Attention, Demographics, and the Stock Market*

Stefano DellaVigna
UC Berkeley
sdellavi@econ.berkeley.edu

Joshua M. Pollet
Harvard University
pollet@fas.harvard.edu

This version: November 21, 2003

Abstract

Do investors pay enough attention to long-term fundamentals? We consider the case of demographic information. Cohort size fluctuations produce forecastable demand changes for age-sensitive sectors, such as toys, bicycles, beer, life insurance, and nursing homes. These demand changes are predictable once a specific cohort is born. We use lagged consumption and demographic data to forecast future consumption demand growth induced by changes in age structure. We find that demand forecasts predict profitability by industry. Moreover, forecasted demand changes 5 to 10 years in the future predict annual industry returns. One additional percentage point of annualized demand growth due to demographics predicts a 4 to 6 percentage point increase in annual abnormal industry stock returns. However, forecasted demand changes over shorter horizons do not predict stock returns. The predictability results are more substantial for industries with higher barriers to entry and with more pronounced age patterns in consumption. A trading strategy exploiting demographic information earns an annualized risk-adjusted return of 5 percent. Our results are consistent with underreaction to information about the distant future.

*We thank John Campbell, Liran Einav, Claudia Goldin, Caroline Hoxby, Michael Jansson, Lawrence Katz, David Laibson, Ronald Lee, Ulrike Malmendier, Ashley Pollet, Jack Porter, Matthew Rabin, Andrei Shleifer, Jeremy Stein, Geoff Tate, Tuomo Vuolteenaho, seminar participants at Trento, Berkeley, Harvard, Stanford, and the participants at the NBER Behavioral Finance Program Meeting for their comments. Jessica Chan, Fang He, Lisa Leung, Shawn Li, Fanzi Mao, Rebbecca Reed, and Terry Yee helped collect the dataset of industries. Christine Ayee and Saurabh Bhargava provided excellent research assistance. We thank Ray Fair and John Wilmoth for making demographic data available to us. For financial support, DellaVigna thanks the CEDA and the Academic Senate in Berkeley.

1 Introduction

According to the theory of efficient financial markets, stock prices should reflect all available information. However, evidence on post-earnings announcement drift and short-horizon momentum effects suggests that stock prices do not fully adjust to new information.

We suggest and implement a novel test of underreaction to information. We examine whether investors respond appropriately to changes in the demographic structure. One unusual feature characterizes these changes—they are forecastable years in advance. Current cohort sizes, in combination with mortality and fertility tables, generate accurate forecasts of future cohort sizes even at long horizons. Different goods have distinctive age profiles of consumption, and therefore forecastable changes in the age distribution produce forecastable shifts in demand across goods. These shifts in demand induce predictable changes in profitability for industries that are not perfectly competitive. Stock returns of companies in these industries should respond to such shifts.

The stock market reaction to forecastable demographic changes is a test of investor attention to determinants of profitability. In particular, we can use demographic variables to estimate separately the response of returns to short-term and long-term changes in profitability. This test is quite different from other tests of predictability based on announcements of quarterly earnings or previous performance information measured by recent returns or accounting ratios. These variables convey information about profitability that is not easily decomposable into a short-term component and a long-term component.

If investors fully incorporate the short-term and long-term effects of demographic variables, the forecastable changes in profitability due to demographic information will be reflected in the stock prices of companies within the related industries. If investors, instead, are inattentive to demographics, this information will be incorporated gradually. As a consequence, demographic variables will predict industry asset returns.

We address two questions. First, how do markets evaluate information about short-term and long-term outcomes? Beyond demographics, other factors that affect long-term profitability include new plant openings and research and development (Hall and Hall, 1992). The evidence in this paper complements the existing results on the response to short-term events, such as earnings surprises. Second, do individuals consider demographic variables when making decisions about job choices, public policy, and portfolio allocations? While this paper focuses on investor behavior, the findings have broader implications for other economic decisions affected by demographics. This question is of particular relevance given the large demographic changes faced by American and European societies.

We illustrate the basic idea of this paper with an example. Assume that a large cohort is born in 2004. This large cohort will increase the demand for school buses as of 2010. As long as the school bus industry is not perfectly competitive, the companies in the industry will enjoy an increase in abnormal profits in 2010. When should stock prices increase in anticipation of higher future profitability? According to the standard analysis with attentive investors, the marginal investor foresees the positive demand shift induced by demographic changes and invests in school bus stocks in 2004. The price of school bus shares should increase in 2004 until the opportunity to receive abnormal returns has dissipated. However, if investors are inattentive to demographic changes, stock prices will not increase sufficiently in 2004 and will move upward subsequently. Therefore, the industry abnormal return between 2005 and 2010 will be predictable using the information available in 2004.

This example motivates a simple test of attention with respect to demographic changes. In a model with attentive investors, forecastable fluctuations in cohort size do not generate predictability for stock returns, because stock prices react immediately to the demographic information. Under the alternative hypothesis of inattention, current demographic information forecasts future stock returns.

In Section 3 we implement this test of attention to demographic information for a set of 47 US industries over the period 1935-2002. We define industries in an effort to separate goods with different age profiles in consumption and yet cover all final consumption goods. Several goods have an obvious association with a demographic cohort. In the life cycle of consumption, books for children are followed by toys and bicycles. Later in the life cycle, individuals consume housing, life insurance, and pharmaceuticals. Sadly, the life cycle ends with nursing homes and funeral homes. Other expenditure categories, like beer, food, and property insurance, have a less obvious association with a specific age group.

The empirical strategy follows six steps, described in detail in Section 3. In the first step, we use current cohort sizes, mortality tables, and fertility rates to forecast future cohort sizes. The forecasted one-year cohort growth rates for the ages 10-70 are very close to the actual ones, with an R^2 of about 80 percent. Forecasts for future births and older generations are substantially less precise, due to unforecastable fluctuations in fertility and old-age mortality.

In the second step, we estimate age-consumption profiles for the 47 goods in the sample. We use historical surveys on consumer expenditure from 1935-36, 1960-61, and 1972-73 to complement the more standard Consumer Expenditure Survey for the years 1983-85. We find that: (i) consumption of most goods depends significantly on the demographic composition of the household; (ii) across goods, the age profile of consumption varies substantially; (iii) for a given good, the age profile is quite stable across the different surveys. These findings support the use of cohort size as a predictor of consumer demand.

In the third step we combine the demographic forecasts with the age profiles of consumption for each good. The output is the good-by-good forecasted demand growth caused by demographic changes. The forecasted one-year growth rates of consumption have an average within-good standard deviations of .61 percent. We identify the 20 industries with the highest within-good standard deviation of growth as the subsample most affected by demographic changes.

In the next three steps, we match the consumption forecasts with accounting information from *Compustat* and stock returns data from *CRSP*. In order to perform this match, we disaggregate the industry classification beyond the 4-digit SIC code level. For example, we separate the SIC codes for book producers 2730-2739 into 4 industries depending on the targeted age group. In the fourth step we examine whether the forecasted consumption growth predicts profitability for companies in an industry. For the subset of demographics-exposed industries, the accounting return on equity increases by approximately 1.5 percentage points for each additional percentage point of contemporaneous demand growth induced by demographics. For the whole sample of industries, these effects are somewhat smaller. The response of profitability to demand changes is more pronounced in industries with a more concentrated industrial structure, a proxy for barriers to entry or market power.

In the fifth step, we test for underreaction to demographic information using stock returns. We regress beta-adjusted annual industry stock returns on measures of medium-term and long-term forecasted demand growth. The medium-term measure is the forecasted annualized growth rate of consumption due to demographics over the next 5 years. The long-term measure is the forecasted annualized growth rate of consumption during years 5 to 10. We find that long-term demand growth forecasts annual stock returns. A one percentage point increase in the annualized demand growth rate due to demographics predicts a 4 to 6 percentage point increase in abnormal industry return. The effect of medium-term demand growth on returns is negative but not statistically significant.

The predictability of returns is more substantial in sectors with high concentration. Industries with concentration ratios above the median exhibit predictability that is approximately twice as large compared to industries with concentration ratios below the median. However, our estimates of this pattern are imprecise. The predictability results are stronger over the last 25 years, during which both consumption and industry returns are likely to be measured more accurately.

Finally, in the sixth step we examine the returns of a trading strategy designed to exploit the predictability of industry-level stock returns using demographic information. We construct a zero-investment portfolio that is long in industries with high forecasted demand growth for years 5 to 10 and short in industries with low forecasted demand growth. For the subset of demographic industries, this portfolio outperforms various factor models by more than 5 percent per year. The outperformance is approximately 50 percent larger during the period from 1975 to 2002. For the sample including all industries, the portfolio outperforms the benchmark portfolios by approximately 2 percentage points per year. A portfolio constructed using only high-concentration industries earns annualized abnormal returns of 4 percentage points.

In Section 4 we interpret these results within the framework of the model of short-sighted investors described in Section 2. We assume that investors only consider information about future profitability within a horizon of h years. For the periods further into the future, investors use a combination of a parametric estimate for the long-term growth and an extrapolation from the near-term forecasts. This model embeds the standard framework as a limiting case as h approaches infinity. Evidence from I/B/E/S, one of the most comprehensive data sets for analyst forecasts, suggests that the horizon h may be between 3 and 5 years. While forecasts of earnings 1, 2, or even 3 years into the future are available for most companies, earnings forecasts beyond the 4 year horizon exist for less than 10 percent of the sample. If analysts do not produce long-term forecasts of earnings, most investors are unlikely to have access to information about long-term profitability.

For a horizon h of approximately 5 years, the model of short-sighted investors suggests that forecasted demand growth 5 to 10 years ahead should predict industry stock returns. Forecastable demographic shifts occurring 5 years in the future are neglected by investors at the beginning of the year. As the year unfolds, investors notice the upcoming shifts and react accordingly. Furthermore, predictability should be more substantial for industries with higher concentration. In the presence of higher barriers to entry, demand changes have a stronger impact on profitability, and consequently, stock returns. Moreover, the model can match the estimated magnitude of the effects.

In Section 4 we briefly discuss limits to arbitrage and consider some alternative interpretations of the results. An alternative attention-based interpretation is that individuals underreact to all slowly-moving variables. We also discuss rational predictability and improvements in data analysis as possible explanations.

The most relevant literature in finance considers the forecastability of returns for individual stocks, industries, or specific portfolios using publicly known information including previous returns (de Bondt and Thaler, 1985; Jegadeesh and Titman, 1993; Moskowitz and Grinblatt 1999; Hong, Touros, and Valkanov, 2003), accounting ratios (Basu 1983; Fama and French 1992), and earnings announcements (Watts 1978, Bernard and Thomas 1989). In particular, stock returns display positive autocorrelation at short horizons (momentum) and positive earning surprises are followed in the subsequent dates by positive abnormal returns (post-earnings announcement drift) in the subsequent weeks.

Two explanations of these phenomena rely on slow diffusion of information (Hong and Stein, 1999) or fluctuations in overconfidence (Daniel, Hirshleifer, and Subrahmanyam, 1998). Since our forecasts use information that has been in the public domain for more than one year and possibly for many years, our findings do not appear to be directly linked to either explanation.

This paper is also related to the literature on the effects of demographic variables on economic outcomes, including social security (Auerbach and Lee, 2001; Gruber and Wise, 1999, and Hurd, 1990), college graduation (Card and Lemieux, 2000 and Bound and Turner, 2003), research and development (Acemoglu and Lin, 2003), macro variables (Fair and Dominguez, 1991), and family choices (Easterlin, 1987). Zarkin (1985) considers the choice of college students to be certified for elementary school and high-school teaching. He finds that students respond strongly to current wages but display essentially no response to future wages predicted by forecastable changes in cohort size.

Another closely linked segment of literature has examined the relationship between demographic changes and the aggregate demand for stocks and bonds. The evidence regarding an effect on the equity premium is mixed (Poterba, 2001; Goyal, forthcoming; Ang and Maddaloni, 2003; Geneakoplos, Magill, and Quinzii, 2002). The test we undertake differs markedly from this literature because it focuses on changes in demand across consumption goods rather than on aggregate shifts in demand for financial assets.

Mankiw and Weil (1989) find that contemporaneous changes in the age structure for the population partially explain the time-series behavior of housing prices. We generalize their approach by analyzing 47 industries and examining stock market returns where, unlike for housing prices, arbitrage should eliminate predictability. While we also find evidence of predictability, the predictability of stock returns is due to future, rather than to contemporaneous demographic information.

Finally, this paper also contributes to the growing empirical literature regarding the role of attention allocation within economics. (Barber and Odean, 2002; Gabaix et al., 2002; Huberman and Regev, 2001). The empirical evidence in this paper suggests that individuals may simplify complex decisions by neglecting long-term information.

The rest of the paper is structured as follows. Section 2 presents a stylized model of the impact of demand changes on stock returns in the presence of short-sighted investors. Section 3 describes the six steps of the empirical analysis, from the forecasts of cohort size to the portfolio performance. Section 4 interprets the empirical results in light of the model from Section 2 and discusses alternative interpretations. Section 5 concludes.

2 A model of inattention

While this Section is focused on the implications of inattention for stock returns, we start by sketching a model of the effects of a demand change on firm profitability. This is important since, as we will show, firm profitability mediates the impact of demand changes on stock prices.

Industrial structure. We model the industrial structure as a two-stage game (Mankiw and Whinston, 1986). In the first period, a set of potential entrants decides whether to pay a fixed cost K and enter an industry. In the second period, the N firms that paid K in the first period choose production levels $\{q_n\}$ in a Cournot game. Each firm has convex costs of production c satisfying $c(0) = 0$,$c'(.) > 0$, and $c''(.) \geq 0$. We consider symmetric equilibria in the second stage where all firms choose the same quantity q. The aggregate demand function for the market is $\alpha D(P)$ where α is a proportional demand shift capturing demographic changes. Aggregate supply Q is equal to Nq; we can write the equilibrium inverse demand function $P = P[Nq/\alpha]$. We assume $P'(Q) < 0$ and $P''(Q) < 0$ for all Q. The first assumption is simply a requirement that demand curves be downward-sloping. The second assumption is a technical requirement that guarantees strict concavity of the profit function and, therefore, the uniqueness of the solution to the profit-maximization problem in the second stage:

$$\max_q \pi(q|N,\alpha) = P\left[\frac{(N-1)\bar{q}+q}{\alpha}\right] q - c(q).$$

Consider first the effect of an short-run increase in demand from α_0 to $\alpha_1 > \alpha_0$. The demand change occurs after the decision to enter has been made, that is, for fixed number of firms N. The firms observe the level of demand α before they choose the production q^*. Given non-decreasing marginal costs, the firms increase production at most proportionally with the demand shift, that is, $q^*(\alpha_0) < q^*(\alpha_1) \leq \alpha_1 q^*(\alpha_0)/\alpha_0$. To see this, consider the first order conditions for the firms:

$$P'\left[\frac{Nq^*}{\alpha}\right]\frac{q^*}{\alpha} + P\left[\frac{Nq^*}{\alpha}\right] - c'(q^*) = 0. \tag{1}$$

If these conditions are satisfied for $q^*(\alpha_0)$ at $\alpha = \alpha_0$, it is easy to check that the left-hand side of (1) is (weakly) negative for $q = \alpha_1 q^*(\alpha_0)/\alpha_0$ at $\alpha = \alpha_1$. Since the objective function is strictly concave, $q^*(\alpha_1) \leq \alpha_1 q^*(\alpha_0)/\alpha_0$ follows. Similarly, the left-hand side of (1) is positive for $q = q^*(\alpha_0)$ at α_1, because, for constant q, an increase in α increases the left-hand side of (1). Using again the concavity of the profit function, we can conclude $q^*(\alpha_0) < q^*(\alpha_1)$. This proves the desired inequality for q^* as a function of α. In turn, this implies $0 < \partial q^*(\alpha)/\partial\alpha \leq q^*(\alpha)/\alpha$.

Second, consider the impact of a demand shift on firm profitability. The derivative of firm profits π with respect to a demand change α is

$$\frac{\partial \pi}{\partial \alpha} = -P'\left(\frac{q^*}{\alpha} - \frac{\partial q^*}{\partial \alpha}\right)\left(\frac{Nq^*}{\alpha}\right) + (P - c'(q^*))\frac{\partial q^*}{\partial \alpha} > 0 \tag{2}$$

where the last inequality makes use of $0 \leq \partial q^*(\alpha)/\partial\alpha \leq q^*(\alpha)/\alpha$, of the assumption $P' < 0$, and of the fact that the price P is higher than marginal cost $c'(q^*)$ by (1). Therefore, profitability is increasing in the demand shift α in the second stage (short-run).

While in the short-run the number of firms N is constant, in the long-run firms enter until profits are zero. Ignoring the integer constraint, the sub-game perfect equilibrium in the first stage of the game implies that $\pi(q^*,N^*,\alpha) - K = 0$. Gross profits π equal the entry cost K. A change in demand α that is known before the entry decision, therefore, does not affect the profits.

In the special case with constant marginal costs $c(q) = cq$, expression (1) implies that in the short-run $\partial q^*(\alpha)/\partial\alpha = q^*(\alpha)/\alpha$. As a consequence, expression (2) simplifies to

$$\frac{\partial\pi}{\partial\alpha} = (P - c)\frac{q^*}{\alpha} = \frac{\pi}{\alpha}.$$

If the firm is in long-run equilibrium before the change in α, the zero-profit condition implies $\pi = K$. It follows that the short-run derivative of profits with respect to small demand shifts is increasing in the entry costs K.

To summarize, long-run profits are independent of demand changes, while short-run profits are increasing in demand changes. Moreover, in the case of constant marginal costs, the effect of demand shifters is increasing in the size of the entry cost.

This model has two main implications. First, a demand change is more likely to affect profits if the entry decision takes longer, that is, if barriers to entry are higher. In this case, even if the demand change is known in advance, potential competitors cannot enter the market. Second, the higher the entry costs, the higher the response of profitability to a demand change. Both implications suggests that the responsiveness of profits to demand changes are likely to be higher in industries with higher concentration. High concentration may reflect either a long time to enter or high entry costs. In Section 3 concentration measures serve as proxies for barriers to entry and entry costs.[1]

Stock Returns. Assume that demand shifts affect profitability to some extent. How should returns of firms in an industry respond if investors are short-sighted? We use log-linear approximations for stock returns (Campbell and Shiller, 1988; and Campbell, 1991) and for accounting return on equity (Vuolteenaho, 2002). Consider a generic expectation operator (not necessarily rational), $\widehat{E}[.]$, with the properties $\widehat{E}_t[ca_{t+j} + b_{t+k}] = c\widehat{E}_t[a_{t+j}] + \widehat{E}_t[b_{t+k}]$ and $a_t = \widehat{E}_t[a_t]$. The unexpected log return can be expressed as a change in expectations about profitability and returns:[2]

$$r_{t+1} - \widehat{E}_t r_{t+1} = \Delta\widehat{E}_{t+1}\sum_{j=0}^{\infty}\rho^j roe_{t+1+j} - \Delta\widehat{E}_{t+1}\sum_{j=1}^{\infty}\rho^j r_{t+1+j} \tag{3}$$

Equation (3) relates the unexpected log return to the change in expectations about the profitability (measured by roe) and returns. In this expression, r_{t+1} is the log return between t and $t+1$, roe_{t+1} is the log of the accounting return on equity between t and $t+1$, $\rho < 1$ is a constant (interpreted as a discount factor) associated with the log-linear approximation, and $\Delta\widehat{E}_{t+1}[.] = \widehat{E}_{t+1}[.] - \widehat{E}_t[.]$ is the change in expectations between periods.

[1] A third implication is that the effect of demand changes of profits is higher if the potential entrants ignore forecastable demographic changes.

[2] The derivation of equation (3) is in Appendix A.

Short-sighted investors have correct short-term expectations but incorrect long-term expectations about profitability. Let $E_t^*[.]$ be the expectation operator for short-sighted investors at time t. Similarly, let $E_t[.]$ be the fully rational expectation operator for period t. Short-sighted investors have rational expectations regarding dividend growth for the first h periods after t, $E_t^* roe_{t+1+j} = E_t roe_{t+1+j}\ \forall\ j<h$. For periods beyond $t+h$, they form incorrect expectations of profitability based on a constant term, $\overline{roe}$, and an extrapolation from the expected (rational) periods $t+1+h-n$ to $t+h$:

$$E_t^* roe_{t+1+j} = w\overline{roe} + (1-w)\sum_{i=1}^{n} \frac{E_t \Delta roe_{t+1+h-i}}{n}\ \forall j \geq h. \tag{4}$$

Finally, we assume that short-sighted investors believe that expected returns are constant: $E_t^* r_{t+1+j} = \bar{r}\ \forall t,\ \forall\ j \geq 0$.

We consider three leading cases of the model. In the limiting case as $h \rightarrow \infty$, investors possess *rational expectations* about future profitability. If h is finite and $w = 1$, then investors exhibit *unconditional inattention*. In this situation, investors expect that the return to equity after period $t+h$ will equal a constant, $\overline{roe}$. If h is finite and $w < 1$, then investors exhibit *inattention with partial extrapolation*. Investors form expectations for the return on equity after period $t+h$ with a combination of a fixed forecast, $\overline{roe}$, and an extrapolation based on the average expected return on equity for the n periods before $t+1+h$.

This model of inattention assumes that investors carefully form expectations about profitability in the immediate future, but adopt rules of thumb to evaluate long-term profitability. In a world with costly information processing, we indeed expect agents to form more accurate forecasts of short-term than of long-term information. The further in the future is the event, the lower is the benefits from formulating a forecast. However, investors make sub-optimal forecasts by neglecting long-term demographic information because they do not realize that demographic variables provide relatively precise information about long-term profitability.

Let $E^*[.]$ characterize the expectations of a representative agent. We can substitute the short-sighted expectations, $E^*[.]$, for the generic operator $\widehat{E}[.]$ in (3) to get

$$\begin{aligned}
r_{t+1} - \bar{r} &= r_{t+1} - E^* r_{t+1} = \Delta E_{t+1}^* \sum_{j=0}^{\infty} \rho^j roe_{t+1+j} - \Delta E_{t+1}^* \sum_{j=1}^{\infty} \rho^j r_{t+1+j} = \\
&= \Delta E_{t+1} \sum_{j=0}^{h-1} \rho^j \Delta d_{t+1+j} + \rho^h \left(E_{t+1} roe_{t+1+h} - w\overline{roe} - (1-w)\sum_{i=1}^{n} \frac{E_t roe_{t+1+h-i}}{n} \right) \\
&\quad +(1-w) \sum_{j=h+1}^{\infty} \rho^j \left(\sum_{i=1}^{n} \frac{E_{t+1} roe_{t+2+h-i}}{n} - \sum_{i=1}^{n} \frac{E_t roe_{t+1+h-i}}{n} \right).
\end{aligned}$$

The last equation presents the 'unexpected' return for short-sighted investors when $E^*[.]$ governs the behavior of the representative agent. Notice that the unexpected return, $r_{t+1} - \bar{r}$, depends on the value of the return on equity only up to period $t+1+h$; as of period $t+1$, later periods are not incorporated.

Taking conditional rational expectations at time t (that is, using $E_t[.]$) and applying the law of iterated expectations, we obtain an expression for return predictability from the perspective of the fully rational investor.

$$
\begin{aligned}
E_t r_{t+1} - \bar{r} &= \rho^h w \left(E_t roe_{t+1+h} - \overline{roe}\right) + \rho^h (1-w) \sum_{i=1}^{n} \frac{E_t \left[roe_{t+1+h} - roe_{t+1+h-i}\right]}{n} \\
&\quad + \frac{\rho^{h+1}}{1-\rho} \frac{(1-w)}{n} E_t[roe_{t+1+h} - roe_{t+1+h-n}] \qquad (5)
\end{aligned}
$$

The expected return between time t and time $t+1$ depends on the sum of three terms. For rational investors ($h \rightarrow \infty$), all terms converge to zero (given $\rho < 1$) and we obtain the standard result of unforecastable returns. For investors with unconditional inattention (h finite and $w=1$), only the first term applies: $E_t r_{t+1} - \bar{r} = \rho^h w \left(E_t roe_{t+1+h} - \overline{roe}\right)$. Returns between year t and year $t+1$ are predictable using the difference between the expected return on equity $h+1$ years ahead and the constant $\overline{roe}$. For inattentive investors with extrapolation (h finite and $w=0$), only the last two terms apply. Returns depend positively on the expected return on equity $h+1$ years ahead and negatively on the expected return on equity in the previous n years.

In general, for inattentive investors (h finite), stock returns between time t and $t+1$ are forecasted positively by the expected return on equity $h+1$ years ahead and negatively by the expected return on equity for the n years before $t+1+h$. Between years t and $t+1$, investors update their expectations by incorporating the expected profitability in period $t+1+h$, which was previously ignored. This information replaces the earlier forecast that was created using $\overline{roe}$ and the expected return on equity between years $t+1+h-n$ and $t+h$. Expected returns are an increasing function of the update about future profitability. This update depends positively on expected profitability in period $t+1+h$ and negatively on $\overline{roe}$ and expected profitability between $t+1+h-n$ and $t+1+h$.

Building on the discussion of the industrial structure, we assume that the return on equity, our profitability measure, responds to contemporaneous demand changes due to demographics. In particular, we model the log return on equity as a linear function of two components, demand growth due to demographics and all other factors:

$$
roe_{t+1+j} = \phi + \theta \Delta c_{t+1+j} + v_{t+1+j} \qquad (6)
$$

where Δc_{t+1+j} is the growth rate of demand due to demographics, θ is the sensitivity of accounting return on equity to demand growth induced by demographics, and v_{t+1+j} captures all other factors. For simplicity, we also assume that $E_{t+1} v_{t+1+j} = 0$. The sensitivity of *roe* to demand shifts, θ, depends on the industrial organization of the industry. For instance, in a perfectly competitive industry with no barriers to entry we expect that $\theta \approx 0$. In the presence of barriers to entry, we expect $\theta > 0$. Substituting expression (6) into equation (5) we obtain

$$\begin{aligned} E_t r_{t+1} - \bar{r} &= A + \rho^h w \theta E_t \Delta c_{t+1+h} + \rho^h (1-w)\theta \sum_{i=1}^{n} \frac{E_t[\Delta c_{t+1+h} - \Delta c_{t+1+h-i}]}{n} \\ &\quad + \frac{\rho^{h+1}}{1-\rho} \frac{(1-w)}{n} \theta E_t[\Delta c_{t+1+h} - \Delta c_{t+1+h-n}] \end{aligned} \tag{7}$$

where A is a constant equal to $\rho^h w \left(\phi - \overline{roe}\right)$. Equation (7) allows us to make the following predictions.

Prediction 1. *If investors are rational ($h \rightarrow \infty$), the expected return, $E_t r_{t+1}$, is independent of expected demand growth, $E_t \Delta c_{t+1+j}$, for any $j \geq 0$.*

Prediction 2. *If investors are inattentive (h finite), the expected return, $E_t r_{t+1}$, is positively related to $E_t \Delta c_{t+1+h}$. Moreover, $\partial E_t r_{t+1} / \partial E_t \Delta c_{t+1+h} = \rho^h \theta \left[1 + (1-w)\,\rho / \left((1-\rho)\,n\right)\right]$..*

Prediction 3. *If investors are inattentive with partial extrapolation (h finite and $w < 1$), the expected return $E_t r_{t+1}$ is negatively related to $E_t \Delta c_{t+1+h-i}$ for all $1 \leq i \leq n$. In addition, $\partial E_t r_{t+1} / \partial E_t \Delta c_{t+1+h} > |\partial E_t r_{t+1} / \partial E_t \Delta c_{t+1+h-i}|$ for all $1 \leq i \leq n$.*

Under the null hypothesis of rational investors, forecastable demographic shifts do not affect stock returns (Prediction 1). Under the alternative hypothesis of inattention, instead, forecastable demand growth in period $t+h+1$ predicts stock returns (Prediction 2). This prediction also links the magnitude of forecastability to the sensitivity of accounting return on equity to demand changes (θ); $\partial E_t r_{t+1} / \partial E_t \Delta c_{t+1+h}$ may be as small as $\rho^h \theta$ (for $w = 1$) or as large as $\rho^h \theta \left[1 + \rho / (1-\rho)\right]$ (for $w = 0$ and $n = 1$). Finally, Prediction 3 states that, if investors extrapolate to some extent using medium-term expectations (for $w < 1$), then demand growth less than $h+1$ periods ahead forecasts returns negatively. The negative relationship between medium-term demand growth and expected returns is smaller in absolute magnitude than the positive relationship between $E_t r_{t+1}$ and $E_t \Delta c_{t+1+h}$.

In this analysis we have made two restrictive assumptions. First, we only consider a representative agent model. An alternative model would consider a model of interactions between inattentive investors and rational agents in the presence of limited arbitrage (DeLong et al., 1990; Shleifer, 2000). We also assume that all investors have a horizon of exactly h periods. While this assumption is clearly not realistic, the model provides intuition for the case in which the horizon varies across individuals and industries. If the horizon varies between h and $h+\tilde{H}$, industry returns would be forecastable using a combination of demand growth rates due to demographics between years $t+h$ and $t+h+\tilde{H}$. The empirical specification in Section 3.6 acknowledges that horizons may vary and that the precision of the data does not permit separate estimates of each relationship between returns and expected consumption growth at a specific horizon. Therefore, in the baseline specification, we form two demand growth forecasts, one for medium-term growth between t and $t+5$, and one for long-term growth between $t+5$ and $t+10$.

3 Empirical analysis

The empirical specification consists of six steps. First, we generate demographic forecasts and estimate age patterns in consumption data for each good. Next, we combine the demographic forecasts with the consumption data to obtain demographic-based forecasts of demand growth by good. We estimate the response of industry profitability and stock returns to forecasted demand changes. Finally, we analyze the performance of a trading strategy designed to exploit demographic information.

3.1 Demographic forecasts

We combine data sources on cohort size, mortality, and fertility rates to form forecasts of subsequent cohort sizes.[3] All the demographic information is disaggregated by gender and one-year-age groups. The cohort size data is from the *Current Population Reports, Series P-25* (US Department of Commerce, Bureau of the Census). The cohort size estimates are for the total population of the United States, including armed forces overseas. We use mortality rates from period life tables for the years 1920-2000 from *Life Tables for the United States Social Security Area 1900-2080.* Finally, we take age-specific birth rates from Heuser (1976) and the *Vital Statistics of the United States: Natality.*

We use demographic information available in year t to forecast the age distribution by gender and one-year age groups for years $u>t$. We assume that fertility rates for the years $u>t$ equal the fertility rates for year t. We also assume that future mortality rates equal mortality rates in year t except for a backward-looking percentage adjustment. We obtain the adjustment by regressing mortality at a particular age for a specific decade on mortality at the same age in the previous decade for each of the last 5 decades before year t. The adjustment coefficient is allowed to differ by 10-year age groups. The estimated percentage improvement in mortality rates for the ages 0-9 is about 25 percent per decade. For the ages 50-59 it is about 9 percent per decade.

Using cohort size in year t and these forecasts of future mortality and fertility rates, we form preliminary forecasts of cohort size for each year $u>t$. To account for net migration, we estimate the average percentage difference between the actual cohort size and the preliminary forecasted cohort size formed the year before. We estimate the percentage difference separately

[3] Additional details regarding our forecasts of cohort size are available in Appendix B.

for each 10-year age group using data from the most recent ten year period prior to year t. We attribute this difference to historical net migration. For the 0-9 age group, the average imputed net migration is about .5 percent per year, while for the 50-59 age group, it is about .1 percent per year. We apply this imputed adjustment for migration to the initial population forecasts made at time t.

We define $\hat{\mathbf{A}}_{g,u|t} = \left[\hat{A}_{g,0,u|t},\ \hat{A}_{g,1,u|t},\ \hat{A}_{g,2,u|t}, ..\right]$ as the future forecasted age distribution. $\hat{A}_{g,j,u|t}$ is the number of people of gender g alive at u with age j forecasted using demographic information available at t. Figure 1a plots the actual series of population aged 30-34 over the years 1930-2002, as well as three forecasts as of 1935, 1955, and 1975. The forecasts track actual cohort sizes quite well, except for very long-term forecasts when they depend on predicting future births. Figure 1b for the age group 70-74 shows that the forecasts for older people are less precise. In Table 1 we regress the actual annual growth of each population group on the forecast of the annual growth rate from the previous year. Each observation is a (gender)*(year of forecast)*(one-year age group) cell. We run the regressions separately by 10-year groups of age. The forecasts for the ages between 1 and 69 are quite accurate, with R^2 measures of approximately .8. The forecasts for the older ages and for the unborn are significantly less precise due to substantial uncertainty about fertility rates and mortality rates for the old.[4]

3.2 Age patterns in consumption

Unlike demographic information, exhaustive information on consumption of different goods is available only after 1980. For the previous years, we use the only surveys available in an electronic format: the *Study of Consumer Purchases in the United States, 1935-1936,* the *Survey of Consumer Expenditures, 1960-1961,* and the *Survey of Consumer Expenditures, 1972-1973.*[5]. We combine these three early surveys with the 1983-1984 cohorts of the ongoing *Consumer Expenditure Survey*[6]. Taken together, these four cross-sections provide information on the age distribution of consumption throughout the past century. Table 2 reports summary statistics on the most important household demographics. Family size decreases over time, while the proportion of urban households increases. The sample sizes and sampling rules differ across surveys. While the post-War surveys cover a representative sample of the US population, the 1935-36 survey includes only married couples and is therefore biased toward younger families. The bottom part of Table 2 presents information on average yearly income and total consumption in 1982-84 dollars.

We cover all major expenditures on final goods included in the survey data. The selected

[4]The forecasting error for the very old should not have a large impact, given the small size of this population group. The size of the unborn population only matters for a limited number of goods.

[5]Costa (1999) discusses the main features of these surveys.

[6]The cohorts in the Consumer Expenditure Survey are followed for four quarters after the initial interview. The data for the fourth cohort of 1984, therefore, includes 1985 consumption data.

level of aggregation attempts to distinguish goods with potentially different age-consumption profiles. For example, within the category of alcoholic beverages, we separate beer and wine from hard liquor expenditures. Similarly, within insurance we distinguish between health, property, and life insurance expenditures. We attempt to define these categories in a consistent way across the survey years. Unfortunately, the categories are coded differently across the four surveys, and consequently, we do not have enough information to construct certain expenditure categories in some surveys. This problem is especially serious for the 1960-61 survey which classifies consumption data into particularly broad categories. Table 3 presents the summary statistics on good-by-good annual household expenditure for each survey year. The expenditures are in 1982-84 dollars.[7] Despite substantial differences across the four surveys in the sample, in the survey procedure, and in the definition of the goods, the mean household expenditure by good category is relatively stable over time.

More importantly, we can compare the age profile of consumption across survey years and across expenditure categories. To illustrate the age profile of selected goods, we use kernel regressions of household consumption on the age of the head of household[8]. Figure 2a, for example, plots normalized[9] expenditure on bicycles and drugs for the 1935-36, 1960-61, 1972-73, and 1983-84 surveys. Across the two surveys, the consumption of bicycles (Figure 2a) peaks between the ages of 35 and 40. At these ages, the household heads are most likely to have 5- to 10-year-old children. The demand for drugs (Figure 2a), instead, is increasing with age, particularly in the later surveys. Older individuals demand more pharmaceutical products. The differences in age profiles occur not just between goods targeted at young generations (e.g., bicycles) and goods targeted to the old (e.g., drugs), but also within broad categories, such as alcoholic beverages. The peak of the age profile of consumption for beer and wine (Figure 2b) occurs about 20 years earlier than the peak of the profile for hard liquor (Figure 2b). These age patterns are similar across the two surveys with data on alcoholic consumption. In another example, purchases of large appliances peak at 25-30 years of age, while purchases of small appliances are fairly constant across the years 25-50. Large appliances are largely associated with the purchase of the first house, while small appliances are purchased on a more regular basis.

Overall, this evidence supports three general statements. First, the amount of consumption for each good depends significantly on the age of the head of household. Patterns of consumption for most goods are not flat with respect to age. Second, these age patterns vary substantially across goods. Some goods are consumed mainly by younger heads (child care and toys), some in middle ages (life insurance and cigars), others at older ages (cruises and

[7]Details about the composition of the various goods are available form the authors upon request.

[8]We use an Epanechnikov kernel with a bandwidth of 5 years of age for all the goods and years.

[9]For each survey-good pair we divide age-specific consumption by the average expenditure across all ages for that particular good.

nursing homes). Third, the age profile of consumption for a given good is quite stable across time. For example, the expenditure on furniture peaks at the ages 25-35, whether we consider the 1935-36, the 1960-61, the 1972-73, or the 1983-84 cohorts. Taken together, the evidence suggests that changes in age structure of the population have the power to affect consumption demand in a substantial and consistent manner.

With this evidence in the background, we now present the methodology we use to estimate age consumption patterns. In order to match the consumption data with the demographic data, we transform the household-level consumption data into individual-level information. We use the variation in demographic composition of the families to extract individual-level information—consumption of the head, of the spouse, and of the children—from household-level consumption data. We use an OLS regression in each of the four cross-sections. Denote by $c_{i,k,t}$ the consumption by household i of good k in year t and by $H_{i,t}$ a set of dummies for the age groups of the head of household i in year t. In particular, $H_{i,t} = [H_{18,i,t}, H_{27,i,t}, H_{35,i,t}, H_{45,i,t}, H_{55,i,t}, H_{65,i,t}]$ where $H_{j,i,t}$ is a dummy equal to 1 if the head of household i at time t is older than j and younger than the next age group. For example, $H_{35,i,t}$ indicates that the head of household i is aged 35 to 44 in year t. The dummy $H_{65,i,t}$ indicates a head older than 65 years of age. Similarly, let $S_{i,t}$ be a set of dummies for the age groups of the spouse. Finally, we add discrete variables $O_{i,t} = [O_{0,i,t}, O_{6,i,t}, O_{12,i,t}, O_{18,i,t}, O_{65,i,t}]$ that indicate the total number of other individuals (children or old relatives) living with the family in year t. For example, $O_{0,i,t} = 2$ indicates that two children aged 0 to 5 live with the family in year t.

The regression specification is

$$c_{i,k,t} = B_{k,t}H_{i,t} + \Gamma_{k,t}S_{i,t} + \Delta_{k,t}O_{i,t} + \varepsilon_{i,k,t}.$$

This OLS regression is run separately for each good k and for each of the four cross-sections t. The purpose is to obtain an estimate of annual consumption of good k for individuals of different ages. For example, the coefficient $B_{35,\text{cars},1960}$ indicates the average total amount that a (single) head aged 35 to 44 spends on cars in 1960.[10]

3.3 Demand forecasts

In the third step of the research design, we combine these age profiles of consumption with the demographic forecasts in order to forecast future demand for different goods. Consider for example a forecast of toys consumption in 1975 as of 1965. Age-by-age, we multiply the forecasted cohort sizes for 1975 by the age-specific consumption of toys estimated on the most

[10]We do not include spouse dummies in the 1935-36 survey (only married couples were interviewed) and in the 1960-61 survey (age of the spouse not reported). In addition, since in the 1935-36 survey the size of sample is only a third to a half as large as the samples in the other surveys, we use broader age groups: 18, 35, 50, and 65.

recent consumption data, the 1960-61 survey. We then aggregate across all the age groups to obtain the forecasted overall demand for toys for 1975.

Formally, we denote by $\hat{A}^b_{g,u|t}$ the aggregation of $\hat{A}_{g,u|t}$ into the same age bins that we used for the consumption data. For example, $\hat{A}^b_{f,35,u|t}$ is the number of females aged 35 though 44 forecasted to be alive in year u as of year t. We combine the forecasted age distribution $\hat{A}^b_{g,u|t}$ with the age-specific consumption coefficients $B_{k,t}$, $\Gamma_{k,t}$, and $\Delta_{k,t}$ for good k. In order to do so, for each age group j we estimate the shares $h_{g,j,t}$,$s\ _{g,j,t}$, and $o_{g,j,t}$ of people in the population. For example, $h_{f,35,t}$ is the number of female heads 35-44 over total number of females aged 35-44 in the most recent consumption survey prior to year t. We then obtain a demographic-based forecast at time t of the demand for good k in year u which we label $\hat{C}_{k,u|t}$:

$$\hat{C}_{k,u|t} = \sum_{g\in\{f,m\}} \sum_{j\in\{0,6,12,18,...,65\}} \hat{A}^b_{g,j,s|t} \left(h_{g,j,t} B_{j,k,t} + s_{g,j,t} \Gamma_{j,k,t} + o_{g,j,t} \Delta_{j,k,t} \right).$$

The coefficients B, Γ, and Δ in this expression are estimated using the most recent consumption survey antecedent to year t with information on good k. This forecast implicitly assumes that the tastes of consumers for different products depend on age and not on cohort of birth. We assume that individuals age 45 in 1975 consume the same bundles of goods that individuals age 45 consumed in 1965. By construction, we hold prices of goods constant when computing future demand.

Table 4 presents the average 1-year forecasted consumption growth for each good during the years 1936-2002. Across all goods and years, the average annual growth due to demographics is 1.26 percent, with a standard deviation of .75 percent. The average within-good standard deviation in one-year growth is 0.61 percent, indicating an appreciable effect of changing demographics on consumption patterns.

We use the within-good standard deviation to identify the expenditure categories that are mostly affected by cohort size changes. The subsample 'Demographic Goods' includes the 20 expenditure categories with the highest standard deviations of forecasted consumption growth one year ahead (Column 4). This sample includes the expenditure on children as well as on funeral homes, cruises, beer (and wine), and cars.

Figure 3 shows the consumption growth due to demographics for three subcategories of the general book category—books for K-12 schools, books for higher education, and other books (mostly fiction). Formally, we plot $\ln \hat{C}_{k,u|1975} - \ln \hat{C}_{k,1975|1975}$ for $u = 1976, 1977, \ldots, 1995$, that is, the cumulated percentage changes in consumption forecasted as of 1975 twenty years into the future. For each of the three goods, we produce forecasts using the age-consumption profiles estimated from each of the four consumption data sets. The demand for K-12 books is predicted to experience a large decline as the baby-bust generation keeps entering schools. The demand for college books is predicted to increase and then decline, as the cohorts entering college are first large (baby boom) and then small (baby bust). Finally, the demand for other books, which is mostly driven by adults age 30 through age 50, is predicted to keep growing as baby-boomers gradually reach these ages. These patterns do not depend on the year (1935-36, 1960-61, 1972-73, or 1983-84) of expenditure survey used to estimate the age-consumption profile for each category. Figure 3 indicates that within the entire book category there is substantial variability in the pattern of demand growth across the subcategories. It also shows that the time-series variation in consumption growth is fairly independent of the consumption survey used to estimate the age profile.

We provide additional evidence that the time-series variation in consumption growth is mostly due to demographic changes. For each good and over the years 1937-2002, we generate one-year consumption forecasts using estimates of the age profile of consumption from the 1935-36 survey for the entire period. We repeat this process using age-consumption estimates from the 1960-61, the 1972-73, and the 1984-84 surveys. We then compute the correlation between these four measures of consumption growth. Using data for all goods and years, the correlations are in the .7 to .8 range (Table 5).The correlations for one specific good—for instance, cars—are higher on average. These correlations confirm that the consumption patterns are similar across surveys. Therefore, the variation in the forecasts does not appear to be driven by differences across surveys. The bottom part of Table 5 addresses a different concern, that is, the importance of error in forecasting demographics. We compute a measure of forecasted growth using actual demographic changes, rather than forecasted demographic changes. The correlation between the two measures is .69, indicating that errors in demographic forecasts are unlikely to have a large role.

3.4 ROE predictability

In the fourth step of the research design, we test whether forecastable demand changes affect profitability by industry, a necessary condition for our attention test. As a measure of profitability we use a transformation of the accounting return on equity (ROE). For each firm, the return on equity at time $t+1$ is defined as the ratio of earnings from the end of fiscal year t to the end of fiscal year $t+1$ (*Compustat* data item 172) to the book value of equity at the end of fiscal year t (*Compustat* data item 60). If data item 172 is missing, we calculate ROE using the clean surplus accounting identity from Vuolteenaho (2002). In order to obtain an industry-level measure of profitability, we compute the average return on equity for all companies in the industry weighted by the book value of equity in year t.

Since some industries require a higher level of disaggregation than provided by the standard 4-digit SIC codes, we create the industry classification ourselves whenever necessary. Using a company-by-company search within the relevant SIC codes we partition the companies into the relevant groups. For example, the SIC code 5092 on 'toys' include both companies producing toys for children and companies manufacturing golf equipment, two goods clearly associated with consumption by different age groups. Appendix Table 1 displays the SIC codes associated with a particular industry. The SIC codes in parentheses are those that are shared by different industries, and therefore require a company-by-company search. For larger industries such as automobiles, oil, and coal, our SIC grouping system yields portfolios that are similar to the industry portfolios generated by Fama and French.

We construct the annual industry return on equity $ROE_{k,t+1}$ using only companies that already belonged to the industry categorization at the end of year t. We exclude companies entering the SIC code classification between t and $t+1$ and we drop companies with negative book values. We control for inter-industry mergers and industry reclassifications by excluding companies that exit the industry between t and $t+1$. The final measure is the log return on equity, $roe_{k,t+1} = \log(1 + ROE_{k,t+1})$. In order to avoid the possibility of accounting outliers driving our results, we winsorize this accounting return measure at the 1 percent and 99 percent level. Table 6 presents summary statistics for the log annual return on equity (mean and standard deviation), the average number of firms included in the industry over time, and the number of years for which the ROE data is available. The average log return ranges from 5 percent (golf) to 26 percent (motorcycle). The within-industry standard deviation of the return is often as high as 15 percent. The longest series has 51 observations, but many series are shorter.

In Table 8 we test the predictability of the one-year industry log return on equity (Table 6) using the forecasted contemporaneous growth rate in consumption due to demographics (Table 4). Denote by $\hat{c}_{k,s|t}$ the natural log of the forecasted consumption of good k in year s forecasted as of year t.The following specification is motivated by equation (6):

$$\log(1 + ROE_{k,t+1}) = \lambda + \eta_k + \varphi_{t+1} + \theta \left[\hat{c}_{k,t+2|t-1} - \hat{c}_{k,t|t-1}\right]/2 + \varepsilon_{k,t+1} \tag{8}$$

The coefficient θ indicates the responsiveness of log return on equity in year $t+1$ to contemporaneous changes in demand due to forecasted demographic changes. Since the measure of cohort size for year $t+1$ refers to the July1 value, approximately in the middle of the fiscal year, we use the average demand growth between July 1 of year $t-1$ and July 1 of year $t+1$ as a measure of contemporaneous demand change. We scale by 2 to annualize this measure. The forecast of consumption growth between years t and $t+1$ uses only demographic and consumption information available up to year $t-1$. This lag ensures that all information should be of public knowledge by year t.[11] We run specifications of (8) both with and without industry and year dummies.

In this panel setting it is unlikely that the error terms across industries are independent since there are shocks that affect multiple industries at the same time. Therefore, we allow for arbitrary correlation across industries at any given time by calculating 'robust' standard

[11] At present, the Bureau of the Census releases the demographic information for July 1 of year t around December of the same year, that is, with less than a year lag.

errors clustered by year, under the assumption that the residuals are independent across years. Formally, if we define Ω as the error covariance matrix and X as the matrix of explanatory variables, then the covariance matrix for the coefficient estimator is $(X'X)^{-1}X'\Omega X(X'X)^{-1}$. If we assume that the errors for each cross-section are independent, then $X'\Omega X = \sum_{t=1}^{T} X_t'\Omega_t X_t$ where X_t is the matrix of explanatory variables and Ω_t is the error covariance matrix for each cross-section. Since Ω_t is unknown we estimate $X_t'\Omega_t X_t$ with $X_t'\widehat{\varepsilon}_t\widehat{\varepsilon}_t'X_t$, and similarly, $X'\Omega X$ using $\sum_{t=1}^{T} = X_t'\widehat{\varepsilon}_t\widehat{\varepsilon}_t'X_t$ where the vector of estimated residuals for each cross-section is denoted $\widehat{\varepsilon}_t$.

In the baseline specification of Table 8 we use the subsample of Demographic Industries. In Column 1 we present the results for specification (8) without industry or year dummies. The estimated coefficient, $\hat{\theta}$, is significantly positive and economically large. A one percent increase in yearly consumption growth due to demographics increases the log return on equity from an average of 12.2 percent to an average of 13.6 percent, an 11 percent increase. The R^2 of the regression is low due to the modest size of demographic changes relative to other determinants of profitability. In Column 2 we introduce industry dummies. In this case, the identification depends only on time-series changes in the growth rates and not on between-industries differences. The estimate for θ does not change. In Column 3 we also introduce year dummies. In this specification, the identification depends on differential time-series in demand changes across industries. The estimated coefficient, $\hat{\theta}$, is still significantly positive and large.

In Columns 4-6 we reestimate the model for the period after 1974. The accounting data for the earlier period is noisier since the accuracy of the industry classification increases with proximity to the present[12]. In addition, the industry-level measure of return on equity is likely to be more precise since the number of companies covered in the accounting data increases substantially over time. In this later time period $\hat{\theta}$ is comparable to the estimates on the overall sample. Finally, in Columns 7 through 12 we reestimate the same models for the whole sample of 47 industries. The point estimates for θ are comparable to the corresponding ones for the subset of Demographic Goods. The standard errors in the whole sample are somewhat larger than in the demographic sample, despite a threefold increase in sample size, suggesting a higher signal-to-noise ratio for the non-demographic industries.

Overall, forecasted demand changes due to demographics have a statistically and economically significant effect on industry-level profitability. It appears that entry and exit by firms into industries does not fully undo the impact of forecastable demand changes on profitability.

[12]The company-level information used to generate, for example, the book subcategories is accurate for the present (2003), but less likely so the earlier in time.

3.5 Industry concentration

The discussion of the industrial organization in Section 2 suggests that the qualitative impact of a demand change on profitability should depend on the market structure. At one extreme, in a perfectly competitive industry with no barriers to entry, the consumers capture all the surplus arising from a positive demand shift. In this scenario, demographic changes do not affect abnormal profits. At the other extreme, a monopolist in an industry with high barriers to entry generates additional profits from a positive demand change. We address this issue by estimating the effect of demand changes on profitability for sectors with different measures of barriers to entry.

As a proxy for barriers to entry and/or market power, we use the concentration ratio C-4, that is, the ratio of industry revenue produced by the 4 largest companies. Starting in 1947 this measure is available from the *Census of Manufacturers* for industrial sectors with 4-digit SIC codes between 2000 and 3999. We create an industry concentration index as the average C-4 ratio for the SIC codes included in the definition in the range 2000-3999. The average is weighted by the aggregate revenue for an SIC code. To avoid industries switching concentration ratio groups over time, we use the earliest measure to classify each industry. Therefore, for almost all industries the concentration ratio measure is taken from the 1947 *Census of Manufacturers*. Unfortunately, concentration ratios are not available for many non-manufacturing industries, such as insurance and utilities, that do not have an SIC code within the appropriate range. Table 7 reports the C-4 ratios for the industries in the sample. There is substantial variation across industries in the concentration measure and .31 is the median concentration ratio.

Table 9 reports the profitability regressions (8) for the subsample of industries with above-median concentration and the subsample of industries with below-median concentration. For the sample of more concentrated industries (Columns 1 through 3) the effect of an increase in demand due to demographics is insignificant in the baseline specification (Column 1), but large and significant for the specification with year and industry dummies (Column 3). The estimates of θ for the more concentrated industries are larger than in the comparable specifications for the whole sample (Table 8, Columns 7-9). For the sample of unconcentrated industries (Columns 4 through 6), instead, the coefficient $\hat{\theta}$ is small and not significantly different from zero. As a robustness check, we estimate the alternative specification

$$\begin{aligned}\log\left(1+ROE_{k,t+1}\right) &= \lambda+\eta_k+\varphi_{t+1}+\theta\left[\hat{c}_{k,t+2|t-1}-\hat{c}_{k,t|t-1}\right]/2 \\ &\quad +\theta^C C_{4k}\left[\hat{c}_{k,t+2|t-1}-\hat{c}_{k,t|t-1}\right]/2+\varsigma C_{4k,t}+\varepsilon_{k,t+1}\end{aligned}$$

where C_{4k} is the (continuous) concentration measure for industry k. The coefficient θ^C captures the extent to which the sensitivity of profits to demand shifts of demand shifts is higher for more concentrated industries. The estimated coefficient, $\hat{\theta}^C$, is indeed positive and significant for the specifications with industry dummies (Column 8) and with the full set of dummies (Column 9). The evidence supports the prediction that the demand changes due to demographics alter profits more substantially in the presence of barriers to entry.

3.6 Return predictability

In the fifth step, we examine the relationship between forecasted demand growth and stock returns. We aggregate firm-level stock return data from *CRSP* to form value-weighted industry-level measures of returns. The aggregation procedure is identical to the methodology used for the profitability measure. The procedure employs SIC codes augmented by specific company-by-company searches.[13] Table 7 displays the summary statistics on one-year value-weighted stock returns (mean and standard deviations), average number of firms, and years covered. The sample of returns is larger than the sample of accounting profitability because returns data is available for a longer time period and for more companies. The average annual log stock return varies from 2.7 percent (nursing homes) to 20 percent (motorcycles). The standard deviation of the yearly stock returns—32 percent on average—is negatively correlated with the number of firms in the industry.

We choose specifications motivated by expression (7) in Section 2 and investigate when stock prices incorporate the forecastable consumption changes generated by demographic variables. In the baseline specification we regress annual returns on the forecasted growth rate of demand due to demographics from t to $t+5$ (the medium-term) and $t+5$ to $t+10$ (the long-term). We beta-adjust industry returns to remove the market-wide macroeconomic shocks. Define $r_{k,u,t}$ to be the natural log of the stock return for good k between year t and year u. The natural log of the market return over the same horizon is given by $r_{m,u,t}$. Further, let $\hat{\beta}_{k,t}$ be the coefficient of a regression of monthly industry returns on market returns over the 48 months previous to year t. The specification of the regression is

$$\begin{aligned} r_{k,t+1,t} - \hat{\beta}_{k,t} r_{m,t+1,t} &= \gamma + \eta_k + \varphi_{t+1} + \delta_0 \left[\hat{c}_{k,t+5|t-1} - \hat{c}_{k,t|t-1}\right]/5 + \\ &\quad + \delta_1 \left[\hat{c}_{k,t+10|t-1} - \hat{c}_{k,t+5|t-1}\right]/5 + \varepsilon_{k,t+1} \end{aligned} \tag{9}$$

Scaling the two consumption growth variables by 5 implies that the coefficients δ_0 and δ_1 represent the average increase in abnormal yearly returns for each one percentage point of additional annualized growth in demographics. Once again, the forecasts of consumption as of time t only use information released available in period $t-1$.

The model in Section 2 suggests that, if the forecast horizon h is shorter than 5 years, the coefficient δ_0 should be positive and δ_1 should be zero. If the forecast horizon is between 5 and 10 years, the coefficient δ_0 should be zero or negative and the coefficient δ_1 should be positive.

[13] The results are qualitatively similar if the industry classification scheme does not include our separate company specific searches.

Finally, if the investors have a horizon greater than 10 years (including rational investors with $h \to\infty$), both coefficients should be zero. A significantly positive coefficient indicates that stock prices adjust as the demographic growth information enters the forecast horizon.

Table 10 presents the estimates of (9). In the benchmark specification (without year or industry dummies) for the sample of 'Demographic Industries' (Column 1), the coefficient on medium-term demographics $\hat{\delta}_0$ equals -2.1 and is not significantly different from zero. The coefficient on long-term demographics $\hat{\delta}_1$ equals 5.7 and is significantly larger than zero. A one percent annualized increase in demand from year 5 to year 10 increases the average abnormal yearly stock return by 5.7 percentage points. The coefficients have the same magnitude when industry fixed effects (Column 2) and year fixed effects (Column 3) are introduced, although $\hat{\delta}_1$ is only marginally significant in these specifications. For a more recent sample (Columns 4 through 6), we observe the same pattern of results with larger magnitudes: the coefficient on medium-term demographics is negative and insignificant, while the coefficient on long-term demographics is large and significant. In the overall sample (Columns 7 through 12), the coefficients have comparable and somewhat smaller magnitudes than in the Demographic Sample, with the same pattern of significance.

Barriers to entry. As we discussed above, testing attention using stock market reaction to demand changes is meaningful only for industries with substantial barriers to entry. In the first six columns of Table 11 we replicate the test of specification (9) separately for industries with concentration ratio C-4 above and below the median. For the sample of more concentrated industries (Columns 1 through 3) the effect of demand growth between $t+5$ and $t+10$ is larger than in the overall sample of all industries, although the estimates are not significantly different from zero. In the sample of less concentrated industries (Columns 4 through 6) there is no significant effect of demand changes on stock returns. In columns 7 through 9 we use the continuous measure of concentration C-4 and interact it with demand growth at the different horizons in the following alternative specification:

$$
\begin{aligned}
r_{k,t+1,t} - \hat{\beta}_{k,t} r_{m,t+1,t} \quad = \quad & \lambda + \eta_k + \varphi_{t+1} + \delta_0 \frac{\hat{c}_{k,t+5|t-1} - \hat{c}_{k,t|t-1}}{5} \\
& + \delta_1 \frac{\hat{c}_{k,t+10|t-1} - \hat{c}_{k,t+5|t-1}}{5} + \delta_0^C C_{4k} \frac{\hat{c}_{k,t+5|t-1} - \hat{c}_{k,t|t-1}}{5} \\
& + \delta_1^C C_{4k} \frac{\hat{c}_{k,t+10|t-1} - \hat{c}_{k,t+5|t-1}}{5} + \varsigma C_{4k,t} + \varepsilon_{k,t+1}
\end{aligned}
$$

The baseline estimate (Column 7) of δ_1^C is large and significantly different from zero. For an industry with a low concentration ratio of .2, the predicted responsiveness of stock returns to long-term demand growth is $\hat{\delta}_1 + .2\hat{\delta}_1^C = 2.5$. For an industry with a high concentration ratio of .6, the predicted responsiveness is $\hat{\delta}_1 + .6\hat{\delta}_1^C = 12.9$. Similarly large magnitudes are suggested by the specifications in Columns 8 and 9, although the results are not significant. If we repeat the analysis for the period from 1975 to 2002 (results not shown), the estimated coefficient, $\hat{\delta}_1^C$, is even larger. The evidence suggests that return predictability is stronger in industries with higher concentration, although the estimates are not precise.

Investor Horizon. We consider a specification of return predictability that is more closely linked with the model of short-sighted investors in Section 2. We estimate the specification

$$\begin{aligned} r_{k,t+1,t} - \hat{\beta}_{k,t} r_{m,t+1,t} &= \lambda + \eta_k + \varphi_{t+1} + \delta_H \left(\hat{c}_{k,t+h+1|t-1} - \hat{c}_{k,t+h|t-1} \right) \\ &\quad + \delta_{H-N} \frac{\hat{c}_{k,t+h|t-1} - \hat{c}_{k,t+h-n+1|t-1}}{n} \\ &\quad + \delta_N \left(\hat{c}_{k,t+h-n+1|t-1} - \hat{c}_{k,t+h-n|t-1} \right) + \varepsilon_{k,t+1} \end{aligned}$$

where h is the investor horizon and n is the length of the extrapolation period. We fix n at 4 years and estimate the model for different investor horizons h. In particular, the coefficient δ_H measures the extent to which consumption growth h years ahead forecasts stock returns. Table 12 reports the results of the estimation on the sample of Demographic Industries[14] and Figure 4 reports the estimate of $\hat{\delta}_H$ for different horizons, including 95% confidence intervals. The estimate, $\hat{\delta}_H$, is positive but small and insignificant for horizons $h = 3$ and $h = 4$, is large and significant for $h = 5$ and $h = 6$, and is negative and not significant for longer horizons. These findings suggest that stock returns are predicted by forecasted demand growth occurring between 5 and 6 years in the future.

3.7 Portfolio returns

The results in the previous step provide evidence of return predictability using long-term demand growth due to demographics. In this final step, we analyze whether rational market participants could exploit the underreaction to long-term demographic information with a trading strategy.

We follow a strategy from 1938 to 2002 for sector indices belonging to the sample of Demographic Industries. Each year the industries are sorted by the forecasted demand growth between years $t + 5$ and $t + 10$. The zero-investment portfolio is long in indices for industries in the top third of the consumption growth measure and short in indices for industries in the bottom third. We compute monthly portfolio returns by equally weighting the relevant industry returns.

We control for market performance by regressing the series on the CRSP value-weighted stock index net of the 1-month Treasury rate. The standard errors are corrected for heteroskedasticity and autocorrelation using the Newey-West estimator with 6 lags[15]. The results in Column 1 of Table 13 indicate that the portfolio earns abnormal return of 5.6 percent. The outperformance remains the same if we also include the size and the book-to-market factors

[14] The pattern of results is similar if all industries are included in the analysis.

[15] The results do not change qualitatively if the lag length for the Newey-West standard errors is 12 or 18.

(Column 2) as well as the momentum factor (Column 3). These magnitudes are consistent with the estimates from the predictability regressions in Table 10. The annualized abnormal return for the portfolio (5.6 %) is approximately equal to the product of $\hat{\delta}_1$ (5.7) from Table 10 (column 1) and the average difference between forecasted demand growth from $t+5$ to $t+10$ for the long and short constituent portfolios (1.2%).

In Columns 4 through 6 we report the abnormal performance of the investment strategy since 1975. During this time period, the portfolio is formed using of a substantially larger set of industries, each industry contains more firms, and the industry classification is likely to be more accurate. For this sample, the portfolio has an average abnormal annualized return of more than 8% per year. Finally, in Columns 7 through 9 we only use the SIC codes for industry classification and we ignore the manual reclassification of companies within each SIC code. Some sectors, such as books for children, books for K-12, and books for college disappear, because they depend on our company-by-company classification. Other industries, while still defined, have a less precise classification; for example, the toy industry now contains golf equipment retailers. For this alternative classification scheme, the portfolio strategy outperforms the market by an annualized return of 4.8%. The outperformance is only slightly weakened by the rougher classification system.

In Table 14 we present the results for a similar zero-investment portfolio for all 47 industries. We restrict the sample to years after 1947, the first year of availability of concentration data. This portfolio earns average annual abnormal returns of 2 to 3 percentage points (Columns 1 through 3). Unlike the other estimates, the outperformance is insignificant after controlling for the 3-factor or 4-factor risk-adjustment procedure. The weaker performance of the portfolio strategy in this sample is roughly consistent with the estimates in Table 11 because the difference between average forecasted consumption growth for the top third and bottom third of industries is only 0.6%.

In Columns 4 through 9 of Table 14 we examine two additional samples. The first sample only includes industries with above-median concentration ratios, while the second only includes industries with below-median concentration ratios. The average abnormal return for the high-concentration sample is approximately 4% and statistically significant, while the portfolio return for the low-concentration sample is approximately 1% and insignificant. This difference is consistent with the earlier results suggesting that abnormal returns are more sensitive to forecasted demand growth in the high-concentration sample.

The average abnormal returns from trading on long-term demographic information appear to be large and statistically significant. The estimates from the predictability regressions and the abnormal returns for the trading strategy are broadly consistent.

4 Attention and other interpretations

Three stylized facts emerge from the empirical analysis of industry stock returns. First, forecastable future demand changes due to demographic variables predict abnormal annual stock returns. Second, while demographic changes in the more distant future ($t+5$ to $t+10$) forecast returns, demographics changes in the near future (t to $t+5$) do not have significant forecasting power. Third, return predictability is stronger in industries with higher concentration ratios (a proxy for high barriers to entry) and with more volatile demand shifts induced by demographics.

The first stylized fact is inconsistent with the predictions of the model for fully rational (attentive) investors. According to prediction 1 in Section 2, if investors are rational, then stock returns should not be forecastable using expected demand changes. However, other models of rational investment behavior might accommodate the empirical evidence. Demographic information could proxy for a state variable that systematically alters the future investment opportunity set. Demographic changes might be an unknown risk factor that is not considered in the standard model. In this setting, return predictability would be rational according to Merton (1973).

Prediction 2 in Section 2 offers a more straightforward explanation of return predictability. If investors neglect information beyond a particular horizon h, then returns at $t+1$ should be predictable using long-term demographic information emerging between $t+h$ and $t+1+h$. The results in Tables 10 and 11 suggest that the horizon h could be between 5 and 10 years. Table 12 and Figure 4 provide a more precise estimate of h. As Figure 4 indicates, only the consumption growth from $t+5$ to $t+6$ and from $t+6$ to $t+7$ predicts stock returns. Since demographic information is measured in July rather than at the end of the year, these findings suggest that investors have a horizon between 4.5 and 6.5 years.

The findings in Table 12 also provide some support for Prediction 3. For a horizon h of 5 or 6 years (Columns 3 and 4), the estimates of δ_{H-N} and δ_N are negative for three out of the four cases, although the estimates are not significantly different from zero. This (weak) finding is suggestive evidence that individuals extrapolate long-term expectations about profitability from short horizon forecasts. In addition, the relative magnitudes of the coefficients, with $|\delta_N| < |\delta_H|$ and $|\delta_{H-N}| < |\delta_H|$, are consistent with Prediction 3 as well.

The model in Section 2 also makes a prediction regarding the magnitude of the coefficient on long-term forecasted demand growth in the return predictability regression. We examine the estimated magnitudes in Table 10, but the following discussion also applies to the estimates in Table 12. The estimate $\hat{\delta}_1 \approx 5$, is approximately 4 times larger than $\hat{\theta} \approx 1.5$, the estimate for the responsiveness of accounting return on equity to forecasted demand growth These magnitudes are no consistent with a model of unconditional inattention ($w = 1$) which predicts that δ_1 should be smaller than θ: $\delta_1 = \rho^h \theta < \theta$. However, a model of inattention with partial extrapolation ($w<\ 1$) can match the estimated magnitude of δ_1. We set the annual discount factor ρ equal to 0.95, the extrapolation weight w equal to 0.5, and the number of periods of extrapolation n equal to 4. For these parameters the model of inattention with partial extrapolation implies $\delta_1 = \theta\rho^h \left[1 + (1-w)\,\rho/\left((1-\rho)\,n\right)\right] \approx 3.5\theta$ when the horizon h is equal to 5 years. The estimated coefficient of stock returns on long-term demand growth $\hat{\delta}_1$, therefore, is consistent with the estimate of the responsiveness of profitability to demand growth, $\hat{\theta}$.

The direction and magnitudes of the estimated coefficients are therefore consistent with investor underreaction to information beyond a horizon of approximately 5 years. The calibration exercise provides indirect evidence of partial extrapolation.

The third stylized fact is readily explained by the industrial organization of the different sectors. For industries with low barriers to entry, demand changes should not have a significant impact on firm profitability. Demand shifts might lead to entry or exit, but profitability and stock returns are unaffected. Similarly, in industries with relatively uniform age profiles of consumption, changes in cohort sizes have a limited impact on demand. As a consequence, profitability and expected stock returns are unaltered. These results are consistent with the empirical relationship between demand shifts and return on equity. Profitability responds strongly to contemporaneous demographic shifts only in industries where measured concentration is high and where the consumption profile depends markedly on age.

Our interpretation of the overall evidence suggests that investors do not pay attention to information beyond a 5 year horizon. We now provide suggestive evidence from analyst forecasts that the estimated investor horizon of 5 years is consistent with the observed horizon of analyst forecasts. In Table 15 we use the I/B/E/S data to document the pattern of earnings forecasts by analysts at different horizons. In Column 1 we consider forecasts made in 1990 and report the number of companies with at least one earning forecast h years into the future. Almost all companies in the sample appear to have earnings forecasts for the next two years. The number of forecasts further in the future, however, decays very quickly with distance. Less than half of the companies have forecasts 3 years ahead and less than 10 percent of the companies have forecasts 5 five years in the future. Forecasts beyond 5 years are not even reported in the data set in 1990. Not surprisingly, the share of firms with forecasts 3, 4, and 5 years ahead is higher among the firms with at least 5 analysts (Columns 2 and 3). However, even in this group the percentage of firms with 5-year-ahead forecasts is only 15 percent. Columns 4 through 6 present similar evidence for analyst forecasts recorded in 2000.

Analysts do not appear to produce forecasts of annual earnings beyond a 5 year horizon. However, it is possible that some analysts produce year-by-year forecasts for longer horizons but this information is not reported to I/B/E/S. In either case, investors do not possess readily available information regarding profitability in the distant future. Given this evidence, it would not be surprising if investors, as a rule of thumb, ignored outcomes more than 5 years in the future.

Ignoring information about the distant future and using a rule of thumb is a good strategy in many circumstances. Long-term patterns, such as consumer taste changes, are often already observed in the short-term data, making long-term information redundant. For other long-term variables, such as GDP growth, the forecasts are surrounded by so much uncertainty that a rule of thumb may be approximately correct. Demographic information is unusual because future long-term demographic variables can be precisely estimated and may differ significantly from their short-term pattern. Therefore, neglecting information about the long-term is costly.

Neglect of slowly-moving variables. A second attention-based interpretation of the results is based on the neglect of slowly-moving variables. In the frenzy of earnings and merger announcements, liquidity-driven orders, and media headlines about world news, investors are likely to disregard variables that display little daily variation. If investors focus on daily changes, they may never notice a change in variables like demographics. Studies on just-noticeable differences (Weber, 1834) suggest a minimum size of a stimulus necessary for detection, let alone to attract attention. Recent studies of visual perception have analyzed conditions under which subjects detect a change between two scenes. Simons, Franconeri, and Reimer (2000) expose one group of subjects to a 12-second movie of a natural scene where one item in the scene gradually changes color. In the control group, subjects watch for 11 seconds the first image of the same movie, and then see the final image. The subjects in the gradual-change condition detect 31% of the color changes, while subjects in the abrupt-change condition detect 92% of the changes.

The experimental evidence from these studies suggests that discrete changes are likely to be detected. Small, continuous changes, instead, are less likely to capture attention. A strategy that ignores variables exhibiting little daily variation is a rational choice for most days. However, neglecting slow-moving variables forever cannot be the optimal strategy. Attentive investors should periodically evaluate all information.

This interpretation correctly predicts that investors neglect demographic information, which has the feature of being always slowly-moving. However, the horizon of forecastability is inconsistent with this hypothesis. This story suggests that demographic information should be incorporated when reflected in earnings announcements, which are discrete events. We would therefore expect that short-horizon, rather than long-horizon, demographic information predicts stock returns.

Limits to arbitrage. We have argued that a model of inattention as in Section 2 can explain the predictability of industry stock returns. However, the predictions of the model hold only if every investor is inattentive. The presence of an even small portion of rational arbitrageurs is sufficient to eliminate any evidence of predictability, in the absence of limits to arbitrage (De Long et al., 1990; Shleifer, 2000). A reasonable explanation of the results in terms of inattention therefore requires limits to arbitrage.

Limits to arbitrage may arise because managers of sector funds are constrained to hold companies within an industry and therefore cannot build a diversified demographic portfolio. Limits to arbitrage may also arise because money-managers are evaluated based on short-term performance. In this setting, money managers may not be able to expose themselves to risk for a long enough period to reap the returns from trading on demographic information. Although this explanation is plausible, the substantial abnormal returns at an annual frequency must be relevant even for professionals with relatively short investment horizons.

Limited computing power. The results may be due to superior forecasting technology that was not available in the past. Although the empirical strategy in this paper is relatively simple, it still requires least-squares regressions on consumption surveys and matrix computations to obtain forecasts of future demand. Both operations would have been very tedious in 1960, and close to impossible in 1930. Against this hypothesis, however, weighs the evidence that the forecastability of returns for the years after 1975 is stronger than for the longer time period (Tables 10 and 13). Yet, in the years since 1975 computing power has become increasingly available.

5 Conclusions

We present evidence relating demographic variables to consumption patterns, industry profitability, and stock returns. Different goods have substantially dissimilar age patterns of consumption and these patterns are remarkably stable through time. While age patterns of consumption are obvious for goods such as children books and nursing homes, other patterns are not as straightforward. For example, the age-consumption profile of liquor peaks 20 years after the profile for beer and wine.

We combine our estimates of consumption by age with forecasts of cohort size by age. Our methodology produces forecasts of demand growth due to demographic changes for 47 different expenditure categories. We match the expenditure categories to industry-level accounting measures and stock market returns. The forecasted growth rates of demand due to demographic changes predict the accounting return on equity. This predictability result is more substantial for industries with larger variations of forecasted demand growth and higher concentration ratios.

We regress industry returns on growth rates of consumption due to demographics. We find that long-term growth rates of demand forecast annual abnormal returns, while medium-term growth rates do not have significant forecasting power. The predictability is more pronounced for the same groups of industries that exhibit a stronger relationship between profitability and forecastable demand growth.

The evidence supports the hypothesis that investors are short-sighted. In particular, investors appear to neglect information about expected profitability beyond a 5-year horizon. This finding is consistent with the near absence of earnings forecasts by analysts at this long horizon.

We have identified a novel form of predictability in financial markets. The finding of underreaction to information about the distant future has implications for other economic decisions beyond portfolio allocation. Voters and consumers may neglect relevant information about long-term outcomes for their decisions. Workers might disregard forecastable future demand changes in their choice of careers (Zarkin, 1985). Managers may neglect long-term demand shifts in their strategic decisions.

Further examination of consumer, investor, and firm response to anticipated events will cast more light on the phenomena presented in this paper.

A Appendix A. Model

We summarize the derivation of equation (3) in Section 2 (Vuolteenaho, 2002). We assume that the market price, M, book equity, B, and dividend payments, D, are positive during any time period. Define m, b, and d as the log transformation of each variable, respectively. We assume the 'clean-surplus identity' between earnings, X, book equity, and dividend payments, that is, $B_{t+1} = B_t + X_{t+1} - D_{t+1}$. Earnings X that are not paid to shareholders as dividends increase book equity.

We define the log stock return, r_{t+1}, and log accounting return on equity, roe_{t+1}, as

$$
\begin{aligned}
r_{t+1} &\equiv \log\left(\frac{M_{t+1} + D_{t+1}}{M_t}\right) && (10)\\
roe_{t+1} &\equiv \log\left(\frac{B_t + X_{t+1}}{B_t}\right) = \log\left(\frac{B_{t+1} + D_{t+1}}{B_t}\right) && (11)
\end{aligned}
$$

The second expression for roe_{t+1} follows from the clean-surplus identity. Finally, we assume that $d_{t+1} - m_{t+1}$ and $d_{t+1} - b_{t+1}$ follow stationary processes. By construction, the unconditional mean of $d_{t+1} - m_{t+1}$, denoted $\overline{d-m}$, is equal to the average log dividend-price ratio. We log-linearize (10) and (11) around the expansion point $\overline{d-m}$:

$$
\begin{aligned}
r_{t+1} &\approx k + \rho m_{t+1} + (1-\rho) d_{t+1} - m_t\\
roe_{t+1} &\approx k + \rho b_{t+1} + (1-\rho) d_{t+1} - b_t
\end{aligned}
$$

with $\rho = \left[1 + \exp(\overline{d-m})\right]^{-1}$ and $k = -\log(\rho) - (1-\rho)(\overline{d-m})$. Ignoring the approximation errors, we subtract the log-linearization for roe_{t+1} from the log-linearization for r_{t+1} to get a difference equation for the log market-to-book ratio:

$$m_t - b_t = \rho(m_{t+1} - b_{t+1}) - r_{t+1} + roe_{t+1} \tag{12}$$

Solving equation (12) forward and imposing the condition $\lim_{j\to\infty} \rho^j (m_{t+j} - b_{t+j}) = 0$, we get

$$\sum_{j=0}^{\infty} \rho^j [roe_{t+1+j} - r_{t+1+j}] = m_t - b_t = \widehat{E}_t \sum_{j=0}^{\infty} \rho^j [roe_{t+1+j} - r_{t+1+j}] \tag{13}$$

where the second equality follows from taking expectations with respect to the operator $\widehat{E}$ and noting $\widehat{E}_t (m_t - b_t) = m_t - b_t$. Substituting the right hand-side of (13) into (12) leads to

$$r_{t+1} - \widehat{E}_t r_{t+1} = \Delta\widehat{E}_{t+1} \sum_{j=0}^{\infty} \rho^j roe_{t+1+j} - \Delta\widehat{E}_{t+1} \sum_{j=1}^{\infty} \rho^j r_{t+1+j},$$

the desired expression for (3).

B Appendix B. Data Appendix

Cohort size adjustment. The initial population profile is from various years of the *Current Population Reports, Series 25*. For the years before 1980, these series lump together all age groups above the age of 84. In order to match the cohort sizes with the mortality rates, we disaggregate this group into 1-year age groups using the relative cohort sizes in 1980. Let $A_{g,j,t}$ be the population size at age j for gender g in year t. For any $t<$ 1980 we impute population sizes for ages 85 to 99 using $A_{g,j,t} = \left(\sum_{j=85}^{99} A_{g,j,t} / \sum_{j=85}^{99} A_{g,j,1980}\right) * A_{g,j,1980}$. This imputation imposes a constant population distribution in each year for ages beyond 84. Therefore, forecasts of population growth for ages beyond 84 will not match the imputed age distribution in the following year. Given the small size of population above 84 years of age (2,197,000 individuals in 1979), this problem is unlikely to matter. However, it does explain our inability to forecast the one-year growth rate of cohort sizes for the elderly population in Table 1.[16]

Mortality rate adjustment. At the end of each decade the Social Security Administration computes mortality rates from period life tables in *Life Tables for the United States Social Security Area 1900-2080.* To adjust for improvements in mortality rates over time, we compute mortality rate adjustment for each ten-year age range using data from the previous 5 decades. Let $q_{g,j,d}$ be the mortality rate for gender g, age j, and decade d from the life tables and let $d(t)$ be the end of the most recent decade before t. If $t = 1951$, then the mortality adjustment for ages 10 to 19 is based on the coefficient ($\kappa_{[10,19],1951}$) from the following regression:

$$q_{g,j,d} = k_{[10,19],1951} * q_{g,j,d-1} + \epsilon_{g,j,d}$$

for all observations with $d \in \{1910, 1920, 1930, 1940, 1950\}$ and $10 \leq j \leq 19$. Therefore, $\hat{q}_{g,j,u|t}$, the forecast from year t of mortality rates at age j in year $u>t$, is given by $\hat{q}_{g,j,u|t} = q_{g,j,d(t)} * \left(\kappa_{z(j),t}\right)^{\frac{u-t}{10}}$, where $z(j)$ is the 10-year age range corresponding to age j.

Fertility. We take the fertility rate by one-year age of the mother from Heuser (1976) and update it for the more recent years using the *Vital Statistics of the United States: Natality*[17].We assume that the forecasted fertility rate $\hat{b}_{j,u|t}$ for women of age j in year u forecasted as of year t equals the actual fertility rate $b_{j,t|t}$ for women of age j in year t: $\hat{b}_{j,u|t} = b_{j,t|t}$.

Cohort size forecast. By combining the initial population profile with the forecasts of mortality and fertility, we produce a preliminary forecast of the population profile with an iterative procedure. Starting with the preliminary population profile $\hat{A}^p_{g,u-1|t} = [\hat{A}^p_{g,0,u-1|t}, \hat{A}^p_{g,1,u-1|t}, \hat{A}^p_{g,2,u-1|t}, ...]$ for year $u-1$, we generate a forecasted population profile for the next year u using two relationships. First, for any $j \geq 1$ we calculate $\hat{A}^p_{g,j,u|t}$ as $\hat{A}^p_{g,j,u|t} = \hat{A}^p_{g,j-1,u-1|t} * (1 - \hat{q}_{g,j-1,u-1|t})$. Second, the forecasted number of newborns in year u is given by $\hat{A}^p_{g,0,u|t} = sr_g * \sum_{j=15}^{49} A^p_{f,j,u-1|t} * \hat{b}_{j,u-1|t}$, where $sr_m = 0.501$ is the average probability that a

[16]In the years before 1940, the series lump together age groups above 74. We apply the same imputation procedure using the age distribution of 1940 up to age 84 and the age distribution of 1980 beyond age 84.

[17]We thank John Wilmoth for providing this series.

newborn will be male ($sr_f = 1 - sr_m$ by construction).

Immigration adjustment. We compute a backward-looking adjustment for net migration by regressing the percentage difference between the actual cohort size and the preliminary forecasted cohort size formed the year before on a constant. We produce these adjustment coefficients separately for each 10-year age group using data from the most recent ten year period prior to year t. For instance, if $t = 1951$, then the immigration adjustment for ages 10 to 19 is based on the coefficient ($\psi_{[10,19],1951}$) from the following regression:

$$\left(A_{g,j,t-i+1} - \hat{A}^p_{g,j,t-i+1|t-i}\right) / \hat{A}^p_{g,j,t-i+1|t-i} = \psi_{[10,19],1951} + \nu_{g,j,t-i}$$

for all observations with $1 \leq i \leq 10$ and $10 \leq j \leq 19$. Therefore, $\hat{A}_{g,j,u|t}$, the forecast of cohort size for gender g and age j in year u as of year t, is given by $\hat{A}_{g,j,u|t} = \hat{A}^p_{g,j,u|t} * \prod_{i=1}^{u-t} \left(1 + \psi_{z(j-i),t}\right)$, where the function z converts $j - i$ to an age range. The forecasted cohort size profile $\hat{A}_{g,u|t} = \left[\hat{A}_{g,0,u|t},\ \hat{A}_{g,1,u|t},\ \hat{A}_{g,2,u|t}, ...\right]$ is the basis for the empirical analysis in the paper.

References

[1] Ang, Andrew and Angela Maddaloni. "Do Demographic Changes Affect Risk Premiums? Evidence from International Data," ECB Working Paper no. 208, January 2003.

[2] Acemoglu, Daron and Joshua Linn. "Market Size in Innovation: Theory and Evidence From the Pharmaceutical Industry", NBER Working Paper 10038, October 2003.

[3] Auerbach, Alan and Ronald Lee, ed. *Demographic Change and Fiscal Policy.* Cambridge University Press, Cambridge, January 2001.

[4] Bakshi, Gurdip S, and Zhiwu Chen. "Baby Boom, Population Aging, and Capital Markets," *Journal of Business,* Vol. 67(2), 165–201, 1994.

[5] Barber, Brad M. and Terrance Odean. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors", UC Berkeley Working Paper, April 2002.

[6] Basu, Sanjoy. "The Relationship Between Earnings Yield, Market Value, and Return for NYSE Common Stocks: Further Evidence", *Journal of Financial Economics,* Vol. 12, 129-156, 1983.

[7] Bell, Felicitie C., Alice H. Wade, and Stephen C. Goss. "Life Tables for the United States Social Security Area 1900-2080", Actuarial Study No. 107, SSA Pub. No. 11-11536.

[8] Bernard, Victor L. and Jacob K. Thomas. "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" *Journal of Accounting Research*, Vol. 27, 1-36, 1989.

[9] Bound, John and Sarah Turner. "Cohort Crowding: How Resources Affect Collegiate Attainment?" mimeo, University of Michigan, Ann Arbor, March 2003.

[10] Campbell, John Y. "A Variance Decomposition for Stock Returns", *Economic Journal,* Vol. 101, 157–179, 1991.

[11] Campbell, John Y. and Robert J. Shiller. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies,* Vol. 1, 195–228, 1988a.

[12] Campbell, John Y. and Robert J. Shiller. "Stock Prices, Earnings, and Expected Dividends", *Journal of Finance,* Vol. 43, 661–676, 1988b.

[13] Card, David, and Lemieux, Thomas. "Dropout and Enrollment Trends in the Post-War Period: What Went Wrong in the 1970s?" in J. Gruber, ed. *An Economic Analysis of Risky Behavior Among Youth.* University of Chicago Press, Chicago, 2000, 439–482.

[14] Costa, Dora. "American Living Standards: Evidence from Recreational Expenditures", NBER Working Paper 7148, May 1999.

[15] Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. "Investor Psychology and Security Market Under- and Overreactions", *Journal of Finance,* Vol. 53(6), 1839-1885, December 1998.

[16] De Long, J. Bradford, Andrei Shleifer. Lawrence H. Summers, Robert J.Waldmann. "Noise Trader Risk in Financial Markets", *Journal of Political Economy,* Vol. 98(4), 703–738, 1990.

[17] DeBondt, Werner F.M. and Thaler, Richard. "Does The Stock Market Overreact?" *Journal of Finance,* Vol. 40, 793–805, 1985.

[18] Easterlin, Richard A. *Birth and Fortune: The Impact of Numbers on Personal Welfare.* University of Chicago Press, Chicago, 1987.

[19] Fair, Ray C. and Kathryn M. Dominguez. "Effects of the Changing U.S. Age Distribution on Macroeconomic Equations", *American Economic Review,* Vol. 81(5), 1276–1294, December 1991.

[20] Fama, Eugene F., and Kenneth R. French. "Permanent and Temporary Components of Stock Prices", *Journal of Political Economy,* Vol. 96, 246–273, 1988a.

[21] Fama, Eugene F., and Kenneth R. French. "The Cross-section of Expected Stock Returns", *Journal of Finance,* Vol. 47, 427–465, 1992.

[22] Gabaix, Xavier and David Laibson "Bounded Rationality and Directed Cognition", Mimeo, MIT, 2002.

[23] Gabaix, Xavier, David Laibson, Guillermo Moloche and Stephen Weinberg."The Allocation of Attention: Theory and Evidence", Mimeo, MIT, 2002.

[24] Geneakoplos, John, Michael J.P. Magill and Martine Quinzii. "Demography and the Long-Run Predictability of the Stock Market", Yale Cowles Foundation Discussion Paper no. 1380, August 2002.

[25] Goyal, Amit. "Demographics, Stock Market Flows, and Stock Returns", *Journal of Financial and Quantitative Analysis,* Forthcoming.

[26] Gruber, Jonathan and David Wise. "Social Security Programs and Retirement Around the World," *Research on Labor Economics,* Vol. 18, 1999, 1-40.

[27] Hall, Brownyn, and Hall, Robert E."The Value and Performance of U.S. Corporations," *Brookings Papers on Economic Activity,* (1), 1-50, 1993.

[28] Heuser, Robert L. "Fertility tables for birth cohorts by color." U.S. Department of Health, Education, and Welfare. DHEW Publication No. 76-1152, 1976.

[29] Hong, Harrison and Jeremy C. Stein. "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets", *Journal of Finance,* Vol. 54(6), 2143-2184, December 1999.

[30] Hong, Harrison, Torous, Walter, and Ross Valkanov. "Do Industries Lead the Stock Market?" Mimeo, UCLA, July 2003.

[31] Huberman, Gur, and Regev, Tomer. "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar", *Journal of Finance*, Vol. 56(1), 387–396, February 2001.

[32] Hurd, Michael D. "Research on the Elderly: Economic Status, Retirement, and Consumption and Saving", *Journal of Economic Literature,* Vol. 28(2), 565-637, June 1990.

[33] Jegadeesh, Narasimhan, and Sheridan Titman. "Returns to Buying Winners and Selling Losers: Implications For Stock Market Efficiency", *Journal of Finance,* Vol. 48, 65-91, 1993.

[34] Lakonishok, Josef, Andrei Shleifer, and Robert. W. Vishny. "Contrarian Investment, Extrapolation, and Risk", *Journal of Finance,* Vol. 49(5), 1541-1578, 1994.

[35] La Porta, Raphael. "Expectations and the Cross-Section of Returns", *Journal of Finance,* Vol. 51, 1715-1742, 1996.

[36] Moskowitz, Tobias J. and Mark Grinblatt. "Do Industries Explain Momentum?" *Journal of Finance,* Vol. 54(4), 1249-1290, August 1999.

[37] Mankiw, Gregory N. and David N. Weil. "The Baby Boom, the Baby Bust, and the Housing Market", *Regional Science and Urban Economics,* Vol. 19, 235-258, 1989.

[38] Mankiw, Gregory N. and Michael D. Whinston. "Free Entry and Social Inefficiency", *RAND Journal of Economics,* Vol. 17(1), 48–58, Spring 1986.

[39] Merton, Robert C. "An Intertemporal Capital Asset Pricing Model", *Econometrica* Vol. 41, 867–87, 1973.

[40] Poterba, James M. "Demographic Structure and Asset Returns", *The Review of Economics & Statistics,* Vol. 83(4), 565-584, 2001.

[41] Shleifer, Andrei. *Inefficient Markets: An Introduction to Behavioral Finance. Clarendon Lectures in Economics.* Oxford and New York: Oxford University Press, 2000.

[42] Simons, Daniel J. and Steven L. Franconeri, Rebecca L. Reimer. "Change Blindness in the Absence of a Visual Disruption",. *Perception,* Vol. 29, 1143-1154, 2000.

[43] Simons, Daniel J. and Levin. "Failure to Detect Change to People in a Real-World Interaction", *Psychonomic Bulletin and Review,* 5(4), 644–649, 1998.

[44] U.S. Department of Commerce, Bureau of the Census. *Census of Manufacturers,* Washington DC: US Government Printing Office, various years.

[45] U.S. Department of Commerce, Bureau of the Census. "Estimates of the Population of the US by Age, Sex, and Race", *Current Population Reports,* Series P-25 (Numbers 311, 519, 917, 965, and 985) Washington DC: US Government Printing Office, various years.

[46] U.S. Department of Labor. Bureau of Labor Statistics. *Consumer Expenditure Survey, 1960-1961.* ICPSR 9035. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 1983

[47] U.S. Department of Labor. Bureau of Labor Statistics *Survey of Consumer Expenditures, 1972-1973.* ICPSR 9034. Inter-university Consortium for Political and Social Research, 1987.

[48] U.S. Department of Labor, Bureau of Labor Statistics and United States Department of Agriculture, Bureau of Home Economics, et al. *Study of Consumer Purchases in the United States, 1935-1936.* ICPSR 8908. Ann Arbor,MI: Inter-university Consortium for Political and Social Research, 1999.

[49] US Department of Health and Human Services, Public Health service, Centers for Disease Control and Prevention, National Center for Health Statistics, *Vital Statistics of The United States 1992: Natality* Vol. 1, 1996.

[50] Vuolteenaho, Tuomo. "What Drives Firm-Level Stock Returns", *Journal of Finance,* Vol. 57(1), 233–264, 2002.

[51] Watts, Ross L. "Systematic 'Abnormal' Returns After Quarterly Earnings Announcements", *Journal of Financial Economics*, Vol. 6, 127-150, 1978.

[52] Weber, E. H. *De Pulsu, Resorptione, Auditu, et Tactu. Annotationes Anatomicae et Physiologicae.* Koehler, Leipzig, 1834.

[53] Zarkin, Gary. "Occupational Choice: An Application to the Market for Public School Teachers", *Quarterly Journal of Economics,* Vol. 100(2), 409–446, 1985.

Figure 1a. Forecasts of Total Population Ages 30-34

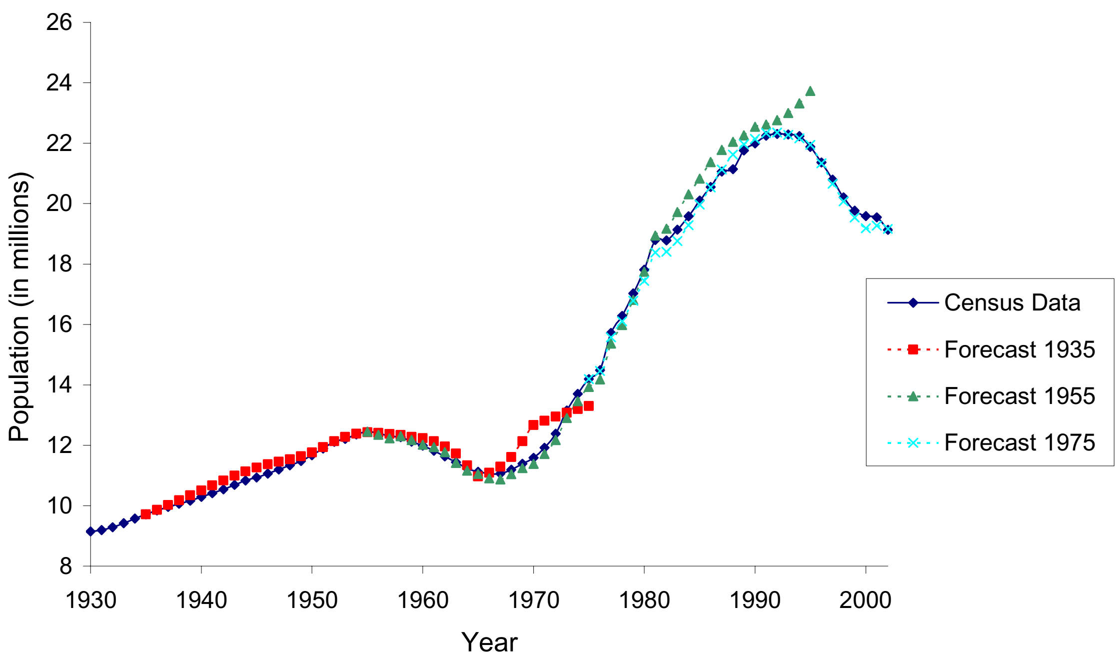
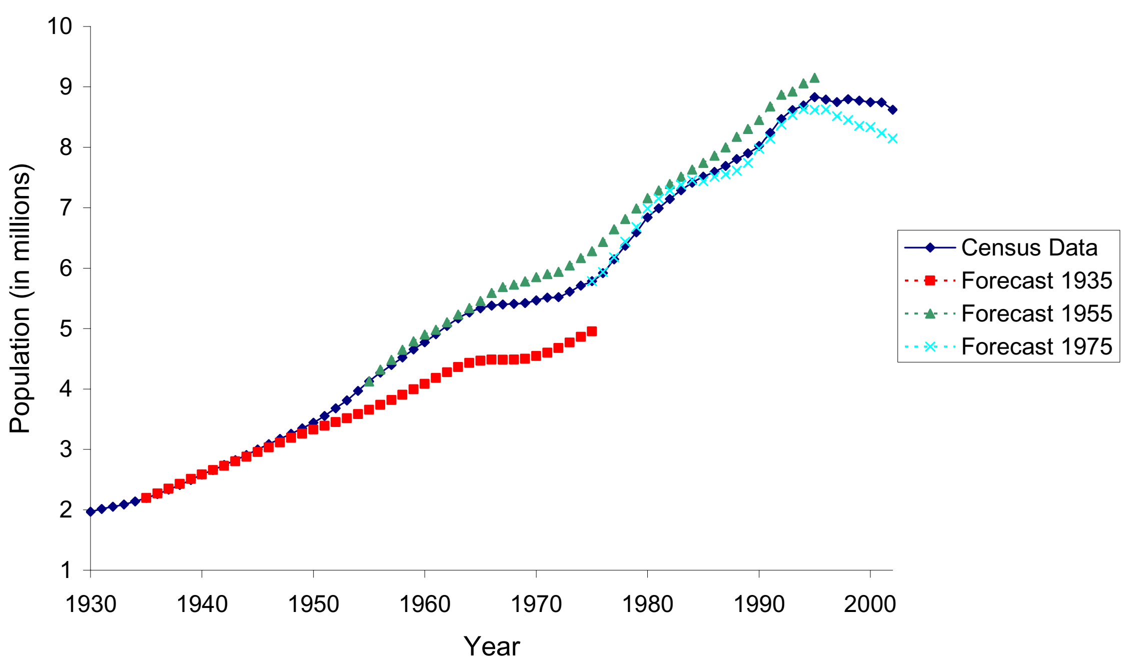


Figure 1b. Forecasts of Total Population Ages 70-74



Notes: Figures 1a and 1b display time series of actual and forecasted cohort size for the age groups 30-34 and 70-74. Each Figure shows the actual time series as well as three different 40-year forecasts as of 1935, 1955, and 1975.

Figure 2a. Age Profile of Consumption for Bicycles and Drugs

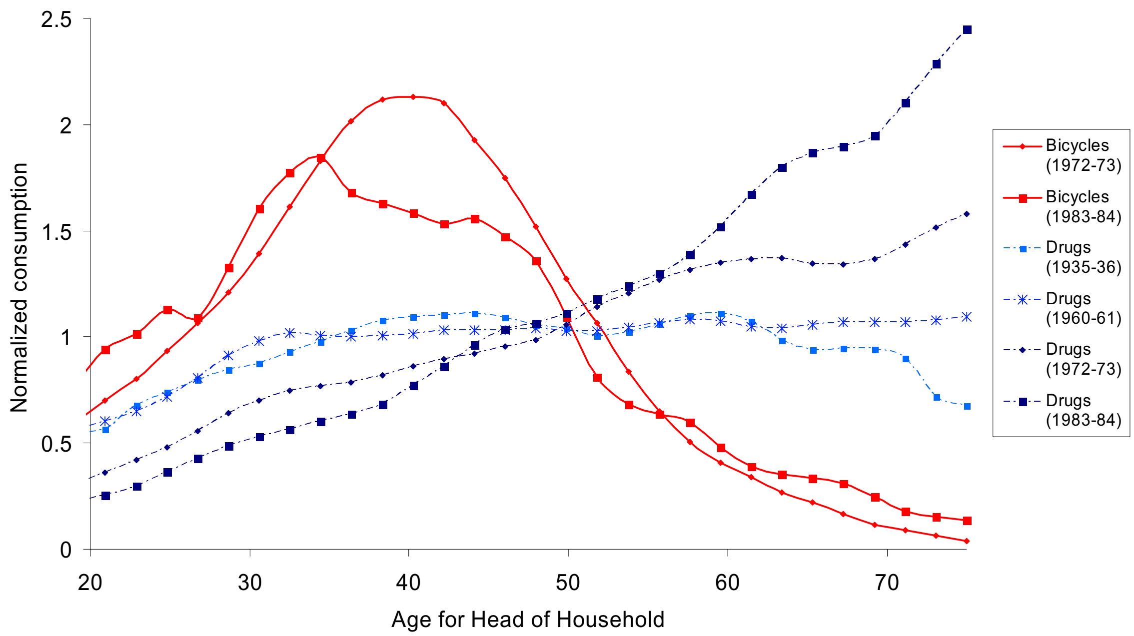
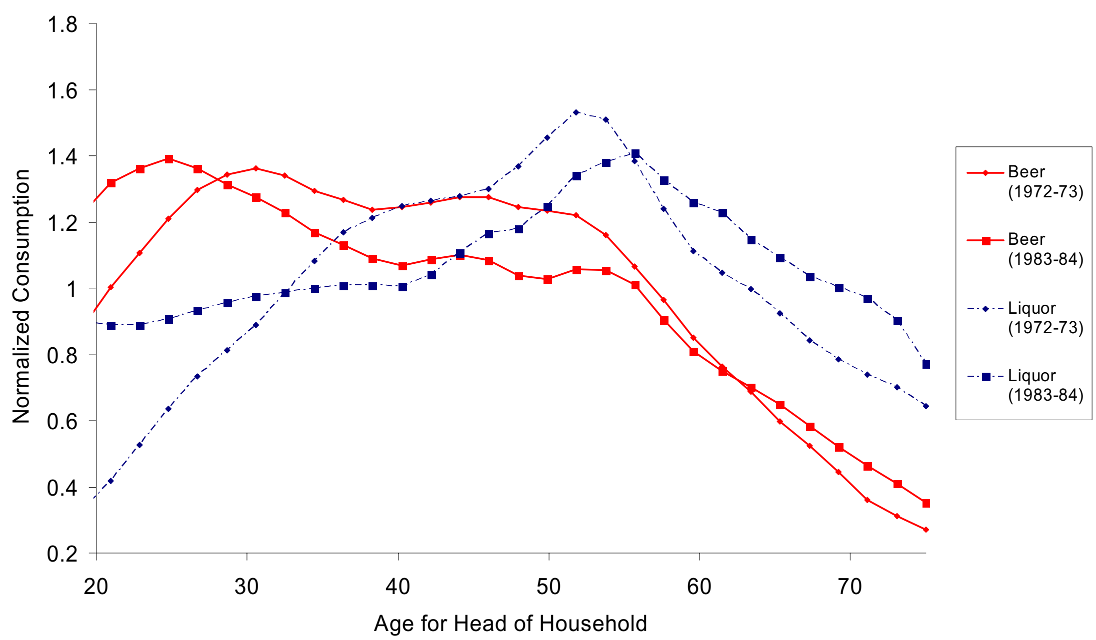
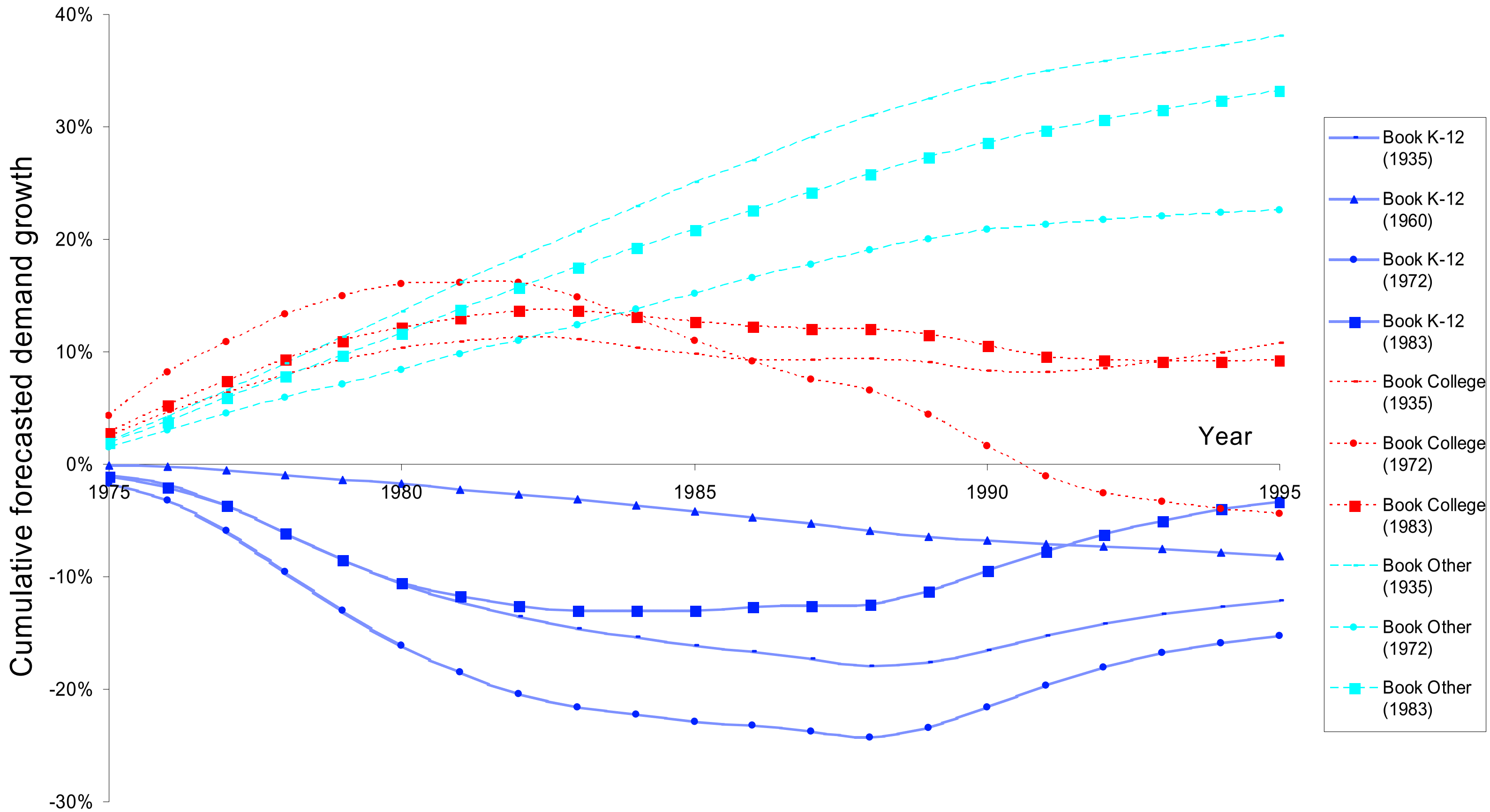


Figure 2b. Age Profile of Consumption for Beer and Liquor



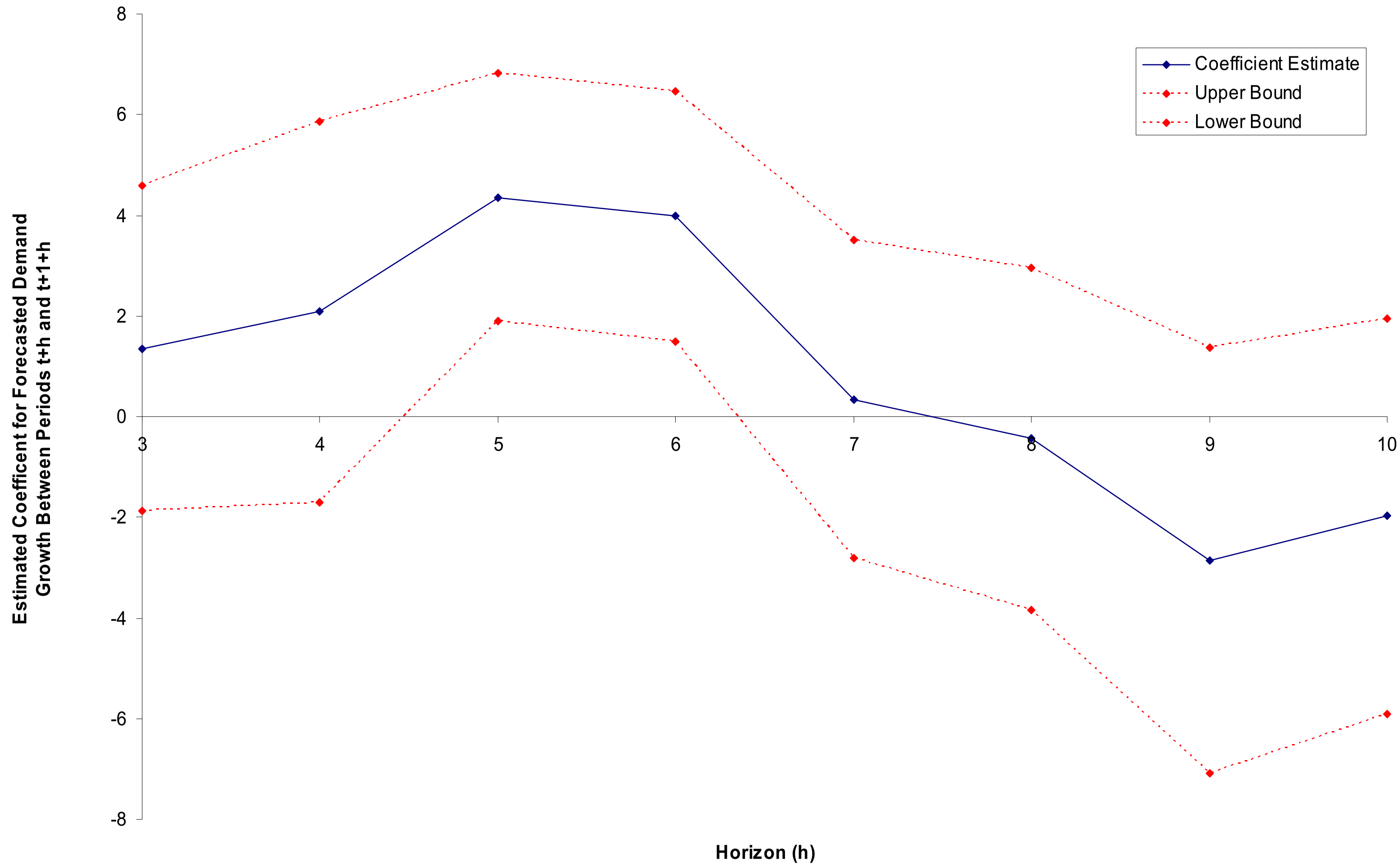
Notes: Figures 2a and 2b display a kernel regression of normalized household consumption for each good as a function of the age for the head of the household. The regressions use an Epanechnikov kernel and a bandwidth of 5 years. Each different line for a specific good uses an age-consumption profile from a different consumption survey. Expenditures are normalized so that the average consumption for all ages is equal to 1 for each survey-good pair. For bicycles and alcohol consumption, no data is available for the 1935-36 and the 1960-61 surveys.

Figure 3. Forecasted Demand Growth for Books



Notes: Figure 3 displays the predicted consumption growth due to forecasted demographic changes for three subcategories of books: books for K-12 schools, books for higher education, and other books (mainly fiction). The forecasts are computed combining the demographic information of year 1975 and age-consumption profiles for the 1935-36, 1960-61, 1972-73, and 1983-84 consumption surveys. Each distinct line for a good uses an age-consumption profile from a different data set. The forecast for higher education books and other books for 1960 is missing since the 1960-61 survey does not record book expenditures with a sufficient level of detail.

Figure 4: Return Predictability Coefficient for Demand Growth Forecasts at Different Horizons



Notes: The estimated coefficient for each horizon is from the regression of abnormal returns at t+1 on forecasted consumption growth between t+h and t+h+1. The estimates are reported in the second row of Table 12. The confidence interval is constructed using robust standard errors.

Table 1. Predictability of Population Growth Rates By Age Group

	Dependent variable: Annual Population growth rates for each age										
	Newborns	Ages 1-9	Ages 10-19	Ages 20-29	Ages 30-39	Ages 40-49	Ages 50-59	Ages 60-69	Ages 70-79	Ages 80-89	Ages 90-99
Constant	0.0135	0.0009	0.0020	0.0012	0.0029	0.0030	0.0029	0.0043	0.0068	0.0296	0.0400
	(0.0039)***	-0.0006	(0.0004)***	(0.0006)*	(0.0006)***	(0.0006)***	(0.0005)***	(0.0004)***	(0.0005)***	(0.0007)***	(0.0008)***
Forecasted annual population growth	0.8540	0.8450	0.9260	0.8379	0.8058	0.8189	0.7997	0.7328	0.6775	0.1133	0.2429
	(0.0905)***	(0.0109)***	(0.0074)***	(0.0112)***	(0.0115)***	(0.0108)***	(0.0111)***	(0.0138)***	(0.0143)***	(0.0099)***	(0.0146)***
R^2	0.4064	0.8342	0.9216	0.8100	0.7894	0.8141	0.7966	0.6808	0.6297	0.0895	0.1743
N	*N* = 132	*N* = 1188	*N* = 1320	*N* = 1320	*N* = 1320	*N* = 1320	*N* = 1320	*N* = 1320	*N* = 1320	*N* = 1320	*N* = 1320

Notes: Reported coefficents from the regression of annual population growth rates by age and gender onto the corresponding 1 year ahead forecasted annual growth rates. The regression specification is $y_{it} = \alpha + \beta x_{it} + \varepsilon_{it}$ where t is a year ranging form 1935 to 2000 and i is a age-gender observation within the relevant age range indicated at the top of each column. The OLS standard errors are in parentheses. Actual population sizes for both sexes between the ages 0 and 99 are from the P-25 Series from the Current Population Reports provided by U.S. Census. Forecasted population sizes for each age-gender observation are calculated using the previous year's P-25 data and mortality rates from the period life table at the beginning of the decade from Life Tables for the United States Social Security Area 1900-2080. The forecasted of number of newborns is calculated by applying birth rates from the previous year to the forecasted age profile of the female population. The actual and estimated growth rates are defined as the difference in the log population for a particular age-gender pair.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2. Summary Statistics: Household Demographics

Consumer Survey	1935-36	1960-61	1972-73	1983-84
Demographic Variables	(1)	(2)	(3)	(4)
Age of Head	44.26	48.28	47.87	44.17
	(12.7)	(15.68)	(17.38)	(18.3)
Male Head	1.00*	0.83	0.78	0.66
	(.)	(.37)	(.42)	(.47)
White Head	0.90	0.88	0.90	0.85
	(.29)	(.32)	(.3)	(.35)
Married Head	1.00*	.77*	0.68	0.52
	(.)	(.42)	(.47)	(.5)
Age of Spouse	40.36	(.)*	42.96*	43.16*
	(12.12)	(.)	(15.1)	(15.54)
No. of Children Living at Home	1.29	1.12	1.05	0.74
	(1.28)	(1.46)	(1.52)	(1.15)
No. of Old People Living at Home	0.06	0.04	0.03	0.03
	(.26)	(.21)	(.18)	(.18)
Family Size	3.76	3.28	2.99	2.57
	(1.59)	(1.87)	(1.86)	(1.6)
Urban Household	0.50	0.75	0.84	0.91
	(.5)	(.43)	(.37)	(.28)
Economic Variables				
Total Income (in $)	11094.56	21144.98*	27347.78*	31262.18*
	(15087.03)	(16164.53)	(28872.33)	(37026.55)
Total Consumption (in $)	10030.84	16792.38	18108.06	17935.47
	(8132.27)	(10247.24)	(11743.3)	(13339.84)
Number of Observations	*N* = 6113	*N* = 13728	*N* = 19975	*N* = 13133

Notes: Columns 1-4 present household-level summary statistics on demographic and economics variables in the consumption surveys. Column 1 refers to the *Study of Consumer Purchases in the United States, 1935-36.* Column 2 refers to the *Survey of Consumer Expenditures, 1960-1961.* Column 3 refers to the *Survey of Consumer Expenditures, 1972-1973.* Column 4 refers to the *Consumer Expenditure Survey, 1983-84.*

***** The information on the age of the spouse is missing in the 1960-61 survey. The variable Age of spouse is defined for 13,534 (in 1972-73) and 6,798 (in 1983-84) observations. In the 1935-36 survey only male married heads are interviewed. The variable Married Head is defined for 13,722 observations in the 1960-61 survey. The variable Total Income is defined for 13,694 observations in 1960-61, 18,861 observations in 1972-73, and 9,230 observations in 1983-84.

Table 3. Summary Statistics: Expenditure by Good

Consumer Survey	**1935-36**		**1960-61**		**1972-73**		**1983-84**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expenditure Category	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Child Care	1.43	(32.36)	(.)	(.)	91.31	(384.58)	117.20	(602.53)
Children's Books	(.)	(.)	(.)	(.)	0.47	(15.59)	2.70	(39.01)
Children's Clothing	7.42	(35.16)	18.56	(65.07)	21.37	(87.63)	38.42	(122.59)
Toys	24.90	(56.37)	(.)	(.)	13.77	(65.22)	75.36	(211.85)
Books -- college text books	12.94	(99.00)	(.)	(.)	20.87	(141.47)	32.50	(129.94)
Books -- general	8.82	(56.52)	(.)	(.)	18.00	(92.56)	37.41	(102.77)
Books -- K-12 school books	25.09	(53.24)	(.)	(.)	5.75	(41.59)	5.15	(30.4)
Movies	84.33	(135.70)	(.)	(.)	101.76	(256.79)	77.44	(168.88)
Newspapers	101.31	(78.90)	147.71	(161.14)	53.16	(70.70)	87.27	(95.45)
Cruises	(.)	(.)	(.)	(.)	2.40	(73.91)	12.79	(334.96)
Dental Equipment	92.26	(220.23)	151.89	(331.08)	148.63	(400.42)	122.33	(396.62)
Drugs	75.18	(138.43)	223.29	(300.52)	109.58	(214.28)	105.30	(219.93)
Health Care (Services)**	338.53	(688.64)	688.70	(890.59)	800.52	(1160.57)	549.19	(1035.64)
Health Insurance**	338.53	(688.64)	688.70	(890.59)	800.52	(1160.57)	549.19	(1035.64)
Medical Equipment**	338.53	(688.64)	688.70	(890.59)	800.52	(1160.57)	549.19	(1035.64)
Funeral Homes and Cemet.	21.03	(248.98)	(.)	(.)	3.24	(95.05)	51.98	(531.13)
Nursing Home Care	18.70	(208.13)	(.)	(.)	14.31	(273.54)	13.84	(298.35)
Construction Equipment*	6414.77	(6228.08)	3657.25	(2899.4)	4608.23	(4033.02)	5607.49	(5299.56)
Floors	37.51	(167.73)	86.83	(358.19)	94.26	(389.43)	59.37	(400.31)
Furniture	87.56	(297.42)	246.19	(578.63)	295.62	(772.49)	277.51	(1078.15)
Home Appliances Big	164.52	(408.67)	231.24	(495.04)	408.62	(666.92)	322.09	(675.65)
Home Appliances Small	15.17	(48.06)	25.01	(65.31)	54.77	(150.70)	61.53	(179.32)
Housewares	18.18	(55.41)	46.01	(121.71)	21.36	(94.45)	31.66	(125.94)
Linens	44.17	(80.35)	108.89	(177.62)	108.02	(238.89)	75.46	(226.54)
Residential Construction*	6414.77	(6228.08)	3657.25	(2899.4)	4608.23	(4033.02)	5607.49	(5299.56)
Residential Development*	6414.77	(6228.08)	3657.25	(2899.4)	4608.23	(4033.02)	5607.49	(5299.56)
Residential Mortgage	217.45	(636.88)	379.23	(735.42)	636.00	(1449.82)	1140.54	(2635.34)
Beer (and Wine)	61.02	(255.37)	525.30	(1116.88)	337.49	(802.86)	508.11	(849.15)
Cigarettes	137.78	(203.99)	299.85	(328.04)	264.14	(365.08)	201.98	(304.69)
Cigars and Other Tobacco	63.36	(133.88)	(.)	(.)	24.90	(110.19)	14.43	(67.44)
Food	3130.90	(2041.04)	4104.13	(2369.29)	3968.45	(2847.73)	3084.30	(2004.85)
Liquor	(.)	(.)	(.)	(.)	19.55	(54.01)	(49.36)	(114.78)
Clothing (Adults)	931.04	(1054.04)	1092.44	(1163.94)	868.30	(989.58)	605.21	(865.95)
Cosmetics	69.53	(96.77)	(.)	(.)	148.58	(243.73)	111.70	(165.3)
Golf	12.80	(99.65)	(.)	(.)	(.)	(.)	(.)	(.)
Jewelry	4.33	(13.33)	(.)	(.)	30.05	(195.)	83.30	(493.15)
Sporting Equipment	21.84	(68.1)	98.29	(254.94)	103.80	(210.47)	80.49	(229.07)
Life Insurance	672.52	(1462.62)	460.57	(838.06)	531.77	(951.55)	240.33	(866.86)
Property Insurance	98.15	(169.49)	329.21	(339.97)	389.85	(431.1)	442.40	(555.45)
Airplanes	(.)	(.)	(.)	(.)	97.26	(353.83)	179.70	(633.14)
Automobiles	764.45	(2105.43)	1002.87	(2437.16)	1571.92	(3323.69)	1729.10	(5085.54)
Bicycles	6.49	(37.03)	(.)	(.)	24.06	(83.33)	11.19	(98.27)
Motorcycles	(.)	(.)	(.)	(.)	36.38	(296.60)	27.06	(331.38)
Coal	205.40	(254.93)	(.)	(.)	11.14	(70.34)	2.84	(42.57)
Oil	480.00	(614.89)	1504.18	(964.36)	893.12	(811.44)	1076.62	(930.53)
Telephone	106.19	(141.12)	253.18	(224.38)	390.99	(339.01)	409.22	(359.85)
Utilities	383.44	(350.99)	1161.90	(792.22)	768.81	(568.66)	1045.84	(832.67)
Number of households	*N* = 6113		*N* = 13728		*N* = 19975		*N* = 13133	

Notes: Columns 1, 3, 5, and 7 present the average yearly household expenditure in the featured category. Columns 2, 4, 6, and 8 present the standard deviation across households. Columns 1 and 2 refer to the *Study of Consumer Purchases in the United States, 1935-36.* Columns 3 and 4 refer to the *Survey of Consumer Expenditures, 1960-1961.* Columns 5 and 6 refer to the *Survey of Consumer Expenditures, 1972-1973.* Columns 7 and 8 refer to the *Consumer Expenditure Survey, 1983-84.*

***** The expenditure for the categories "Construction Equipment", "Residential Construction" and "Residential Development" is given by the imputed rent estimated from the value of the dwelling of residence.

****** The expenditure for the categories "Health Cares (Services)", "Health Insurance" and "Medical Equipment" is given by the total expenditure in health insurance, physicians, and hospitals.

Table 4. Summary Statistics For Predicted Demand Growth Rates

Expenditure Category	Grouping	Avg. Predicted 1-Year Growth Rate of Consumption	Std. Dev. Predicted 1-Year Growth Rate of Consumption	Demographic goods	Number of Observations
	(1)	(2)	(3)	(4)	(5)
Child Care	Children	0.0139	(0.0164)	Yes	65
Children's Books	Children	0.0164	(0.0165)	Yes	28
Children's Clothing	Children	0.0138	(0.0109)	Yes	65
Toys	Children	0.0138	(0.0069)	Yes	65
Books -- college text books	Media	0.0105	(0.0123)	Yes	65
Books -- general	Media	0.0122	(0.0049)	No	65
Books -- K-12 school books	Media	0.0116	(0.0152)	Yes	65
Movies	Media	0.0115	(0.0071)	Yes	65
Newspapers	Media	0.0127	(0.0037)	No	65
Cruises	Health	0.0167	(0.0057)	Yes	28
Dental Equipment	Health	0.0117	(0.0040)	No	65
Drugs	Health	0.0137	(0.0022)	No	65
Health Care (Services)	Health	0.0132	(0.0028)	No	65
Health Insurance	Health	0.0132	(0.0028)	No	65
Medical Equipment	Health	0.0132	(0.0028)	No	65
Funeral Homes and Cemet.	Senior	0.0187	(0.0069)	Yes	53
Nursing Home Care	Senior	0.0144	(0.0046)	No	65
Construction Equipment	House	0.0107	(0.0174)	Yes	65
Floors	House	0.0129	(0.0046)	No	65
Furniture	House	0.0114	(0.0067)	Yes	65
Home Appliances Big	House	0.0112	(0.0044)	No	65
Home Appliances Small	House	0.0117	(0.0040)	No	65
Housewares	House	0.0128	(0.0045)	No	65
Linens	House	0.0131	(0.0037)	No	65
Residential Construction	House	0.0107	(0.0174)	Yes	65
Residential Development	House	0.0130	(0.0032)	No	65
Residential Mortgage	House	0.0135	(0.0052)	No	65
Beer (and Wine)	Perishable	0.0116	(0.0066)	Yes	65
Cigarettes	Perishable	0.0104	(0.0053)	No	65
Cigars and Other Tobacco	Perishable	0.0134	(0.0020)	No	65
Food	Perishable	0.0122	(0.0021)	No	65
Liquor	Perishable	0.0155	(0.0041)	No	28
Clothing (Adults)	Clothing	0.0122	(0.0055)	Yes	65
Cosmetics	Clothing	0.0118	(0.0059)	Yes	65
Golf	Clothing	0.0127	(0.0072)	Yes	65
Jewelry	Clothing	0.0121	(0.0057)	Yes	65
Sporting Equipment	Clothing	0.0111	(0.0053)	No	65
Life Insurance	Insurance	0.0129	(0.0034)	No	65
Property Insurance	Insurance	0.0138	(0.0029)	No	65
Airplanes	Transport	0.0166	(0.0038)	No	28
Automobiles	Transport	0.0115	(0.0055)	Yes	65
Bicycles	Transport	0.0110	(0.0104)	Yes	65
Motorcycles	Transport	0.0107	(0.0058)	Yes	28
Coal	Utilities	0.0124	(0.0025)	No	65
Oil	Utilities	0.0119	(0.0034)	No	65
Telephone	Utilities	0.0125	(0.0035)	No	65
Electricity	Utilities	0.0124	(0.0025)	No	65
Total Consumption		0.0125	(0.0041)	No	65

Notes: Complete list of expenditure categories, grouping in 10 broader groups (Column 1), average predicted one-year demand growth rate due to demographic changes (Column 2), standard deviation of predicted demand growth rate due to demographic change (Column 3), definition of subsample "*Demographic Goods*" (Column 4), number of years with demand growth estimates for each type of expenditure (Column 5). The subsample "Demographic Goods" is formed by the 20 expenditure categories with the highest within-good standard deviation of consumption growth due to demographics.

Table 5. Correlation between Forecasts of 1-Year Consumption Growth

Goods:	Consumption surveys:	1935-36 Survey	1960-61 Survey	1972-73 Survey	1983-85 Survey
All expenditure categories	1935-36 Survey	1 (*N* = 2730)			
	1960-61 Survey	0.6591 (*N* = 2015)	1 (*N* = 2015)		
	1972-73 Survey	0.721 (*N* = 2665)	0.8611 (*N* = 2015)	1 (*N* = 2990)	
	1983-85 Survey	0.7400 (*N* = 2665)	0.8119 (*N* = 2211)	0.8609 (*N* = 2990)	1 (*N* = 2990)

		1935-36 Survey	1960-61 Survey	1972-73 Survey	1983-85 Survey
Expenditure on Cars	1935-36 Survey	1 (*N* = 65)			
	1960-61 Survey	0.8998 (*N* = 65)	1 (*N* = 65)		
	1972-73 Survey	0.8765 (*N* = 65)	0.9915 (*N* = 65)	1 (*N* = 65)	
	1983-85 Survey	0.8873 (*N* = 65)	0.9835 (*N* = 65)	0.9918 (*N* = 65)	1 (*N* = 65)

	Source of demogr. data:	Actual Demographic Data	Forecasted Demographic Data
All expenditure categories	Actual Demogr. Data	1 (*N* = 2730)	
	Forecasted Demographic Data	0.6858 (*N* = 2730)	1 (*N* = 2730)

Notes: This table presents simple correlation coefficients among different versions of the forecasted 1-year growth rate of consumption due to demographic changes. The forecasts are made using 2-year old information. The sample includes the 47 expenditure categories over the years 1938-2002. The first set of correlations varies the source of the consumption coefficients. The consumption forecasts in the first column are obtained using the age profile of consumption estimated on data from the *Study of Consumer Purchases in the United States, 1935-36* . The consumption forecasts in the second column use data from the *Survey of Consumer Expenditures, 1960-1961.* The third column uses data from the *Survey of Consumer Expenditures, 1972-1973.* and the last column uses the *Consumer Expenditure Survey, 1983-84* .

The first set of correlations is for the whole sample of 47 expenditure categories, while the second set uses only the expenditure on cars. The last set of correlations holds constant the age profile of consumption and varies the source of demographic data. The first column uses the actual demographic data from the P-25 Bureau of Census Series. The second column uses forecasted demographic data to compute growth rates.

Table 6. Summary Statistics: Compustat Accounting Data

	Log Yearly Return on Equity			
	(1)	(2)	(3)	(4)
Industry Category	**Mean**	**Std. Dev.**	**No. Years**	**No. Firms**
Child Care	0.1130	(0.1371)	28	2.28
Children's Books	0.0808	(0.0856)	21	1.86
Children's Clothing	0.1645	(0.0891)	39	1.92
Toys	0.1164	(0.0768)	38	8.61
Books -- college text books	0.1983	(0.0603)	23	1.87
Books -- general	0.1260	(0.0558)	39	6.82
Books -- K-12 school books	0.1393	(0.0455)	35	2.17
Movies	0.0853	(0.1261)	51	15.40
Newspapers	0.1578	(0.1041)	51	13.62
Cruises	0.2000	(0.0769)	15	3.40
Dental Equipment	0.0883	(0.1765)	39	2.68
Drugs	0.1871	(0.0218)	51	74.13
Health Care (Services)	0.1188	(0.0611)	33	35.64
Health Insurance	0.1002	(0.0435)	30	9.23
Medical Equipment	0.1395	(0.0313)	51	48.21
Funeral Homes and Cemet.	0.0671	(0.1032)	39	2.23
Nursing Home Care	0.0724	(0.0907)	32	11.53
Construction Equipment	0.1228	(0.1000)	39	20.28
Floors	0.0831	(0.0373)	45	4.74
Furniture	0.0993	(0.0297)	51	14.52
Home Appliances Big	0.1540	(0.0618)	51	17.75
Home Appliances Small	0.1572	(0.0321)	51	4.29
Housewares	0.1007	(0.0775)	37	2.84
Linens	0.1085	(0.1002)	36	3.81
Residential Construction	0.0730	(0.0929)	38	10.44
Residential Development	0.0676	(0.0494)	39	37.08
Residential Mortgage	0.1300	(0.1714)	36	9.24
Beer (and Wine)	0.1200	(0.0357)	51	6.31
Cigarettes	0.1648	(0.0444)	51	3.83
Cigars and Other Tobacco	0.2383	(0.2109)	51	4.50
Food	0.1321	(0.0229)	51	158.62
Liquor	0.1057	(0.0766)	51	4.85
Clothing (Adults)	0.1313	(0.0331)	51	42.35
Cosmetics	0.2336	(0.1436)	46	8.74
Golf	0.0528	(0.1124)	29	3.79
Jewelry	0.0889	(0.0510)	39	8.72
Sporting Equipment	0.1241	(0.1174)	51	5.50
Life Insurance	0.0967	(0.0761)	38	11.89
Property Insurance	0.1204	(0.0693)	26	16.11
Airplanes	0.1105	(0.0589)	51	33.75
Automobiles	0.1274	(0.0899)	51	52.29
Bicycles	0.0791	(0.1159)	34	1.32
Motorcycles	0.2645	(0.1517)	16	0.90
Coal	0.0700	(0.1062)	44	6.43
Oil	0.1121	(0.0378)	51	140.60
Telephone	0.0761	(0.0655)	51	16.19
Electricity	0.1053	(0.0275)	43	131.25
Total Consumption	0.1222	(0.0194)	51	2956.82

Notes: For each company ROE in year *t+1* is the ratio of earnings (Compustat data172) in year *t+1* to the book value of equity in year *t* (Compustat data60). The analogous industry measure of ROE is the weighted-average of ROE using the book value of equity at *t* as weights. In the table we transform the industry measure by taking the log of 1 plus the industry ROE. Companies entering or exiting the industry between year *t* and year *t+1* are excluded from the weighted average. Column 2 reports the within-industry standard deviation of the measure in Column 1. Also featured are the number of years for which the data is available (Column 3) and the average number of firms in the industry (Column 4).

Table 7. Summary Statistics: CRSP Data and Concentration Ratios

	Value Weighted Annual Log Stock Return				Concentration Ratio C-4	
	(1)	(2)	(3)	(4)	(5)	(6)
Industry Category	**Mean**	**Std. Dev.**	**No. Years**	**No. Firms**	**Largest 4 Firms**	**Year Measured**
Child Care	0.0935	(0.4258)	29	3.48	(.)	(.)
Children's Books	0.0380	(0.3460)	24	2.17	0.180	1947
Children's Clothing	0.0698	(0.3433)	41	2.93	0.120	1963
Toys	0.0707	(0.4430)	41	11.98	0.390	1947
Books -- college text books	0.1468	(0.2947)	41	2.00	0.180	1947
Books -- general	0.1079	(0.2449)	41	8.49	0.180	1947
Books -- K-12 school books	0.1158	(0.2756)	39	2.77	0.180	1947
Movies	0.1122	(0.3026)	65	22.40	(.)	(.)
Newspapers	0.1387	(0.2662)	65	15.09	0.257	1947
Cruises	0.1525	(0.3037)	17	3.82	(.)	(.)
Dental Equipment	0.0587	(0.3551)	65	3.18	0.400	1947
Drugs	0.1288	(0.1915)	65	93.20	0.280	1947
Health Care (Services)	0.1122	(0.3415)	35	55.74	(.)	(.)
Health Insurance	0.0913	(0.2211)	41	13.59	(.)	(.)
Medical Equipment	0.1518	(0.2276)	65	59.68	0.484	1963
Funeral Homes and Cemet.	0.0353	(0.4921)	41	2.61	0.260	1947
Nursing Home Care	0.0277	(0.4255)	34	16.97	(.)	(.)
Construction Equipment	0.1112	(0.2392)	41	24.22	0.420	1963
Floors	0.0845	(0.3582)	65	6.18	0.400	1992
Furniture	0.0986	(0.2654)	65	15.15	0.260	1947
Home Appliances Big	0.1129	(0.3028)	65	20.77	0.400	1947
Home Appliances Small	0.1305	(0.2513)	54	5.46	0.410	1963
Housewares	0.0826	(0.3114)	41	3.24	0.554	1947
Linens	0.0961	(0.5502)	38	4.58	0.303	1947
Residential Construction	0.0617	(0.4567)	41	12.80	(.)	(.)
Residential Development	0.0662	(0.3130)	41	51.78	(.)	(.)
Residential Mortgage	0.0781	(0.3753)	41	14.49	(.)	(.)
Beer (and Wine)	0.1142	(0.2287)	65	8.55	0.256	1947
Cigarettes	0.1266	(0.2148)	65	5.22	0.900	1947
Cigars and Other Tobacco	0.1288	(0.2143)	65	6.09	0.749	1947
Food	0.1150	(0.1643)	65	185.37	0.325	1947
Liquor	0.1349	(0.2244)	65	5.86	0.750	1947
Clothing (Adults)	0.1024	(0.2608)	65	49.51	0.093	1947
Cosmetics	0.1082	(0.2993)	65	9.14	0.240	1947
Golf	0.0346	(0.3963)	30	5.63	(.)	(.)
Jewelry	0.1055	(0.3464)	41	11.22	0.130	1947
Sporting Equipment	0.0839	(0.3839)	65	6.75	0.240	1947
Life Insurance	0.1150	(0.2748)	40	35.08	(.)	(.)
Property Insurance	0.1050	(0.2085)	63	15.75	(.)	(.)
Airplanes	0.1128	(0.2727)	65	39.86	0.590	1958
Automobiles	0.1092	(0.2372)	65	65.82	0.353	1947
Bicycles	0.0334	(0.4254)	36	1.50	0.420	1947
Motorcycles	0.2014	(0.3710)	21	1.48	0.420	1947
Coal	0.1012	(0.2473)	65	10.02	(.)	(.)
Oil	0.1147	(0.1744)	65	170.20	0.300	1992
Telephone	0.0791	(0.2324)	65	25.28	(.)	(.)
Electricity	0.0961	(0.1708)	65	144.38	(.)	(.)
Total Consumption	0.1063	(0.1583)	65			

Notes: The measure of value-weighted yearly stock return in year *t+1* is the average yearly stock return for all companies belonging to the industry between December 31 in year *t* and December 31 in year *t+1* (Column 1). The average is value-weighted by the market capitalization at the end of year *t*. Column 2 reports the within-industry standard deviation. Also featured are the number of years for which the data is available (Column 3) and the average number of firms in the industry (Column 4). The measure of concentration ratio C-4 is the ratio of revenue produced by the largest 4 companies over the total industry revenue. The year of measurement is the first year of availabiliy of data. The source is the *Bureau of Manufacturers*. The measure is the average across all the 4-digit SIC codes that define the industry, weighted by the revenue in the sector. The measure is missing for industries with no SIC codes within the manufacturing range (2000-3999).

Table 8. Predictability of Return on Equity Using Demographic Changes

	Dependent Variable: Annual Log Return on Equity (ROE) at *t+1*											
	Demographic Industries						All Industries					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.1054	0.2533	0.2998	0.0994	0.169	0.1855	0.1123	0.1842	0.2061	0.1002	0.17	0.1425
	(0.0076)***	(0.0366)***	(0.0374)***	(0.0110)***	(0.0203)***	(0.0230)***	(0.0076)***	(0.0198)***	(0.0197)***	(0.0108)***	(0.0206)***	(0.0230)***
Forecasted annualized demand growth between t and $t+2$	1.4031	1.4792	1.5014	1.6199	2.4933	1.3676	0.8543	1.271	1.622	1.4988	2.4119	2.5716
	(0.3683)***	(0.4065)***	(0.4777)***	(0.5472)***	(0.6171)***	(0.5982)**	(0.4275)*	(0.4475)***	(0.4950)***	(0.5875)**	(0.6752)***	(0.6249)***
Industry Fixed Effects		X	X		X	X		X	X		X	X
Year Fixed Effects			X			X			X			X
Year >= 1975				X	X	X				X	X	X
Clustering by Year	X	X	X	X	X	X	X	X	X	X	X	X
R^2	0.0121	0.2377	0.312	0.0105	0.2156	0.2785	0.0025	0.1985	0.2437	0.0054	0.2762	0.3186
N	$N = 720$	$N = 720$	$N = 720$	$N = 491$	$N = 491$	$N = 491$	$N = 1852$	$N = 1852$	$N = 1852$	$N = 1217$	$N = 1217$	$N = 1217$

Notes: Columns 1 through 12 report the coefficients of OLS regressions of log yearly return on equity at *t+1* (Table 6) on the forecasted annualized demand growth due to demographics between years *t* and *t+2* (Table 4). The forecast is made using information available as of year *t-1*, The subset "Demographic Industries" denotes the 20 industries in Table 4 with the highest within-industry standard deviation of 1-year consumption growth due to demographics. Robust standard errors clustered by year in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9. Predictability of Return on Equity and Industry Concentration

	Dependent Variable: Annual Log Return on Equity (ROE) at *t+1*								
	Concentration C-4 above median			Concentration C-4 below median			All industries		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.1283	0.2522	0.2256	0.1273	0.0966	0.1527	0.1129	0.0926	0.0956
	(0.0124)***	(0.0369)***	(0.0415)***	(0.0090)***	(0.0166)***	(0.0160)***	(0.0129)***	(0.0166)***	(0.0149)***
Forecasted annualized demand growth between *t* and *t+2*	0.9008	1.6205	3.0112	0.751	-0.0977	0.69	0.6914	-0.8853	-0.2149
	(0.7320)	(0.9064)*	(1.1046)***	(0.5331)	(0.5423)	(0.5318)	(0.7882)	(0.7864)	(0.7577)
C-4* (Forecasted annualized demand growth between *t* and *t+2*)							0.3965	4.8237	6.4129
							(2.0002)	(2.3069)**	(2.2175)***
Concentration C-4							0.0415	(.)	(.)
							(0.0290)		
Industry Fixed Effects		X	X		X	X		X	X
Year Fixed Effects			X			X			X
Clustering by Year	X	X	X	X	X	X	X	X	X
R^2	0.002	0.167	0.2619	0.0029	0.2236	0.3168	0.0104	0.1901	0.2556
N	*N* = 600	*N* = 600	*N* = 600	*N* = 619	*N* = 619	*N* = 619	*N* = 1219	*N* = 1219	*N* = 1219

Notes: Columns 1 through 6 report the coefficients of OLS regressions of log yearly return on equity at *t+1* (Table 6) on the forecasted annualized demand growth due to demographics between year *t* and year *t+2* (Table 4). The forecast is made using information available as of year *t-1*. Columns 1 through 3 report the results for the subsample of industries with concentration-ratio 4 higher than .31. Columns 4 through 6 report the results for the subsample of industries with concentration-ratio 4 lower than or equal to .31. Columns 7 through 9 report the results for the whole sample of industries for the years subsequent to the first measure of concentration (usually, 1947). Details on the concentration ratio measure are in Table 7 and in the text. Robust standard errors clustered by year in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10. Predictability of Stock Returns Using Demographic Changes

	Dependent Variable: Beta-Adjusted Log Industry Stock Returns at *t+1*											
	Demographic Industries						All Industries					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-0.0489	0.1139	0.1617	-0.0736	0.0701	0.3037	-0.0408	-0.0386	-0.0631	-0.0728	0.0802	0.0887
	(0.0298)	(0.0755)	(0.0554)***	(0.0439)	(0.0830)	(0.0760)***	(0.0321)	(0.0693)	(0.0671)	(0.0539)	(0.0852)	(0.0724)
Forecasted annualized demand growth between *t* and *t+5*	-2.0824	-2.0187	-3.2459	-0.1502	0.6275	-2.9931	-1.8678	-1.5081	-2.2978	-0.8513	-0.0211	-1.5213
	(2.2462)	(2.6404)	(2.6327)	(2.9863)	(3.4489)	(2.8127)	(2.2618)	(2.5684)	(1.9819)	(3.8302)	(4.3824)	(2.5084)
Forecasted annualized demand growth between *t+5* and *t+10*	5.6818	5.8781	5.3486	6.5133	8.9855	7.0119	4.7628	4.7721	3.7751	6.1532	8.3018	4.6291
	(2.3039)**	(2.9910)*	(2.8852)*	(2.3684)**	(3.7848)**	(3.6344)*	(2.2627)**	(2.8335)*	(2.5150)	(2.2261)**	(3.3103)**	(3.3166)
Industry Fixed Effects		X	X		X	X		X	X		X	X
Year Fixed Effects			X			X			X			X
Year >= 1975				X	X	X				X	X	X
Clustering by Year	X	X	X	X	X	X	X	X	X	X	X	X
R^2	0.0124	0.0362	0.2707	0.0204	0.0581	0.2342	0.0059	0.0291	0.188	0.0114	0.0568	0.1924
N	*N* = 809	*N* = 809	*N* = 809	*N*=517	*N*=517	*N*=517	*N* = 2207	*N* = 2207	*N* = 2207	*N* = 1273	*N* = 1273	*N* = 1273

Notes: Columns 1 through 12 report the coefficients of OLS regressions of log yearly beta-adjusted industry stock returns at *t+1* (Table 6) on the forecasted annualized demand growth due to demographics (Table 4). The forecasts are made using information available as of year *t-1*. The industry betas for year *t* are obtained regressing monthly industry returns on market returns for the 48 months prior to year *t*. The coefficients on the forecasted annual demand growth are normalized by the number of years of the forecast (5 for both coefficients). The coefficient indicates the typical increase for the annual industry abnormal log stock return due to an annualized one percentage point increase in consumption due to demographics over the years 0 to 5 (or 5 to 10). The subset, "Demographic Industries", denotes the 20 industries in Table 4 with the highest within-industry standard deviation of 1-year consumption growth due to demographics. Robust standard errors clustered by year in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11. Predictability of Stock Market Returns and Industry Concentration

	Dependent Variable: Beta-Adjusted Log Industry Stock Returns at *t+1*								
	Concentration C-4 above median			Concentration C-4 below median			All industries		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	-0.0655	0.1006	0.0874	-0.0325	-0.1004	-0.2422	-0.02	-0.008	-0.1317
	(0.0445)	(0.0830)	(0.0819)	(0.0332)	(0.0785)	(0.0750)***	(0.0343)	(0.0470)	(0.1288)
Forecasted annualized demand growth between *t* and *t+5*	-2.4259	-1.1324	-0.0537	-0.8806	-1.7419	-1.349	2.3257	0.5094	-0.0924
	(3.3231)	(3.7758)	(4.8284)	(2.8128)	(3.0366)	(2.8279)	(3.7504)	(4.1884)	(3.6969)
Forecasted annualized demand growth between *t+5* and *t+10*	7.7342	6.7608	2.1029	2.6769	3.5693	0.2046	-2.7272	-1.5445	-1.8516
	(4.0789)*	(4.8396)	(5.0673)	(2.2668)	(2.9461)	(3.0067)	(3.0176)	(4.2445)	(4.1772)
C-4 * (Forecasted annualized demand growth between *t* and *t+5*)							-14.1878	-8.0682	-3.8927
							(10.0214)	(11.3104)	(11.5375)
C-4 * (Forecasted annualized demand growth between *t+5* and *t+10*)							26.0341	22.4902	11.3368
							(11.2365)**	(13.8781)	(13.2240)
Concentration C-4							-0.0944	(.)	(.)
							(0.0917)		
Industry Fixed Effects		X	X		X	X		X	X
Year Fixed Effects			X			X			X
Clustering by Year	X	X	X	X	X	X	X	X	X
R^2	0.0142	0.0503	0.2314	0.0022	0.0258	0.2559	0.0112	0.0386	0.2115
N	*N* = 642	*N* = 642	*N* = 642	*N* = 667	*N* = 667	*N* = 667	*N* = 1309	*N* = 1309	*N* = 1309

Notes: Columns 1 through 9 report the coefficients of OLS regressions of log yearly beta-adjusted industry stock returns at *t+1* (Table 6) on the forecasted annualized demand growth due to demographics between t and t+5 and between *t+5* and *t+10* (Table 4). The forecast is made using information available as of year *t-1*. The industry betas for year t are obtained regressing monthly industry returns on market returns for the 48 months previous to year t. The coefficients on the forecasted annual demand growth are normalized by the number of years of the forecast, 5. The coefficient indicates the increase in log industry abnormal stock return due to an annualized one percentage point increase in consumption due to demographics. Columns 1 through 3 report the results for the subsample of industries with concentration-ratio 4 higher than .31. Columns 4 through 6 report the results for the subsample of industries with concentration-ratio 4 lower than or equal to .31. Columns 7 through 9 report the results for the whole sample of industries for the years subsequent to the first measure of concentration (usually, 1947). Details on the concentration ratio measure are in Table 7 and in the text. Robust standard errors clustered by year in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 12. Predictability of Returns and Investor Horizon

	Dependent Variable: Beta-Adjusted Log Industry Stock Returns at *t+1*							
	Demographic Industries							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-0.0309	-0.0278	-0.0482	-0.0472	-0.0521	-0.0573	-0.0586	-0.0383
	(0.028)	(0.027)	(0.0281)*	(0.0277)*	(0.0276)*	(0.0278)**	(0.0278)**	(0.031)
Forecasted demand growth between t+h and t+h+1	1.354	2.0816	4.3543	3.9855	0.3423	-0.4441	-2.8515	-1.9746
	(1.618)	(1.888)	(1.2340)***	(1.2424)***	(1.577)	(1.703)	(2.116)	(1.964)
Forecasted demand growth between t+h and t+h+1	-0.8917	2.5814	-1.0795	1.4895	5.0982	5.8307	4.1846	1.7997
	(2.230)	(2.087)	(1.965)	(1.983)	(2.3659)**	(1.8085)***	(2.525)	(2.860)
Forecasted demand growth between t+h-4 and t+h-3	1.0334	-3.2254	-0.3213	-2.079	-1.8216	-1.3017	2.5913	4.6876
	(1.596)	(1.3805)**	(1.992)	(1.313)	(1.436)	(1.396)	(1.4073)*	(1.3442)***
Horizon (h)	3	4	5	6	7	8	9	10
R-squared	0.0026	0.0086	0.0118	0.0161	0.0133	0.0146	0.0146	0.0152
N	*N = 809*	*N = 809*	*N = 809*	*N = 809*	*N = 809*	*N = 809*	*N = 809*	*N = 809*

Notes: Columns 1 through 8 report the coefficients of OLS regressions of log yearly beta-adjusted industry stock returns at *t+1* (Table 6) on the forecasted annualized demand growth due to demographics between t+h and t+1+h for different horizons. All forecasts are constructed using information available as of year *t-1*. The industry betas for year t are obtained regressing monthly industry returns on market returns for the 48 months previous to year t. The coefficients on the forecasted annual demand growth are normalized by the number of years of the forecast. The coefficient represents the average increase of the log industry abnormal stock return due to an annualized one percentage point increase in consumption due to demographics. Robust standard errors clustered by year in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 13. Performance of the Zero-Investment Portfolio for Demographic Industries

	Dependent Variable: Monthly Return on the Zero-Investment Portfolio								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.0557	0.0556	0.0567	0.0837	0.0828	0.0816	0.0474	0.0499	0.0495
	(0.0189)***	(0.0191)***	(0.0209)***	(0.0273)***	(0.0288)***	(0.0318)**	(0.0205)**	(0.0214)**	(0.0230)***
VW Index Excess Return	-0.1263	-0.1034	-0.1045	-0.1103	-0.0720	-0.0712	-0.1246	-0.1137	-0.1139
(VWRF)	(0.0501)**	(0.0537)*	(0.0537)*	(0.0700)	(0.0847)	(0.0864)	(0.0504)**	(0.0491)**	(0.0467)**
Size Factor Return		-0.1020	-0.1030		-0.1403	-0.1411		-0.0620	-0.0617
(SMB)		(0.0537)	(0.0734)		(0.0902)	(0.0919)		(0.0844)	(0.0798)
Value Factor Return		-0.0132	0.0108		-0.0466	-0.0486		-0.0330	-0.0321
(HML)		(0.0715)	(0.0704)		(0.0991)	(0.1065)		(0.1009)	(0.0965)
Momentum Factor Return			-0.0086			-0.0085			-0.0031
(UMD)			(0.0675)			(0.0806)			(0.0839)
Year >= 1975				X	X	X			
SIC Classification Only							X	X	X
R^2	0.0139	0.0173	0.0174	0.0097	0.0191	0.0192	0.0127	0.0142	0.0142
N	*N = 780*	*N = 780*	*N = 780*	*N = 336*	*N = 336*	*N = 336*	*N = 780*	*N = 780*	*N = 780*

Notes: Columns 1 through 9 report the coefficients of OLS regressions of the zero-investment portfolio monthly returns on different sets of monthly benchmark factors. The zero-investment portfolio is long industries with high predicted long-term demand growth and short industries with low predicted long-term demand growth. VWRF is the return on the CRSP value-weighted stock index minus the 1-month treasury rate. SMB and HML are the returns on the Fama-French factor-mimicking portfolios for size and book-to-market, respectively. UMD is the return on the factor-mimicking portfolio for momentum. For Columns 7 through 9 the classification of companies into industries only uses Standard Industry Classification (SIC) codes instead of SIC codes in conjunction with the authors' company-by-company classfication using historical information. The constant has been annualized to make its interpretation more straightforward. Heteroskedasticity and autocorrelation consistent standard errors are calculated using the Newey-West estimator with 6 lags (in parentheses).

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 14. Performance of the Zero-Investment Portfolio for All Industries

	Dependent Variable: Monthly Return on the Zero-Investment Portfolio								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.0315	0.0209	0.0173	0.0458	0.0457	0.0316	0.0109	0.0030	0.0121
	(0.0132)**	(0.0136)	(0.0140)	(0.0189)**	(0.0196)**	(0.0190)*	0.0175	(0.0160)	(0.0180)
VW Index Excess Return	-0.0700	-0.0214	-0.0194	-0.0106	0.0369	-0.0444	-0.1115	-0.0530	-0.0581
(VWRF)	(0.0326)**	(0.0333)	(0.0329)	(0.0395)	(0.0402)	(0.0402)	(0.0374)***	(0.0381)*	(0.0378)*
Size Factor Return		-0.0890	-0.0870		-0.1502	-0.1430		-0.0558	-0.0606
(SMB)		(0.0439)**	(0.0439)*		(0.0557)***	(0.0528)***		(0.0582)	(0.0544)
Value Factor Return		0.1658	0.1719		-0.0058	-0.0286		0.2416	0.2263
(HML)		(0.0472)***	(0.0474)***		(0.0679)	(0.0635)		(0.0615)***	(0.0506)***
Momentum Factor Return			0.0292			0.1086			-0.0729
(UMD)			(0.0361)			(0.0434)**			(0.0409)*
Concentration Ratio > 0.4				X	X	X			
Concentration Ratio <= 0.4							X	X	X
Year >= 1947	X	X	X	X	X	X	X	X	X
R^2	0.0114	0.0486	0.0500	0.0001	0.0132	0.0147	0.0167	0.0511	0.0562
N	*N=672*	*N=672*	*N=632*	*N=672*	*N=672*	*N=672*	*N=672*	*N=672*	*N=672*

Notes: Columns 1 through 9 report the coefficients of OLS regressions of the zero-investment portfolio monthly returns on different sets of monthly benchmark factors. The zero-investment portfolio is long industries with high predicted long-term demand growth and short industries with low predicted long-term demand growth. Long-term demand growth is measured from years 5 to 10 and the constituent portfolios of the strategy are rebalanced every year. VWRF is the return on the CRSP value-weighted stock index minus the 1-month treasury rate. SMB and HML are the returns on the Fama-French factor-mimicking portfolios for size and book-to-market, respectively. UMD is the return on the factor-mimicking portfolio for momentum. The concentration ratio measure is the first available data from the Census of Manufacturers for the ratio of revenue for the largest 4 firms to total industry revenue taken. Since this measure does not exist before 1947 the sample does not include data before 1947. The constant has been annualized to make its interpretation more straightforward. Heteroskedasticity and autocorrelation consistent standard errors are calculated using the Newey-West estimator with 6 lags (in parentheses).

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 15. Analyst Forecasts of Earnings at Different Time Horizons

	Number of companies with at least one forecast for the fiscal period					
	(1)	(2)	(3)	(4)	(5)	(6)
Fiscal Year 1	4144	2003	2141	5425	2767	2658
Fiscal Year 2	3992	2002	1990	5326	2763	2563
Fiscal Year 3	1961	1435	526	3164	2087	1077
Fiscal Year 4	537	478	59	485	269	216
Fiscal Year 5	337	317	20	173	121	52
Fiscal Year 6	2	0	2	3	0	3
Fiscal Year 7	1	0	1	1	0	1
Analysts >= 5 for FY1		X			X	
Analysts < 5 for FY1			X			X
Year of Forecast	1990	1990	1990	2000	2000	2000

Notes: The Table reports the number of companies in the I/B/E/S data set with at least one analyst forecast for the Fiscal Year h, where h is the horizon of the forecast. For example, the row Fiscal Year 2 denotes the availability of analysts making forecasts for the Fiscal Year 2 years ahead. The sample for columns 2 and 5 is restricted to companies with at least 5 analysts making forecasts for fiscal year 1 and the sample for columns 3 and 6 is restricted to companies with fewer than 5 analysts making forecasts for fiscal year 1. Columns 1 through 3 are formed using forecasts made in 1990. Columns 4 through 6 are formed using forecasts ade in 2000. Columns 2 and 5 restrict the sample to companies with at least 5 analysts in the base year. Columns 3 and 6 restrict the sample to companies with less than 5 analysts in the base year.

Appendix Table 1: Industries and their Standard Industrial Classification (SIC) Codes

Expenditure Category	Grouping	Standard Industrial Classification Codes
Child Care	Children	8350-8359
Children's Books	Children	(2730-2739)
Children's Clothing	Children	2360-2369, 5640-5649, (5130, 5137)
Toys	Children	(3940), 3941-3948, (3949), (5090), 5092, (5940), 5945, (6711), (7990)
Books -- college text books	Media	(2730-2739)
Books -- general	Media	5942, (2720-2739, 5192)
Books -- K-12 school books	Media	(2720-2739)
Movies	Media	7810-7819, 7820-7849
Newspapers	Media	2710-2729, (2730-2739, 5192)
Cruises	Health	4480-4489, (4410, 4411, 7990, 7999)
Dental Equipment	Health	3843, 8020-8029, (3840, 5047, 8090)
Drugs	Health	2830-2839, 5120-5129 (8090)
Health Care (Services)	Health	8000-8019, 8030-8049, (8050-8059), 8060-8071, (8072), 8080-8089, (8090-8092)
Health Insurance	Health	6320-6329
Medical Equipment	Health	3840-3842, 3844-3849, 5047, (5040, 5120-5129, 8090)
Funeral Homes and Cemet.	Senior	3995, 7260-7269, (3990, 6550, 6553)
Nursing Home Care	Senior	8050-8059, (6510, 6513, 6798, 8080-8089, 8360-8361)
Construction Equipment	House	3531, 5031-5039, 5210-5259, (3530, 5080, 5082)
Floors	House	2270-2279, 5713, (5020, 5710, 5719)
Furniture	House	2510-2519, 5021, 5712 (5020, 5710, 5719)
Home Appliances Big	House	3631-3633, 3639, 5720-5729 (3630, 3651, 5060, 5075, 5078)
Home Appliances Small	House	3634, (3630, 3645, 5020, 5023, 5060)
Housewares	House	3262, 3263, 3914, (3260, 3269, 3910, 5944, 5719)
Linens	House	2391-2392, 5714, (2390, 5020, 5710, 5719)
Residential Construction	House	1520-1529, (1540-1549)
Residential Development	House	6513, 6530-6539, 6552, (1520-1529, 6510, 6550)
Residential Mortgage	House	6160-6169
Beer (and Wine)	Perishable	2082, 2083, 2084, 5181, (2080, 2084, 2085, 5180, 5182, 5813)
Cigarettes	Perishable	2100-2119
Cigars and Other Tobacco	Perishable	2120-2199
Food	Perishable	0100-0299, 2000-2079, 2086, 2087, 2090-2099, 5140-5149, 5400-5499, 5812 (5810)
Liquor	Perishable	2085 (2080, 2084, 5180, 5182, 5810, 5813, 5920-5921)
Clothing (Adults)	Clothing	2310-2349 5136, 5137, 5610-5619, (5130), 5136
Cosmetics	Clothing	2844, 7231, (2840, 5120, 5122, 5130)
Golf	Clothing	(2320, 2329, 3940, 3949, 5090, 5130, 5940, 7990, 7999)
Jewelry	Clothing	3911, 3915, 5944, (3910, 5090, 5094, 5940)
Sporting Equipment	Clothing	3949, 5941, (2320, 2329, 2390, 3940-3948, 5090-5091, 5130, 5940, 5945, 7999)
Life Insurance	Insurance	6310-6319
Property Insurance	Insurance	6330-6339
Airplanes	Transport	3720-3729, 4511-4512, (4510, 4513)
Automobiles	Transport	3010-3019, 3710-3719, 5010-5019, 5510-5529
Bicycles	Transport	(3710, 3750-3759, 3714, 5090)
Motorcycles	Transport	(3750-3759, 3571)
Coal	Utilities	1200-1299
Oil	Utilities	1300-1399, 2910, 2911
Telephone	Utilities	4810-4811, 4813-4819
Utilities	Utilities	4910-4959

Notes: Complete list of expenditure categories (Column 1) with SIC industry classification (Column 2). Each expenditure category is associated to two sets of codes. The first set of codes (not in parentheses) corresponds to the 4-digit SIC codes that are uniquely identified with one category. The second set of codes (in parentheses) identifies the SIC codes that are explicitly associated with multiple categories or have a large number of misclassified companies. Randomly selected companies within each SIC code are searched to determine if an SIC code has many mis-classified companies or multiple expenditure categories. All companies in each SIC code listed in parentheses are subjected to an internet search to determine its expenditure category classification. If the internet search can not identify the specific category for one of these companies, then the company is excluded from our analysis.