Demographic Changes, Financial Markets, and the Economy

Robert D. Arnott[§]
Research Affiliates

Denis B. Chaves[†]
Research Affiliates

Forthcoming in the Financial Analysts Journal

[§] Robert Arnott is chairman of Research Affiliates, LLC; 620 Newport Center Drive, Suite 900, Newport Beach, CA 92660; web: http://www.rallc.com/; email: arnott@rallc.com.
[†] Denis Chaves is a senior researcher at Research Affiliates, LLC; 620 Newport Center Drive, Suite 900, Newport Beach, CA 92660; web: http://www.rallc.com/; email: chaves@rallc.com.

ABSTRACT

It seems natural that the shifting composition of a nation's population ought to influence GDP growth and perhaps also capital markets returns. As the baby boomers have aged, many people have studied past demographic data in an effort to extract indications for the future influence of the boomers on many aspects of the economy. We extend this body of literature by analyzing the effect of demographic changes on three measures of great importance for countries all over the world: real per capital PPP-adjusted GDP growth, stock market excess returns, and bond market excess returns.

We confirm what others have already demonstrated, but we extract markedly more statistical significance by adapting a polynomial curve-fitting technique pioneered by Fair and Dominguez (1991), to this new purpose. In our work, we find that a growing roster of young adults (age 15–49) is very good for GDP growth, a growing roster of older workers is a little bad for GDP growth, and a growing roster of young children or senior citizens is very bad for GDP growth.

We find surprisingly powerful results when we apply the same technique for exploring the links between demography and capital markets returns, net of the strong and well-documented effects of valuation and yield levels. Stocks perform best when the roster of people age 35–59 is particularly large, and when the roster of people age 45–64 is fast-growing. Bonds follow a similar pattern, with an age-shift: they're best when the roster of people age 50–69 is growing quickly. We carry out three different forms of robustness checks, each of which provides statistical significance in different ways: applying different country weights, testing alternative demographic variables, and confirming GDP results on out-of-sample countries.

It would be dangerous to forecast the future based on these results. Tacitly, we would be assuming that past relationships between demography and either GDP growth or capital market returns will hold unaltered in the future. However, given the high levels of statistical significance in the historical relationships, it is too tempting to resist exploring the possible implications for future GDP growth and capital market returns. These implications—with all the caveats that must necessarily be offered—are sobering, to say the least.

"Demography is Destiny."

19th Century, attributed to Auguste Comte

1. Introduction

Demography is one of the rare social sciences in which forecasts—at least for the short run—have startlingly little uncertainty. Today's 40-year-olds are next year's 41-year-olds. We can count them; we know the likely mortality for 40-year-olds and the likely rate of immigration and emigration of this age cadre. Looking 10 years into the future, there is significant wiggle room in the number of people under 10, some wiggle room in the number of people over 70, depending on the progress of medical science, and surprisingly little wiggle room in the number of people age 10-70, barring war, pestilence or other catastrophes. As the baby boomers have aged, many people have studied past demographic data in an effort to extract indications for the future influence of the boomers on many aspects of the economy, from housing prices to consumer preferences to retirement plans.

While the genesis for our own work has the same roots—curiosity about the potential impact of aging baby boomers—we decided to pursue a broader course of study, spanning decades of data and dozens of countries. We concentrated on three areas in which demographic shifts might influence the economy: real per capita GDP growth, stock market returns, and bond market returns.

We do not strive to extend the theory of demography in this paper, but rather apply new empirical techniques to the study of the effect of demographic changes on GDP and capital market returns. Our most important extensions of past work in this field lie in two areas. First, we sought to extract more statistical significance by looking at data from many countries over many years. Second, instead of fitting regressions against broad and *ad hoc* demographic cohorts, we fit a polynomial to the regression coefficients between demographic age groups and GDP growth, as well as stock and bond returns. The use of polynomials is intuitive, as it satisfies two important criteria: parsimony—only a small number of parameters are required—and continuity across age groups—behavior should change in a reasonably smooth way from one age cadre to the next.

Two core principles influence our research design.

First, we believe that models for GDP growth are less interesting than models for per capita, PPP-adjusted[1] real GDP growth. All three of these modifiers to GDP growth are important. After all, in a country with 3 percent population growth, 3 percent GDP growth means no growth at all for the average citizen, hence our reliance on per capita data. The same logic holds for 10 percent per capita GDP growth in a country with 10 percent inflation, hence our focus on real per capita GDP growth. The

[1] Purchasing Power Parity (PPP) adjusts income or GDP to reflect the cost of goods and services in an economy. If a good quality hotel room and an excellent dinner for four each costs $30 in Urumqi, China, and a similar room and dinner each costs $300 in Chicago, then a smaller GDP will buy more consumer goods and services in Urumqi than in Chicago.

PPP adjustment creates a fairer global comparison by focusing on the domestic purchasing power of the average citizen for the consumption basket that matters for each country.[2]

Second, stock and bond returns are measured as excess returns relative to domestic cash returns rather than as simple annualized returns. We do this for two very simple reasons. Stock and bond excess returns over cash can fairly be compared around the world because, for the currency-hedged investor, arbitrage equalizes the returns for domestic cash within relatively narrow bounds. Also, by looking at excess returns over domestic cash, we strip out inflation differences, crudely but reasonably effectively: cash yields rarely differ from domestic rates of inflation by more than a couple of percentage points.

Our approach leads to some simple and compelling demographic curves that conform nicely to our intuition about how people behave at various stages in their lives. This approach also delivers surprisingly tight confidence intervals surrounding these polynomial curves. What we learn is unsurprising, except in the statistical significance of our findings, and in their tacit implications for the years ahead.

As a robustness test, we run our core regressions—comparing demographic profiles with GDP growth and stock and bond excess returns—on the basis of both equal-weighting and GDP-weighting the countries in the regression. These results are almost identical, which gives us added confidence in the accuracy of our findings. We run our core regressions on both a country's demographic profile and the rate of change in its demographic profile, again using both weighting schemes. These results differ in a fashion that is intuitive: the rate of change models should lead the demographic profile models, and they do.[3] Finally, we also run our GDP regressions on a separate group of tier-two economies, again with very similar results.

WHY SHOULD DEMOGRAPHIC CHANGES AFFECT FINANCIAL MARKETS AND THE ECONOMY?

As a theoretical motivation, we break down the per capita total output of goods and services in an economy first by age groups and then as a product of per worker productivity and per capita number of workers:

$$\frac{\text{Total output}}{\text{population}} = \sum_j \text{productivity per worker in age group } j \cdot \frac{\text{number of workers in age group } j}{\text{population}}$$

We note that it seems tautological that a growing workforce—or even a growing population for this matter—should be better for GDP growth than a shrinking workforce. To isolate these effects, we focus on *per capita* GDP growth. In this construct, the relevant measure becomes the size of the working-age

[2] In other recent writings, we go so far as to suggest that GDP is a poor measure of prosperity. In Arnott (2011) we direct a spotlight on GDP *net of deficit spending* as a measure of "Structural GDP," and GDP *net of all government spending* as a measure of "Private Sector GDP." Both seem much more relevant than debt-laden top-line GDP. We consciously chose not to add this additional complication to our GDP measure. While we think it's a purer measure of national prosperity and growth, it is not yet accepted by the broad mainstream and, therefore, would trigger controversy that would distract from the main implications of our research.

[3] The rate of change of an age cadre must be positive before that age cadre can have above average weight in the population. This "lead effect" is evident in the results of our regression models.

cadre (generally, people age 20-60) as a percentage of the total population: If the working-age cadre is growing faster than the broad population, that should provide a tail-wind to per capita PPP-adjusted real GDP; if slower, then our GDP measure should face a headwind.[4]

This simple equation allows us to identify two channels through which demographic changes can influence a country's GDP. First, assume that productivity varies significantly across age group. As a large age group makes its way into the workforce, then into the more productive stage of its life, and subsequently into retirement, total output per capita should first increase and then decrease as a consequence. This effect is known as a "demographic dividend" in the demography literature. Imagine the waves created by a rock thrown into a lake and their paths as they move away from the point of contact.

Second, anecdotal evidence tells us that most entrepreneurs, inventors, and innovators are young adults. Nobel Prizes, for instance, are usually awarded to older scientist/researchers, but for contributions made years before when they were much younger. Kanazawa (2003) studies scientists, musicians, and painters and shows that productivity peaks at ages between 30 and 40—the only exception are authors, who tend to reach their peaks after the age of 40 but still before the age of 50. Therefore, we also expect to identify a higher increase in productivity across all age groups, and consequently in total output per capita, in countries with a relatively higher share of younger cadres.

Either way, we expect that per capita GDP growth is strongest in populations dominated by young adults, and in populations in which the young adult population is growing quickly.

We stress that we are using growth rates in output, as opposed to simply output, and for this reason the effects we identify should peak at ages relatively earlier than pre-retirement. A worker's contribution to total output likely peaks when she has more experience, but its contribution to *growth* in total output is highest when she is in the process of acquiring that experience.[5]

A theoretical analysis of the effects of demographic changes on financial markets is much more complicated and would require more space than we have available here. Not to mention that many excellent papers on this topic already exist. Therefore, and because our focus here is on trying to advance the empirical side of the literature, here we present some of the findings of such models and summarize their intuition.

The main argument made by skeptics of the demographic effects on financial markets is that rational and forward-looking agents would incorporate any slow-moving and predictable changes in age

[4] This observation has direct relevance to the developed world today. The working age cadre has been growing faster than the overall population for roughly the last 50 years, by ¼ to ½ percentage points per year. This reverses for most of the developed world for the coming 30 years. Shouldn't this fact, *alone*, reduce the natural per capita real GDP growth by about ½ percent to 1 percent per year relative to what we've learned to expect? This simple fact has been overlooked by much of the media, and academia, in their ruminations on our recent inability to match the growth rates of past decades.

[5] This is, of course, a sad acknowledgement for one of the authors, enduring his relentless decline in his late 50s and his increasing reliance on the innovative spirit of younger adults, like his coauthor, in his early 30s. Our results might suggest that one of the authors is saving and investing based on the surging GDP contribution of the other author. Neither author would entirely reject this interpretation!

distributions into their information set and act accordingly. As a consequence, age effects would be insignificant and in practice not observable. To address and dismantle this criticism, some authors (Abel [2003] and Geanakoplos, Magill, and Quinzii [2004], among others) develop models to explain how demography can affect stock and bond returns even in the face of fully informed, rational agents. The intuition is simple, as explained by the IMF (2004) in their World Economic Outlook: "This is because only living generations trade in financial markets at a point in time, meaning that differences in the demand and supply of financial assets—a reflection of differences in size across generations—cannot be arbitraged away ahead of time."

Brooks (2000), for instance, solves an overlapping generations (OLG) model with forward-looking agents, a risky asset and a riskless one-period bond. Different age cohorts trade with each other and during baby booms consumption is relatively higher, while savings is relatively lower, effectively pushing up returns on both stocks and bonds. During population busts, the opposite effect dominates. Moreover, agents shift their investments from stocks to bonds as they approach retirement.[6]

The characteristics of most of these models are easily summarized and reflect common wisdom. Young adults, often in the process of starting a family, will rarely be major contributors to the quest for savings, investments, and capital accumulation.[7] As they look past their and their children's immediate needs to their own eventual retirements, they begin to invest, first in stocks, then in bonds. As they slide into retirement, they begin selling assets in order to buy goods and services that they no longer produce, either directly on their own investments, or indirectly through their pension benefits. They tend to liquidate their riskiest assets, the stocks, before their less-risky bond assets.

LITERATURE REVIEW

Analyzing the relationship between demographic changes and the economy, whether qualitatively or quantitatively, has been a topic of interest for centuries.[8] Accordingly, summarizing all the relevant work would be an impossible task. We focus on what is current in the literature, with an emphasis on empirical tests, comparing the existing literature with our own application of new methods to these data.

One of the first studies linking demography and financial assets is Mankiw and Weil (1989). They show a strong relationship between the baby boom in the United States, the subsequent increase in housing demand in the early 1970s, and a substantial impact on housing prices over the following 20 years. Their

[6] His simulations show effects on the order of 14–39 basis points per year; our empirical estimates and forecasts are materially stronger than this. Nonetheless, there is also a large literature on the equity premium puzzle, summarized by Mehra (2008), showing that market frictions, and in particular borrowing constraints as in Constantinides, Donaldson, and Mehra (2002), can exacerbate returns on risky and long-term securities.

[7] Nor should they. They typically will have far more human capital than investment capital. Why invest at a young age? Or, more provocatively, why *not* borrow against the future human capital in order to smooth out the lifetime consumption expectations? Sadly, we seem not to shake this pattern as we age, continuing to borrow and spend, even as our human capital dwindles.

[8] In 1798 Thomas Malthus published the first of six editions of his famous treatise: *An Essay on the Principle of Population*. He argued that the combination of linear growth in food production with geometrical growth in population would result in famine. His work is seen by many as discredited, based on two centuries of rising population and rising *per capita* health and lifestyles. Still, we should acknowledge that his work was the first serious examination of the connections between demography and health or wealth. And there are many who view him as merely being early—very early—in his prognostications.

forecast that housing prices would decline sharply after 1990, following the baby bust, did not materialize on schedule, though demographic effects may have magnified the recent collapse of the housing bubble.[9]

Arnott and Casscells (2003, 2004) discuss the demographic changes that will occur in the United States in the coming decades, and explore their implications to capital markets and possible solutions. The important finding in their work is the observation that the entitlements problem is not a financial problem exacerbated by a failure to prefund these obligations. Rather, it is a support ratio problem, tied to demography pure and simple: prefunding does not create—in advance—the goods and services that will change hands. Regardless of prefunding, goods and services must still change hands, from those who produce the goods and services to those who no longer do so. This work also speculates on the likely implications for capital market returns, consistent with results that have been seen subsequently in the United States, Japan, and parts of Western Europe. Finally, they explore and largely dismiss the common arguments that immigration or productivity gains can offset the pressures associated with pending demographic shifts.

Bosworth, Bryant, and Burtless (2004) survey the literature on analyzing and forecasting prices (returns) of financial assets and identify two approaches. The first one uses microeconomic data on goods or asset holdings,[10] together with forecasts of age distributions, to predict how future demand and prices (returns) will evolve (e.g., Poterba [2001]). DellaVigna and Pollet (2007), for instance, define what they call "age-sensitive sectors, such as toys, bicycles, beer, life insurance, and nursing homes" and estimate how the demand for products or services in these sectors will de/increase. They find that their demand forecasts predict both profitability and stock returns by industry five to ten years in the future. The second approach, followed in this paper and also surveyed by Davis and Li (2003), estimates the direct time-series relationship between prices (returns) and demographic variables. Notable examples in this strand include, but are not limited to, Yoo (1994), Bakshi and Chen (1994), Lindh and Malmberg (1999), Ang and Maddaloni (2003), Goyal (2004), and DellaVigna and Pollet (2007).

Yoo (1994) estimates multivariate time-series regressions of annual U.S. stock, corporate, and government bond returns on shares of total population for age groups 25–34, 35–44, 45–54, 55–64, and 65+. His strongest results are a negative relationship between short- and medium-term government bonds and age group 45–54. However, the statistical significance is weak for almost all coefficients in all five age groups. He also estimates the regressions with three- and five-year centered moving averages and finds a significant increase both in terms of statistical significance and fit, supporting our claim that long horizons provide a better test for low frequency population changes.

Bakshi and Chen (1994) propose a hypothesis that relative risk aversion is positively correlated with age, and model the utility function of the representative consumer as a function of aggregate consumption

[9] Green and Hendershott (1996) object to Mankiw and Weil's finding that housing demand starts declining after the age of 35. Researchers and commentators have also pointed to a series of possible causes for the surge in housing prices in the last few decades that preceded the current crisis: loose monetary policy by the Fed, government housing policy, lack of proper screening by lenders, or a mismanagement of financial risks spread by the use of derivatives (CDOs, CDSs, etc.). We also note, among our friends that demand for improved housing often continues well into one's 50s.
[10] The Survey of Consumer Finances is a good example.

and average age of the population. Their tests use Euler equations as well as a two-factor model based on consumption growth and percentage change in average age.[11] They find strong support for their life-cycle risk aversion hypothesis and a positive and statistically strong relationship between U.S. stock excess returns and growth in the average U.S. population age.

Ang and Maddaloni (2003) use both a long sample (1900–2001) with five countries and a short one (1970–2000) with 15 countries to study the relationship between excess stock returns (at one-, two-, and five-year horizons) and log changes in the following demographic variables: average age of the population over 20 years old, fraction of adults over 65 years old, and percentage of people in the 20–64 age group. In pooled regressions, their results display a strong and *negative* effect for the fraction of retirees in the population (65+). Interestingly, the authors find an opposite and *positive* result in isolated regressions for the United States and the United Kingdom. To explain this difference in results, they conduct further tests and show that the effect of the 65+ age group is stronger in countries with well-developed social security systems and less developed financial markets. Finally, and most important for this paper, they also show that "pooling data from five countries gives almost the same power as increasing the sample size of the United States by five times," validating the strong results we find here using an extended sample of countries.

Lindh and Malmberg (1999), on the other hand, study the effects of log age group shares (15–29, 30–49, 50–64, and 65+) on five-year growth rates in GDP per worker using a sample of OECD countries (1950–1990). They find significant and *positive* coefficients for the age group 50–64, and significant and *negative* coefficients for retirees (65+). When talking about their choice of age groups, the authors make the following comment: "The youngest age group, children aged 0–14, had to be dropped in order to avoid high degrees of linear dependency among the age variables. Some arbitrariness in the definition of the age group variables cannot be avoided." These two points highlight the need for an approach that uses all available information and that defines the demographic variables less arbitrarily and more systematically. Indeed, given our findings, we wonder if any study that ignores young people could be important.

Our most important identification tool, and most significant improvement over previous papers, is our use of the econometric methodology pioneered by Fair and Dominguez (1991). They use a polynomial to analyze the relationship between a changing U.S. age distribution and economic variables such as consumption, housing investment, money demand, and labor force participation. This methodology consists in forcing the regression coefficients on age groups into a polynomial and has at least two advantages. First, it allows us to include *all age groups* in the regression while avoiding the statistical issues created by the high correlation between them. Second, the interpretation of the results is more intuitive, as will become clear later. Here we closely follow Higgings (1998), who studies the effect of demographic changes on savings, investment, and the current account balance, but we analyze the implications for financial markets and economic growth.

[11] See Cochrane (2005) for an introduction to consumption-based asset pricing, Euler equations, and factor models.

In the empirical arena, one of the strongest criticisms comes from Poterba (2001). His main point is that "statistical tests based on the few effective degrees of freedom" lack the power to "find robust evidence of such relationships in the time series data." Because we *agree* with this statement, and recognize that demographic variables are both persistent and slow-moving in nature, we take a number of steps to increase the power of our tests. Some of our remedies include:

a) Relying on *five-year* rates of GDP growth and capital market returns, instead of *annual* data;
b) Controlling for starting valuation levels, GDP levels, and business cycle measures so that the demographic effects can be viewed in isolation from these other powerful effects;
c) Using a large cross section of countries; and
d) Including the information from all age groups in the regressions, as mentioned above.

The result is a substantial improvement in the statistical significance of our findings, as compared with previous studies of demographic effects on the capital markets or on GDP growth.

2. Data and variables

We draw data from many sources. The demographic and GDP data is deepest, with demographic age profiles and GDP available from the United Nations (UN) and the Penn World Table on well over 200 countries, typically well-documented back to 1950. The stock and bond data is solid for only the developed and largest emerging economies, with comparatively thin data before 1970. Still, we can extract over 200 observations in 22 countries for our main regressions, and over 1,600 observations in 176 countries for our extended group of countries, using non-overlapping five-year returns or growth rates.

Population data by five-year age cadres come from the UN's Population Division.[12] For most countries it starts in 1950, continues in five-year intervals, and includes projections until 2050. For future years, a total of eight different variants are available to forecast combinations of trends in fertility, mortality, and international migration. For our forecasts for future GDP growth and capital market excess returns, we choose the medium variant available online, which assumes a mid-range expectation for fertility and normal trends on mortality and migration.[13]

Financial variables come from a combination of different sources, but mostly from Global Financial Data. Total return indexes for stocks are complemented using data from Claus Parum[14] (Denmark), the Central Statistics Office[15] (Ireland), and Wydler (1989) (Switzerland). Ten-year yields, when missing, are proxied by five- or seven-year yields, while any missing three-month yields are proxied by central bank discount rates.

[12] http://www.un.org/esa/population/unpop.htm.
[13] Some other possible variants include low or high fertility trends, mortality rates constant at 2005–2010 levels and zero migration as of 2005–2010.
[14] http://staff.cbs.dk/parum/.
[15] http://www.cso.ie/.

Measures of per capita and PPP-adjusted GDP are available from the Penn World Table (Heston, Summers, and Aten [2009]). Version 6.3 includes annual observations from 1950 until 2007 and assigns a quality grade ranging from "A" (best) to "D" (worst) to each country "to signal the relative reliability of the estimates." (See Johnson, et al., (2009) for a discussion of problems and concerns about this dataset.) Data availability and reliability—missing observations originate mainly from the financial variables—restrict our main sample to 22 developed countries, of which 16 have a quality grade of "A" and 6 have a grade of "B."[16] However, given the broad coverage of the population and GDP data, we are able to execute robustness checks—based on GDP growth only, not stock or bond excess returns—for a sample of roughly 175 additional countries.

Our dependent variables are annualized five-year non-overlapping growth rates, denoted by $r_{i,t}$ and identified individually when necessary. Stock and bond returns are in excess of the domestic (same-country) bill return. Growth in PPP-adjusted real GDP per capita measures economic activity as the average citizen might perceive it. Unlike most other papers, we choose five-year returns mainly for two reasons: demographic data are more widely available in five-year intervals, and demographic changes occur slowly, giving low frequency data a better shot at identifying the effects of interest for this paper. The intersection of data sources leaves us with approximately 200–250 non-overlapping observations in our main tests, and over 1,600 in our robustness tests. As it happens, this broad database turns out to provide ample degrees of freedom, delivering surprising statistical significance for our results.

The explanatory variables of interest are either the percentage of total population by age group, $s_{i,t}^{(j)}$, or changes thereof $\Delta s_{i,t}^{(j)}$.[17] The age groups range from 0–4, $s_{i,t}^{0-4}$, through 70+, $s_{i,t}^{70+}$, yielding a total of 15 five-year demographic variables. We use dividend yields, $DY_{i,t}$, three-month yields, $3M_{i,t}$, and 10-year yields, $10Y_{i,t}$, as control variables in our regressions. Taking into account the initial valuation (or stage of the business cycle) in our regressions is vital, given that temporary swings in prices or economic activity might interfere with the estimation. For example, if stock market returns are high, inclusive of an initial dividend yield that is high, we could not have confidence that we are extracting the demographic effects separately from well-known valuation effects. The demographic results are relatively independent of the valuation measure one chooses. For instance, in our robustness checks we are forced to use the log of the ratio between consumption and GDP, $\log(C_{i,t}/\mathrm{GDP}_{i,t})$, given that interest rates are simply not available for most countries.

These data are deep enough, in our view, to draw some important conclusions about *past* linkages between demography and GDP growth and the capital markets. Applying these results to *predict* the prospective influence of demography on the economy or on the capital markets is a bit riskier: past is not prologue. Nonetheless, while our confidence in the forward linkages is less than our confidence in the past linkages, the potential implications of these results are sobering.

[16] Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, United Kingdom, and United States.

[17] Tacitly, by looking at shares of the population, which must always sum to 1.0, and changes in shares, which must always sum to 0.0, we choose to ignore overall growth rates. Our assumption is, therefore, that overall population growth, while it may materially affect overall real GDP growth, will not affect *per capita* real GDP growth. As an untested hypothesis, this is worth exploring in future research.

3. Results and Discussion

Ideally, one would like to estimate the *joint* effect of all 15 age groups $s_{i,t}^{(j)}$, $j \in \{0\text{–}4, \ldots, 70+\}$, or their changes $\Delta s_{i,t}^{(j)}$,on stock and bond excess returns (or growth in GDP per capita):

$$r_{i,t} = a + \gamma X_{i,t-1} + b_1 s_{i,t}^{(1)} + \cdots + b_N s_{i,t}^{(N)} + \varepsilon_{i,t}, \tag{1}$$

where $X_{i,t-1}$ represents control variables, such as interest rates or valuation ratios. The problem with this approach is that the demographic variables are highly correlated and would generate the usual multicollinearity problems.[18] Moreover, in our case, the estimation of Equation (1) is *impossible!* Because the maximum number of non-overlapping five-year observations for each country (12 in 60 years) is less than the number of demographic regressors (15 age groups), the covariance matrix of $s_{i,t}^{(j)}$ is singular.[19] The usual solution in the literature is to include only a limited number of broad age groups or to combine them in some *ad hoc* way—introducing a risk of data-mining—or to pool multiple countries, which still leads to an unwieldy regression, with far too many independent variables.

We choose a different approach.

Our approach, following Fair and Dominguez (1991) and Higgings (1998), forces the demographic coefficients, b_j, to satisfy a polynomial of order k, allowing us to incorporate the information available in the demographic profile. The details of the methodology are in the Appendix. After imposing this restriction, we obtain k transformed demographic variables, $\gamma_{i,t}^{(j)}$, whose coefficients, D_j, can be estimated with:[20]

$$r_{i,t} = a + \lambda X_{i,t-1} + D_1 \gamma_{i,t}^{(1)} + D_2 \gamma_{i,t}^{(2)} + \cdots + D_k \gamma_{i,t}^{(k)} + \varepsilon_{i,t}. \tag{2}$$

Because k ranges between two and four (see below), this approach provides a reduction in the number of demographic variables. Equation (2) and its k polynomial coefficients provide the ingredients for regression diagnosis. The 15 implied coefficients by age group in Equation (1) convey all the economic intuition.

Note that unlike the control variables, $X_{i,t-1}$, the age group shares, $s_{i,t}^{(j)}$, and their changes, $\Delta s_{i,t}^{(j)}$, are concurrent to the dependent variables. This timing is important because we are exploring demographic effects on financial assets, which presumably occur mainly via shifts in the demand for financial assets.

[18] See Greene (2008), for instance, for a detailed discussion of multicollinearity problems.

[19] Many of our readers are probably familiar with this issue in the context of portfolio optimization. The availability of data relative to the sheer number of assets in the investment universe renders the estimation and inversion of covariance matrices a challenge. A common solution is to adopt a factor structure on the asset returns or their covariance matrix, which is akin in spirit to the solution used here, because both impose a limit on the number of parameters to be estimated.

[20] A polynomial of order k is completely specified with $k+1$ parameters, but because the demographic shares add up to a constant, $\sum_{i=1}^{15} s_t^i = 1$ or $\sum_{i=1}^{15} \Delta s_t^i = 0$, the methodology imposes the extra restriction $\sum_{i=1}^{N} b_i = 0$ to avoid the perfect correlation with the constant term in the regression. See the Appendix for details.

The same logic applies to our concurrent timing on GDP and demographic variables, because any demographic effects on GDP will presumably occur mainly via changes in the active population and their relative contribution to GDP. Moreover, this data *can* be concurrent because we know the demographic profile for each country over the next 5 to 10 years with some precision, especially for the future working-age and retirement-age cadres, who we can simply count today.

The choice of the polynomial degree for the demographic variables in each case strikes a compromise between parsimony and statistical power. We recognize that searching over a large number of options is meaningless, because larger values of k provide no reduction in dimensionality, and also create a concern for data mining. For these reasons, we entertain only three options: polynomials of order k equal to two, three, or four. In other words, we limit our tests to quadratic, cubic, and quartic polynomials. Because models with lower degrees are special cases of their high-degree counterparts, we rely on Wald tests for nested models to gauge which order of polynomial is most appropriate.

For simplicity and parsimony, in each regression we include only one control variable, $X_{i,t-1}$, out of the following three: dividend yield, 10-year yield, and 3-month yield. These variables play a central role in many papers as forecasters of stock and bond returns and of economic growth (e.g., Campbell [1987], Keim and Stambaugh [1986], and Ang, Piazzesi, and Wei [2006]). The demographic coefficients, the main point of interest for *this* paper, are relatively invariant to the choice of controls. Other alternatives, such as yield spreads, produce very similar results because demographic effects are *low-frequency* by nature and any variables that reflect medium-term movements in prices or economic cycles are effective here, wholly consistent with past research.

Some skeptics might argue that other factors not included in the regressions can contribute to growth in per capita GDP or returns of stocks or bonds. Our first answer is that comprehensive data for other variables—both over time and across countries—practically doesn't exist. Second, the financial yields we include in the regressions serve as controls or proxies for some of these variables, because lower initial prices reflect higher risk aversion or higher risk premiums, among others. Third, the effects we capture have a low frequency by construction, making the argument for high-frequency variables, such as liquidity, less plausible. Finally, we note that missing variables always have the potential to alter the results of any regression, but the different robustness checks we run make us confident that our results are not spurious.

We present a total of six sets of results for combinations of the three dependent variables—excess stock and bond returns and growth in GDP per capita—and the two demography regressors—shares of total population and their changes.

DEMOGRAPHIC SHARES

We start with the first set of results, using shares of total population by age group, $s_{i,t}^{(j)}$, as the regressors. **Table 1** presents the results from Equation (2) and **Figure 1** plots the *implied* polynomial coefficients for each of the age cadres relative to Equation (1). The first three columns, labeled "Demographic Shares," show coefficients and t-statistics (in parentheses) for three regressions using different dependent variables: five-year stock and bond excess returns, and five-year growth in GDP per capita. All regressions are estimated using pooled ordinary least squares (OLS) and include country dummies (or fixed effects) to control for any unobserved characteristics particular to each country. Standard errors are corrected for heteroskedasticity and cross-sectional correlation (time clusters).[21]

The coefficients on the dividend yield and 10-year yield in the first two regressions are positive and statistically significant, as is the case in most of the literature: higher yields signal high risk aversion, depressed prices, and higher expected returns for stocks and bonds. In the third regression we find a negative and statistically significant relationship between GDP growth and the three-month yield. The intuition for this relationship is simple. The Federal Reserve keeps the short yield at lower levels during recessions, which are inevitably followed by periods of faster growth in GDP.

We start the discussion of the demographic results with the choice of k, the polynomial degree. The last two rows of Table 1 present *p*-values of Wald test comparisons between cubic and quadratic polynomials ($k = 3 \rightarrow k = 2$), and between quartic and cubic versions ($k = 4 \rightarrow k = 3$). In the regression for bonds, we can see that a polynomial of fourth degree is the best choice because both tests strongly reject the null of statistical equivalence between the nested models. In the case of stocks, the choice is a cubic polynomial because the tests cannot reject a difference between $k = 3$ and $k = 4$ (*p*-value of 28 percent). In the case of GDP growth, a parabola is the best choice because both *p*-values are above 10 percent.

The fit of all three regressions is relatively strong, with R^2s in the order of 30 percent. The statistical significance of the demographic polynomial coefficients is usually not as high as those for the financial variables, yet most of the *t*-statistics are close to or above three. This is expected because most of the variation in prices and economic activity occurs at a medium frequency and is captured by the yields. Given the difficulty in evaluating the economic significance and intuition of the demographic effects in Table 1, we plot the implied coefficients for each age group in Figure 1, along with 90 percent two-sided confidence intervals (shaded areas).

The graphs in Figure 1 and **Figure 2** deserve some explanation. Consider the first graph in Figure 1. The solid line shows the implied coefficient linking the size of each five-year demographic age cadre in the population, as an independent variable, with the corresponding stock market excess return for the *concurrent* five-year span in the same country, as our dependent variable. We see a negative link for the first six age cadres, from age 0–29, with statistical significance at the 90 percent level for the three age cadres from age 5–19. This means that, excluding any valuation effects based on the starting dividend yields, a larger share of the population in these age cadres historically conforms to lower stock market returns (excess returns over cash). Note that this polynomial hits a trough, with a coefficient of –0.7 for the 10–14 age cadre. This means that a 1 percent higher concentration in this age cadre results in annual stock market returns which are 0.7 percent lower. That's a big number for 1 percent change in

[21] Because the dependent variables are non-overlapping returns and growth rates, we believe cross-sectional correlation is the most important deviation from traditional OLS assumptions. We have also tried corrections for time-series correlation (country clusters) and both time and country clusters. These resulted in small and occasional differences, but the qualitative conclusions remained unchanged.

the size of one five-year age cadre. These effects are economically meaningful, especially considering how fast the age structure is changing in most developed countries and that these results reflect annual returns for a five-year span. Note that the coefficient for the age cohort 70+ is slightly worse than negative 1.5, which is important in some of the faster-aging countries.

OLS estimation assigns the same weight to all countries in the sample. To make sure our conclusions hold unaltered in larger countries—the United States, for instance, has PPP-adjusted GDP that is roughly 100 times that of New Zealand—we estimate the same regressions using pooled weighted least squares (WLS). The natural weighting choice—weighting by GDP—creates the opposite imbalance, however.

Using the reciprocal of the Herfindahl Index (HI) as a measure of the "effective" number of observations in a dataset, we find that our sample shrinks from 22 equally weighted countries to only about 5 GDP-weighted countries.[22] For this reason we choose to weight the regressions by square root of PPP-adjusted GDP, measured in 2005 international dollars. This choice represents a balance and still leaves us with 14 "effective" countries. The dashed lines show the results for these regressions. It is reassuring that every one of these lines on Figures 1 and 2 lies well within the confidence bands for the equal-weight regressions. This is our first important robustness test, which helps to confirm the merits of our results.

As we compare the three graphs in Figure 1, the most striking fact about the polynomials is the similarity between the demographic effects on stocks and bonds and, at the same time, how different they are from the demographic effects on per capita GDP growth. While larger groups in the early 20s are clearly beneficial for economic growth, they hurt the performance of stocks and bonds. This may be due to the relatively lower savings and higher debt levels of young adults, perhaps in an attempt to smooth consumption over their lives. Perhaps young adults are more rational than the older generation thinks! The opposite holds for the middle-aged groups in their 50s, affecting GDP growth negatively (already, even before they stop working!) and financial assets positively: it would appear that productivity starts to decline and savings to accelerate, long before retirement.

One common theme emerges from all three graphs: large populations of retirees (65+) seem to erode financial markets performance as well as economic growth. This makes perfect sense: retirees are disinvesting, in order to buy goods and services that they no longer produce, and they are no longer contributing goods and services into the macro economy. This effect is less pronounced for bonds, presumably as these are sold later in retirement than stocks, conforming to widespread popular financial advice.

Using the age group polynomials to estimate demographic effects has several benefits. The first one, as mentioned before, is the possibility to estimate the *joint* effect of *all* demographic variables in the

[22] The Herfindahl Index is calculated as the weighted average of individual weights, or $\mathrm{HI} = \sum_{i=1}^{N} w_i^2$. It is widely used as a measure of competition among firms within a country, market, or industry, and ranges between *1/N* (perfect competition) and *1* (monopoly). The reciprocal of this index, ranging between *N* to *1*, serves as an indicator of the hypothetical number of equally powerful firms in a country, market, or industry. If two countries have equal weight, HI equals 0.5, so 1/HI equals 2 … with equal weighting, our effective number of countries matches the actual number. If two countries are weighted 90/10, the former utterly dominates the regression, and HI equals 0.82 … so, we have effectively only 1.2 (1/HI) countries in our regression.

regressions. This allows us to make use of all the available information in the age distributions, even after imposing a perhaps restrictive polynomial structure on the coefficients. Second, instead of using *ad hoc* age group selections as others have done (ages 35–45 for instance), the resulting polynomial tells us exactly how to select and weight the age groups in order to obtain the most powerful explanatory variables.

One concern that afflicts other demographic studies is the persistence of—some might even suggest a risk of non-stationarity in—$s_{i,t}^{(j)}$, the size of each age cadre as a percentage of the population. One answer is that these variables are limited to the interval [0,1], but we also note that:

a) We use 60 years of data for most countries. Although this might not look like much in terms of demography, significant changes in population structure occurred in most countries during the period from 1950–2010. Indeed, Japan has changed from being the *youngest* country in the developed world to the *oldest*, with its median age *doubling* from 22 to 44 in just 60 years.
b) With a large cross section of countries, we increase the power in our tests by exploring the varying stages in the demographic transition process experienced by different nations.[23]
c) The polynomial curves show very significant differences between the implied coefficients for various age groups—the coefficients on younger and older cohorts, for instance, have opposing signs. This is a strong indication that the empirical methodology also explores variation across the different age groups—$s_{i,t}^{(m)}$ versus $s_{i,t}^{(n)}$—which are arguably less persistent and more volatile.

In the next section we repeat the same analysis using changes in the demographic shares of each age group, $\Delta s_{i,t}^{(j)}$, which exhibit more independence, both relative to neighboring five-year age cadres and relative to prior and subsequent five-year spans than the demographic shares, $s_{i,t}^{(j)}$.[24] We find similar results using these alternative regressors, which largely avoid any risk of non-stationarity. We hope that these results will alleviate any possible remaining criticisms.

CHANGES IN DEMOGRAPHIC SHARES

As a robustness test, we look at changes in the demographic shares. The last three columns in Table 1, labeled "Changes in Demographic Shares," along with the graphs in Figure 2, present the results for the rates of change in each demographic variable, $\Delta s_{i,t}^{(j)}$.

The regressions for stocks and bonds have R^2s that are similar to those presented earlier, but we see a decrease to 17 percent in the case of GDP growth. The degree of the polynomial is the same for bond

[23] Skeptics might still argue that the countries in our sample do not provide independent tests because they are mostly based on developed countries and are, therefore, highly correlated. We answer this point in two ways. First, we refer the readers to Ang and Maddaloni (2003), who show that using a sample of five countries increases the power of their tests by almost the same amount as augmenting their 20th century U.S. sample by a factor of five. Second, we extend our test for GDP growth to other less developed countries in the section on robustness checks, and find relatively similar but noisier results (unsurprising, given the lower quality of the data).

[24] Consider that a large age cadre ($s_{i,t}^{(j)}$ larger than average) can be associated with positive change ($\Delta s_{i,t}^{(j)}>0$), if the age cadre is increasing to its peak, or with negative change ($\Delta s_{i,t}^{(j)}<0$) if the age cadre is fading from its peak.

returns and GDP growth, but for stock returns the Wald test strongly favors a quartic over a cubic curve (*p*-value of zero). The magnitude and statistical significance of the financial variables are almost the same as in the previous tests.

Comparing the demography coefficients with those from the previous section (first three columns in Table 1) reveals some interesting facts. The demographic coefficients are similar in proportion across the two regressions for GDP growth, but noticeably different in the case of stock and bond returns. To understand whether these changes affect our intuition about the demographic effects, we need to look at the implied coefficients in Figure 2. The magnitudes of the coefficients have increased roughly by a factor of two, reflecting the lower volatility of $\Delta s_{i,t}^{(j)}$, when compared to $s_{i,t}^{(j)}$. The polynomials for stocks and bonds have the same pattern as in the previous section, but seem to shift to the right. This makes intuitive sense because the peak rate of change in the size of a given age cadre will lead the peak *size* of that cadre ($\Delta s_{i,t}^{(j)}$ leads $s_{i,t}^{(j)}$), in most cases by perhaps a decade or more. Thus, the polynomial curves should exhibit peaks at later ages for $\Delta s_{i,t}^{(j)}$ than for $s_{i,t}^{(j)}$.

Out-of-Sample Robustness Check

A third and more conventional set of robustness checks is to test our hypothesis on a new set with *different* countries. Unfortunately, outside of the developed markets, there is little reliable financial data beyond what is used in the previous sections. Nevertheless, the two main GDP datasets—the UN (population) and Penn World Table (national income accounts)—span about 200 countries. For this reason, the tests here are restricted to economic activity only. Stock and bond market data is simply unavailable or too sparse to be useful.

To get around the lack of financial yields as control variables, we use the log of the ratio between consumption and GDP, $\log(C_{i,t-1}/GDP_{i,t-1})$, as a proxy for business cycles. Cochrane (1994) shows that this ratio forecasts GDP growth because consumption is nearly a random walk. Finally, there is also the important issue regarding the quality of the data for the remaining countries not used in our main tests. Johnson, et al. (2009) show that the vast majority of these countries have a quality grade of “C” or “D” (worst), as assigned by the authors of the Penn World Table, and that the standard errors of the estimates and the likelihood of revisions are much larger for the smaller countries and economies. For this reason, we should expect wider confidence bands.

Many of these countries are also very small, with unreliable data and undiversified economies. For this reason, we follow our earlier approach and estimate the regressions in this section using pooled WLS, where each country is again weighted by the square root of their 2005 GDP, measured in international dollars (PPP-adjusted). This approach assigns more weight to larger economies and provides more accurate estimates, without reducing the small countries to irrelevance.[25]

[25] If we use equal-weighting, China has the same weighting as Kiribati. So, even though we have 175 countries, we have no confidence that our models will have relevance for the major markets. If we use GDP-weighting, then the four BRIC countries comprise very nearly half of the sample. The “effective” number of countries is about 16 because of the heavy reliance on the biggest countries. An inviting compromise is to split the difference, with square root of GDP weighting. China’s GDP is 100 times

Table 2 shows that this sample includes 175 *new* countries and over 1,600 observations. The dependent variable—non-overlapping five-year growth rate in GDP per capita—and demographic variables—either $s_{i,t}^{(j)}$ or $\Delta s_{i,t}^{(j)}$—are still the same as in the previous sections, but now we add $\log(C_{i,t-1}/GDP_{i,t-1})$ as a substitute for $DY_{i,t}$ and $3M_{i,t}$, both of which are unavailable in most of these smaller economies. The statistical fit, as expected, is much lower, with R^2s at 3 and 4 percent. The Wald tests recommend a parabola and a cubic polynomial for $s_{i,t}^{(j)}$ and $\Delta s_{i,t}^{(j)}$, respectively.

In the regression using $s_{i,t}^{(j)}$ in the first column, the demographic coefficients D_1 and D_2 are the same as their counterparts in Table 1, up to two decimal places. This is also confirmed by **Figure 3**, where we plot the corresponding GDP polynomials from our primary tests (the bottom charts on Figures 1 and 2) as a dotted line. The first graph depicts a remarkable similarity between the two parabolas, showing that our results using demographic shares as forecasters also hold in a much larger sample.

The second regression has demographic results that differ from our main results, so we rely on the implied coefficients profile in Figure 3 to visually confirm a connection with the earlier results. The general curve is broadly similar, though less significant in statistical terms. The comparison in the second graph is less precise than in the first—the dotted line for the primary GDP model in Figure 2 falls mostly outside of the 90 percent confidence range for the polynomials on this new test (though largely within the joint 90 percent confidence range on differences, which is not shown).

The plot reveals the same positive relationship between GDP growth and middle-aged groups, and negative relationship between GDP growth and dependent groups, children, and retirees. Also, the transition ages, or roots of the polynomials, are only 10 years apart. One possible interpretation of this result is simple: most of these countries are very poor and very young. Perhaps a few additional oldsters may add stability and patience to a demography that might otherwise be pretty volatile![26]

4. Forecasts and Implications

What does the future hold? To state the obvious, past is not prologue. Indeed, the assumption that past statistical significance implies prospective forecasting accuracy has been the downfall of many quantitative models. So, we must share these indicative results with this very important caveat: high in-sample statistical significance does not guarantee that these models can be trusted to accurately predict the future. In these "forecasts," we tacitly assume that past relationships between demography and GDP growth, or between demography and capital market returns, will hold unaltered in the future. The prospective scenarios that flow from this work are *entirely possible*, but we dare not assume that they are highly likely. Accordingly, we should not view these daunting results as anything more than a cautionary tale.

that of Haiti, so it gets 10 times the weight. The effective number of countries is about 75 because we rely only lightly on the smaller countries, and yet the BRICs still get 14 percent weight out of the 175 countries. With this compromise, we can have some confidence that our models are useful for the major markets.

[26] We acknowledge that this may be an *ex post* rationalization of the modest differences between these results and our core results.

The number of factors that could affect the outcomes of individual countries is too large to even try to enumerate here, so our predictions should be taken not with a grain but a shakerful of salt. Still, the combination of the thoroughly researched projections by the UN Population Division for the years ahead and the statistical significance of our estimated relationships create an opportunity that is too tempting to resist. To the best of our knowledge this is the first paper to make systematic projections for GDP growth and financial markets based on demographic changes for practically every country in the world.[27]

Our forecasts are based on Equation (2). Because the regressions include country dummies or fixed effects, we rewrite them as

$$r_{i,t} - \bar{r}_i = \lambda\left(X_{i,t-1} - \bar{X}_i\right) + D_1\left(\gamma_{i,t}^{(1)} - \bar{\gamma}_i^{(1)}\right) + D_2\left(\gamma_{i,t}^{(2)} - \bar{\gamma}_i^{(2)}\right) + \cdots + D_k\left(\gamma_{i,t}^{(k)} - \bar{\gamma}_i^{(k)}\right) + \varepsilon_{i,t}. \quad (3)$$

Further, we use only the demographic variables to construct our predictions,

$$\mathrm{E}_t[r_{i,t+j} - \bar{r}_i] = D_1 \cdot \mathrm{E}_t[\gamma_{i,t+j}^{(1)} - \bar{\gamma}_i^{(1)}] + D_2 \cdot \mathrm{E}_t[\gamma_{i,t+j}^{(2)} - \bar{\gamma}_i^{(2)}] + \cdots + D_k \cdot \mathrm{E}_t[\gamma_{i,t+j}^{(k)} - \bar{\gamma}_i^{(k)}], \quad (4)$$

where $\bar{r}_i$, for instance, represents the time-series average of $r_{i,t}$ for country i. These forecasts focus singularly on the demographic component, setting aside all starting valuation effects. Unlike yields and interest rates, the population variables are projected by the UN far into the future, allowing us to make extended forecasts.

It's important to note that all of our predictions should be seen as abnormal GDP growth or abnormal stock and bond returns. These are deviations from long-term means relative to *past individual country* performance. We chose not to include $\bar{r}_i$ in our forecasts for two reasons. First, we could only include them for a small number of countries due to data availability. Second, the statistical precision around these historical means is small, introducing even more uncertainty into our figures. This decision means that comparisons across countries should be made carefully and under the perhaps reasonable assumption that nations have relatively similar long-term expectations for $\bar{r}_i$.

Because the population data come in five-year intervals, we calculate decade-long forecasts for 2011–2020 as the geometric average between the forecasts for two time periods: 2011–2015 and 2016–2020. All forecasts use the coefficients in Table 1. **Figure 4** (GDP), **Figure 5** (stocks), and **Figure 6** (bonds) show a total of four graphs each. For each figure, the top two use projected demographic shares, $E_t[s_{i,t+j}^{(n)}]$, and the bottom two use changes in projected demographic shares, $E_t[\Delta s_{i,t+j}^{(n)}]$. Each figure has two types of graphs: maps with color-coded results for all countries and bar plots for the 22 developed market countries in our main tests. The bar chart also serves to provide a sense of scale for the colored maps on the left. The countries are first divided into positive (blue) and negative (red) groups and then into three subgroups with approximately the same number of countries within each subgroup.

[27] We do recognize that most countries do not have well established and functioning financial markets. However, the questions posed here are still interesting, given that the answers to them inform us about deeper aspects of the population structure of different countries. For instance, one can still learn a lot about the demographic pressures on national savings, investment, or productivity.

We see two common themes in the developed countries: negative values for GDP growth and positive values for abnormal bond excess returns. This is intuitive given their current demographic stage: A rising number of retirees combined with low birth rates create significant pressure on output (negatively) and savings (positively). The net effect on stock excess returns varies across countries; in a handful of cases we also find divergence between the two graphs in the same figure, most dramatically for Singapore and Germany. This divergence between the forecasts using demographic shares and changes in demographic shares is mainly due to unusually rapid variations in some age groups, creating extreme values for $E_t[s_{i,t+j}^{(n)}]$ or $E_t[\Delta s_{i,t+j}^{(n)}]$, relative to their historical means.[28] These effects are amplified in the case of stocks, given their higher volatility and the relatively larger differences between the coefficient profiles in Figures 1 and 2. *We do not favor one set of results over the other.*

Not surprisingly, there is a much wider range in the global forecasts. In the case of GDP growth the implied forecasts are bleak, with rare exceptions including India and most African countries given their large working age cohorts and very few senior citizens. Bonds exhibit the opposite results. In this case most of the aging world has a fast-rising roster of middle-aged potential savers, leading to positive predictions for bonds, while Africa presents mostly negative forecasts due to a large number of younger borrowers relative to older savers.

We recognize that many of the emerging market results, for stock and bond markets, are hypothetical as many of these countries do not have well-developed stock or bond markets. Still, these results are ever more relevant and interesting as capital markets begin to take shape in many of these markets.

The most interesting and extreme cases are found for stocks. Japan, Finland, and Sweden have a dangerous combination of very low birth rates and an exploding number of retirees, giving them a strong demographic headwind, which has been much in the news lately. This work perhaps adds scientific rigor to those discussions. The results for Canada, the United States, and central Europe are mixed, with slightly negative or positive projections depending on the measure used. The European periphery, including Ireland, Portugal, Spain, and Greece, has slightly better forecasts for stock and bond returns, showing that there is hope, even in light of their recent troubles. Here, again, the emerging economies present mixed results. Latin America (including Mexico) and Asia fare relatively well, with mostly positive returns. Central Sub-Saharan Africa has too many young people and too few saving-age adults for stock markets in these countries to fare well, while Paraguay has too many old people and too few savers. Finally, the BRIC countries seem to have a bright future ahead of them, at least for the next 10 years.

5. Conclusion

[28] A plot of the population distribution for these two countries (not reported here) reveals some very interesting properties. Germany was more affected by World War II than any of his European neighbors. Thus, while it will experience similar developments in age groups 0–39, the structural changes in ages 40+ will be significantly different. Singapore has experienced very intense migration, which gives it one of the fastest changing demographic structures in the world across all ages.

This paper studies the effect of demographic changes on three measures of great importance for countries all over the world: real per capita PPP-adjusted GDP growth, stock market excess returns, and bond market excess returns. We use an econometric technique that has not been used for this purpose; it allows us to use all the available information in population profiles by fitting a polynomial curve on the regression coefficients of demographic age groups. Our main regressions use a large panel spanning over 60 years and 22 countries (and approximately 200 countries in some special cases).

By force-fitting a polynomial to our demographic variables, we extract surprising statistical significance with implications that align nicely with intuition. Children are not immediately helpful to GDP: they don't contribute to it. Nor do they help stock and bond market returns in any meaningful way: their parents are likely disinvesting to pay their support. Young adults are the driving force in GDP growth; they are the sources of innovation and entrepreneurial spirit. But, they are not yet investing as they are overspending against their future human capital. Middle adults are the engine for capital markets returns; they are in their prime for income, savings, and investments. And senior citizens contribute to neither GDP growth nor stock and bond market returns, as they disinvest to buy goods and services that they no longer produce.

In the closing part of the paper we provide "forecasts" for GDP growth, as well as abnormal stock and bond market excess returns, for almost every country in the world. Our projections are bleak in the case of GDP growth for most of the developed world. The developed countries, and Japan in particular, have alarmingly low birth rates and high number of retirees. But, especially for the younger of the developed economies, the picture is relatively good for bonds and mixed in the case of stocks. These "forecasts" must be seen for what they are: out-of-sample extrapolations of some reasonably powerful in-sample relationships, to be viewed as more cautionary than predictive.

Finally, we would like to stress again that these are long-term trends only, estimated by assuming past relationships between demographic structures and measures of economic growth and capital markets. Population profiles change very slowly and these forecasts are by no means immutably bleak scenarios. The time to respond gets shorter every year, but countries have many tools to try and act on these forecasts, to position themselves for more benign circumstances. This is a topic that is outside the scope of our work, but could certainly fill many more papers.

Appendix

CONSTRUCTION OF THE DEMOGRAPHIC POLYNOMIAL

The effect of each age group, $s_t^{(i)}$, or changes thereof, $\Delta s_t^{(i)}$, can be estimated with the following regression:

$$r_t = a + \gamma X_{t-1} + b_1 s_t^{(1)} + \cdots + b_N s_t^{(N)} + \varepsilon_t, \tag{A1}$$

where r_t is the rate of return on stocks, bonds, or GDP, and X_{t-1} represents any control variables, such as interest rates or valuation ratios. We constrain the demographic coefficients b_i to satisfy a polynomial of order k:

$$b_i = D_0 + D_1 i + D_2 i^2 + \cdots + D_k i^k \qquad i = 1, \ldots, N \tag{A2}$$

Substituting these back into the regression Equation (A1) gives us:

$$r_t = a + \gamma X_{t-1} + \sum_{j=0}^{k} D_j 1^j s_t^{(1)} + \cdots + \sum_{j=0}^{k} D_j N^j s_t^{(N)} + \varepsilon_t,$$

and rearranging:

$$r_t = a + \gamma X_{t-1} + D_0 + D_1 \sum_{i=1}^{N} i s_t^{(i)} + D_2 \sum_{i=1}^{N} i^2 s_t^{(i)} + \cdots + D_k \sum_{i=1}^{N} i^k s_t^{(i)} + \varepsilon_t. \tag{A3}$$

Notice that the demographic shares add up to a constant, $\sum_{i=1}^{N} s_t^{(i)} = 1$ or $\sum_{i=1}^{N} \Delta s_t^{(i)} = 0$, at all points in time. To avoid multicollinearity with the intercept a, we impose the restriction $\sum_{i=1}^{N} b_i = 0$ on the demographic coefficients. Translating this restriction into the polynomial coefficients we obtain:

$$D_0 = -\frac{1}{N}\left(D_1 \sum_{i=1}^{N} i + D_2 \sum_{i=1}^{N} i^2 + \cdots + D_k \sum_{i=1}^{N} i^k \right).$$

Finally, the substitution of D_0 back into (A3) shows how to obtain the restricted coefficients of the demographic polynomial, D_j, with a regression:

$$r_t = a + \gamma X_{t-1} + D_1 \sum_{i=1}^{N} \left(i s_t^{(i)} - \frac{i}{N} \right) + D_2 \sum_{i=1}^{N} \left(i^2 s_t^{(i)} - \frac{i^2}{N} \right) + \cdots + D_k \sum_{i=1}^{N} \left(i^k s_t^{(i)} - \frac{i^k}{N} \right) + \varepsilon_t.$$

To reconstruct the original demographic coefficients we use equation (A2).

References

Abel, Andrew. 2003. "The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security." *Econometrica*, vol. 71, no. 2 (March):551–78.

Ang, Andrew, and Angela Maddaloni. 2003. "Do Demographic Changes Affect Risk Premiums? Evidence from International Data." Working Paper W9677, NBER (May).

Ang, Andrew, Monika Piazzesi, and Min Wei. 2006. "What Does the Yield Curve Tell Us About GDP Growth?" *Journal of Econometrics*, vol. 131, no. 1 (March):359–403.

Arnott, Robert D. 2011. "Does *Unreal* GDP Drive our Policy Choices?" Forthcoming.

Arnott, Robert D., and Anne Casscells. 2003. "Demographics and Capital Market Returns." *Financial Analysts Journal*, vol. 59, no. 2 (March/April):20–29.

Arnott, Robert D., and Anne Casscells. 2004. "Will We Retire Later and Poorer?" *Journal of Investing*, vol. 13, no. 2 (Summer):25–36.

Bakshi, Gurdip S., and Zhiwu Chen. 1994. "Baby Boom, Population Aging, and Capital Markets." *Journal of Business*, vol. 67, no. 2 (April):165–202.

Bosworth, Barry P., Ralph C. Bryant, and Gary Burtless. 2004. "The Impact of Aging on Financial Markets and the Economy: A Survey." Working Paper 2004-23, Center for Retirement Research at Boston College Working Paper 2004-23 (October).

Brooks, Robin. 2000. "What Will Happen to Financial Markets When the Baby Boomers Retire?" Working Paper, IMF (April).

Campbell, John Y. 1987. "Stock Returns and the Term Structure." *Journal of Financial Economics*, vol. 18, no. 2 (June):373–99.

Cochrane, John H. 1994. "Permanent and Transitory Components of GNP and Stock Prices." *Quarterly Journal of Economics*, vol. 109, no. 1 (February):241–265.

———. 2005. *Asset Pricing, Revised Edition.* Princeton, NJ: Princeton University Press.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra. 2002. "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle." *Quarterly Journal of Economics*, vol. 117, no. 1 (February):269–296.

Davis, E. Philip, and Christine Li. 2003. "Demographics and Financial Asset Prices in the Major Industrial Economies." Working Paper, Brunel University (March).

DellaVigna, Stefano, and Joshua M. Pollet. 2007. "Demographics and Industry Returns." *American Economic Review*, vol.97, no. 5 (December): 1667–1702.

Fair, Ray C., and Kathryn M. Dominguez. 1991. "Effects of the Changing U.S. Age Distribution on Macroeconomic Equations." *American Economic Review*, vol. 81, no. 5 (December):1276–1294.

Geanakoplos, John, Michael Magill, and Martine Quinzii. 2004. "Demography and the Long-Run Predictability of the Stock Market." *Brookings Papers on Economic Activity: 1* (Brookings Institution):241–307.

Goyal, Amit. 2004. "Demographics, Stock Market Flows, and Stock Returns." *Journal of Financial and Quantitative Analysis*, vol. 39, no. 1 (March):115–142.

Green, Richard, and Patric H. Hendershott. 1996. "Age, Housing Demand, and Real House Prices." *Regional Science and Urban Economics*, vol. 26, no. 5 (August):465–480.

Greene, William H. 2008. *Econometric Analysis.* NJ: Prentice Hall.

Heston, Alan, Robert Summers, and Bettina Aten. 2009. *Penn World Table Version 6.3*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania (August).

Higgings, Matthew. 1998. "Demography, National Savings, and International Capital Flows." *International Economic Review*, vol. 39, no. 2 (May):343–369.

International Monetary Fund. 2004. "How Will Demographic Change Affect the Global Economy?" *World Economic Outlook*, Chapter 3 (September):151.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2009. "Is Newer Better? Penn World Table Revisions and Growth." Working Paper No. 191, Center for Global Development Working Paper (September).

Kanazawa, Satoshi. 2003. "Why Productivity Fades with Age: The Crime-Genius Connection." *Journal of Research in Personality*, vol.37, no. 4 (August) : 257-272.

Keim, Donald B., and Robert F. Stambaugh. 1986. "Predicting Returns in the Stock and Bond Markets." *Journal of Financial Economics*, vol. 17, no. 2 (December):357–390.

Lindh, Thomas, and Bo Malmberg. 1999. "Age Structure Effects and Growth in the OECD, 1950–1990." *Journal of Population Economics*, vol. 12, no. 3:431–449.

Mankiw, N. Gregory, and David N. Weil. 1989. "The Baby Boom, the Baby Bust, and the Housing Market." *Regional Science and Urban Economics*, vol. 19, no. 2 (May):235–258.

Mehra, Rajnish. 2008. *Handbook of the Equity Risk Premium.* Amsterdam: Elsevier.

Poterba, James M. 2001. "Demographic Structures and Asset Returns." *Review of Economics and Statistics*, vol. 83, no. 4 (November):565–584.

Wydler, Daniel. 1989. "Swiss Stocks, Bonds, and Inflation, 1926–1987." *Journal of Portfolio Management*, vol. 15, no. 2 (Winter):27–32.

Yoo, Peter S. 1994. "Age Distributions and Returns of Financial Assets." Working Paper 1994-002A, Federal Reserve Bank of St. Louis (February).

Tables and Figures

Table 1 – Regression Results, 1950-2010

This table reports coefficients and *t*-statistics (in parentheses) for 6 separate panel regressions, one in each column. The dependent variables are annualized five-year growth rates of GDP per capita, excess stock returns and excess bond returns. The demographic regressors are shares of total population, $s_{i,t}^{(j)}$, in the first three columns and changes thereof, $\Delta s_{i,t}^{(j)}$, in the last three columns. $D_1 - D_4$ are the coefficients of the polynomials that approximate the demographic coefficients. All regressions include country fixed effects. The last two rows report the *p*-values of Wald tests for comparisons between nested models, opposing a parabola and a cubic polynomial, and cubic and quartic polynomials.

	Demographic Shares			Change in Demographic Shares		
	Stocks	**Bonds**	**GDP**	**Stocks**	**Bonds**	**GDP**
Dividend Yield	3.39			3.38		
	(6.23)			(6.30)		
10-year Yield		0.66			0.62	
		(4.98)			(4.07)	
3-month Yield			-0.15			-0.19
			(2.97)			(4.14)
D_1 (x1)	-0.81	-1.56	0.09	3.09	-0.02	0.21
	(1.63)	(3.33)	(3.77)	(2.90)	(0.04)	(3.56)
D_2 (x10)	1.69	3.47	-0.07	-10.27	-1.83	-0.14
	(2.47)	(3.10)	(4.39)	(3.57)	(1.42)	(3.51)
D_3 (x100)	-0.83	-2.72		11.81	2.96	
	(3.05)	(2.74)		(4.25)	(2.58)	
D_4 (x1000)		0.69		-4.22	-1.17	
		(2.37)		(4.77)	(3.32)	
R^2	28%	34%	30%	31%	30%	17%
Observations	203	241	255	203	241	255
Countries	22	22	22	22	22	22
***k=3* → *k=2* (%)**	0.3	0.0	25.7	7.7	3.4	13.4
***k=4* → *k=3* (%)**	28.4	1.8	12.4	0.0	0.1	53.6

Table 2 – GDP Growth Regression Results (Robustness Check), 1950-2010

This table repeats the per capita GDP growth tests in Table 1 using a different sample of countries. See Table 1 for details.

	Demographic Shares	Changes in Demographic Shares
log(*C/GDP*)	1.67	2.02
	(1.25)	(1.58)
D_1 (x1)	0.09	-0.09
	(4.29)	(1.19)
D_2 (x10)	-0.07	0.31
	(4.76)	(2.15)
D_3 (x100)		-0.18
		(2.49)
R^2	4%	3%
Observations	1640	1640
Countries	176	176
***k=3* → *k=2* (%)**	17.9	1.3
***k=4* → *k=3* (%)**	40.6	40.3

Figure 1 – Effect of Demographic Shares

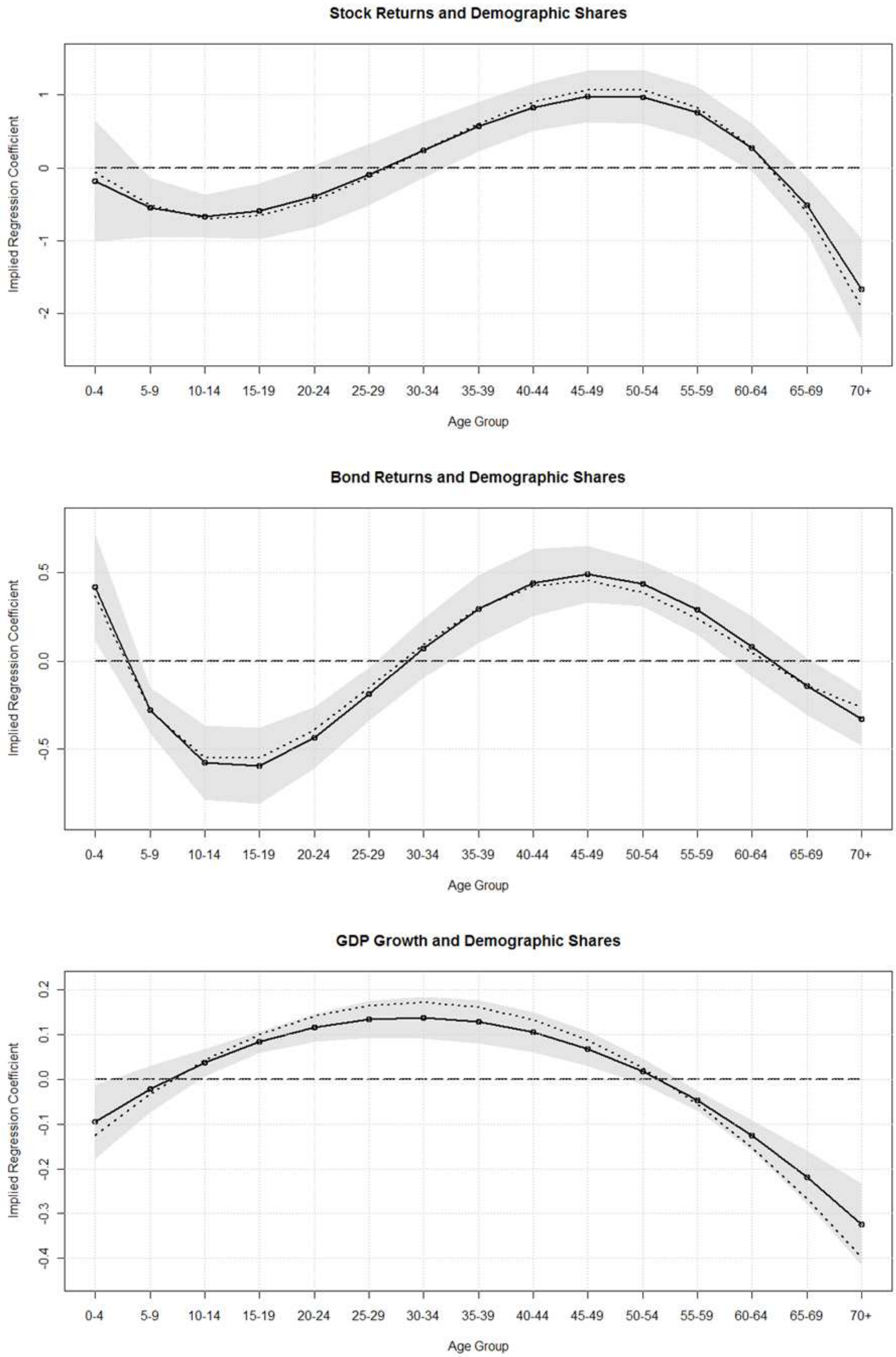
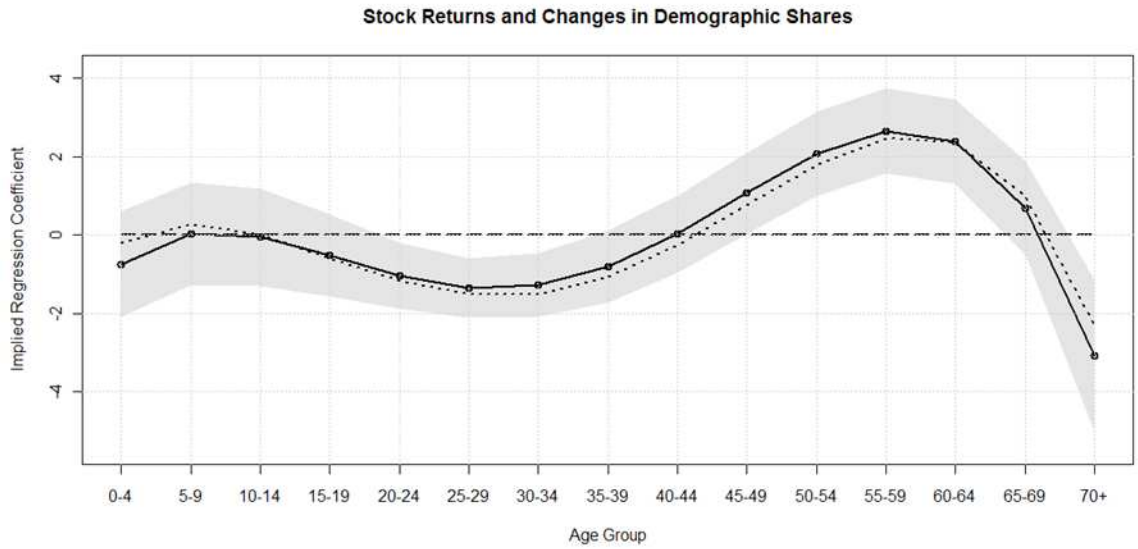
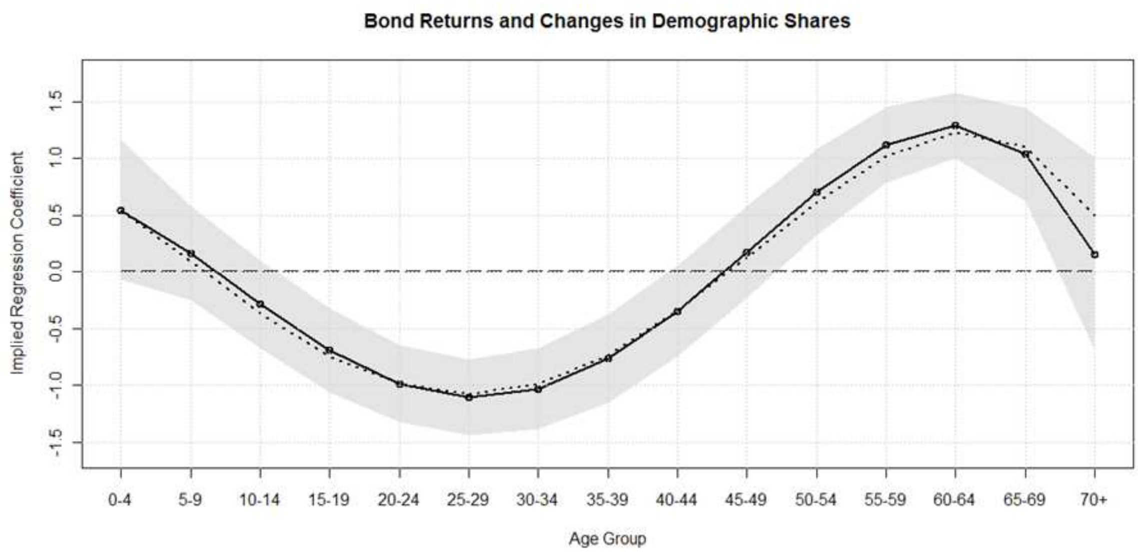


Figure 2 – Effect of Changes in Demographic Shares





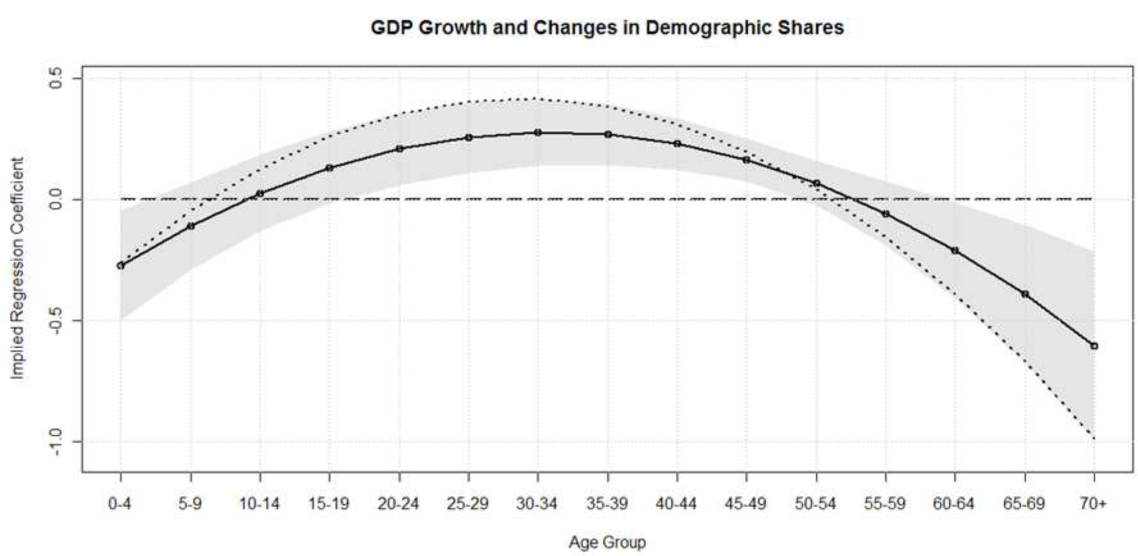


Figure 3 – Robustness Checks

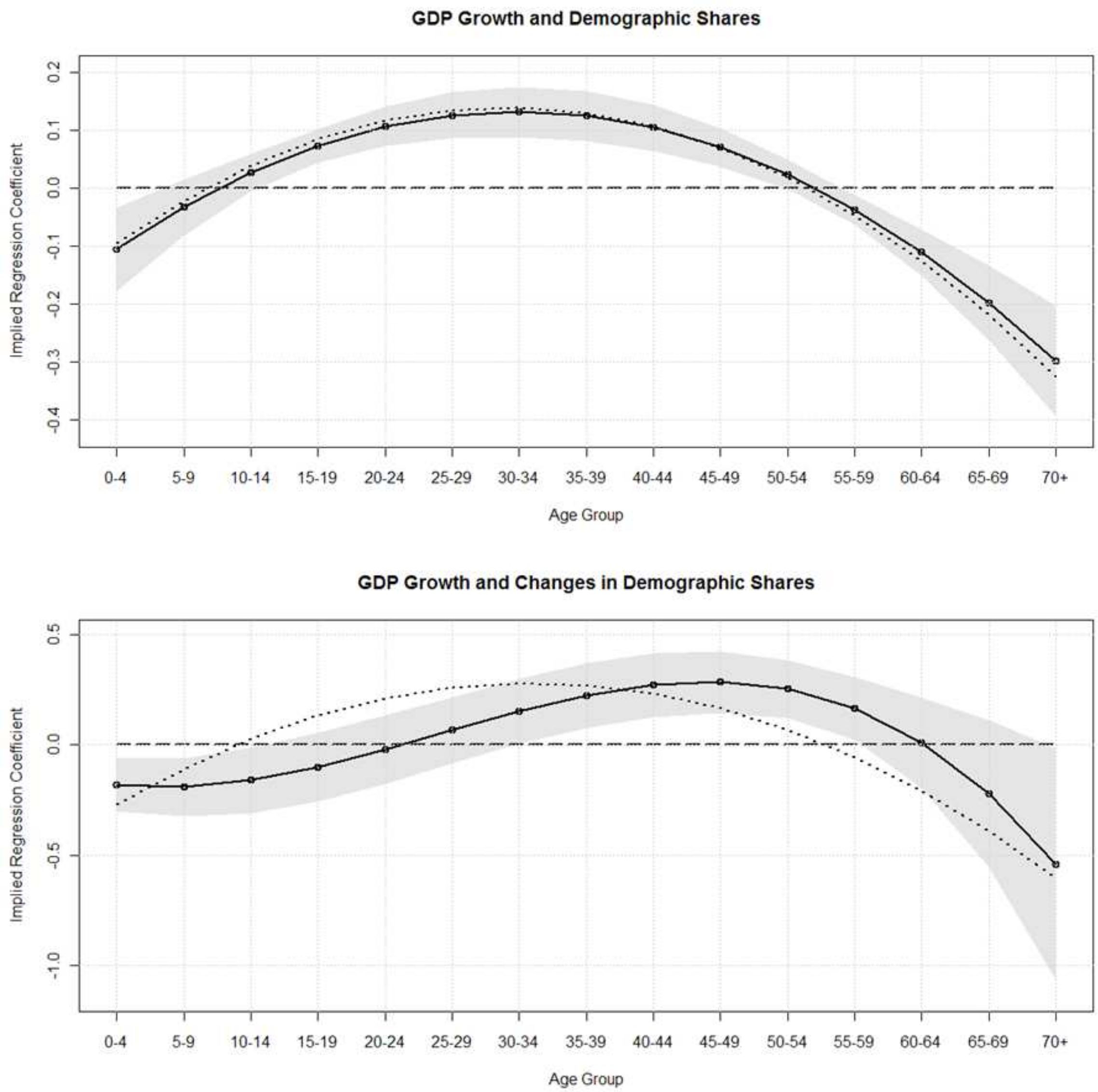
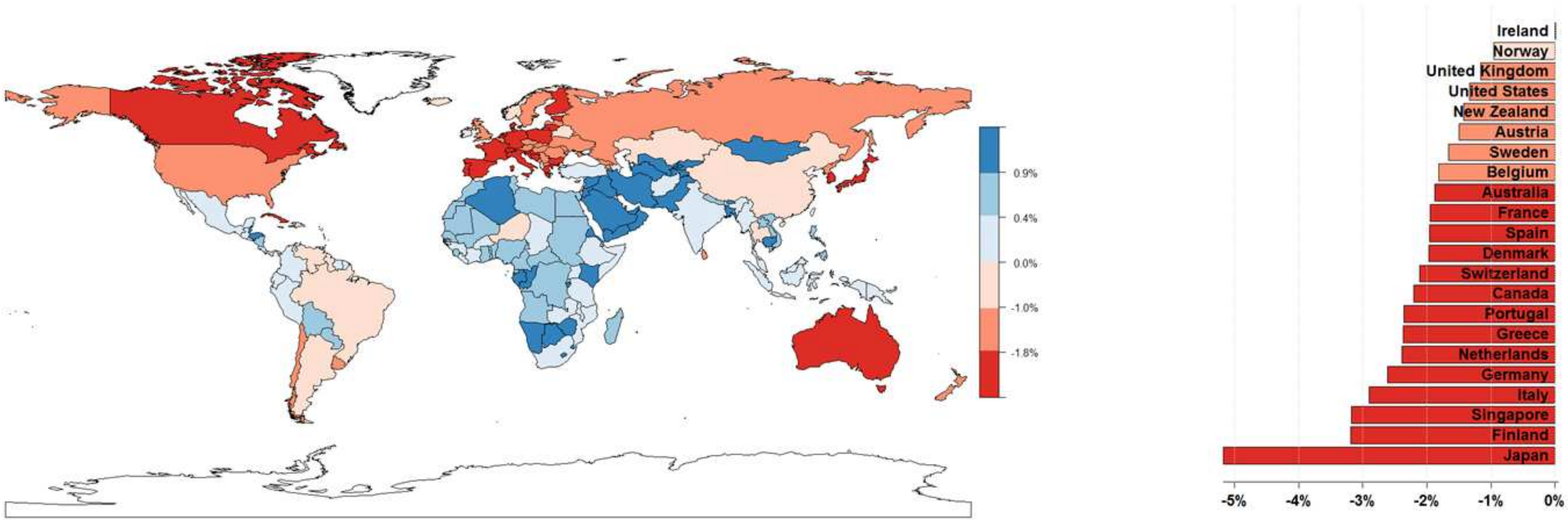


Figure 4 – Annualized GDP Growth Forecasts, 2011-2020

Forecasts using Demographic Shares



Forecasts using Changes in Demographic Shares

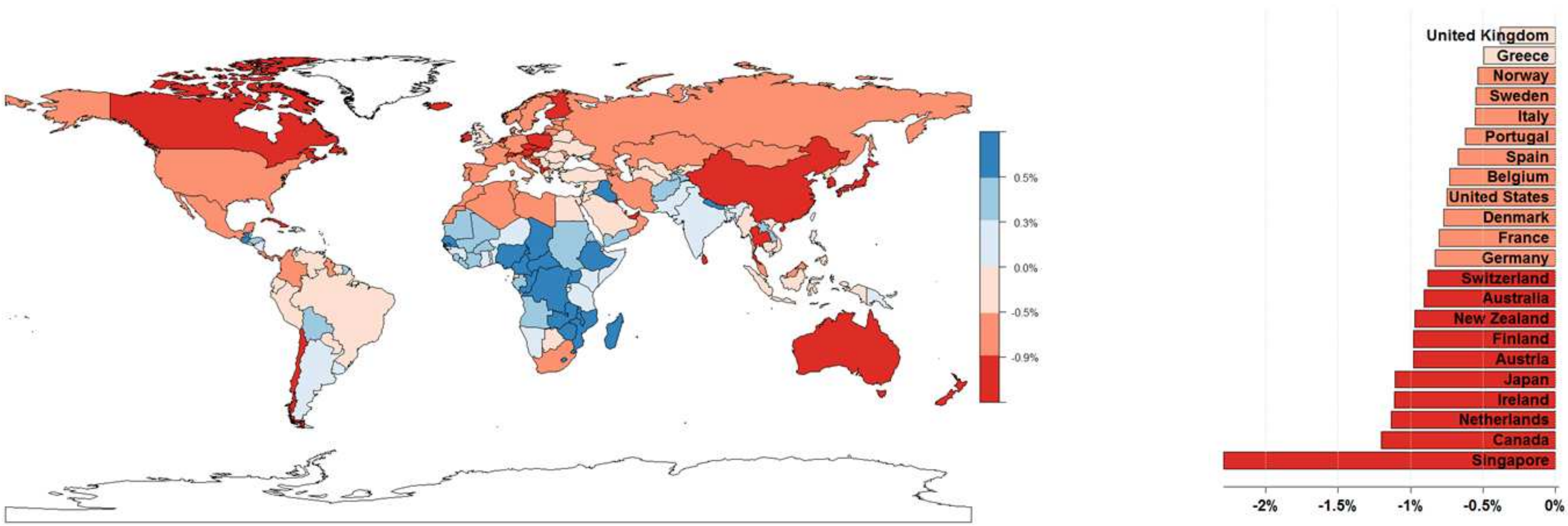
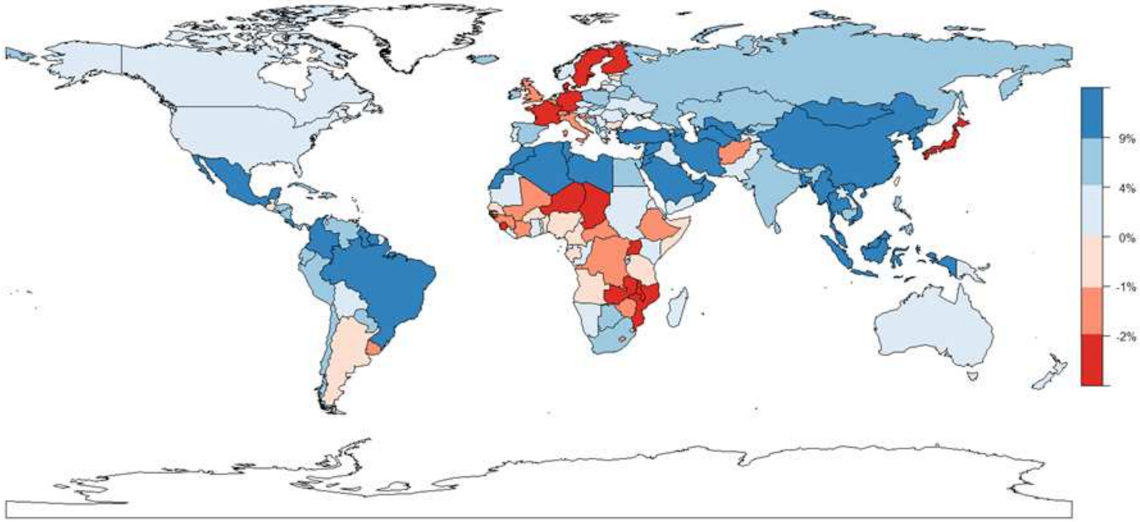
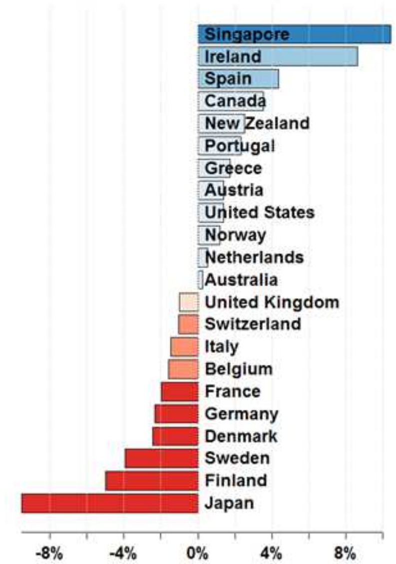


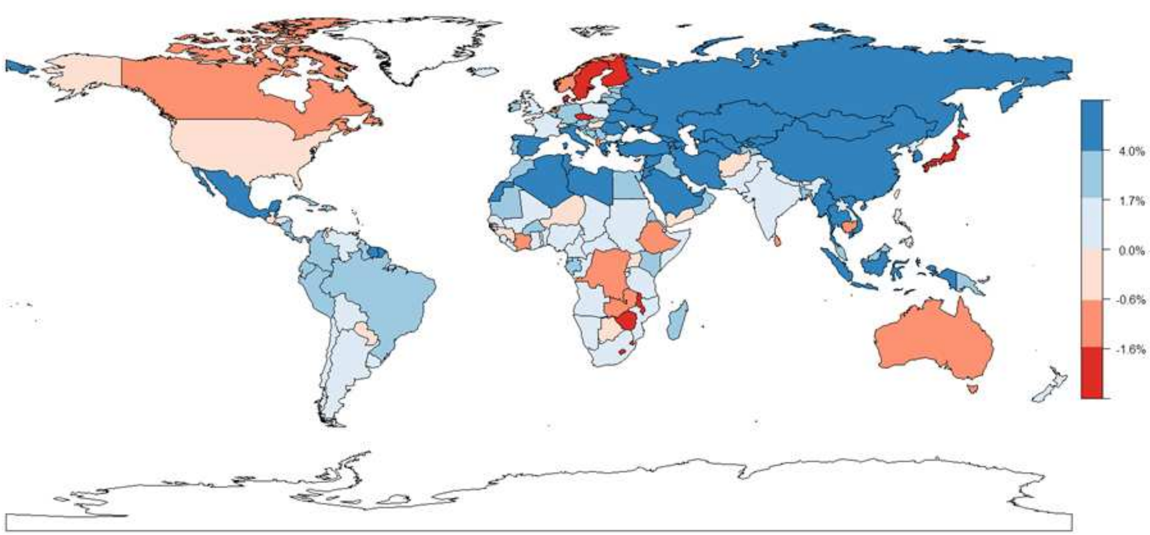
Figure 5 – Annualized Stock Market Forecasts, 2011-2020

Forecasts using Demographic Shares





Forecasts using Changes in Demographic Shares



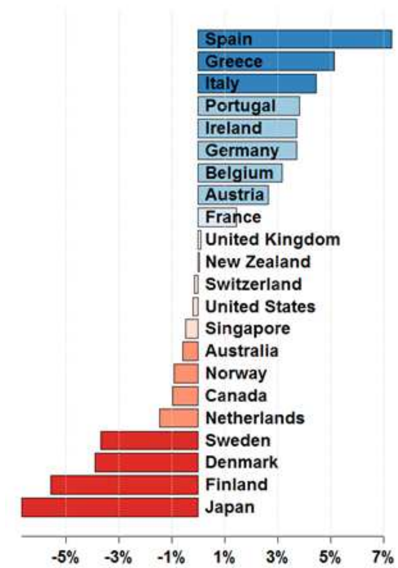
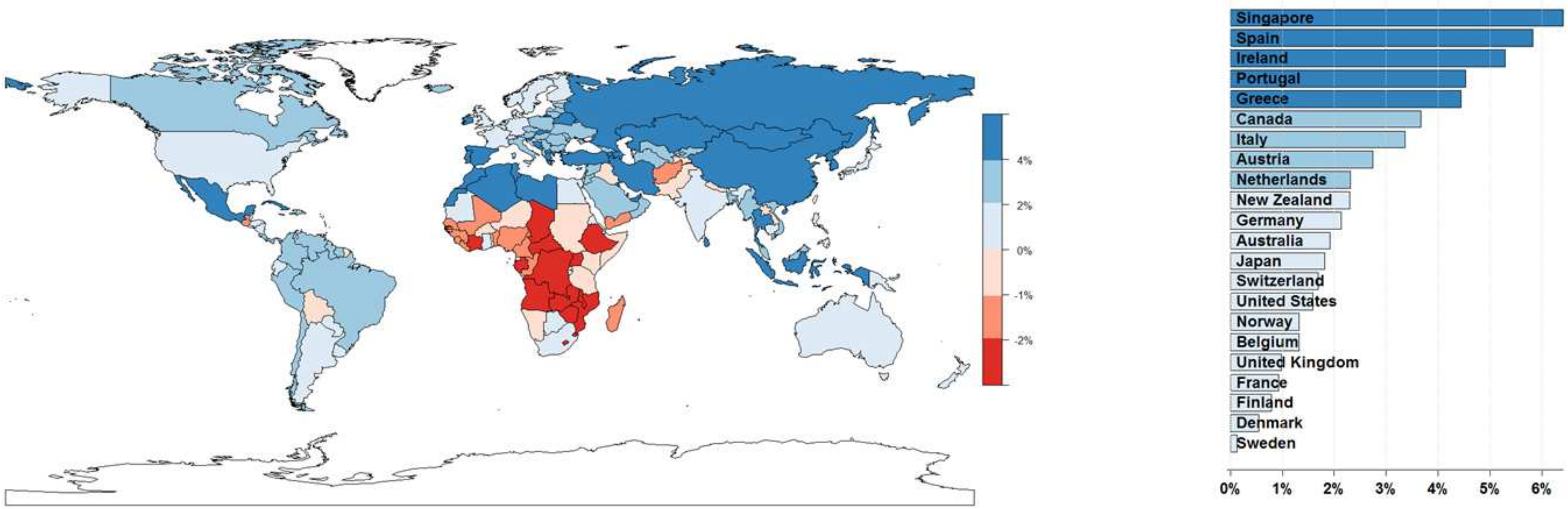


Figure 6 – Annualized Bond Market Forecasts, 2011-2020

Forecasts using Demographic Shares



Forecasts using Changes in Demographic Shares