Internet Appendix for "Understanding the Risk-Return Relation: The Aggregate Wealth Proxy Actually Matters"

Scott Cederburg* Michael S. O'Doherty†

December 4, 2017

Abstract

This document provides supplementary material to the paper "Understanding the Risk-Return Relation: The Aggregate Wealth Proxy Actually Matters." Appendix A shows analytically that stock-based tests of the ICAPM generally produce biased estimates of the coefficient of relative risk aversion when stocks are an imperfect proxy for aggregate wealth. Appendix B introduces a Monte Carlo simulation exercise to characterize the magnitude of the bias in stock-based tests of the ICAPM and examine the power of alternative tests to identify a positive risk-return relation. Appendix C provides additional detail on the data for asset-class values and returns used in constructing the aggregate wealth proxy in the paper. Appendix D analyzes the relation between measurement error in our aggregate wealth proxy and bias in estimated risk-return coefficients. Appendix E includes supplementary empirical results and robustness tests.

*Eller College of Management, University of Arizona, Tucson, AZ 85721. Email: cederburg@email.arizona.edu.
†Trulaske College of Business, University of Missouri, Columbia, MO 65211. Email: odohertym@missouri.edu.

A Derivation of Risk-Return Coefficient for Stocks

In this appendix, we derive equation (12) in the paper which relates the risk-return coefficient from stock-based tests, γ_s, to the coefficient of relative risk aversion, γ, when stocks are an imperfect proxy for aggregate wealth. We specifically show that the coefficient of relative risk aversion does not generally govern the tradeoff between expected stock returns and stock variance. Instead, the true risk-return relation for stocks is substantially less pronounced.

For the purpose of deriving an analytical solution, we assume that aggregate wealth and stock market returns jointly follow the scalar BEKK bivariate GARCH model of Engle and Kroner (1995) given in equations (8) to (10) in the paper. If the ICAPM holds and $E_{t-1}(R_{s,t}) = \gamma\sigma_{s,aw,t}$, the coefficient from a predictive regression of stock market returns on stock market variance is

$$\begin{aligned}\gamma_s &= \frac{Cov(R_{s,t}, \sigma^2_{s,t})}{Var(\sigma^2_{s,t})}\\ &= \frac{Cov\left(\gamma\sigma_{s,aw,t} + \varepsilon_{s,t}, \frac{Cov(\sigma^2_{s,t},\sigma_{s,aw,t})}{Var(\sigma_{s,aw,t})}\sigma_{s,aw,t} + \nu_t\right)}{Var(\sigma^2_{s,t})},\end{aligned} \tag{A1}$$

where ν_t is the residual from an unconditional regression of $\sigma^2_{s,t}$ on $\sigma_{s,aw,t}$. Because $E_{t-1}(R_{s,t}) = \gamma\sigma_{s,aw,t}$ then $Cov(R_{s,t}, \nu_t) = 0$, such that[1]

$$\begin{aligned}\gamma_s &= \frac{Cov\left(\gamma\sigma_{s,aw,t} + \varepsilon_{s,t}, \frac{Cov(\sigma^2_{s,t},\sigma_{s,aw,t})}{Var(\sigma_{s,aw,t})}\sigma_{s,aw,t} + \nu_t\right)}{Var(\sigma^2_{s,t})}\\ &= \frac{\gamma\frac{Cov(\sigma^2_{s,t},\sigma_{s,aw,t})}{Var(\sigma_{s,aw,t})}Var(\sigma_{s,aw,t})}{Var(\sigma^2_{s,t})}\\ &= \frac{\gamma Cov(\sigma^2_{s,t}, \sigma_{s,aw,t})}{Var(\sigma^2_{s,t})}.\end{aligned} \tag{A2}$$

Given the structure of the scalar BEKK GARCH model, we can solve for $Var(\sigma^2_{s,t})$ and $Cov(\sigma^2_{s,t}, \sigma_{s,aw,t})$ as functions of the underlying parameters. First, define $\Omega^* = \frac{\Omega}{1-\alpha^2-\delta^2}$. Note that

$$Var(\sigma^2_{s,t}) = E(\sigma^4_{s,t}) - E(\sigma^2_{s,t})^2, \tag{A3}$$

and

$$E(\sigma^2_{s,t}) = E(\varepsilon^2_{s,t}) = \Omega^*_{s,s}. \tag{A4}$$

[1] Keeping with the notation in the paper that $\sigma^2_{s,t}$ is known to the representative investor at time $t-1$, the ν_t shock occurs in time $t-1$ such that it is uncorrelated with the unexpected stock market return $\varepsilon_{s,t}$. Further, ν_t is orthogonal to $\sigma_{s,aw,t}$ by construction, so $Cov(E_{t-1}(R_{s,t}), \nu_t) = 0$. The statement that $Cov(R_{s,t}, \nu_t) = 0$ results from these two facts.

Solving for $E(\sigma^4_{s,t})$ produces

$$\begin{aligned}
E(\sigma^4_{s,t}) &= E(((1-\alpha^2-\delta^2)\Omega^*_{s,s} + \alpha\varepsilon^2_{s,t-1}\alpha + \delta\sigma^2_{s,t-1}\delta)^2) \\
&= (1-\alpha^2-\delta^2)^2\Omega^{*2}_{s,s} + \alpha^4 E(\varepsilon^4_{s,t-1}) + \delta^4 E(\sigma^4_{s,t-1}) + 2(1-\alpha^2-\delta^2)\Omega^*_{s,s}\alpha^2 E(\varepsilon^2_{s,t-1}) \\
&\quad + 2(1-\alpha^2-\delta^2)\Omega^*_{s,s}\delta^2 E(\sigma^2_{s,t-1}) + 2\alpha^2\delta^2 E(\varepsilon^2_{s,t-1}\sigma^2_{s,t-1}) \\
&= (1-\alpha^2-\delta^2)^2\Omega^{*2}_{s,s} + 3\alpha^4 E(\sigma^4_{s,t}) + \delta^4 E(\sigma^4_{s,t}) + 2(1-\alpha^2-\delta^2)\Omega^{*2}_{s,s}\alpha^2 \\
&\quad + 2(1-\alpha^2-\delta^2)\Omega^{*2}_{s,s}\delta^2 + 2\alpha^2\delta^2 E(\sigma^4_{s,t}), \qquad \text{(A5)}
\end{aligned}$$

such that

$$E(\sigma^4_{s,t}) = \frac{(1-\alpha^4-\delta^4-2\alpha^2\delta^2)\Omega^{*2}_{s,s}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2}. \qquad \text{(A6)}$$

Finally,

$$\begin{aligned}
Var(\sigma^2_{s,t}) &= E(\sigma^4_{s,t}) - E(\sigma^2_{s,t})^2 \\
&= \frac{(1-\alpha^4-\delta^4-2\alpha^2\delta^2)\Omega^{*2}_{s,s}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2} - \frac{(1-3\alpha^4-\delta^4-2\alpha^2\delta^2)\Omega^{*2}_{s,s}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2} \\
&= \frac{2\alpha^4\Omega^{*2}_{s,s}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2}. \qquad \text{(A7)}
\end{aligned}$$

To solve for $Cov(\sigma^2_{s,t}, \sigma_{s,aw,t})$, note that

$$Cov(\sigma^2_{s,t}, \sigma_{s,aw,t}) = E(\sigma^2_{s,t}\sigma_{s,aw,t}) - E(\sigma^2_{s,t})E(\sigma_{s,aw,t}), \qquad \text{(A8)}$$

where

$$E(\sigma^2_{s,t})E(\sigma_{s,aw,t}) = \Omega^*_{s,s}\Omega^*_{s,aw}, \qquad \text{(A9)}$$

and

$$\begin{aligned}
E(\sigma^2_{s,t}\sigma_{s,aw,t}) &= E(((1-\alpha^2-\delta^2)\Omega^*_{s,s} + \alpha\varepsilon^2_{s,t-1}\alpha + \delta\sigma^2_{s,t-1}\delta)((1-\alpha^2-\delta^2)\Omega^*_{s,aw} + \alpha\varepsilon_{aw,t-1}\varepsilon_{s,t-1}\alpha + \delta\sigma_{s,aw,t}\delta)) \\
&= (1-\alpha^2-\delta^2)^2\Omega^*_{s,s}\Omega^*_{s,aw} + 3\alpha^4 E(\sigma^2_{s,t-1}\sigma_{s,aw,t-1}) + \delta^4 E(\sigma^2_{s,t-1}\sigma_{s,aw,t-1}) \\
&\quad + 2(\alpha^2+\delta^2)(1-\alpha^2-\delta^2)\Omega^*_{s,s}\Omega^*_{s,aw} + 2\alpha^2\delta^2 E(\sigma^2_{s,t-1}\sigma_{s,aw,t-1}). \qquad \text{(A10)}
\end{aligned}$$

Rearranging,

$$E(\sigma^2_{s,t}\sigma_{s,aw,t}) = \frac{(1-\alpha^4-\delta^4-2\alpha^2\delta^2)\Omega^*_{s,s}\Omega^*_{s,aw}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2}. \qquad \text{(A11)}$$

Thus,

$$
\begin{aligned}
Cov(\sigma^2_{s,t}, \sigma_{s,aw,t}) &= E(\sigma^2_{s,t}\sigma_{s,aw,t}) - E(\sigma^2_{s,t})E(\sigma_{s,aw,t}) \\
&= \frac{(1-\alpha^4-\delta^4-2\alpha^2\delta^2)\Omega^*_{s,s}\Omega^*_{s,aw}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2} - \frac{(1-3\alpha^4-\delta^4-2\alpha^2\delta^2)\Omega^*_{s,s}\Omega^*_{s,aw}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2} \\
&= \frac{2\alpha^4\Omega^*_{s,s}\Omega^*_{s,aw}}{1-3\alpha^4-\delta^4-2\alpha^2\delta^2}. \qquad \text{(A12)}
\end{aligned}
$$

Substituting the solutions for $Cov(\sigma^2_{s,t}, \sigma_{s,aw,t})$ and $Var(\sigma^2_{s,t})$ into equation (A2),

$$
\begin{aligned}
\gamma_s &= \frac{\gamma Cov(\sigma^2_{s,t}, \sigma_{s,aw,t})}{Var(\sigma^2_{s,t})} \\
&= \gamma\frac{\Omega^*_{s,aw}}{\Omega^*_{s,s}}. \qquad \text{(A13)}
\end{aligned}
$$

Noting that $\Omega^*_{s,aw} = \rho\sigma_s\sigma_{aw}$, where ρ is the unconditional correlation between stock market and aggregate wealth returns,

$$
\begin{aligned}
\gamma_s &= \gamma\frac{\Omega^*_{s,aw}}{\Omega^*_{s,s}} \\
&= \gamma\rho\frac{\sigma_{aw}}{\sigma_s}. \qquad \text{(A14)}
\end{aligned}
$$

B Model Simulation

Appendix A demonstrates that empirical tests that relate the expected stock market return to conditional stock market variance produce biased estimates of the coefficient of relative risk aversion. In this appendix, we use a Monte Carlo simulation exercise to investigate the quantitative magnitude of this bias in the risk-return coefficient and to consider the power of various tests to identify a positive relation between risk and return.

We simulate time series of returns for the stock market and aggregate wealth portfolios from a diagonal BEKK bivariate GARCH model. In particular, return dynamics are given by

$$\begin{bmatrix} R_{aw,t} \\ R_{s,t} \end{bmatrix} = \pi_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma_t), \tag{B1}$$

where π_t is the expected return vector. The process for Σ_t is

$$\Sigma_t = \Omega + \alpha(\varepsilon_{t-1}\varepsilon'_{t-1})\alpha + \delta\Sigma_{t-1}\delta, \tag{B2}$$

and the diagonal BEKK bivariate GARCH model restricts the α and δ parameters to be diagonal matrices. We consider two alternative scenarios for specifying the expected return vector, π_t, in equation (B1). The first case assumes that stocks are a perfect proxy for aggregate wealth, such that $E_{t-1}(R_{s,t}) = \gamma\sigma^2_{s,t}$ under the ICAPM. Prior literature uses the stock market as a proxy for aggregate wealth, such that this case provides a baseline for the assumed properties of risk-return tests in these studies. In the second and more realistic case, the stock market portfolio is an imperfect proxy for aggregate wealth. The ICAPM therefore implies that $E_{t-1}(R_{aw,t}) = \gamma\sigma^2_{aw,t}$ and $E_{t-1}(R_{s,t}) = \gamma\sigma_{s,aw,t}$ in this scenario. As such, we are able to directly investigate the effect of using an imperfect aggregate wealth proxy by comparing across the two scenarios.

The BEKK model parameters in the simulation match values estimated using historical data for the stock market and the aggregate wealth portfolio that we develop in Section 4 in the paper. In particular, we estimate the model using monthly data over the 1953 to 2013 period using our aggregate wealth portfolio and the stock market portfolio. The covariance process has parameters

$$\alpha = \begin{bmatrix} 0.2837 & 0.0000 \\ 0.0000 & 0.3175 \end{bmatrix}, \tag{B3}$$

$$\delta = \begin{bmatrix} 0.9449 & 0.0000 \\ 0.0000 & 0.9315 \end{bmatrix}, \tag{B4}$$

and

$$\Omega \times 10^4 = \begin{bmatrix} 0.1173 & 0.2789 \\ 0.2789 & 0.7400 \end{bmatrix}. \tag{B5}$$

For a range of γ from 0 to 10 (with increments of 0.1), we simulate 10,000 draws of 732-month sample periods of the returns and conditional covariance matrix of aggregate wealth and stock returns. In each draw, we initially set the covariance matrix to its long-run expectation, $(I - \alpha - \delta)^{-1}\Omega(I - \alpha - \delta)^{-1}$, and then simulate a 500-month warm-up period. The first value of Σ_t in the 732-month sample is set to the ending value from the warm-up period.

Given the simulated series of stock market and aggregate wealth returns, we examine three risk-return tests using two-step approaches. In the first step of each test, the relevant conditional variance is estimated using a GARCH(1,1) process. The second step in each test is a predictive regression for returns given the fitted variance predictive variable. More specifically, in the scenario in which stocks are a perfect proxy for aggregate wealth, we regress stock returns on the fitted conditional stock market variance from a GARCH(1,1) model. For the case in which stocks are an imperfect proxy, we examine two tests. First, we relate stock returns to the GARCH(1,1) fitted conditional stock market variance. This test produces a biased estimate of the coefficient of relative risk aversion and is comparable to tests used in the prior literature. Second, we regress aggregate wealth portfolio returns on the fitted conditional aggregate wealth variance from a GARCH(1,1) process, which produces an unbiased estimate of the coefficient of relative risk aversion according to the ICAPM.

To assess statistical significance of coefficient estimates in our simulation, we implement a bootstrap approach within each of the 10,000 Monte Carlo draws that closely matches the method from our empirical tests in the paper. Specifically, given a draw of the time series of returns and fitted conditional variance for a given test, we form 1,000 bootstrapped 732-month samples. We estimate the risk-return coefficient for each bootstrap sample and record the proportion of estimates that are below zero in each iteration of the simulation, which corresponds to bootstrap p-value for the one-sided test that the risk-return coefficient is less than zero. We then assess test power by examining how often this p-value is below the 5% level across iterations in the Monte Carlo simulation.

We begin our analysis by considering tests of the risk-return relation that predict stock returns using stock variance as in equation (6) in the paper. Figure B.1 shows the distribution of γ_s estimates for the cases in which stocks are a perfect proxy (solid line) or imperfect proxy (dotted line) for aggregate wealth. We focus here on the simulations for which the true coefficient of relative risk aversion is set to five. The γ_s estimates from the predictive regression are concentrated around this true value when stocks are a perfect proxy. Further, the estimated risk-return coefficient is positive in over 99% of draws, and the γ_s estimates are statistically significant at the 5% level in a one-tailed test in about 80% of the simulated samples. Thus, the test has desirable properties if stocks are a perfect proxy for aggregate wealth.

Figure B.1 also shows results from the more realistic setting in which stocks are an imperfect proxy. The risk-return coefficient from the predictive regression in equation (6) in the paper is a biased estimate of γ in this case, consistent with the discussion above. The average estimate is 1.94 even as the true γ is 5.00, so the bias is large and economically meaningful. Further, test power is substantially reduced by the bias. A 5% statistical rejection in a one-tailed test occurs in only 25% of simulated samples. The standard errors for the coefficient estimates from the perfect and imperfect proxy cases are similar, such that the observed reduction in power for this test is largely attributable to the downward bias in the risk-return coefficient estimate.

We next turn to a more comprehensive investigation of test power. In related work, Lundblad (2007) examines tests of the risk-return relation for stocks. He assumes that the stock market is a perfect proxy for aggregate wealth and that the true coefficient of relative risk aversion is two. Under these assumptions, Lundblad (2007) finds that the commonly used tests suffer from low power to identify a significantly positive risk-return relation.

Figure B.2 shows the power of the predictive regression in equation (6) in the paper to identify a positive risk-return relation for the perfect and imperfect proxy cases. We consider test power for a range of γ from 0 to 10. In particular, the figure plots the proportion of simulations in which the bootstrap p-value is less than the 5% level for the one-tailed test of the null hypothesis $\gamma_s \leq 0$ versus the alternative $\gamma_s > 0$. The solid line plots test power when stocks are a perfect proxy for aggregate wealth and the dashed line represents the imperfect proxy case.

Figure B.2 demonstrates that test power is substantially reduced when stocks are an imperfect proxy for aggregate wealth. In the case in which stocks are a perfect proxy, the predictive regression produces significant evidence of a positive risk-return relation in 28% of simulations under Lundblad's (2007) assumption of $\gamma = 2$. For higher levels of risk aversion, the test's ability to detect a relation between risk and return increases substantially. For example, a 5% rejection is obtained in 65% of draws when $\gamma = 4$ and 89% of draws for $\gamma = 6$. Thus, in the setting studied by Lundblad

(2007) and prior literature, this test has relatively high power to identify a significant relation between expected stock returns and conditional stock variance even with a moderate sample size as long as γ is reasonably large.

On the other hand, we find that test power is much lower when stocks are an imperfect proxy. That is, the probability of identifying a significant positive relation between the expected stock return and conditional stock variance is substantially reduced relative to the perfect-proxy case over the reasonable range of γ. In comparison to the cases above, the predictive regression produces significant evidence of a positive risk-return relation in only 11% of draws when $\gamma = 2$, 19% of simulations if $\gamma = 4$, and 31% of draws for $\gamma = 6$. Figure B.2 thus shows that finding significant evidence of a positive relation between expected stock returns and conditional stock market risk is likely to be difficult in the more realistic case in which stock returns are not exactly equal to returns on aggregate wealth.

Finally, we consider the properties of risk-return tests when stocks are an imperfect proxy. In this case, the ICAPM implies that expected returns are given by equation (11) in the paper. We examine two predictive regressions in this setting: (i) regress aggregate wealth returns on conditional aggregate wealth variance and (ii) regress stock market returns on conditional stock market variance. Despite the fact that the first of these regressions produces an unbiased coefficient of relative risk aversion estimate, Figure B.3 shows that the power of this test is not substantially different from the power of the second, biased regression. Although the coefficient estimates are substantially larger for the aggregate wealth tests, the standard errors are also larger such that the rejection rates from both tests are similar. Intuitively, the low power of the aggregate wealth risk-return test relative to the perfect-proxy case for a given γ in Figure B.2 results from the comparatively low variation in expected returns when premiums are based on exposure to the risk of aggregate wealth rather than stocks. This test is therefore useful for recovering unbiased estimates of γ but is not expected to substantially increase the statistical evidence for a positive risk-return relation.

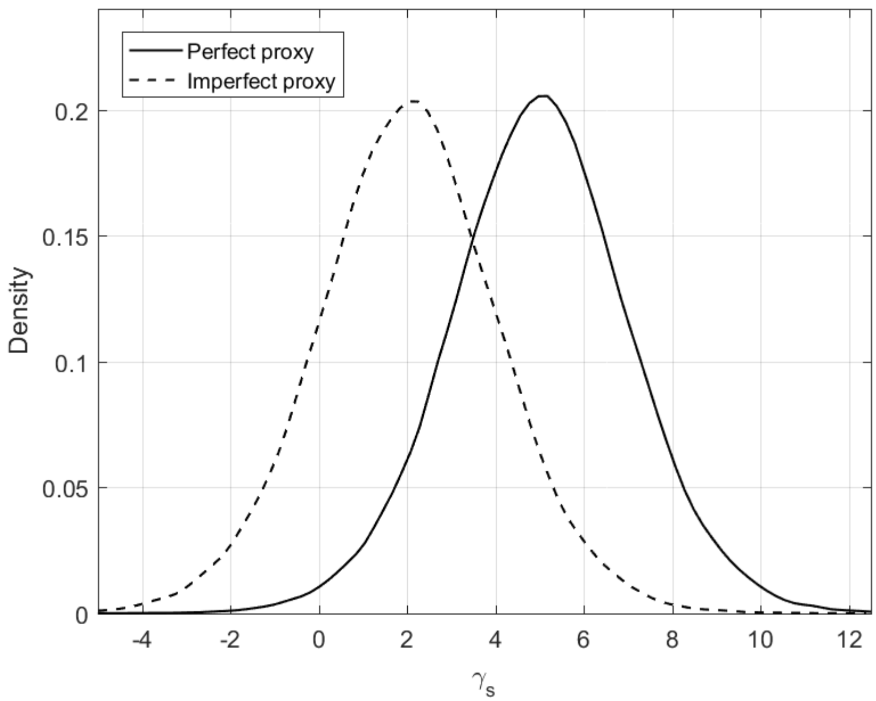

Figure B.1: Monte Carlo Densities of the Risk-Return Coefficient.
The figure shows distributions of the estimated risk-return coefficient from a Monte Carlo simulation in which the true coefficient of relative risk aversion equals five. The solid (dashed) line represents the case in which stocks are a perfect (imperfect) proxy for aggregate wealth. Returns are generated assuming that the conditional covariance matrix of stock market and aggregate wealth returns follows a diagonal BEKK bivariate GARCH process. Conditional stock market variance is estimated in each draw using a GARCH(1,1) model. Estimates of γ_s are from the predictive regression $R_{s,t} = \mu + \gamma_s \hat{\sigma}^2_{s,t} + \eta_{s,t}$, and the figure plots distributions of $\hat{\gamma}_s$ for the two scenarios over 10,000 simulations of 732-month sample periods.

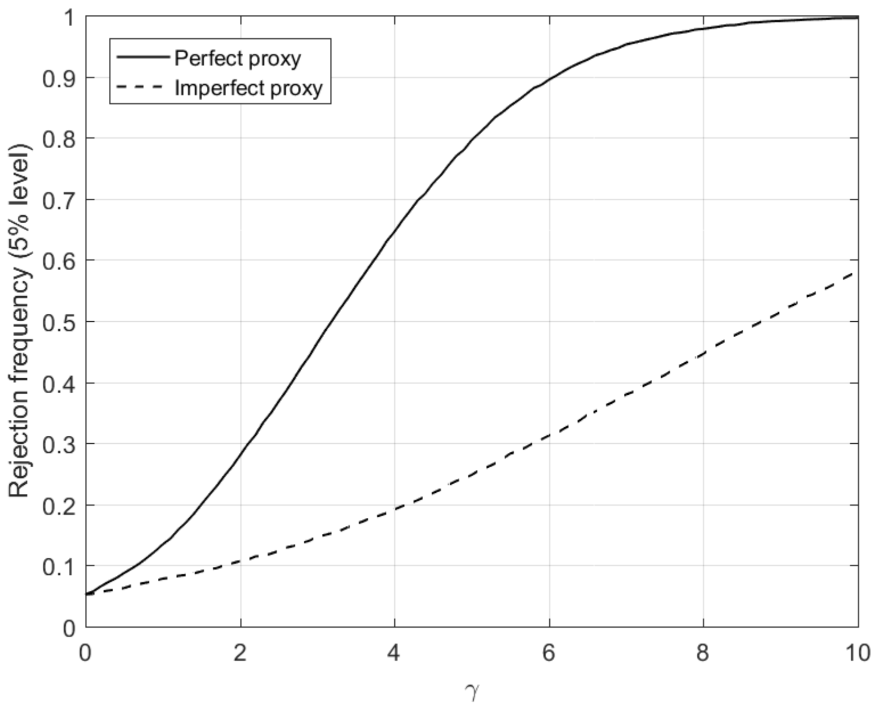

Figure B.2: Monte Carlo Simulations: Comparison of Test Power for the Perfect Proxy and Imperfect Proxy Cases.
The figure shows the probability of obtaining a statistical rejection at the 5% significance level for the one-sided test of the null hypothesis that the risk-return parameter is less than zero using a Monte Carlo simulation. The solid (dashed) line shows the case in which stocks are a perfect (imperfect) proxy for aggregate wealth. The true coefficient of relative risk aversion varies from 0 to 10. The Monte Carlo simulation generates 10,000 draws of 732-month sample periods of stock market and aggregate wealth returns, which are generated assuming that the conditional covariance matrix of stock market and aggregate wealth returns follows a diagonal BEKK bivariate GARCH process. Conditional stock market variance is estimated in each draw using a GARCH(1,1) model. We measure the statistical significance of the γ_s estimates from the predictive regression $R_{s,t} = \mu + \gamma_s \hat{\sigma}^2_{s,t} + \eta_{s,t}$ in each simulation draw using a pairs bootstrap approach with 1,000 bootstrap samples of 732-month time series. The figure plots the proportion of Monte Carlo draws in which the bootstrap p-value is less than 0.05.

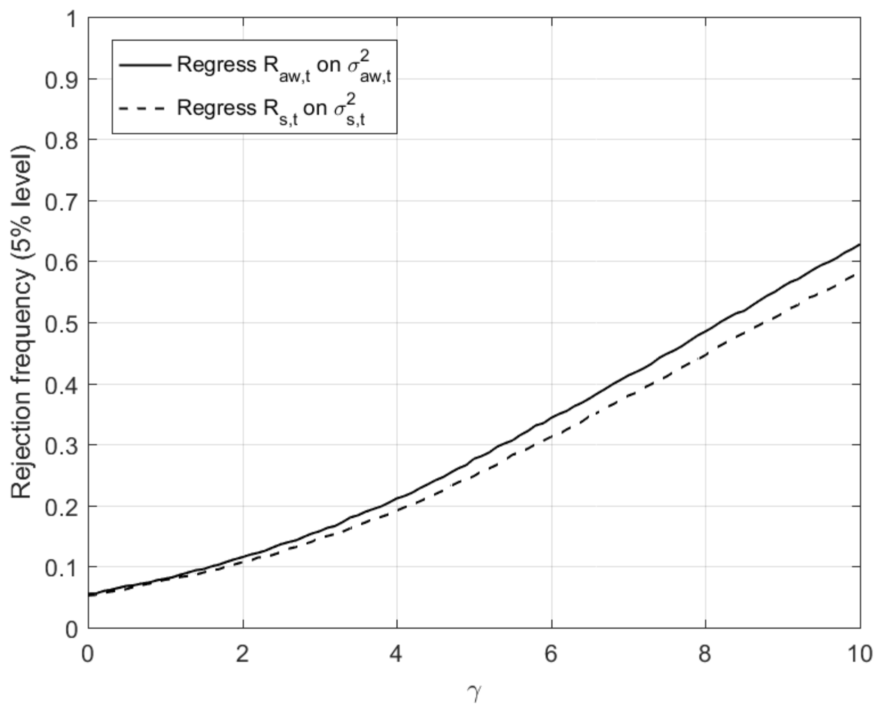

Figure B.3: Monte Carlo Simulations: Comparison of Test Power for Alternative Approaches when Stocks are an Imperfect Proxy for Aggregate Wealth.
The figure shows the probability of obtaining a statistical rejection at the 5% significance level for the one-sided test of the null hypothesis that the risk-return parameter is less than zero when stocks are an imperfect proxy for aggregate wealth using a Monte Carlo simulation. The solid line shows results for a test relating expected aggregate wealth returns to conditional aggregate wealth variance and the dashed line represents a test relating expected stock returns to conditional stock market variance. The true coefficient of relative risk aversion varies from 0 to 10. The Monte Carlo simulation produces 10,000 draws of 732-month sample periods of stock market and aggregate wealth returns, which are generated assuming that the conditional covariance matrix of stock market and aggregate wealth returns follows a diagonal BEKK bivariate GARCH process. The conditional variances of stock market and aggregate wealth returns are estimated in each draw using GARCH(1,1) models. We measure the statistical significance of the γ and γ_s estimates from predictive regressions of returns on conditional variance estimates in each simulation draw using a pairs bootstrap approach with 1,000 bootstrap samples of 732-month time series. The figure plots the proportion of Monte Carlo draws in which the bootstrap p-value is less than 0.05 for each test.

C Data Description

This appendix provides additional details on the construction of our aggregate wealth return series. Section C.1 provides the data sources for component returns, Section C.2 outlines the estimation of portfolio weights using the Federal Reserve Flow of Funds tables, and Section C.3 discusses the primary assets excluded from our specification of aggregate wealth.

C.1 Aggregate Wealth Component Returns

This section details the data sources for returns on each of the nine component asset classes included in the aggregate wealth portfolio. Unless otherwise noted, we use monthly returns for each asset class over the period January 1953 to December 2013.

- **Corporate Equity:** Corporate equity returns are based on the CRSP value-weighted portfolio. We obtain this data from Kenneth French's website.

- **Noncorporate Equity:** Noncorporate equity returns are based on the returns of small publicly traded stocks. We use returns of the smallest decile portfolio from portfolios sorted on market capitalization from Kenneth French's website.

- **Deposits:** Deposits returns are based on the U.S. 30-day Government Bond return series (Table A-14) from the Ibbotson SBBI Classic Yearbook.

- **Treasury Debt:** Treasury debt returns are calculated using returns data for U.S. Intermediate-Term Government bonds (Table A-10) and U.S. Long-Term Government bonds (Table A-6) from the Ibbotson SBBI Classic Yearbook. The Federal Reserve Flow of Funds tables report overall household ownership of Treasury securities. We therefore form a portfolio of the Intermediate-Term and Long-Term Government bond series to proxy for the returns on household government bond holdings. The weights on the Intermediate-Term and Long-Term series are based on the total face values of shorter- versus longer-term maturity bonds, where bonds with less (greater) than 10 years to maturity are classified as intermediate-term (long-term) bonds. Historical outstanding Treasury debt is from the Monthly Statement of the Public Debt, and archived reports are available on the U.S. Department of the Treasury website (http://www.treasurydirect.gov/govt/reports/pd/mspd/mspd.htm). We calculate the weights of Intermediate-Term and Long-Term bonds using the December report for each year and apply these weights to calculate the government bond return series for the following year.

- **Agency Debt:** Agency debt returns are based on the Bank of America Merrill Lynch U.S. Agency Index available through Bloomberg. This return series starts in January 1978.

- **Municipal Debt:** Municipal debt returns are based on the Bank of America Merrill Lynch U.S. Municipal Securities Index available through Bloomberg. This return series starts in January 1989.

- **Corporate Debt:** Corporate debt returns are based on the U.S. Long-Term Corporate Bond return series (Table A-5) from the Ibbotson SBBI Classic Yearbook.

- **Defined Benefit Pensions:** We approximate defined benefit pension returns using a portfolio of corporate and government bonds, with weights determined by the relative values of private pension plans for the corporate bond series and local, state, and federal government pension plans for the government bond series. The corporate bond returns are from the U.S. Long-Term Corporate Bond return series (Table A-5), and the government bond returns are from the U.S. Intermediate-Term Government Bond return series (Table A-10) from the Ibbotson SBBI Classic Yearbook. In estimating weights, we use the value of total financial assets in defined benefit plans for private pension funds (Flow of Funds Table L.117), state and local government employee retirement funds (Table L.118), and federal government employee retirement funds (Table L.119).

- **Real Estate:** Residential real estate returns are estimated using information from the Housing component of the Consumer Price Index (CPI). For the period from December 1952 to June 1985, the CPI calculation included a cost index of Homeownership which reflected the cost of home purchase. This data series is used by Fama and Schwert (1977a) and Stambaugh (1982), with Fama and Schwert (1977a) noting that the CPI Homeownership data "seems to be the best available quality adjusted index of transaction prices for real estate." CPI data for this period are taken from the monthly Consumer Price Index report (http://fraser.stlouisfed.org/publication/?pid=67&tid=21) and CPI Detailed Report (http://fraser.stlouisfed.org/publication/?pid=58&tid=21). Prior to January 1964, the Homeownership index was not directly listed in these reports. Thus, we infer its value from the reported Housing index and its subcomponents along with the CPI basket weights. See http://fraser.stlouisfed.org/docs/publications/bls/bls_no1517.pdf for basket weights in this period. For the period from July 1985 to December 2013, the CPI is based on Owner's Equivalent Rent rather than cost of home purchase. The Owner's Equivalent Rent series is obtained from the Federal Reserve Bank of St. Louis website (http://research.stlouisfed.org/fred2/). Monthly returns are calculated as the percentage change in the index value.

C.2 Aggregate Wealth Component Weights

Individual asset-class weights are assigned in proportion to their respective aggregate values as reported in Table B.100 (Balance Sheet of Households and Nonprofit Organizations) of the quarterly Flow of Funds release. As described below, we also make use of Tables B.100.e (Balance Sheet of Households and Nonprofit Organizations with Equity Detail), L.116 (Private and Public Pension Funds), L.117.c (Private Pension Funds: Defined Contribution Plans), and L.121 (Mutual Funds) in assigning estimates of holdings of indirectly held assets to the appropriate asset classes. Given the quarterly frequency of the Flow of Funds releases, we estimate aggregate wealth portfolio weights at the beginning of each quarter and apply these weights in computing returns over the subsequent three months. The procedure for estimating the aggregate holdings of each asset class is summarized below.

- **Corporate Equity:** Corporate equity holdings are directly held equity shares at market value (Table B.100.e) plus indirectly held equity shares at market value (Table B.100.e).
- **Noncorporate Equity:** Noncorporate equity holdings are reported directly as equity in noncorporate business (Table B.100).
- **Deposits:** Deposits are reported directly in Table B.100.
- **Treasury Debt:** Treasury debt holdings are direct holdings of Treasury securities (Table B.100) plus household share of mutual funds times Treasury securities in mutual funds (Table L.121) plus defined contribution plan share of mutual funds times Treasury securities in mutual funds (Table L.121) plus Treasury securities in defined contribution plans (Table L.117.c). The household share of mutual funds is mutual fund shares (Table B.100) divided by total financial assets of mutual funds (Table L.121). The defined contribution plan share of mutual funds is mutual fund shares (Table L.117.c) divided by total financial assets of mutual funds (Table L.121).
- **Agency Debt:** Agency debt holdings are direct holdings of agency- and GSE-backed securities (Table B.100) plus household share of mutual funds times agency- and GSE-backed securities in mutual funds (Table L.121) plus defined contribution plan share of mutual funds

times agency- and GSE-backed securities in mutual funds (Table L.121) plus agency- and GSE-backed securities in defined contribution plans (Table L.117.c).

- **Municipal Debt:** Municipal debt holdings are direct holdings of municipal securities (Table B.100) plus household share of mutual funds times municipal securities in mutual funds (Table L.121) plus defined contribution plan share of mutual funds times municipal securities in mutual funds (Table L.121).
- **Corporate Debt:** Corporate debt holdings are direct holdings of corporate and foreign bonds (Table B.100) plus household share of mutual funds times corporate and foreign bonds in mutual funds (Table L.121) plus defined contribution plan share of mutual funds times corporate and foreign bonds in mutual funds (Table L.121) plus corporate and foreign bonds in defined contribution plans (Table L.117.c).
- **Defined Benefit Pensions:** Defined benefit pension assets are pension entitlements (Table B.100) minus household defined contribution plan assets (Table L.116).
- **Real Estate:** Real estate assets are real estate for households and nonprofit organizations (Table B.100) minus home mortgages (Table B.100).

C.3 Exclusions from the Aggregate Wealth Specifications

This section itemizes individual asset classes excluded from the proxy for aggregate wealth used in our empirical tests.

- **Human Capital:** Returns and values of human capital are notoriously difficult to estimate, and various assumptions about labor cash flow dynamics and discount rates can produce alternative estimates which differ substantially from an economic perspective. Depending on assumptions, researchers have made the case that returns on the stock market and human capital are strongly positively related (e.g., Campbell (1996), Baxter and Jermann (1997), and Benzoni, Collin-Dufresne, and Goldstein (2007)), not closely related (e.g., Fama and Schwert (1977b), Jagannathan and Wang (1996), and Lustig, Van Nieuwerburgh, and Verdelhan (2013)), or even negatively related (e.g., Lustig and Van Nieuwerburgh (2008)). Due to the difficulty of measuring both the value and returns to human capital, our paper focuses primarily on the components of aggregate wealth that are measurable. In Appendix E, we examine an aggregate wealth series that incorporates Jagannathan and Wang's (1996) measure of human capital returns. Although accounting for human capital using this measure or one of the alternative proxies may change the estimate for the coefficient of relative risk aversion in our tests, it seems unlikely that including labor wealth would change the broader conclusion that tests of the ICAPM differ substantially across the stock market and aggregate wealth portfolios.

- **Durable Goods:** Stambaugh (1982) considers durable goods in his aggregate wealth portfolio proxies. Reliable monthly returns on durables are not available, and durables are also a relatively small proportion of total wealth.

- **International Assets:** We restrict our analysis to the asset holdings of U.S. households, whereas the aggregate wealth portfolio potentially includes the holdings of all households in the world. Theoretically, if the U.S. market is segmented from foreign markets then the assumption of a U.S.-only aggregate wealth portfolio is valid for studying risk and return. A related literature studies the International CAPM by examining equity markets across countries and finds evidence of interrelated markets (e.g., Chan, Karolyi, and Stulz (1992) and Pástor, Sinha, and Swaminathan (2008)), which suggests that international assets may be relevant for characterizing the risk-return relation. From a practical perspective, however, comparable worldwide data on household asset holdings and reliable returns data on alternative asset classes are not readily available. We thus restrict our analysis to the U.S. market.

D Measurement Error

Our main tests rely on a market proxy that weights the returns of several asset classes to form a more comprehensive measure of aggregate wealth relative to the stock market. In the process of developing our proxy, we use information about the values and returns of the asset classes that may contain measurement error. We include an analysis of the robustness of our results to alternative proxies in Appendix E. In this appendix, we take a different approach by analyzing the potential effect of measurement error in our aggregate wealth proxy on the bias in the estimated risk-return coefficients from our tests.

Defining the true aggregate wealth portfolio return as $R_{aw,t}$, the returns of our aggregate wealth proxy ($R_{p,t}$) can be expressed as

$$R_{p,t} = R_{aw,t} + \nu_{p,t}. \tag{D1}$$

In this setting, $\nu_{p,t}$ captures all forms of measurement error in our proxy, including error that results from misestimated weights, mismeasured returns, and omitted asset classes. We argue in the paper that using the stock market as a proxy for aggregate wealth is likely to produce a substantial downward bias in the estimated risk aversion coefficient. In this appendix, we compare the potential magnitude of the bias using stocks to the bias produced by potential measurement error in our aggregate wealth proxy.

Equation (12) of the paper implies that the bias in the estimated coefficient of relative risk aversion using stocks is

$$\text{Bias(Stocks)} = \left(\rho \frac{\sigma_{aw}}{\sigma_s} - 1\right)\gamma, \tag{D2}$$

where ρ is the correlation between the returns on the stock market and aggregate wealth, and σ_s and σ_{aw} are the unconditional standard deviations of the stock market and aggregate wealth portfolios, respectively. The bias in the coefficient of relative risk aversion estimate associated with our aggregate wealth proxy is

$$\text{Bias(Proxy)} = \left(\rho_{p,aw} \frac{\sigma_{aw}}{\sigma_p} - 1\right)\gamma, \tag{D3}$$

where $\rho_{p,aw}$ is the correlation between our aggregate wealth proxy and the true aggregate wealth portfolio and σ_p is the standard deviation of our aggregate wealth proxy. Given equations (D2) and (D3), the absolute magnitude of the bias in the risk-return coefficient estimate will be smaller using our proxy compared to stocks if

$$\rho_{p,aw}\frac{\sigma_{aw}}{\sigma_p} + \rho\frac{\sigma_{aw}}{\sigma_s} < 2. \tag{D4}$$

We first assess the characteristics of the measurement error term $\nu_{p,t}$ that are required to make our proxy perform worse than stocks in terms of bias in the estimated risk-return coefficient. In particular, we investigate combinations of the standard deviation of $\nu_{p,t}$ and the correlation between $R_{aw,t}$ and $\nu_{p,t}$ that may produce a larger bias for our measure relative to stocks. The true aggregate wealth portfolio is unobservable, but the correlation between the returns of our proxy and the stock market of 0.92 produces a correlation inequality,

$$\rho_{p,aw}\rho - \sqrt{(1-\rho_{p,aw}^2)(1-\rho^2)} \leq 0.92 \leq \rho_{p,aw}\rho + \sqrt{(1-\rho_{p,aw}^2)(1-\rho^2)}. \tag{D5}$$

This inequality, along with the following equations,

$$Var(R_{p,t}) = Var(R_{aw,t}) + Var(\nu_{p,t}) + 2\rho_{\nu,aw}\sigma(R_{aw,t})\sigma(\nu_{p,t}) \tag{D6}$$

and

$$Cov(R_{p,t}, R_{aw,t}) = Var(R_{aw,t}) + 2\rho_{\nu,aw}\sigma(R_{aw,t})\sigma(\nu_{p,t}), \tag{D7}$$

allow us to find all feasible combinations of $\sigma(\nu_{p,t})$ and $\rho_{\nu,aw}$ that satisfy the necessary conditions. Specifically, we first form a grid over ρ, $\rho_{\nu,aw}$, and $\sigma(\nu_{p,t})$. We then calculate $\sigma(R_{aw,t})$ and $\rho_{p,aw}$ using equations (D6) and (D7). If $\rho_{p,aw}$ satisfies the correlation inequality in equation (D5), we calculate the bias condition in equation (D4). Finally, for each value of $\rho_{\nu,aw}$, we record the smallest value of $\sigma(\nu_{p,t})$ for which the bias is larger with our measure compared to stocks (i.e., $\rho_{p,aw}\frac{\sigma_{aw}}{\sigma_p} + \rho\frac{\sigma_{aw}}{\sigma_s} > 2$) for some value of ρ.

Figure D.1 plots the lower bound on $\sigma(\nu_{p,t})$ as a function of $\rho_{\nu,aw}$. Our aggregate wealth proxy produces a smaller bias than stocks for all combinations of parameters when $\rho_{\nu,aw} > -0.799$, so the graph shows the lower bound for $\sigma(\nu_{p,t})$ over a range of $\rho_{\nu,aw}$ from -1.000 to -0.799. The smallest values of the lower bound occur for near-perfect negative correlations between the aggregate wealth portfolio returns and measurement error, and the bound achieves a minimum of 0.81% per month when $\rho_{\nu,aw} = -0.996$. Thus, our proxy will always outperform stocks in estimating the risk-return coefficient as long as the monthly standard deviation of measurement error is less than 0.81% when measurement error and the true aggregate wealth portfolio are nearly perfectly negatively correlated. For higher values of $\sigma(\nu_{p,t})$, other parameters determine whether stocks or our proxy would produce a smaller bias in the estimate of γ. The lower bound on $\sigma(\nu_{p,t})$ increases as $\rho_{\nu,aw}$ becomes less negative, and all correlations above -0.875 require the standard deviation of measurement error to exceed 1.00% per month in order for stocks to be a better aggregate wealth measure compared to our proxy. Thus, our aggregate wealth proxy is better than stocks for estimating γ unless measurement error is relatively large and highly negatively correlated with aggregate wealth portfolio returns.

We can also cast these results in terms of the true characteristics of the aggregate wealth portfolio. We use equations (D5) to (D7) and find the combinations of ρ and $\sigma(R_{aw,t})$ for which either our proxy or stocks dominate the alternative measure in terms of producing the smallest bias in the estimated coefficient of relative risk aversion. Figure D.2 shows the regions for which our proxy outperforms (gray), the stock market outperforms (black), and either proxy may perform better depending on other parameters (white). Our proxy is always superior when $\sigma(R_{aw,t})$ is less than 2.7% per month. The region in which the stock market works better is concentrated in high values of $\sigma(R_{aw,t})$ and ρ, such that the true aggregate wealth portfolio must be highly volatile and highly correlated with stocks in order for stocks to be a better aggregate wealth proxy compared to our measure. Estimates from the prior literature (e.g., Stambaugh (1982) and Lustig, Van Nieuwerburgh, and Verdelhan (2013)) are located in the gray region of the graph, in which our aggregate wealth proxy produces a smaller bias compared to stocks.

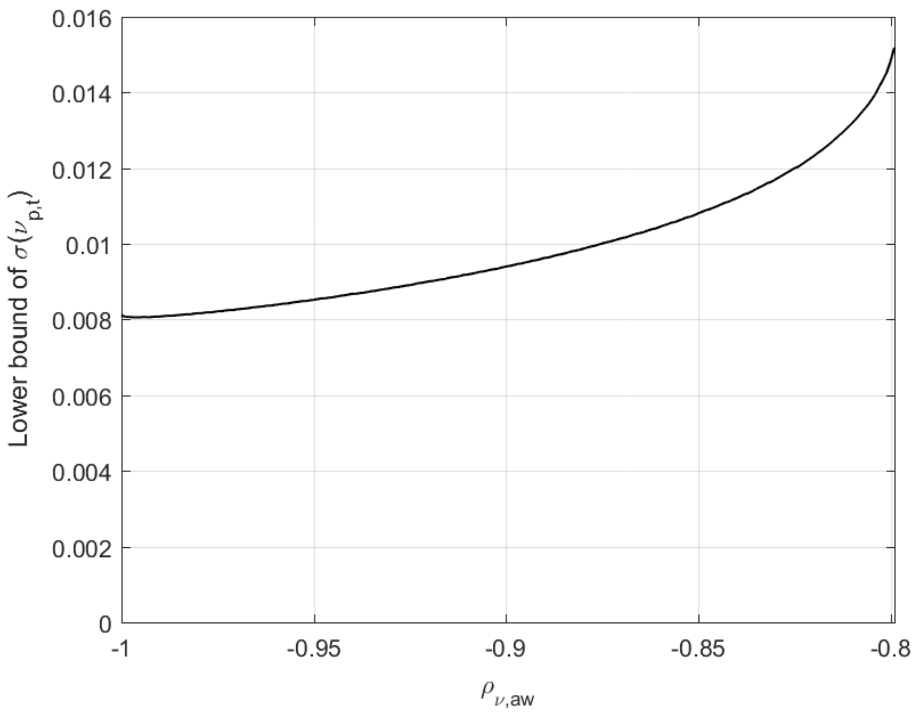

Figure D.1: Properties of Error in Our Proxy that Ensure Lower Bias than Stocks. The figure plots the smallest value of the monthly standard deviation of measurement error in our aggregate wealth proxy for which it is possible that using the stock market in risk-return tests could produce a lower magnitude of bias in the estimated coefficient of relative risk aversion. The lower bound of the standard deviation of measurement error is shown as a function of the correlation between the measurement error term and the return on the true aggregate wealth portfolio. Our proxy produces a smaller magnitude of bias for any standard deviation of measurement error below the plotted line and for any $\rho_{\nu,aw}$ above -0.799. The plotted lower bound is the smallest $\sigma(\nu_{p,t})$ for which stocks can produce a lower magnitude of bias for some value of the correlation between the returns of the stock market and aggregate wealth portfolios.

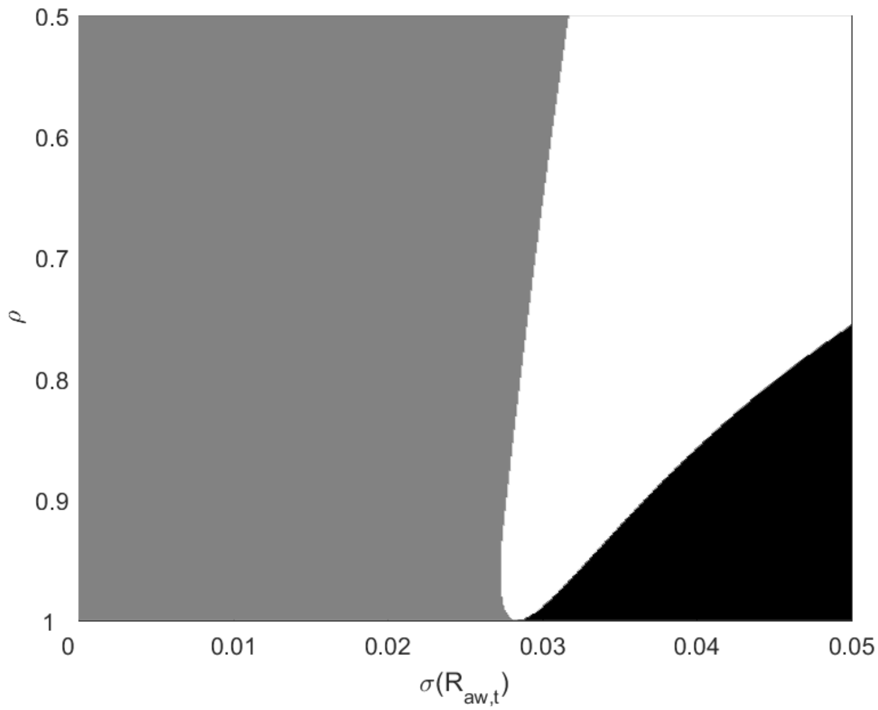

Figure D.2: Comparison of Magnitude of Bias for Our Proxy and Stocks across Aggregate Wealth Characteristics.

The figure shows the regions of the parameter space for which either our proxy or the stock market dominates the alternative measure in terms of producing the smallest magnitude of bias in the estimated coefficient of relative risk aversion. The parameter space is given by the monthly standard deviation of returns on the true aggregate wealth portfolio and the correlation between the returns of the stock market and the true aggregate wealth portfolio. Our proxy produces a smaller bias compared to stocks for any combination of aggregate wealth parameters in the gray region, whereas stocks produce a smaller bias in the black region. Results in the white region are mixed depending on the monthly standard deviation of measurement error in our proxy and the correlation between measurement error and true aggregate wealth portfolio returns.

E Robustness Tests

This appendix investigates the robustness of our empirical findings. Section E.1 considers several modifications to the aggregate wealth portfolio used in our main empirical tests. Section E.2 examines the impact of using alternative models to estimate market volatility on the relation between market return and variance presented in Table 2 in the paper. Sections E.3 and E.4 examine the robustness of our covariance-based tests reported in Tables 3 and 4, respectively. Section E.5 presents additional results to supplement our main asset pricing tests in Table 5.

E.1 Robustness of Results to Specification for Aggregate Wealth

An important theme of our paper is that researchers should be cognizant of the specification for the aggregate wealth proxy used in time-series tests of the ICAPM. We outline our baseline approach to estimating the returns for aggregate wealth in Section 4 in the paper but recognize that other reasonable alternatives are possible. In this section, we examine the robustness of our results to the specification of the market proxy. Table E.1 shows results corresponding to the risk-return tests in Section 6 of the paper for 12 alternative aggregate wealth portfolios. For each proxy, we report estimates of γ from the two-step regression using the conditional variance of aggregate wealth returns (i.e., Panel B of Table 2 in the paper), the two-step regression using the conditional covariance between stocks and aggregate wealth (i.e., Table 3), and the two-step regression using the conditional covariance between individual asset classes and aggregate wealth (i.e., Table 4). For the first two models, we also report the p-value associated with the one-sided test that the risk-return coefficient is less than the corresponding risk-return coefficient estimated from stocks (i.e., Panel A of Table 2). For brevity, we focus on the empirical specifications that include a lag of the dependent variable as an explanatory variable.

Panel A of Table E.1 shows the estimates for our base case of aggregate wealth. These estimates correspond to those in Tables 2 to 4 in the paper. In Panel B of Table E.1, we use alternative proxies for the returns of various asset classes. Case 2 uses the CRSP value-weighted index as a proxy for noncorporate equity returns. Moskowitz and Vissing-Jørgensen (2002) estimate private company returns over the 1990 to 1998 period and find a similarity between private firm returns and stock market returns, so we investigate the impact of using the CRSP index rather than small stocks to proxy for noncorporate equity. In Case 3, we use the Ibbotson Long-Term Corporate bond index as a proxy for defined-benefit pension returns rather than a weighted average of corporate and Treasury bond returns. Cases 4 and 5 include alternative measures for real estate returns.

Our base case uses CPI housing price data throughout the sample period. In Case 4, we begin to measure real estate returns using percentage changes in the deseasonalized Case–Shiller Home Price Index beginning in February 1987 when the data become available (http://us.spindices.com/index-family/real-estate/sp-case-shiller). Case 5 uses the FTSE NAREIT US Real Estate Index series for all real estate investment trusts (REITs) to proxy for real estate returns after the CPI reporting methods change in July 1985 (http://www.reit.com/investing/index-data). The index is initially available beginning in January 1972. The NAREIT index is about ten times more volatile than the CPI real estate series, which may reflect the relatively high levels of leverage employed by many REITs. To produce a comparable series, we form a portfolio of REITs and a 30-day Treasury bill such that the variance of this portfolio is equal to the variance of the real estate returns calculated from CPI data during the period of overlap in the two data samples. In particular, a portfolio that invests 11.4% in the REIT index and 88.6% in the risk-free asset matches the monthly standard deviation for the CPI return series of 0.65% over the period January 1972 to June 1985.

Panel C of Table E.1 includes aggregate wealth portfolio specifications in which the portfolio weights or considered asset classes differ from the base case. Case 6 uses the gross real estate asset value to weight real estate rather than the asset value net of mortgage liabilities. Cases 7 to 10 each exclude one or more asset classes. Cases 7 and 8 exclude real estate and defined benefit pensions, respectively, due to the relative difficulty of producing accurate monthly returns for these asset classes, and Case 9 excludes both series. Case 10 removes deposits, Treasury debt, agency debt, municipal debt, and corporate debt from the aggregate wealth portfolio. Some researchers have suggested that certain types of debt may not contribute to net wealth (e.g., Barro (1974)), so we investigate robustness to excluding all types of debt assets.

Case 11 in Panel D of Table E.1 is based on an aggregate wealth portfolio in which the portfolio weights of component asset classes are lagged by one year relative to the base case. Several quarterly figures from the Federal Reserve Flow of Funds tables are interpolated by the Federal Reserve based on annual observations. Although the representative investor is presumed to know the true weights, lagging all portfolio weights by one year eliminates any potential effects from lookahead bias induced by this interpolation procedure.

Finally, Case 12 in Panel E incorporates human capital in the definition of aggregate wealth. We base our measurement of monthly human capital returns on the method of Jagannathan and Wang (1996). Specifically, we measure labor income as the difference in total personal income and dividend income from Table 2.6 of the National Income and Product Accounts from the Bureau of

Economic Analysis. We then calculate the labor return using the formula,

$$R_t^{\text{labor}} = [L_{t-1} + L_{t-2}]/[L_{t-2} + L_{t-3}], \tag{E1}$$

where L_{t-1} is labor income for month $t-1$. Return data are available beginning in April 1959. In order to form the aggregate wealth portfolio, we also require an estimate of the aggregate value of human capital at the beginning of each quarter to calculate portfolio weights. We follow Jagannathan and Wang (1996) by making the assumption that labor income has both a constant growth rate (g) and a constant discount rate (r). Given this assumption, the value of human capital in month t is $L_t/(r-g)$. The quantity $(r-g)$ is unobservable, and we set it to 2.83%. Setting $(r-g)$ to this value causes the average weight of human capital in the aggregate wealth portfolio to be equal to 2/3, which is a reasonable value according to Campbell (1996).

Throughout Table E.1, we generally find results that are consistent with our inferences from the tests in Section 6 of the paper. The coefficient of relative risk aversion estimates exceed 4.20 in all cases, and 24 out of 36 coefficients are between seven and ten. The differences between the risk-return coefficient estimates for aggregate wealth and stocks are economically large, and 20 out of 24 are statistically significant at the 10% level. Overall, our inferences remain unchanged across a wide variety of aggregate wealth specifications.

E.2 Alternative Variance Estimates

Our tests in Table 2 in the paper relate expected returns to conditional variance estimates from a GARCH model using a two-step approach. Other studies (e.g., French, Schwert, and Stambaugh (1987), Baillie and DeGennaro (1990), and Lundblad (2007)) have adopted a similar empirical framework by testing the ICAPM's predicted risk-return relation for stocks using a GARCH model with mean effects (GARCH-M). Moreover, several alternatives to GARCH have been introduced in the literature to capture various features of return volatility. Two representative models are exponential GARCH (EGARCH) and fractionally integrated GARCH (FIGARCH). In this section, we examine the robustness of our results to using GARCH-M, EGARCH, and FIGARCH models of the volatility processes for stocks and aggregate wealth.

The GARCH-M model assumes that excess returns for market portfolio i follow

$$R_{i,t} = \mu + \gamma\sigma_{i,t}^2 + \lambda R_{i,t-1} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim N(0, \sigma_{i,t}^2), \tag{E2}$$

where the conditional variance $\sigma_{i,t}^2$ is given by

$$\sigma_{i,t}^2 = \omega + \alpha\varepsilon_{i,t-1}^2 + \delta\sigma_{i,t-1}^2. \tag{E3}$$

Equation (E2) allows the conditional expected portfolio return to depend on the conditional return variance, which makes the GARCH-M model a natural candidate for testing ICAPM predictions. As shown in equation (E3), conditional volatility evolves according to a GARCH(1,1) process.

EGARCH, introduced by Nelson (1991), is designed to allow for asymmetric responses of volatility to positive and negative return shocks. In particular, the log conditional variance is given by the process

$$\ln(\sigma_{i,t}^2) = \omega + \alpha\left[\left(\left|\frac{\varepsilon_{i,t-1}}{\sigma_{i,t-1}}\right| - E\left|\frac{\varepsilon_{i,t-1}}{\sigma_{i,t-1}}\right|\right) + \theta\frac{\varepsilon_{i,t-1}}{\sigma_{i,t-1}}\right] + \delta\ln(\sigma_{i,t-1}^2). \tag{E4}$$

The α parameter governs the importance of return shocks on volatility, the θ parameter allows volatility to respond differently to positive and negative shocks, and δ captures the persistent nature of variance. The asymmetry in responses is often viewed as a desirable feature of volatility models to help fit the "leverage effect," which is the tendency for volatility to increase after negative returns.

Baillie, Bollerslev, and Mikkelsen (1996) introduce FIGARCH to incorporate a long memory into conditional variance estimates. Volatility in the FIGARCH model follows

$$\sigma_{i,t}^2 = \omega + (1 - \beta L - \phi L(1-L)^d)\varepsilon_{i,t}^2 + \beta\sigma_{i,t-1}^2, \tag{E5}$$

where L is the lag operator. The current conditional volatility depends on lagged squared innovations with the influence of each lag decaying at a slow hyperbolic rate rather than the exponential decay implied by a standard GARCH model. The ϕ and d parameters govern the importance and rate of decay of lags, and β produces persistence in volatility.

We estimate GARCH-M model parameters in one step using MLE. For the EGARCH and FIGARCH models, we investigate the relations between expected returns and conditional volatility using the two-step approach from Table 2 in the paper. That is, we first estimate conditional variance $\hat{\sigma}_{i,t}^2$ for each model using MLE and then regress returns on the conditional volatility estimates.

Panel A of Table E.2 reports the GARCH-M estimates for the stock market and aggregate wealth portfolios. The GARCH(1,1) process parameters for each portfolio are quite close to the corresponding values in Table 2 in the paper. The risk-return coefficients for stocks in Cases 1 and 2 are 2.80 and 2.94. These estimates are slightly smaller than the corresponding figures in Table 2 and both are only marginally significant at the 10% level. As such, the statistical evidence for a positive risk-return relation for the stock market is somewhat weaker in the GARCH-M model relative to the two-step method in the paper. In contrast, evidence for a positive risk-return tradeoff for the aggregate wealth portfolio is slightly stronger using the GARCH-M approach. The estimates of γ in Cases 3 and 4 are 9.00 and 9.56, respectively, and each coefficient is significant at the 5% level.

Table E.2, Panel B examines the risk-return relation for the stock market and aggregate wealth portfolios using EGARCH to estimate variance. The large negative θ parameters for each portfolio imply that negative returns tend to produce larger increases in volatility compared to positive returns. In the specifications with no lagged returns, returns appear to be virtually unrelated to conditional variance estimates from the EGARCH model for stocks and insignificantly negatively related to volatility for aggregate wealth. The risk-return coefficient for stocks is 0.12 (standard error of 2.05), whereas the estimate of γ is -3.89 for aggregate wealth (standard error of 5.22). Including a lagged return has a substantial effect on inferences, however. After the lagged return is included, the risk-return coefficient for stocks is 1.79 (standard error of 2.01) with a p-value of 0.18. Aggregate wealth produces a γ estimate of 6.20 (standard error of 5.42) which is marginally insignificant at the 10% level. Results for EGARCH are therefore somewhat mixed, but the observed magnitudes of the risk-return coefficients are similar to those observed in Table 2 in the paper once we account for nonsynchronous return measurement.

The estimated risk-return relations using FIGARCH conditional variance estimates are shown in Panel C of Table E.2. The FIGARCH estimates for stocks indicate that volatility does not have a long memory. Resulting estimates are similar to the GARCH specifications from Table 2 in the paper. With no lagged return, the risk-return coefficient is 3.00 (p-value of 0.01). The aggregate wealth portfolio also has estimated variance parameters which do indicate long memory in volatility. The γ estimates are 5.12 (p-value of 0.12) and 6.59 (p-value of 0.06) in the two aggregate wealth specifications, so there is some evidence of a positive risk-return relation using FIGARCH to estimate volatility. Overall, the results in Table E.2 show that tests of the ICAPM have some sensitivity to the specification of volatility, which is consistent with the findings in prior literature. However, the risk-return coefficient estimates from the aggregate wealth portfolio are almost uniformly larger than the corresponding estimates from stocks which provides additional support for our basic conclusions.

E.3 Alternative Covariance Estimates: Stocks

In Section 6.2 in the paper, we examine the time-series relation between excess stock returns and the conditional covariance between stocks and aggregate wealth. In the first stage of this analysis, we model the conditional variance-covariance matrix of aggregate wealth returns and stock returns using a diagonal BEKK bivariate GARCH model. In this section, we examine the robustness of these results to using two alternative models of the conditional covariance process.

Our first approach is based on a scalar BEKK bivariate GARCH model for stocks and aggregate wealth. This model is a restricted version of the model in equation (B2) in which the α and δ parameters are scalars. Each element of the covariance matrix thus shares common GARCH parameters.

Our second method for estimating conditional covariances is based on the dynamic conditional correlation (DCC) model introduced by Engle (2002). DCC is a multivariate GARCH model which specifies a dynamic process for the conditional correlation between two series while the variance process for each series is GARCH. We model the conditional correlation between the stock market and aggregate wealth portfolio as

$$\rho_{s,aw,t} = (1 - \phi - \beta)\bar{\rho}_{s,aw} + \phi \frac{\varepsilon_{s,t-1}}{\sigma_{s,t-1}} \frac{\varepsilon_{aw,t-1}}{\sigma_{aw,t-1}} + \beta \rho_{s,aw,t-1}, \tag{E6}$$

and the GARCH parameter estimates for stocks and aggregate wealth are equal to those from Panel A of Table 2 and Panel B of Table 2, respectively. Given the DCC model estimates, we compute the conditional covariance as the product of conditional correlation and standard deviation terms, $\hat{\sigma}_{s,aw,t} = \hat{\rho}_{s,aw,t}\hat{\sigma}_{s,t}\hat{\sigma}_{aw,t}$.

After computing first-stage estimates of the conditional covariance between stocks and aggregate wealth using either the scalar BEKK GARCH or DCC method, we estimate the second-stage return regression in equation (18) in the paper to assess the risk-return tradeoff in the stock market. These results are provided in Table E.3. In each case, the estimates of the reward for risk are nearly identical to those presented in Panel B of Table 3 based on the diagonal BEKK GARCH approach. The estimates of γ in Table E.3 range from 6.98 to 7.75, and each coefficient is significantly greater than zero at the 10% level. As such, our findings in Table 3 appear robust to alternative modeling approaches for the conditional covariance process used to forecast stock market returns.

E.4 Alternative Covariance Estimates: Multiple Assets

Our next set of robustness checks are related to the risk-return tests in Section 6.3 using a panel of asset-class returns and covariances. For our primary results presented in Table 4 in the paper, we estimate the conditional covariance between each asset class and aggregate wealth separately using a diagonal BEKK bivariate GARCH model following equation (B2) and estimate the conditional variance of aggregate wealth using a GARCH(1,1) model. We then estimate the risk-return tradeoff via the panel regression in equation (20). In this section, we demonstrate the robustness of these results to alternative methods for obtaining the first-stage covariance estimates. The results are shown in Table E.4.

In Cases 1 and 2 (Cases 3 and 4), we estimate the conditional covariances between each asset class and the aggregate wealth portfolio jointly using a scalar (diagonal) BEKK multivariate GARCH model. These models are estimated from the demeaned multivariate excess return series for the aggregate wealth portfolio and the individual asset classes. We assume that the vector of demeaned excess returns in month t has a multivariate normal distribution with mean zero and variance-covariance matrix Σ_t. The BEKK GARCH model specifies that the conditional variance-covariance matrix evolves according to

$$\Sigma_t = \Omega + \alpha(\varepsilon_{t-1}\varepsilon'_{t-1})\alpha + \delta\Sigma_{t-1}\delta, \tag{E7}$$

where Ω is a symmetric $N \times N$ matrix. The GARCH parameters α and δ are scalars in the scalar BEKK GARCH model and $N \times N$ diagonal matrices in the diagonal BEKK GARCH model. Because the models used in Cases 1-4 involve the joint estimation of the variance-covariance matrix across assets, we exclude the asset classes with return series that do not cover the full 1953 to 2013 sample period (i.e., agency debt and municipal debt) from these tests. In Cases 5 and 6, we estimate the conditional covariance between each asset class and aggregate wealth separately using the DCC model outlined in Section E.3.

Table E.4 shows that our findings in Table 4 are robust to these alternative approaches to modeling conditional covariances across asset classes. The estimates of γ in Table E.4 range from 5.41 to 7.73, and five of the six coefficients are significant at the 5% level (the remaining coefficient is significant at the 10% level).

As in Table 4, we also test the null hypothesis that the intercepts are jointly equal to zero in each case. This test follows the double bootstrap design outlined in Section 6.3 of the paper, except that the first-level and second-level bootstrap samples are based on 1,000 rather than 10,000 draws to reduce computational cost. The results from these tests are in line with those from Table 4. We reject the null hypothesis of zero intercepts at the 5% level for the designs without a lagged dependent variable in the second-stage regression (i.e., Cases 1, 3, and 5). However, we fail to reject the null that the intercepts are jointly equal to zero in each of the remaining three cases.

E.5 Supplementary Asset Pricing Tests

Our main asset pricing results appear in Section 6.4 in the paper. We present a series of additional tests below to highlight the portfolio choice and asset pricing implications of using a broader proxy for aggregate wealth.

E.5.1 Portfolio Choice Applications

The primary empirical prediction of the CAPM regarding portfolio choice is that the market portfolio is mean-variance efficient. Although this condition is unlikely to be exactly satisfied in sample, time-series tests of the CAPM center around evaluating whether or not combining a given test asset (or set of test assets) with the specified market proxy leads to a statistically significant improvement in Sharpe ratio (e.g., Gibbons, Ross, and Shanken (1989)). In the paper, we demonstrate that our more comprehensive proxy for the market portfolio achieves a higher Sharpe ratio relative to the stock market proxy popular in prior literature. This result provides initial evidence that our market proxy is likely to be superior for portfolio choice and asset pricing applications.

Table E.5 formally evaluates whether or not individual assets classes are able to expand the investment opportunity set relative to each of the two market proxies (i.e., stocks and aggregate wealth) in the paper. We specifically estimate the following time-series regression using monthly data on excess returns for each combination of asset class (j) and market proxy ($i \in \{s, aw\}$):

$$R_{j,t} = \alpha_j + \beta_j R_{i,t} + e_{j,t}. \tag{E8}$$

We estimate these regressions individually for each of the following seven asset classes: noncorporate equity, Treasury debt, agency debt, municipal debt, corporate debt, pensions, and real estate. A significant estimate of α_j implies that the asset class on the left hand side of the regression expands the investment opportunity set relative to a given market proxy, in contrast to the prediction under the null hypotheses that $\alpha_j = 0$. We also consider regressions of each version of the market portfolio on the other to assess wether a given proxy “spans” the investment opportunities under the alternative specification.

Panel A of Table E.5 reports regression results using stocks as the market proxy. In the first column, we see that the aggregate wealth portfolio earns an alpha of 0.06% per month relative to stocks, and this estimate is statistically significant at the 10% level (p-value of 0.05). Thus, our more comprehensive measure of aggregate wealth expands the investment opportunity set relative to the stocks-only proxy that is popular in most CAPM applications. In the regressions for the seven individual asset classes, the intercepts are significantly different from zero at the 5% level in five cases. As such, there is considerable evidence in Panel A that the stock market portfolio fails to be mean-variance efficient in relation to the additional assets considered in the paper.

Panel B presents the corresponding regression estimates for the aggregate wealth market proxy. In the spanning test, the stock market portfolio earns a statistically insignificant alpha of just -0.02% per month (p-value of 0.73). In addition, only one alpha estimate across the seven individual asset classes (i.e., the intercept for agency debt) is positive and statistically significant at the 5% level. Taken as a whole, the results in Table E.5 suggest that the more comprehensive aggregate wealth proxy introduced in the paper strongly outperforms the stock market in portfolio choice applications.

E.5.2 Time-Series Tests

The tests in Table 5 in the paper apply a two-step, panel regression approach to examine how the specification of aggregate wealth impacts the ICAPM's ability to explain differences in average returns across decile portfolios sorted on size, book-to-market equity, and prior stock returns (i.e., momentum). In this section, we evaluate the performance of various models in pricing these test assets using standard time-series regression methods. We are particularly interested in how inferences might differ when we substitute our broader aggregate wealth portfolio return for the stock market portfolio return as an explanatory variable in a given asset pricing model.

As in Section 6.4, we consider several combinations of asset pricing models and sets of test portfolios. We specifically examine three alternative groups of value-weighted decile portfolios formed on size, book-to-market ratio, and momentum. We compare the ability of the following four asset pricing models to explain average returns for these test assets: (i) CAPM with stocks as the market proxy, (ii) CAPM with aggregate wealth as the market proxy, (iii) Fama-French (1993) three-factor model with stocks as the market proxy, and (iv) Fama-French (1993) three-factor model with aggregate wealth as the market proxy. Specifications (i) and (iii) correspond to how the CAPM and three-factor model are typically implemented in the literature, and comparing these specifications to models (ii) and (iv) highlights the asset pricing implications of our aggregate wealth portfolio. The CAPM regressions are identical to equation (E8), where $R_{j,t}$ now represents the excess return for equity portfolio j in month t. The Fama-French regressions are given by

$$R_{j,t} = \alpha_j + \beta_j R_{i,t} + s_j SMB_t + h_j HML_t + e_{j,t}, \tag{E9}$$

where SMB_t and HML_t are the monthly excess returns for the size and value factors, respectively.

We estimate the time-series regressions separately for each decile portfolio. The primary prediction for a well-specified asset pricing model is $\alpha_j = 0$. For a given model and set of test assets, Panel A of Table E.6 reports the regression intercepts for the top and bottom decile portfolios. We order the portfolios such that "decile 1" ("decile 10") corresponds to the set of stocks with lower (higher) expected alpha based on prior literature. Time-series tests commonly focus on the ability of a given model to explain spreads in abnormal returns across portfolios, so we also examine the alpha for a strategy that takes a long position in decile 10 and a short position in decile 1 in each case. Panel B presents mean absolute regression intercepts across the ten decile portfolios, bootstrap p-values for tests that the ten intercepts are jointly equal to zero, and average time-series regression R^2 values. The joint test is based on the following F statistic:

$$F = \frac{T-n-k}{n} \left[1 + \hat{\mu}_k' \hat{\Omega}_k^{-1} \hat{\mu}_k\right]^{-1} \hat{\alpha}' \hat{\Omega}^{-1} \hat{\alpha}, \tag{E10}$$

where T is the number of sample months, $n = 10$ is the number of test portfolios, k is the number of factors, $\hat{\mu}_k$ is a $k \times 1$ vector of factor mean returns, $\hat{\Omega}_k$ is the factor covariance matrix, $\hat{\alpha}$ is an $n \times 1$ vector of regression intercepts, and $\hat{\Omega}$ is the covariance matrix of regression residuals. We compute F for our sample and compare it to the bootstrap distribution under the null hypothesis of $\alpha = 0$. To estimate this distribution, we impose $\alpha_j = 0$ for all j by jointly resampling monthly adjusted excess portfolio returns and factor returns, where the adjusted excess return for portfolio j is $R^*_{j,t} = R_{j,t} - \hat{\alpha}_j$. The p-value in Panel B reflects the likelihood of observing an F statistic in the resamled data that exceeds the sample value of F.

The results in Table E.6 suggest that our aggregate wealth proxy positively impacts the CAPM's ability to explain returns on size-sorted portfolios. For example, the spread in alphas between the top and bottom size deciles is 0.22% per month (p-value of 0.09) relative to the traditional version of the CAPM in Case 1. In Panel B, the mean pricing error across the ten size portfolios is 0.12% per month and the F test produces a rejection of the null hypothesis that the ten intercepts are jointly zero at the 1% level. Case 7 substitutes the aggregate wealth portfolio for the stock market portfolio as the explanatory return in the CAPM tests. In contrast to Case 1, the high-minus-low alpha spread is essentially zero with a p-value of 0.51. The average pricing error in this case is also reduced to just 0.03% per month, and the corresponding p-value for the joint test is 0.88.

Introducing a broader market proxy in the time-series CAPM regressions also helps to account for patterns in average returns associated with book-to-market equity. The spread in CAPM alphas between value stocks and growth stocks in Case 3 is 0.39% per month (p-value of 0.01) relative to the traditional CAPM. This difference is reduced in magnitude by 23% in Case 9, although the alpha estimate of 0.30% per month remains statistically significant (p-value of 0.03). The mean pricing error across the ten book-to-market deciles also drops from 0.15% per month in Case 3 to 0.12% in Case 9. Moreover, the F statistic from the joint test in Case 9 with the broader proxy for aggregate wealth is only marginally significant (p-value just above 0.05).

In contrast, none of the four models considered exhibits particularly impressive performance in explaining average returns for the momentum portfolios. Across Cases 5-6 (stock market proxy) and 11-12 (aggregate wealth proxy), the high-minus-low portfolio alpha estimates are economically large and strongly significant. The F tests also produce rejections of the null that alphas are jointly zero for the momentum deciles, with p values of 0.00 in all four cases.

E.5.3 Cross-Sectional Tests

Table E.7 reports results from cross-sectional asset pricing tests for the same set of models considered in Table E.6. These tests provide additional perspective on the importance of the aggregate wealth specification in asset pricing applications. Furthermore, the cross-sectional regressions also allow us to evaluate the impact of including a broader market proxy on the explanatory power of other asset pricing factors (e.g., Fama and French's (1993) size and value factors, SMB and HML).

We examine each model's ability to capture variation in average return across the following two sets of test assets: (i) ten size portfolios, ten book-to-market portfolios, and ten momentum portfolios and (ii) ten size portfolios, 17 industry portfolios, and three bond portfolios. The size, book-to-market, and momentum deciles in group (i) are identical to the value-weighted test assets used in Table E.6. However, in the cross-sectional tests, we now require each model to price these assets simultaneously. We also introduce the portfolios in group (ii) to address Lewellen, Nagel, and Shanken's (2010) critique of evaluating asset pricing models solely on their ability to explain returns for anomaly-sorted portfolios in cross-sectional tests. This specific combination of test portfolios is similar in spirit to the one used in Campbell (1996). The 17 industry portfolios are value weighted, and these return series are from Kenneth French's website. Returns data for the three industry portfolios are from the Ibbotson SBBI Classic Yearbook. We include the series for long-term government bonds, intermediate-term government bonds, and long-term corporate bonds. We convert all returns in the asset pricing tests to excess returns by subtracting the 30-day Government Bond index return from Ibbotson.

The cross-sectional regression model is given by

$$\bar{R}_j = \lambda_0 + \lambda_1 \hat{\beta}_j + \lambda_2 \hat{s}_j + \lambda_3 \hat{h}_j + u_j, \tag{E11}$$

where $\bar{R}_j$ is the average excess return for portfolio j and $\hat{\beta}_j$, $\hat{s}_j$, and $\hat{h}_j$ are portfolio j's estimated factor loadings for the market, size, and value factors, respectively. We consider two models—the CAPM and the Fama-French three-factor model. In each case, we compare the performance of these models with stocks and aggregate wealth as the market proxy. The factor loadings for each portfolio are estimated using time-series regressions over the full sample period.

Panel A of Table E.7 reports parameter estimates for each model, including the zero beta rate (i.e., λ_0) and factor prices of risk (i.e., λ_1, λ_2, and λ_3). A well-specified model should have an estimate of λ_0 that is statistically insignificant and economically small, estimated risk prices that are in line with theoretical predictions, and pricing errors (i.e., u_j for $j = 1, ..., n$) near zero. In Panel B, we compare pricing errors across models by considering the cross-sectional regression R^2 and mean absolute pricing error (MAPE). The latter is given by

$$\text{MAPE} = \frac{1}{n} \sum_{j=1}^{n} |\hat{u}_j|. \tag{E12}$$

Following Goyal (2012), we also consider a joint test of whether all the pricing errors are zero. In our application, the factors are all traded portfolios, and the test statistic is

$$T(1 + \hat{\lambda}' \hat{\Omega}_k^{-1} \hat{\lambda}) \hat{u}' \hat{\Omega}^{-1} \hat{u} \sim \chi^2_{n-k}, \tag{E13}$$

where $\hat{\lambda}$ is a vector of the factor risk prices, $\hat{\Omega}_k$ is the factor covariance matrix, $\hat{u}$ is the vector of pricing errors, and $\hat{\Omega}$ is the covariance matrix of residuals from the first-stage, time-series regressions.

Cases 1-4 in Table E.7 examine the ability of the models to price size, book-to-market, and momentum portfolios. Prior literature suggests that the CAPM performs poorly in explaining differences in average returns across these test assets, and we see additional evidence of this conclusion in Case 1. The estimated zero-beta rate is 1.02% per month, and the point estimate for the price of market risk is actually negative. Moreover, the pricing errors are large as evidenced by the MAPE of 0.17% per month and the low cross-sectional R^2 value of 3.09%. Case 3 presents CAPM results using the broader aggregate wealth proxy. This version of the model performs somewhat better than the traditional CAPM specification in Case 1. The cross-sectional regression intercept falls to 0.62% per month, the price of market risk turns positive (but remains statistically insignificant), and the MAPE is lower at 0.16% per month. Cases 2 and 4 add the size and value factors to the specifications in Cases 1 and 3, respectively. Not surprisingly, incorporating these factors leads to meaningful increases in R^2 values and large drops in MAPE. The risk prices for SMB and HML are both positive, and the estimates for the size factor are statistically significant. These models still appear somewhat misspecified, however, as the zero-beta rate exceeds 2% per month in each case. For all four models, the χ^2 test leads to a rejection of the null hypothesis that the pricing errors are jointly equal to zero across test assets.

Cases 5-8 show results for the size, industry, and bond portfolios. Both versions of the CAPM (i.e., Cases 5 and 7) appear to explain patterns in returns for these test assets. The price of market risk is positive and statistically significant in each case, and each of these two models produces an economically reasonable estimate for the zero-beta rate of 0.22% per month. However, the model in Case 7 using our broader definition of aggregate wealth delivers superior pricing performance relative to the traditional CAPM specification in Case 5. For example, the cross-sectional R^2 value is 63.1% in Case 7 versus 61.3% in Case 5, and the MAPE is also better at 0.08% versus 0.09%. The MAPE in Case 7 is actually slightly lower than the average absolute pricing errors produced by the two versions of the three-factor model (i.e., Cases 6 and 8) among these test portfolios. Finally, although the aggregate wealth version of the CAPM in Case 7 is formally rejected in the χ^2 test, the χ^2 value of 52.4 is substantially lower than the values produced by the models in Cases 5, 6, and 8.

Table E.1: Robustness Tests for the Specification of the Aggregate Wealth Portfolio, 1953–2013.
The table demonstrates the robustness of our results to the definition of aggregate wealth. The table presents results corresponding to the two-step regressions using the conditional variance of aggregate wealth returns (i.e., Panel B of Table 2), the two-step regressions using the conditional covariance between stocks and aggregate wealth (i.e., Table 3), and the two-step regressions using the conditional covariance between individual asset classes and aggregate wealth (i.e., Table 4). All of the empirical specifications include a return lag for the dependent variable. The table presents estimates of γ and the associated standard errors (in parentheses) using the various approaches. For each of the first two methods, the table also shows a bootstrap p-value ($p(\gamma \leq \gamma_s)$) for the one-sided test that the the risk-return coefficient is less than the corresponding risk-return coefficient estimated from stocks. For comparison, Panel A reports results corresponding to our original specification of the aggregate wealth portfolio. Panel B presents results corresponding to aggregate wealth portfolios with alternative proxies for individual component returns. Panel C provides results corresponding to aggregate wealth portfolios with modified estimates of asset-class weights. Panel D shows results corresponding to the original specification of the aggregate wealth portfolio with asset-class weights lagged by 12 months to account for reporting delays and interpolation in the Federal Reserve Flow of Funds data. Panel E reports results corresponding to a measure of aggregate wealth that accounts for human capital.

		$\sigma^2_{aw,t}$		$\sigma_{j,aw,t}$, Stocks		$\sigma_{j,aw,t}$, All
Case	Aggregate wealth modification	γ	$p(\gamma \leq \gamma_s)$	γ	$p(\gamma \leq \gamma_s)$	γ
	Panel A: Base specification for aggregat wealth					
(1)	Base case	8.523	0.066	7.628	0.072	7.534
		(4.435)		(4.636)		(3.946)
	Panel B: Modifications to return proxies					
(2)	Noncorporate equity: CRSP value-weighted index	8.359	0.112	6.325	0.095	8.828
		(5.679)		(4.081)		(4.108)
(3)	DB pensions: Ibbotson long-term corporate bonds	7.502	0.099	7.314	0.092	7.205
		(4.280)		(4.739)		(3.857)
(4)	Real estate: Case-Shiller index returns	8.160	0.080	7.645	0.079	7.480
		(4.468)		(4.754)		(3.935)
(5)	Real estate: NAREIT index returns	8.967	0.058	7.951	0.064	7.761
		(4.649)		(4.754)		(3.957)
	Panel C: Modifications to portfolio weights					
(6)	Gross real estate holdings	8.843	0.073	8.255	0.073	7.928
		(4.732)		(5.073)		(4.186)
(7)	Exclude real estate	7.163	0.066	6.140	0.074	6.034
		(3.424)		(3.628)		(3.041)
(8)	Exclude DB pensions	7.468	0.056	5.424	0.112	5.953
		(3.673)		(3.372)		(3.004)
(9)	Exclude real estate and DB pensions	6.004	0.063	4.210	0.180	4.636
		(2.691)		(2.542)		(2.241)
(10)	Exclude deposits and debt securities	7.609	0.043	6.084	0.062	6.305
		(3.499)		(3.444)		(3.069)
	Panel D: Lagged portfolio weights					
(11)	Base case with weights lagged by 12 months	7.783	0.065	7.091	0.075	7.259
		(4.057)		(4.337)		(3.716)
	Panel E: Account for human capital					
(12)	Add human capital returns	9.877	0.379	29.507	0.071	24.389
		(21.583)		(19.071)		(14.893)

Table E.2: The Risk-Return Tradeoff with Alternative Volatility Specifications, 1953–2013.
The table presents evidence on the relation between excess return and volatility for the stock market portfolio and the aggregate wealth portfolio using a GARCH-M specification (Panel A) and two-step estimation approach (Panels B and C). For Panel A, we estimate the GARCH-M parameters using maximum likelihood estimation (MLE). We report standard errors in parentheses and a p-value for the one-sided test of the null hypothesis that the-risk-return parameter is less than zero in brackets. The R^2 value in each case is the total R^2 for the GARCH-M model. The first-stage results in Panel B (Panel C) are EGARCH(1,1) (FIGARCH(1,1)) parameter estimates for each portfolio using demeaned excess returns. We estimate the EGARCH(1,1) and FIGARCH(1,1) parameters using MLE. The second-stage return regression for stocks (aggregate wealth) is given by $R_{s,t} = \mu + \gamma_s \hat{\sigma}^2_{s,t} + \lambda R_{s,t-1} + \eta_{s,t}$ ($R_{aw,t} = \mu + \gamma \hat{\sigma}^2_{aw,t} + \lambda R_{aw,t-1} + \eta_{aw,t}$). For these regressions, we report bootstrap standard errors in parentheses and a bootstrap p-value for the one-sided test of the null hypothesis that the-risk-return parameter is less than zero in brackets. The R^2 value in each case is the second-stage regression R^2.

Panel A: GARCH-M specification

	GARCH(1,1) parameters			Return regression parameters			
Case	$\omega \times 10^4$	α	δ	$\mu \times 10^2$	γ	λ	R^2 (%)
			Stock market portfolio				
(1)	0.997	0.114	0.838	0.217	**2.800**		0.34
	(0.411)	(0.027)	(0.032)	(0.385)	**(2.168)**		
					[0.098]		
(2)	0.955	0.113	0.841	0.154	**2.938**	0.076	1.12
	(0.394)	(0.027)	(0.031)	(0.397)	**(2.291)**	(0.045)	
					[0.100]		
			Aggregate wealth portfolio				
(3)	0.264	0.110	0.822	0.029	**8.996**		0.24
	(0.094)	(0.026)	(0.040)	(0.170)	**(4.776)**		
					[0.030]		
(4)	0.239	0.112	0.826	−0.034	**9.556**	0.194	3.33
	(0.087)	(0.026)	(0.039)	(0.167)	**(4.925)**	(0.042)	
					[0.026]		

Panel B: EGARCH(1,1) specification

	First-stage EGARCH(1,1) parameters				Second-stage return regression parameters			
Case	ω	α	δ	θ	$\mu \times 10^2$	γ	λ	R^2
				Stock market portfolio				
(5)	−0.817	0.199	0.872	−0.856	0.556	**0.117**		0.00
	(0.238)	(0.071)	(0.037)	(0.346)	(0.371)	**(2.049)**		
						[0.478]		
(6)					0.173	**1.791**	0.109	0.96
					(0.385)	**(2.011)**	(0.043)	
						[0.182]		
				Aggregate wealth portfolio				
(7)	−1.130	0.145	0.858	−1.021	0.437	**−3.886**		0.13
	(0.132)	(0.045)	(0.017)	(0.338)	(0.181)	**(5.220)**		
						[0.773]		
(8)					0.002	**6.198**	0.202	3.33
					(0.199)	**(5.417)**	(0.044)	
						[0.126]		

(continued)

Table E.2—*Continued*

Panel C: FIGARCH(1,1) specification								
	First-stage FIGARCH(1,1) parameters				Second-stage return regression parameters			
Case	$\omega \times 10^4$	ϕ	d	β	$\mu \times 10^2$	γ	λ	R^2
				Stock market portfolio				
(9)	2.543	0.000	0.810	0.738	−0.058	**2.997**		0.89
	(0.905)	(0.000)	(0.151)	(0.137)	(0.288)	**(1.339)**		
						[0.014]		
(10)					−0.092	**2.930**	0.084	1.60
					(0.287)	**(1.341)**	(0.045)	
						[0.016]		
				Aggregate wealth portfolio				
(11)	0.583	0.316	0.369	0.553	0.092	**5.124**		0.25
	(0.324)	(0.135)	(0.213)	(0.198)	(0.164)	**(4.369)**		
						[0.121]		
(12)					−0.020	**6.591**	0.180	3.47
					(0.158)	**(4.168)**	(0.042)	
						[0.056]		

Table E.3: Robustness Tests for the Risk-Return Tradeoff for the Stock Market Portfolio: Covariance Approach, 1953–2013.

The table presents evidence on the relation between stock market portfolio return and conditional covariance between stock returns and aggregate wealth returns using a two-step estimation approach. Each panel shows results from the following second-stage regression of stock market excess return, $R_{s,t}$, on the conditional covariance between stock returns and aggregate wealth returns, $\hat{\sigma}_{s,aw,t}$, and lagged excess stock returns, $R_{s,t-1}$: $R_{s,t} = \mu + \gamma\hat{\sigma}_{s,aw,t} + \lambda R_{s,t-1} + \eta_{s,t}$. We report bootstrap standard errors in parentheses and a bootstrap p-value for the one-sided test of the null hypothesis that $\gamma \leq 0$ in brackets. The R^2 value in each case is the second-stage regression R^2. In Panel A, the conditional covariance estimates are from a scalar BEKK GARCH model. We estimate these parameters using the demeaned bivariate excess return series for the aggregate wealth portfolio and the stock market portfolio. We assume that the vector of demeaned excess returns in month t, $[R_{aw,t} - \bar{R}_{aw} \quad R_{s,t} - \bar{R}_s]'$, has a bivariate normal distribution with mean zero and variance-covariance matrix Σ_t. The conditional variance-covariance matrix evolves according to $\Sigma_t = \Omega + \alpha(\varepsilon_{t-1}\varepsilon_{t-1}')\alpha + \delta\Sigma_{t-1}\delta$, where Ω is a symmetric 2×2 matrix and α and δ are scalars. In Panel B, the conditional covariance estimates are from a dynamic conditional correlation (DCC) model for the stock market and aggregate wealth. The conditional correlation between these two series is given by $\rho_{s,aw,t} = (1 - \phi - \beta)\bar{\rho}_{s,aw} + \phi\frac{\varepsilon_{s,t-1}}{\sigma_{s,t-1}}\frac{\varepsilon_{aw,t-1}}{\sigma_{aw,t-1}} + \beta\rho_{s,aw,t-1}$. The conditional covariance estimate is given by the product of the conditional correlation and standard deviation terms, $\hat{\sigma}_{s,aw,t} = \hat{\rho}_{s,aw,t}\hat{\sigma}_{s,t}\hat{\sigma}_{aw,t}$, where the standard deviations are estimated from GARCH(1,1) models.

	Return regression parameters			
Case	$\mu \times 10^2$	γ	λ	R^2 (%)
Panel A: Scalar BEKK GARCH approach				
(1)	0.010	**7.012**		0.41
	(0.373)	**(4.726)**		
		[0.071]		
(2)	−0.081	**7.501**	0.090	1.22
	(0.362)	**(4.539)**	(0.045)	
		[0.050]		
Panel B: DCC approach				
(3)	0.022	**6.978**		0.37
	(0.401)	**(5.171)**		
		[0.090]		
(4)	−0.091	**7.745**	0.092	1.21
	(0.384)	**(4.900)**	(0.044)	
		[0.057]		

Table E.4: Robustness Tests for the Risk-Return Tradeoff Across Asset Classes: Covariance Approach, 1953–2013.
The table presents evidence on the risk-return relation across asset classes using a two-step estimation approach. In each case, the second-stage regression of excess returns for individual asset classes and the aggregate wealth portfolio on the conditional covariance between asset-class returns and aggregate wealth returns and lagged excess returns is given by $R_{j,t} = \mu_j + \gamma \hat{\sigma}_{j,aw,t} + \lambda_j R_{j,t-1} + \eta_{j,t}$. In Cases 1 and 2 (Cases 3 and 4), the conditional covariance estimates are from a scalar (diagonal) BEKK multivariate GARCH model. In Cases 5 and 6, the conditional covariance estimates are from pairwise dynamic conditional correlation (DCC) models for each asset class and aggregate wealth. For each model, we also estimate the conditional variance for the aggregate wealth portfolio using the GARCH(1,1) model in Panel B of Table 2. In Panel A, we report intercepts for each asset class, and the corresponding bootstrap standard errors are shown in parentheses. For each model, Panel B presents an estimate of the risk-reward coefficient, γ, with the corresponding bootstrap standard error in parentheses. A bootstrap p-value for the one-sided test of the null hypothesis that $\gamma \leq 0$ is given in brackets. Panel B also provides a bootstrap p-value, $p(\mu_1 = \cdots = \mu_9 = 0)$, for a test of the null hypothesis that the intercepts are jointly equal to zero. The R^2 value in each case is the second-stage regression R^2. We exclude agency debt and municipal debt from the analyses in Cases 1-4 as these return series do not cover the full 1953 to 2013 sample period.

	Scalar BEKK-GARCH approach		Diagonal BEKK-GARCH approach		Pairwise DCC approach	
Case	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Intercepts from second-stage return regressions						
	$\mu_j \times 10^2$					
Corporate equity	0.154	0.096	0.154	0.097	0.011	−0.089
	(0.250)	(0.251)	(0.250)	(0.250)	(0.350)	(0.339)
Noncorporate equity	0.228	0.009	0.222	0.003	0.043	−0.236
	(0.344)	(0.335)	(0.347)	(0.338)	(0.476)	(0.454)
Treasury debt	0.101	0.081	0.102	0.082	0.110	0.088
	(0.063)	(0.062)	(0.063)	(0.061)	(0.062)	(0.061)
Agency debt	n/a	n/a	n/a	n/a	0.160	0.118
					(0.067)	(0.067)
Municipal debt	n/a	n/a	n/a	n/a	0.190	0.170
					(0.094)	(0.095)
Corporate debt	0.084	0.059	0.084	0.059	0.077	0.044
	(0.101)	(0.097)	(0.101)	(0.097)	(0.106)	(0.101)
DB pensions	0.160	0.134	0.160	0.134	0.162	0.136
	(0.060)	(0.060)	(0.060)	(0.060)	(0.061)	(0.060)
Real estate	−0.098	−0.077	−0.098	−0.077	−0.092	−0.075
	(0.037)	(0.035)	(0.037)	(0.034)	(0.037)	(0.034)
Aggregate wealth	0.081	0.024	0.081	0.024	0.021	−0.057
	(0.118)	(0.116)	(0.118)	(0.116)	(0.162)	(0.156)
Panel B: Additional details for second-stage return regressions						
Risk aversion coefficient						
γ	5.410	5.503	5.513	5.598	7.107	7.725
	(3.012)	(2.919)	(3.063)	(2.967)	(4.412)	(4.213)
	[0.037]	[0.029]	[0.037]	[0.030]	[0.056]	[0.033]
Joint test of intercepts						
$p(\mu_1 = \cdots = \mu_9 = 0)$	0.023	0.170	0.022	0.168	0.028	0.201
Model fit						
R^2 (%)	1.11	4.92	1.11	4.92	1.06	4.87
Control for lagged returns						
Include $R_{j,t-1}$?	No	Yes	No	Yes	No	Yes

Table E.5: Time-Series CAPM Regressions for Individual Asset Classes, 1953–2013.
The table presents results from time-series regressions of excess returns for individual asset classes on excess returns for stocks (Panel A) and aggregate wealth (Panel B). The regression model is given by $R_{j,t} = \alpha_j + \beta_j R_{i,t} + e_{j,t}$, where $R_{j,t}$ is the monthly excess return for asset class j and $R_{i,t}$ is the monthly excess return for market portfolio i. For each regression, the table presents parameter estimates with the corresponding bootstrap standard errors in parentheses. A bootstrap p-value for the two-sided test of the null hypothesis that $\alpha_j = 0$ is shown in brackets. The R^2 value in each case is the time-series regression R^2.

	Spanning tests for market proxies		Individual asset-class regressions						
	Aggregate wealth portfolio	Corporate equity	Noncorporate equity	Treasury debt	Agency debt	Municipal debt	Corporate debt	DB pensions	Real estate
Panel A: Stock market proxy									
Alpha, $\alpha_j \times 10^2$	0.056	0.000	0.201	0.129	0.181	0.211	0.112	0.126	−0.036
	(0.029)	n/a	(0.141)	(0.060)	(0.065)	(0.091)	(0.093)	(0.061)	(0.015)
	[0.052]	n/a	[0.149]	[0.031]	[0.006]	[0.020]	[0.231]	[0.040]	[0.013]
Beta, β_j	0.407	1.000	1.095	0.035	0.041	0.050	0.138	0.060	−0.002
	(0.009)	n/a	(0.040)	(0.016)	(0.017)	(0.024)	(0.027)	(0.017)	(0.003)
R^2 (%)	83.82	100.00	60.58	0.89	1.98	2.08	5.76	2.55	0.07
Panel B: Aggregate wealth proxy									
Alpha, $\alpha_j \times 10^2$	0.000	−0.022	−0.012	0.100	0.152	0.170	0.062	0.094	−0.040
	n/a	(0.064)	(0.089)	(0.059)	(0.064)	(0.091)	(0.089)	(0.060)	(0.015)
	n/a	[0.733]	[0.895]	[0.091]	[0.018]	[0.061]	[0.487]	[0.116]	[0.007]
Beta, β_j	1.000	2.057	2.894	0.168	0.192	0.229	0.444	0.230	0.006
	n/a	(0.048)	(0.057)	(0.037)	(0.046)	(0.059)	(0.056)	(0.037)	(0.008)
R^2 (%)	100.00	83.82	83.85	4.10	7.50	7.34	11.81	7.38	0.09

Table E.6: Time-Series CAPM Regressions for Anomaly Portfolios, 1953–2013.
The table presents results from time-series asset pricing regressions for value-weighted decile portfolios sorted on size, book-to-market, and momentum. We consider two market proxies—stocks and aggregate wealth. For each market proxy and set of decile portfolios, we estimate ten CAPM regressions ("CAPM") and ten Fama-French (1993) three-factor model regressions ("FF3"). The CAPM regressions are given by $R_{j,t} = \alpha_j + \beta_j R_{i,t} + e_{j,t}$, where $R_{j,t}$ is the monthly excess return for decile portfolio j and $R_{i,t}$ is the monthly excess return for market portfolio i. The Fama-French regressions are given by $R_{j,t} = \alpha_j + \beta_j R_{i,t} + s_j SMB_t + h_j HML_t + e_{j,t}$. For each set of anomaly portfolios, Decile 10 (1) represents the group of stocks with higher (lower) expected returns based on prior literature. In Panel A, we report intercepts for the top decile, bottom decile, and their difference. The corresponding bootstrap standard errors are shown in parentheses, and the figure in brackets is a bootstrap p-value for the one-sided test of the null hypothesis that $\alpha_{10} - \alpha_1 \leq 0$. Panel B reports the average absolute regression intercept across decile portfolios and also provides a bootstrap p-value, $p(\alpha_1 = \cdots = \alpha_{10} = 0)$, for a test of the null hypothesis that the intercepts across the decile portfolios are jointly equal to zero. The R^2 value reported in the table is the average time-series regression R^2 for the ten decile portfolios.

	Market proxy: Stock market portfolio						Market proxy: Aggregate wealth portfolio					
	Size portfolios		Book-to-market portfolios		Momentum portfolios		Size portfolios		Book-to-market portfolios		Momentum portfolios	
Model	CAPM	FF3	CAPM	FF3	CAPM	FF3	CAPM	FF3	CAPM	FF3	CAPM	FF3
Case	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Intercepts from time-series regressions												
	$\alpha_j \times 10^2$											
Decile 1	−0.019	0.034	−0.121	0.109	−0.935	−1.082	−0.008	0.087	−0.123	0.161	−1.025	−1.008
	(0.034)	(0.017)	(0.068)	(0.049)	(0.162)	(0.163)	(0.082)	(0.055)	(0.104)	(0.072)	(0.170)	(0.173)
Decile 10	0.201	−0.051	0.273	−0.201	0.525	0.618	−0.012	−0.049	0.180	−0.129	0.465	0.677
	(0.140)	(0.065)	(0.123)	(0.073)	(0.119)	(0.109)	(0.089)	(0.059)	(0.123)	(0.098)	(0.128)	(0.125)
Decile 10−Decile 1	0.220	−0.084	0.394	−0.310	1.461	1.701	−0.004	−0.135	0.303	−0.290	1.491	1.684
	(0.167)	(0.065)	(0.169)	(0.084)	(0.233)	(0.241)	(0.145)	(0.065)	(0.165)	(0.084)	(0.231)	(0.238)
	[0.090]	[0.905]	[0.010]	[1.000]	[0.000]	[0.000]	[0.510]	[0.982]	[0.033]	[1.000]	[0.000]	[0.000]
Panel B: Additional details for time-series regressions												
Absolute pricing error												
Mean $\lvert\alpha_j\rvert \times 10^2$	0.117	0.029	0.154	0.070	0.259	0.319	0.031	0.061	0.122	0.082	0.266	0.306
Joint test of pricing errors												
$p(\alpha_1 = \cdots = \alpha_{10} = 0)$	0.005	0.024	0.008	0.008	0.000	0.000	0.877	0.178	0.052	0.004	0.000	0.000
Model fit												
Mean R_j^2 (%)	83.47	96.06	82.02	89.72	78.70	80.54	83.96	89.01	71.09	80.70	67.01	72.37

Table E.7: Cross-Sectional Regressions, 1953–2013.

The table presents cross-sectional asset pricing tests for the CAPM and Fama-French (1993) three-factor model. We consider two market proxies—stocks and aggregate wealth. We also consider the following two sets of test assets: (1) ten size portfolios, ten book-to-market portfolios, and ten momentum portfolios and (2) ten size portfolios, 17 industry portfolios, and three bond portfolios. The cross-sectional regression model is given by $\bar{R}_j = \lambda_0 + \lambda_1 \hat{\beta}_j + \lambda_2 \hat{s}_j + \lambda_3 \hat{h}_j + u_j$, where $\bar{R}_j$ is the average excess return for portfolio j and $\hat{\beta}_j$, $\hat{s}_j$, and $\hat{h}_j$ are portfolio j's estimated factor loadings for the market, size, and value factors, respectively. Panel A reports parameter estimates with the associated Shanken (1992) standard errors in parentheses. These parameter estimates are expressed in percentage per month. For each model, Panel B presents the cross-sectional regression R^2, the average absolute regression residual in percentage per month, and results from a χ^2 test of whether the pricing errors are jointly equal to zero.

	Test assets: Size, B/M, & momentum portfolios				Test assets: Size, industry, and bond portfolios			
Case	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Parameter estimates from cross-sectional regressions								
Intercept	1.020	2.093	0.623	2.282	0.219	0.251	0.219	0.205
	(0.296)	(0.377)	(0.232)	(0.421)	(0.093)	(0.088)	(0.099)	(0.099)
Market (Stocks)	−0.338	−1.462			0.435	0.399		
	(0.341)	(0.406)			(0.194)	(0.186)		
Market (Aggregate wealth)			0.021	−0.653			0.200	0.205
			(0.119)	(0.194)			(0.089)	(0.094)
SMB		0.223		0.262		0.148		0.136
		(0.112)		(0.113)		(0.113)		(0.113)
HML		0.142		0.143		−0.155		−0.155
		(0.110)		(0.110)		(0.142)		(0.141)
Panel B: Additional details for cross-sectional regressions								
R^2 (%)	3.09	38.82	0.11	42.19	61.30	63.77	63.11	64.56
Mean absolute pricing error (%)	0.170	0.118	0.157	0.113	0.089	0.080	0.078	0.079
$T^2(\chi^2_{n-k})$	261.53	238.48	87.88	249.99	110.18	134.61	52.42	85.16
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.005]	[0.000]

References

Baillie, Richard T., Tim Bollerslev, and Hans Ole Mikkelsen, 1996, Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 74, 3–30.

Baillie, Richard T., and Ramon P. DeGennaro, 1990, Stock returns and volatility, *Journal of Financial and Quantitative Analysis* 25, 203–214.

Barro, Robert J., 1974, Are government bonds net wealth?, *Journal of Political Economy* 82, 1095–1117.

Baxter, Marianne, and Urban J. Jermann, 1997, The international diversification puzzle is worse than you think, *American Economic Review* 87, 170–180.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2007, Portfolio choice over the life-cycle when the stock and labor markets are cointegrated, *Journal of Finance* 62, 2123–2167.

Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.

Chan, K.C., G. Andrew Karolyi, and René M. Stulz, 1992, Global financial markets and the risk premium on U.S. equity, *Journal of Financial Economics* 32, 137–167.

Engle, Robert, 2002, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics* 20, 339–350.

Engle, Robert, and Kenneth F. Kroner, 1995, Multivariate simultaneous GARCH, *Econometric Theory* 11, 122–150.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–53.

Fama, Eugene F., and G. William Schwert, 1977a, Asset returns and inflation, *Journal of Financial Economics* 5, 115–146.

——— , 1977b, Human capital and capital market equilibrium, *Journal of Financial Economics* 4, 95–125.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 87, 1121–1152.

Goyal, Amit, 2012, Empirical cross-sectional asset pricing: A survey, *Financial Markets and Portfolio Management* 26, 3–38.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset-pricing tests, *Journal of Financial Economics* 96, 175–194.

Lundblad, Christian, 2007, The risk return tradeoff in the long run: 1836-2003, *Journal of Financial Economics* 85, 123–150.

Lustig, Hanno, and Stijn Van Nieuwerburgh, 2008, The returns on human capital: Good news on Wall Street is bad news on Main Street, *Review of Financial Studies* 21, 2097–2137.

Lustig, Hanno, Stijn Van Nieuwerburgh, and Adrien Verdelhan, 2013, The wealth-consumption ratio, *Review of Asset Pricing Studies* 3, 38–94.

Moskowitz, Tobias J., and Annette Vissing-Jørgensen, 2002, The returns to entrepreneurial investment: A private equity premium puzzle?, *American Economic Review* 92, 745–778.

Nelson, Daniel B., 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59, 347–370.

Pástor, Ľuboš, Meenakshi Sinha, and Bhaskaran Swaminathan, 2008, Estimating the intertemporal risk-return tradeoff using the implied cost of capital, *Journal of Finance* 63, 2859–2897.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–55.

Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, *Journal of Financial Economics* 10, 237–268.