NBER WORKING PAPER SERIES

ATTENTION, DEMOGRAPHICS, AND THE STOCK MARKET

Stefano DellaVigna
Joshua M. Pollet

Working Paper 11211
http://www.nber.org/papers/w11211

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2005

We thank George Akerlof, Colin Camerer, John Campbell, David Card, Zhiwu Chen, Liran Einav, Ed Glaeser, Claudia Goldin, Amit Goyal, Caroline Hoxby, Gur Huberman, Michael Jansson, Lawrence Katz, David Laibson, Ronald Lee, Ulrike Malmendier, Ignacio Palacios-Huerta, Ashley Pollet, Jack Porter, James Poterba, Matthew Rabin, Joshua Rauh, Andrei Shleifer, Jeremy Stein, Geoffrey Tate, Tuomo Vuolteenaho, Jeff Wurgler, seminar participants at Bocconi, Columbia GSB, Emory, Haas, Kellogg, Harvard, Ohio State University, Stanford (Economics Department and GSB), Trento, UC Berkeley, UI Urbana-Champaign, and the participants at the NBER Behavioral Finance Program Meeting, at the NBER Summer Insitute on Aging, at the WFA 2004, and at the ASSA Meetings 2004 for their comments. Jessica Chan, Fang He, Lisa Leung, Shawn Li, Fanzi Mao, Rebbecca Reed, and Terry Yee helped collect the dataset of industries. Dan Acland, Saurabh Bhargava and Christine Yee 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. Both authors thank the NSF for support through grant SES-0418206. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Attention, Demographics, and the Stock Market
Stefano DellaVigna and Joshua M. Pollet
NBER Working Paper No. 11211
March 2005
JEL No. G1, J1, D0

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 5 to 10 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 to 7 percent. We present a model of underreaction to information about the distant future that is consistent with the findings.

Stefano DellaVigna
UC, Berkeley
Department of Economics
549 Evans Hall #3880
Berkeley, CA 94720-3880
and NBER
sdellavi@econ.berkeley.edu

Joshua M. Pollet
University of Illinois at Urbana Champaign
340 Wohlers Hall
1206 S. Sixth St.
Champaign, IL 61820
pollet@uiuc.edu

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 implement a novel test of underreaction to information based on demographic variables. We examine whether investors respond appropriately to changes in the demographic structure of the United States.

One unusual feature characterizes demographic 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 for various goods. These shifts in demand induce predictable changes in profitability for industries that are not perfectly competitive. Consequently, the timing of the stock market reaction to these predictable demand shifts is a test of investor attention to determinants of profitability.

We illustrate the 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. If 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 returns be abnormally high in anticipation of greater future profitability?

The timing of abnormally high returns depends on the attention horizon of the marginal investor. According to the standard analysis with perfect attention, the marginal investor foresees the positive demand shift induced by demographic changes and purchases school bus stocks in 2004. The price of school bus shares increases in 2004 until the opportunity to receive abnormal returns in the future dissipates. In this case, forecastable changes in profitability do not predict stock returns after 2004.

Alternatively, under a particular form of inattention, agents neglect information about future profitability until it is embedded in realized profits. In this case, stock returns in the school bus industry are abnormally high in 2010, six years after the release of information. A second, more realistic form of inattention has investors that incorporate information about future profitability only up to a fixed horizon. For instance, investors may have a four-year horizon, since analyst forecasts for profitability are typically available for the next four years. In this case, stock returns are abnormally high in 2006, two years after the release of information. For both forms of inattention, formalized in Section 2, demographic information available in 2004 predicts industry abnormal returns between 2005 and 2010.

This example motivates a simple test of attention with respect to future outcomes. 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. If investors, instead, are inattentive to information about future profitability, demographic variables predict industry asset returns.

This test can provide an estimate of the attention horizon. Since cohort size is predictable both in the short-term and in the long-term, we can estimate separately the impact of short-term and long-term profitability changes on stock returns. This examination of investor attention is different from other tests of predictability based on announcements of quarterly earnings or performance information measured by previous returns (DeBondt and Thaler, 1985) or accounting ratios (Fama and French, 1992). These variables convey information about profitability that is not easily decomposable into short-term and long-term components.

In Section 3 we perform the attention test. We analyze stock return predictability in a set of 48 US industries over the period 1935-2003. 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 life, individuals consume housing, life insurance, and pharmaceuticals. The life cycle ends with nursing homes and funeral homes. Other expenditure categories, like clothing, food, and property insurance, have a less obvious association with a specific age group.

The empirical strategy follows six steps. In the first step, we use current cohort sizes, mortality tables, and fertility rates to forecast future cohort sizes. The forecasted cohort growth rates over the next ten years are close to the actual growth rates over the same horizon, as well as to the corresponding Census projections.

In the second step, we estimate age-consumption profiles for the 48 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 1983-84. 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. In each year, we identify the 20 industries with the highest forecasted standard deviation of consumption growth. This subsample, labeled Demographic Industries, is most likely to be 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 contemporaneous profitability for companies in an industry. For the Demographic Industries, the accounting return on equity increases by 1.5 to 2.5 percentage points for each additional percentage point of contemporaneous demand growth induced by demographics. The results are comparable for the whole sample of industries. The point estimates are larger in industries with a more concentrated industrial structure, but the estimates are not statistically different from zero. Finally, when we separate industries by target demographic groups, we find that much of the identification comes from industries selling products to the young, although the point estimates are similar for industries producing goods for adults and the elderly.

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 short-term and long-term forecasted demand growth. The short-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 5 to 10 percentage point increase in abnormal industry return. The effect of short-term demand growth on returns is negative but not statistically significant. The effects are comparable for the whole sample of industries.

The predictability of returns depends on industry concentration, a proxy for barriers to entry or market power. Industries with above-median concentration ratios exhibit predictability that is approximately twice as large as in the whole sample, while industries with below-median concentration ratios exhibit no predictability. We also analyze the relationship between stock returns and forecasted demand growth at different horizons. We find that demand growth 5 to 8 years ahead is the strongest predictor of returns.

Finally, in the sixth step we present an alternative measure of the stock return predictability due to demographics. We construct a zero-investment portfolio that is long in industries with high absolute and relative long-term forecasted growth and short in industries with low absolute and relative long-term forecasted growth. For the Demographic Industries, this portfolio outperforms various factor models by more than 7 percent per year. For the sample including all industries, the portfolio (marginally) outperforms the benchmark portfolios by approximately 3 percentage points per year. A portfolio constructed using only high-concentration industries earns annualized abnormal returns of over 8 percentage points.

In Section 4 we interpret these results within the framework of a model with inattentive 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 for analysts 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 4 years exist for less than 10 percent of the sample.

For a horizon h of approximately 5 years, the model of short-sighted investors matches the stylized facts in this paper. 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. 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, on stock returns. Moreover, a calibrated version of the model is consistent with the magnitudes of our findings.

We also consider alternative interpretations of the results. In particular, we discuss rational predictability, poor estimation of systematic risk, persistent regressors, generated regressors, asset manager horizon, and neglect of slowly-moving variables as possible explanations.

This paper contributes to the literature on the role of attention allocation in economics and finance (Barber and Odean, 2002; Gabaix, Laibson, Moloche, and Weinberg, 2004; Hirshleifer, Lim, and Teoh, 2004; Huberman and Regev, 2001; Peng and Xiong, 2004). The evidence in this paper suggests that individuals may simplify complex decisions by neglecting long-term information.

A related literature in financial economics analyzes the positive autocorrelation of stock returns at short horizons (Jegadeesh and Titman, 1993; Moskowitz and Grinblatt 1999; Hong, Touros, and Valkanov, 2003) and the post-earnings announcement drift (Watts 1978, Bernard and Thomas 1989). Three behavioral explanations of these phenomena rely on underreaction due to slow diffusion of information (Hong and Stein, 1999), fluctuations in overconfidence (Daniel, Hirshleifer, and Subrahmanyam, 1998), and investor sentiment (Barberis, Shleifer, and Vishny 1998). Since our forecasts use information that has been in the public domain for at least one year and possibly decades, our findings suggests that underreaction may persist for years.

This paper also extends the literature on the effect of demographic variables on corporate decisions and stock market behavior. Pharmaceutical companies introduce new drugs in response to predictable demand increases induced by demographics (Acemoglu and Linn, 2004). The evidence on the relationship between cohort size and the equity premium is mixed (Bakshi and Chen, 1994; Poterba, 2001; Geneakoplos, Magill, and Quinzii, 2002). Compared to the literature on the equity premium, our paper focuses on changes in demand across consumption goods rather than on aggregate shifts in demand for financial assets.

Finally, Mankiw and Weil (1989) find that contemporaneous cohort size partially explains the time-series behavior of housing prices. We generalize their approach by analyzing 48 industries and examining stock market returns where, unlike for housing prices, arbitrage should eliminate predictability. While we also find evidence of predictability, stock returns are predicted by forecasted demand growth in the distant future, rather than by contemporaneous demand growth.

The rest of the paper is structured as follows. Section 2 presents a 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 forecast 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 first examine the effects of demand changes on firm profitability. The model of the industrial organization serves to illustrate some simple comparative statics results, which follow also under more general models.

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 the uniqueness of the solution to the profit-maximization problem in the second stage. Let $\bar{q}$ be the average production choice of the competitors, then the second stage maximization problem is

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

Consider first the effect of a short-run increase in demand from α_0 to $\alpha_1 > \alpha_0$. The demand change occurs after the entry decision 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 to 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 in equilibrium:

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

If these conditions are satisfied for $q^*\left(\alpha_0\right)$ at $\alpha=\alpha_0$, it is easy to check that the left-hand side of (1) is (weakly) negative for $q=\alpha_1 q^*\left(\alpha_0\right)/\alpha_0$ at $\alpha=\alpha_1$. Since the objective function is strictly concave, $q^*\left(\alpha_1\right)\leq\alpha_1 q^*\left(\alpha_0\right)/\alpha_0$ follows. Similarly, the left-hand side of (1) is positive for $q=q^*\left(\alpha_0\right)$ 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 $q^*\left(\alpha_0\right)<q^*\left(\alpha_1\right)$. This proves the desired inequality for q^* as a function of α. In turn, this implies $0<\partial q^*\left(\alpha\right)/\partial\alpha\leq q^*\left(\alpha\right)/\alpha$.

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

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

where the last inequality makes use of $0\leq\partial q^*\left(\alpha\right)/\partial\alpha\leq q^*\left(\alpha\right)/\alpha$, of the assumption $P'<0$, and of the fact that the price P is higher than marginal cost $c'\left(q^*\right)$ by (1). Therefore, profitability is increasing in the demand shift α in the second stage, that is, in the 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 $c\left(q\right)=cq$, expression (1) implies that in the short-run $\partial q^*\left(\alpha\right)/\partial\alpha=q^*\left(\alpha\right)/\alpha$. As a consequence, expression (2) simplifies to

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

If the industry 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 a demand shift 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 and therefore the short-run equilibrium is more likely to apply. For example, with strong brand loyalty market entry may require a multi-year advertising campaign. 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. 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), $\widehat{E}[\cdot]$, 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 (see derivation in Appendix A):

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

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 $\Delta\widehat{E}_{t+1}[\cdot] = \widehat{E}_{t+1}[\cdot] - \widehat{E}_t[\cdot]$ is the change in expectations between periods.

Short-sighted investors have correct short-term expectations but incorrect long-term expectations about profitability. Let $E_t^*[\cdot]$ be the expectation operator for short-sighted investors at time t. Similarly, let $E_t[\cdot]$ 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) average log return on equity for periods $t+1+h-n$ to $t+h$:

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

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 $h \to \infty$, 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

[1] In addition, if the potential entrants ignore forecastable demographic changes, then the impact of demand changes on profits is larger.

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 profitability in the more distant future. In a world with costly information processing, these rules of thumb could be approximately optimal. The short-term forecasts embed most of the available information about profitability in the distant future. However, investors disregard useful information when they neglect long-term demographic variables. They do not realize that these demographic variables provide relatively precise forecasts of profitability even at long horizons.

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 $t+1$ the later periods are not incorporated.

Taking conditional rational expectations at time t (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.

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

The expected return between time t and time $t+1$ depends on the sum of three terms. For rational investors ($h \to \infty$), 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: $E_t r_{t+1} - \bar{r} = \rho^h \left(E_t roe_{t+1+h} - \overline{roe} \right)$. 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 on 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:

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

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 ($h \to \infty$), the expected return, $E_t r_{t+1}$, is independent of expected future 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 expected future demand growth h periods ahead, $E_t \Delta c_{t+1+h}$. 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); the value of $\partial E_t r_{t+1} / \partial E_t \Delta c_{t+1+h}$ 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 short-term expectations (for $w < 1$), then demand growth less than $h+1$ periods ahead forecasts returns negatively. This occurs because investors overreact to information in the near future. The negative relationship between short-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 make the unrealistic assumption that all investors have a horizon of exactly h periods. If the horizon instead varied between h and $h + \tilde{H}$, industry returns would be forecastable using demand growth rates due to demographics between years $t+h$ and $t + h + \tilde{H}$. The empirical specification in Section 3.7 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, we form two demand growth forecasts, one for short-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. Then, we estimate the response of industry profitability and stock returns to forecasted demand changes. Finally, we evaluate 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. (Additional details are in Appendix B.1) 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 update this information using the *Vital Statistics of the United States: Natality* (US Department of Health and Human Services).

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 10-19 is about 20 percent per decade. For the ages 40-49 it is about 10 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 for each 10-year age group using data from the most recent five-year period prior to year t. We attribute this difference to historical net migration and other systematic deviations generated by the forecasting methodology. For the 10-19 age group, the average imputed net migration is about .4 percent per year, while for the 40-49 age group it is approximately .05 percent per year. We apply this imputed adjustment for migration to the initial population forecasts made at time t.

We define $\hat{A}_{g,u|t} = \left[\hat{A}_{g,0,u|t},\ \hat{A}_{g,1,u|t},\ \hat{A}_{g,2,u|t}, ...\right]$ 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. $A_{g,j,u}$ is the actual cohort size of gender g alive at u with age j. 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 that depend on predicting future births. Figure 1b for the age group 70-74 shows that the forecasts for older people are less precise.

Table 1 evaluates the precision of our demographic forecasts. We focus the analysis on the same forecast horizons employed in our tests of return predictability: a short-term forecast over the next 5 years and a long-term forecast 5 to 10 years in the future. In Column 1 we regress the actual population growth rate over the next 5 years, $\log A_{g,j,t+5} - \log A_{g,j,t}$, on the forecasted growth rate over the same horizon, $\log \hat{A}_{g,j,t+5|t} - \log \hat{A}_{g,j,t|t}$. Each observation is a (gender)x(one-year age group)x(year of forecast) cell; this specification includes all age groups and years between 1937 and 1997.[2] The R^2 of 0.83 and the regression coefficient close to 1 indicate that the forecasts are quite accurate. The precision of the forecasts is comparable for the cohorts between 0 and 18 years of age ($R^2 = .82$, Column 2) but lower for the cohorts between 65 and 99 years of age ($R^2 = .56$, Column 3). Columns 4 to 6 show the corresponding results for forecasts 5 to 10 years in the future. The precision of these long-term forecasts is only slightly inferior to the precision of the short-term forecasts for the total sample (Column 4) and for the 65+ age group (Column 6). However, the accuracy of our forecasts is substantially lower for the cohorts up to age 18 (Column 5) because a large fraction of the forecasted cohorts are unborn as of year t.

Overall, our demographic forecasts predict cohort size growth quite well over the horizons of interest. These forecasts also parallel the Census Bureau population forecasts created using data from the 2000 Census. In Column 7 we regress the official forecast for population growth for the next 5 years, $\log \hat{A}^C_{g,j,2005|2000} - \log \hat{A}^C_{g,j,2000|2000}$, on our forecast, $\log \hat{A}_{g,j,2005|2000} - \log \hat{A}_{g,j,2000|2000}$, for age groups between 0 and 99. This regression has an R^2 of .85 and a coefficient estimate slightly greater than 1. Column 8 reports similarly precise results for forecasted demographic growth between 2005 and 2010.

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.*[3] We combine these three early surveys with the 1983-1984 cohorts of the ongoing *Consumer Expenditure Survey*[4]. These four cross-sections provide information on the age distribution of consumption throughout the past century. Appendix Table 1 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 Appendix Table 1 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 level of aggregation attempts to distinguish goods with potentially different age-consumption

[2]Cohort age groups older than 75 before 1940 or older than 85 from 1940 to 1979 are excluded from this analysis because the actual cohort sizes are imputed (see Appendix B.1).

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

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

profiles. For example, within the category of alcoholic beverages, we separate beer and wine from hard liquor expenditures. Similarly, within insurance we distinguish among health, property, and life insurance expenditures. We attempt to define these categories in a consistent way across the survey years. Unfortunately, some surveys do not provide enough information to construct certain expenditure categories. This problem is especially serious for the 1960-61 survey which classifies consumption data into broad categories. Table 2 presents the summary statistics on good-by-good annual household expenditure for each survey year. The expenditures are in 1982-84 dollars.[5] 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.

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 annual consumption on the age of the head of household[6]. Figure 2a, for example, plots normalized[7] 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 45. At these ages, the heads of household are most likely to have children between the ages of 5 and 10. 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 that have 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 (results not shown). Large appliances are largely associated with the purchase of the first house, while small appliances are purchased on a more regular basis.

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 household heads (child care and toys), some by heads in middle age (life insurance and cigars), others by older heads (cruises and 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 ages 25-35, whether we

[5]Appendix B.2 provides additional information about the consumption data. Further details are also available from the authors upon request.

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

[7]For each survey-good pair we divide age-specific consumption for good k by the average consumption across all ages for good k.

consider the 1935-36, the 1960-61, the 1972-73, or the 1983-84 cohorts. Taken as a whole, the evidence suggests that changes in age structure of the population have the power to influence consumption demand in a substantial and consistent manner.

With this evidence in the background, we 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. We 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 indicator variables 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 equal to 1 if the head of household i in year t is at least as old as j and younger than the next age group. For example, if $H_{35,i,t} = 1$ then the head of household i is aged 35 to 44 in year t. The variable $H_{65,i,t}$ indicates that the age of the head of household is greater than or equal to 65. Similarly, let $S_{i,t}$ be a set of indicator variables 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 count the total number of other individuals (children or old relatives) living with the family in year t. For instance, if $O_{0,i,t} = 2$, then 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 estimated separately for each good k and for each of the four cross-sections t. The purpose is to obtain estimates of annual consumption of good k for individuals at different ages. For example, the coefficient $B_{35,\text{cars},1960}$ is the average total amount that a (single) head aged 35 to 44 spends on cars in 1960.[8]

3.3 Demand forecasts

In the third step of the research design, we combine the estimated age profiles of consumption with the demographic forecasts in order to forecast future demand for different goods. For example, consider a forecast of toys consumption in 1975 made as of 1965. For each age group, we multiply the forecasted cohort sizes for 1975 by the age-specific consumption of toys estimated on the most recent consumption data as of 1965, that is, the 1960-61 survey. Next,

[8] We do not include the set of spouse variables in the 1935-36 survey (only married couples were interviewed) and in the 1960-61 survey (the age of the spouse was not reported). Since the size of sample for the 1935-1936 survey is only a third to a half as large as the sample sizes for the other surveys, for this survey we use broader age groups for the head variables: 18, 35, 50, and 65.

we aggregate across all the age groups to obtain the forecasted overall demand for toys for 1975.

Formally, let $\hat{A}^b_{g,u|t}$ be 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 perform this operation, we estimate the shares $h_{g,j,t}$, $s_{g,j,t}$, and $o_{g,j,t}$ of people in the population for each age group j. For instance, $h_{f,35,t}$ is the number of female heads 35-44 divided by the total number of females aged 35-44 in the most recent consumption survey prior to year t. We 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 prior 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 of age 45 in 1975 consume the same bundles of goods that individuals of age 45 consumed in 1965. By construction, we hold the prices of each good constant at its level in the most recent consumption survey prior to year t.[9]

Figure 3 shows the results of the consumption forecasts for three subcategories of the general book category—books for K-12 schools, books for higher education, and other books (mostly fiction). We plot the predicted cumulative demand growth from 1975 to 1995 using the information available in 1975 from the expression $\ln \hat{C}_{k,u|1975} - \ln \hat{C}_{k,1975|1975}$ for $u = 1975, 1976, ..., 1995$. For each of the three goods, we produce forecasts using the age-consumption profiles estimated from each of the three consumption data sets that record detailed expenditure for books. The demand for K-12 books is predicted to experience a decline as the baby-bust generation continues to enter schools, followed by an increase. 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 between the ages of 30 and 50, is predicted to grow substantially as members of the baby-boom generation gradually reach these ages. These patterns do not depend on the year of expenditure survey (1935-36, 1972-73, or 1983-84) used to estimate the age-consumption profile for each category. We draw two main conclusions. First, within the entire book category there is substantial variability in the pattern of demand growth across the subcategories. Second, the time-series variation in consumption growth appears to be fairly independent of the consumption survey used to estimate the age profile.

[9]See Appendix B.2 for information on the calculation of forecasted demand growth rates for construction machinery and residential construction.

While we cannot present the same detailed information for all goods, we report the consumption forecasts at three points in time. Columns 2, 4, and 6 of Table 3 summarize the five-year predicted growth rate due to demographics, $\ln \hat{C}_{k,t+5|t-1} - \ln \hat{C}_{k,t|t-1}$, respectively for years $t = 1950$, $t = 1975$, and $t = 2000$. In each case, data from the most recent consumer expenditure survey is used. In 1950, child-related expenditures are predicted to grow very quickly due to the boom in births starting in 1947. Demand for housing-related goods is relatively low due to the small size of cohorts born in the 1930s. In 1975, the demand for child care and toys is low due to the small size of the 'Baby Bust' generation. The demand for most adult-age commodities is predicted to grow at a high rate (1.5-2 percent a year) due to the entry of the 'Baby Boom' generation into prime consumption age. In 2000 the demand for child-related commodities is relatively low. The aging of the 'Baby Boom' generation implies that the highest forecasted demand growth is for goods consumed later in life, such as cigars, cosmetics, and life insurance.

Table 3 also categorizes goods by their sensitivity to demographic shifts. For example, the demand for oil and utilities is unlikely to be affected by shifts in the relative cohort sizes, while the demand for bicycles and motorcycles depends substantially on the relative size of the cohorts aged 15-20 and 20-30, respectively. We construct a forward-looking measure of Demographic Industries using information available at time $t-1$ to identify the goods where demographics shifts are likely to have the most impact. In each year t, we compute the standard deviation of the one-year consumption forecasts up to 15 years ahead given by $\left(\ln \hat{C}_{k,t+s+1|t-1} - \ln \hat{C}_{k,t+s|t-1}\right)$ for $s = 0, 1, ..., 15$. We define the set of Demographic Industries as the 20 industries with the highest standard deviation of growth[10]. Column 3 shows that in 1950 the Demographic Industries are associated with high demand by children (child care, toys) and by young adults, such as housing. The classification is similar in the later years 1975 (Column 5) and 2000 (Column 7). Finally, Column 8 summarizes the percentage of years in which an industry belongs to the Demographic Industries subsample.

Since the demand forecasts use the most recent demographic and consumption information, the forecasts for later years use different consumption surveys than the forecasts for earlier years. We verify that the impact of using different consumption surveys is limited. For each good and over the years 1939-2003, we generate annual consumption growth forecasts $\ln \hat{C}_{k,t+1|t-1} - \ln \hat{C}_{k,t|t-1}$ using estimates of the age profile of consumption from the 1935-36 survey. We repeat this process using age-consumption estimates from the 1960-61, the 1972-73, and the 1984-84 surveys. Next, we compute the correlation between these four measures of consumption growth. Using data for all goods and years, the correlations are in the .65 to .8 range (results not shown). These correlations confirm that the consumption patterns are

[10] A previous version of this paper used the standard deviation of annual forecasted demand growth over the years 1939-2003 to classify the Demographic Industries. The results are similar, but the current definition has the advantage that the classification in year t only uses information available up to year $t-1$.

similar across surveys.

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 through 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 ('toys') includes both companies producing toys for children and companies manufacturing golf equipment, two goods clearly associated with consumption by different age groups. Appendix Table 2 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. (See Appendix B.3 for details)

We construct the annual industry return on equity $ROE_{k,t+1}$ as the weighted average of ROE for the companies in industry k. We use the book value for each company in year t as the weights and drop companies with negative book values. 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.[11] Columns 1 through 4 of Table 4 present 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 sample is limited to the years in which the consumption data is also available. The average log return ranges from 4 percent (golf) to 26 percent (motorcycle). The within-industry standard deviation of the return ranges between 2 percent (drugs) and 15 percent (cigars). The longest series have 52 observations, but many series are shorter. The average number of firms per industry varies between 1.2 (motorcycle) and 167 (food).

In Table 5 we test the predictability of the one-year industry log return on equity (Table

[11] The results are qualitatively similar for the unwinsorized measure.

4) using the forecasted contemporaneous growth rate in consumption due to demographics (Table 3). 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):

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

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 July 1 value, approximately in the middle of the fiscal year, we use the average demand growth between July 1 of year t and July 1 of year $t+2$ as a measure of contemporaneous demand change. We scale by 2 to annualize this measure. The forecast of consumption growth between years t and $t+2$ uses only demographic and consumption information available up to year $t-1$. This lag ensures that all information should be public knowledge by year t.[12] We run specification (8) both with and without industry and year fixed effects.

In this panel setting it is unlikely that the errors from the regression are uncorrelated across industries and over time because there are persistent shocks that affect multiple industries at the same time. We allow for arbitrary correlation across industries at any given time by calculating standard errors clustered by year. In addition, we correct these standard errors to account for autocorrelation in the error structure.

More formally, let X be the matrix of regressors, θ the vector of parameters, and ε the vector of errors. The panel has T periods and K industries. Under the appropriate regularity conditions, $\sqrt{\frac{1}{T}}(\hat{\theta} - \theta)$ is asymptotically distributed $N(0, (X'X)^{-1} S(X'X)^{-1})$ where $S = \Gamma_0 + \sum_{q=1}^{\infty}(\Gamma_q + \Gamma_q')$ and $\Gamma_q = E[(\sum_{k=1}^{K} X_{kt}\varepsilon_{kt})'(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q})]$. The matrix Γ_0 captures the contemporaneous covariance, while the matrix Γ_q captures the covariance structure between observations that are q periods apart. While we do not make any assumptions about contemporaneous covariation, we assume that $X_{kt}'\varepsilon_{kt}$ follows an autoregressive process given by $X_{kt}'\varepsilon_{kt} = \rho X_{kt-1}'\varepsilon_{kt-1} + \eta_{kt}'$ where $\rho < 1$ is a scalar and $E[(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q})'(\sum_{k=1}^{K} \eta_{kt})] = 0$ for any $q > 0$.

These assumptions imply $\Gamma_q = \rho^q\Gamma_0$ and therefore, $S = [(1+\rho)\,/\,(1-\rho)]\Gamma_0$. (See derivation in Appendix C) The higher the autocorrelation coefficient ρ, the larger the terms in the matrix S. Since Γ_0 and ρ are unknown, we estimate Γ_0 with $\frac{1}{T}\sum_{t=1}^{T} X_t'\hat{\varepsilon}_t\hat{\varepsilon}_t'X_t$ where X_t is the matrix of regressors and $\hat{\varepsilon}_t$ is the vector of estimated residuals for each cross-section. We estimate ρ from the pooled regression for each element of $X_{kt}'\hat{\varepsilon}_{kt}$ on the respective element of $X_{kt-1}'\hat{\varepsilon}_{kt-1}$.

We use the set of Demographic Industries for the years 1974-2003 as the baseline sample for the paper. As discussed above, the Demographic Industries are more likely to be affected

[12] 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.

by demographic demand shifts. As for the time period, data accuracy is higher over the more recent time period in at least two respects. First, the number of companies included in the accounting and return data increases substantially over time, and in particular it almost doubles in 1974 with the introduction of Nasdaq data into CRSP. Second, the accuracy of the industry classification increases with proximity to the present[13].

In Column 1 of Table 5 we present the results of specification (8) for the baseline sample without industry or year fixed effects. The estimated coefficient, $\hat{\theta} = 1.85$, 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 11.0 percent to an average of 12.8 percent, a 16 percent increase. The R^2 of the regression is low due to the modest role of demographic changes relative to other determinants of profitability. In this specification, as well as in the subsequent specifications, controlling for autocorrelation is important: the estimated $\hat{\rho}$ equals approximately .5, resulting in a correction coefficient $(1+\hat{\rho})/(1-\hat{\rho}) = 3$. In Column 2 we introduce industry indicators. In this case, the identification depends only on time-series changes in the growth rates and not on between-industry differences. The estimate for θ is significantly positive and larger than in Column 1, with $\hat{\theta} = 2.86$. In Column 3 we introduce year indicators as well. In this specification, the identification depends on differential time-series in demand changes across industries. The estimated coefficient, $\hat{\theta}$, has a similar magnitude as in Column 1, and is also statistically significantly different from zero.

In Columns 4-6 we reestimate the model for the whole time period 1939-2003. The estimates for θ are lower than our baseline results, but still economically large and significant (except in Column 4). Finally, in Columns 7 through 12 we reestimate the same models for the whole sample of 48 industries. The point estimates for θ are somewhat lower than the corresponding ones for the subset of Demographic Industries, but are still large and significant in most specifications. The standard errors in the whole sample are larger than those for the Demographic Industries, despite a threefold increase in sample size, suggesting a lower 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.

3.5 Industry concentration

The 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

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

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 how the impact of demand changes on profitability varies with 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 fraction 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 by taking the average C-4 ratio for the SIC codes included in the industry 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 concentration measures as of 1972. 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. Among the 32 industries with concentration data (Column 9 in Table 4), the median C-4 ratio is .35.

For the subsample of industries with above-median concentration (Columns 1 and 2 of Table 6), the magnitude of the coefficient $\hat{\theta}$, capturing the impact of demographics on profitability, is similar to the benchmark estimates (Table 5), but is not significant. For the sample of unconcentrated industries (Columns 3 and 4), the coefficient $\hat{\theta}$ is fifty percent smaller and is also not significantly different from zero. In an alternative specification, we estimate the regression

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

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 is higher for more concentrated industries. The estimated coefficient, $\hat{\theta}^C$, is positive and large but not significant (Columns 5 and 6). Over the period from 1939 to 2002 (results not shown), the estimated effects are larger, but are also mostly insignificant. Therefore, we find inconclusive evidence regarding the prediction that the demand changes due to demographics alter profits more substantially in the presence of barriers to entry.

3.6 Age groups

Our results suggest that demographic shifts affect industry profitability. Are these effects driven by profitability shifts for the industries targeting children and the elderly? Do they mainly depend on more subtle shifts in the demand for goods for adults? We address these questions by separating industries in three broad groups, which we label Young, Adult, and

Elderly. The Young group includes all the industries under the Children grouping (Appendix Table 2), books for college, books for K-12, and bicycles. The Elderly group includes the Health grouping and the Senior grouping. The Adult group includes the other 33 industries.

In Columns 7 through 12 of Table 6 we replicate specification (8) relating profitability to contemporaneous consumption growth for each of the three groups. For the Young group of industries (Columns 7 and 8), we find a significant and large effect of demand shifts on profitability with and without industry fixed effects. The estimated coefficient $\hat{\theta}$ is comparable to the coefficient for the Demographic Industries. The standard error of $\hat{\theta}$ is also close to the one estimated for the Demographic Industries, despite the fact that the regression in Columns 7 and 8 has only one third as many observations. The R^2 in the specification without industry fixed effects is .064, substantially larger than the R^2 of .015 over the Demographic Industries. The small group of industries selling mainly to the young provides a quite precise estimation of the profitability effects.

In the groups of industries selling to Adults (Column 9 and 10) and the Elderly (Column 11 and 12), the estimated effect of demographics on profitability is positive and large (at least with industry fixed effects) but imprecisely estimated. In both cases, the effect is only marginally significant when industry indicators are included. The lower significance level relative to the Young group is not due to lower point estimates, but to standard errors that are two to three times as large. The demographic shifts for these industries are less drastic and do not allow for as precise an estimation of the effects on profitability.

3.7 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. Columns 5 through 8 of Table 4 display 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. As above, the sample is limited to the years in which the consumption data is also available. The average annual log stock return varies from 2.7 percent (bicycles) to 19.4 percent (motorcycles). The standard deviation of the yearly stock returns—30 percent on average—is negatively correlated with the number of firms in the industry. The longest series run for 65 years, and the average number of firms in an industry ranges from 1.4 (motorcycles) to 180.7 (food).

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 short-term) and $t+5$ to $t+10$ (the long-term). We beta-adjust industry returns to remove the market-wide shocks[14]. Define $r_{k,u,t}$ to be the natural log of the stock return for good k between the end of year t and the end of 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.[15] The specification of the regression is

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

Since the consumption growth variables are scaled by 5, the coefficients δ_0 and δ_1 represent the average increase in abnormal yearly returns for one percentage point of additional annualized growth in demographics. Once again, the forecasts of consumption as of time t only use information 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. 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 information enters the forecast horizon.

Columns 1 through 3 of Table 7 present the estimates of (9) for the sample of Demographic Industries during the years 1974-2003. In the specification without year or industry indicators (Column 1), the coefficient on short-term demographics, $\hat{\delta}_0$, equals -0.8 and is not significantly different from zero. The coefficient on long-term demographics, $\hat{\delta}_1$, equals 10.1 and is significantly larger than zero. A one percentage point annualized increase in demand from year 5 to year 10 increases the average abnormal yearly stock return by 10.1 percentage points. In this specification as well as in the subsequent specifications, the estimated $\hat{\rho}$ equals approximately 0.2, resulting in a correction coefficient $(1+\hat{\rho})/(1-\hat{\rho}) = 1.5$. The coefficients have the same magnitude and significance when industry fixed effects (Column 2) are introduced. The introduction of year fixed effects (Column 3) lowers the estimated $\hat{\delta}_1$ to a still large and marginally significant estimate of 6.0. In the longer sample (Columns 4 through 6), we observe a similar pattern of results, with smaller coefficients: the coefficient on short-term demographics is negative and insignificant, while the coefficient on long-term demographics is positive and marginally significant (except in Column 6). The estimated coefficients for the sample of all industries (Columns 7 through 12) are slightly smaller than the estimates for the Demographic Industries, with the same pattern of significance. While the coefficient

[14]The results are essentially the same if we use net returns instead of abnormal returns.

[15]We require a minimum of 30 observations for the estimation of β.

δ_0 on short-term demand forecast is not statistically significant, the point estimate is always negative.

Barriers to entry. As we discussed above, testing attention using stock market reaction to demand changes is more meaningful for industries with substantial barriers to entry. In the first six columns of Table 8 we replicate specification (9) separately for industries with C-4 concentration ratio above and below the median. For the industries with above-median concentration (Column 1) the coefficient $\hat{\delta}_1$ on demand growth between $t+5$ and $t+10$ is significantly positive and larger than in the overall sample of all industries; the coefficient remains large and is marginally significant with industry fixed effects (Column 2). For the industries with below-median concentration (Columns 3 and 4) the point estimates are only a third as large and there is no significant relationship between demand changes and stock returns. As an alternative specification, in columns 5 and 6 we interact the continuous measure of concentration C-4 with demand growth at the different horizons:

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

The baseline estimate (Column 5) of $\delta_1^C = 60.9$ 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 $\hat{\delta}_1 + .2\hat{\delta}_1^C = 1.5$. For an industry with a high concentration ratio of .5, the predicted responsiveness is $\hat{\delta}_1 + .5\hat{\delta}_1^C = 19.8$, a very large response. We find similarly large magnitudes in the specifications with industry fixed effects (Column 6). Over the period from 1939 to 2003 (results not shown), the estimated effects are smaller but the pattern of statistical significance is similar. The evidence suggests that return predictability is stronger in industries with higher concentration.

Age Groups. In Columns 7 through 12 of Table 8 we replicate specification (9) for each of the three main demographic sub-groups. For the Young group of industries (Columns 7 and 8), we find a significant and large effect of long-term demand shifts on stock returns with and without industry fixed effects. The estimates for $\hat{\delta}_1$, and the standard errors, are comparable to the estimates for the sample of Demographic Industries, and the R^2 of .062 is substantially higher. The small group of industries selling mainly to the young provides a quite precise estimation of the forecastability of returns.

In the Adult group of industries (Column 9 and 10) we also obtain a large effect of long-term demand shifts on stock returns. The estimated coefficient $\hat{\delta}_1$ is even higher than the estimate in the sample of Demographic Industries and is significantly different from zero. This second group of industries also contributes to the predictability findings. Finally, in the Elderly group of industries, the estimates are much more imprecise, with standard errors three times as large as in the other two groups. The slow-moving demographic shifts for these industries may not allow for a precise estimation of the effect of forecasted demand on returns.

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

$$r_{k,t+1,t} - \hat{\beta}_{k,t} r_{m,t+1,t} = \lambda + \eta_k + \varphi_{t+1} + \delta_H(\hat{c}_{k,t+h+1|t-1} - \hat{c}_{k,t+h|t-1}) + \varepsilon_{k,t+1}$$

on the sample of Demographic Industries[16] over the years 1974-2003, for investor horizon h between 0 and 13 years. The coefficient δ_H measures the extent to which consumption growth h years ahead forecasts stock returns in year $t+1$ (Figure 4). The coefficient δ_H on contemporaneous demand growth ($h = 0$ or $h = 1$) is small and insignificant. The coefficient increases with the horizon h and becomes significantly positive, reaching the peak value of 9.42 at the horizon h of 7 years. The coefficient then decreases for larger h, becoming half as large for $h = 10$ and insignificant for h past 11 years. These findings suggest that stock return predictability is highest for forecasted demand growth occurring 5 to 8 years in the future.

3.8 Portfolio returns

These results provide evidence of return predictability using long-term demand growth due to demographics. We now analyze whether rational market participants could exploit the underreaction to long-term demographic information with a trading strategy. This provides another measure of the predictability of stock returns induced by underreaction to demographics.

We follow a strategy from 1974 to 2003 for sector indices belonging to the sample of Demographic Industries. We create the zero-investment portfolio by double-sorting the group of industries at the beginning of each year, as suggested by the model. In the presence of inattention with partial extrapolation, both $E_t[\Delta c_{t+1+h}]$ and $E_t[\Delta c_{t+1+h} - \Delta c_{t+1+h-n}]$ will positively predict stock returns. Therefore, we first sort the industries into two equal groups based on long-term forecasted demand growth, $\Delta\hat{c}_{LR} \equiv \hat{c}_{t+10} - \hat{c}_{t+5}$. Next, within each of these two groups we sort the industries into two equal sub-groups based on the difference between long-term and short-term forecasted growth, that is, $\Delta\hat{c}_{LR-SR} \equiv (\hat{c}_{t+10} - \hat{c}_{t+5}) - (\hat{c}_{t+5} - \hat{c}_t)$. The zero-investment portfolio is long in industries with high predicted long-term growth $\Delta\hat{c}_{LR}$ and high $\Delta\hat{c}_{LR-SR}$, and is short in industries with low predicted long-term growth $\Delta\hat{c}_{LR}$ and low $\Delta\hat{c}_{LR-SR}$. The portfolio is designed to exploit both inattention to long-term information—measured by $\Delta\hat{c}_{LR}$— and extrapolation—measured by $\Delta\hat{c}_{LR-SR}$.

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 one-month Treasury rate. The standard errors are corrected for heteroskedasticity and autocorrelation using the Newey-West estimator with 6 lags[17]. The results in Col-

[16] The results are similar if all industries are included in the analysis.

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

umn 1 of Table 9 indicate that the portfolio earns a monthly abnormal return of .71 percent.[18] The outperformance remains essentially the same if we also include the size and the book-to-market factors (Column 2), as well as the momentum factor (Column 3). These magnitudes are consistent with the estimates from the predictability regressions in Table 7. The annualized abnormal return for the portfolio (8.5%) is only slightly lower than the product of $\hat{\delta}_1$ (10.1) from Table 7 (Column 1) and the average difference between forecasted demand growth from $t + 5$ to $t + 10$ for the long and short constituent portfolios (1 percentage point).

In Columns 4 through 6 we report the abnormal performance of the investment strategy over the longer time period 1939-2003. For this sample the portfolio has an average abnormal annualized return of about 5% per year. This outperformance is significant with a 1-factor or a 3-factor model (Columns 4 and 5), and is marginally significant with a 4-factor model (Column 5). The lower abnormal returns over this longer time period are consistent with the OLS findings in Table 7. During the early years of this sample period the portfolio is formed using a substantially smaller set of industries, and each industry contains fewer firms.

In Columns 7 through 12 we report the performance of the long and the short portfolio underlying the zero-investment portfolio of Columns 1 through 3. In general, the outperformance of the zero-investment portfolio depends more heavily on the long portfolio.

In Table 10 we present the results for a similar zero-investment portfolio constructed using all 48 industries over the years 1974-2003. This portfolio earns average annual abnormal returns of about 3 percentage points (Columns 1 through 3). Unlike the other estimates, the outperformance is only marginally significant after controlling for the 3-factor risk-adjustment procedure and insignificant after controlling for the 4 factors. The weaker performance of the portfolio strategy in this sample is roughly consistent with the OLS results in Table 8. The difference between average forecasted consumption growth for the industries in the long portfolio and the industries in the short portfolio is only 0.5 percentage points.

In Columns 4 through 9 of Table 10 we split the overall sample into above-median and below-median concentration industries. The average abnormal return for the high-concentration sample is over 7 percent per year and is statistically significant. The portfolio return for the low-concentration sample, instead, is approximately 1 percent per year and is insignificant. Abnormal returns are more sensitive to forecasted demand growth for more concentrated industries, a finding consistent with the OLS results (Table 8).

We also examine whether the outperformance depends uniformly on large and small companies within an industry. In Columns 10 through 12 we replicate the portfolio results of Columns 4 through 6, except that here the industry returns refer to the returns for the largest company in the industry. The levels of outperformance are similar to the those found in Columns 4 through 6, suggesting that a portfolio strategy can be successfully implemented even for stocks with relatively low trading costs.

[18]The average monthly return (without a market control) is .66 percent (s.e. .26).

The average abnormal returns from trading on demographic information, therefore, are large and statistically significant. The estimates from the predictability regressions and the abnormal returns for the trading strategy are broadly consistent with one another.

4 Interpretation of the results

4.1 Attention interpretation

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, demographic 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.

Prediction 2 in Section 2 offers a 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 7 and 8 suggest that the horizon h could be between 5 and 10 years. Figure 4 shows that stock return predictability is highest using predicted consumption growth between 5 and 8 years ahead. 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 7.5 years.

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 (Table 7). The estimate $\hat{\delta}_1 \approx 10$, is approximately 5 times larger than $\hat{\theta} \approx 2$, the estimate for the responsiveness of accounting return on equity to forecasted demand growth These magnitudes are not 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. For example, set the annual discount factor ρ equal to 0.96, the extrapolation weight w equal to 0.25, and the number of periods of extrapolation n equal to 4. For these parameters the model of inattention with partial extrapolation implies $\delta_1 = \theta\rho^h \left[1 + (1-w)\,\rho/\left((1-\rho)\,n\right)\right] \approx 5\theta$ 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 negative (although statistically insignificant) point estimate for δ_0 in Table 7 is also consistent with partial extrapolation according to Prediction 3.

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.

Our interpretation of the overall evidence is that investors do not pay attention to information beyond a horizon of approximately 5 years. This estimated horizon for investors is consistent with the observed horizon of analyst forecasts. In Table 11 we use the I/B/E/S data to document the availability 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 have earnings forecasts for the next two years. The number of forecasts further in the future, however, decays 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 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.

According to I/B/E/S, analysts do not produce forecasts of annual earnings beyond a 5 year horizon. While such long-term forecasts may be available in privately-held data sources, investors are unlikely to possess readily available information regarding profitability in the distant future. Given this evidence, it is not surprising that investors tend to ignore outcomes more than 5 years in the future.

Ignoring information about the distant future, after all, is a reasonable rule of thumb 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 neglecting the long-term future is approximately correct. While such a rule may work well in general, its implementation is costly when applied to demographic information. Long-term demographic variables can be precisely estimated and may differ significantly from their short-term pattern.

4.2 Alternative interpretations

Rational predictability. 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).

Poor estimation of systematic risk. For the specifications in Tables 7 and 8 the industry beta is estimated using the previous 48 months of industry returns. If the actual beta increases for industries with high demand growth rates 5 to 10 years in the future, then the estimated beta understates the actual systematic risk. This estimation problem could explain the observed outperformance. To test for this, we regress annual changes in estimated beta, $\hat{\beta}_{k,t+1} - \hat{\beta}_{k,t}$, on forecasted short-term and long-term demand growth. We find no evidence of a relationship between changes in estimated beta and long-term demand growth.[19]

Persistent regressors. The predictability results in Tables 7 and 8 could suffer from bias from persistent regressors. Following Stambaugh (1999), assume that the demand growth due to demographics, denoted x, follows an AR(1) process, $x_t = \theta + \rho x_{t-1} + v_t$, with $|\chi| < 1$. Denote by σ_v^2 the variance of v and denote by $\sigma_{\varepsilon v}$ the covariance between v_t and ε_t, the error term in (9). In this case, Stambaugh shows that the bias in the estimate of $\hat{\delta}_1$ is equal to $E\left(\hat{\delta}_1 - \delta_1\right) = \left(\sigma_{\varepsilon v}/\sigma_v^2\right) E\left(\hat{\rho} - \rho\right)$.

To evaluate the seriousness of this problem, we estimate $\hat{\rho}$ and $\hat{v}_{k,t}$ by a panel regression of the 5- to 10-year growth rate due to demographics $x_{k,t}$ on its lagged value $x_{k,t-1}$. We include industry fixed effects and assume that $\rho_k = \rho$ for each industry k. We obtain a point estimate for $\hat{\rho}$ of .9546, with a standard error of .0102. We use this to generate the series for $\hat{v}_{k,t}$. We then regress the estimated errors $\hat{\varepsilon}_{k,t}$ from the return regression (including industry indicators) on the series $\hat{v}$, again including industry fixed effects. We obtain an estimate for $\sigma_{\varepsilon v}/\sigma_v^2$ of -4.7539, with standard error 4.3368. First, this estimate is not statistically different from zero and, consequently, we fail to reject the null hypothesis of no bias. Second, since the bias $E\left(\hat{\rho} - \rho\right)$ would be negative and bounded below by $\left(\hat{\rho} - 1\right)$, the point estimate for the bias $E\left(\hat{\delta}_1 - \delta_1\right)$ is approximately $(-.04) * (-4.7) = .188$, a small correction relative to the 10.1 estimate for $\hat{\delta}_1$. The persistence of regressors does not appear to be a main concern in our setting.

Generated regressors. In the predictability regressions, the forecasted demand growth rates are estimates created from demographic and consumption data. In general, the standard errors should be corrected for the uncertainty in these preliminary estimates. However, Pagan (1984) shows that the standard errors do not require adjustment under the null hypothesis that the generated regressors do not have any predictive power—the null hypothesis evaluated in the paper.

[19] These results are available from the authors upon request.

Asset manager horizon. Money managers are usually evaluated based on short-term performance. These managers may not be able to expose themselves to risk for a long enough period to reap the returns from trading on long-term information. However, the trading strategy on demographics has substantial abnormal returns even at an annual frequency. These returns should be relevant even for professionals with relatively short investment horizons.

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 may disregard variables that display little daily variation, like demographics. Studies on just-noticeable differences (Weber, 1834) suggest a minimum size of a stimulus necessary for detection, let alone to attract attention. Demographic information may therefore be neglected until the information is incorporated in earnings announcements, which are discrete events. This hypothesis could explain the stock return forecastability, but not its horizon. This story suggests that short-horizon, rather than long-horizon, demographic information should predict stock returns.

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 childrens 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 48 different expenditure categories over 65 years. We match the expenditure categories to industry-level accounting measures and stock market returns. The forecasted demand growth due to demographics predicts the contemporaneous industry-level 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 short-term growth rates do not have significant forecasting power. This predictability result is more pronounced for those groups of industries that exhibit a stronger relationship between profitability and forecastable demand growth.

The evidence supports the hypothesis that investors are inattentive at longer horizons. 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 horizon.

We have identified a novel form of predictability in financial markets based on long-term demographic information. The evidence in this paper complements the existing results on the response of stock returns to short-term events, such as earnings surprises. Our findings have 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 in 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, $B_{t+1} = B_t + X_{t+1} - D_{t+1}$. Earnings that are not paid to shareholders as dividends increase book equity. We define log stock return, r_{t+1}, and log accounting return on equity, roe_{t+1}, as

$$r_{t+1} \equiv \log\left[\left(M_{t+1} + D_{t+1}\right)/M_t\right], \tag{10}$$

$$roe_{t+1} \equiv \log\left[\left(B_t + X_{t+1}\right)/B_t\right] = \log\left[\left(B_{t+1} + D_{t+1}\right)/B_t\right]. \tag{11}$$

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 $\rho = [1 + \exp(\overline{d-m})]^{-1}$ 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

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

The second equality follows from taking expectations with respect to operator $\widehat{E}$ and noting $\widehat{E}_t\left(m_t - b_t\right) = m_t - b_t$. Substituting the right hand-side of (13) into (12) leads to (3):

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

B Appendix B. Data

B.1 Demographic forecasts

Cohort size adjustment. The cohort size data is from 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 the group of age 85+ 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 $A_{g,j,t} = (\sum_{j=85}^{99} A_{g,j,t} / \sum_{j=85}^{99} A_{g,j,1980}) * A_{g,j,1980}$. This imputation[20] 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 issue is unlikely to matter.

Mortality rate adjustment. We use the 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 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 $z\left(j\right)$ 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.* 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 present population profile with the forecasts of mortality and fertility, we produce a preliminary forecast of the future 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 age $j \geq 1$ 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 (age 0) is given by $\hat{A}^p_{g,0,u|t} = sr_g * \sum_{j=14}^{49} \hat{A}^p_{f,j,u-1|t} * \hat{b}_{j,u-1|t}$, where $sr_m = 0.501$ is the average probability that a 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 five-year period prior to year t.[21] 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 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 $0 \leq$

[20]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.

[21]For the age group 0-9, we allow for a separate adjustment coefficient for age 1, and we do not adjust the forecast for the unborn (age 0).

$i \leq 5$ 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.[22] 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.

B.2 Consumption data

Expenditure categories. The dependent variable in the regressions in Section 3.2 is the yearly expenditure, $c_{i,k,t}$, on each category k listed in Appendix Table 2. In particular, the automobile and motorcycle categories include expenditures on both new and used vehicles. The coal category includes expenditure on both coal and electricity. The health care and medical equipment categories are estimated using total expenditure on health, including health insurance. The health insurance category, instead, is limited to health insurance expenditure. The residential mortgage category is estimated using expenditure on mortgage interest. The utilities category includes expenditure on electricity, water, and natural gas.

Housing. The residential development category is estimated using the housing value. For some of the observations, the information on housing value is not available for renters. In this case, we compute an implicit conversion rate from yearly rent to housing value for the sample for which both measures are available, and apply it to the yearly rent value. The conversion rate from yearly rent to housing value equals 1/.028 in 1936-37, 1/.088 in 1972-73, and 1/.076 in 1983-84. Since the conversion rate for 1960-61 cannot be computed, we use the rate for 1972-73. Table 2 reports the annualized rental value. The expenditures for residential construction and for construction equipment, which depend on changes in the housing stock, rather than on levels, is computed differently. First, we compute the forecasted housing value $\hat{C}_{\text{housing},u|t}$ for year u, given information of year t. Then we compute the forecasted demand for residential construction and construction equipment as $\hat{C}_{\text{housing},u|t} - \hat{C}_{\text{housing},u-1|t} + .1\hat{C}_{\text{housing},u-1|t}$, that is, the change in the forecasted housing stock plus housing depreciation.

Other issues. The value of income and housing in the 1960-61 survey is reported in discrete categories. We assign it the mean value in the interval, and 1.5 times the value for the top category. Housing value is top-coded in the 1983-84 survey. We use the 1972-73 category to compute the appropriate adjustment coefficient of 1.53. Finally, in the 1983-84 survey some households are interviewed for fewer than 4 quarters. We compute an annualized consumption value for these records.

B.3 Industry classification

The industry classification system is designed to satisfy three basic criteria: (i) aggregate goods with a relatively homogeneous age profile of consumption; (ii) define categories that are consistent over time; (iii) minimize deviations from the Standard Industrial Classification (SIC). These criteria lead to 48 industries (Appendix Table 2) belonging to three groups.

[22] The forecasts for the unborn are obtained by applying the adjustment coefficient to the mothers, computing the forecasted number of births, and aging the cohort.

Standard industries—such as oil, telephone, and health insurance—are constructed from a list of 4-digit SIC codes. For example, the health insurance industry is defined by the SIC codes 6320-6329. A company belongs to industry k in year t if its SIC code for year t coincides with one of the listed codes for industry k. In Appendix Table 2 these industries are characterized by the absence of codes in parentheses. The classification for these industries closely resembles the Fama-French classification.

Searched industries—such as toys, cruises, and furniture—are also constructed on the basis of a list of 4-digit SIC codes. In addition, we eliminate the companies in these SIC codes whose core business does not belong in the industry (from our standpoint). For example, we eliminate golf equipment manufacturers and retailers from the toys industry. Further, we define a list of additional SIC codes and identify companies in these codes that belong to the industry. The searched industries are identifiable in Appendix Table 2 by the presence of SIC codes without parentheses (the basic codes) and with parentheses (the additional codes).

Reclassified industries—the book industry subcategories, as well as golf, motorcycles, and bicycles—are not uniquely associated with any SIC codes. Companies in these industries are identified from within a list of SIC codes. For example, in order to construct the four book categories, we search the SIC codes 2730-2739 and determine the companies whose core business consists of books for children, books for K-12, etc. In Appendix Table 2 these expenditure categories only have SIC codes in parentheses.

C Appendix C. Standard errors

Define $\Gamma_q = E\left[\left(\sum_{k=1}^{K} X_{kt}\varepsilon_{kt}\right)'\left(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q}\right)\right]$ and assume $X'_{kt}\varepsilon_{kt} = \rho X'_{kt-1}\varepsilon_{kt-1} + \eta'_{kt}$, where $\rho < 1$ is a scalar and $E\left[\left(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q}\right)'\left(\sum_{k=1}^{K} \eta_{kt}\right)\right] = 0 \ \forall \ q > 0$. Then,

$$\begin{aligned}
\Gamma_q &= E\left[\left(\sum_{k=1}^{K}\left(\rho^q X_{kt-q}\varepsilon_{kt-q} + \sum_{j=1}^{q}\rho^{q-j}\eta_{kt-q+j}\right)\right)'\left(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q}\right)\right] \\
&= \rho^q E\left[\left(\sum_{k=1}^{K}\rho^q X_{kt-q}\varepsilon_{kt-q}\right)'\left(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q}\right)\right] + \\
& \quad E\left[\left(\sum_{j=1}^{q}\rho^{q-j}\sum_{k=1}^{K}\eta_{kt-q+j}\right)'\left(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q}\right)\right] \\
&= \rho^q E\left[\left(\sum_{k=1}^{K}\rho^q X_{kt-q}\varepsilon_{kt-q}\right)'\left(\sum_{k=1}^{K} X_{kt-q}\varepsilon_{kt-q}\right)\right] = \rho^q\Gamma_0.
\end{aligned}$$

Using the relationship for Γ_q, we obtain

$$S = \Gamma_0 + 2\sum_{q=1}^{\infty}\rho^q\Gamma_0 = \left(\left(2\sum_{q=0}^{\infty}\rho^q\right) - 1\right)\Gamma_0 = \left(\frac{2}{1-\rho} - \frac{1-\rho}{1-\rho}\right)\Gamma_0 = \left(\frac{1+\rho}{1-\rho}\right)\Gamma_0.$$

References

[1] Acemoglu, Daron and Joshua Linn. "Market Size in Innovation: Theory and Evidence From the Pharmaceutical Industry", *Quarterly Journal of Economics*, Vol. 199, 1049–1090, 2004.

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

[3] Barber, Brad M. and Terrance Odean. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors", mimeo, 2002.

[4] Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. "A model of investor sentiment", *Journal of Financial Economics,* Vol. 49, 307-343, 1998.

[5] 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.

[6] 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.

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

[8] 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, 1988.

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

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

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

[12] DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, Robert J. Waldmann. "Noise Trader Risk in Financial Markets", *Journal of Political Economy,* Vol. 98, 703-738, 1990.

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

[14] Gabaix, Xavier, David Laibson, Guillermo Moloche and Stephen Weinberg."Information Acquisition: Experimental Analysis of a Boundedly Rational Model", mimeo, 2004.

[15] 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, 2002.

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

[17] Hirshleifer, David, Sonya S. Lim, and Siew H. Teoh. "Disclosure to an Audience with Limited Attention", mimeo, 2004.

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

[19] Hong, Harrison, Walter Torous, and Ross Valkanov. "Do Industries Lead the Stock Market?", mimeo, 2003.

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

[21] 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.

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

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

[24] Mankiw, N. Gregory and Michael D. Whinston. "Free Entry and Social Inefficiency", *RAND Journal of Economics,* Vol. 17, 48-58, 1986.

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

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

[27] Newey, Whitney K. and Kenneth D. West. "Hypothesis Testing with Efficient Method of Moments Estimation", *International Economic Review,* Vol. 28, 777-787, 1987.

[28] Pagan, Adrian R., "Econometric Issues in the Analysis of Regressions with Generated Regressors", *International Economic Review,* Vol. 25, 221-247, 1984.

[29] Peng, Lin and Wei Xiong. "Limited Attention and Asset Prices", mimeo, 2004.

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

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

[32] Stambaugh, Robert F. "Predictive regressions", *Journal of Financial Economics,* Vol. 54, 375-421, 1999.

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

[34] 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.

[35] 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.

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

[37] 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.

[38] 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. Inter-university Consortium for Political and Social Research, 1999.

[39] Vuolteenaho, Tuomo. "What Drives Firm-Level Stock Returns", *Journal of Finance,* Vol. 57, 233-264, 2002.

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

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

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

Figure 1a. Forecasted and Actual Population Ages 30-34

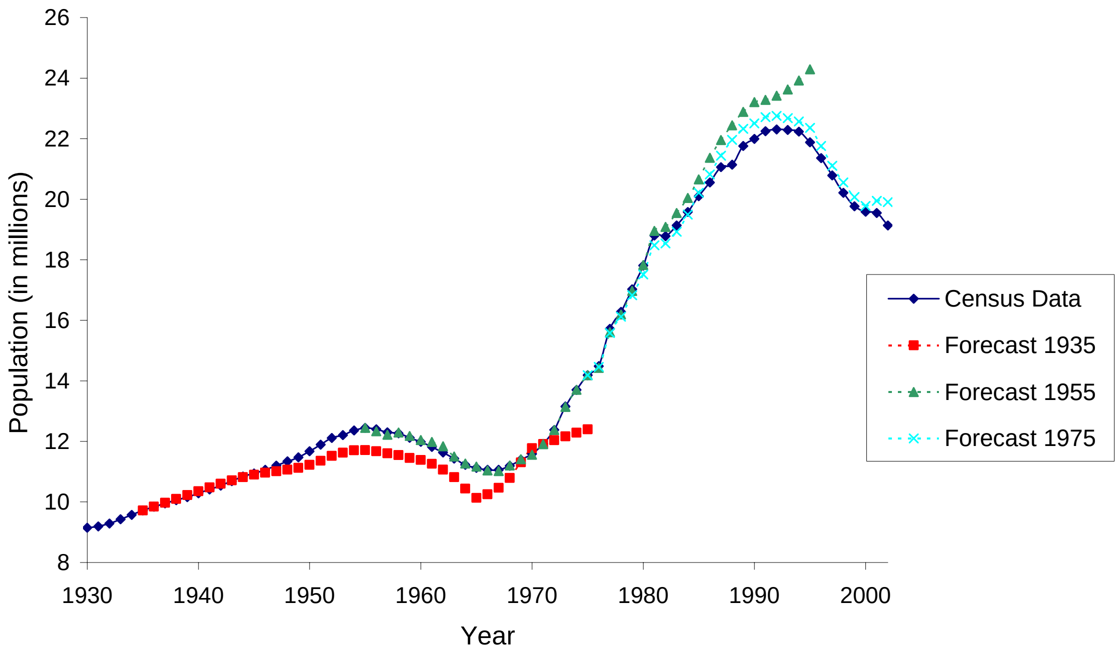
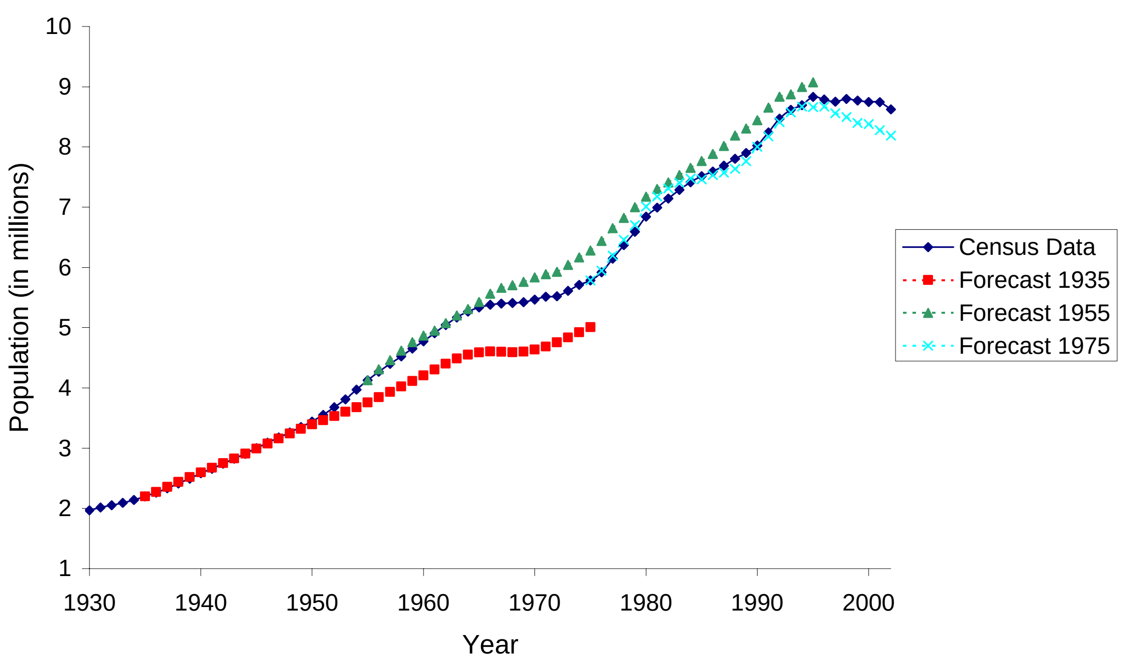


Figure 1b. Forecasted and Actual Population Ages 70-74



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

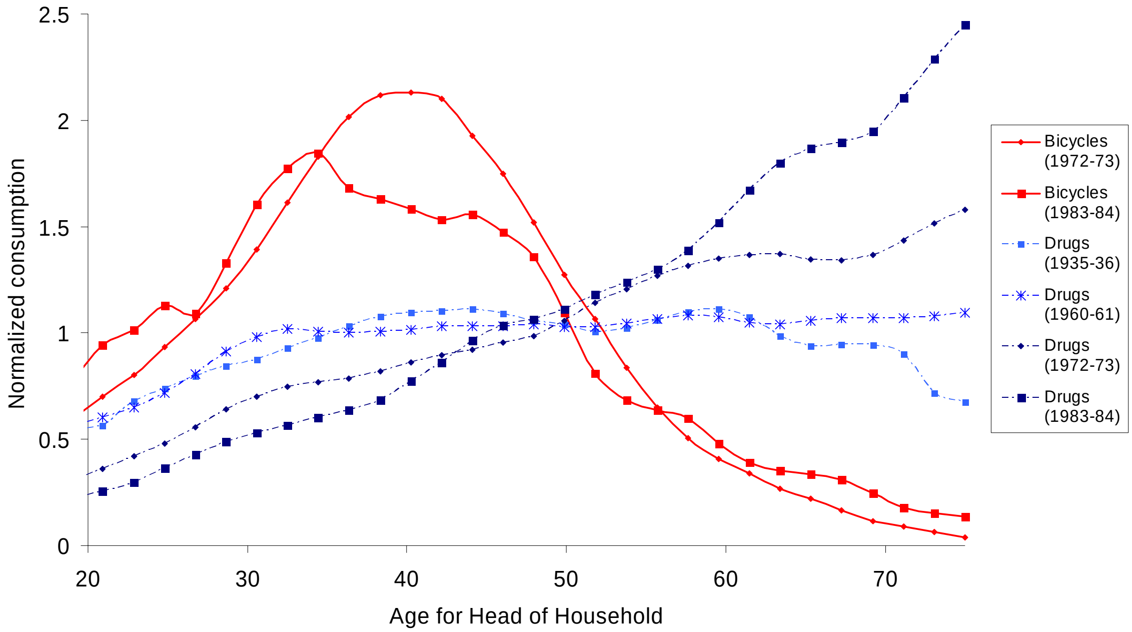
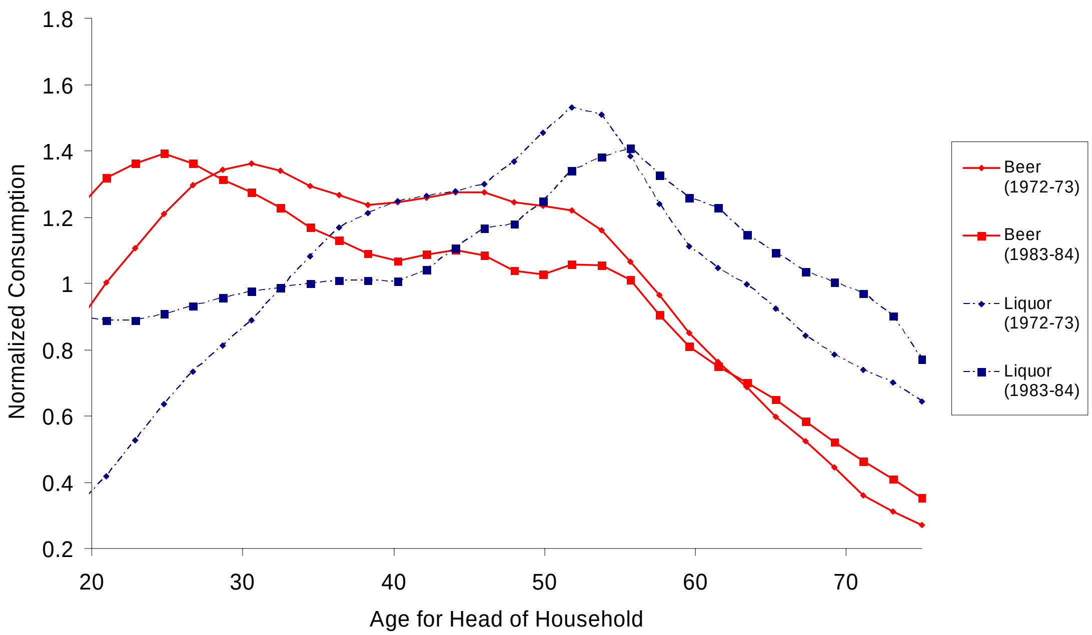
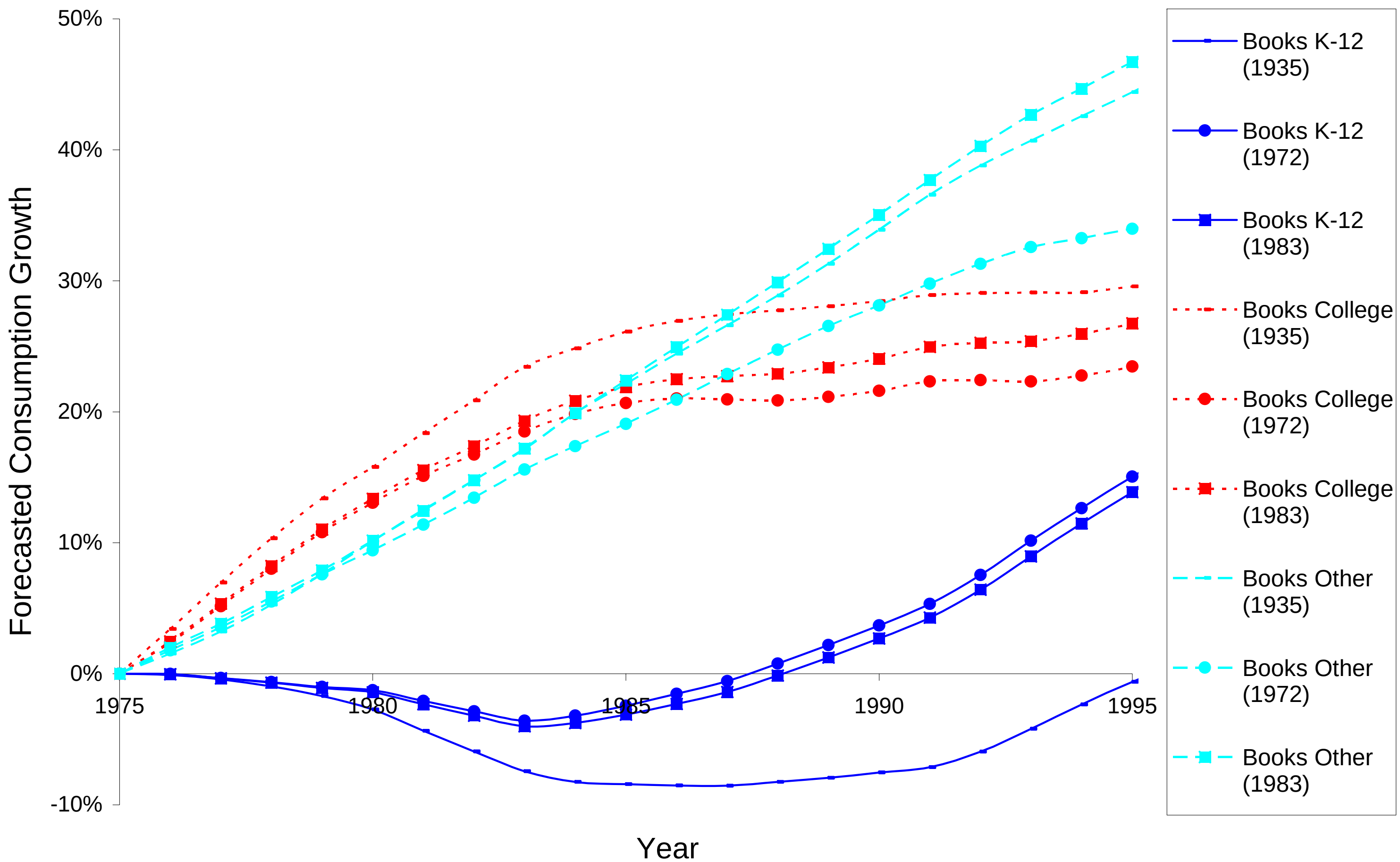


Figure 2b. Age Profile of Consumption for Beer and Liquor



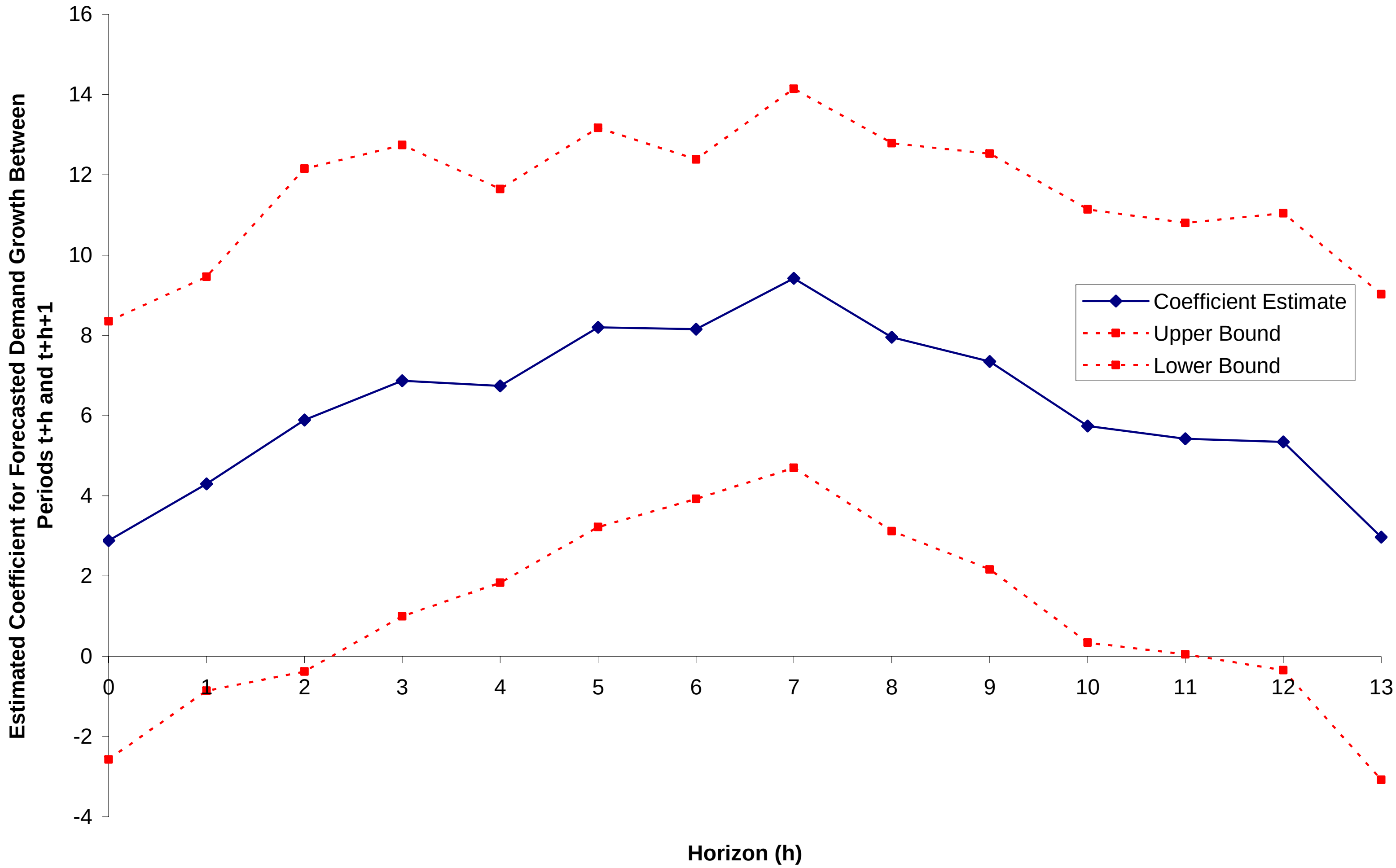
Notes: Figures 2a and 2b display kernel regressions of normalized household consumption for each good as a function of the age for the head of the household. The regressions use an Epanechnikov kernel and a bandwidth of 5 years. Each different line for a specific good uses an age-consumption profile 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, 1972-73, and 1983-84 consumption surveys. Each distinct line for a good uses an age-consumption profile from a different data set. Forecasts for book expenditure in 1960 are 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 a univariate regression of abnormal returns at t+1 on forecasted consumption growth between t+h and t+h+1. The confidence interval is constructed using robust standard errors.

Table 1. Predictability of Population Growth Rates By Cohort

Dependent Variable:	Actual population growth for each cohort						Census projection of population growth	
	0 to 5 years ahead			5 to 10 years ahead			0 to 5 yrs	5 to 10 yrs
	Ages 0-99 (1)	Ages 0-18 (2)	Ages 65+ (3)	Ages 0-99 (4)	Ages 0-18 (5)	Ages 65+ (6)	Ages 0-99 (7)	Ages 0-99 (8)
Constant	0.0060	-0.0079	0.0283	0.0125	0.0044	0.0397	-0.0286	-0.0270
	(0.0005)***	(0.0012)***	(0.0019)***	(0.0007)***	(0.0022)**	(0.0020)***	(0.0049)***	(0.0043)***
Forecasted population growth: 0 to 5 yrs	0.9024	0.9121	0.7574				1.1705	
	(0.0039)***	(0.0088)***	(0.0124)***				(0.0354)***	
Forecasted population growth: 5 to 10 yrs				0.8413	0.7097	0.6798		1.1155
				(0.0056)***	(0.0177)***	(0.0143)***		(0.0320)***
R^2	0.8312	0.8188	0.5593	0.6928	0.4250	0.4665	0.8466	0.8601
N	*N* = 11000	*N* = 2356	*N* = 2940	*N* = 10000	*N* = 2166	*N* = 2590	*N* = 200	*N* = 200

Notes: Reported coefficents from the regression of actual population growth rates on our forecasted growth rates in Columns (1) through (6). In Columns (7) through (9) we report coefficients from the regression of Census projections of population growth rate as of 2000 on our forecasted growth rates. In Columns (1) through (3) and in column (7) the growth rates refer to the next 5 years. In Columns (4) through (6) and in column (8) the growth rates refer to the period between 5 and 10 years ahead. The regression specification is $y_{it} = a + bx_{it} + e_{it}$ 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. Age is defined by one year cells. 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 number of newborns is calculated by applying birth rates from the previous year to the forecasted age profile of the female population. The Census projection of population growth rate is calculated using data from the Census website. 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: Expenditure by Good

Consumer Survey	1935-36		1960-61		1972-73		1983-84	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Yearly Expenditure	**Mean (in $)**	**Std. Dev. (in $)**	**Mean (in $)**	**Std. Dev. (in $)**	**Mean (in $)**	**Std. Dev. (in $)**	**Mean (in $)**	**Std. Dev. (in $)**
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	77.52	(56.61)	147.71	(161.14)	36.73	(49.14)	55.98	(62.84)
Magazines	23.80	(39.58)	(.)	(.)	16.44	(42.31)	31.29	(58.25)
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	48.65	(145.31)	298.47	(317.23)	467.57	(521.43)	284.22	(494.05)
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*	1796.14	(1743.86)	3218.38	(2551.48)	4083.81	(3574.06)	4304.69	(4068.31)
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	(150.7)	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*	1796.14	(1743.86)	3218.38	(2551.48)	4083.81	(3574.06)	4304.69	(4068.31)
Residential Development*	1796.14	(1743.86)	3218.38	(2551.48)	4083.81	(3574.06)	4304.69	(4068.31)
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	(296.6)	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" refers to the imputed annual rent estimated from the value of the dwelling of residence. See Appendix B for details on the construction of the housing expenditure.

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

Table 3: Summary Statistics For Predicted Demand Growth Rates

Expenditure Category	No. Years	Forecasted 0-5 Growth	Demogr. Industry	Forecasted 0-5 Growth	Demogr. Industry	Forecasted 0-5 Growth	Demogr. Industry	% Dem. Industry
		1950		1975		2000		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Child Care	65	0.0268	Yes	0.0001	Yes	-0.0035	Yes	100%
Children's Books	28	.	.	.	.	0.0036	Yes	40%
Children's Clothing	65	0.0157	Yes	0.0226	Yes	0.0087	No	97%
Toys	65	0.0270	Yes	0.0044	Yes	0.0051	No	89%
Books -- college text books	65	-0.0083	Yes	0.0270	Yes	0.0133	Yes	100%
Books -- general	65	0.0064	No	0.0205	Yes	0.0077	No	88%
Books -- K-12 school books	65	0.0241	Yes	-0.0087	Yes	0.0075	Yes	100%
Movies	65	-0.0006	Yes	0.0232	Yes	0.0093	No	49%
Newspapers	65	0.0077	No	0.0174	No	0.0119	No	12%
Magazines	65	0.0042	No	0.0206	Yes	0.0097	No	35%
Cruises	28	.	.	.	.	0.0118	No	12%
Dental Equipment	65	0.0046	No	0.0138	No	0.0111	No	20%
Drugs	65	0.0111	No	0.0167	No	0.0137	No	11%
Health Care (Services)**	65	0.0108	No	0.0173	No	0.0114	No	20%
Health Insurance	65	0.0053	No	0.0168	No	0.0125	Yes	11%
Medical Equipment**	65	0.0108	No	0.0173	No	0.0114	No	17%
Funeral Homes and Cemet.	53	0.0243	Yes	.	No	0.0152	Yes	37%
Nursing Home Care	65	0.0104	No	0.0198	Yes	0.0107	Yes	83%
Construction Equipment*	65	0.0060	Yes	0.0200	Yes	0.0092	Yes	98%
Floors	65	0.0065	No	0.0177	No	0.0118	Yes	51%
Furniture	65	0.0007	Yes	0.0201	Yes	0.0077	No	71%
Home Appliances Big	65	0.0043	Yes	0.0169	No	0.0091	No	37%
Home Appliances Small	65	0.0050	No	0.0153	No	0.0108	No	22%
Housewares	65	0.0061	No	0.0192	Yes	0.0115	Yes	31%
Linens	65	0.0082	No	0.0170	No	0.0107	No	31%
Residential Construction*	65	0.0060	Yes	0.0200	Yes	0.0092	Yes	100%
Residential Development*	65	0.0088	No	0.0168	No	0.0107	No	12%
Residential Mortgage	65	0.0146	No	0.0164	Yes	0.0036	No	52%
Beer (and Wine)	65	0.0035	Yes	0.0209	No	0.0081	No	68%
Cigarettes	65	0.0009	Yes	0.0178	No	0.0108	No	43%
Cigars and Other Tobacco	65	0.0104	No	0.0141	No	0.0140	Yes	6%
Food	65	0.0090	No	0.0145	No	0.0104	No	0%
Liquor	28	.	.	.	No	0.0120	No	3%
Clothing (Adults)	65	0.0031	Yes	0.0197	Yes	0.0106	Yes	51%
Cosmetics	65	0.0018	Yes	0.0222	Yes	0.0129	No	34%
Golf	65	0.0014	Yes	0.0217	Yes	0.0123	Yes	68%
Jewelry	65	0.0031	Yes	0.0189	Yes	0.0110	Yes	31%
Sporting Equipment	65	0.0031	Yes	0.0183	No	0.0069	Yes	45%
Life Insurance	65	0.0081	No	0.0140	No	0.0129	Yes	37%
Property Insurance	65	0.0081	No	0.0177	No	0.0110	No	8%
Airplanes	28	.	.	.	.	0.0118	Yes	3%
Automobiles	65	0.0032	Yes	0.0199	Yes	0.0086	No	31%
Bicycles	65	0.0193	Yes	0.0027	Yes	0.0010	Yes	88%
Motorcycles	28	.	.	.	.	0.0090	Yes	40%
Coal	65	0.0097	No	0.0149	No	0.0112	No	3%
Oil	65	0.0062	No	0.0161	No	0.0105	No	0%
Telephone	65	0.0075	No	0.0185	No	0.0104	No	11%
Utilities	65	0.0084	No	0.0149	No	0.0114	No	6%

Notes: Complete list of expenditure categories, with number of years of availability of data (Column 1) and average predicted five-year demand growth rate due to demographic changes in 1950 (Column 2), in 1975 (Column 4), and in 2000 (Column 6). Table also indicates whether the industry belongs to the subsample of *Demographic Industries* in 1950 (Column 3), in 1975 (Column 5), and in 2000 (Column 7). Each year the subset *Demographic Industries* includes the 20 industries with the highest standard deviation of forecasted annual consumption growth over the next 15 years. Column 8 presents percentage of years in which expenditure category belongs to the subsample of "Demographic Industries".

Table 4. Summary Statistics: Compustat Data, CRSP Data and Concentration Ratios

	Log Yearly Return on Equity				Value Weighted Annual Log Stock Return				Conc. Ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Industry Category	**Mean**	**Std. Dev.**	**# Years**	**# Firms**	**Mean**	**Std. Dev.**	**# Years**	**# Firms**	**Largest 4 Firms**
Child Care	0.116	(0.123)	29	2.76	0.104	(0.422)	30	3.47	(.)
Children's Books	0.077	(0.093)	20	2.05	0.066	(0.286)	22	2.27	0.202
Children's Clothing	0.160	(0.091)	40	2.08	0.076	(0.342)	42	2.93	0.170
Toys	0.110	(0.076)	39	9.74	0.075	(0.438)	42	12.10	0.337
Books: college texts	0.196	(0.060)	24	2.00	0.146	(0.291)	42	1.98	0.202
Books: general	0.126	(0.054)	40	7.23	0.115	(0.246)	42	8.45	0.202
Books: K-12 texts	0.139	(0.045)	36	2.22	0.116	(0.276)	39	2.77	0.202
Movies	0.073	(0.113)	52	17.81	0.114	(0.304)	65	22.51	(.)
Newspapers	0.178	(0.081)	50	10.44	0.137	(0.257)	65	10.38	0.197
Magazines	0.097	(0.068)	40	6.25	0.127	(0.291)	42	7.81	(.)
Cruises	0.194	(0.077)	16	3.63	0.176	(0.309)	18	3.78	(.)
Dental Equipment	0.091	(0.125)	41	3.05	0.064	(0.356)	65	3.21	0.350
Drugs	0.184	(0.021)	52	84.75	0.127	(0.190)	65	97.60	0.282
Health Care (Services)	0.115	(0.063)	34	42.06	0.115	(0.337)	36	55.67	(.)
Health Insurance	0.099	(0.043)	31	11.45	0.096	(0.220)	42	14.00	(.)
Medical Equipment	0.141	(0.030)	52	55.79	0.149	(0.225)	65	61.89	0.374
Funeral Homes, Cemet.	0.068	(0.109)	28	2.75	0.118	(0.511)	30	2.93	0.250
Nursing Home Care	0.071	(0.091)	33	13.54	0.046	(0.433)	35	17.11	(.)
Construction Equip.	0.124	(0.092)	40	21.58	0.119	(0.242)	42	23.90	0.430
Floors	0.082	(0.040)	46	5.17	0.081	(0.356)	65	6.17	0.400
Furniture	0.099	(0.029)	52	15.69	0.093	(0.260)	65	15.38	0.166
Home Appliances Big	0.147	(0.070)	52	19.58	0.115	(0.305)	65	20.97	0.632
Home Appliances Small	0.153	(0.050)	52	4.73	0.136	(0.253)	55	5.49	0.430
Housewares	0.099	(0.075)	38	2.97	0.091	(0.313)	42	3.21	0.575
Linens	0.100	(0.107)	37	3.97	0.101	(0.544)	39	4.51	0.263
Residential Const.	0.079	(0.094)	39	11.87	0.075	(0.460)	42	12.71	(.)
Residential Develop.	0.066	(0.049)	40	41.68	0.071	(0.311)	42	51.57	(.)
Residential Mortgage	0.137	(0.145)	37	11.59	0.092	(0.382)	42	14.55	(.)
Beer (and Wine)	0.122	(0.040)	52	7.04	0.111	(0.227)	65	8.69	0.519
Cigarettes	0.169	(0.045)	52	4.04	0.128	(0.216)	65	5.12	0.930
Cigars, Other Tobacco	0.194	(0.148)	52	4.75	0.127	(0.214)	65	5.92	0.656
Food	0.132	(0.023)	52	167.19	0.114	(0.163)	65	180.75	0.360
Liquor	0.126	(0.111)	27	4.41	0.145	(0.147)	28	5.32	0.470
Clothing (Adults)	0.128	(0.034)	52	44.00	0.103	(0.263)	65	48.11	0.158
Cosmetics	0.221	(0.112)	47	9.49	0.110	(0.299)	65	9.43	0.380
Golf	0.037	(0.115)	30	4.27	0.051	(0.401)	31	5.58	(.)
Jewelry	0.087	(0.051)	40	9.45	0.116	(0.349)	42	11.21	0.203
Sporting Equipment	0.120	(0.104)	52	6.38	0.083	(0.383)	65	6.91	0.280
Life Insurance	0.096	(0.072)	39	13.23	0.120	(0.273)	41	34.49	(.)
Property Insurance	0.112	(0.054)	30	27.40	0.096	(0.189)	65	23.46	(.)
Airplanes	0.097	(0.068)	27	41.59	0.124	(0.212)	28	48.82	0.621
Automobiles	0.126	(0.084)	52	57.08	0.108	(0.235)	65	66.20	0.807
Bicycles	0.070	(0.118)	35	1.40	0.027	(0.421)	37	1.49	0.650
Motorcycles	0.258	(0.115)	18	1.22	0.194	(0.364)	22	1.45	0.650
Coal	0.069	(0.103)	45	6.87	0.112	(0.248)	65	9.91	(.)
Oil	0.111	(0.038)	52	156.87	0.117	(0.175)	65	172.25	0.300
Telephone	0.078	(0.050)	52	18.46	0.086	(0.240)	65	25.80	(.)
Electricity	0.102	(0.033)	44	161.34	0.097	(0.171)	65	146.15	(.)

Notes: The measure of 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 industry measure of ROE is the average of ROE weighted by the book value of equity in year t. Column 1 displays the log of 1 plus the industry ROE. 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 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 5). The average is value-weighted by the market capitalization at the end of year *t*. Columns 6 through 8 are parallel to Columns 2 through 4.

The Concentration Ratio measure (Column 9) is the ratio of revenue produced by the largest 4 companies over the total industry revenue in 1972. The data 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 5. Predictability of Return on Equity Using Demographic Changes

	Dependent Variable: Annual Log Return on Equity (ROE) at *t+1*											
Sample	**Demographic Industries**						**All Industries**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.0850	0.1385	0.1258	0.1029	0.0785	0.1291	0.0907	0.0301	0.0217	0.1071	0.0537	0.0912
	(0.0146)***	(0.0228)***	(0.0194)***	(0.0138)***	(0.0216)***	(0.0184)***	(0.0172)***	(0.0408)	(0.0427)	(0.0147)***	(0.0268)**	(0.0243)***
Forecasted annualized demand growth between *t* and *t+2*	1.8523	2.8637	1.8805	1.1597	1.8248	2.3046	1.8145	2.8426	2.0261	1.0270	1.5416	1.8037
	(0.8010)**	(0.8169)***	(0.7900)**	(0.7937)	(0.7261)**	(0.8099)***	(1.0518)*	(1.0101)***	(0.8369)**	(0.9361)	(0.8735)*	(0.8525)**
Industry Fixed Effects		X	X		X	X		X	X		X	X
Year Fixed Effects			X			X			X			X
Sample: 1974 to 2003	X	X	X				X	X	X			
Sample: 1939 to 2003				X	X	X				X	X	X
R^2	0.0149	0.2522	0.3240	0.0075	0.2474	0.3350	0.0081	0.2703	0.3201	0.0036	0.2223	0.2768
N	*N* = 540	*N* = 540	*N* = 540	*N* = 825	*N* = 825	*N* = 825	*N* = 1334	*N* = 1334	*N* = 1334	*N* = 1940	*N* = 1940	*N* = 1940

Notes: Columns 1 through 12 report the coefficients of OLS regressions of log yearly return on equity at *t+1* on the forecasted annualized demand growth due to demographics between years *t* and *t+2*. The forecast is made using information available as of year *t-1.* The coefficients on the forecasted annual demand growth are normalized by the number of years of the forecast, 2. The coefficient indicates the typical increase in log industry return on equity (an accounting measure of profitability) due to an annualized one percentage point increase in consumption due to demographics. Each year the subset *Demographic Industries* includes the 20 industries with the highest standard deviation of forecasted annual consumption growth over the next 15 years. Standard errors are clustered by year and then scaled by a function of the autocorrelation coefficient estimated from the sample orthogonality conditions. A more thorough description of the concentration ratio measure and the standard errors is available in the text.

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

Table 6. Predictability of Return on Equity, Industry Concentration, and Industry Target Demographic

	Dependent Variable: Annual Log Return on Equity (ROE) at *t+1*											
Sample	**C-4 > median**		**C-4 <= median**		**All Industries**		**Young**		**Adults**		**Elderly**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.1095	0.2395	0.0998	0.1415	0.1101	0.2269	0.0832	0.1213	0.1115	0.0235	0.0451	0.1246
	(0.0507)**	(0.0515)***	(0.0165)***	(0.0166)***	(0.0340)***	(0.0507)***	(0.0182)***	(0.0160)***	(0.0355)***	(0.0404)	(0.0478)	(0.0296)***
Forecasted annualized demand growth between *t* and *t+2*	2.4718	2.1281	1.0367	0.8910	-0.1939	-0.9128	2.6081	2.1431	0.3740	3.3089	4.9537	5.2951
	(3.1927)	(3.4464)	(0.9309)	(0.9119)	(2.1168)	(2.0971)	(0.9981)***	(0.9072)**	(2.1658)	(1.8841)*	(3.4053)	(2.8365)*
C-4 x (Forecasted annualized demand growth between *t* and *t+2*)					5.3765	6.9705						
					(8.2430)	(8.4480)						
Industry Fixed Effects		X		X		X		X		X		X
R^2	0.0080	0.2915	0.0052	0.1870	0.0183	0.2728	0.0638	0.1397	0.0002	0.2896	0.0279	0.2882
N	*N* = 413	*N* = 413	*N* = 416	*N* = 416	*N* = 829	*N* = 829	*N* = 176	*N* = 176	*N* = 942	*N* = 942	*N* = 216	*N* = 216

Notes: Columns 1 through 10 report the coefficients of OLS regressions of log yearly return on equity at *t+1* on the forecasted annualized demand growth due to demographics between year *t* and year *t+2* from 1974 until 2003. The forecast is made using information available as of year *t-1* . The coefficients on the forecasted annual demand growth are normalized by the number of years of the forecast, 2. The coefficient indicates the typical increase in log industry return on equity (an accounting measure of profitability) due to an annualized one percentage point increase in consumption due to demographics.

The concentration ratio measure C-4 is the ratio of revenue for the largest 4 firms to total industry revenue, from the Census of Manufacturers conducted in 1972. Columns 1 and 2 report the results for the subsample of industries with a concentration ratio higher than the median of .35. Columns 3 and 4 report the results for the subsample of industries with a concentration ratio lower than or equal to the median. Columns 5 and 6 report the results for the whole sample of industries, with an interaction term between the concentration ratio and forecasted annualized demand growth. In columns 5 and 6 the concentration ratio is an unreported control variable. Columns 7 through 12 report the results for different subsets of industries based on the age group most likely to actually consume the various products. Standard errors are clustered by year and then scaled by a function of the autocorrelation coefficient estimated from the sample orthogonality conditions. A more thorough description of the concentration ratio measure and the standard errors is available in the text.

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

Table 7. Predictability of Stock Returns Using Demographic Changes

Dependent Variable: Beta-Adjusted Log Industry Stock Returns at *t+1*

Sample	Demographic Industries						All Industries					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-0.1089	0.0547	0.2944	-0.0603	-0.0723	0.0488	-0.0906	-0.0851	0.0896	-0.0428	-0.0461	0.0050
	(0.0665)	(0.1180)	(0.0871)***	(0.0397)	(0.0459)	(0.0396)	(0.0608)	(0.0712)	(0.0662)	(0.0370)	(0.0414)	(0.0332)
Forecasted annualized demand growth between *t* and *t+5*	-0.8231	-0.5300	-2.1120	-1.5905	-1.4977	-2.6454	-1.8324	-1.1504	-2.0459	-1.8714	-1.9479	-2.9338
	(4.6208)	(4.4568)	(3.2943)	(2.8413)	(3.0650)	(2.7822)	(4.5509)	(4.9956)	(3.0883)	(2.7358)	(2.9038)	(2.3055)
Forecasted annualized demand growth between *t+5* and *t+10*	10.1148	11.1968	5.9619	5.8350	5.9943	4.7945	9.3010	11.0032	5.2254	5.1586	5.4245	4.3055
	(3.6036)***	(3.5706)***	(3.5673)*	(3.4108)*	(3.4099)*	(2.7378)	(3.1419)***	(3.5824)***	(3.8584)	(3.0857)*	(3.2318)*	(2.7595)
Industry Fixed Effects		X	X		X	X		X	X		X	X
Year Fixed Effects			X			X			X			X
Sample: 1974 to 2003	X	X	X				X	X	X			
Sample: 1939 to 2003				X	X	X				X	X	X
R^2	0.0325	0.1201	0.3237	0.0110	0.0752	0.3262	0.0186	0.0595	0.1944	0.0060	0.0272	0.1888
N	*N* = 565	*N* = 565	*N* = 565	*N* = 916	*N* = 916	*N* = 916	*N* = 1385	*N* = 1385	*N* = 1385	*N* = 2272	*N* = 2272	*N* = 2272

Notes: Columns 1 through 12 report the coefficients of OLS regressions of log yearly beta-adjusted industry stock returns at t+1 on the forecasted annualized demand growth due to demographics between t and t+5 and between t+5 and t+10. The forecasts are made using information available as of year *t-1* . The industry betas for year t are obtained by 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 for both coefficients). The coefficient indicates the typical increase in log industry abnormal stock return due to an annualized one percentage point increase in forecasted consumption due to demographics. Each year the subset *Demographic Industries* includes the 20 industries with the highest standard deviation of forecasted annual consumption growth over the next 15 years. Standard errors are clustered by year and then scaled by a function of the autocorrelation coefficient estimated from the sample orthogonality conditions. A more thorough description of the concentration ratio measure and the standard errors is available in the text.

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

Table 8. Predictability of Stock Market Returns, Industry Concentration, and Industry Target Demographics

	Dependent Variable: Beta-Adjusted Log Industry Stock Returns at *t+1*											
Sample	**C-4 > median**		**C-4 <= median**		**All Industries**		**Young**		**Adults**		**Elderly**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-0.0702	0.0515	-0.0521	-0.0455	0.0032	-0.0073	-0.0951	0.0381	-0.0806	-0.0818	-0.0930	-0.0610
	(0.0762)	(0.1049)	(0.0575)	(0.0530)	(0.0642)	(0.1213)	(0.0658)	(0.0821)	(0.0783)	(0.0849)	(0.1439)	(0.1190)
Forecasted annualized demand growth between *t* and *t+5*	-11.8684	-8.6716	0.4621	0.5679	9.0352	6.7764	0.7729	-0.9288	-7.6901	-2.5510	-1.6711	1.3835
	(6.9137)*	(6.5962)	(4.5074)	(4.4957)	(6.4745)	(6.3223)	(4.3876)	(4.5646)	(6.3541)	(6.6393)	(18.8003)	(21.5727)
Forecasted annualized demand growth between *t+5* and *t+10*	18.7337	14.2421	3.0260	5.0031	-10.7217	-5.6762	9.8249	11.4864	15.0172	12.3324	7.7081	5.5776
	(7.8139)**	(7.526982)*	(3.2129)	(3.5418)	(5.5232)*	(5.9098)	(4.5912)**	(5.0212)**	(5.4828)***	(5.3214)**	(18.3862)	(18.8958)
C-4 x (Forecasted annualized demand growth between *t* and *t+5)*					-40.5195	-29.0355						
					(18.0495)**	(16.4257)*						
C-4 x (Forecasted annualized demand growth between *t+5* and *t+10)*					60.9772	44.0491						
					(20.9165)***	(21.7250)**						
Industry Fixed Effects		X		X		X		X		X		X
R^2	0.0207	0.0524	0.0035	0.0530	0.0138	0.0554	0.0618	0.1070	0.0160	0.0534	0.0041	0.0321
N	*N* = 425	*N* = 425	*N* = 445	*N* = 445	*N* = 870	*N* = 870	*N* = 193	*N* = 193	*N* = 969	*N* = 969	*N* = 223	*N* = 223

Notes: Columns 1 through 12 report the coefficients of OLS regressions of log yearly beta-adjusted industry stock returns at *t+1* on the forecasted annualized demand growth due to demographics between t and t+5 and between *t+5* and *t+10* from 1974 until 2003. The forecast is made using information available as of year *t-1*. The industry betas for year t are obtained by 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 for both coefficients). The coefficient indicates the typical increase in log industry abnormal stock return due to an annualized one percentage point increase in consumption due to demographics.

The concentration ratio measure C-4 is the ratio of revenue for the largest 4 firms to total industry revenue, from the Census of Manufacturers conducted in 1972. Columns 1 and 2 report the results for the subsample of industries with a concentration ratio higher than the median. Columns 3 and 4 report the results for the subsample of industries with a concentration ratio lower than or equal to the median. Columns 5 and 6 report the results for the whole sample of industries with two interaction terms between the industry concentration ratio and forecasted consumption growth from t to t+5 and from t+5 to t+10. In columns 5 and 6 the concentration ratio is an unreported control variable. Columns 7 through 12 report the results for different subsets of industries based on the age group most likely to actually consume the various products. Standard errors are clustered by year and then scaled by a function of the autocorrelation coefficient estimated from the sample orthogonality conditions. A more thorough description of the concentration ratio measure and the standard errors is available in the text.

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

Table 9. Performance of the Zero-Investment Portfolio for Demographic Industries

	Dependent Variable: Monthly Return on the Zero-Investment Portfolios											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.0071	0.0057	0.0066	0.0048	0.0043	0.0042	0.0081	0.0027	0.0042	0.0009	-0.0030	-0.0024
	(0.0026)***	(0.0025)**	(0.0028)**	(0.0021)**	(0.0022)**	(0.0022)*	(0.0031)***	(0.0028)	(0.0030)	(0.0024)	(0.0020)	(0.0019)
VW Index Excess Return	-0.1037	-0.0783	-0.0900	-0.0135	-0.0494	-0.0489	0.9514	1.0840	1.0648	1.0551	1.1623	1.1547
(VWRF)	(0.0753)	(0.0772)	(0.0790)	(0.0580)	(0.0579)	(0.0576)	(0.0779)***	(0.0810)***	(0.0778)	(0.0571)***	(0.0637)***	(0.0596)***
Size Factor Return		0.1902	0.1961		0.1997	0.2004		0.5480	0.5577		0.3578	0.3616
(SMB)		(0.1104)*	(0.1097)*		(0.0955)**	(0.0956)**		(0.1680)***	(0.1499)***		(0.1304)***	(0.1214)***
Value Factor Return		0.1723	0.1483		0.0673	0.0691		0.6850	0.6453		0.5126	0.4971
(HML)		(0.1043)*	(0.1049)		(0.1054)	(0.1013)		(0.1516)***	(0.1303)***		(0.1139)	(0.1005)***
Momentum Factor Return			-0.0897			0.0075			-0.1478			-0.0581
(UMD)			(0.0871)			(0.0732)			(0.1052)			(0.0520)
Sample: 1974 to 2003	X	X	X				X	X	X	X	X	X
Sample: 1939 to 2003				X	X	X						
High Predicted Growth							X	X	X			
Low Predicted Growth										X	X	X
R^2	0.0084	0.0257	0.0306	0.0001	0.0115	0.0115	0.4412	0.5611	0.5695	0.6059	0.6732	0.6746
N	N = 360	N = 360	N = 360	N = 780	N = 780	N = 780	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360

Notes: Columns 1 through 12 report the coefficients of OLS regressions of the zero-investment portfolio monthly returns on different sets of monthly benchmark factors. We create the zero-investment portfolio by double sorting the select group of demographic industries at the beginning of each year. First, we sort the industries into two equal groups based on long-term predicted demand growth. Next, within each of these two groups we sort the industries into two equal sub-groups based on the difference between predicted long-term and short-term demand growth. In columns 1 through 6, the zero-investment portfolio is long in industries with high predicted long-term demand growth and high long-term minus short-term predicted demand growth and short in industries with low predicted long-term demand growth and low long-term minus short-term predicted demand growth. Columns 1 through 3 report results from 1974 to 2003 and columns 4 through 6 report results from 1939 to 2003.
In columns 7 through 9, the zero-investment portfolio is long in industries with high predicted long-term demand growth and high long-term minus short-term predicted demand growth and short in the 1-month treasury rate. In columns 10 through 12, the zero-investment portfolio is long in the 1-month treasury rate and short in industries with low predicted long-term demand growth and low long-term minus short-term predicted 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. Heteroskedasticity and autocorrelation consistent standard errors are calculated using the Newey-West estimator with 6 lags (in parentheses). The constant is interpreted as the average monthly abnormal return for the investment strategy.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10. 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)	(10)	(11)	(12)
Constant	0.0031	0.0028	0.0024	0.0057	0.0074	0.0064	0.0002	0.0007	0.0010	0.0063	0.0076	0.0060
	(0.0016)*	(0.0017)*	(0.0017)	(0.0025)**	(0.0026)***	(0.0026)**	(0.0029)	(0.0032)	(0.0039)	(0.0034)*	(0.0037)**	(0.0036)*
VW Index Excess Return	-0.1290	-0.1069	-0.1009	-0.1148	-0.1210	-0.1087	-0.0342	-0.0401	-0.0435	-0.0553	-0.0418	-0.0208
(VWRF)	(0.0392)***	(0.0516)**	(0.0504)**	(0.0675)*	(0.0761)*	(0.0769)	(0.0916)	(0.1021)	(0.1032)	(0.0802)	(0.0934)	(0.0934)
Size Factor Return		-0.0288	-0.0318		-0.2932	0.2994		-0.0673	-0.0656		-0.3168	-0.3274
(SMB)		(0.0828)	(0.0827)		(0.1162)**	(0.1039)***		(0.1333)	(0.1400)		(0.1389)**	(0.1217)***
Value Factor Return		0.0562	0.0684		-0.1611	-0.1358		-0.0511	0.0580		-0.1103	-0.0671
(HML)		(0.0802)	(0.0784)		(0.1002)	(0.0914)		(0.1656)	(0.1894)		(0.1322)	(0.1165)
Momentum Factor Return			0.0458			0.0944			-0.0256			0.1612
(UMD)			(0.0453)			(0.0723)			(0.1302)			(0.0837)*
Concentration Ratio > 0.35				X	X	X				X	X	X
Concentration Ratio <= 0.35							X	X	X			
Large Company Only										X	X	X
$\mathbf{R^2}$	0.0374	0.0415	0.0452	0.0124	0.0515	0.0582	0.0007	0.0021	0.0024	0.0017	0.0268	0.0382
N	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360	N = 360

Notes: Columns 1 through 12 report the coefficients of OLS regressions of the zero-investment portfolio monthly returns from 1974 to 2003 on different sets of monthly benchmark factors. We create the zero-investment portfolio by double sorting all industries at the beginning of each year. First, we sort the industries into two equal groups based on long-term predicted demand growth. Next, within each of these two groups we sort the industries into two equal sub-groups based on the difference between predicted long-term and short-term demand growth. The zero-investment portfolio is long in industries with high predicted long-term demand growth and high long-term minus short-term predicted demand growth and short in industries with low predicted long-term demand growth and low long-term minus short-term predicted 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.
The concentration ratio measure is the ratio of revenue for the largest 4 firms to total industry revenue, from the 1972 Census of Manufacturers. In columns 1 through 3, we report results for all industries. In columns 4 through 6, we report results for industries with above-median concentration ratios. In columns 7 through 9, we report results for industries with below-median concentration. In columns 10 through 12, we report results using the returns for the company with the largest market capitalization at the beginning of the month instead of value-weighted returns for each industry with above-median concentration ratios. The constant is interpreted as the average monthly abnormal return for the investment strategy. 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 11. 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. 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 1982-84 $)	12879.05	21092.61*	27347.78*	23725.39*
	(15532.21)	(16178.67)	(28872.33)	(21230.03)
Total Consumption (in 1982-84 $)	10211.25	16792.38	18108.06	17935.47
	(8092.03)	(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 economic variables in the consumption surveys. Standard deviations are in parentheses. 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 variable White Head is defined for 5,435 observations in the 1935-36 survey. The information on the age of the spouse is missing in the 1960-61 survey, is defined for 13,534 observations in 1972-73 and for 6,798 observations in 1983-84. 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 and 19,974 observations in the 1972-73 survey. The variable Urban Household is defined for 13,727 observations in the 1960-61 survey. The variable Total Income is defined for 5,266 observations in 1935-36, 13,694 observations in 1960-61, 18,861 observations in 1972-73, and 9,230 observations in 1983-84. Finally, the variable Total Consumption is defined for 6,005 observations in the 1935-36 survey.

Appendix Table 2. 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	2710-2719, (5192)
Magazines	Media	2720-2729, (2730-2739, 5192)
Cruises	Health	4480-4481, (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 Industry grouping (Column 2) and SIC industry classification (Column 3). Each expenditure category is associated with 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 their expenditure category classification. If the internet search cannot identify the specific category for one of these companies, then the company is excluded from our analysis.