap (millions of currency units)
Salary growth = 0 per cent per annum.

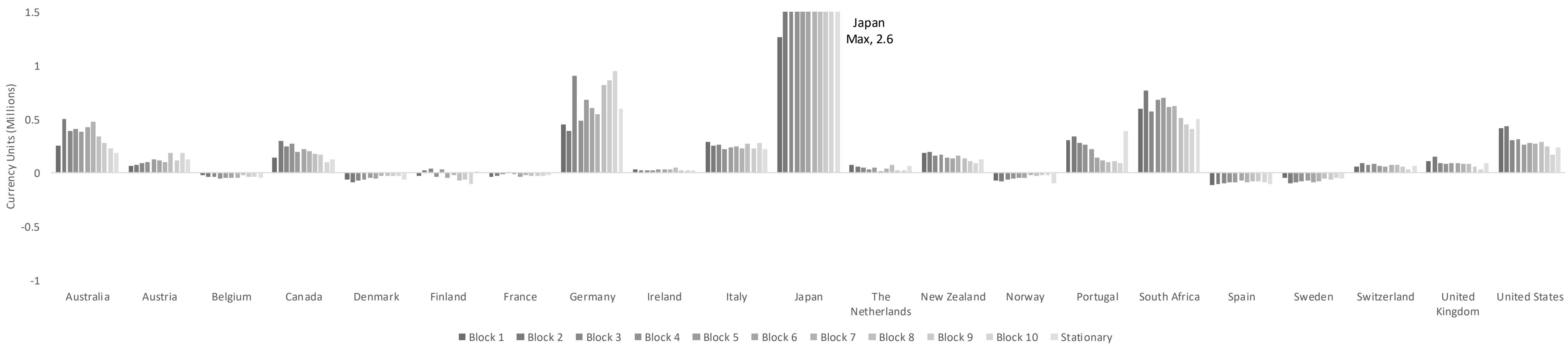


Table 10: Real data 0 per cent salary growth – differences in RMSE between given simulation strategy and trimmed bootstrap (millions of currency units)

Country	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	0.55***	1.21***	0.78***	0.76**	0.78***	0.80***	0.91***	0.66***	0.55***	0.43***	0.47***
Austria	0.15***	0.19***	0.26	0.28	0.45***	0.27***	0.28***	0.64***	0.34***	0.77***	0.39***
Belgium	0.02	0.06	0.01	-0.03	0.01	0.00	-0.01	0.03	-0.02***	-0.02***	-0.02**
Canada	0.31***	0.66***	0.71***	0.69***	0.43***	0.48***	0.45***	0.49***	0.43***	0.23***	0.34***
Denmark	-0.03***	-0.06***	-0.04	-0.04*	-0.02***	-0.03***	-0.01**	-0.01**	-0.01**	-0.02***	-0.04***
Finland	0.25***	0.37***	0.34	0.22	0.42***	0.07**	0.06***	0.01	0.00	-0.04***	0.38***
France	-0.01	0.05	0.26	0.46	0.15**	0.04	0.05**	0.06	0.03	0.02	0.03**
Germany	2.34***	1.48***	8.79	2.28	3.76***	2.85***	2.95***	5.45***	5.90***	5.77***	2.69***
Ireland	0.07***	0.05***	0.06**	0.05*	0.06***	0.07***	0.07***	0.13***	0.08***	0.07***	0.06***
Italy	0.82***	0.71***	0.84	0.55	0.62***	0.57***	0.57***	0.62***	0.56***	0.81***	0.62***
Japan	3.11***	7.33***	6.93	7.53	7.93***	4.80***	3.69***	3.26***	4.61***	5.31***	5.64***
The Netherlands	0.23***	0.21***	0.18**	0.08	0.09***	0.08***	0.09***	0.24***	0.07***	0.07***	0.11***
New Zealand	0.30***	0.33***	0.26***	0.24***	0.23***	0.20***	0.25***	0.21***	0.16***	0.13***	0.20***
Norway	-0.03***	-0.05***	-0.04	-0.03*	-0.01**	-0.02***	0.00	0.00	0.00	0.00	-0.06***
Portugal	0.83***	0.89***	0.70	0.77	0.80***	0.32***	0.33***	0.22***	0.30***	0.21***	0.90***
South Africa	1.16***	1.71***	1.21***	1.68***	3.64***	1.33***	1.78***	0.97***	0.79***	0.68***	1.06***
Spain	-0.01	-0.03*	-0.02	-0.02	-0.03*	-0.01	-0.06***	-0.03***	-0.04***	-0.06***	-0.06***
Sweden	-0.03***	-0.07***	-0.06**	-0.06**	-0.05***	-0.07***	-0.06***	-0.04***	-0.05***	-0.04***	-0.04***
Switzerland	0.08***	0.25***	0.17	0.25**	0.13***	0.16***	0.16***	0.16***	0.13***	0.05***	0.11***
United Kingdom	0.16***	0.29***	0.16***	0.18***	0.17***	0.18***	0.16***	0.18***	0.11***	0.06***	0.14***
United States	0.94***	1.09***	0.77***	0.67***	0.55***	0.55***	0.54***	0.66***	0.55***	0.37***	0.48***

Key: Positive values (trimmed bootstrap superior) in black, negative values (trimmed bootstrap inferior) in red.

* significant at alpha = 0.10; ** significant at alpha = 0.05; *** significant at alpha = 0.01.

Table 11: Real data instances of trimmed bootstrap displaying stochastic dominance and almost stochastic dominance when compared to the other tested simulation methodologies. Wage growth of 0 per cent per annum

	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Stationary
Australia	0.002	0.001	0.01	0.001	0	0.013	0.009	0.02	0.003	0	0.002
Austria	0.004	0.005	0.003	0.006	0.017	0.024	0.021	0.002	0.001	0.004	0.003
Belgium	0.525	0.66	0.586	0.754	0.703	0.691	0.668	0.631	0.641	0.726	0.617
Canada	0	0.002	0.003	0	0.002	0	0.002	0.023	0.004	0.05	0
Denmark	0.57	0.661	0.59	0.599	0.571	0.61	0.528	0.499	0.503	0.567	0.589
Finland	0.651	0.583	0.527	0.614	0.504	0.564	0.488	0.533	0.513	0.579	0.546
France	0.483	0.497	0.509	0.503	0.513	0.56	0.51	0.504	0.504	0.508	0.51
Germany	0.014	0.005	0	0.001	0.001	0	0	0	0.005	0	0.001
Ireland	0.224	0.395	0.296	0.302	0.34	0.275	0.392	0.304	0.313	0.397	0.372
Italy	0.001	0.004	0.004	0.001	0.001	0	0.002	0.006	0	0.001	0
Japan	0.003	0	0	0.002	0	0	0	0.003	0	0.001	0.001
The Netherlands	0.242	0.29	0.27	0.341	0.295	0.458	0.338	0.344	0.341	0.358	0.299
New Zealand	0	0.004	0.005	0.004	0.001	0.006	0.004	0	0.005	0.004	0
Norway	0.638	0.613	0.567	0.545	0.515	0.572	0.452	0.427	0.406	0.447	0.65
Portugal	0.014	0	0.001	0.001	0.071	0.024	0.019	0.055	0.001	0.068	0
South Africa	0	0	0	0	0	0.001	0.002	0.001	0	0	0
Spain	0.796	0.672	0.715	0.697	0.683	0.666	0.676	0.666	0.642	0.748	0.815
Sweden	0.568	0.759	0.722	0.637	0.663	0.749	0.732	0.733	0.736	0.765	0.728
Switzerland	0.012	0.088	0.103	0.165	0.03	0.423	0.03	0.053	0.087	0.23	0.162
United Kingdom	0.003	0.025	0.149	0.136	0.101	0.141	0.229	0.14	0.31	0.317	0.101
United States	0	0.003	0	0	0.004	0.017	0	0	0	0.009	0

KEY: Stochastic Dominance | Almost Stochastic Dominance

The analysis presented in this study has, for the large part, illustrated that the trimmed bootstrap is a good technique for the majority of countries. However, it can underperform if future long-term returns surpass historical observations by a great extent. Although such returns may appear low probability events, Figure 5 shows that they did occur for a number of countries in the out-of-sample period 1974–2013. Given that future researchers will be looking to determine what the optimal simulation methodology for their particular market is, there is a need to provide such researchers with some guidance in their selection. Assistance is provided in two ways. First, we provide a snapshot of the out-of-sample testing results in an easy-to-follow Table 12, which provides the 'winning' technique for each country. Following this analysis, we describe in the subsequent section why the trimmed bootstrap is likely to be the best bootstrap method for the future.

Table 12 illustrates the 'best' technique (i.e., least MAE) for each country as suggested by the out-of-sample test results. The three different MAE analyses undertaken in this study, that is, the MAE for the nominal data with 4 per cent wage growth, the real returns with 2 per cent per annum wage growth, and the real returns with 0 per cent per annum wage growth, are presented here. This table provides researchers with the best nonparametric technique for their consideration using the out-of-sample results from 1974–2013.

However, an issue with the results in Table 12 is that the 'winning' methods come from examining data which is 40-years old. As we add to the number of observations on which the trimmed bootstrap is based, it is expected the likelihood of more-extreme results will reduce. The in-sample period from 1900 to 1973 means that the trimming of 40-year return paths used in the trimmed bootstrap tests was based on the distribution of just 35 rolling 40-year returns. However, by the end of 2013 there were 75 rolling

40-year returns available so the historic distribution of 40-year returns looks very different at the end of 2013 from how it looked at the end of 1973. These changed circumstances, addressed in the next section, provide researchers with the tools to make considered judgements about the appropriate bootstrap method to employ in each country. This analysis utilises a forward-looking approach employing the full dataset, which allows researchers to apply their own subjective views to determine which method is best for their particular market.

Table 12: Table of the best performing empirical simulation techniques from out-of-sample tests

	Nominal Data 4pc	Real Data 2pc	Real Data 0pc
Australia	Trimmed	Trimmed	Trimmed
Austria	Trimmed	Trimmed	Trimmed
Belgium	Trimmed	Block 4	Block 4
Canada	Trimmed	Trimmed	Trimmed
Denmark	Stationary	Block 2	Block 2
Finland	Trimmed	Block 10	Block 10
France	Trimmed	Block 1	Block 1
Germany	Trimmed*	Trimmed	Trimmed
Ireland	Stationary	Trimmed	Trimmed
Italy	Block 3	Trimmed	Trimmed
Japan	Trimmed	Trimmed	Trimmed
The Netherlands	Trimmed	Trimmed	Trimmed
New Zealand	Block 4	Trimmed	Trimmed
Norway	Stationary	Stationary	Stationary
Portugal	Trimmed	Trimmed	Trimmed
South Africa	Block 2	Trimmed	Trimmed
Spain	Stationary	Block 1	Block 1
Sweden	Block 3	Block 2	Block 2
Switzerland	Trimmed	Trimmed	Trimmed
United Kingdom	Block 4	Trimmed	Trimmed
United States	Trimmed	Trimmed	Trimmed

*Germany had no significant results, however, the trimmed bootstrap had smaller MAE against all other techniques

Forward-looking analysis

Determining the accuracy of the models using MAE in the previous section has a novel bonus. The methodology enables an analysis into which technique would be optimal for any level of retirement wealth outcome in the future. This analysis can be done by running each simulation technique on the entire dataset (that is, data from 1900–2013) to simulate the next 40-year period. These simulated outcomes can then be compared to a series of hypothetical outcomes to determine for what range of outcomes the trimmed bootstrap will be superior for each market. For example, we could assume the next 40-year actual terminal wealth outcome is 500,000 currency units, and then determine which technique in each market would be optimal (in the sense of having the lowest MAE) for this outcome when simulating using the full dataset and the retirement assumptions. This process can be repeated for a range of hypothetical terminal wealth outcomes, to provide information that can be used to guide researchers' choices of bootstrap technique.

Table 13 provides this forward-looking analysis, and examines incremental 100,000 currency unit hypothetical future outcomes from 0 to 2,000,000 units for the real returns case, with 0 per cent per annum wage growth. The currency unit outcome may be difficult to interpret; hence, for completeness the table also provides an annualised internal rate of return (IRR) estimate for each outcome. Further, to relate the results back to specific historical experiences within each country, the historical range from rolling real historical outcomes is placed over the table in the form of three lines under the same wage assumptions. The red line is the historical minimum outcome, the blue line represents the historical median outcome, and the dark green line signifies the historical maximum under the same set of assumptions. The crossover point (where the trimmed becomes inferior), along with the implied probability of this crossover, denoted Pr(W>Co) for each country, is presented at the bottom of the table. Note that by inferior we mean that at least one of the other 11 bootstraps has a lower MAE than the trimmed MAE (even if the difference between them is not significant). The implied probability is simply derived from the number of instances where historical real wealth outcomes were larger than this crossover amount in currency units.

Table 13 shows that the trimmed bootstrap displays dominance over all other techniques in every country in the DMS database if the terminal wealth outcome in 40-years falls at or below 700,000 currency units, or is less than a 6.39% real IRR. The methodology outperforms all other techniques to the upper-end of 2 million currency units, an equivalent internal rate of return of just over 10 per cent per annum, in seven countries (Finland, Germany, Japan, The Netherlands, Portugal, South Africa and Sweden).

With reference to the crossover point (where the trimmed was no longer superior to all other bootstraps), there were 13 countries which were beyond the historical maximum outcome. This suggests that the use of any other technique in these countries is forecasting a 40-year period that surpasses the best we have had in those countries in 114 years of history. This would be a difficult assumption to defend, and it is believed that the trimmed bootstrap should be the only technique employed in these markets. In the event that the future surpasses the historical upper extreme, this becomes a practical win for the plan member (having enjoyed the best 40-year return period in history) versus a theoretical research issue of not having utilised the best simulation technique.

The crossover point is also well above the median historical outcome in every case. In fact, the percentage of the historical rolling wealth outcomes that have exceeded the crossover point (Pr(W>Co)) can be seen, from the final row of Table 13, to range from 0 per cent for 13 countries to 14.9 per cent for Norway. Even in

Norway's case, choosing an alternative to the trimmed bootstrap implies that the researcher is expecting an outcome more extreme than has been observed 85.1 per cent of the time historically in Norway. Such a choice could only be justified if the researcher had some basis for expecting that the Norwegian market in the next 40 years would perform better than it has in most of the last 114 years. Without such a basis, the trimmed bootstrap would be the logical choice.

Overall, the trimmed bootstrap has been the superior estimator for the majority of likely outcomes throughout the world. The combination of Table 12 and Table 13 represents the main contribution to the literature in this paper. Table 13 in particular can be used to guide a researcher's choice of bootstrap technique. We trust the results of this study will ensure the widespread adoption of the trimmed bootstrap among pension finance researchers who conduct research using empirical stochastic simulations around the world.

Table 13: Winning simulation technique subject to MAE under a range of possible terminal wealth outcomes using real returns from 1900-2013 to simulate 40-year return paths under the 0 per cent wage growth assumption

Terminal Wealth	Australia	Austria	Belgium	Canada	Denmark	Finland	France	Germany	Ireland	Italy	Japan	The Netherlands	New Zealand	Norway	Portugal	South Africa	Spain	Sweden	Switzerland	United Kingdom	United States	IRR Equivalent
0	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	-
100,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	-1.96%
200,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	1.39%
300,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	3.12%
400,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	4.27%
500,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	5.13%
600,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	5.82%
700,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	6.39%
800,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	6.88%
900,000	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	7.31%
1,000,000	Trimmed	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	7.69%
1,100,000	Trimmed	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	8.03%
1,200,000	Trimmed	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	Trimmed	Block 10	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	8.34%
1,300,000	Trimmed	Trimmed	Block 9	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	8.62%
1,400,000	Block 10	Trimmed	Block 9	Trimmed	Block 10	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	8.88%
1,500,000	Block 10	Trimmed	Block 9	Trimmed	Stationary	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Block 10	Block 10	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	9.12%
1,600,000	Block 10	Trimmed	Block 9	Trimmed	Stationary	Trimmed	Block 10	Trimmed	Block 9	Block 10	Trimmed	Trimmed	Block 10	Block 5	Trimmed	Trimmed	Block 10	Trimmed	Trimmed	Trimmed	Trimmed	9.35%
1,700,000	Block 10	Block 10	Block 9	Trimmed	Stationary	Trimmed	Block 10	Trimmed	Block 9	Block 10	Trimmed	Trimmed	Block 10	Block 5	Trimmed	Trimmed	Block 10	Trimmed	Stationary	Block 10	Trimmed	9.56%
1,800,000	Block 10	Block 10	Block 9	Block 10	Stationary	Trimmed	Block 9	Trimmed	Block 9	Block 10	Trimmed	Trimmed	Block 10	Block 5	Trimmed	Trimmed	Block 10	Trimmed	Block 5	Block 10	Trimmed	9.76%
1,900,000	Block 10	Block 10	Block 9	Block 10	Stationary	Trimmed	Block 9	Trimmed	Block 9	Block 10	Trimmed	Trimmed	Block 10	Block 5	Trimmed	Trimmed	Block 10	Trimmed	Block 5	Block 10	Block 10	9.94%
2,000,000	Block 10	Block 10	Block 9	Block 5	Block 1	Trimmed	Block 9	Trimmed	Block 9	Block 10	Trimmed	Trimmed	Block 1	Block 5	Trimmed	Trimmed	Block 10	Trimmed	Block 5	Block 10	Block 10	10.12%
Cross over	1,365,000.00	1,610,000.00	942,000.00	1,778,000.00	944,000.00	4,050,000.00	1,083,000.00	7,000,000.00	1,568,000.00	1,223,000.00	2,850,000.00	2,030,000.00	1,160,000.00	752,000.00	3,000,000.00	2,220,000.00	840,000.00	2,600,000.00	1,680,000.00	1,675,000.00	1,815,000.00	10.29%
Pr(W>Co)	0.135	0.000	0.108	0.000	0.108	0.000	0.068	0.000	0.081	0.000	0.000	0.000	0.014	0.149	0.000	0.000	0.095	0.000	0.000	0.000	0.000	10.45%

Minimum Historical Outcome

Median Historical Outcome

Maximum Historical Outcome

5.5 Conclusion

We have used an out-of-sample test, to examine the trimmed bootstrap in 21 countries, employing both nominal and real return data. The results show that for the majority of countries, the trimmed bootstrap is the superior simulation technique. The instances where the trimmed bootstrap was found to be inferior experienced fundamental shifts in their political and economic prosperity between the in-sample and out-of-sample periods under analysis: shifts that are unlikely to occur in the future. This statement was supported by a forward-looking analysis which examined a range of likely outcomes, suggesting that the trimmed bootstrap would be superior in all countries which achieved less than 6.39% real IRR in the future.

This study corroborates the findings from our previous study and provides further evidence that the trimmed bootstrap technique is a superior technique for pension finance researchers. While a select number of countries did not record the trimmed bootstrap as a superior technique in the out-of-sample analysis, it is argued the reasons for this are less likely to occur in the future. Further, the study provides a summary table which provides pension finance researchers across the world with their particular market's best performing technique as suggested by the results of this study. The study also showed that if you select one of the alternative bootstrap methods over the trimmed bootstrap, then you are implicitly assuming that the next 40 years will be better than most of the 75 rolling 40-year outcomes observed over the last 114 years. Such an assumption may be hard to justify. It is hoped this information provides researchers with better tools to provide more accurate information to inform effective policy, and with more advice for plan members to maximise wealth in retirement.

References

Aisen, A., & Veiga, F. (2006). *Political Instability and Inflation Volatility Working Paper*. International Monetary Fund.

Antolin, P., Payet, S., & Yermo, J. (2010). Assessing Default Investment Strategies in Defined Contribution Pension Plans. *OECD Journal: Financial Market Trends, 2010*(1), 1-30. doi: http://www.oecd.org/finance/financial-markets/46010869.pdf

Bateman, H., & Kingston, G. (2012). The future of financial advice and MySuper. *JASSA*(3), 49-53.

Bayoumi, T. (1990). *Output, Employment and Financial Sanctions in South Africa*. International Monetary Fund Retrieved from https://www.imf.org/external/pubs/cat/longres.aspx?sk=3077.0.

Bodie, Z., Merton, R., & Samuelson, W. (1992). Labor supply flexibility and portfolio choice in a life cycle model *Journal of Economic Dynamics & Control, 16*, 427-449.

Bruno, S., & Chincarini, L. (2010). A historical examination of optimal real return portfolios for non-US investors. *Review of Financial Economics, 19*(4), 17.

Cairns, A., Blake, D., & Dowd, K. (2006). Stochastic lifestyling: Optimal dynamic asset allocation for defined contribution pension plans. *Journal of Economic Dynamics & Control, 30*, 843-877.

Dimson, E., Marsh, P., & Staunton, M. (2002). *Triumph of the Optimist: 101 years of global investment returns*: Princeton University Press.

Doran, B. & Bornholt, G. (2017). The Trimmed Bootstrap: An Empirical Simulation Technique for Pension Finance Researchers. *Journal of Retirement*, 4(3), 77-95.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics, 7*(1), 1-26.

Evans, L., Grimes, A., Wilkinson, B., & Teece, D. (1996). Economic reform in New Zealand 1984-95: The pursuit of efficiency. *Journal of Economic Literature, 34*(4), 46.

Howard, N. (2012). MySuper. *Taxation in Australia, 47*(6), 324-325.

Jacobs, D., & Rush, A. (2015). *Why is wage growth so low?* : RBA Bulletin - June Quarter Retrieved from http://www.rba.gov.au/publications/bulletin/2015/jun/pdf/bu-0615-2.pdf.

Johnson, V. D., & Dickinson, E. (2015). Internations Norms and the End of Apartheid in South Africa. *Safundi: The Journal of South African and American Studies, 16*(4), 355-377.

Kingston, G., & Thorp, S. (2005). Annuitization and asset allocation with HARA Utility. *Journal of Pension Economics and Finance, 4*(3), 225-248.

Künsch, H. R. (1989). The jacknife and the bootstrap for general stationary observations. *Annals of Statistics, 17*(3), 1217-1241.

Laverty, A. (2007). Impact of Economic and Political Sanctions on Apartheid. Retrieved March 28, 2016

Leshno, M., & Levy, H. (2002). Preferred by 'All' and Preferred by 'Most' Decision Makers: Almost Stochastic Dominance. *Management Science, 48*(8), 1074-1085.

MacDonald, K. L., Bianchi, R. J., & Drew, M. E. (2012). KiwiSaver and Retirement Adequacy. *Australiasian Accounting Business and Finance Journal, 6*(4), 61-78.

Martin-Acena, P., Martinez Ruiz, E., & Pons, M. (2012). War and Economics: Spanish civil war finances revisited. *European Review of Economic History, 16*(2), 144-165.

Mathers, C. D., Stevens, G. A., Boerma, T., White, R., A., & Tobias, M. (2015). Causes of international increases in older age life expectancy. *The Lancet, 385*(9967), 540-548.

OECD. (1994). *Economic Survey of New Zealand*. Paris.

OECD. (2013). Pensions at a Glance 2013: OECD and G20 Indicators. *OECD Publishing*. doi: http://dx.doi.org/10.1787/pension_glance-2013-en

Pfau, W. D., & Kitces, M. E. (2014). Reducing Retirement Risk with a Rising Equity Glide Path. *Journal of Financial Planning, 27*(1), 38-45.

Politis, D. N., & Romano, J. P. (1994). The Stationary Bootstrap. *Journal of American Statistical Association, 89*(428), 1303-1313.

Rashbrooke, G. (2009). Simple, effective and (relatively) inexpensive: New Zealand retirement provision in the international context. *Social Policy Journal of New Zealand*(36), 97-110.

Scutt, D. (Producer). (2016, August 1). ANZ: Low wage growth in Australia may be here to stay. *Business Insider Australia*.

Sweeting, P. (2009). Tax-Efficient Pension Choices in the UK. *Annals of Actuarial Science, 4*(2), 177-197.