Necessary Evidence for a Risk Factor's Relevance*

Alex Chinco,† Samuel M. Hartzmark,‡ and Abigail B. Sussman§

February 19, 2021

Abstract

Textbook models assume that investors are trying to insure against bad states of the world associated with specific risk factors when investing. This is a testable assumption and we develop a survey framework for doing so. Our framework can be applied to any risk factor. We demonstrate the approach using consumption growth, making our results applicable to most modern asset-pricing models. Participants respond to changes in the mean and volatility of stock returns consistent with textbook models, but we find no evidence that they view an asset's correlation with consumption growth as relevant to investment decisions.

Keywords: Risk Factors, Expected Returns, Asset Pricing

*We would like to thank Elena Asparohova, Nicholas Barberis, Justin Birru, Ing-Haw Cheng, James Choi, Tony Cookson, William Diamond, Teodor Dyakov, Xavier Gabaix, Stefano Giglio, Niels Gormsen, Robin Greenwood, Arpit Gupta, Lars Hansen, Xing Huang, Alex Imas, Bryan Kelly, Ralph Koijen, Ben Matthies, Marina Niessner, Lubos Pastor, Kelly Shue, Amir Sufi, David Solomon, Doron Sonsino, Johannes Stroebel, Jessica Wachter, George Wu, Jeff Wurgler, and Anthony Zhang as well as seminar participants at Michigan State, Michigan, Chicago Booth, TCU, Utah, Georgia State, Yale, Rochester, Baruch, Tinbergen Institute, Oxford, University of Southern Carolina, University of Washington, Wharton, Colorado, University of Alberta, Virtual Experimental Finance Workshop, Western Finance Association, and NBER behavioral Finance for extremely helpful comments and suggestions. We would also like to thank Abigail Bergman, Nicholas Herzog, and Kari Greenswag for excellent research assistance. Becky White and Jasmine Kwong provided invaluable help with data collection. A prior version was circulated under the title *Risk-Factor Irrelevance*.

†University of Chicago, Booth School of Business. alex.chinco@chicagobooth.edu

‡University of Chicago, Booth School of Business and NBER. samuel.hartzmark@chicagobooth.edu

§University of Chicago, Booth School of Business. abigail.sussman@chicagobooth.edu

1 Introduction

The textbook explanation for why some assets have higher expected returns than others involves "extending the principles behind fire and casualty insurance to investment portfolios (Cochrane, 1999)." While investors in textbook models prefer assets with high average returns and low volatility, they also want assets that provide insurance against certain bad future states of the world. All else equal, they are willing to pay more for assets that tend to realize high returns in these bad future states, giving these assets higher current prices and lower expected future returns.

Different models argue that investors worry about different kinds of bad future states. When a model argues that X is a priced risk factor, it is saying that asset prices are moving because investors are trying to insure the specific kind of bad future state associated with drops in X (whatever that may be). The textbook approach to testing such a model involves looking for a purely econometric relationship between average returns and correlations with X. While such evidence is consistent with X being priced, it is not sufficient to prove the case (Lewellen, Nagel, and Shanken, 2010).

This paper proposes a new and complementary approach. We begin with the most common interpretation of factor models in the literature, which is that they reflect investors' economic goals and motivations.[1] According to this interpretation, asset prices reflect the "strategic behavior of investors who wish to hedge against future adverse changes" (Campbell and Viceira, 1999) in particular risk factors. Investors must value insurance against these risk factors, view their portfolio as a means to get this insurance, and strategically invest to do so. We highlight that this motivation represents a testable hypothesis that is not typically directly tested. We develop a survey-based framework for testing it.

Survey experiments are the ideal tool for this purpose. While an investor can buy an asset that happens to provide insurance without knowing it, investors cannot all agree on the equilibrium price of this insurance unless it is commonly understood ahead of time. Home owners can typically explain what they are paying for when they buy fire insurance. Drivers can typically explain what they are paying for when they buy car insurance. Similarly, if asset markets are "in reality big insurance markets (Cochrane, 1999)", a well-designed survey should be able to provide evidence that investors typically view their portfolio as a way to buy insurance and construct their investment strategy with an eye towards achieving this goal.

The framework we develop in this paper can be used to evaluate the relevance of nearly any risk factor. We demonstrate it using a case study that focuses on consumption growth (CCAPM; Lucas, 1978). We choose consumption growth because the hedging of consumption growth is the core problem of most modern asset-pricing models, or, put differently, "all factor models are derived as specializations of the consumption-based model (Cochrane, 2001, §9.3)." Thus, by studying consumption growth, our results represent a test of much of modern asset pricing. While we find strong evidence that investors care about

[1] Below we discuss alternative interpretations of these models and the implications of our results under each interpretation.

the mean and volatility of returns, we find no evidence that investors view their stock-market portfolio as a way to insure shocks to their consumption. This result persists across multiple different tests, methods, and participant pools.

Our two-part survey-based framework begins by examining investment decisions. We ask each participant multiple questions about how they would allocate an endowment between a portfolio of stocks and a riskless bond based on data describing stock returns and consumption growth. To proxy for consumption growth, we show each participant the actual historic time series of GDP growth and describe it to participants using the term *'economic growth'*.[2] Each participant sees the same economic-growth time series in every question they answer. However, we simulate the stock-returns in each question using randomly chosen mean, volatility, and consumption-growth correlation parameters. To remove the need to form expectations, we show participants the parameters and ask them to assume these parameters hold in the future.

We survey a wide range of investor types: finance professionals (including some professional traders), Mechanical Turkers, Booth MBA students, and a group of clients at an invite only meeting of one of the world's largest asset managers. While this asset manager must remain anonymous, this last group of participants represents prototypical sophisticated investors.

We designed our survey experiment to make it as easy as possible for all participants to follow textbook logic if it was something they would naturally do. We gave them intuitive instructions, removed all superfluous information, and presented data that is usually difficult to find in a straightforward manner. Our experiment focuses exclusively on two economic concepts (economic growth and stock-market returns) and three statistical parameters (means, volatilities, and correlations). Participants in our online versions had to pass a comprehension test demonstrating that they understood the definitions of each concept and parameter.

We find that participants respond to the changes in the mean and volatility of returns as predicted by textbook theory (e.g., Markowitz, 1952; Hicks, 1939). They invest more in stocks when average stock returns are higher (p-value $< 0.01\%$), and they invest less when stock returns are more volatile (p-value $< 0.01\%$). These results demonstrate that participants are thinking about and responding to risk and return. This suggests that they understand their investment task and strongly respond to changes in some parameters consistent with textbook logic.

However, participants do not meaningfully respond to changes in consumption-growth correlations (p-value $= 99.9\%$). We examine scenarios ranging from stock returns being uncorrelated with consumption growth, $\rho = 0.00$, to having a correlation of $\rho = 0.45$. This is an economically large increase in correlations according to textbook models. To accommodate this increase in correlations, investors in the Lucas (1978) model should demand an additional 11% risk premium per year, investors

[2]To test the robustness of our findings, we also run surveys using nine other terms one might use to describe "bad times". In the paper, we use the term *'consumption growth'* when describing the risk factor in all these related surveys.

in a habit-formation model should demand an additional 8% (Campbell and Cochrane, 1999), and investors in a long-run risks model (Bansal and Yaron, 2004) should demand an additional 20% (see Section 5.4). Despite these predictions and despite finding a t-statistic above 14 in response to a 4% change in expected returns, participants do not change their allocations in response to correlations.

While these results suggest that participants ignore an asset's correlation with consumption growth when investing, it is possible they are thinking about this parameter in a manner not captured by our empirical setup. To address this concern, the second part of our framework directly asks participants how they made their investment decisions in the task described above. First, we ask whether a participant considered each parameter (mean, volatility, correlation with consumption growth). Second, for each parameter a participant considered, we ask what direction they favored for this parameter when forming portfolios. For example, if a participant said he considered return volatility, we then asked if he was trying to invest more in stocks when returns were more volatile or less volatile.

Consistent with the first part of our framework, participants report caring about the mean and volatility of stock returns when investing. By contrast, most participants (57%) stated that the correlation between stock returns and consumption growth did not play any role in their decision making. Furthermore, among the 43% of participants who did report thinking about correlations, roughly 3 out of 4 reported increasing their demand for stocks when stock returns were *more* correlated with consumption growth—the opposite of what a textbook investor is assumed to do. Across all participants, only 11% reported thinking about consumption-growth correlations in a manner consistent with textbook theory.

Our approach helps address common concerns about using surveys to study financial markets. First, we minimize concerns that our results are specific to a given sample by administering the survey to a diverse array of participants. We find consistent results in populations ranging from MTurkers to asset managers. Second, we minimize concerns that participants are confused or that our framework lacks statistical power by including questions about means and volatilities as a point of comparison for questions about correlations. Strong responses to means and volatilities consistent with traditional theories suggest that participants are responding meaningfully to the survey and that the survey has sufficient power to detect a response. Third, we minimize concerns that participants are responding to a different question than the one being asked (e.g., as in attribute substitution; Kahneman and Frederick, 2002) by focusing on a concrete investment-allocation decision and asking subsequent questions about that decision.

Finally, our framework minimizes concerns that participants are providing responses based on expectations of how they should respond—e.g., resulting from experimenter demand (Schwarz, 1999) or social-desirability concerns (Grimm, 2010). In the first part of our experiment, we do not directly question participants about the core parameters (mean, volatility, correlation). Instead, we examine investment-allocation decisions for which there is no clear or suggested correct answer. When we ask participants about their reasoning in the second part, we ask not only whether they considered a given

parameter but also the direction that they preferred. In terms of direction, there again is no clear or suggested correct answer. Importantly, to the extent that concerns remain regarding a particular question type, we provide converging evidence for our results across multiple question types, each of which supports the paper's main conclusions.

Another benefit of using surveys is that it is straightforward to run robustness checks to evaluate alternative hypotheses. For example, there could be confusion as to what words a layperson would use to describe the theoretical concept of consumption growth. We use the term *'economic growth'* in our baseline version, but we find the same null response to changes in correlations when we use the terms *'gross domestic product (GDP)'*, *'industrial production'*, *'aggregate consumption'*, *'personal wealth'*, *'personal income'*, *'house prices'*, *'personal consumption'*, *'personal spending'*, and *'material standard of living'*. Participants could be confused about how to interpret the numeric value of a correlation (even though a definition is provided, participants must pass a comprehension check, and some participants are professional traders). Yet, we find similar results when we include additional details on the scale of a correlation, when we show participants a scatterplot rather than a time-series chart, when we remove the graph entirely, when we expand the range of correlations to include negative values, and when we use words (high, medium, low, or none) rather than numbers to describe the correlation.

The framework we propose has a number of attributes that make it suited to be a workhorse framework for testing asset-pricing models. It is fast, inexpensive, consistent with best practices of survey design, and can be used to evaluate nearly any risk factor. That being said, there are of course alternative approaches one could take to provide evidence of risk-factor relevance. We explore the robustness of our base framework by examining two alternative research designs.

The first alternative approach is based on a free-response question format. This format allows participants to describe their investment considerations with minimal constraints or experimenter influence. Participants can demonstrate a desire to insure specific risk factors, including those we did not specifically address in the prior experiments. We provided participants a prompt about an investment decision and asked them to describe the considerations they thought were important in their own words. Responses to such an open-ended question can be difficult to quantify and need not be direct tests of a proposed model. To overcome this challenge, we first asked a pilot group of participants to respond to a similar prompt. Based on these responses, we created a series of multiple-choice follow-up questions to encompass the range of ideas expressed by participants in the pilot. The main study consists of participants who first responded to a free-response prompt and then self-categorized these free responses by answering multiple-choice follow-up questions. This procedure allowed us to classify participant responses and quantify the relevance of a given risk factor.

Participants in this open-ended survey reported that they care about the mean and volatility of stock returns when investing, but they did not report viewing the stock market as a way to insure consumption shocks. Reading through the free responses alone, we found no response indicating a desire to use their

portfolio as insurance in any way. While many participants report that their personal financial well-being and the state of the economy are important considerations, follow-up questions reveal that they do not view these variables as risks to be insured through the stock market.

The fact that we find similar results to our baseline framework using this open-ended design allays a variety of concerns about the baseline results. Our main survey-based framework is designed to focus participants' attention on the key parameters in textbook models.[3] This design choice is unlikely to be driving our results since we find similar results when using a broad open-ended prompt that does not meaningfully restrict participants' responses or direct their focus towards any specific variable. Further, the open-ended treatment does not explicitly mention any economic concepts or statistical parameters. Finding similar results in this free-form setting ameliorates concerns that the baseline findings can be ascribed to problems specific to how we present these ideas.

The second alternative approach involves looking at the information typically provided to investors. If investors were interested in insuring consumption shocks, a variety of information providers would have incentives to prominently display information relevant to these shocks. Thus, the fact that *The Wall Street Journal* does not routinely announce or discuss correlations with consumption growth should give us pause. It could be that investors already know these numbers, leaving no need for the media to report them. To explore this possibility, we look at sources that would display relevant information even if the information was widely known, such as mutual-fund prospectuses. Even if everyone is aware of a fund's risk-factor correlations, the fund must still discuss them in its prospectus. Presumably most investors know the stock market can go up or down, yet every prospectus we examined discusses this risk. By contrast, not a single fund lists a correlation with any aggregate risk factor in its prospectus. We also study professional risk-assessment tools and educational documents produced by regulatory agencies. We find no evidence that these investor groups seek information about how their portfolio could serve as insurance against consumption drops, making such an insurance motive less likely to be a fundamental motivation for investing.

Researchers use and interpret asset-pricing models in different ways, and the implications both of the models and of survey evidence vary accordingly. The interpretation of consumption-based asset-pricing models favored by much of the literature is that they capture the economic motivation guiding investor choices. For example, in a recent review article, Cochrane (2017) writes that these models do not simply "'explain' facts or events ex post" but rather encode "explicit measures of fearful outcomes. . . that quantitatively account for asset pricing facts." Under this interpretation, a necessary condition for risk-factor relevance is that investors are able to answer a well-designed survey in a manner consistent with the corresponding factor model.

A key contribution of this paper is to emphasize the importance of survey evidence for testing

[3]This attention should, if anything, induce a bias towards finding that participants do consider risk factors in their investment decisions. Thus we think our baseline treatment is a relatively conservative test.

risk-based asset-pricing models under the standard interpretation.[4] For any given survey (such as the ones in this paper), if the evidence is not consistent with a proposed model it could be that the survey is flawed. An appropriate response under this interpretation is to demonstrate the shortcoming using a different well-designed survey. A paper which found the opposite results using a survey design that better captured investor motives would validate a core contribution of this paper by highlighting the importance of survey evidence for demonstrating risk-factor relevance. If such a survey existed, it would enhance our understanding of survey design, and have implications for future model development. We believe that our framework is effective and that our results are robust, but if the current design falls short, we hope future researchers will design new surveys to demonstrate it.

An alternative interpretation is that risk factors are ex post descriptors of market outcomes. Under this interpretation, investors behave *'as if'* (Friedman, 1953) they are making choices based on a factor model even though the model does not represent their economic motivation. As such, under this interpretation, investors will be unable to provide direct survey evidence consistent with a model. The meaning and use of factor models is much more limited under this interpretation (see Section 5.1 for further discussion). Thus, a contribution of this paper is to demonstrate that the meaning and use of factor models should be much more limited than the expansive interpretation favored by the literature in the absence of evidence that investors are trying to insure a particular risk factor.

Our results underscore the importance of understanding not only whether a model is empirically successful but also why. When an asset's correlation with X does not explain average returns, there are two possibilities to consider. It could be that X is not a relevant risk factor, or it could be that X is a relevant risk factor in spite of its empirical shortcomings. Researchers can use our framework to figure out which possibility is more likely to be correct. If investors are not actively considering X, a researcher should discard it. Whereas, if investors are actively considering X, it could be valuable for a researcher to add features to his model to accommodate X as a priced risk factor.

The aim of the paper is to change how economists evaluate factor models. A researcher proposing that X is a priced risk factor can apply our survey-based framework to strengthen his claim. A researcher proposing a particular behavioral bias or trading friction as an explanation for asset prices can apply our framework to address concerns about competing risk-based explanations. Consumers of this research should be skeptical that any X is a priced risk factor without direct evidence of risk-factor relevance from the investor's perspective.

Contribution to the Literature

There are many models that predict either the time series (Cochrane, 2017) or the cross-section (McLean and Pontiff, 2016) of returns. The standard approach to disentangling these models is to use

[4]This is a noteworthy attribute of risk-factor models that is not universal to asset-pricing models in general. For example, a model in which prices are set by naïve investors who are making a behavioral error cannot be tested in the same way. Such investors should not be able to answer direct questions about an error that they do not know that they are making.

advanced econometrics (Chinco, Neuhierl, and Weber, 2020; Feng, Giglio, and Xiu, 2020; Freyberger, Neuhierl, and Weber, 2020; Harvey and Liu, 2020; Bryzgalova, 2017; Kozak, Nagel, and Santosh, 2018b). Instead, we propose a survey-based framework to test whether investors follow the core economic logic behind a given factor model.

A number of papers have examined biases in how correlations are perceived (Jennings, Amabile, and Ross, 1982; Matthies, 2018; Ungeheuer and Weber, 2019; Laudenbach, Ungeheuer, and Weber, 2019). Our setting attempts to remove the influence of such a channel. We provide the values of all relevant parameters and investigate what participants do with this information.

Researchers have examined how misunderstanding correlations can bias decision making (Enke and Zimmermann, 2017; Levy and Razin, 2015). Researchers have also demonstrated that participants often do not appropriately account for the correlations between assets (Eyster and Weizsacker, 2016; Kallir and Sonsino, 2009; Matthies, 2018). Further still, there is a macro literature exploring investor inattention (Gabaix, 2014, 2016). Our paper adds to this literature by showing that an additional way to test a factor model is to determine whether investors are motivated to consider key correlations.

This paper complements a literature on decision making and rational expectations in experimental settings (Plott and Sunder, 1988; Smith, Suchanek, and Williams, 1988). A number of papers examine whether people behave like classic mean-variance investors (Bossaerts and Plott, 2004; Huber, Palan, and Zeisberger, 2019; Kroll, Levy, and Rapoport, 1988), and recent work studies the Lucas (1978) model in laboratory settings (Asparouhova, Bossaerts, Roy, and Zame, 2016; Crockett, Duffy, and Izhakian, 2019). While related, these papers use experiments to study market outcomes. By contrast, our framework studies the logic investors use to arrive at these outcomes.

Prior work has looked at how people frame financial decisions—e.g., over individual positions (Odean, 1998), across positions (Frydman, Hartzmark, and Solomon, 2017), and over portfolios (Hartzmark, 2014). There are survey papers asking questions about consumption growth (Di Maggio, Kermani, and Majlesi, 2020) and investors' views on popular finance models (Choi and Robertson, 2020; Giglio, Maggiori, Stroebel, and Utkus, 2020; Liu, Peng, Xiong, and Xiong, 2020). We use a survey-based framework that shows, contrary to textbook logic, investors' framing of their portfolio decision does not involve consumption-growth correlations as an input.

2 Survey Design

Textbook models make a core claim about how investors view assets markets. Investors in these models prefer assets with stable returns that are high on average. But they also worry about not having enough money during particular kinds of bad future states of the world. So, all else equal, they are willing to pay more for assets that tend to realize good returns in these bad future states, giving such assets low expected returns in the future. Our paper emphasizes that it is possible to test this core claim and that surveys are the ideal tool for doing so. This section describes the survey-based framework we

develop to test risk-factor relevance.

2.1 Instructions

This section describes our baseline survey administered via an online platform. We discuss the minor modifications we made for different participant populations and robustness checks below. We include the text of all survey variants in the Internet Appendix. The Internet Appendix also contains links to the Qualtrics QSF files we used to run each survey so the results can be easily replicated and new variants easily explored. In addition, we have written step-by-step instructions that researchers unfamiliar with the method can use to quickly and easily create online survey experiments (Bergman, Chinco, Hartzmark, and Sussman, 2020).

When designing the survey, our goal was to make it as easy as possible for participants to follow textbook logic if they were naturally inclined to do so. We used only three statistical concepts (means, volatilities, and correlations) as well as only two economic concepts (the stock market and economic growth). We provided non-technical definitions for every concept used. We used plain language and provided intuitive descriptions that participants could easily understand.

Upon entering the survey and completing a consent form, participants were given instructions explaining that they would be making investment decisions based on economic growth and stock market performance. The instructions included intuitive definitions of key terms:

> *Economic growth refers to how well the economy as a whole is doing. It is commonly reported as Gross Domestic Product (GDP) which is a measure of the goods and services produced in the US economy. The information about the stock market is for a mutual fund that passively invests in a broad blue-chip stock market index, such as the S&P 500 or the Dow Jones Index. The value of the mutual fund reflects the value of its investments, so when the stocks it invests in have a higher price, the value of the mutual fund will be higher.*

We did not use the technical terms of *'consumption'* and the bundle of *'all risky investments'* to describe the key economic concepts in our survey. Instead we chose terms that participants were likely to understand and that intuitively captured the concepts the models are meant to speak to. We did this to avoid confusion that more technical terms would likely create for participants. Particularly with the term *'consumption'*, we felt that its popular usage often has a negative connotation (e.g., akin to *'expenses'*) rather than a positive one.[5]

In addition, there is debate as to what real-world measure captures the theoretical concepts of *'consumption growth'* (for a review see Ludvigson, 2013) and *'all risky investments'* (e.g., see Roll, 1977). For the bundle of all risky investments, finance academics often proxy for this variable using the aggregate stock market, though this does not directly map to the concept in the model (Roll, 1977). We frame our question as related to the stock market, but we do not use the most common academic measure

[5]The list of synonyms for *'consumption'* in the Oxford Dictionary includes *'expending'*, *'depletion'*, *'exhaustion'*, *'waste'*, *'squander'*, and *'draining'*.

of the CRSP value weighted measure of performance. Instead, we ask about the S&P 500 and the Dow Jones Index because they are the most likely to be reported and discussed, though they represent inferior measures of market performance (Hartzmark and Solomon, 2020).

Similar to risky investments, there is active debate as to what the conceptual measure of consumption growth in the model relates to in the real world and how to measure it. Measures include non-durable consumption (Mehra and Prescott, 1985), GDP (Barro, 2006), and garbage (Savov, 2011) among others. We chose the term *'economic growth'* and describe it using the term *'GDP'* in our baseline treatment because these are the most commonly discussed and reported terms in popular discussions.[6] In robustness checks, we explore many alternative labels for the proxy for consumption growth. When describing our results in both this baseline version and the alternative labeling, we will use *'consumption growth'* as shorthand for this collection of related state variables meant to proxy for the theoretical concept of consumption growth.

Each participant was informed that they would be seeing annualized numeric values for the mean, volatility, and correlation between two time series as well as a graph showing the cumulative performance of both stock returns and the economy as a whole. To make sure that it was clear what this meant, we provided each participant with the definition of mean, volatility, and correlation:

> *When the average per year is higher you should expect greater increases in value in a given year, corresponding to steeper increases in the line displayed. When uncertainty is higher, you should expect greater swings, for example higher highs and lower lows are more likely than if uncertainty is low. When a correlation is higher, this means that if one series goes up, the other is more likely to go up too, and if it goes down, the other is also more likely to go down.*

To ensure that participants understood the concepts, they were asked to answer several multiple-choice comprehension questions about the definitions of economic growth, average growth, uncertainty, and correlation. We only included data from the subset of online participants who correctly answered every comprehension question correctly on their first try in our empirical analysis.[7]

The first portion of our survey asks participants to allocate a $1,000 endowment between stocks and bonds based on information about both economic growth and stock returns provided in the question. Specifically, we ask them to allocate money between a "mutual fund that passively invests in a broad blue-chip stock market index, such as the S&P 500 or the Dow Jones Index," and "a bond earning 2%." See Figure 1 for a sample question. Participants in the online versions see 10 questions of this type where each question was randomly selected from a larger set of 36 possible choices.

[6] Other survey applications attempting to examine consumption-based asset-pricing models make similar choices (e.g., Giglio, Maggiori, Stroebel, and Utkus, 2020).

[7] Finance professionals continued the experiment only if they correctly answered all of the comprehension questions correctly on the first try (493 participants). All MTurkers completed the full survey irrespective of their responses to the comprehension check, but we only use data from the participants (322 in the baseline version) who passed the comprehension checks on their first try.

For each question, a participant observed a time series of cumulative stock returns and a time series of economic growth as well as summary statistics describing the mean, volatility, and correlation between these two time series. Participants were told that the numeric values provided to them are stable predictors of future returns for that particular fund. We also told participants that each round of investment allocations was unrelated to the last.

After participants finished this first set of questions, we asked them a series of follow-up questions about the economic reasoning behind their portfolio decisions. We started by asking participants whether they considered a parameter (mean, volatility, correlation) at all. Then, for each parameter a participant considered, we asked how they were using this parameter when forming portfolios. For example, if a participant reported caring about the volatility of stock returns, we then asked whether he was trying to invest more in stocks when stock market volatility was higher or lower.

2.2 Design Choices

We aimed to use research methods that would most accurately assess participant motivations. Specifically, we aimed to identify whether investors view a proposed risk factor as relevant. While seemingly simple, there are multiple ways that one could examine such a question. As with any tool, there are better and worse ways to implement a survey design. Poor econometric technique can lead to erroneous conclusions from data; whereas, a well-executed analysis can lead to meaningful insights. The same is true of surveys. In this section, we discuss why we designed the survey as we did and why we believe this design provides accurate and interpretable responses.

First, we examine a simple investment-allocation decision setting that allows us to elicit as much information as possible from our participants while avoiding a number of potential pitfalls. In our stylized setting, we can define all relevant terms to make sure the concepts are understood. Further, in this setting we can capture behavior that an economist would describe as consumption hedging irrespective of how the participant would describe it. We provide a concrete investment-decision task and then ask participants both to complete the investment allocation and to directly explain how they made the decision in the experiment.

One common concern in survey design is attribute substitution (Kahneman and Frederick, 2002). When asked a complicated question, participants may give an answer to an easier related question. If we were to give participants a complex question, such as a question about how important a particular strategy is for their own personal investment decisions (which involves many variables such as intra-household decision-making, preferences over non-financial aspects of assets, biases favoring certain stocks over others, etc), participants may instead respond to a simpler question, such as a question about whether that strategy sounds reasonable. By designing a simple concrete task and creating a question focused on that task, our setting makes it as easy as possible for participants to answer the intended question.

We are also careful not to ask leading questions, which can produce experimenter-demand effects (Schwarz, 1999). This bias occurs when a researcher's choice of questions influences participants'

answers. Suppose a participant does not actually try to insure consumption risk in the stock market and has never thought about doing so. When asked about such a strategy directly, the participant might anticipate that a researcher would not be asking this question unless it was widely used or recommended. As a result, he might falsely claim this is how he invests.

The investment-decision questions in the first part of our framework largely side-step this concern since it is not clear how a participant should be incorporating the information provided. However, when directly asking participants about their economic reasoning, we use a two-stage question so that participants cannot identify the experimenter's intent, thereby avoiding introducing bias into their responses. In the first stage, we ask whether participants considered a parameter (e.g., *'Did you think about mean stock returns when investing?'*). Then, if the participant reported considering a given parameter, we ask a bi-directional question about how they invested based on this parameter in the second stage (e.g., *'When mean stock returns were higher, did you try to hold more or less stock?'*). Using this procedure, even if participants believe the researcher would like them to care about a given parameter, it is not clear which direction they should report desiring. Thus, we jointly examine consideration and direction of each parameter as our main outcome variables.

By using multiple question types (i.e., investment allocation and economic reasoning) and examining whether responses are consistent at the population level across question types, our framework provides converging evidence. To the extent that responses are consistent across question types, arriving at the same conclusions using two different questioning strategies should help ameliorate concerns about using any single technique. Further, using multiple question types also allows us to cross-validate responses within an individual, by examining whether a given participant's answers are internally consistent, across question types. Consistent evidence within participants further suggests that our survey experiment produces meaningful responses to the questions being asked.

We use consumption growth as a case study in this paper to demonstrate our framework, but our survey design can evaluate the relevance of almost any candidate risk factor.[8] For most candidate risk factors, a researcher should be able to take our framework, replace consumption growth with a new risk factor, and run the same survey experiments. However, we recognize that some risk factors might have idiosyncrasies requiring further tailoring of our approach. There is an active literature studying how to design surveys which we cannot do justice to here. As such, we emphasize that researchers should abide by current best practices in survey design when deviating from our benchmark survey. In Bergman, Chinco, Hartzmark, and Sussman (2020), we provide a step-by-step guide that researchers can use to run their own online survey experiments to test risk-factor relevance along with discussion of the survey design literature we think is likely to be relevant.[9]

[8] Section 5.3 demonstrates how the framework can be used to test cross-sectional factors from Fama and French (1993).
[9] We strongly encourage economists interested in developing a new survey design to consult colleagues who are well versed in survey design.

2.3 Participant Populations

We surveyed four distinct participant populations. The goal was to solicit responses from a broad swath of the investing population to minimize the possibility that there was an important investor type we did not reach. As such, we examined a range of investors in terms of sophistication and wealth.[10]

The first participant pool is comprised of people who work in the finance or banking industries. We used CloudResearch, a service that specializes in connecting researchers with unique and hard-to-reach sample populations, to recruit these participants.[11]

We present summary statistics for the 493 finance professionals who completed our survey in panel (a) of Table 1. This participant pool was fairly wealthy, with 56% earning more than \$100*k* per year. 45% of the finance professionals we surveyed were under the age of 40. Furthermore, these participants tended to invest their own money, with 90% reporting that they owned either individual stocks or mutual funds. We also asked participants about their job function, and 28% of this population stated that their job involved investing in financial securities.

The second population we examine is drawn from Amazon's Mechanical Turk (MTurk) marketplace. For enhanced data quality, all surveys using MTurk participants leveraged a different feature of the CloudResearch platform for participant recruitment and screening. Research examining this platform finds that participants recruited through MTurk, who are commonly referred to as *'MTurkers'*, tend to perform similarly on tasks (Casler, Bickel, and Hackett, 2013) and perform better on attention checks (Hauser and Schwarz, 2016) than traditional participant pools recruited in labs. MTurkers also tend to be more diverse (Paolacci and Chandler, 2014).

We present summary statistics for the 322 MTurkers who completed our baseline survey in panel (b) of Table 1. MTurkers have lower incomes when compared to the finance professionals, with only 13% of MTurkers earning more than \$100*k* per year. MTurkers also tended to be younger than the finance professionals, with 72% being under 40 years of age. A lower fraction of MTurkers owned financial securities, but even in this sample 65% reported owning either stocks or a mutual fund.

The third population we examine is drawn from MBA students at the University of Chicago, Booth School of Business. Our sample consisted of 308 participants, of which 38% reported having previously worked in the finance industry. We gave our survey to students enrolled in various MBA courses at the business school, including sections of a core investments class. Panel (c) of Table 1 contains summary statistics for this participant pool.

[10] Specific investor types might play a special role in markets, but they do not play such a role in textbook asset-pricing models. The canonical Lucas (1978) model studies a representative investor as do popular models built on top of it such as habit formation (Campbell and Cochrane, 1999), long-run risks (Bansal and Yaron, 2004), rare disasters (Rietz, 1988; Barro, 2006), etc. These models do not give any guidance as to which specific investor populations we should survey. If it is critical that we survey a particular investor population when testing a model, then the identity of these investors should be embodied in the model somehow.

[11] CloudResearch has access to more than 50 million online panelists worldwide. See `https://www.cloudresearch.com/products/prime-panels/`.

The fourth population we examine represents clients attending a conference of a large asset manager. Conference attendees at this invite only conference were mostly wealthy investors and portfolio managers. This sample consists of 93 participants who completed our survey. A condition of the survey was that we cannot disclose the name of the asset manager or their clients, but we can assume that conference attendees fit the textbook description of a sophisticated investor and generally manage large sums of money on their clients behalf. Panel (d) of Table 1 contains summary statistics for this participant pool.

2.4 Survey Variations

Sections 2.1 and 2.2 describe the online version of the survey experiment administered to finance professionals and MTurkers. We ran the MBA-student survey using pen and paper during a class break. At the asset-manager conference, we gave our survey using tablets at a designated booth.

We presented the MBA-student and asset-manager samples with abbreviated instructions and definitions. We did not ask comprehension checks due to time constraints and because both groups are likely to be familiar with the basic concepts presented in the survey. The goal of the design was to parsimoniously present the same information as in the online survey experiment to a group of people that have more knowledge of financial markets. The Internet Appendix includes examples of instructions from each of these versions.

For the MBA-student and asset-manager participant pools, we also reduced the number of questions to fit page limits and time constraints. The MBA students saw 5 investment-decision questions. The asset-manager sample saw 4 investment-decision questions. Half of the investors at the asset managers' conference were asked a percent-allocation question, dividing a 100% allocation across the two options rather than stating the number of dollars explicitly. The results were materially similar.

2.5 Data Simulation

In the first part of our survey-based framework, we ask participants questions about how they would invest an endowment based on time-series data about economic growth and stock returns. The economic-growth time series is the same across all questions and all participants. This time series represents seasonally-adjusted quarterly US GDP, $\Delta \log C_t \stackrel{\text{def}}{=} \log(GDP_t) - \log(GDP_{t-1})$, from 1980Q1 to 2018Q4 (i.e., $T = 156$ observations in total).[12]

We simulate stock returns for each question using a randomly selected combination of the following parameter values:

$$\begin{aligned} \mu_R &\in \{4\%, 6\%, 8\%\} \\ \sigma_R &\in \{10\%, 15\%, 20\%\} \\ \rho &\in \{0.00, 0.15, 0.30, 0.45\} \end{aligned}$$

[12] The realized consumption-growth series is quite smooth. So, for a subset of participants, we added orthogonalized noise to the time series to make it easier to see comovement between this time series and stock returns. Results are the same with or without noise, so we pool the samples in the paper and include results split by sample in the Internet Appendix.

μ_R denotes the mean annualized stock return, σ_R denotes annualized stock-return volatility, and ρ denotes the correlation between stock returns and economic growth.

For each set of parameter values, we first draw $T = 156$ iid realizations $\Delta \log Z_t \overset{\text{iid}}{\sim} \text{N}(0, 1)$. Then, we orthogonalize these 156 random draws with respect to the realized set of 156 consumption growth shocks, $\Delta\widetilde{\log} Z_t \overset{\text{def}}{=} (\Delta \log Z_t - \widehat{\mathbb{E}}[\Delta \log Z_t | \Delta\widetilde{\log} C_t]) \,/\, \widehat{\mathbb{S}\text{d}}[\Delta \log Z_t | \Delta\widetilde{\log} C_t]$, where $\Delta\widetilde{\log} C_t \overset{\text{def}}{=} (\Delta \log C_t - \widehat{\mathbb{E}}[\Delta \log C_t]) \,/\, \widehat{\mathbb{S}\text{d}}[\Delta \log C_t]$ denotes the consumption-growth shocks. We do this to avoid the error-in-variables problem caused by our finite sample period, which contains 156 quarters. If we skip this step, the resulting error-in-variables problem is not large, but since we can control everything about our experimental setting, we try to remove all avoidable sources of error from our results.

We simulate stock returns using the formula below:

$$R_t = \mu_R + \sigma_R \times (\rho \cdot \Delta\widetilde{\log} C_t + \sqrt{1 - \rho^2} \cdot \Delta\widetilde{\log} Z_t)$$

Because of the orthogonalization step, the resulting stock-return time series has a mean of exactly μ_R, a volatility of exactly σ_R, and a correlation with consumption growth of exactly ρ. We show participants a line labeled *'stock market'* representing the cumulative returns to investing \$1 in this portfolio in 1980 on a log scale. See Figure 1. For each set of parameter values, we run the simulation using 5 different random-number seeds to make sure that our results are not driven by some chance feature of a particular random realization.

Under standard calibrations, the range of correlations we examine is economically large. The CCAPM says an asset's expected excess return should be

$$\begin{aligned} \mu_R &= \gamma \times \mathbb{C}\text{ov}[R, \Delta \log C] \\ &= \gamma \times (\rho \cdot \sigma_R \cdot \sigma_{\Delta \log C}) \end{aligned}$$

where γ denotes investors' coefficient of relative risk aversion. The typical excess return on the stock market is $\mu_R = 6\%$ per year, and annualized return volatility is roughly $\sigma_R = 16\%$ (Cochrane, 2001, §1.4). Thus, to match the equity premium, we need to assume $\gamma \approx 100$ (Campbell, 2003). So, according to the CCAPM, when $\sigma_{\Delta \log C} = 1.6\%$ investors should view a mutual fund with a 6% per year average returns as underpriced and have high demand for its shares when $\rho = 0.00$; whereas, they should see this same 6%-per-year fund as overpriced when $\rho = 0.45$:

$$\underbrace{100 \times (0.00 \cdot 16\% \cdot 1.6\%) = 0\%}_{\substack{\text{CCAPM-implied expected excess return} \\ \text{in low correlation setting (underpriced)}}} < \quad 6\% \text{ per year} \quad < \underbrace{11.5\% = 100 \times (0.45 \cdot 16\% \cdot 1.6\%)}_{\substack{\text{CCAPM-implied expected excess return} \\ \text{in high correlation setting (overpriced)}}}$$

In other words, moving from $\rho = 0.00$ to $\rho = 0.45$ should cause a CCAPM investor to increase the expected excess return he demands from zero to roughly double the sample average.

A potential concern with this argument is that the variation in correlations we examine is economically meaningful because the relative risk aversion of $\gamma = 100$ is too high. Lowering relative risk aversion to a level more consistent with that found in other settings, say, $\gamma = 10$ would reduce how much CCAPM investors respond to changes in consumption-growth correlations, bringing the Lucas (1978) model more in line with our findings. With that said, when $\gamma = 10$ the models suffers from a constellation of empirical failures commonly referred to as the equity premium puzzle Mehra and Prescott (1985). Since a calibration of $\gamma = 10$ does not appear relevant for understanding asset prices, it is not clear why it should be relevant for understanding correlations.

To address the equity-premium puzzle, researchers constructed consumption-based models, such as habit formation (Campbell and Cochrane, 1999) and long-run risks (Bansal and Yaron, 2004), which can better match the data under more plausible parameter values. These models are not alternatives to the Lucas (1978) model; rather they represent special cases which magnify the effect of consumption risk on asset prices. As such, changes in exposure to consumption shocks have large effects in these models. In Section 5.4 we show that in standard calibrations, the swing in correlations we examine imply 8% and 20% increases in expected returns in habit and long-run risks models respectively. Thus standard asset-pricing models calibrated to match empirical patterns suggest that this range of correlations is economically significant.

3 Main Results

According to textbook models, prices differ because intelligent forward-looking investors worry about not having enough money during certain kinds of bad future states of the world. Assets that are less correlated with the associated risk factors are more likely to have positive returns during these bad future states. Such assets offer better insurance. Investors in textbook models recognize this fact and are willing to pay more for these asset today, giving them lower expected returns going forward.

The standard interpretation is that, for such a model to explain why markets fluctuate, investors must be following this logic. They must prefer assets that have lower risk-factor correlations and adjust their demand based on this information. This is not what we find for consumption growth. We find no evidence that investors trade like textbook investors are supposed to trade or think like textbook investors are supposed to think. We find no evidence that consumption growth is a relevant risk factor.

3.1 Investment Decisions

We begin by exploring how participants' investment decisions change as we vary the mean, volatility, and correlation between stock returns and consumption growth. We do this by estimating regressions of the following form:

$$\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{mean}_{i,q} + \hat{\gamma} \cdot \texttt{volatility}_{i,q} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$$

The dependent variable, $\texttt{stockFrac}_{i,q}$, is the fraction of the \$1,000 endowment that the ith participant invested in stocks when answering the qth question. The variables $\texttt{mean}_{i,q}$, $\texttt{volatility}_{i,q}$, and $\texttt{correlation}_{i,q}$ represent the mean, volatility, and correlation with consumption growth used to simulate the stock returns for that question. We estimate all t-statistics and p-values for the investment-decision regressions using standard errors clustered by participant.

Baseline Results

Table 2 reports the results of these regressions. Column (1) regresses the fraction invested in the stock market on only the mean stock return. We estimate a slope coefficient of $\hat{\beta} = 3.24$, which is statistically significant at the 1% level. This coefficient implies that participants increased the fraction invested in stocks by $12.96\% = (8\% - 4\%) \times 3.24$ in response to a move from the lowest mean return, 4%, to the highest mean return, 8%. Across our four participant pools, people invested about 60% of their endowment in stocks on average. Thus, a $(8\% - 4\%) = 4\%$ change in expected returns increases the proportion allocated to stocks by about $12.96\% \,/\, 60\% \approx 20\%$.

Column (2) repeats the regression using stock-return volatility rather than mean stock returns as the sole right-hand-side variable. We estimate a slope coefficient of $\hat{\gamma} = -0.61$, which is again statistically significant at the 1% level. This coefficient implies that a 10% drop in stock-return volatility—i.e., a move from the highest volatility regime, 20%, to the lowest, 10%—results in a 6.1% increase in participants' stock investment. This corresponds to a $6.1\% \,/\, 60\% \approx 10\%$ reduction in the average amount invested in the stock market. The results in columns (1) and (2) of Table 2 suggest that participants respond to changes in the mean and volatility of stock returns. Moreover, they do so consistent with a textbook investor who likes higher means and dislikes higher volatilities.

The results are quite different when examining correlations in column (3). This column repeats the regression using correlation with consumption growth as the sole right-hand-side variable. We find no measurable change in participants' behavior in response to a change in the correlation between stock returns and consumption growth. The estimated coefficient is $\hat{\delta} = 0.0000235$ (rounded to 0.00) with a t-statistic of 0.00 and a p-value of 99.9%. In addition to being statistically insignificant, this point estimate is economically small. In response to a $\Delta\rho = 0.45$ increase in correlations, participants decrease their allocation to stocks by $0.001\% = (0.45 - 0) \times 0.0000235$. This is one tenth of one basis point. Our participants do not adjust their demand in response to changes in the correlation between stock returns and consumption growth, the canonical risk factor in textbook models.

Our results on means and volatilities suggest that if participants were behaving in accordance with the predictions of textbook models with respect to correlations, our setting would have sufficient power to easily detect it. When examining a change of expected returns of 4%, our experiment identifies strong effects with t-statistics above 14. As discussed in Section 5.4, standard calibrations of the CCAPM, habit based models (Campbell and Cochrane, 1999) and models of long-run risks (Bansal and Yaron, 2004) suggest that the equivalent return compensation for the swing in correlations in our experiment range

from about 8% to over 20%. In other words, textbook investors would view the return equivalent of the range of correlations we observe as 200% to 500% larger than the range of expected returns we explore. Given we find a double digit t-statistic on this smaller range of returns, there should not be a statistical issue identifying textbook responses to correlations.

In column (4) of Table 2, we include all three of these right-hand-side variables in the same regression and find nearly identical results. Participants adjust their demand in response to changes in the mean and volatility of stock returns but not in response to changes in the correlation between stock returns and consumption growth.[13]

To examine whether these results are driven by participant-specific attributes, we add participant fixed effects in column (5) of Table 2. The coefficient on the correlation parameter hardly changes from column (3). Another concern is that participants might change their behavior over the course of the experiment. To account for this, we introduce question-order fixed effects in column (6) and find similar results. Finally, column (7) adds both participant and question-order fixed effects which again results in unchanged point estimates.

The results in Table 2 show that, on average, participants strongly respond to changes in the mean and volatility of stock returns but ignore changes in the correlation between stock returns and consumption growth. It remains possible, however, that these pooled results hide the behavior of a subset of participants which acts differently. To address this concern, we re-estimate the coefficients in column (3) of Table 2 on each participant pool and present the results in Table 3. Column (2) in this table adds participant-pool fixed effects to the specification in column (3) of Table 2 to capture that fact that there are differences in the average fraction invested in the stock market across participant pools. The estimated slope coefficient, $\hat{\delta} = 0.00$, is the same to two decimal points as that in column (3) of Table 2 and suggests these differences are not driving our results.

In columns (3)-(6) of Table 3, we re-estimate the regression separately for each participant pool. We find a coefficient on the correlation parameter of $\hat{\delta} = 0.00$ for the finance professionals, 0.00 for the MTurkers, 0.07 for MBA students, and -0.03 for the asset-manager sample. MBA students have the only point estimate that is marginally statistically significant, but it has a positive sign rather than the negative sign predicted by textbook models. In addition to being generally statistically insignificant, these point estimates are all at least two orders-of-magnitude smaller than the point estimate on mean returns in column (1) of Table 2.

We also examine our results based on participant characteristics. Figure 2 graphs regression coefficients estimated over various subgroups of our participant pools. Every subgroup of participants we examine exhibits similar behavior. Old and young participants; male and female participants; participants with incomes greater than \$100k and those with incomes less than \$100k; participants who think they

[13] Finding similar results to the univariate regressions in this specification is expected since we randomly assign the mean, volatility, and correlation parameters used to simulate the stock returns in each question.

invest wisely and those who do not think they invest wisely all ignore changes in correlation. We repeat the analysis for all participants who state that they own either stocks or mutual funds. We repeat the analysis for the subset of 136 finance professionals in our sample who stated that their job involved trading financial securities. Financial professionals who trade securities for a living do not adjust their demand in response to changes in correlations. In every sub-sample, participants react to changes in means; they react to changes in volatilities; they do not react to changes in correlations.

Survey Variants

In our baseline treatment, we label our variable of interest *'economic growth'* and describe it to participants as "commonly reported as gross domestic product (GDP)." Our goal was to translate the theoretical concept at the heart of textbook asset-pricing models into straightforward intuitive language a layperson would understand (See Section 2.2). With that said, participants could be trying to hedge something similar to economic growth which they think of in different terms. If such a channel were at work, we would wrongly be calling consumption growth irrelevant due to labeling.

Table 4 shows results analogous to column (4) in Table 2 for survey variants where the variable meant to proxy for consumption is labeled and defined differently. First, we examine different terms meant to capture aggregate consumption growth. We label the variable as *'gross domestic product (GDP)'* (see Barro, 2006; Giglio, Maggiori, Stroebel, and Utkus, 2020), *'industrial production'*, and *'aggregate consumption'*.[14] Next we explore variants of personal consumption rather than aggregate consumption by using the terms *'personal wealth'*, *'personal income'*, *'house prices'*, and *'personal consumption'*. Choi and Robertson (2020) asks questions related to consumption-based models and describes the relevant variable as *'spending'* and *'material standing of living'*. We repeat the exercise using these terms in the final two columns. We note that if we were to uncover strong patterns for some of these terms, but not others, this would be important in understanding asset-pricing models and likely lead to new insights as to what they were capturing.

With that said, we do not find a strong response to any of these terms. The coefficient on $\texttt{correlation}_{i,q}$ is statistically indistinguishable from zero in every column. This list of terms represents the cumulative suggestions from a wide array of seminar and conference participants. Finding similar responses to the collective suggestions of a wide swath of academic finance makes it less likely that our results are driven by our choice of words to describe the concept of consumption.

Another potential concern is that the way that information about correlations was conveyed to participants was not understood. For example, it could be that participants are unable to interpret the numeric value of a correlation. There are a priori reasons that make it unlikely that innumeracy explains our results. Online participants passed a comprehension test and reacted to numerical changes in the

[14]Note: We chose not to use the term *'consumption'* in our baseline framework because this is largely a technical term. In common usage it often carries a negative connotation. In our definition of consumption we include the clarification that more consumption is actually a positive thing.

mean and volatility of stock returns. Further, Booth MBA students and professional asset managers are likely to have received statistical education relating to topics such as correlations.

In Table 5, we explore concerns related to the display format of the correlation changes. While we provide a definition of correlation in our baseline, it could be that this was insufficient. In column (1) we show results from a treatment with a more detailed description of correlation, including an explicit statement that the range of a correlation falls between negative one and one. It could be the case that people simply cannot understand numeric values for correlations, so column (2) shows results from a treatment where we remove all numeric correlations and replace them with the words {*'none'*, *'low'*, *'medium'*, *'high'*}. Finally, it could be that we explore an insufficient range of correlations, so column (4) shows results from a treatment with correlations randomly drawn from $\rho \in \{-0.45, -0.30, -0.15, 0.00, 0.15, 0.30, 0.45\}$. We find similar results to our baseline in all these treatments. The coefficient on $\texttt{correlation}_{i,q}$ is still statistically indistinguishable from zero in each.

Another potential concern is that it is unclear whether participants were interpreting the question as asking about allocating a marginal \$1,000 to their existing portfolio or an entire portfolio consisting of \$1,000. The fact that we find similar results when examining MTurkers (to whom \$1,000 might not be viewed as marginal) and when examining asset managers (to whom this is likely viewed as a relatively small amount) suggests that this confusion is unlikely to drive our results. However, we directly address this concern with a separate treatment in which we emphasize that the \$1,000 endowment should be viewed as a marginal decision. Column (4) in Table 5 shows that we find the same results after being explicit about this distinction.

Our experiments include instructions clearly stating that the stock market is simulated and explaining that we provide participants with the relevant parameters that dictate its future movement. We present information in this way to abstract away from the complex problem of forming expectations about future market outcomes (e.g., Malmendier and Nagel, 2011; Greenwood and Shleifer, 2014; Gennaioli, Ma, and Shleifer, 2016; Coibion, Gorodnichenko, and Kumar, 2018). It could be that participants are not following our instructions and are applying their own preconceived notions about stock-market dynamics in our setting. If this were the case though, it is unclear why there would not be a similar concern about means and variances, for which we see strong results consistent with textbook models. Further, we find similar results across each of our participant pools, which suggests that, if preconceived notions were driving the results, then these preconceived notions about market dynamics would have to be similar across these populations. However, it seems unlikely that all the investors in the wide range of populations we examine share a common prior or what that prior would be.

We directly address this concern in two new treatments. The first places additional emphasis on the fact that the stock market is simulated and that the parameters provided are what participants should use to predict future returns, not their prior beliefs. In addition to this emphasis, we include a further comprehension-check question that participants must answer correctly to make sure the instructions were

read and understood. To the extent that investors have strong priors about correlations, it seems likely that these priors are about the stock market and that prior beliefs should be weaker for an arbitrary individual stock. Thus, as a second treatment, we ask about investing in a hypothetical stock, not the stock market. Columns (5) and (6) in Table 5 shows that, when we do this, our results are unchanged.

Finally, there could be concerns related to the time-series plot which we display. We included this plot because it is something that the financial press often includes. A scatterplot would do more to visually highlight a correlation, though such a graph is rarely displayed.[15] To examine whether the display of such a graph influences decisions, we run a new survey in which we show a scatterplot instead of a time-series graph.[16] Column (6) in Table 5 shows that, if we present participants with scatterplots rather than time-series graphs, they behave in the opposite way from a textbook investor. The coefficient on $\texttt{correlation}_{i,q}$ is positive ($\hat{\delta} = 0.15$) and statistically significant at the 1% level. It could be that participants were mislead by the graph altogether. Column (2) in Table 5 shows that, if we ask participants the same investment-decision questions but remove the time-series graph, participants are again less likely to follow textbook asset-pricing logic. So, instead of investing less in stocks when stock returns are more correlated with consumption growth as a textbook investor would do, participants invest more in stocks when stock returns are more correlated. If anything, participants' investment decisions are less consistent with textbook theory when we remove the graph.

We go to great lengths to communicate the correlation between stock returns and consumption growth to our participants. Every survey variant we study delivers similar results: participants respond to changes in the mean and volatility of stock returns in the manner that textbook models say they should. However, this is not the case for consumption-growth correlations.

3.2 Economic Reasoning

The results above suggest participants do not adjust their demand for stocks in response to changes in the correlation between stock returns and consumption growth, the canonical risk factor in textbook asset-pricing models. But perhaps participants are considering consumption risk in a manner that is not captured by our regressions. To address this possibility, we follow up the investment-decision questions by asking participants directly about the economic reasoning behind their choices.

This research design also helps address the concern that the lack of responsiveness to correlations in portfolio allocations is due to immutable prior beliefs or certain biased perceptions of correlations. Under the first explanation, participants with a strong prior would disregard the numbers describing correlations that are displayed to them. Alternatively, participants with a bias would perceive a different correlation than the number displayed.[17] Either way, participants would care about correlations, but the

[15] The fact that media outlets do not generally choose this graphical format, which accentuates correlations, already suggests that this statistic is less likely to be relevant to investors.

[16] The scatterplot graphs quarterly economic growth relative to quarterly stock returns.

[17] We note that under some biases, the first portion of our experiment would still likely yield significant results in the direction

investment-decision regressions would not capture this because the relevant correlations would be investors' prior beliefs or biased perception, not the displayed parameters. While such investors would not change their allocation decisions in the first portion of our experiment, they should still report caring about correlations in a manner consistent with textbook models in the second portion.[18]

We first ask participants whether they considered mean stock returns, stock-return volatility, and/or the correlation between consumption growth and stock returns when making their investment decisions. Let $\texttt{consider}_i$ be an indicator variable which is equal to one if the ith participant reported thinking about a parameter when making their investment decisions. The first row of Table 6 shows that 77% of participants considered average stock returns when making their investment decisions, 59% of participants considered stock-return volatility, but only 43% of participants considered the correlation between stock returns and consumption growth. Thus, most participants did not consider the correlation between stock returns and consumption growth when investing even after being given its numeric value and being asked directly about it. Textbook theory suggests this parameter should be a central object of interest for investors, yet this parameter was not even considered by 57% of participants.

In the second stage, we ask participants who said they considered a given parameter about the direction in which they used this information. Let $\texttt{textbookLogic}_i$ be an indicator variable that is equal to one if the ith participant reported thinking about a parameter using textbook asset-pricing logic. This variable equals zero if a participant does not consider the variable at all. The third row of Table 6 examines whether the participants who considered a variable did so in a manner consistent with textbook theory. It shows that 76% of the participants who considered average stock returns when making their investment decisions tried to buy more stocks when average stock returns were higher. Likewise, 72% of the participants who considered stock-return volatility tried to buy more stocks when this parameter was lower. In short, three out of four participants who considered the mean and volatility of stock returns when investing did so in the textbook direction.

Column (3) in Table 6 shows that the opposite is true for consumption-growth correlations. Of the participants who did consider the correlation, most did so in the opposite direction of what textbook models would suggest—i.e., three out of four participants who considered consumption-growth correlations tried to buy *more* stocks when stock returns were *more* correlated with consumption growth. This means that three out of four participants who reported considering the correlation between stock

predicted by textbook asset-pricing models. For example, one potential bias is perceiving correlations to be more moderate than they actually are (e.g., Enke and Zimmermann, 2017). Suppose that participants perceived correlations to be half as large as they actually are. Participants should still react to the attenuated correlation values that they perceive as predicted by textbook models. And the effect of a $\Delta\rho = 0.45/2 = 0.225$ increase in correlations should be a 5.8% point increase in expected returns. This is larger than the $(8\% - 4\%) = 4\%$ difference in expected returns we examine in our survey, which participants respond to with a t-statistic of 14. Under other biases, such as viewing all non-negative correlations as having the same value (e.g., Matthies, 2018), we would not find significant results in the first portion of our experiment for the reasons described above.

[18] Similarly, in the open-ended treatment, they should report caring about correlations and we should find evidence of investors demanding to see information about correlations in the field data. We do not find such evidence in Section 4.

returns and consumption growth were trying to hold more stocks when stock returns were a worse hedge against bad economic times. The results suggest the non-responsiveness of participants' demand to changes in risk-factor correlations in the first portion of our experiment simply reflects how participants think about their investment decisions.

Table 7 repeats the analysis in Table 6 separately for each participant pool to show that the results are not driven by the less financially sophisticated participants in our sample. Column (3) shows that only 35% of the investors at the asset manager's conference (33 out of 93 participants) said that they considered the stock market's correlation with consumption growth when making their investment decisions. Column (9) shows that less than half of those 33 investors (48% or 16 participants) said that they tried to invest more when holding stocks was a better hedge against drops in consumption. Across all participant pools, investors tended to not think about correlations or to do so in a manner inconsistent with textbook theory, mirroring the results from the first part of our framework.

Table 9 further shows that the results are the same when we relabel *'economic growth'* as any of the alternatives that we previously explored. In every case, participants were more likely to report thinking about the mean and volatility of stock returns when forming their portfolio. And conditional on doing so, they were more likely to think about the mean and volatility of stock returns like a textbook investor should. Similar to the previous results, among the roughly 1/3 of participants who reported thinking about the risk-factor correlation at all, most did so using the opposite of textbook logic. Of the participants who thought about a risk-factor correlation at all, four out of five reported thinking about the correlation between stock returns and *'gross domestic product (GDP)'*, *'industrial production'*, *'aggregate consumption'*, *'personal wealth'*, *'personal income'*, *'house prices'*, *'personal consumption'*, *'personal spending'*, or *'material standard of living'* tried to buy more stocks when stocks were more correlated with this risk factor.

One of the benefits of our two-part survey design is that it allows us to examine the consistency of a specific participant's responses across multiple question types. In doing so, we provide evidence as to whether certain responses are likely to represent meaningful data. If the same person offers inconsistent responses to different questions about the same idea, then his responses likely reflect noise inherent in the survey design. However, if survey participants give consistent answers to these different questions, then these answers likely reflect participants' true motivations.

For example, in Table 6 we find that 11% of participants report thinking about consumption-growth correlations like a textbook investor should. One possibility is that these participants strongly follow the textbook logic and perhaps are particularly important for driving asset prices. Another possibility is that this 11% is largely noise.

To examine this question, in Table 8 we review each participant's portfolio-allocation decisions as a function of his responses to the follow-up questions. Each entry in Table 8 is a regression of the fraction invested in stocks on the indicated parameter using data on the specific subpopulation. The first and

second rows report results only for those participants who said they either did not consider or did consider a particular parameter. The first row of columns (1) and (2) shows that participants adjusted their demand for stocks in response to changes in the mean and volatility of stock returns even when they said they were not explicitly considering these parameters. When participants did report thinking about either the mean or volatility of stock returns, their portfolio response was even stronger as shown in the second row of those two columns. In contrast, the first two rows of column (3) show that there was no change in participants' demand in response to changes in consumption-growth correlations regardless of whether they reported thinking about the parameter. In other words, participants' reports of economic motives appear to map onto their investment decisions for mean and volatility, but not for correlation.

The third and fourth rows of Table 8 report results only for participants who considered a given parameter in a manner consistent or inconsistent with textbook logic. The third row reports results for participants who stated they did not think about the parameter using textbook logic. The fourth row reports results for those who told us they thought about the parameter like a textbook investor would. In columns (1) and (2), the demand responsiveness to the mean and volatility of stock returns is stronger when participants reported that they were using textbook logic. When a participant said he was trying to buy more stock when mean returns were higher in the second portion of our experiment, he did so in the first portion of our experiment, $4.06\% \gg 1.18\%$. Likewise, participants who said they were thinking about stock-return volatility and trying to buy less stock when this volatility was higher again followed through on their reported aims, $-1.17\% \ll 0.05\%$.

By contrast, the demand responsiveness to correlation changes is nearly zero for all participants who told us they considered the parameter. The participants who told us they were trying to use textbook logic have demand that is indistinguishable from those who told us they were trying to take the opposite approach. This suggests that the respondents who report thinking like a textbook investor, do not behave as such in the portfolio-allocation decisions. Thus the responses in the second portion of our experiment indicating that correlations motivated investment decisions likely represent noise, rather than a strongly held desire to insure consumption.

The results of our framework provide converging evidence that is difficult to square with textbook models. These models argue that differences in asset prices are the result of investors trying to insure consumption shocks. We find no evidence that investors trade like textbook investors are supposed to trade or think like textbook investors are supposed to think.

4 Alternative Approaches

As discussed in Section 2, we think our baseline framework is consistent with best practices in survey design and provides a rigorous test of whether investors view their portfolio as a consumption hedge. Further it is inexpensive, fast, and easy to modify to examine alternative risk factors. However, any research design could have shortcomings that lead to erroneous conclusions. Thus, in this section we

explore two alternative methodologies to provide further evidence that our main results reflect investors lack of desire to hedge risk factors and not some artifact of our main survey-based framework.

4.1 Open-Ended Survey

The first alternative approach involves running a survey in which participants answer an open-ended question about what matters to them when making an investment decision. Open-ended questions encourage deeper thinking and allow participants to elaborate on their choices (Behr, Kaczmirek, Bandilla, and Braun, 2012; O'Cathain and Thomas, 2004; Singer and Couper, 2017). The baseline framework draws attention to certain economic and statistical concepts in an abstract setting. An open-ended framework allows us to examine participants' motivations without highlighting specific concepts within the survey (Reja, Manfreda, Hlebec, and Vehovar, 2003). If an investor is trying to insure against a risk factor, it should be natural for him to describe this goal when asked. If the investor has other considerations in mind, an open-ended question will allow us to identify these other considerations as well.

We constructed this treatment to focus on testing a risk factor, but the iterative method outlined below can be applied more broadly to understand investors' goals and motivations. Based on initial prompts, it would be possible to construct different questions which allow a deeper understanding of how investors are approaching a given problem and what it is that they are trying to solve. While beyond the scope of this paper, we think this is an interesting and important direction for future research that can allow researchers to probe existing hypotheses as well as to uncover new ones.

Survey Design

The survey design asks an open-ended question about an investment decision:

> *Please imagine that you have some money that you could either invest in the stock market or keep in savings. You are currently deciding how much to invest in the stock market. Please take a moment to think about this decision.*
>
> *For example, what factors are typically most important to you when making investment decisions in your daily life? What information would you want to help you make your decision? What elements of the decision would be most important to you? What are the main goals you would be hoping to achieve through this choice?*
>
> *Using the space below, enter the first thoughts that come to mind.*

While the responses to this prompt are informative on their own (and we analyze them below), the free-response format makes it difficult to systematically interpret and quantify these responses. Participants can express any thoughts they have, which means a given response need not be consistent, inconsistent, or even related to a model under consideration. Thus, we require a further step in our experiment to ensure the results are quantifiable and can map to an asset-pricing model of interest.

We do so using follow-up questions that occur after the free response. In order to create these

questions, we first ran a pre-test where 50 participants responded to the open-ended prompt. Participants were provided with separate lines for up to 10 distinct thoughts and were required to list at least 3.[19] The goal was to understand the broad categories that participants were likely to refer to. We then used this understanding to create follow-up questions representing the full range of likely responses. In the main survey, participants could then classify their own responses within these categories. To create the categories, a research assistant, blind to our hypothesis, coded responses to the pre-test by creating a category for any response mentioned at least three times. This produced six broad categories: *'stock market uncertainty'*, *'expected stock market returns'*, *'personal financial well-being'*, *'economic/political climate'*, *'investment-related costs'*, and *'expert guidance/advice'*. Based on these categories and the specific thoughts described within them, we created follow-up questions asking people whether and how they considered that category in their response.

Thus our main experiment involved two parts, which were completed by 285 participants on MTurk. First, participants answered the open-ended prompt about investments. After doing so, participants selected as many or as few of these six categories as they thought were captured in their response. There was also an option to select *'none of the above'* and a text box where participants could describe what they meant.[20] Allowing participants to self-code their own responses ensures that we are not misinterpreting their responses and is consistent with recent research methods for interpreting open-ended responses (Johnson, Häubl, and Keinan, 2007; Sussman and Oppenheimer, 2020). After selecting a given category, participants were prompted with further questions specific to each one. The Internet Appendix includes all questions for all of the categories.

To illustrate the questioning, we use personal financial well-being as an example. Many of the free-response answers related to personal well-being along a variety of dimensions, such as personal savings, income level, and home values among others. If a subject classified their free response as related to their personal financial well-being, they were then prompted to select the option that best describes their response which included:[21]

- Overall wealth: Do I have enough money to invest in the stock market?
- Uncertainty: Is the stock market too uncertain for my personal financial well-being?
- Relationship with the stock market: How will changes in my personal financial well-being correspond to changes in stock market value?
- Other (please specify)

A participant investing with a goal of insurance will click the *'relationship with the stock market'* option. If a participant clicked on the relationship option they were then asked, when considering the

[19] In the pilot we provided distinct lines to make it more straightforward to categorize individual responses. In the main experiment, participants entered their response into one text box.

[20] The six categories were counterbalanced and *'none of the above'* always appeared as the last option.

[21] The questions (and the follow-up questions) were counterbalanced and *'Other (please specify)'* appeared as the last option with a textbox to describe what they meant.

relationship, which best describes when they would invest more money in stocks:

- Similar: Invest more in stocks if changes in my personal financial well-being move in the *same* direction as the stock market
- Different: Invest more in stocks if changes in my personal financial well-being move in the *opposite* direction as the stock market
- Other (please specify)

While this structure roughly corresponds to the statistical concept of correlation in the baseline experiment, the question in this treatment is purely conceptual and uses no statistical terms. If similar results to the baseline treatment are found using this methodology, it provides further support that the baseline results are not driven by confusion about statistical terms.

Results

Examining the free responses from the first portion of the experiment, we found no evidence consistent with a strong insurance motive. Only two participants used the word *'insurance'* in their responses, and neither response related to using investments as insurance.[22] Only one participant used the word *'correlation'*, and again he did not use this term in any way related to hedging risk. As far as we can tell, none of the responses of the 285 participants directly talked about trying to insure consumption shocks or shocks to any other risk factor.

We present the results of the self-classification portion of our open-ended survey in Table 10. Columns (1) and (2) show that the majority of participants classified at least one of their responses as having something to do with their *'personal financial well-being'* (68%), with *'stock-market uncertainty'* (64%), or with *'expected stock-market returns'* (56%). Significantly fewer investors reported caring about the *'economic/political climate'* (20%), *'expert guidance/advice'* (20%), or *'investment-related costs'* (19%). Only 3 participants indicated that our categories did not capture their motives by selecting the *'none of the above'* option.

Similar to our baseline results and consistent with textbook finance theory, we find that investors report strongly caring about expected returns and return volatility. Further, we find that they do so in the direction that textbook theory predicts.[23] Column (4) shows that 85% of the participants who classified their response as related to expected returns were trying to invest more when returns were higher. Likewise, Column (3) shows that 77% of the participants who classified their response as related to return volatility were trying to invest more when volatility was lower. Consistent with our baseline results, we again find strong evidence for these two core assumptions of textbook theory.

However, just as in our baseline experiment, the free-response survey provides no evidence that participants are trying to insure shocks to consumption when investing in the stock market. While

[22] Both responses related to being worried about having enough money to be able to afford health insurance.

[23] This is consistent with our reading of the free-response answers. Many answers broadly referenced wanting to invest more if returns were expected to be high and being afraid to invest more if they thought volatility would be high.

the majority of participants indicated that their responses were related to their *'personal financial well-being'*, most responses related to their level of wealth, or how uncertain they viewed it to be. Only 14% of participants, 40 in total, reported thinking about the relative timing of stock-market returns. Of these 40 participants, most reported trying to invest more in stocks when stock returns were more correlated with changes in their own finances, the exact opposite of what a textbook investor would do. Out of 285 total participants, only 7 reported caring about correlation between their personal financial situation and the stock market in a manner consistent with textbook models.

Only 58 of the 285 participants reported considering the *'economic/political climate'*. Of this group, again, most reported caring about the expected level of the variable or its uncertainty. Only 24 respondents said that timing of changes in the variable relative to changes in stock returns was important to them. Of these, only 3 reported that they would invest more when these two variables moved in opposite directions as dictated by textbook financial theory.

The research design of the open-ended format is quite different from our baseline framework, but the results are materially the same. Of the 285 participants, we found 10 who reported thinking about the timing of some variables in a manner consistent with an insurance motive. Using this open-ended methodology we find no evidence that an insurance motive guides how people invest in the stock market. These results make it less likely that our baseline results were driven by concerns specific to that design.

4.2 Information Demand

The second alternative research design we explore is a systematic investigation of the information displayed to investors. Investors who care about risk-factor correlations would presumably like this information to be presented as clearly as possible. Further, one might expect there to be widespread discussion of how best to present this information and tools to accomplish this insurance goal. Examining a variety of information sources, we find that data on risk factors are not easily accessible, that there is minimal discussion of risk-factor insurance, and that tools to implement such insurance are not widely available. From a revealed-preference perspective this suggests that real-world investors do not demand this information.

If an investor from your favorite asset-pricing model were to peruse a popular financial-news source, he would likely be puzzled by the lack of information related to correlations with macroeconomic variables. Suppose that investors did not care only about correlations with consumption growth, but instead cared about both correlations with consumption growth and correlations with slow moving habit or with estimates of long-run risks. There is no reason why such correlations could not also be widely reported. Thus the absence of such statistics or any discussion of them in newspapers suggests that it is unlikely that investors' core strategy is related to a risk-factor insurance motive.

While puzzling, it could be that real-world investors do care about risk-factor correlations in spite of the fact that they are not widely reported. To examine whether the information demanded by investors suggests that this is plausible, we studied a variety of sources to search for evidence that investors view

risk-factor correlations as an important component of their investment decisions. Financial authorities provide educational documents to investors describing the risks that investors might consider relevant for their investments. We examined documents from FINRA, the SEC, the Financial Conduct Authority, and the Ontario Security Commission.[24] While these documents list many risks, they provide no discussion about how investors might insure aggregate shocks by holding assets whose returns are less correlated with these shocks. For example, in *'Investment Risk, Explained'* provided by FINRA, there are nine specific sources of risk listed for investors to consider. The discussions surrounding each of these risks pertain to uncertainty and volatility. There is no discussion of hedging or insuring shocks, either to consumption or to any other risk factor.

If the more relevant group is professional investors, it is possible that such documents simply reflect the viewpoints of uneducated investors. Professional investors use a variety of risk-assessment tools, such as those included in Bloomberg. While these products contain a number of options to assess various aspects of portfolio risk, none calculates correlations with macro risk factors as a default input.[25] If professional investors viewed such correlations as an important aspect of how they form their own portfolios or their clients' portfolios, it would be quite surprising that the tools they used lacked this basic feature.

To analyze a more systematic source of information, we examined mutual-fund prospectuses. A fund is required to report its investment objectives and risks regardless of how newsworthy these objectives and risks are. Funds also have discretion to highlight a variety of other potential aspects of the fund. For example, each Vanguard fund includes a *'plain talk'* section in its prospectus that attempts to explain investing concepts or strategies using straightforward language. Thus, if a fund thought that its correlation with an aggregate risk factor was an important component of investors' decision making, it could and should present information about this statistic in its prospectus. If a fund had correlations with macroeconomic risks that would make investors want to buy more of the fund, then it would likely say so in its prospectus to drive flows. If a fund's legal department believed that there was some possibility of being sued by an investor who viewed a correlation with a risk factor as a relevant risk, then including it in the list of potential risks would be an obvious step to minimize liabilities from such a suit.

Mutual funds do not report these numbers or discuss their correlations with macro risk factors in their prospectuses. For example, the Vanguard 500 (VFIAX), which has nearly $500*b* in assets under

[24] For an example from each agency, see *Investment Risk, Explained* provided by FINRA (https://www.finra.org/investors/insights/investment-risk), *What is Risk* provided by the SEC (www.investor.gov/introduction-investing/investing-basics/what-risk), *Assessing suitability: Establishing the risk a customer is willing and able to take and making a suitable investment selection* provided by the FCA (www.fca.org.uk/publication/finalised-guidance/fsa-fg11-05.pdf), and *9 types of investment risk* from the Ontario Security Commission (www.getsmarteraboutmoney.ca/invest/investing-basics/understanding-risk/types-of-investment-risk/).

[25] We examined documentation from Bloomberg and Factset. We also discussed the details of these tools with a number of industry participants at a variety of different large financial firms. None reported any standard option in any of the available tools which would be relevant for hedging macro risk factors, such as aggregate consumption growth.

management, stated its investment objective as tracking a benchmark index. It did not discuss any aggregate risk factors or its correlation with such variables. Under principle risks, the fund lists stock-market risk, "which is the chance that stock prices overall will decline," as well as investment-style risk, "which is the chance that returns from large-capitalization stocks will trail returns from the overall stock market," but it never talks about exposure to aggregate risk factors. There is further discussion of risks later in the prospectus, but this is largely related to volatility: "stock markets tend to move in cycles, with periods of rising prices and periods of falling prices." Funds report a variety of statistics about their past performance such as fees, taxes, distributions, and performance. The Vanguard 500 fund's prospectus reports 171 numeric values in tables and figures, with even more values in the text. None of these numbers corresponds to the correlation between the fund's returns and a macro risk factor.

We systematically reviewed the mutual-fund prospectuses of the largest 25 US mutual funds, which jointly held about $\$3.7t$ at the time. Table 11 summarizes the results. We ranked US open-ended funds listed on Morningstar Direct as of July 30, 2019 based on their share-class asset value. We examined each fund's prospectus for five characteristics related to risk-factor correlations. Did a fund report a numerical value for the correlation between the fund's performance and any macroeconomic variable? Did a fund graph its performance with any macroeconomic variable? In the section on risks, did a fund list its return correlation with any macroeconomic variable other than the stock market itself? In the section on objectives, did a fund list an objective related to its correlation with a macroeconomic variable, such as exposure to an aggregate risk? Finally, we searched the text of each prospectus for the words *'covary'*, *'covariance'*, *'correlate'*, and *'correlation'*, counting the number of times these words appear in the document.

Mutual-fund prospectuses lack information about how a fund's returns covary with aggregate risk factors. None of the funds report numeric or graphical information relating their performance to macroeconomic fundamentals. No fund lists its correlation with macroeconomic outcomes as an investment risk, and no fund lists hedging an aggregate risk factor as an investment objective. Perhaps the closest thing we find is that some funds list reasons why the market may be volatile. For example, the Fidelity Contra Fund (FCNTX) warns its investors that "stock markets are volatile and can decline significantly in response to adverse issuer, political, regulatory, market, or economic developments." Notably, these are descriptions of why there may be volatility in returns not of how the fund's returns might be correlated with such variables. Our word search reveals that 22 of 25 funds fail to use any word related to correlation or covariance in their prospectus. The 3 prospectuses that do contain one of these words use them in a way that is unrelated to macroeconomic risk, as they reference the fund's tracking error or its relationship to derivative securities.

We note that the evidence in this section should be interpreted with caution in that we largely focused on settings that would report aggregate correlations if investors were concerned with them. It is of course possible that investors focus on their own personal consumption or other idiosyncratic risks. If this were

the case though, it is unclear whether mutual funds would discuss such concerns.[26] Given these issues we think the evidence in this section is best considered in conjunction with the experiments in this paper, as it is not conclusive on its own.

Textbook investors view the stock market as a way to insure bad states of the world. However, risk-factor correlations are almost never reported in the financial news, are not discussed in important fund documents, and are not used as the default settings in professional risk-analysis tools. Their absence from these sources suggests that investors do not demand this information. It suggests this information is not relevant to their investment decisions.

5 Discussion

When we use our survey-based framework to study consumption growth, we find no evidence that investors view this canonical risk factor as relevant to their portfolio decisions. This section aims to put this finding into the appropriate context. Researchers interpret factor models in different ways. We start by discussing what researchers can learn from our framework under these various interpretations. Next we talk about concerns posed by equilibrium effects. We then describe how our results connect to the existing literature in cross-sectional asset pricing and macro-finance. Finally, we outline how researchers can use our approach to guide model development going forward.

Our discussion relates to a model's ability to explain and predict asset-pricing data, but clearly there are other important uses of models. For example, models can be used to make normative statements about what investors should be attending to in their investments even if they are not currently doing so. While such an insight would not be relevant for understanding how asset prices move (if people are not actually engaging in such behavior), they can offer important insights for building future financial products or emphasizing issues for investor education.

5.1 Model Interpretations

How survey evidence relates to testing asset pricing models depends on the interpretation of why these models are written and how they are used. When a researcher claims that exposure to a risk factor explains the data, he could mean a variety of different things. This section discusses the distinct interpretations of these models and the implications of survey based evidence under each.

Model Reflects The Economic Problem Investors Are Trying To Solve

The first interpretation we discuss, and the one we largely focus on in this paper, is that these asset-pricing models are meant to accurately reflect the economic problem investors are trying to solve. We focus on this interpretation because this is the one favored by the literature. Understanding the

[26] With that said, it seems likely that there would be other financial products tailored to satisfy investors' demand for information about idiosyncratic risks. These products seem largely absent from financial markets. For example, products could be tailored based on professions. Different products might be offered for professions with high exposure to financial markets, such as people working in finance, or low exposure to financial markets, such as government employees.

economic mechanism behind why prices move is important in its own right. Further, a model that reflects the problem that investors are actually trying to solve is more likely to make accurate predictions in novel as-yet-unseen market environments.[27] To summarize the purpose of asset-pricing models, in a recent review article Cochrane (2017) writes that "the challenge is not one of telling stories or 'explaining' facts…ex post" but rather one of finding "explicit measures of fearful outcomes…that quantitatively account for asset pricing facts."[28]

In addition to such explicit statements, much of the discussion, motivation and interpretation of results from these models is only coherent if the models capture the economic problem being solved by agents. For example, perhaps the most common description of the equity premium puzzle (Mehra and Prescott, 1985) is that the risk aversion needed to match the data is implausibly high relative to evidence on how humans respond to risk. However, if asset-pricing models are not meant to reflect actual human behavior, then there is no need for researchers to choose a model's risk-aversion parameter based on estimates of how humans actually behave.[29] Models in this literature are often motivated based on human psychology including examples such as human beings wanting to "keep up with the Joneses" (Abel, 1990) or that hedge-fund managers would "sooner die than fly commercial again" (Campbell and Cochrane, 1999). These motivations only make sense if models are meant to capture human behavior.

Before seeing our results, it would be reasonable to think that factor models capture the economic problem investors are trying to solve. There are examples in other settings where people use financial markets to hedge a given risk, and in these situations investors are typically able to provide evidence that they are implementing a strategy to do so. For example, investors trading oil futures know that they are doing so as insurance against future changes to oil prices and that today's price is influenced by exposure to future oil shocks.[30] When the CEO of Southwest airlines was asked about why they were active in the oil-futures market, he stated that the company "loaded up years ago on hedges against higher fuel prices."[31] This risk is commonly understood, which is why it is priced.

But the fact that some asset markets operate like insurance markets does not imply that all markets operate on the same principles or that all insurance goes through asset markets. While investors could

[27] Whereas, the predictions of a 'just so' model are likely to break down out of sample (Manski, 1995).

[28] There are many other similar quotes throughout the asset-pricing literature. For instance, Lewellen, Nagel, and Shanken (2010) states that the fact that returns can be expressed as $\boldsymbol{R} = \boldsymbol{BF} + \boldsymbol{e}$, where $\boldsymbol{R}$ stands for returns and $\boldsymbol{F}$ stands for a risk factor, on its own "has no economic content since an appropriate $\boldsymbol{F}$ can always be found; for example, any K portfolios that span the tangency portfolio would work."

[29] Friedman (1953) considers modeling the distribution of leaves on a tree *'as if'* this distribution were optimally chosen by the tree to capture the most sunlight. It would not make sense to reject such an *'as if'* model because, say, the tree would have to have an implausibly high level of risk aversion when compared with humans to fit the observed leaf distribution.

[30] Other examples include the following: Cheng and Xiong (2014) states that "commodity futures markets have had a long history of assisting commodity producers to hedge their commodity-price risks." Further, "exchange rates are a major source of uncertainty for multinationals (Jorion, 1990)" and FX forward markets exist so firms can hedge this risk. Likewise, "sovereign CDS contracts function as insurance contracts that allow investors to buy protection against the event that a sovereign defaults (Longstaff, Pan, Pedersen, and Singleton, 2011)."

[31] Bailey, J. "Southwest Airlines gains advantage by hedging long-term oil contracts." *New York Times.* Nov 28, 2007.

attempt to use their portfolio for auto insurance, most investors purchase this product elsewhere. Further, just because there is a risk that a theoretical investor would want to insure against, this does not mean that actual investors are doing so.

Under this standard interpretation, investors should be able to describe which risk factors they are trying to insure against. This represents a testable necessary condition of any factor model. Thus this paper provides evidence against using a broad class of asset-pricing models with this interpretation.

It is of course possible that a given survey, such as the one in this paper, is flawed and leads to erroneous conclusions. We have discussed why we think our design is consistent with best practices and is a good test of these models. If there were a better test, it would be a valuable contribution for future research to construct a well designed survey that illustrated why the results of the current study are flawed. Doing so would not only provide evidence for the necessary condition of the models, but it also would likely illustrate mechanisms not currently understood, and thus missed by this paper. Providing such survey evidence would lead to new insights that future models could build upon. Further, under the economic motive interpretation of models, it should be possible to do so.

Model is an 'As If' Description

An alternative interpretation is that factors models do not capture the motives of individuals. Instead, they assume the data arise from the collective actions of individuals behaving as if they were governed by the underlying mechanism (Friedman, 1953). Individuals could be doing anything, but as long as assets with more consumption-risk exposure also have higher average returns in the data, we would say that consumption growth is a relevant risk factor under this interpretation.[32]

For some research questions, simply estimating this empirical relation and matching the data is illuminating. For example, there are literatures that examine purely empirical representations of the stochastic discount factor and that study latent factor models without making any claim as to what these factors represent (e.g., Fama and French, 1993; Engle, 1982; Ang, Hodrick, Xing, and Zhang, 2006; Kozak, Nagel, and Santosh, 2018a; Kelly, Pruitt, and Su, 2019).

Models used in such a fashion are typically judged based on their empirical performance. The empirical performance of consumption-based asset-pricing models is significantly worse than their purely empirical counterparts (Ross, 1976; Nagel, 2013). An asset-pricing model can be helpful because it fits the data or because it captures the essence of how investors price assets. The inferior empirical performance of consumption-based models makes it difficult to justify their use based on empirical performance alone, so their contribution is typically viewed as reflecting the economic mechanism at

[32]Friedman's assertion was not universally embraced at the time or at present. For example, Samuelson (1963) colorfully illustrates the logical inconsistency of Friedman's approach. He derisively refers to the idea as the *'F-twist'* because he did not want to sully Friedman's name with the scathing critique. Samuelson writes that "Good science discerns regularities and simplicities that are there in reality." Moreover, other subfields of economics have largely rejected Friedman's approach. For example, Larry Katz recently tweeted that an essential rule for "research in labor economics" is to "talk to economic actors (workers, employers, . . .) – social scientists can talk to our actors unlike physicists."

work in real-world financial markets.[33]

If one argues for using factor models based on *'as if'* reasoning, it is important to recognize that there are fewer uses of models under this interpretation. If the goal in writing a model is to understand why a pattern exists in the data, such an understanding can only occur when a model accurately captures the economic behavior explaining why. The literature has argued that models can help guide empirical work, but this guidance is meaningful primarily when a model captures what real-world investors are trying to do.[34] It is common for empirical researchers to use factor models to compute risk-adjusted returns, but interpreting such an analysis as controlling for risk only makes sense if investors are trying to insure against the associated bad states of the world. Further, if the model is not meant to capture behavior, it does not make sense to motivate it or describe it in terms of human behavior. Thus we contribute to the literature by ruling out these common uses of factor models in the absence of evidence of a risk factor's relevance.

A less stringent interpretation of Friedman (1953) is that it says it is acceptable to abstract away from real-world complexities that are not first-order important. Simplifying assumptions are an important part of any model by definition, so the relevant question becomes: Is a given assumption merely a simplification, or does removing the assumption remove a core aspect of the problem? As Friedman himself puts it: "The important problem... is to specify the circumstances under which the formula works, or, more precisely, the general magnitude of the error in its predictions."

Friedman (1953) gives three examples illustrating how he thinks about the answer. The first illustrates that one can assume away aspects of a problem that are not first-order important by discussing when it is appropriate or not for a physicist to assume away the effect of air friction. When modeling a falling compact ball, the effect of air friction can be safely ignored. The same cannot be said when modeling a falling feather. Our paper examines the core strategy investors implement in a factor model, so this strategy cannot be ignored and treated like an innocuous simplifying assumption.

The second example is meant to illustrate the disciplining influence of economic constraints by examining how a tree orients its leaves relative to the sun. In this example, evolutionary constraints lead to nearly optimal outcomes even though trees clearly lack rational decision-making abilities. This represents one of the more popular applications of the Friedman critique, with Becker (1962) illustrating that budget constraints can lead markets populated by irrational actors to behave as if they were populated by rational ones. However, given that consumption-based asset-pricing models function through expectations of future outcomes, it is unclear what economic constraints would induce investors to behave as if they were trying to insure consumption risk. Thus it is again unclear whether this example

[33] Cochrane (2017) writes that, "in explaining which models become popular throughout economics, tractability, elegance, and parsimony matter more than probability values of test statistics."

[34] Cochrane (2005) states that "we can always construct a reference portfolio that perfectly fits all asset returns... The only content to empirical work in asset pricing is what constraints the author put on his fishing expedition... The main fishing constraint is that the factor portfolios are in fact mimicking portfolios for some well-understood macroeconomic risk."

is relevant for our setting.

The third example, which is likely most relevant for our setting, involves an expert billiard player who makes shots as if he understands the underlying mathematical formulas even though he does not. A formal model would assume the billiards player makes his shots by optimizing complex mathematical formulas, while Friedman claims the player "just figures it out" and "rubs a rabbit foot."[35]

While an expert billiards player will certainly have developed an intuitive feel for the game, there are still many aspects of his thought process that a researcher would be able to identify directly. He would be able to state that he was playing billiards, the rules of the game and the placement of the balls on the table. Based on these characteristics he would be able to describe the strategy he was trying to implement and the desired outcome of his next shot. Thus, even an intuitive player should be able to describe the game he is playing, his goal in playing the game, the strategy he is using to achieve that goal, and the values of the relevant input variables to that strategy.

Our paper shows that people are not shown correlations, do not consider correlations when they are provided, do not respond to correlations when making investments, and do not state that insurance of any kind is a goal in a free response. This is akin to an expert billiards player who does not know the rules of the game, does not understand the strategy he is trying to implement, and does not know the position of the balls on the table. A situation where the goal, the strategy, and the relevant parameters are unknown decreases the plausibility that an alternative decision rule leading to *'as if'* outcomes exists.

Moreover, even under an *'as if'* interpretation, survey evidence likely still will provide important insights. Suppose that investors are pricing assets based on some alternative decision rule and that a factor model approximates the outcomes of this alternative rule. Understanding the alternative decision rule that investors are actually using would likely lead to new and important insights for asset-pricing theory. The best way to learn about such an alternative decision rule that investors are using is to ask them about it.

5.2 Equilibrium Effects

In general equilibrium models, it is often possible to express the influence of certain variables without including the variable itself. For example, investors in Campbell and Viceira (1999) have an optimal demand rule that is a decreasing function of an asset's correlation with consumption growth in partial equilibrium. However, in general equilibrium, the authors show that it is possible to express this optimal demand rule without any correlation parameters by substituting in the budget constraint. Thus an investor in such a model would be able to describe outcomes related to their correlation based demand rule, without having to cite correlations themselves.

[35] We note that surveys would be helpful here in understanding what expert billiard players actually understand. When you ask a professional billiard player about what he is doing, you often get responses based on the complicated math. e.g., there is a popular book on the subject called *The Science of Pocket Billiards* (Koehler, 1995). There is simply no substitute for asking people what they are doing; this is true both of investors and of pool sharks.

A benefit of using surveys is that we are not restricted to equilibrium outcomes. By examining outcomes off the equilibrium path, this breaks such a variable-elimination argument. Put another way, Campbell and Viceira (1999) investors should jump at the opportunity to buy an asset with high average returns and a negative correlation with consumption growth. They will never encounter such assets in equilibrium, but the logic of the model dictates how they should react if they did. We surveyed investors about what they would do in such a situation. Unlike the theoretical investors in Campbell and Viceira (1999), we find no evidence that real-world investors desire such assets.

5.3 Fama-French Factors

One of the richest sources of empirical success in asset pricing lies in explaining cross-sectional patterns in returns. One of the benefits of our framework is that it is simple to apply to almost any proposed risk factor. We demonstrates this by applying our framework to the three most studied risk factors from the cross-sectional asset-pricing literature.

Specifically, we focus on the excess return on the market, the return to a small-minus-big (SMB) size factor, and the return to a high-minus-low (HML) value factor Fama and French (1993). Labeling these sources of return predictability as risk factors implies that they represent non-diversifiable risks investors would like to insure against. Thus, all else equal, an asset that is more correlated with one of these risk factors will offer worse insurance and investors should be less inclined to hold it.[36]

To examine cross-sectional variation, we change our survey from asking about investing in the aggregate stock market to asking about investing in a generic mutual fund. We do so because it would be difficult to examine whether investors viewed the market as a risk factor if they could only invest in the market. Instead of economic growth, we present investors with monthly excess returns from one of the three factors from Ken French's website.[37] We tell participants that the mutual fund returns are simulated using different parameter values each period. We provide definitions for each, and outside of these changes re-run our experiment in a similar manner to our baseline treatment.

Table 12 reports results for the three Fama-French factors. Each column represents the estimated coefficients from regressing the fraction invested in the mutual fund on the fund's average returns, its return volatility, and the correlation between the fund's returns and a particular Fama-French factor. For each factor, we ran a separate survey on a different population of MTurkers.

Similar to our main results, participants strongly respond to changes in the mean and volatility of a

[36] Another interesting extension would be to use our framework to directly explore multifactor models. The Fama-French factors are designed to be largely uncorrelated with one another, but in other instances the correlation across factors could be important (e.g., for value and momentum). It is possible to extend our framework to include information on the multiple risk factors and the relation between these factors. We leave this for future research.

[37] This is a test of whether investors view the given portfolio as something to be insured against. However, it is not clear that a real-world investor would understand that a portfolio's high average returns (e.g., value) are due to its higher exposure to an underlying risk (e.g., correlation with recessions). There is no commonly agreed on economic source of risk for the Fama-French factors. If a future researcher proposes a source of risk to account for them, he should use our framework to test whether investors want to insure against it.

mutual fund's returns. The coefficients on $\texttt{mean}_{i,q}$ and $\texttt{volatility}_{i,q}$ are statistically significant, economically large, and directionally consistent with textbook models. However, we find no evidence of the negative coefficient on $\texttt{correlation}_{i,q}$ predicted by theory. If anything, participants invest more in a mutual fund when its returns are more correlated with one of these factors.

We find similar results when we look at the economic-reasoning portion of our framework in Table 13. Just like before, participants are more likely to report thinking about the mean and volatility of a mutual fund's returns than about its correlation with any of the three Fama-French factors. Moreover, column (9) shows that participants who reported thinking about exposure to a Fama-French factor typically did so using the opposite of textbook logic.

The idea that return predictability must represent compensation for risk is so ingrained in academic finance that empirical regularities are called a risk factors almost without thought. This occurs even though there is evidence that at least some of this predictability seems consistent with mispricing (McLean and Pontiff, 2016). This semantic issue underscores that academic finance tends to focus solely on the necessary, but not sufficient, econometric relation when labeling cross-sectional return predictability a risk factor. We suggest using a more agnostic term, such as predictable returns, when a pattern is first discovered in the data. A strong empirical relation on its own should not be considered sufficient evidence for the risk factor label. Evidence needs to be provided, such as that from our framework, that investors view these returns as compensation for exposure to a risk factor before the risk factor label can be accurately applied.

5.4 Centrality of Consumption Hedging

The framework we develop in this paper can be used to evaluate the relevance of nearly any proposed risk factor. To demonstrate the framework, however, we had to select a specific risk factor for our case study. We chose consumption growth as our main variable of interest. Consumption growth is the sole state variable of the CCAPM, but this is not the main reason why we selected it. In this section, we discuss how results for this variable have implications for most modern asset-pricing models.

To address the empirical failings of the CCAPM, modern macro-finance models (e.g., habit formation (Campbell and Cochrane, 1999), long-run risks (Bansal and Yaron, 2004), rare disasters (Rietz, 1988; Barro, 2006), heterogeneous agents (Constantinides and Duffie, 1996), etc) introduce new mechanisms that amplify the influence of consumption. The idea is that, if exposure to consumption growth cannot fully account for why markets fluctuate, then exposure to consumption growth interacted with an additional state variable can. According to Cochrane (2017), "each of them [the new models] boils down to a generalization of marginal utility or discount factor, most of the same form $M_{t+1} = \delta \cdot (\frac{C_{t+1}}{C_t})^{-\gamma} \cdot X_{t+1}$" where X_{t+1} represents the new state variable of each model. Empirically, consumption growth is not very volatile, so in essence these new state variables serve to amplify the core concern of consumption hedging to better match the data.

Before discussing the specific models, it is worth emphasizing that consumption-based models

represent the current dominant paradigm of asset pricing. While opinions in the field vary as to whether this focus on consumption-based asset pricing is warranted, the revealed preferences of the field show that such models represent the majority of asset-pricing papers currently being published and circulated at the most prestigious outlets. To provide empirical evidence of this, we examined all papers from recent NBER asset-pricing meetings (five meetings during 2019-2020) as well as those published in the Journal of Finance (six issues during 2020) to see if they included a model which implied consumption risk should be priced.[38] For the NBER, we found that more than 80% of papers with models and more than 40% of all papers implied consumption risk should be priced. For the Journal of Finance, we found that 60% of papers with asset-pricing models and more than 20% of all papers (including non-asset-pricing papers) did the same.

Habit Formation

The Campbell and Cochrane (1999) model studies a representative investor with power utility, $U_t \stackrel{\text{def}}{=} \frac{(C_t - H_t)^{1-\gamma}-1}{1-\gamma}$, over consumption in excess of a slow-moving benchmark, H_t—i.e., habit; the level of consumption investors have become accustomed to, $\log H_t \stackrel{\text{def}}{=} \lambda \cdot \sum_{\ell=0}^{\infty} \phi^{\ell} \cdot \log C_{t-\ell}$ where $\lambda > 0, \phi \in (0,1)$. The idea is to make drops in consumption following booms more painful to investors. The key state variable in this model is investors' surplus-consumption ratio, $X_t \stackrel{\text{def}}{=} \frac{C_t - H_t}{C_t}$. The stochastic discount factor (SDF) is then given by $\Delta \log M_{t+1} = \log \delta - \gamma \cdot \Delta \log C_{t+1} - \gamma \cdot \Delta \log X_{t+1}$.

ICAPM logic seems to indicate that average stock returns could be high either because they covary with consumption growth or because they covary with growth in the surplus-consumption ratio:

$$\mathbb{E}[R_{t+1}] - R_f \approx \gamma \times \mathbb{C}\text{ov}[\Delta \log C_{t+1}, R_{t+1}] + \gamma \times \mathbb{C}\text{ov}[\Delta \log X_{t+1}, R_{t+1}]$$

But *'either/or'* is not the right conjunction. The second term on the right hand side is not independent of the first. If investors are not trying to insure drops in consumption (if first term is zero), they cannot be trying to insure drops in surplus consumption (second term must be zero). The surplus-consumption ratio is not a separate risk factor. It is a way of amplifying the effects of consumption risk.

In fact, expected returns in Campbell and Cochrane (1999) can be rewritten as $\mathbb{E}[R_{t+1}] - R_f \approx \frac{\gamma}{X_t} \times \mathbb{C}\text{ov}[\Delta \log C_{t+1}, R_{t+1}]$. Thus, habit formation will increase the influence of consumption risk so long as the surplus-consumption ratio, X_t, tends to be less than one. Campbell and Cochrane (1999) use $\phi = 0.87$ in their calibration. With $\sigma_{\Delta \log C} = 1\%$ per year and $\gamma = 10$, the average surplus consumption ratio is $\mathbb{E}[X_t] = \sigma_{\Delta \log C} \cdot \sqrt{\frac{\gamma}{1-\phi}} \approx 0.088$. So, in an external habit model with $\sigma_R = 16\%$ per year, the effect of changing exposure to consumption risk is:

$$\partial_{\rho}(\mathbb{E}[R_{t+1}] - R_f) = \frac{10}{0.088} \times 0.01 \cdot 0.16 \approx 0.18$$

[38] This methodology was meant to be conservative as it ignores any paper without a model, even though some of these papers are testing consumption-based asset-pricing concepts. See the Internet Appendix for additional details.

These calculations imply that the $\Delta\rho = 0.45$ increase we study in our framework should increase the expected return on the stock market by $0.18 \times 0.45 \approx 8\%$. The fact that participants in our study do not adjust their demand in response to such correlation changes is inconsistent with this model.

Long-Run Risk

The Bansal and Yaron (2004) long-run-risk model uses a different preference specification and state variable, but the result is the same. The effects of shocks to the new state variable on asset prices cannot exist if investors do not want to insure their exposure to consumption-growth shocks. One way to see this is to notice that the long-run-risk model is formally equivalent to a model where investors are ambiguity averse with respect to parameters of the consumption-growth process (Hansen and Sargent, 2008; Epstein and Schneider, 2010; Bidder and Dew-Becker, 2016).

Let P_t denote the current price of an asset whose payout is aggregate consumption in the following period, and let $1/\alpha$ denote investors' elasticity of intertemporal substitution (EIS). The long-run risk model says that the equity premium will be determined as follows:

$$\mathbb{E}[R_{t+1}] - R_f \approx \gamma \times \mathbb{C}\text{ov}[\,\Delta \log C_{t+1}, R_{t+1}\,] + \text{f}(\gamma, \alpha) \times \mathbb{C}\text{ov}[\,\log(P/C)_{t+1}, R_{t+1}\,]$$

where $\text{f}(\gamma, \alpha) \leq 0$ comes from the Campbell and Shiller (1988) approximation of $(P/C)_t$. Thus, ICAPM logic seems to suggest that expected returns could be high either because stock returns covary with consumption growth or because they covary with the aggregate price-to-consumption ratio.

But again *'either/or'* is not the right conjunction. The second term on the right hand side is not independent of the first. When $1/\alpha \approx 1$, things simplify to

$$\mathbb{E}[R_{t+1}] - R_f \approx \gamma \times \mathbb{C}\text{ov}[\,\Delta \log C_{t+1}, R_{t+1}\,] + (1-\gamma) \times \mathbb{C}\text{ov}[\,\log(P/C)_{t+1}, R_{t+1}\,]$$

Prices are forward looking, so changing an asset's correlation with consumption growth also changes its correlation with the price-to-dividend ratio. We estimate that $\log(P/C)_{t+1} = 3.61 - 30.26 \cdot \Delta \log C_{t+1} + \varepsilon_{t+1}$ using quarterly data from Q1 1948 through Q2 2020. Substituting in this linear approximation to $\log(P/C)_{t+1}$ simplifies things even further:

$$\mathbb{E}[R_{t+1}] - R_f \approx \{\gamma - (1-\gamma) \cdot 30.26\} \times \mathbb{C}\text{ov}[\,\Delta \log C_{t+1}, R_{t+1}\,]$$

Using $\gamma = 10$, $\sigma_{\Delta \log C} = 1\%$, and $\sigma_R = 16\%$ gives $\partial_\rho(\mathbb{E}[R_{t+1}] - R_f) \approx 0.45$. Thus, the $\Delta\rho = 0.45$ increase that we study in our survey-based framework would increase annual expected excess returns by $0.45 \times 0.45 \approx 20\%$ in the long-run risk model. The fact that participants in our survey do not respond to such correlation changes is inconsistent with this model.

Rare Disasters

There is a class of models built on top of the CCAPM framework that our results do not directly speak to: rare-disaster models à la Rietz (1988), Barro (2006), and Gabaix (2012). We find that investors are not trying to insure normal-times variation in consumption growth, but we do not directly test whether investors want to insure themselves against extreme shocks to consumption.

That being said, other survey evidence already suggests that real-world investors do not think like investors in a rare-disasters model. For example, Giglio, Maggiori, Stroebel, and Utkus (2020) finds that "expected returns and the subjective probability of rare disasters are negatively related, both within and across investors"—i.e., investors do not demand a disaster-risk premium as predicted by a rare-disasters model. Researchers could use the survey-based framework we develop in this paper to more directly test whether investors follow the economic logic behind this model.

Heterogeneous Agents

Most of the discussion in our paper (and in the literature) relates to representative-agent models, but our paper also has implications for heterogeneous-agent models. For example, in the Constantinides and Duffie (1996) model, heterogeneous investors try to insure shocks to their own personal income. When this hedging demand is aggregated, the model predicts that aggregate consumption growth should look like a priced risk factor. We directly show in Tables 4 and 9 that participants are not trying to insure shocks to their own personal consumption, income, wealth, spending, or standard of living.

Investors likely differ along a variety of dimensions, whether it be preferences, wealth, expectations, or something else. While models that focus on these various dimensions have been proposed, most heterogeneous agent models assume that investors are still trying to insure risk factors, be they aggregate or personal. Our results suggest this approach is unlikely to accurately represent how investors actually approach their portfolio decisions.

5.5 Model Development

The results in this paper provide direction for researchers interested in writing new asset-pricing models. We find that investors' desire to insure consumption shocks is unlikely to explain why asset prices move, so models that add complexity to this basic idea are similarly unlikely to be correct.

However, suppose our results had suggested otherwise and participants had strongly responded to changes in the correlation between stock returns and consumption. Further suppose participants had reported thinking about an asset's correlation with consumption growth as suggested by the CCAPM. Such results would not have solved the CCAPM's empirical shortcomings, but they would have supported the literature's standard approach to dealing with these flaws. They would have implied that investors cared about consumption-growth correlations in a more complicated way than is captured by the CCAPM. So models introducing new risk-factor correlations to the CCAPM, such as habit formation (Campbell and Cochrane, 1999) and long-run risk (Bansal and Yaron, 2004), would have a strong

foundation to build on. However, this is not what we find.

That being said, investors are not necessarily irrational or wrong just because they do not follow CCAPM logic. If deep-pocked asset managers at an invite only conference run by a large well-known fund are not trying to insure consumption shocks when investing in the stock market, it does not mean that these investors are doing the wrong thing. It means that economists are using the wrong model. Our results strongly support the textbook assumption that investors view the mean and variance of returns as first order important. The open-ended framework we develop can be used to understand how investors are considering these variables and what else is influencing their decisions. Future model development will be enhanced if surveys similar to the ones used in this paper are used to identify and corroborate investor motives.

6 Conclusion

Under standard interpretations, if X is a priced risk factor, then investors must be *trying* to insure themselves against bad future shocks captured by drops in X. The typical investor should be aware of this goal. This is a necessary condition that must hold for any X to be a relevant risk factor and we develop a survey-based framework for testing it. We apply our framework to evaluate the relevance of aggregate consumption growth and find no evidence that investors prefer assets with lower consumption-growth correlations or adjust their portfolio holdings based on this reasoning.

Unlike settings that attempt to uncover behavioral biases that investors are unaware of, studying rational risk-based models represent the most simple application of surveys. The standard interpretation is that "the average returns from multifactor or market-timing strategies are earned as compensation for holding real aggregate risks that the average investor is anxious not to hold. (Cochrane, 1999)" If a theory says equilibrium prices are moving because most investors are intentionally trying to achieve a certain goal, then it should be straightforward to find some investors who answer questions about this goal consistent with the theory.

This makes the identification of what investors are trying to accomplish with their portfolios ideally suited to being explored using the tools developed in this paper. While this paper focuses on the testing of specific asset-pricing models, the iterative procedure developed in Section 4.1 can be tailored to explore why people are making investment decisions. Better understanding this motivation should aid in developing new insights and models about asset prices.

Our results do not imply that investors do not care about consumption shocks nor do they imply that investors do not understand insurance. Our results show that investors are not trying insure consumption shocks when investing in the stock market.[39] There are a number of mechanisms that could account for this finding. Do investors not want consumption insurance? Do they not view their stock-market

[39] "Tracing a common factor in returns to an economic state variable does not in itself imply that the state variable is of special hedging concern to investors. (Fama and French, 1996)"

portfolio as a way of acquiring this insurance?[40] Do they not understand insurance in this setting? Understanding which of these motivations accounts for the findings in this paper is an important question for future research.

Going forward, when a researcher proposes a new risk factor, X, consumers of this research should ask for evidence that investors actually think and trade based on the logic of his model. Economists typically apply a strict hierarchy of explanations. If there is a risk-based explanation for an asset-pricing phenomenon, then this explanation is viewed as correct even if there are other more parsimonious models. Our results cast doubt on this hierarchy. Researchers should be skeptical of risk-based explanations in the absence of any supporting evidence that the model actually captures how investors price assets. And researchers promoting behavioral biases or trading frictions can apply our framework to address concerns about competing risk-based explanations.

References

Abel, A. (1990). Asset prices under habit formation and catching up with the Joneses. *American Economic Review*.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance*.

Asparouhova, E., P. Bossaerts, N. Roy, and W. Zame (2016). "Lucas" in the laboratory. *Journal of Finance*.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset-pricing puzzles. *Journal of Finance*.

Barro, R. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*.

Becker, G. (1962). Irrational behavior and economic theory. *Journal of Political Economy*.

Behr, D., L. Kaczmirek, W. Bandilla, and M. Braun (2012). Asking probing questions in web surveys: Which factors have an impact on the quality of responses? *Social Science Computer Review*.

Bergman, A., A. Chinco, S. Hartzmark, and A. Sussman (2020). Survey curious? Start-up guide and best practices for online survey experiments. *Working Paper*.

Bidder, R. and I. Dew-Becker (2016). Long-run risk is the worst-case scenario. *American Economic Review*.

Bossaerts, P. and C. Plott (2004). Basic principles of asset-pricing theory: Evidence from large-scale experimental financial markets. *Review of Finance*.

Bryzgalova, S. (2017). Spurious factors in linear asset-pricing models. *Working Paper*.

[40] Investors do not insure against consumption shocks in the stock market. But, then again, they do not buy auto and dental insurance in the stock market either.

Campbell, J. (2003). Consumption-based asset pricing. *Handbook of the Economics of Finance*.

Campbell, J. and J. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock-market behavior. *Journal of Political Economy*.

Campbell, J. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*.

Campbell, J. and L. Viceira (1999). Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics*.

Casler, K., L. Bickel, and E. Hackett (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*.

Cheng, I.-H. and W. Xiong (2014). Financialization of commodity markets. *Annual Review of Financial Economics*.

Chinco, A., A. Neuhierl, and M. Weber (2020). Estimating the anomaly base rate. *Journal of Financial Economics*.

Choi, J. and A. Robertson (2020). What matters to individual investors? Evidence from the horse's mouth. *Journal of Finance*.

Cochrane, J. (1999). Portfolio advice for a multi-factor world. *Chicago Federal Reserve, Economic Perspectives*.

Cochrane, J. (2001). *Asset pricing*. Princeton University Press.

Cochrane, J. (2005). Financial markets and the real economy. *Foundations and Trends in Finance*.

Cochrane, J. (2017). Macro-finance. *Review of Finance*.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2018). How do firms form their expectations? New survey evidence. *American Economic Review*.

Constantinides, G. and D. Duffie (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy*.

Crockett, S., J. Duffy, and Y. Izhakian (2019). An experimental test of the Lucas asset-pricing model. *Review of Economic Studies*.

Di Maggio, M., A. Kermani, and K. Majlesi (2020). Stock-market returns and consumption. *Journal of Finance*.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*.

Enke, B. and F. Zimmermann (2017). Correlation neglect in belief formation. *Review of Economic Studies*.

Epstein, L. and M. Schneider (2010). Ambiguity and asset markets. *Annual Review of Financial Economics*.

Eyster, E. and G. Weizsacker (2016). Correlation neglect in portfolio choice: Lab evidence. *Working Paper*.

Fama, E. and K. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*.

Fama, E. and K. French (1996). Multifactor explanations of asset-pricing anomalies. *Journal of Finance*.

Feng, G., S. Giglio, and D. Xiu (2020). Taming the factor zoo. *Journal of Finance*.

Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics non-parametrically. *Review of Financial Studies*.

Friedman, M. (1953). *Essays in positive economics*. University of Chicago Press.

Frydman, C., S. Hartzmark, and D. Solomon (2017). Rolling mental accounts. *Review of Financial Studies*.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics*.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *Quarterly Journal of Economics*.

Gabaix, X. (2016). Behavioral macroeconomics via sparse dynamic programming. *Working Paper*.

Gennaioli, N., Y. Ma, and A. Shleifer (2016). Expectations and investment. *NBER Macroeconomics Annual*.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2020). Five facts about beliefs and portfolios. *American Economic Review*.

Greenwood, R. and A. Shleifer (2014). Expectations of returns and expected returns. *Review of Financial Studies*.

Grimm, P. (2010). Social desirability bias. *Wiley international encyclopedia of marketing*.

Hansen, L. and T. Sargent (2008). *Robustness*. Princeton University Press.

Hartzmark, S. (2014). The worst, the best, ignoring all the rest: The rank effect and trading behavior. *Review of Financial Studies*.

Hartzmark, S. and D. Solomon (2020). Reconsidering returns. *Working Paper*.

Harvey, C. and Y. Liu (2020). False (and missed) discoveries in financial economics. *Journal of Finance*.

Hauser, D. and N. Schwarz (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject-pool participants. *Behavior Research Methods*.

Hicks, J. (1939). *Value and capital*. Oxford University Press.

Huber, J., S. Palan, and S. Zeisberger (2019). Does investor risk perception drive asset prices in markets? Experimental evidence. *Journal of Banking & Finance*.

Jennings, D., T. Amabile, and L. Ross (1982). Informal covariation assessment: Data-based vs. theory-based judgements. *Judgment Under Uncertainty: Heuristics and Biases*.

Johnson, E., G. Häubl, and A. Keinan (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology*.

Jorion, P. (1990). The exchange-rate exposure of US multinationals. *Journal of Business*.

Kahneman, D. and S. Frederick (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and Biases: The Psychology of Intuitive Judgment*.

Kallir, I. and D. Sonsino (2009). The neglect of correlation in allocation decisions. *Southern Economic Journal*.

Kelly, B., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*.

Koehler, J. (1995). *The science of pocket billiards*. Sportology Publications.

Kozak, S., S. Nagel, and S. Santosh (2018a). Interpreting factor models. *Journal of Finance*.

Kozak, S., S. Nagel, and S. Santosh (2018b). Shrinking the cross-section. *Journal of Financial Economics*.

Kroll, Y., H. Levy, and A. Rapoport (1988). Experimental tests of the separation theorem and the capital asset-pricing model. *American Economic Review*.

Laudenbach, C., M. Ungeheuer, and M. Weber (2019). How to alleviate correlation neglect. *Working Paper*.

Levy, G. and R. Razin (2015). Correlation neglect, voting behavior, and information aggregation. *American Economic Review*.

Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset-pricing tests. *Journal of Financial Economics*.

Liu, H., C. Peng, W. Xiong, and W. Xiong (2020). Resolving the excessive trading puzzle: An integrated approach based on surveys and transactions. *Working Paper*.

Longstaff, F., J. Pan, L. Pedersen, and K. Singleton (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*.

Lucas, R. (1978). Asset prices in an exchange economy. *Econometrica*.

Ludvigson, S. (2013). Advances in consumption-based asset pricing: Empirical tests. *Handbook of the Economics of Finance*.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*.

Manski, C. (1995). *Identification problems in the social sciences*. Harvard University Press.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*.

Matthies, B. (2018). Biases in the perception of covariance. *Working Paper*.

McLean, D. and J. Pontiff (2016). Does academic research destroy stock-return predictability? *Journal of Finance*.

Mehra, R. and E. Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics*.

Nagel, S. (2013). Empirical cross-sectional asset pricing. *Annual Review of Financial Economics*.

O'Cathain, A. and K. Thomas (2004). "Any other comments?" Open questions on questionnaires—a bane or a bonus to research? *BMC medical research methodology*.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance*.

Paolacci, G. and J. Chandler (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*.

Plott, C. and S. Sunder (1988). Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*.

Reja, U., K. Manfreda, V. Hlebec, and V. Vehovar (2003). Open-ended vs. close-ended questions in web questionnaires. *Developments in Applied Statistics*.

Rietz, T. (1988). The equity risk-premium: A solution. *Journal of Monetary Economics*.

Roll, R. (1977). A critique of the asset-pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics*.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*.

Samuelson, P. (1963). Problems of methodology: Discussion. *American Economic Review*.

Savov, A. (2011). Asset pricing with garbage. *Journal of Finance*.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*.

Singer, E. and M. Couper (2017). Some methodological uses of responses to open questions and other verbatim comments in quantitative surveys. *Methods, Data, Analyses*.

Smith, V., G. Suchanek, and A. Williams (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*.

Sussman, A. and D. Oppenheimer (2020). The effect of effects on effectiveness: A boon-bane asymmetry. *Cognition*.

Ungeheuer, M. and M. Weber (2019). The perception of dependence, investment decisions, and stock prices. *Journal of Finance*.

A Figures

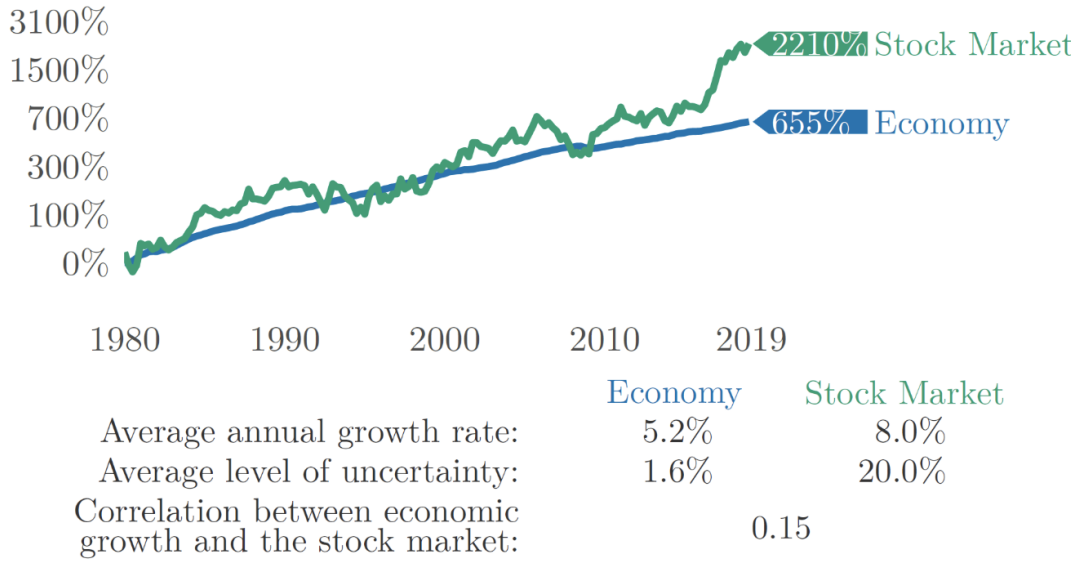

Figure 1. Sample Investment-Decision Question. This figure shows a sample question about investment decisions from the first part of our survey experiment.

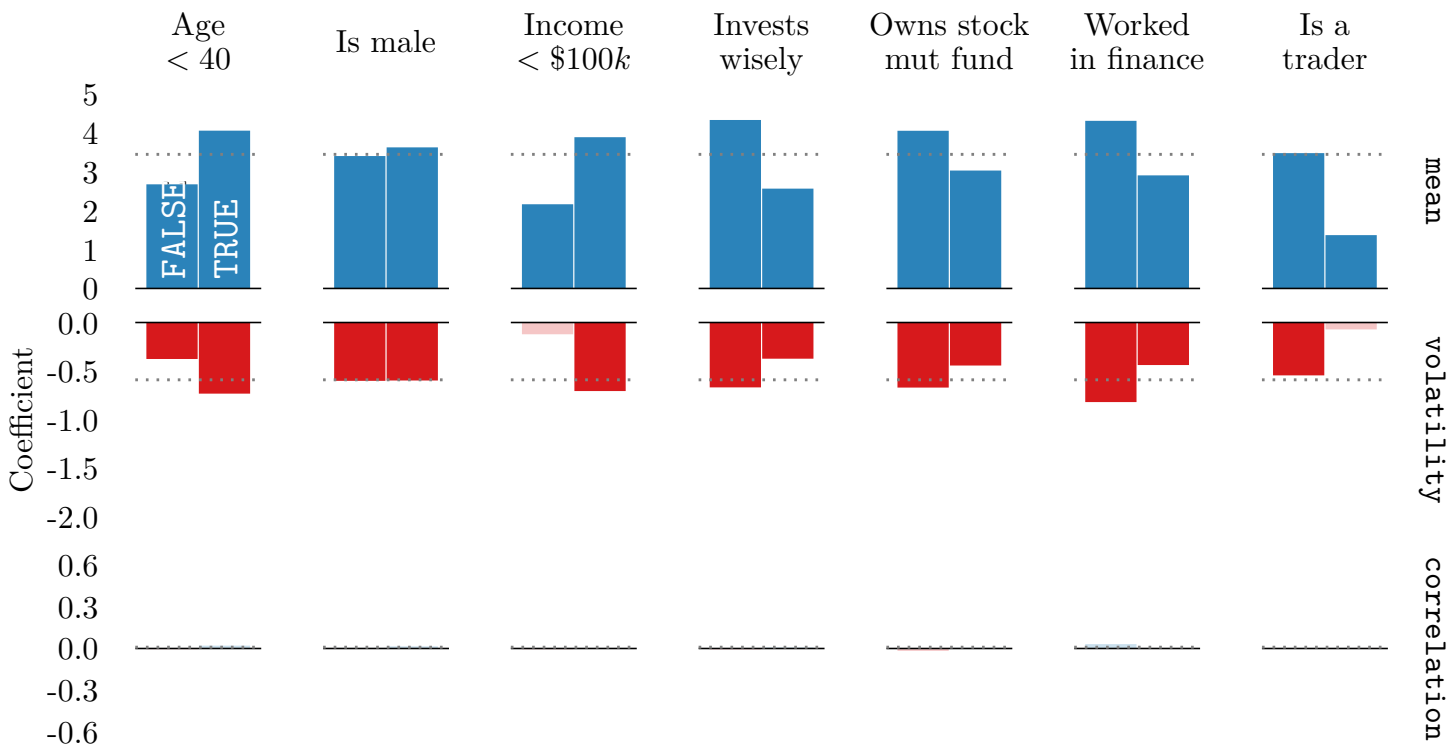

Figure 2. Investment Decisions by Participant Characteristics. This figure reports regression results corresponding to column (4) in Table 2 for different subsets of our participant pools. From top to bottom, each set of three bars represents the slope coefficients from the regression $\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{mean}_{i,q} + \hat{\gamma} \cdot \texttt{volatility}_{i,q} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$. The dependent variable is the fraction invested in stocks, $\texttt{stockFrac}_{i,q}$. The right-hand-side variables correspond to the average stock returns, $\texttt{mean}_{i,q}$, stock-return volatility, $\texttt{volatility}_{i,q}$, and the correlation between stock returns and consumption growth, $\texttt{correlation}_{i,q}$, used to simulate the data for each question. The y-axis is scaled so that a change in the given parameter of low to high would match the scale of the y-axis for the mean graphs. Opaque bars are significant at the 5% level using standard errors clustered by participant. Transparent bars are insignificant. Blue bars denote positive values. Red bars denote negative values. The horizontal dotted gray lines correspond to coefficient values from Table 2 column (4).

B Tables

a) Finance Professionals (N = 493)	#	Avg	Sd	Min	Med	Max
Stock Fraction		0.56	0.24	0.00	0.50	1.00
Age < 40	221	0.45				
Is male	215	0.44				
Income < \$100k	217	0.44				
Owns stock or mutual funds	442	0.90				
Is a trader	136	0.28				
b) MTurkers (N = 322)						
Stock Fraction		0.58	0.27	0.00	0.60	1.00
Age < 40	232	0.72				
Is male	210	0.65				
Income < \$100k	281	0.87				
Own stocks or mutual funds	209	0.65				
Worked in finance	29	0.09				
c) MBA Students (N = 308)						
Stock Fraction		0.67	0.30	0.00	0.70	1.00
Age < 40	304	0.99				
Is male	183	0.59				
Worked in finance	118	0.38				
d) Asset Managers (N = 93)						
Stock Fraction		0.67	0.25	0.00	0.70	1.00

Table 1. Summary Statistics. This table presents summary statistics describing the four participant pools in our survey experiment. Stock Fraction: average fraction of their endowment that each participant invests in stocks; computed using data at the participant×question level. All remaining rows are computed at the participant level. #: represents the number of participants who answered *'Yes'*. Asset managers were not asked background questions prior to taking our survey due to time constraints.

Dependent Variable: $\texttt{stockFrac}_{i,q}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\texttt{mean}_{i,q}$	3.24***			3.22***			
	(14.52)			(14.51)			
$\texttt{volatility}_{i,q}$		−0.61***		−0.60***			
		(10.07)		(9.97)			
$\texttt{correlation}_{i,q}$			0.00	0.00	0.01	0.00	0.01
			(0.00)	(0.16)	(0.69)	(0.06)	(0.68)
Participant FE					✓		✓
Question FE						✓	✓
# Obs	10,062	10,062	10,062	10,062	10,062	10,062	10,062
Adj. R^2	0.04	0.01	0.00	0.05	0.43	0.01	0.44

Table 2. Investment Decisions. This table shows how participants' investment decisions vary with average stock returns, $\texttt{mean}_{i,q}$, stock-return volatility, $\texttt{volatility}_{i,q}$, and the correlation between stock returns and consumption growth, $\texttt{correlation}_{i,q}$. This table uses observations on all participant pools. Each column reports the results of a different regression of the form $\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{mean}_{i,q} + \hat{\gamma} \cdot \texttt{volatility}_{i,q} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$. The dependent variable is the fraction invested in stocks, $\texttt{stockFrac}_{i,q}$. Columns (5) and (7) include participant fixed effects. Columns (6) and (7) include question-order fixed effects. The numbers in parentheses are t-stats computed using standard errors clustered by participant. *, **, and *** indicate statistically significant estimates at the 10%, 5%, and 1% levels.

Dependent Variable: $\texttt{stockFrac}_{i,q}$

	All Participants		Finance Professionals	MTurkers	MBA Students	Asset Managers
	(1)	(2)	(3)	(4)	(5)	(6)
$\texttt{correlation}_{i,q}$	0.01	0.00	0.00	0.00	0.07*	−0.03
	(0.69)	(0.33)	(0.11)	(0.12)	(1.88)	(0.38)
Participant FE	✓		✓	✓	✓	✓
Pool FE		✓				
# Obs	10,062	10,062	4,930	3,220	1,540	372
Adj. R^2	0.43	0.02	0.37	0.43	0.51	0.38

Table 3. Investment Decisions by Participant Pool. This table shows how the investment decisions of different types of participants change in response to an asset's return correlation with consumption growth, $\texttt{correlation}_{i,q}$. Each column reports the results of a different regression of the form $\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$. The dependent variable is the fraction invested in stocks, $\texttt{stockFrac}_{i,q}$. Columns (1) and (2) report results using all participant pools. Column (3) reports results only for finance professionals. Column (4) reports results only for MTurkers. Column (5) reports results only for MBA students. And, column (6) reports results only for asset managers attending the investor conference. Columns (1), (3), (4), (5), and (6) include participant fixed effects. Column (2) includes pool fixed effects. The numbers in parentheses are *t*-stats computed using standard errors clustered by participant. *, **, and *** indicate statistically significant coefficient estimates at the 10%, 5%, and 1% levels.

Dependent Variable: $\texttt{stockFrac}_{i,q}$

	GDP	Industrial Production	Aggregate Consumption	Personal Wealth	Personal Income	House Prices	Personal Consumption	Personal Spending	Standard of Living
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\texttt{mean}_{i,q}$	3.16^{***}	2.71^{***}	3.49^{***}	4.35^{***}	6.34^{***}	2.82^{***}	4.09^{***}	4.34^{***}	3.52^{***}
	(5.89)	(5.13)	(6.78)	(8.16)	(10.15)	(4.92)	(8.99)	(9.83)	(8.51)
$\texttt{volatility}_{i,q}$	−0.09	−0.15	-0.29^{*}	-0.51^{***}	-0.71^{***}	-0.75^{***}	-0.67^{***}	-0.65^{***}	-0.71^{***}
	(0.52)	(0.82)	(1.80)	(3.37)	(4.40)	(4.55)	(4.53)	(4.84)	(5.28)
$\texttt{correlation}_{i,q}$	0.01	0.06	0.04	0.04	−0.02	0.04	0.04	0.00	0.01
	(0.34)	(1.45)	(1.00)	(1.15)	(0.35)	(0.97)	(1.16)	(0.07)	(0.46)
# Obs	1,490	1,700	1,560	1,860	1,110	1,340	2,450	2,290	2,520
Adj. R^2	0.04	0.03	0.05	0.08	0.16	0.05	0.07	0.07	0.06

Table 4. Investment Decisions Using Related Variables. This table shows how participants' investment decisions vary with average stock returns, $\texttt{mean}_{i,q}$, stock-return volatility, $\texttt{volatility}_{i,q}$, and the correlation between stock returns and a risk factor, $\texttt{correlation}_{i,q}$, when that variable is labeled as *'gross domestic product (GDP)'*, *'industrial production'*, *'aggregate consumption'*, *'personal income'*, *'personal wealth'*, *'house prices'*, *'personal consumption'*, *'personal spending'*, and *'material standard of living'*. Each column reports the results of a different regression of the form $\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{mean}_{i,q} + \hat{\gamma} \cdot \texttt{volatility}_{i,q} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$. The dependent variable is the fraction invested in stocks, $\texttt{stockFrac}_{i,q}$. Each column uses data from a separate survey run on a separate population of MTurkers. Numbers in parentheses are *t*-stats computed using standard errors clustered by participant. *, **, and *** indicate statistically significant estimates at the 10%, 5%, and 1% levels.

Dependent Variable: $\texttt{stockFrac}_{i,q}$

	Text Only	$\rho \in [-0.45, 0.45]$	Additional ρ Instructions	Marginal Decision	Stable Predictors	Single Stock	Scatterplot	No Graph
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\texttt{mean}_{i,q}$	4.11***	3.33***	3.52***	3.72***	4.02***	6.25***	1.31***	1.93***
	(6.52)	(7.35)	(6.82)	(7.97)	(8.82)	(11.63)	(2.89)	(4.64)
$\texttt{volatility}_{i,q}$	−0.58***	−0.13	−0.71***	−0.51***	−0.65***	−0.52***	−0.74***	−0.37**
	(3.35)	(1.04)	(4.51)	(3.94)	(4.51)	(4.02)	(3.99)	(2.26)
$\texttt{correlation}_{i,q}$	0.05	0.02	0.01	−0.01	−0.03	0.02	0.15***	0.18***
	(1.29)	(0.92)	(0.18)	(0.29)	(0.89)	(0.69)	(3.85)	(4.61)
# Obs	1,290	2,410	1,800	2,510	2,140	1,960	1,520	1,870
Adj. R^2	0.07	0.04	0.06	0.05	0.07	0.13	0.03	0.03

Table 5. Investment Decisions Under Treatment Variations. This table shows how participants' investment decisions vary when we provided them with information about the correlation between stock returns and consumption growth in a different format. Each column reports the results of a different regression of the form $\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{mean}_{i,q} + \hat{\gamma} \cdot \texttt{volatility}_{i,q} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$ estimated using data from a separate survey run on a different population of MTurkers. Each of these surveys represents a slight variation on our baseline survey. The dependent variable is the fraction invested in stocks, $\texttt{stockFrac}_{i,q}$. In column (1), participants saw $\rho \in \{$*'none'*, *'low'*, *'medium'*, *'high'*$\}$ reported as text rather than $\rho \in \{0.00, 0.15, 0.30, 0.45\}$. In column (2), participants saw questions in which data were simulated using correlations which were sometimes negative. Correlation values were drawn at random from $\rho \in \{-0.45, -0.30, -0.15, 0.00, 0.15, 0.30, 0.45\}$. In column (3), participants received additional instructions on how to interpret a correlation coefficient. In column (4), participants received additional instructions emphasizing that they were to treat the \$1,000 endowment as a marginal investment decision. In column (5), participants received additional instructions emphasizing that the parameter values shown to them were stable over time and would apply to future investments. In column (6), participants answered questions about investing in a single stock rather than a broad value-weighted mutual fund. In column (7), participants saw a scatterplot rather than a time-series plot. In column (8), participants saw the numeric values of the mean, volatility, and correlation but no time-series graphs. Numbers in parentheses are t-stats computed using standard errors clustered by participant. *, **, and *** indicate statistically significant estimates at the 10%, 5%, and 1% levels.

	Mean μ_R (1)	Volatility σ_R (2)	Correlation ρ (3)
Pr[$\texttt{consider}_i$]	0.77 [0.04]	0.59 [0.07]	0.43 [0.03]
Pr[$\texttt{textbookLogic}_i$]	0.58 [0.05]	0.44 [0.08]	0.11 [0.01]
Pr[$\texttt{textbookLogic}_i \mid \texttt{consider}_i$]	0.76^{+++} [0.02]	0.72^{+++} [0.03]	0.24^{---} [0.03]

Table 6. Economic Reasoning. This table depicts the rate at which participants reported thinking about average stock returns (μ_R), stock-return volatility (σ_R), and the correlation between stock returns and consumption growth (ρ) when making their investment decisions. $\texttt{consider}_i$ is an indicator variable for whether the ith participant thought about a parameter at all. $\texttt{textbookLogic}_i$ is an indicator variable for whether the ith participant thought about this parameter using textbook asset-pricing logic. If $\texttt{consider}_i$ = *false*, then $\texttt{textbookLogic}_i$ = *false* as well. This table uses observations on all participant pools. Numbers in square brackets are standard errors clustered by participant pool. In the bottom row, we use $^{+}$, $^{++}$, and $^{+++}$ to indicate probabilities greater than 0.50 with statistical significance at the 10%, 5%, and 1% levels; whereas, we use $^{-}$, $^{--}$, and $^{---}$ to indicate probabilities less than 0.50 at the same significance levels.

Dependent Variable:	consider$_i$			textbookLogic$_i$			textbookLogic$_i$ | consider$_i$		
	μ_R	σ_R	ρ	μ_R	σ_R	ρ	μ_R	σ_R	ρ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
isFinancePro$_i$	0.72	0.52	0.38	0.56	0.36	0.10	0.78^{+++}	0.69^{+++}	0.27^{---}
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.01]	[0.02]	[0.03]	[0.03]
isMTurker$_i$	0.77	0.53	0.48			0.08			0.17^{---}
	[0.02]	[0.03]	[0.03]			[0.02]			[0.04]
isMBAstudent$_i$	0.88	0.77	0.48	0.69	0.58	0.11	0.78^{+++}	0.76^{+++}	0.23^{---}
	[0.02]	[0.02]	[0.03]	[0.03]	[0.03]	[0.02]	[0.03]	[0.03]	[0.04]
isAssetManager$_i$	0.62	0.54	0.35	0.37	0.40	0.17	0.59^{+}	0.74^{+++}	0.48
	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.03]	[0.05]	[0.05]	[0.06]
# Obs	1,216	1,216	1,216	894	894	1,216	683	543	525
Adj. R^2	0.77	0.61	0.44	0.60	0.46	0.11	0.77	0.72	0.26

Table 7. Economic Reasoning by Participant Pool. This table depicts the rate at which different participant pools reported thinking about average stock returns (μ_R), stock-return volatility (σ_R), and the correlation between stock returns and consumption growth (ρ) when making their investment decisions. Each column reports results for a separate regression of the form: $\texttt{y}_i = \hat{\alpha} \cdot \texttt{isFinancePro}_i + \hat{\beta} \cdot \texttt{isMTurker}_i + \hat{\gamma} \cdot \texttt{isMBAstudent}_i + \hat{\delta} \cdot \texttt{isAssetManager}_i + \hat{\varepsilon}_{i,q}$. The dependent variable is an indicator variable capturing whether/how the *i*th participant thought about a parameter. consider$_i$ is an indicator variable for whether the *i*th participant thought about a parameter at all. textbookLogic$_i$ is an indicator variable for whether the *i*th participant thought about this parameter using textbook asset-pricing logic. If consider$_i$ = *false*, then textbookLogic$_i$ = *false* as well. The right-hand-side variables are indicators for which population the *i*th participant belongs to. Numbers in square brackets are standard errors. In columns (7)-(9), we use $^{+}$, $^{++}$, and $^{+++}$ to indicate probabilities greater than 0.50 with statistical significance at the 10%, 5%, and 1% levels; whereas, we use $^{-}$, $^{--}$, and $^{---}$ to indicate probabilities less than 0.50 at the same significance levels.

Dependent Variable: $\texttt{stockFrac}_{i,q}$

	$\texttt{mean}_{i,q}$	$\texttt{volatility}_{i,q}$	$\texttt{correlation}_{i,q}$
	(1)	(2)	(3)
$\texttt{consider}_i = \textit{false}$	2.03***	−0.37***	−0.01
	(4.30)	(4.34)	(0.59)
$\texttt{consider}_i = \textit{true}$	3.62***	−0.77***	0.01
	(14.36)	(9.16)	(0.67)
$(\texttt{consider}_i = \textit{true}) \,\&\, (\texttt{textbookLogic}_i = \textit{false})$	1.18**	0.05	0.03
	(2.22)	(0.26)	(1.12)
$(\texttt{consider}_i = \textit{true}) \,\&\, (\texttt{textbookLogic}_i = \textit{true})$	4.06***	−1.17***	−0.03
	(11.68)	(10.38)	(0.71)

Table 8. Investment Decisions by Economic Reasoning. This table presents regression results showing how participants' investment decisions vary with average stock returns, $\texttt{mean}_{i,q}$, stock-return volatility, $\texttt{volatility}_{i,q}$, and the correlation between stock returns and consumption growth, $\texttt{correlation}_{i,q}$. This table uses data on all participant pools. Each entry in the table represents the estimated slope coefficient, $\hat{\beta}$, of a separate regression of the form $\texttt{stockFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{x}_{i,q} + \hat{\varepsilon}_{i,q}$ where $\texttt{x}_{i,q} \in \{\texttt{mean}_{i,q}, \texttt{volatility}_{i,q}, \texttt{correlation}_{i,q}\}$ using the specified sub-population for a given row. The dependent variable is the fraction invested in stocks, $\texttt{stockFrac}_{i,q}$. $\texttt{consider}_i$ is an indicator variable for whether the *i*th participant thought about a parameter at all. $\texttt{textbookLogic}_i$ is an indicator variable for whether the *i*th participant thought about this parameter using textbook asset-pricing logic. If $\texttt{consider}_i = \textit{false}$, then $\texttt{textbookLogic}_i = \textit{false}$ as well. The first row shows investment-decision results only for those participants who did not report considering a parameter in the second part of our survey. The second row shows analogous results for only those participants who did report considering the parameter. The third row shows results for the subset of participants who reported thinking about the parameter but did so using the opposite of textbook logic. And the fourth row shows results for the participants who thought about a parameter using textbook logic. Numbers in parentheses are *t*-stats computed using standard errors clustered by participant. *, **, and *** indicate statistically significant estimates at the 10%, 5%, and 1% levels.

Dependent Variable:	consider$_i$			textbookLogic$_i$			textbookLogic$_i$|consider$_i$		
	μ_R	σ_R	ρ	μ_R	σ_R	ρ	μ_R	σ_R	ρ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDP	0.68 [0.05]	0.56 [0.04]	0.40 [0.04]	0.60 [0.05]	0.49 [0.04]	0.06 [0.02]	0.88^{+++} [0.03]	0.88^{+++} [0.04]	0.15^{---} [0.04]
Industrial Production	0.62 [0.04]	0.48 [0.04]	0.38 [0.04]	0.54 [0.04]	0.41 [0.03]	0.02 [0.01]	0.87^{+++} [0.03]	0.85^{+++} [0.04]	0.06^{---} [0.03]
Aggregate Consumption	0.72 [0.03]	0.46 [0.04]	0.28 [0.03]	0.60 [0.04]	0.43 [0.04]	0.05 [0.02]	0.82^{+++} [0.04]	0.94^{+++} [0.03]	0.18^{---} [0.04]
Personal Wealth	0.69 [0.04]	0.53 [0.03]	0.37 [0.04]	0.60 [0.04]	0.44 [0.03]	0.04 [0.02]	0.87^{+++} [0.03]	0.83^{+++} [0.04]	0.12^{---} [0.05]
Personal Income	0.65 [0.04]	0.48 [0.05]	0.28 [0.04]	0.64 [0.04]	0.41 [0.05]	0.05 [0.03]	0.99^{+++} [0.01]	0.85^{+++} [0.05]	0.19^{---} [0.09]
House Prices	0.68 [0.05]	0.54 [0.05]	0.31 [0.04]	0.63 [0.04]	0.52 [0.05]	0.06 [0.02]	0.92^{+++} [0.02]	0.96^{+++} [0.02]	0.20^{---} [0.06]
Personal Consumption	0.70 [0.03]	0.56 [0.03]	0.32 [0.03]	0.58 [0.03]	0.49 [0.03]	0.04 [0.01]	0.83^{+++} [0.03]	0.88^{+++} [0.02]	0.13^{---} [0.04]
Personal Spending	0.69 [0.03]	0.51 [0.04]	0.36 [0.03]	0.61 [0.03]	0.46 [0.03]	0.03 [0.01]	0.89^{+++} [0.03]	0.90^{+++} [0.03]	0.10^{---} [0.04]
Standard of Living	0.68 [0.03]	0.55 [0.03]	0.37 [0.03]	0.58 [0.03]	0.48 [0.03]	0.08 [0.02]	0.85^{+++} [0.02]	0.88^{+++} [0.03]	0.23^{---} [0.06]

Table 9. Economic Reasoning Using Related Variables. This table depicts the rate at which participants reported thinking about average stock returns (μ_R), stock-return volatility (σ_R), and the correlation between stock returns and different aggregate risk factors (ρ) when making their investment decisions. Each row uses data from a different survey run on a separate population of MTurkers. The survey is the same as our main treatment except that, instead of *'economic growth'*, the aggregate risk factor shown to participants is labeled as *'gross domestic product (GDP)'*, *'industrial production'*, *'aggregate consumption'*, *'personal wealth'*, *'personal income'*, *'house prices'*, *'personal consumption'*, *'personal spending'*, or *'material standard of living'*. consider$_i$ is an indicator variable for whether the *i*th participant thought about a parameter at all. textbookLogic$_i$ is an indicator variable for whether the *i*th participant thought about this parameter using textbook asset-pricing logic. If consider$_i$ = *false* then textbookLogic$_i$ = *false* as well. Numbers in square brackets are standard errors clustered by participant pool. In columns (7)-(9), row we use $^{+}$, $^{++}$, and $^{+++}$ to indicate probabilities greater than 0.50 with statistical significance at the 10%, 5%, and 1% levels; whereas, we use $^{-}$, $^{--}$, and $^{---}$ to indicate probabilities less than 0.50 at the same significance levels.

	$\texttt{consider}_i$		$\texttt{textbookLogic}_i \mid \texttt{consider}_i$	
	#	Fraction	#	Fraction
	(1)	(2)	(3)	(4)
Personal financial well-being	195	0.68		
Overall wealth	90	0.32		
Uncertainty	59	0.21		
Relationship with stock market	40	0.14	7	0.18^{---} [0.06]
Stock market uncertainty	184	0.64	142	0.77^{+++} [0.03]
Expected stock market returns	160	0.56	136	0.85^{+++} [0.03]
Economic/political climate	58	0.20		
Expected outcome	19	0.07		
Uncertainty	14	0.05		
Relationship with stock market	24	0.08	3	0.13^{---} [0.06]
Expert guidance/advice	56	0.20		
Investment-related costs	53	0.19		
None of the above	3	0.01		

Table 10. Open-Ended Survey. This table reports results from the second part of our open-ended survey. The first part asked 285 MTurkers to list considerations that are important for stock-market investing. The second part then asked each participant whether he was referring to any of 6 broad categories in his responses: *'stock market uncertainty'*, *'expected stock market returns'*, *'personal financial well-being'*, *'economic/political climate'*, *'investment-related costs'*, and *'expert guidance/advice'*. Within the *'personal financial well-being'* category, we also asked participants to further classify their response into three subcategories: *'Do I have enough?'*, *'Are my finances are too uncertain?'*, and *'How will changes in my personal finances correspond to changes in stock market returns?'* Within the *'economic/political climate'* category, we asked participants to further classify their response into three subcategories: *'I expect the future climate to be good/bad.'*, *'I am too uncertain about the future climate.'*, and *'What is the relationship with the stock market?'* Columns (1) and (2) report the number and fraction of all participants who classified at least one of their responses as referring to a given category. Textbook models argue that investors try to buy more stocks when stock returns are higher in bad times, less volatile, and higher on average. Columns (3) and (4) give the number and fraction of participants who reported following textbook logic in follow-up questions conditional on referring to a category in their response. Numbers in square brackets are standard errors. We use $^{+}$, $^{++}$, and $^{+++}$ to indicate probabilities greater than 0.50 with statistical significance at the 10%, 5%, and 1% levels; whereas, we use $^{-}$, $^{--}$, and $^{---}$ to indicate probabilities less than 0.50 at the same significance levels.

		TNA ($*bil*)		Mentions of	Mentions related to other macro variables	
		Share Class (1)	Fund (2)	correlat(ion|e) covar(iance|y) (3)	§Investment Risks (4)	§Investment Objectives (5)
Vanguard 500 Index, Adm	VFIAX	276	483	0	0	0
Vanguard Total Stock Market Index, Adm	VTSAX	225	814	0	0	0
Fidelity 500 Index	FXAIX	198	198	4	0	0
Vanguard Total Stock Market Index, Instl Pl	VSMPX	170	814	0	0	0
Vanguard Total International Stock Index, Inv	VGTSX	146	382	0	0	0
Vanguard Total Stock Market Index, I	VITSX	140	814	0	0	0
Vanguard Total Stock Market Index, Inv	VTSMX	139	814	0	0	0
Vanguard Institutional Index, Instl Pl	VIIIX	115	229	0	0	0
Vanguard Institutional Index, I	VINIX	114	229	0	0	0
Vanguard Total International Stock Index, Instl Pl	VTPSX	112	382	0	0	0
Vanguard Total Bond Market II Index, Inv	VTBIX	107	182	0	0	0
Vanguard Total Bond Market Index, Adm	VBTLX	100	229	0	0	0
Fidelity Contrafund	FCNTX	95	122	0	0	0
Amer Funds Gr Fund of America, A	AGTHX	91	196	0	0	0
Vanguard Wellington, Adm	VWENX	89	105	0	0	0
Vanguard Total Bond Market II Index, I	VTBNX	75	182	0	0	0
Vanguard Total International Stock Index, Adm	VTIAX	75	382	0	0	0
Amer Funds Income Fund of America, A	AMECX	74	111	0	0	0
Amer Funds American Balanced, A	ABALX	72	150	0	0	0
Amer Funds Europa Gr, R6	RERGX	71	162	0	0	0
Dodge & Cox Stock	DODGX	71	71	1	0	0
Vanguard 500 Index, Instl Select	VFFSX	70	483	0	0	0
PIMCO Income Institutional	PIMIX	67	128	3	0	0
Amer Funds Cap Income Builder, A	CAIBX	65	105	0	0	0
Vanguard Intermediate-Term Tax-Exempt, Adm	VWIUX	65	68	0	0	0
		2821	3734	8	0	0

Table 11. Mutual-Fund Prospectuses. This table describes how the 25 largest US mutual funds talk about risk-factor correlations in their prospectuses. Columns (1) and (2) report total net assets by share class and at the fund level. Column (3) reports the number of times *'correlation'*, *'correlate'*, *'covariance'*, or *'covary'* appear in a fund's prospectus. Columns (4) and (5) report the number of times a fund mentions its exposure to a macroeconomic variable in the Investment Risks or Investment Objectives sections of its prospectus.

Dependent Variable: $\texttt{mutualFundFrac}_{i,q}$

	Market	Size	Value
	(1)	(2)	(3)
$\texttt{mean}_{i,q}$	3.65***	3.35***	4.20***
	(7.81)	(8.84)	(10.80)
$\texttt{volatility}_{i,q}$	−0.51***	−0.40***	−0.74***
	(3.43)	(2.85)	(5.36)
$\texttt{correlation}_{i,q}$	0.04	0.11***	0.05
	(1.23)	(3.41)	(1.76)
# Obs	2,350	2,150	2,230
Adj. R^2	0.05	0.06	0.09

Table 12. Investment Decisions Using Fama-French Risk Factors. This table shows how participants' investment decisions vary with average mutual-fund returns, $\texttt{mean}_{i,q}$, the volatility of mutual-fund returns, $\texttt{volatility}_{i,q}$, and the correlation between the mutual fund's returns and each of the three Fama-French risk factors, $\texttt{correlation}_{i,q}$. Market corresponds to the excess return on the market portfolio. Size corresponds the small-minus-large size factor. Value corresponds to the high-minus-low value factor. Each column reports the results of a different regression of the form $\texttt{mutualFundFrac}_{i,q} = \hat{\alpha} + \hat{\beta} \cdot \texttt{mean}_{i,q} + \hat{\gamma} \cdot \texttt{volatility}_{i,q} + \hat{\delta} \cdot \texttt{correlation}_{i,q} + \hat{\varepsilon}_{i,q}$. The dependent variable is the fraction invested in the mutual fund, $\texttt{mutualFundFrac}_{i,q}$. Each column uses data from a separate survey run on a separate population of MTurkers. Numbers in parentheses are t-stats computed using standard errors clustered by participant. *, **, and *** indicate statistically significant estimates at the 10%, 5%, and 1% levels.

Dependent Variable:	consider$_i$			textbookLogic$_i$			textbookLogic$_i$ | consider$_i$		
	μ_R	σ_R	ρ	μ_R	σ_R	ρ	μ_R	σ_R	ρ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Market	0.71	0.51	0.31	0.59	0.45	0.05	0.83^{+++}	0.88^{+++}	0.15^{---}
	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]	[0.01]	[0.03]	[0.03]	[0.04]
Size	0.72	0.44	0.24	0.65	0.37	0.04	0.90^{+++}	0.84^{+++}	0.17^{---}
	[0.02]	[0.04]	[0.03]	[0.03]	[0.03]	[0.01]	[0.02]	[0.04]	[0.05]
Value	0.65	0.48	0.35	0.57	0.45	0.04	0.88^{+++}	0.94^{+++}	0.13^{---}
	[0.03]	[0.03]	[0.03]	[0.03]	[0.04]	[0.01]	[0.03]	[0.02]	[0.03]

Table 13. Economic Reasoning About Fama-French Risk Factors. This table depicts the rate at which participants reported thinking about average mutual-fund returns (μ_R), the volatility of mutual-fund returns (σ_R), and the correlation between the mutual fund's returns and each of the three Fama-French risk factors (ρ) when making their investment decisions. Each row uses data from a separate survey run on a separate population of MTurkers. Market corresponds to the excess return on the market portfolio. Size corresponds the small-minus-large size factor. Value corresponds to the high-minus-low value factor. consider$_i$ is an indicator variable for whether the *i*th participant thought about a parameter at all. textbookLogic$_i$ is an indicator variable for whether the *i*th participant thought about this parameter using textbook asset-pricing logic. If consider$_i$ = *false* then textbookLogic$_i$ = *false* as well. Numbers in square brackets are standard errors clustered by participant pool. In columns (7)-(9), we use $^{+}$, $^{++}$, and $^{+++}$ to indicate probabilities greater than 0.50 with statistical significance at the 10%, 5%, and 1% levels; whereas, we use $^{-}$, $^{--}$, and $^{---}$ to indicate probabilities significantly less than 0.50.