Financial Analysts Journal
Volume 68 • Number 1

FAJ

Demographic Changes, Financial Markets, and the Economy

Robert D. Arnott and Denis B. Chaves

Using a large sample of countries and 60 years of data, the authors found a strong and intuitive link between demographic transitions and both GDP growth and capital market returns. Unlike previous researchers, who used ad hoc *and restrictive demographic variables, the authors imposed a smooth and parsimonious polynomial curve across all age groups. They also performed robustness checks and produced forecasts for the coming decade, with all the necessary caveats.*

Demography is destiny.
—Auguste Comte (1798–1857)

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 for this age cadre. Looking 10 years into the future, we can see some uncertainty in the number of people under 10 and in the number of people over 70 (depending on the progress of medical science), but surprisingly little uncertainty in the number of people aged 10–70—barring war, pestilence, or other catastrophes. As the Baby Boomers have aged, many researchers have studied past demographic data in an effort to extract indications of the future influence of Boomers on various aspects of the economy, from housing prices to consumer preferences to retirement plans.

Although the genesis of our study has the same roots—curiosity about the potential impact of aging Baby Boomers—we decided to pursue a broader course of study that spans 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 strove not to extend the theoretical relationships between demographic changes and financial markets or the economy but, rather, to apply new empirical techniques. 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 polynomials to the regression coefficients between demographic age groups and both GDP growth and stock and bond returns. The use of polynomials is intuitive because it satisfies two important criteria: parsimony (only a small number of parameters are required) and continuity across age groups (behavior should change reasonably smoothly from one age cadre to the next).

Two core principles influenced our research design. First, we deemed models for GDP growth less interesting than models for real per capita PPP-adjusted GDP growth.[1] All three of these modifiers of 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—thus, our focus on real per capita GDP growth. The PPP adjustment creates a fairer global comparison by emphasizing the domestic purchasing power of the average citizen for the consumption basket that matters for each country.[2]

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

Robert D. Arnott is CEO and chairman of Research Affiliates, LLC, Newport Beach, California. Denis B. Chaves is a senior researcher at Research Affiliates, LLC, Newport Beach, California.

Effects of Demographic Changes on Financial Markets and the Economy

The relationship between per capita GDP growth and demographic changes can be motivated by a decomposition of per capita total output of goods and services in an economy into age groups and then into a product of per worker productivity and per capita number of workers:

$$\frac{\text{Total output}}{\text{Population}} = \sum_j \text{Productivity per worker in age group } j \left(\frac{\text{Number of workers in age group } j}{\text{Population}}\right).$$

That a growing workforce—or even a growing population, for that matter—should be better for GDP growth than a shrinking workforce seems tautological. To isolate these effects, we can focus on *per capita* GDP growth. In this construct, the relevant measure becomes the size of the working-age cadre (generally, people aged 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 tailwind to real per capita PPP-adjusted GDP; if slower, then our GDP measure should face a headwind.[3]

This simple equation allows us to identify two channels through which demographic changes can influence a country's GDP. First, let us assume that productivity varies significantly across age groups. 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 per capita output should first increase and then decrease. In the demography literature, this effect is known as a *demographic dividend*: 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. For instance, Nobel Prizes are usually awarded to older scientists and researchers for contributions made years earlier, when they were much younger. Kanazawa (2003) found that the productivity of scientists, musicians, and painters peaks between 30 and 40—the only exception is authors, whose productivity tends to peak between 40 and 50. Therefore, we would also expect to identify a higher increase in productivity across all age groups—and thus in total per capita output—in countries with a relatively higher share of younger cadres.

Another reason that young adults have more impact on GDP growth than do middle-aged and older adults is even simpler. Because we are using growth rates in output as opposed to merely output, the effects that we identify should peak at ages considerably earlier than pre-retirement. A worker's contribution to total output likely peaks when she has more experience, but her contribution to *growth* in total output is highest when she is in the process of acquiring that experience.[4] For each of us, the biggest jump in our contribution to GDP occurs as we transition from nonworking adolescents into gainfully employed 20-somethings. Another, often smaller, jump in our contribution to GDP occurs as we mature into our 30s. By our 40s, the evidence of real wages would suggest that most of us are at or approaching our peak contribution to GDP, with a falling contribution to GDP in our 50s and 60s.

Either way, we would expect per capita GDP growth to be strongest in populations dominated by young adults and in populations in which the young-adult population is growing quickly. This circumstance is indeed the case for most emerging economies and is assuredly *not* the case for most developed economies.

Establishing a theoretical link between financial markets and demographic changes is a complex task and would require more space than is available here, not to mention that many excellent papers on this topic already exist. Moreover, the focus of our study was on trying to advance the empirical side of the literature. Therefore, as a motivation for our empirical analysis, we offer a summary of the intuition and findings in the theoretical literature.

Brooks (2000), for instance, solved an overlapping-generations model with forward-looking agents, a risky asset, and a riskless one-period bond. Various age cohorts trade with each other. During baby booms, consumption is relatively higher and savings 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.[5]

In contrast, 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 distributions into their information set and act accordingly. Thus, age effects would be insignificant and, in practice, unobservable. To address and dismantle this criticism, some researchers (see, e.g., Abel 2003; Geanakoplos, Magill, and Quinzii 2004) developed 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 International Monetary Fund (2004) in its *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. (p. 151)

To summarize, the findings of most of these models 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.[6] As they look past their own and their children's immediate needs to their eventual retirements, they begin to invest—first in stocks, then in bonds. As they slide into retirement, they begin to sell assets in order to buy goods and services that they no longer produce—either directly, through their own investments, or indirectly, through their pension benefits. They tend to liquidate their riskiest assets (stocks) before their less risky assets (bonds).

Review of the Empirical Literature

Analyzing the relationship between demographic changes and the economy, whether qualitatively or quantitatively, has been a topic of interest for centuries.[7] Accordingly, summarizing all the relevant research would be an impossible task. In this short review, we focus on the recent literature, with an emphasis on empirical tests, comparing it with our own application of new methods to the data.

Mankiw and Weil (1989) conducted one of the first studies to link demography and financial assets. They showed a strong relationship between the U.S. Baby Boom, the subsequent increase in housing demand in the early 1970s, and a substantial impact on housing prices over the next 20 years. Their forecast that housing prices would decline sharply after 1990, following the Baby Bust, did not materialize on schedule, although demographic effects may have magnified the recent collapse of the housing bubble.[8]

Arnott and Casscells (2003, 2004) discussed the demographic changes that will likely occur in the United States in the coming decades, explored their implications for capital markets, and examined possible solutions. Their important finding was that the entitlements problem is not a financial problem exacerbated by a failure to prefund these obligations but, rather, 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. Arnott and Casscells also speculated on the likely implications for capital market returns, consistent with the results that have subsequently been seen in the United States, Japan, and parts of western Europe. Finally, they examined and largely dismissed the common arguments that immigration or productivity gains can offset the pressures associated with pending demographic shifts.

Bosworth, Bryant, and Burtless (2004) surveyed the literature on analyzing and forecasting prices (returns) of financial assets and identified two approaches. The first approach uses microeconomic data on goods or asset holdings,[9] together with forecasts of age distributions, to predict how future demand and prices (returns) will evolve (see Poterba 2001). For example, DellaVigna and Pollet (2007, p. 1667) defined what they called "age-sensitive sectors, such as toys, bicycles, beer, life insurance, and nursing homes," and estimated how the demand for products or services in those sectors will decrease or increase. They found that their demand forecasts predict both profitability and stock returns by industry 5–10 years in the future. The second approach—followed in our study and surveyed by Davis and Li (2003)—estimates the direct time-series relationship between prices (returns) and demographic variables. Notable studies 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) estimated multivariate time-series regressions of annual U.S. stock, corporate bond, and government bond returns on shares of total population for the 25–34, 35–44, 45–54, 55–64, and 65+ age groups. His strongest results were a negative relationship between short- and medium-term government bonds and the 45–54 age group. The statistical significance, however, was weak for almost all coefficients in all five age groups. He also estimated the regressions with three- and five-year centered moving averages and found a significant increase in terms of both statistical significance and fit, which supports our claim that long horizons provide a better test for low-frequency population changes.

Bakshi and Chen (1994) hypothesized that relative risk aversion is positively correlated with age and modeled the utility function of the representative consumer as a function of aggregate consumption and average age of the population. In their tests, they used Euler equations and a two-factor model based on consumption growth and percentage change in average age.[10] They found strong support for their life-cycle risk aversion hypothesis and a positive and statistically significant relationship between U.S. stock excess returns and growth in the average age of the U.S. population.

Using both a long sample (1900–2001; 5 countries) and a short sample (1970–2000; 15 countries), Ang and Maddaloni (2003) studied the relationship between excess stock returns (at one-, two-, and five-year horizons) and log changes in three demographic variables: average age of the population over 20, fraction of adults over 65, and percentage of people in the 20–64 age group. In pooled regressions, their results displayed a strong and *negative* effect for the fraction of retirees in the population (65+). Interestingly, they found an opposite and *positive* result in isolated regressions for the United States and the United Kingdom. To explain this difference in results, they conducted additional tests and found that the effect for the 65+ age group is stronger in countries with well-developed social security systems and less developed financial markets. Finally, and most important for our study, they also found that "pooling data from five countries gives almost the same power as increasing the sample size of the United States by five times" (p. 10), which validates the strong results that we found by using an extended sample of countries.

Using a sample of countries in the Organisation for Economic Co-Operation and Development (1950–1990), Lindh and Malmberg (1999) studied the effects of log age group shares (15–29, 30–49, 50–64, and 65+) on five-year growth rates in GDP per worker. They found significant and *positive* coefficients for the 50–64 age group and significant and *negative* coefficients for retirees (65+). With respect to their choice of age groups, they commented that "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" (p. 435). These concerns raised by Lindh and Malmberg 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 whether any study that ignores young people could be important.

Our most important means of identification—and most significant improvement over prior research—is our use of the econometric methodology pioneered by Fair and Dominguez (1991). They used a polynomial to analyze the relationship between a changing U.S. age distribution and such economic variables as consumption, housing investment, money demand, and labor force participation. This methodology, which forces the regression coefficients on age groups into a polynomial, has at least two advantages: (1) It permits the inclusion of *all age groups* in the regression while avoiding the statistical issues created by the high correlation between them, and (2) the interpretation of the results is more intuitive. In that regard, we closely followed Higgins (1998), who studied the effect of demographic changes on savings, investment, and the current account balance, but we analyzed the implications for financial markets and economic growth.

In the empirical arena, Poterba (2001, p. 565) lodged one of the strongest criticisms, maintaining 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, we took a number of steps to increase the power of our tests:

1. Relied on *five-year* rates of GDP growth and capital market returns instead of *annual* data
2. Controlled for starting valuation levels, GDP levels, and business cycle measures so that the demographic effects could be viewed in isolation from these other powerful effects
3. Used a large cross section of countries
4. Included the information from all age groups in the regressions, as previously mentioned

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

Data and Variables

We drew our data from many sources. The demographic and GDP data are deepest, with demographic age profiles and GDP on well over 200 countries from the UN and the Penn World Table, typically well documented back to 1950.[11] The stock and bond data are solid for only the developed economies and the largest emerging economies, with comparatively thin data before 1970. Using nonoverlapping five-year returns or growth rates, we were able to extract more than 200 observations in 22 countries for our main regressions and more than 1,600 observations in 176 countries for our extended group of countries.

We obtained population data by five-year age cadres with a half-decade frequency from the UN Population Division.[12] For most countries, the data start in 1950 and include projections through 2050 in five-year intervals. For future years, 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 chose the medium variant available online, which assumes a midrange expectation for fertility and normal trends in mortality and migration.[13]

We obtained our financial variables from various sources, primarily Global Financial Data. We complemented total return indices for stocks with data from Claus Parum[14] (Denmark), the Central Statistics Office[15] (Ireland), and Daniel Wydler (1989; Switzerland). We proxied any missing 10-year yields with 5- or 7-year yields and any missing three-month yields with central bank discount rates.

We used measures of real per capita GDP from Version 6.3 of the Penn World Table (see Heston, Summers, and Aten 2009). This version includes annual observations over 1950–2007 and assigns a quality grade ranging from A (best) to D (worst) to each country "to signal the relative reliability of the estimates" (p. 18 of the Data Appendix).[16] Data availability and reliability—missing observations are mainly an issue with the financial variables—restricted our main sample to 22 developed countries,[17] of which 16 had a quality grade of A and 6 a B. Given the broad coverage of the population and GDP data, however, we were able to conduct robustness tests—with respect to GDP growth only, not stock or bond excess returns—for a sample of 176 additional countries.

Our dependent variables are annualized, five-year, nonoverlapping 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. Real per capita GDP growth measures economic activity as the average citizen might perceive it. Unlike most previous researchers, we chose five-year returns, chiefly for two reasons: (1) Demographic data are more widely available in five-year intervals, and (2) demographic changes occur slowly, giving low-frequency data a better shot at identifying the effects of interest in our study. The intersection of data sources left us with approximately 200–250 nonoverlapping observations in our main tests and more than 1,600 in our robustness tests. This broad database provided ample degrees of freedom and delivered surprisingly strong statistical significance for our results.

Our explanatory variables are the percentage of total population by age group, $s_{i,t}^{(j)}$, and changes therein, $\Delta s_{i,t}^{(j)}$.[18] Our 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 used 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 was vital, given that temporary swings in prices or economic activity might affect the estimation. For example, if stock market returns were high following a high initial dividend yield, we could not be confident that we were properly separating the demographic effects from well-known valuation effects. Our demographic results, however, are relatively robust with regard to the choice of the valuation measure. For instance, in one of our robustness tests, we were forced to use the log of the ratio of consumption to GDP, $\log(C_{i,t}/GDP_{i,t})$, given that interest rates for most countries are unavailable.

In our view, these data are deep enough to draw some important conclusions about *past* links between demography, on the one hand, and GDP growth and the capital markets, on the other. Applying our 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, although our confidence in the forward links is weaker than our confidence in the past links, the potential implications of our results are sobering.

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 on per capita GDP growth):

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

where $X_{i,t-1}$ represents such control variables as interest rates and valuation ratios. The problem with this approach is that the demographic variables are highly correlated and would generate the usual multicollinearity problems.[19] Moreover, in our case, the estimation of Equation 1 is *impossible*! Because the maximum number of nonoverlapping 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.[20] In the literature, the usual solution 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 with far too many independent variables, which still leads to unwieldy regressions.

We chose a different approach. Following Fair and Dominguez (1991) and Higgins (1998), we forced the demographic coefficients, b_j, to satisfy a polynomial of order k, allowing us to incorporate the information in the entire demographic profile. See Appendix A for details of the methodology. After imposing this restriction, we obtain k transformed demographic variables, $\gamma_{i,t}^{j}$, whose coefficients, D_j, can be estimated with[21]

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

Because k ranges between 2 and 4 (discussed later), this approach reduces 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 with the dependent variables. This timing is important because we are exploring demographic effects on financial assets, which presumably occur primarily via shifts in the demand for financial assets. 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 contributions to GDP. Moreover, these data *can* be concurrent because we know the demographic profile for each country over the next 5–10 years with some precision, especially for the future working-age and retirement-age cadres, which we can simply count today.

In each case, the choice of polynomial degree for the demographic variables strikes a compromise between parsimony and statistical power. Searching a large number of options is meaningless because larger values of k provide no reduction in dimensionality and also create a concern about data mining. For these reasons, we entertained only three options: polynomials of order k equal to 2, 3, or 4. In other words, we limited our tests to quadratic, cubic, and quartic polynomials. Because models with lower degrees are special cases of their high-degree counterparts, we relied on Wald tests for nested models to gauge which order of polynomial was most appropriate.

For simplicity and parsimony, in each regression we included only one control variable, $X_{i,t-1}$, out of these three: dividend yield, 10-year yield, and three-month yield. These variables play a central role in the literature as forecasters of stock and bond returns and of economic growth (see, e.g., Keim and Stambaugh 1986; Campbell 1987; Ang, Piazzesi, and Wei 2006). Demographic coefficients—the main point of interest in our study—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 were effective in our study, wholly consistent with prior research.

Some skeptics might argue that other factors not included in the regressions could contribute to growth in per capita GDP or stock and bond returns. Our first response is that comprehensive data for other variables—both over time and across countries—practically do not exist. Second, the financial yields that we included in the regressions served as controls or proxies for some of these variables because lower initial prices reflect higher risk aversion or higher risk premiums, among other things. Third, the effects that we captured 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 various robustness tests that we ran make us confident that our results are not spurious.

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

Demographic Shares. For the first set of results, we used shares of total population by age group, $s_{i,t}^{(j)}$, as the regressors. **Table 1** reports the results from Equation 2, and **Figure 1** plots the *implied* polynomial coefficients for each of the age cadres relative to Equation 1. The three columns of Table 1 under "Demographic Shares" show coefficients and t-statistics for regressions using each of our three dependent variables: five-year stock and bond excess returns and five-year growth in per capita GDP. We estimated all regressions by using pooled ordinary least-squares (OLS) analysis and included country dummies (or fixed effects) to control for any unobserved characteristics peculiar to each country. We corrected standard errors for heteroscedasticity and cross-sectional correlation (time clusters).[22]

The coefficients of the dividend yield and the 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

Table 1. Regression Results, 1950–2010

	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 (×1)	–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 (×10)	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 (×100)	–0.83	–2.72		11.81	2.96	
	(3.05)	(2.74)		(4.25)	(2.58)	
D_4 (×1,000)		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
$k = 3 \rightarrow k = 2$	0.3%	0.0%	25.7%	7.7%	3.4%	13.4%
$k = 4 \rightarrow k = 3$	28.4%	1.8%	12.4%	0.0%	0.1%	53.6%

Notes: This table reports coefficients and *t*-statistics (in parentheses) for six separate panel regressions, one in each column. Standard errors are corrected for heteroscedasticity and cross-sectional correlation (time clusters). The dependent variables are annualized five-year growth rates of per capita GDP, 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 therein, $\Delta s_{i,t}^{(j)}$, 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.

found a negative and statistically significant relationship between GDP growth and the three-month yield. The intuition for this relationship is simple: The Fed keeps the short yield at lower levels during recessions, which are inevitably followed by periods of faster growth in GDP.

We begin the discussion of the demographic results with the choice of *k*, the polynomial degree. The last two rows of Table 1 present the *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 the 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 $k = 4$ (a *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 on the order of 30 percent. The statistical significance of the demographic polynomial coefficients is usually not as high as that of the financial variables, yet most of the *t*-statistics are close to or above 3. This outcome 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 plotted the implied coefficients for each age group in Figure 1, together with 90 percent two-sided confidence intervals (shaded areas).

The graphs in Figure 1 and **Figure 2** deserve some explanation. Consider Panel A in Figure 1. The solid line shows the implied coefficient that links 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 a dependent variable. We can see a negative link for

Figure 1. Effect of Demographic Shares

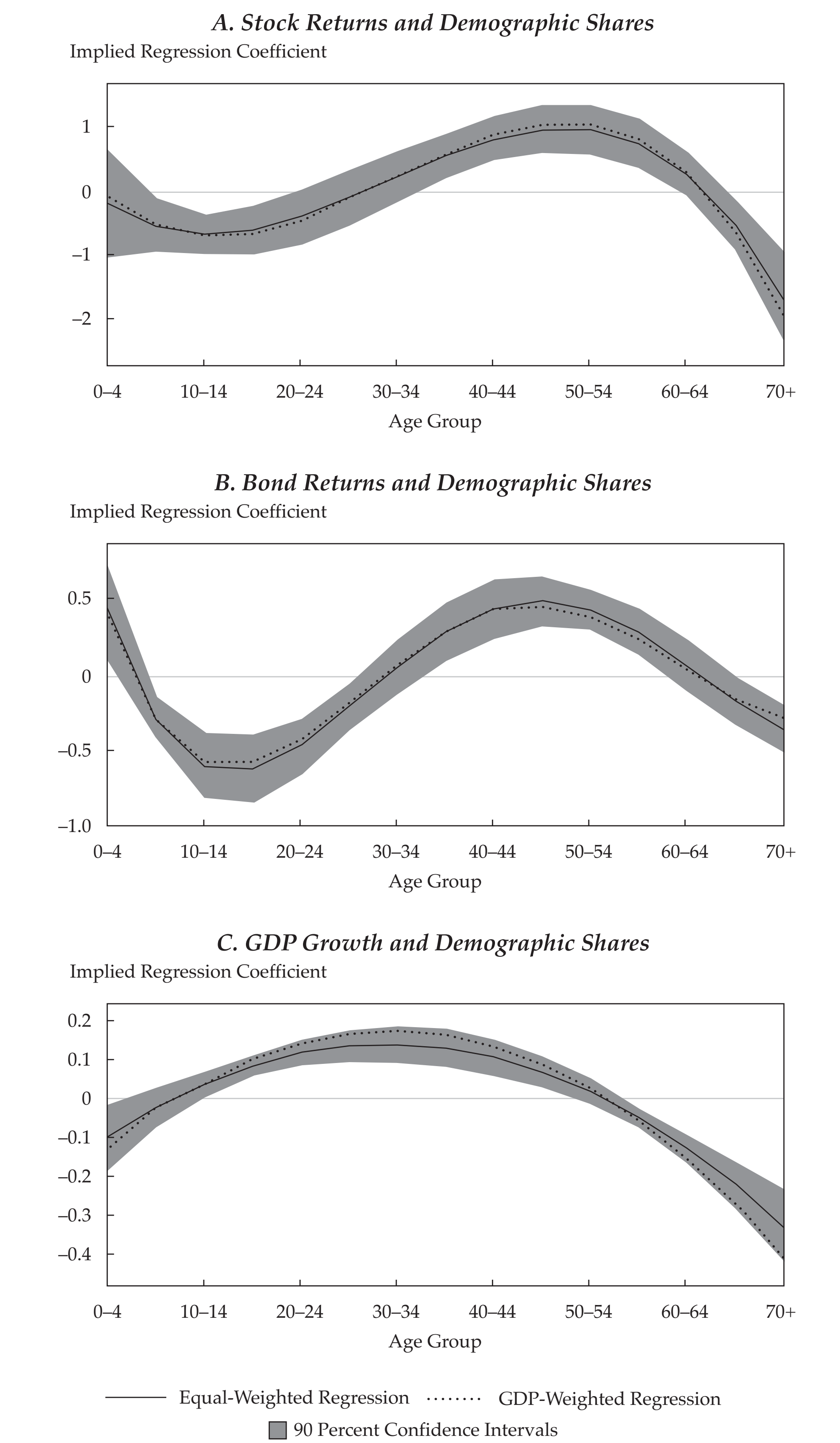
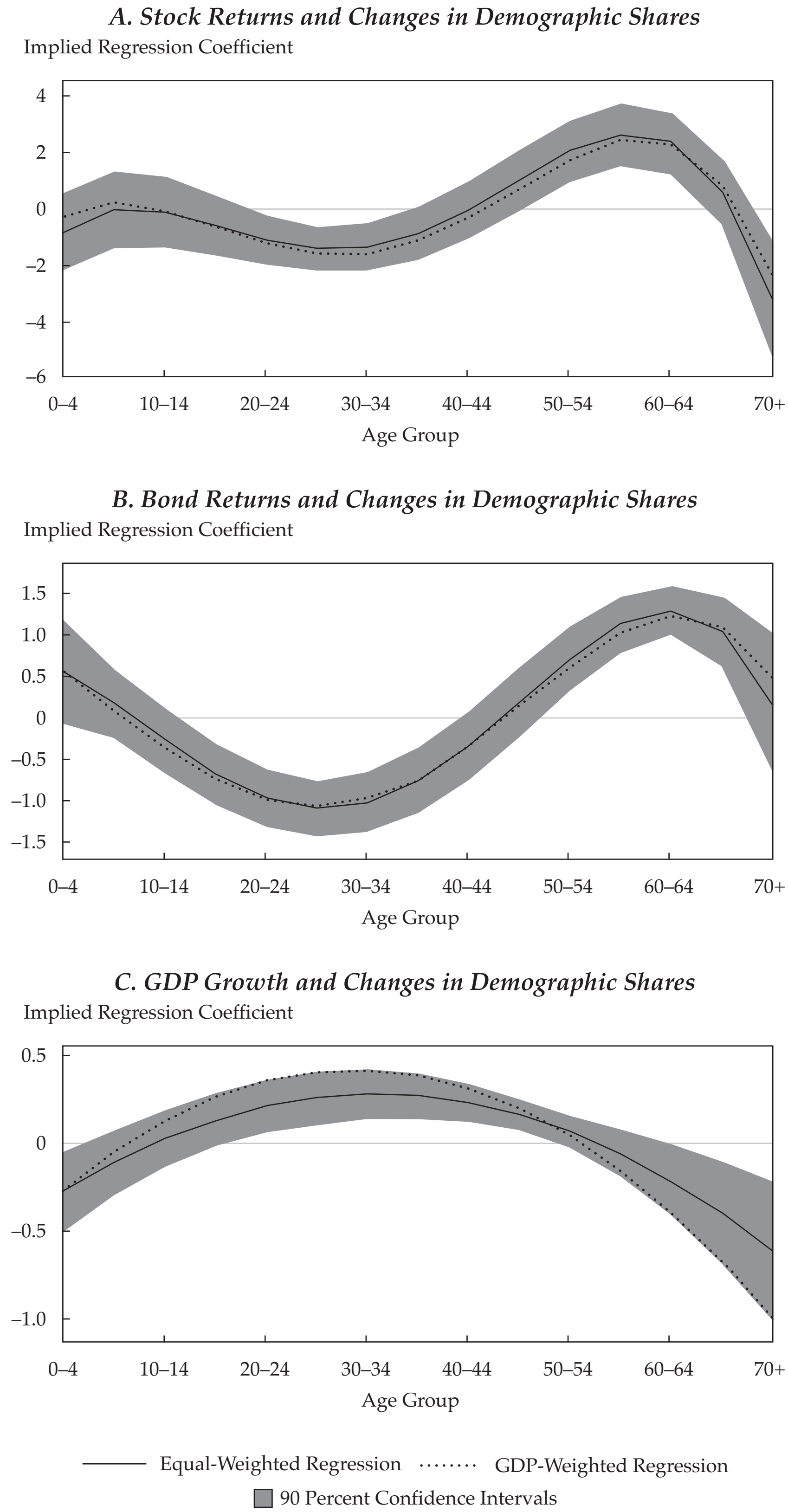


Figure 2. Effect of Changes in Demographic Shares



the first six age cadres (0–29), with statistical significance at the 90 percent level for the three age cadres from 5 to 19. This finding 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 for the 10–14 age cadre, with a coefficient of –0.7, which means that a 1 percent higher concentration in this age cadre results in 0.7 percent lower annual stock returns. That is a big number for a 1 percent change in 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 over five years. Note also that the coefficient for the 70+ age cohort is slightly less than –1.5, which is important in some of the faster-aging countries.

The OLS estimation implicitly assigned the same weight to all countries in our sample. To ensure that our conclusions held unaltered in larger countries (e.g., the United States has a PPP-adjusted GDP that is roughly 100 times that of New Zealand), we estimated the same regressions by using pooled weighted least-squares (WLS) analysis. Weighting by GDP (the natural choice), however, would create the opposite imbalance.

Using the reciprocal of the Herfindahl Index (HI) as a measure of the "effective" number of observations in a dataset, we found that our sample shrank from 22 equally weighted countries to only about 5 GDP-weighted countries.[23] Thus, we chose to weight the regressions by the square root of PPP-adjusted GDP, measured in 2005 international dollars. This choice represented a balance and still left us with 14 "effective" countries. The dotted lines show the results for these regressions. That every one of these lines in Figures 1 and 2 lies well within the confidence bands for the equal-weight regressions is reassuring. Thus, our first important robustness test has helped confirm the merits of our results.

Comparing the three panels in Figure 1, we can see that 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 real per capita GDP growth. Although larger age groups in the early 20s clearly benefit economic growth, they hurt the performance of stocks and bonds. This effect may be due to the relatively lower savings and higher debt levels of young adults, perhaps in an attempt to smooth consumption over their lifetimes (young adults may be more rational than the older generation thinks!). The opposite holds for the middle-aged groups in their 50s, which affect GDP growth negatively (already, even before they stop working!) and financial assets positively: Apparently, productivity starts to decline and savings to accelerate long before retirement.[24]

One common theme emerges from all three panels in Figure 1: Large populations of retirees (65+) seem to erode the performance of financial markets as well as economic growth. This finding 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 macroeconomy. This effect is less pronounced for bonds, presumably because they are sold later in retirement than stocks, in conformity with widespread financial advice.

Using the age group polynomials to estimate demographic effects has several benefits. The first benefit, as mentioned earlier, is the possibility of estimating the *joint* effect of *all* demographic variables in the regressions, which allows us to use all the available information in the age distributions, even after imposing a (possibly) restrictive polynomial structure on the coefficients. Second, instead of using *ad hoc* age group selections as others have done (e.g., 35–45), we can use the resulting polynomials to tell 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 nonstationarity 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 the following:

1. We used 60 years of data for most countries. Although this horizon might seem unimportant in terms of demography, significant changes in population structure occurred in most countries over 1950–2010. Indeed, Japan went from being the youngest country in the developed world to being the oldest, with its median age *doubling*, from 22 to 44, in just 60 years.
2. With a large cross section of countries, we increased the power of our tests by exploring the varying stages in the demographic transition process experienced by various nations.[25]
3. The polynomial curves show very significant differences among the implied coefficients for the various age groups (e.g., the coefficients of younger and older cohorts have opposite signs). This finding is a strong indication that the empirical methodology also explores variations across the age groups—$s_{i,t}^{(m)}$ versus $s_{i,t}^{(n)}$—which are arguably less persistent and more volatile.

Changes in Demographic Shares. We repeated the same analysis on changes in the demographic shares of each age group, $\Delta s_{i,t}^{(j)}$, which exhibit more independence than the demographic shares, $s_{i,t}^{(j)}$, both relative to neighboring five-year age cadres and relative to prior and subsequent five-year spans.[26] Using these alternative regressors, which largely avoid any risk of nonstationarity, we found similar results, which we hope will answer any remaining criticisms.

The last three columns in Table 1, under "Change in Demographic Shares," together 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 reported earlier, but we can see a decrease to 17 percent in the case of GDP growth. The degree of the polynomial is the same for both bond returns and GDP growth; for stock returns, however, the Wald test strongly favors a quartic over a cubic curve (a *p*-value of zero). The magnitude and statistical significance of the financial variables are almost the same as in the previous tests.

Comparing the demographic coefficients with those from the previous tests (the 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 are 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 2, reflecting the lower volatility of $\Delta s_{i,t}^{(j)}$ when compared with $s_{i,t}^{(j)}$. The polynomials for stocks and bonds have the same pattern as before but seem to shift to the right. This finding 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 as much as a decade or more.[27] 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 Tests. We conducted a third and more conventional series of robustness tests by checking our hypothesis on a new set of *different* countries. Unfortunately, apart from the developed markets, there is a dearth of reliable financial data beyond the data we used in the previous tests. Nevertheless, the two main datasets in our GDP tests—the UN (population) and the Penn World Table (national income accounts)—span about 200 countries. Thus, our robustness tests were restricted to economic activity only; stock and bond market data are simply unavailable or too sparse to be useful.

To get around the lack of financial yields as control variables, we used the log of the ratio of consumption to GDP, $\log(C_{i,t-1}/GDP_{i,t-1})$, as a proxy for business cycles. Cochrane (1994) showed 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, Larson, Papageorgiou, and Subramanian (2009) showed 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. Therefore, we should expect wider confidence bands.

Many of these countries are very small, with unreliable data and undiversified economies. We thus followed our earlier approach and estimated the regressions by using pooled WLS analysis, whereby each country is weighted by the square root of its 2005 GDP, measured in PPP-adjusted international dollars. This approach assigns more weight to larger economies and provides more accurate estimates, without reducing the small countries to irrelevance.[28]

Table 2 shows that this sample includes 176 *new* countries and 1,640 observations. The dependent variable (nonoverlapping five-year growth

Table 2. GDP Growth Regression Results (Robustness Tests), 1950–2010

	Demographic Shares	Change in Demographic Shares
log(C/GDP)	1.67	2.02
	(1.25)	(1.58)
D_1 (×1)	0.09	–0.09
	(4.29)	(1.19)
D_2 (×10)	–0.07	0.31
	(4.76)	(2.15)
D_3 (×100)		–0.18
		(2.49)
R^2	4%	3%
Observations	1,640	1,640
Countries	176	176
$k = 3 \rightarrow k = 2$	17.9%	1.3%
$k = 4 \rightarrow k = 3$	40.6%	40.3%

Note: This table repeats the per capita GDP growth tests in Table 1 with a different set of countries (see Table 1 for details).

rate in real per capita GDP) and the demographic variables—either $s_{i,t}^{(j)}$ or $\Delta s_{i,t}^{(j)}$—are the same as before, but we add $\log(C_{i,t-1}/GDP_{i,t-1})$ as a substitute for $3M_{i,t}$, which is unavailable in most of these smaller economies. The statistical fit, as expected, is much lower, with R^2s at 4 and 3 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 for $s_{i,t}^{(j)}$ (first column of Table 2), the demographic coefficients D_1 and D_2 are the same as their counterparts in Table 1, up to two decimal places. This finding is also confirmed by **Figure 3**, which plots the corresponding GDP polynomials from our primary tests (Panel C in Figures 1 and 2) as a dotted line. Panel A in Figure 3 depicts a remarkable similarity between the two parabolas, showing that our results with demographic shares as forecasters also hold in a much larger sample.

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

Figure 3. Robustness Tests

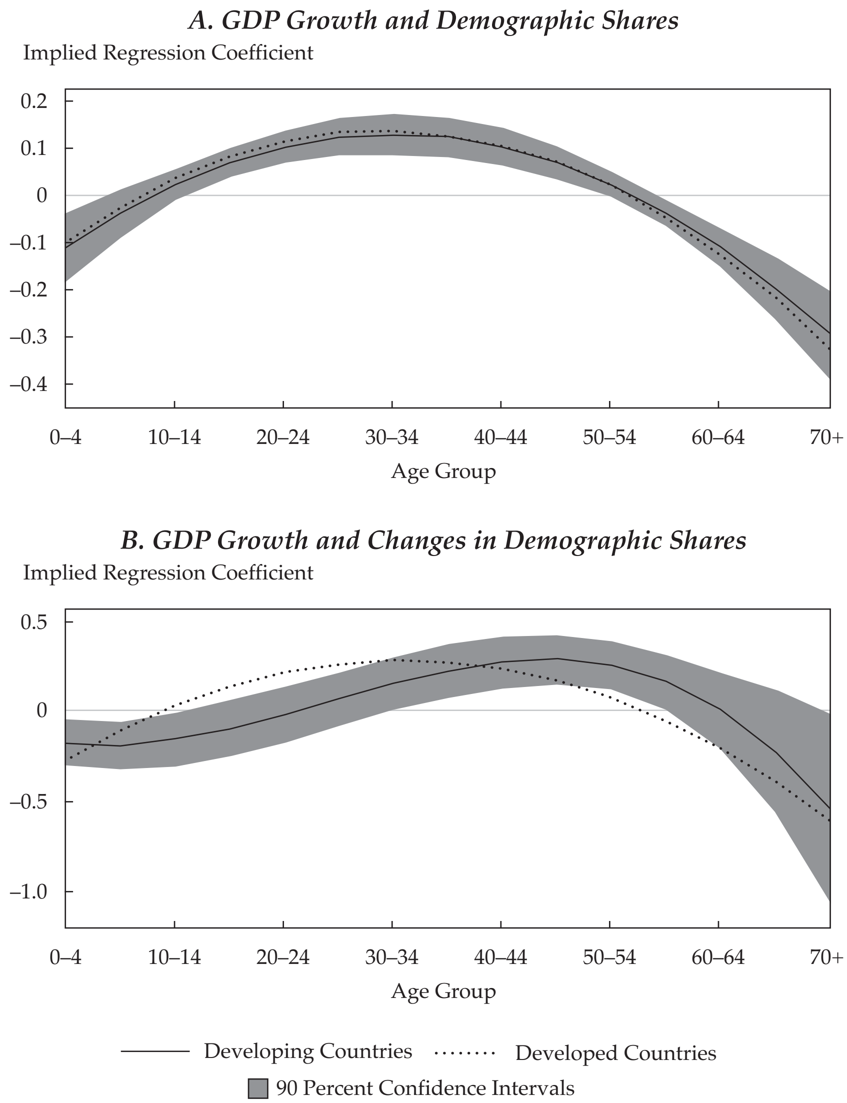


Figure 3 reveals the same positive relationship between GDP growth and middle-aged groups and the same negative relationship between GDP growth and children and retirees. In addition, the transition ages, or roots of the polynomials, are only 10 years apart. One possible interpretation of this result is simple: Most of these countries are very poor and very young. Perhaps a few additional oldsters might add stability and patience to a demography that might otherwise be fairly volatile![29]

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 predict the future accurately. 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.

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 with 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 creates an opportunity that is too tempting to resist. To the best of our knowledge, our study is the first to make systematic projections for GDP growth and financial markets on the basis of demographic changes for practically every country in the world.[30]

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

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

Further, we used only the demographic variables to construct our predictions:

$$\begin{aligned} E_t\left(r_{i,t+j} - \bar{r}_i\right) = D_1 E_t\left[\gamma_{i,t+j}^{(1)} - \bar{\gamma}_i^{(1)}\right] \\ + D_2 E_t\left[\gamma_{i,t+j}^{(2)} - \bar{\gamma}_i^{(2)}\right] \\ + \cdots + D_k E_t\left[\gamma_{i,t+j}^{(k)} - \bar{\gamma}_i^{(k)}\right], \end{aligned} \quad (4)$$

where, for example, $\bar{r}_i$ 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.

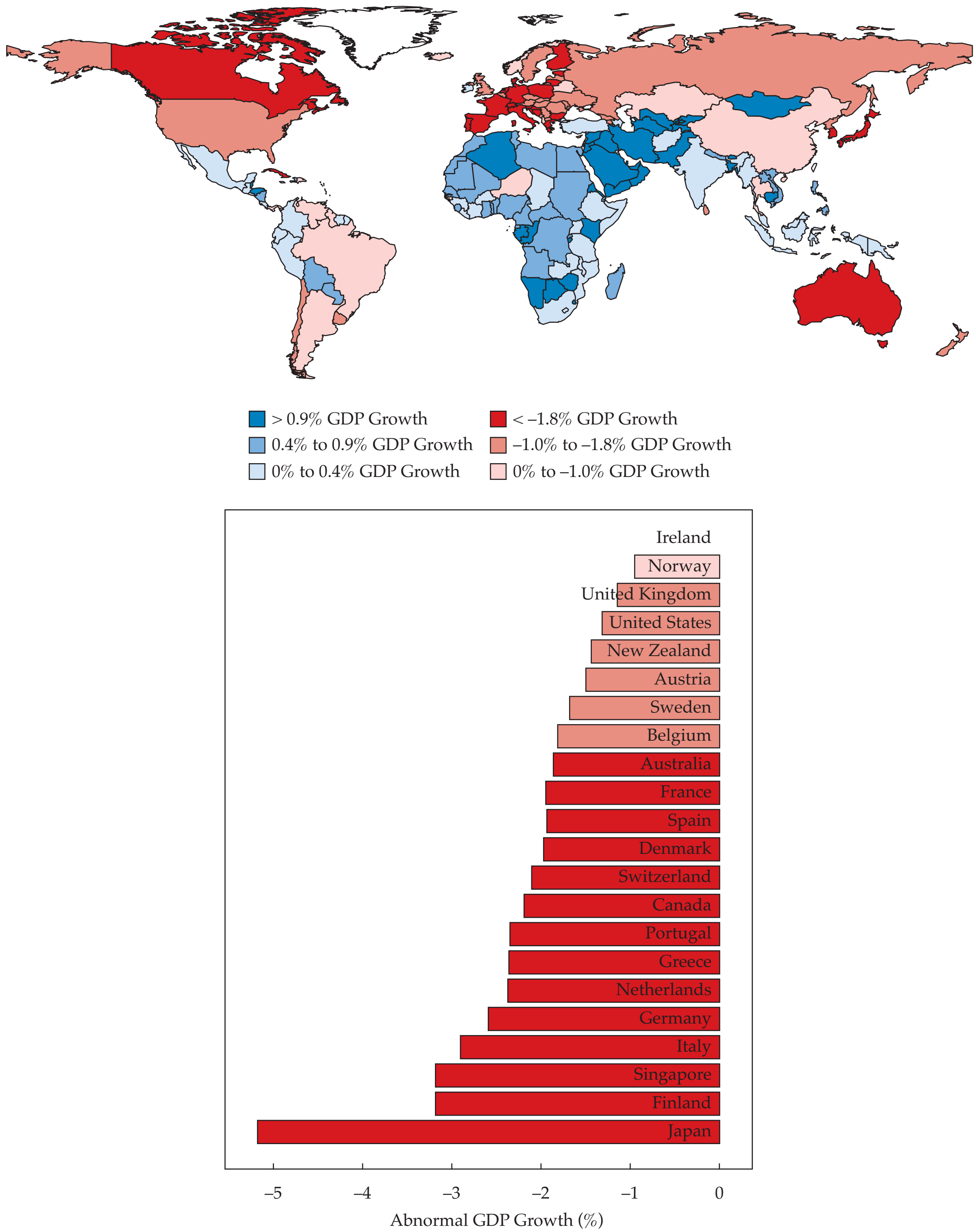
All our predictions should be seen as abnormal GDP growth or abnormal stock and bond returns. These are deviations from long-term means relative to *past individual country* performance. We chose not to include $\bar{r}_i$ in our forecasts for two reasons. First, we could include them for only a small number of countries owing 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 are in five-year intervals, we calculated decade-long forecasts for 2011–2020 as the geometric average between the forecasts for two periods: 2011–2015 and 2016–2020. For all our forecasts, we used the coefficients in Table 1. In **Figure 4** (GDP), **Figure 5** (stocks), and **Figure 6** (bonds), Panel A depicts projected demographic shares, $E_t\left[s_{i,t+j}^{(n)}\right]$, and Panel B depicts changes in projected demographic shares, $E_t\left[\Delta s_{i,t+j}^{(n)}\right]$. Each figure contains two types of graphs: color-coded maps showing results for all countries and bar charts for the 22 developed economies in our main tests. The bar charts also provide a sense of scale for the color-coded maps. The countries are first divided into positive (blue) and negative (red) groups and then into three subgroups, with approximately the same number of countries in each subgroup.

We see two common themes in the developed countries: negative values for GDP growth and positive values for abnormal excess bond returns. This finding is intuitive given their current demographic stage: A rising number of retirees combined with low birth rates creates significant pressure on output (negatively) and savings (positively). The net effect on excess stock returns varies

Figure 4. Annualized GDP Growth Forecasts, 2011–2020

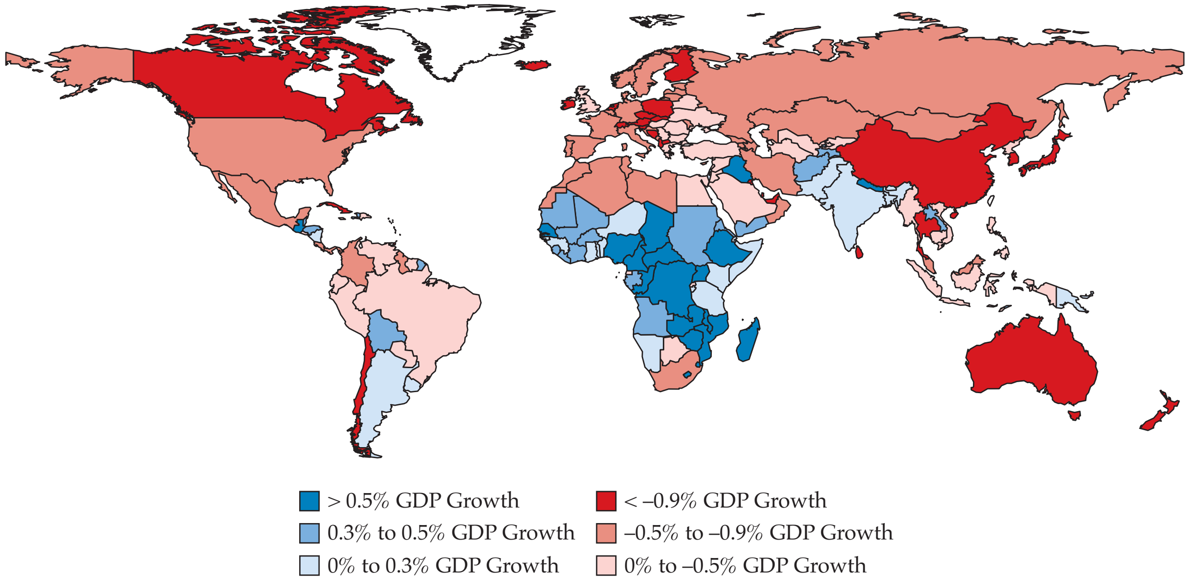
A. Forecasts Using Demographic Shares



(continued)

Figure 4. Annualized GDP Growth Forecasts, 2011–2020 (continued)

B. Forecasts Using Changes in Demographic Shares



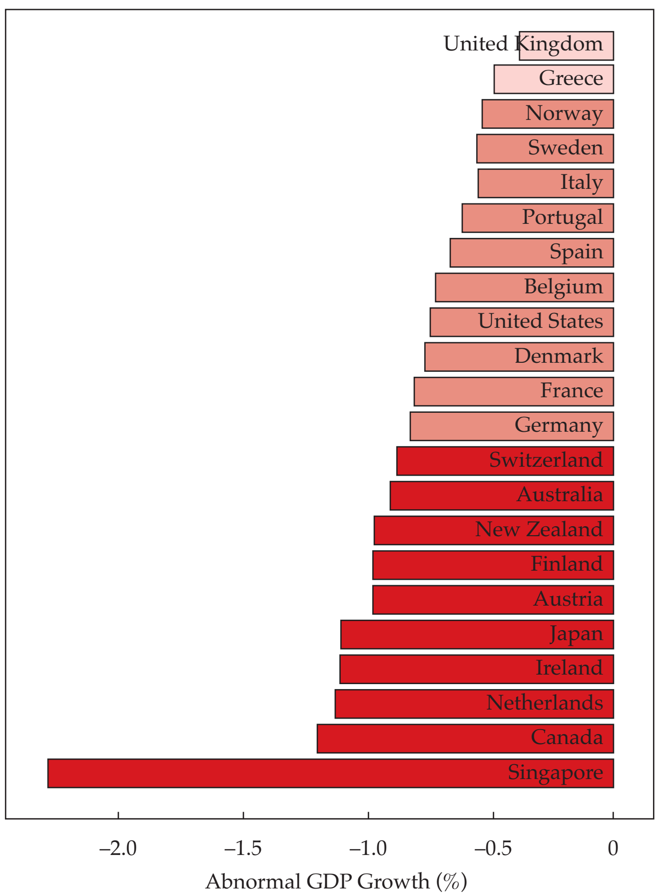
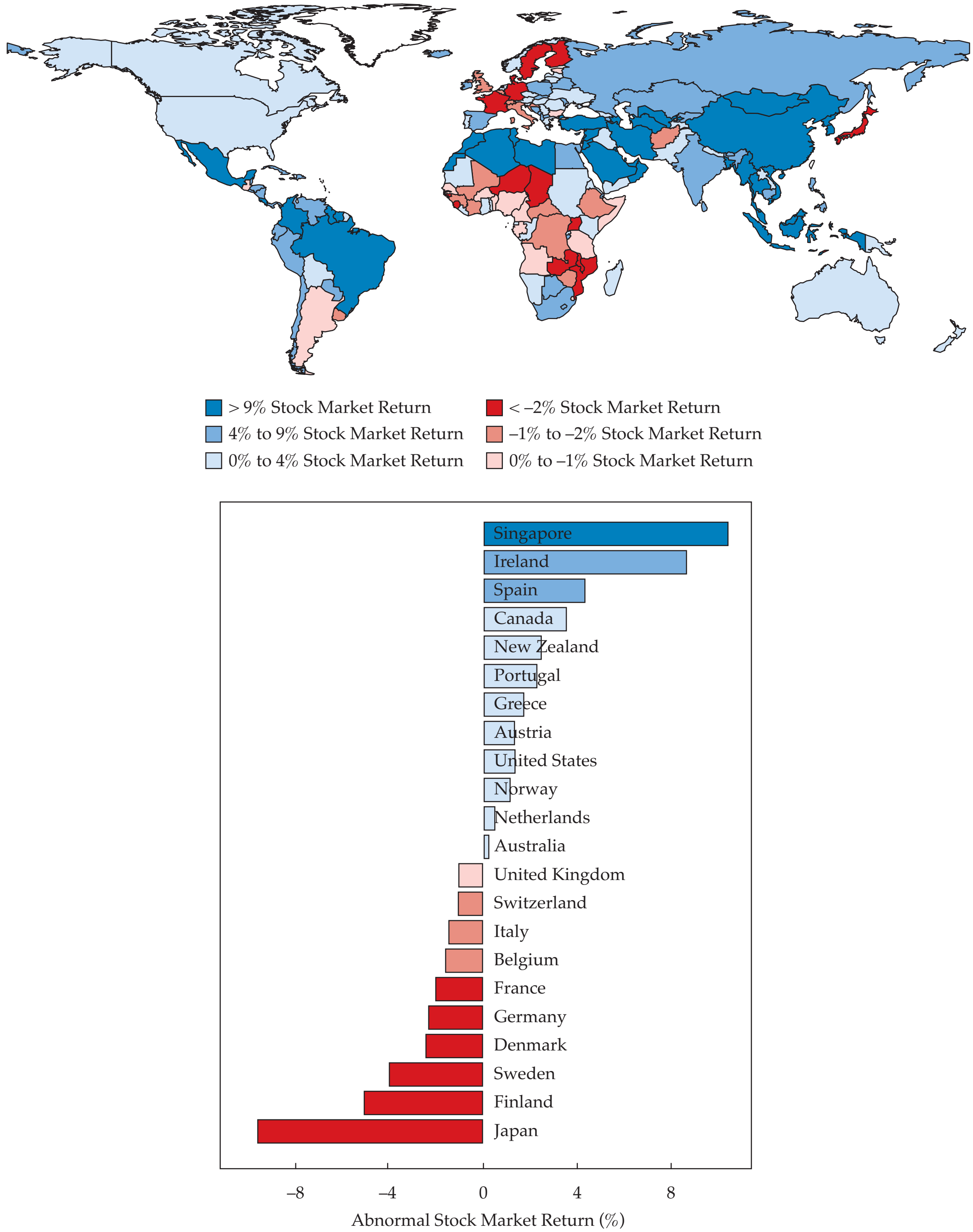


Figure 5. Annualized Stock Market Forecasts, 2011–2020

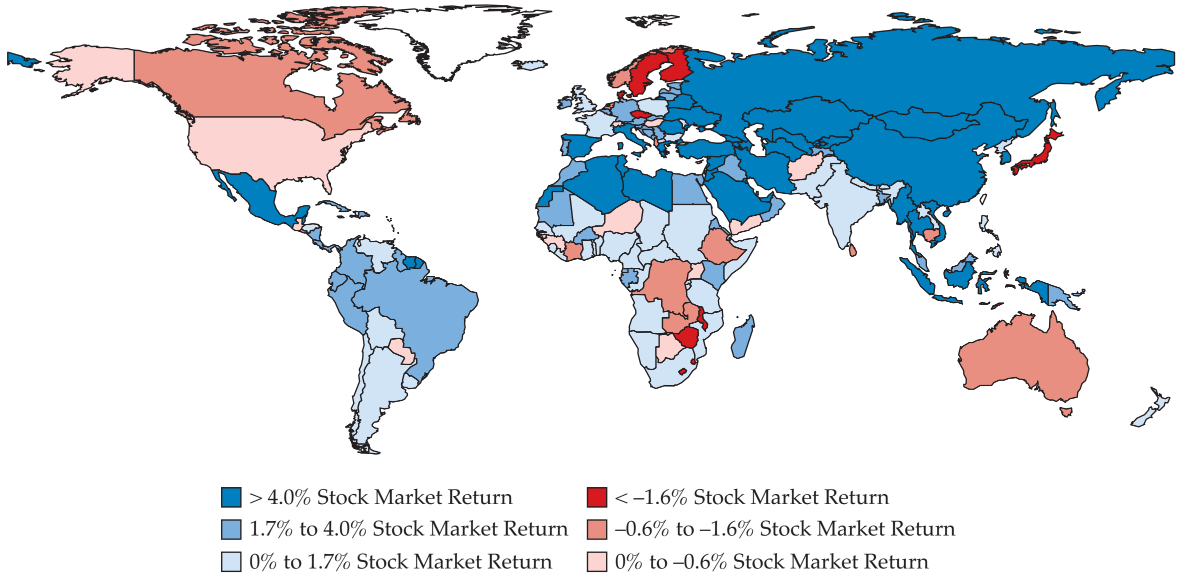
A. Forecasts Using Demographic Shares



(continued)

Figure 5. Annualized Stock Market Forecasts, 2011–2020 (continued)

B. Forecasts Using Changes in Demographic Shares



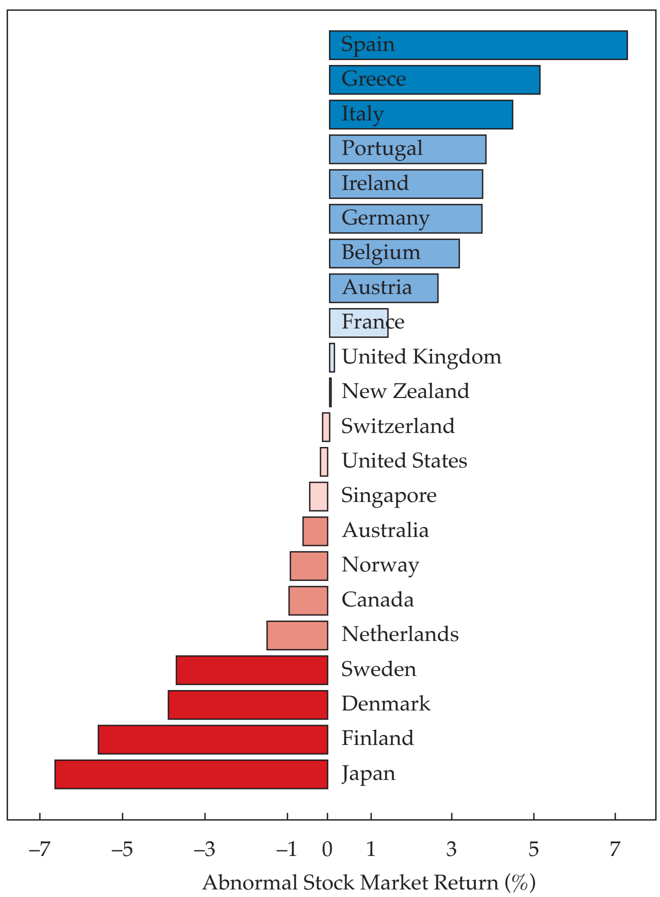
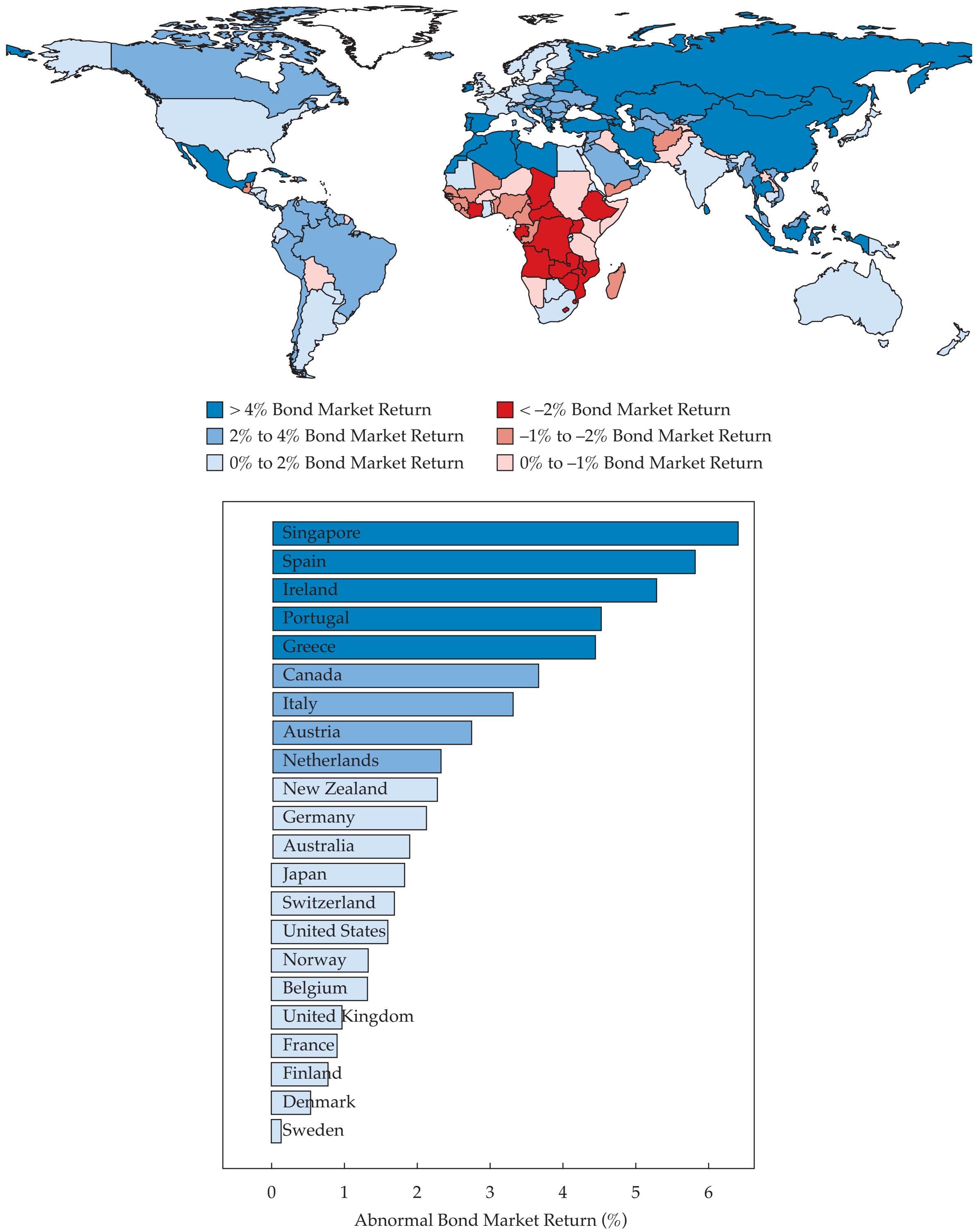


Figure 6. Annualized Bond Market Forecasts, 2011–2020

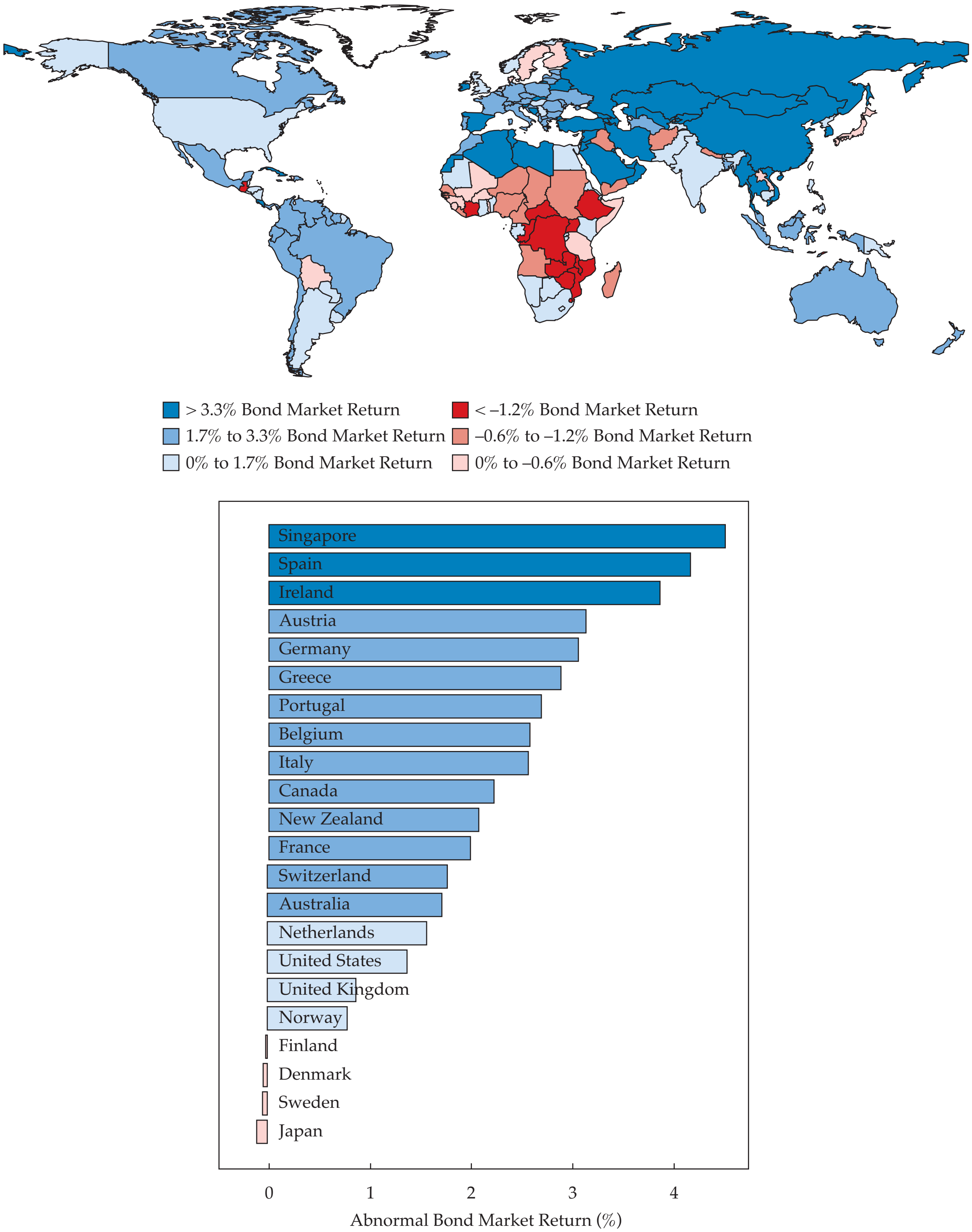
A. Forecasts Using Demographic Shares



(continued)

Figure 6. Annualized Bond Market Forecasts, 2011–2020 (continued)

B. Forecasts Using Changes in Demographic Shares



across countries; in a handful of cases, we also found divergence between the two graphs in the same figure, most dramatically for Singapore and Germany. This divergence between the forecasts with demographic shares and those with changes in demographic shares is mainly due to unusually rapid variations in some age groups, creating extreme values for $E_t\left[s_{i,t+j}^{(n)}\right]$ or $E_t\left[\Delta s_{i,t+j}^{(n)}\right]$, relative to their historical means.[31] 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 that include India and most African countries, given their large working-age cohorts and very small number of senior citizens. Bonds exhibit the opposite results: Most of the aging world has a fast-rising roster of middle-aged potential savers, leading to positive predictions for bonds, whereas Africa presents mostly negative forecasts owing to a large number of younger borrowers relative to older savers.

We recognize that many of the emerging-market results, for both stock and bond markets, are hypothetical because 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 countries.

The most interesting and extreme cases concern 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. Our findings may add 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 for those countries despite their recent troubles. Again, the emerging economies present mixed results. Latin America (including Mexico) and Asia fare relatively well, with mostly positive returns. Much of sub-Saharan Africa has too many young people and too few savings-age adults for stock markets in the region to fare well. Finally, the BRIC countries (Brazil, Russia, India, and China) seem to have a bright future ahead of them, at least for the next 10 years.

Conclusion

We studied 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 used an econometric technique that had never been used for this purpose; it allowed us to use all the available information in population profiles by fitting a polynomial curve on the regression coefficients of demographic age groups. For our main regressions, we used a large panel spanning more than 60 years and 22 countries (and 176 countries in some special cases).

By force-fitting a polynomial to our demographic variables, we extracted surprisingly strong statistical significance, with implications that align nicely with intuition. Children are not immediately helpful to GDP. They do not 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; they are overspending against their future human capital. Middle-aged adults are the engine for capital market 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; they disinvest to buy goods and services that they no longer produce.

We offered "forecasts" of GDP growth and of abnormal stock and bond market excess returns for almost every country in the world. With respect to GDP growth, our projections are bleak for most of the developed world. The developed countries—Japan, in particular—have alarmingly low birth rates and high numbers of retirees. But the picture is relatively bright for bonds and mixed for stocks, especially for the younger of the developed economies. These forecasts must be seen for what they are—namely, out-of-sample extrapolations of some reasonably powerful in-sample relationships, to be viewed as more cautionary than predictive.

Finally, we again stress 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 use in acting on these forecasts in order to position themselves for more benign circumstances. This topic is outside the scope of our article but could certainly be the subject of many future studies.

This article qualifies for 1 CE credit.

Appendix A. Construction of the Demographic Polynomial

The effect of each age group, $s_t^{(i)}$, or changes therein, $\Delta s_t^{(i)}$, 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 or bonds, or the growth rate in 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:

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

where $i = 1, \ldots, N$. Substituting these coefficients back into Equation A1 gives us

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

which we can rearrange as

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

Note 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, substituting D_0 back into Equation A3 shows us how to obtain the restricted coefficients of the demographic polynomial, D_j, with a regression:

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

We use Equation A2 to reconstruct the original demographic coefficients.

Notes

1. Purchasing power parity (PPP) adjusts income or GDP to reflect the cost of goods and services in an economy. If a high-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.
2. In other recent research, the suggestion has even been made that GDP is a poor measure of prosperity. Arnott (2011) spotlighted 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 traditional, debt-laden GDP. Although we think it is a purer measure of national prosperity and growth, we chose not to add this complication to our GDP measure because it is not yet accepted by the broad mainstream and would thus trigger controversy that would distract from the main implications of our study.
3. This observation has direct relevance for the developed world today. For roughly the last 50 years, the working-age cadre has been growing faster than the overall population by a quarter to a half percentage point a year. For most of the developed world, this dynamic will be reversed for the next 30 years. Should not this fact *alone* reduce the natural per capita real GDP growth by about a half to a whole percentage point a year relative to what we have 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.
4. This observation is a sad acknowledgment for one of the authors of this article as he endures his relentless decline in his late 50s and relies increasingly on the innovative spirit of younger adults like his coauthor, who is in his early 30s. Our results might suggest that one of the authors is saving and investing on the basis of the surging GDP contribution of the other author. Neither author would entirely reject this interpretation!
5. Brooks's simulations show effects of around 14–39 bps a year, materially weaker than our empirical estimates and forecasts. Nonetheless, the extensive literature on the equity premium puzzle, summarized by Mehra (2008), shows that market frictions—especially borrowing constraints (see Constantinides, Donaldson, and Mehra 2002)—can exacerbate returns on risky and long-term securities.
6. Nor should they be. They will typically 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.

7. 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 and geometrical growth in population would result in famine. Given the intervening two centuries of rising population and rising *per capita* longevity and living standards, his work is seen by many as discredited. Still, we should acknowledge that his study was the first serious examination of the connections between demography and health or wealth. And some view him as merely being early—very early—in his prognostications.
8. Green and Hendershott (1996) objected to Mankiw and Weil's finding that housing demand begins to decline after the age of 35. Researchers and commentators have also pointed to a variety of possible causes for the surge in housing prices in the last few decades before the recent crisis: the Fed's loose monetary policy, government housing policy, lack of proper screening by lenders, and the mismanagement of financial risks spread by the use of derivatives (collateralized debt obligations, credit default swaps, etc.). We also note that among our friends, the demand for improved housing often continues well into the 50s.
9. The U.S. Federal Reserve Board's Survey of Consumer Finances is a good example of a source of such data.
10. See Cochrane (2005) for an introduction to consumption-based asset pricing, Euler equations, and factor models.
11. See http://pwt.econ.upenn.edu/.
12. See www.un.org/esa/population/unpop.htm.
13. Other possible variants include low- or high-fertility trends, mortality rates constant at 2005–10 levels, and zero migration as of 2005–2010.
14. See http://staff.cbs.dk/parum/.
15. See www.cso.ie/.
16. For a discussion of problems and concerns with this dataset, see Johnson, Larson, Papageorgiou, and Subramanian (2009).
17. Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, and the United States.
18. 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 tacitly chose to ignore overall growth rates. Thus, we assumed that although overall population growth may materially affect overall real GDP growth, it will not affect *per capita* real GDP growth. This untested hypothesis is worth exploring in future research.
19. For a detailed discussion of multicollinearity problems, see Greene (2008).
20. 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 for the asset returns or their covariance matrix, which is akin in spirit to the solution we used in that both impose a limit on the number of parameters to be estimated.
21. 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 Appendix A for details.
22. Because the dependent variables are nonoverlapping returns and growth rates, we believe that cross-sectional correlation is the most important deviation from traditional OLS assumptions. We also corrected for time-series correlation (country clusters) and for both time and country clusters, which resulted in small and occasional differences, but our qualitative conclusions remained unchanged.
23. The Herfindahl Index is calculated as the weighted average of individual weights, or $\text{HI}=\sum_{i=1}^{N} w_i^2$. It is widely used as a measure of competition among companies in a country, market, or industry and ranges between $1/N$ (perfect competition) and 1 (monopoly). The reciprocal of this index, ranging between N and 1, serves as an indicator of the hypothetical number of equally powerful companies in a country, market, or industry. If two countries have equal weight, HI = 0.5, and so 1/HI = 2; with equal weighting, the effective number of countries matches the actual number. If two countries are weighted 90/10, the former utterly dominates the regression and HI = 0.82; so, we effectively have only 1.2 (1/HI) countries in our regression.
24. A controversial paradox in the field of demography stems from the fact that economic growth seems to be unrelated to the performance of financial markets (see, among many others, Dimson, Marsh, and Staunton 2002). Our study may help resolve this paradox. The age cadres that are most associated with GDP growth are young adults; the age cadres that are most associated with stock and bond excess returns are middle-aged adults. The young adults help GDP leap forward but are not yet serious about investing. The middle-aged, aware of their approaching retirements, are bidding stocks and bonds higher, but most are no longer advancing rapidly in their careers, whereby they once helped facilitate rapid GDP growth. Therefore, to the extent that demography drives GDP growth or capital markets, the impact is decidedly asynchronous, with about a 20-year spread between peak impact on GDP and peak impact on stock or bond excess returns.
25. Skeptics might argue that because most of the countries in our sample are developed and thus highly correlated, we were unable to conduct independent tests. We address this point in two ways. First, we refer the reader to Ang and Maddaloni (2003), who showed that using a sample of five countries increased the power of their tests by almost the same amount as augmenting their 20th century U.S. sample by a factor of 5. Second, in checking for robustness (discussed later in the article), we extended our test for GDP growth to other, less developed countries and found relatively similar but noisier results (unsurprising, given the lower quality of the data).
26. Note that a large age cadre—$s_{i,t}^{(j)}$ larger than average—can be associated with positive change $\left[\Delta s_{i,t}^{(j)} > 0\right]$ if it is rising to its peak or with negative change $\left[\Delta s_{i,t}^{(j)} < 0\right]$ if it is falling from its peak.
27. The rate of change of an age cadre must be positive before that age cadre can have above-average weight in the population.
28. If we used equal weighting, China would have the same weight as Kiribati. So, even though we would have 176 countries, we would have no confidence that our models were relevant to the major markets. If we used GDP weighting, then Brazil, Russia, India, and China (BRIC) would compose almost half the sample. The "effective" number of

countries is about 16 because of our heavy reliance on the biggest countries. An inviting compromise is to split the difference, using the square-root-of-GDP weighting: China's GDP is 100 times that of Haiti, so it gets 10 times the weight. The effective number of countries is about 75 because we rely on the smaller countries only lightly, and yet BRIC still get a 14 percent weight vis-à-vis the 176 countries. With this compromise, we can have some confidence that our models are useful for the major markets.

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

30. Although we recognize that most countries do not have well-established and well-functioning financial markets, the questions posed here are still interesting because the answers to them inform us about deeper aspects of the population structure of various countries. For instance, one can learn a lot about the demographic pressures on national savings, investments, and productivity.

31. 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 its neighbors. Thus, although it will experience similar developments in the 0–39 age group, 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 age groups.

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):551–578.

Ang, Andrew, and Angela Maddaloni. 2003. "Do Demographic Changes Affect Risk Premiums? Evidence from International Data." NBER Working Paper 9677 (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. 1–2 (March–April):359–403.

Arnott, Robert D. 2011. "Does *Unreal* GDP Drive Our Policy Choices?" *Fundamentals* (April):1–3.

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

———. 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." CRR Working Paper 2004-23, Center for Retirement Research at Boston College (October).

Brooks, Robin. 2000. "What Will Happen to Financial Markets When the Baby Boomers Retire?" Working paper, International Monetary Fund (April).

Campbell, John Y. 1987. "Stock Returns and the Term Structure." *Journal of Financial Economics*, vol. 18, no. 2 (June):373–399.

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.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2002. *Triumph of the Optimists*. Princeton, NJ: Princeton University Press.

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*, no. 1:241–307.

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.* Upper Saddle River, 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).

Higgins, 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):137–180.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2009. "Is Newer Better? Penn World Table Revisions and Growth." Working Paper 191, Center for Global Development (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 (August):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).

