Disagreement and the Cross-Section of Expected Returns

Namhee Matheson *

This version : November 13, 2017

*PhD candidate at Department of Finance BI Norwegian Business School : Nydalsveien 37 0484 Oslo phone: +47 46 41 07 64 e-mail: namhee.matheson@bi.no I thank Ilan Cooper, Paul Ehling, Bruno Gerard, Christian Heyerdahl-Larsen, Patrick Konermann, Paul Whelan (discussant), Costas Xiouros (advisor), and Fernando Zapatero (advisor) for helpful comments and suggestions. All remaining errors are my own.

ABSTRACT

This paper studies the role of heterogeneous beliefs about a non-fundamental process, which is independent of the aggregate consumption risk, in explaining the cross-section of expected returns. In the model, trading on the non-fundamental process results in redistribution of wealth and consumption along the path of the non-fundamental process. In equilibrium, a consumption-weighted sentiment determines the price of risk of the non-fundamental process, and this measure serves as a conditioning variable for tests of the cross-section. To test the model-based predictions, I develop a novel approach that exploits the close link between the cross-sectional heterogeneity in consumption and heterogeneous beliefs to identify such non-fundamental process. I find that the factor model has a significant ability to explain the cross-section of expected returns, and pessimism is associated with a high market price of risk.

Keywords: Disagreement, heterogeneous beliefs, consumption asset pricing, cross-section of equity returns

JEL classification: G10, G12, G13, E21

I. Introduction

Under the CAPM and CCAPM, the aggregate measures of fundamentals, the return on the wealth portfolio and the growth rate of aggregate consumption, respectively, are sufficient measure of the risks faced by investors. The long list of literature reports that this statement is inconsistent with cross-sectional asset pricing data, as highlighted by Harvey, Liu, and Zhu (2016), which documents 316 anomalies, discovered by academia and proposed as potential factors, and 40 risk factors, even only accounting for the factors motivated by theory.

This paper explores a general mechanism that generates such an abundance of factors and proposes a standard consumption asset pricing model with one extension, in which individual investors hold heterogeneous beliefs about a non-fundamental, a process unrelated to the aggregate consumption growth. The model has a tractable and interpretable implication for asset prices; a consumption-weighted sentiment, a measure that summarizes market expectations for the price of risk on the non-fundamental, serves as a conditioning variable for tests of the cross-section. I show, theoretically and empirically, that the conditional factor model implied by the model has a significant ability to explain the cross-sectional dispersion in expected returns, and fluctuations in aggregate pessimism are associated with a high market price of risk.

To illustrate the role of heterogeneous beliefs in explaining the cross-section of expected returns, I consider a general equilibrium model, in which investors disagree about dynamics of a non-fundamental variable. Markets are complete, which facilitates the investors to make their portfolio decisions based on their beliefs by trading on the non-fundamental process. As a motivating example, consider a new industry, which does not currently produce tangible goods but may have a substantial growth opportunity in the future. Because there is no known history, investors may find it difficult to accurately estimate its perspective, which naturally leads to disagreements among the investors. When they receive a noisy signal and are overconfident, investors would make their portfolio decision based on this signal.

In essence, trading among the individual investors results in the redistribution of wealth and consumption along the path of the non-fundamental process. I show that fluctuations in the cross-sectional consumption distribution introduce the non-fundamental as a new risk factor and induce a time-variation in its risk price. That is, the financial market may fluctuate even when there is no news about the fundamental. Precisely, the market price of risk on the non-fundamental is determined by a consumption-weighted sentiment, a tractable and intuitive measure that aggregates individual sentiments taking equilibrium effects into account. Aggregate pessimism, regarding this measure, predicts a higher market price of risk.

I define a consumption-weighted sentiment as a consumption-weighted average of individual sentiments, where an individual sentiment quantifies a deviation of the growth rates under the subjective belief from that of the actual process. The intuition for this measure is two folds. First, the consumption share is an adequate weight to summarize market expectations in equilibrium. This is because a wealthier individual can hold more significant quantities of assets and thus have more influence on the equilibrium price. As such the consumption share can be considered as an effective weight that adjusts for equilibrium demand. Second, in a sense that a consumption-weighted sentiment is a single valid measure on how optimistic/pessimistic the economy is, its relationship to risk premium can be understood by analogy with homogeneous beliefs economy, where investors collectively hold the same belief. Again in the new industry example, an investor, skeptical about the perspective of this industry, would require a higher risk premium to hold any asset of firms exposed to this industry. The finding that the aggregate pessimism in the heterogeneous economy is associated with a higher market price of risk can be understood by following this intuition.

Given the model predictions, it naturally follows that a consumption-weighted sentiment, due to the ICAPM (Merton, 1973)[1], serves as a conditioning variable for tests of the cross-section. That is, the conditional two-factor model: the CCAPM augmented with a non-fundamental factor, describes the expected excess returns. To help to organize the empirical study, I express the conditional factor model as an unconditional model, which is achieved by introducing a scaled factor, the non-fundamental factor scaled by the conditional variable in addition to two factors. As such the unconditional model implies that the cross-sectional dispersion in expected returns is determined not only by its correlation with aggregate consumption growth and the non-fundamental factor but also by its correlation with the scaled factor.

[1]See also Cochrane (1996); Ferson and Harvey (1999); Lettau and Ludvigson (2001); Santos and Veronesi (2006).

The empirical challenge to test the model-based predictions is that the non-fundamental process and the individual sentiment are not observable. To address this issue, I develop a novel approach that exploits the close link between the cross-sectional heterogeneity in consumption and heterogeneous beliefs, the underlying source of this fluctuation according to my model. The main idea is that the cross-sectional consumption data, which is observable, carry useful information about the latent variables: the non-fundamental process and the individual sentiment.

The procedure is summarized as follows. First, the approach uses household-level consumption data provided by Consumer Expenditure Survey (CEX) and constructs a synthetic balanced panel of ten groups based on their consumption experience. This procedure is designed to overcome the limitation of CEX data that each household participates only four consecutive quarters. Second, I estimate the latent variables by principal component analysis on the synthetic panel. In particular, I interpret the first principal component as the empirical proxy for the non-fundamental process, while the factor loading obtained from a regression of group-level consumption growth rate on this factor as the individual sentiment of each group.

The principal component measures the co-movement in the panel, and thus, naturally serves the purpose of capturing the underlying factor that drives the cross-section of consumption. For the intuition for the factor loadings, consider two optimistic investors, of whom the first is more optimistic than the second, but who are identical otherwise. Both investors optimally invest in assets, which are positively correlated with the non-fundamental process as doing so this portfolio decision provides a favorable profit opportunity from their perspective. By the same reason, the first investor prefers more exposure than the second. As a result, the wealth (and consumption) of both investors rises when the positive state of the non-fundamental process is realized, but the first investor's consumption is more sensitive than the second. That is optimistic beliefs induce a positive correlation between the consumption growth rates and the factor, and thus a positive factor loading in the regression while the factor loading for the first investor is expected to be greater.

I test the model-based prediction in Fama and French 25 portfolios as test assets and find that the non-fundamental factor scaled by the consumption-weighted sentiment considerably explains the cross-sectional spread in average returns. The results suggest that an asset's risk is substantially determined by its correlation with the non-fundamental factor conditional on the sentiment of the economy, even after controlling for Fama and French 3 factors. In particular, I show that a stock earns higher average returns because they are more highly correlated with the non-fundamental factor in bad times; when the economy becomes pessimistic, and risk premium is high. For example, the expected returns of firms, which have significant exposure to the new industry, would be high when the market is pessimistic about this industry. The result gives support to the view that the cross-sectional dispersion in expected returns is attributed to covariance with common non-diversifiable risk factor instead of firm characteristics. I show that the results are robust to alternative numbers of groups and test assets.

The structure of the paper is as follows. Section II discuss the related literature. Section III presents the model. Section IV describes the data and the required empirical settings. Section V presents the identification strategy and Section VI the main empirical results. Conclusion follows in Section VII. The Appendix provides all proofs and empirical specification.

II. Related Literature

This paper is related to the vast literature on equilibrium effects of heterogeneous beliefs. Earlier theoretical works include Harris and Raviv (1993), Detemple and Murthy (1994), Zapatero (1998), and Basak (2000). The most closely related papers are Basak (2000), who studies the equilibrium model that agents' heterogeneous beliefs about non-fundamental processes introduce non-fundamental risk and Jouini and Napp (2007), who show that the asset pricing quantities are determined by consumption share weighted forecasting errors about fundamental risk [2]. My model complements these papers by explicitly showing that disagreement about a non-fundamental risk leads to a conditional two-factor model and this conditional model can be tested by an unconditional factor model augmented with a scaled factor. David (2008) and Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2016) discuss the relationship between the cross-sectional consumption distribution and investor disagreement. My empirical strategy differs from these papers

[2] Detemple and Murthy (1994) are the first to demonstrate that equilibrium prices have a wealth-weighted average structure, and subsequent papers implicitly show this characteristic.

by extracting the factor itself from the cross-sectional consumption data.

The majority of the empirical papers on heterogeneous beliefs and the cross-section of returns focus on a link between empirical proxies of disagreement and returns. Diether, Malloy, and Scherbina (2002) show that stocks with higher dispersion of analysts' earnings forecasts generate lower future returns and suggest that the evidence is consistent with the hypotheses that prices will reflect the optimistic views when there are short-selling constraint e.g. Miller (1977); Harrison and Kreps (1978); Scheinkman and Xiong (2003); Gallmeyer and Hollifield (2008). Li (2016) finds that conditional on high macro disagreement states, high macro beta stocks earn lower future returns relative to low macro beta stocks. Gao, Lu, Song, and Yan (2017) find a positive relationship between the expected returns and their macro-disagreement betas measured based on the Blue Chip Economic Indicators (BCEI) survey in a variety of U.S. asset markets. This paper differs from these studies in two aspects. First, the empirical specification in this paper use only the consumption data. Second, and more important, this paper focuses on the underlying risk factor investors disagree about, and the empirical test does not impose the type of factors.

Another important line of literature (e.g., Brav, Constantinides, and Geczy (2002); Constantinides and Ghosh (2017)) also studies the link between the cross-sectional heterogeneity in consumption and the cross-section of expected returns, but in incomplete market economies. The idea is that individuals face exogenously given idiosyncratic risk such as labor income shocks and are unable to insure against them, and this affects the way they value financial assets. The economic channel I study in this paper is the opposite: the complete capital markets and disagreement induces investors to trade more and thus endogenously creates cross-sectional heterogeneity in consumption. This paper is also related to the previous literature that uses micro household level data from CEX instead of aggregate consumption data to study asset pricing implication. Malloy, Moskowitz, and Vissing-Jørgensen (2009) use CEX data to identify the stockholder and show that the aggregate consumption of the stockholder better explains the asset pricing. Xiouros and Zapatero (2010) infers the heterogeneity of risk aversion from CEX data.

III. The Model

I consider a pure exchange endowment economy (Lucas, 1978), in which investors hold heterogeneous expectations about a non-fundamental variable (Basak, 2000). Because of the disagreement, investors speculate with each other in capital markets, which drives cross-sectional heterogeneity in consumption plans. As a result, the non-fundamental risk is priced, and this generates a close link between cross-sectional heterogeneity in consumption and the stochastic discount factor.

A. Economic Environment

The uncertainty is represented by a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ on which is defined a two-dimensional Brownian motion $B_t = (B_{c,t}, B_{z,t})'$ (where $'$ denotes transpose). There is a single consumption good in exogenous supply C_t, whose dynamics is

$$\frac{dC_t}{C_t} = g_c dt + \sigma_c dB_{c,t} \quad \text{with } C_0 = 1 \tag{1}$$

where the mean growth rate g_c and the volatility parameter $\sigma_c > 0$ are constants. In addition to the aggregate endowment process, investors observe a non-fundamental process Z_t, which I define as a variable independent of dC_t. The evolution Z_t follows

$$\frac{dZ_t}{Z_t} = g_z dt + \sigma_z dB_{z,t} \quad \text{with } Z_0 = 1 \tag{2}$$

where g_z is an expected rate of change, $\sigma_z > 0$ is a constant volatility parameter, and $B_{z,t}$ is independent of $B_{c,t}$. There is a dynamically complete competitive market of contingent claims, such that claims on each realization of B are traded, i.e., Arrow-Debreu securities exist for each state of the world. The results do not depend on the type of securities used to implement the complete market. However, I assume that there are three types of security, of which the first represents a claim to the consumption stream with a net supply C_t, the second is zero net supply security, and the last is a risk-free bond available in zero net supply.

B. Investors

There are K types of price-taking investors, indexed by $i = 1, \cdots, K$. Investors hold heterogeneous expectations about the non-fundamental variable. In particular, investors of type i perceives the drift term to be $g_{i,z} = g_z + \sigma_z \delta_i$ instead of the true mean g_z, where δ_i represents a normalized measure of the forecasting error made by type i. Intuitively, the absolute value of δ_i quantifies degree of mistake and the sign of δ_i indicates sentiment of investor, e.g., a positive value associated with the optimistic investors. For this reason, I refer this variable as the individual sentiment. Throughout the analysis, I treat investors' beliefs as exogenously given [3]. In the following, I refer to the subjective belief of investor i in terms of probability measure $\mathbb{P}_i$ in contrast to the objective measure $\mathbb{P}$. The difference in beliefs can be characterized by the Radon-Nikodym derivative (the likelihood ratio) $\eta_{i,t} \equiv E_t\left[\frac{d\mathbb{P}_i}{d\mathbb{P}}\right]$.

As a motivating example, consider a new industry, which does not currently produce tangible goods but may have a substantial growth opportunity in the future. Because there is no known history, investors may find it difficult to accurately estimate its perspective, which naturally leads to disagreements among the investors. While they receive common information from a source such as news articles, investors may differ in the way they interpret the information as in (Harris and Raviv, 1993), e.g., based on their background knowledge and experience.

The examples also include macro variables such as inflation and unemployment rates. In reality, they may be partially correlated with the aggregate consumption growth. A partially correlated process can be easily accommodated by introducing a Brownian motion correlated with $dB_{c,t}$ which can be decomposed into two independent Brownian motions. For this case, $dB_{z,t}$ is interpreted as an orthogonal component. Lastly, the non-fundamental in the model may not directly reflect one particular variable but can be interpreted as a latent component that drives the common component of a collection of variables investors disagree about.

Investors $i = 1, \cdots, K$ have initial wealth W_i and have constant relative risk aversion (CRRA)

[3]The different subjective beliefs may come from an optimal learning problem of investors with different priors as in for example, Harris and Raviv (1993), Detemple and Murthy (1994), and Basak (2000). In addition, investors may interpret the informativeness of signal differently as in for example, Scheinkman and Xiong (2003), Buraschi and Jiltsov (2006),Xiong (2013), Dumas, Kurshev, and Uppal (2009), and Buraschi, Trojani, and Vedolin (2014).

utility over her respective consumption $C_{i,t}$. Investor i maximizes the expected life time utility under her own probability expectation

$$\max_{C_i} E_0^i \left[\int_0^\infty u(C_{i,t})\, dt \right] \quad \text{with } u(C_{i,t}) \equiv e^{-\rho t} \frac{C_{i,t}^{1-\gamma}}{1-\gamma} \tag{3}$$

subject to the budget constraint

$$E_0^i \left[\int_{t=0}^\infty \xi_{i,t} C_{i,t}\, dt \right] \leq W_i. \tag{4}$$

where E^i and $\xi_{i,t}$ denote the expectation and the state density process under investor i's belief. That is, investors make their portfolio decision based on their own belief, regardless of opinions of other investors.

C. Equilibrium

Observing the realization of Z_t, the perceived shocks to dZ_t/Z_t by the type i investors is

$$dB_{z,t}^i = \frac{1}{\sigma_z} \left(\frac{dZ_t}{Z_t} - g_{i,z} \right). \tag{5}$$

Because the individual beliefs should be consistent with the observed value, the above equation and Equation (2) imply that

$$dB_{z,t}^i = -\delta_i dt + dB_{z,t}. \tag{6}$$

By Girsanov theorem, $B_{z,t}^i$ is a standard Brownian motion under $\mathbb{P}_i$ and the Radon-Nikodym derivative $\eta_{i,t}$ is defined as

$$\eta_{i,t} \equiv e^{\delta_i B_{z,t} - \frac{1}{2}\delta_i^2 t}. \tag{7}$$

The equilibrium allocation in the economy with heterogeneous beliefs is characterized as the solution of the optimization problem of a social planner (see Constantinides (1982); Dumas (1989); Detemple and Murthy (1994); Wang (1996); Basak (2000); Gallmeyer and Hollifield (2008)), whose utility function is a weighted sum of the utilities of individual agents in the economy: that is, the social planner solves:

$$U(C_t;\boldsymbol{\lambda_t}) := \max_{\{C_{i,t}\}_{i=1}^K} \sum_{i=1}^K (\lambda_{i,t} u(C_{i,t})) \tag{8}$$

subject to the market clearing condition

$$\sum_{i=1}^K C_{i,t} = C_t \tag{9}$$

where $\boldsymbol{\lambda_t} \equiv (\lambda_{1,t}, \cdots, \lambda_{K,t})$, $\lambda_{i,t} > 0$ is the utility weight for investor i. Without loss of generality, I normalize $\boldsymbol{\lambda_t}$ by setting $\lambda_{1,t} = 1$. The following proposition shows the equilibrium consumption allocation $\{C_{i,t}\}_{i=1}^K$ and the state price densities $\{\xi_{i,t}\}_{i=1}^K$.

PROPOSITION 1: *In equilibrium, the stochastic weighing process is* $\lambda_{i,t} = \zeta_i \eta_{i,t}$*, where* $\eta_{i,t}$ *given in Equation* (7) *and* ζ_i *is determined by the budget constraint in Equation* (4)*.*

For $i = 1, \cdots, K$*, the equilibrium consumption allocation* $C_{i,t} = \omega_{i,t} C_t$ *is characterized by the consumption share process* $\omega_{i,t}$

$$\omega_{i,t} = \frac{\lambda_{i,t}^{1/\gamma}}{\sum_{i=1}^K \lambda_{i,t}^{1/\gamma}} \tag{10}$$

with $\lambda_{1,t} = 1$*.*

The marginal utility of the social planner is given by

$$U(C_t;\boldsymbol{\lambda_t})' = e^{-\rho t} C_t^{-\gamma} \left(\sum_{i=1}^K \lambda_{i,t}^{1/\gamma} \right)^{\gamma}. \tag{11}$$

Keeping the initial wealth distribution constant, Radon-Nikodym derivative (or the likelihood ratio) $\eta_{i,t}$ determines the impact of heterogeneous beliefs on the consumption allocations and the state prices. Equation (10) implies that the innovation of $\eta_{i,t}$ directly affects individual consumption, thus driving cross-sectional heterogeneity in consumption plans. In particular, if an investor believes a certain state is more likely than others, then it is the state of nature, in which her consumption share of an investor increases. The intuition behind this result is that investors choose consumption plans such a way to consume more in a more likely state from her perspective at the cost of consuming less in the less likely state. Because the social planner's utility depends on $\lambda_{i,t}$ and the innovation of $\lambda_{i,t}$ is determined by the likelihood ratio $\eta_{i,t}$, the state price density is thus directly determined by the distribution of $\eta_{i,t}$.

For brevity, I introduce two notations that refer to recurring terms in what follows: 1) the consumption-weighted sentiment is defined as

$$\mathcal{E}_t(\delta_i) \equiv \sum_{i=1}^{K} (\delta_i \omega_{i,t}) \tag{12}$$

and 2) the consumption-weighted variance of forecasting errors is defined as

$$V_t(\delta_i) \equiv \left[\sum_{i=1}^{K} (\delta_i^2 \omega_{i,t}) - \left(\sum_{i=1}^{K} (\delta_i \omega_{i,t}) \right)^2 \right]. \tag{13}$$

Applying Itô 's lemma on Equation (11) yields the following proposition that describes the dynamics of state price density.

PROPOSITION 2: *Denote $U(C_t; \boldsymbol{\lambda_t})'$ with ξ_t. Under the objective probability measure $\mathbb{P}$, the state price density is given as*

$$\frac{d\xi_t}{\xi_t} = -r_{f,t}dt - \theta_c dB_{c,t} - \theta_{z,t} dB_{z,t} \tag{14}$$

where

$$r_{f,t} = \rho + \gamma g_c - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2 + \frac{\gamma-1}{2\gamma} V_t(\delta_i) \tag{15}$$

$$\theta_c = \gamma \sigma_c \tag{16}$$

$$\theta_{z,t} = -\mathcal{E}_t(\delta_i). \tag{17}$$

The essence of Proposition 2 is that disagreement about Z introduces an additional risk factor beyond the aggregate consumption risk even when Z presents no news about the fundamental. This is in contrast to the case of disagreement about the aggregate consumption process, which does not add a factor, but alters the market price of the aggregate consumption risk. Importantly, the market price of risk on the non-fundamental is determined by a consumption-weighted sentiment $\mathcal{E}_t(\delta_i)$, which induces time-variation in the risk price. In particular, the market price of risk is positive if and only if this measure is negative; that is when the market is pessimistic regarding this measure. The fundamental reason behind these result is that disagreement motivates investors to trade, which in turn lead to the redistribution of wealth and consumption along the path of the non-fundamental process as highlighted in Proposition 1. As a consequence, fluctuations in the cross-sectional consumption distribution introduce the non-fundamental as a new risk factor and induce a time-variation in its risk price.

The intuition for $\mathcal{E}_t(\delta_i)$ is two folds. In the end, I argue that a consumption-weighted sentiment is an intuitive measure that aggregates individual sentiments taking equilibrium effects into account, and thus summarizes market expectations for the price of risk on the non-fundamental. First, the consumption share is an adequate weight to summarize market expectations in equilibrium. This is because a wealthier individual can hold more significant quantities of assets and thus have more influence on the equilibrium price. As such the consumption share can be considered as an effective weight that adjusts for equilibrium demand. Second, a consumption-weighted sentiment is a single valid measure on how optimistic/pessimistic the economy is. Thus, its relationship to risk premium can be understood by analogy with homogeneous beliefs economy, where investors collectively hold the same belief (the intuition suggested by Cvitanić and Malamud (2011)).

Consider a homogeneous economy populated only by pessimistic investors and denote their collective sentiment by $\delta < 0$. The general structure of the model allows that one can obtain the equilibrium quantities in this economy by substituting δ for all investors in Proposition 2. The resulting market price of risk on the non-fundamental risk is $\theta_{z,t} = -\delta$. This observation suggests that pessimistic belief is associated with a positive price of risk. In the presence of heterogeneous beliefs, the outcome can be read as an extension of the homogeneous case by putting the consumption-weighted mean of δ_i in place of δ. For instance, if a negative shock is realized, the wealth (and consumption share) of pessimistic investors and these investors relatively dominate the economy; that is, as an economy becomes more pessimistic, the price of risk increases.

The heterogeneous beliefs also have an implication on the risk-free rate. On the top of the standard result in CRRA utility, the last term in Equation (15) presents an additional variation of risk-free rate introduced by the heterogeneous belief, which is proportional to $V_t(\delta_i)$. Because a variance is by definition non-negative, the sign of $V_t(\delta_i)$ is unambiguously determined only by the magnitude of the risk aversion coefficient γ. In particular, for lower (higher) risk aversion $\gamma < 1$ ($\gamma > 1$), the risk-free rate monotonically decreases (increase) in $V_t(\delta_i)$. In the case for log utility ($\gamma = 1$), heterogeneous beliefs present no effect on risk-free rate. These results are consistent with the previous literature e.g., Varian (1985); Jouini and Napp (2007); Buraschi and Whelan (2012).

Given the equilibrium risk-free rate and the price of risk, Equation (1) and (10) imply that log consumption dynamics of investor i follows

$$d \log C_{i,t} = \frac{1}{\gamma}\left(-\frac{1}{2}\delta_i^2 + r_{f,t} - \rho - \frac{1}{2}\gamma^2\sigma_c^2 + \frac{1}{2}\theta_{z,t}^2\right) dt + \sigma_c dB_{c,t} + \frac{1}{\gamma}\left(\delta_i - \mathcal{E}_t(\delta_i)\right) dB_{z,t}. \qquad (18)$$

To highlight the role of disagreement in explaining the cross-sectional heterogeneity in consumption plans, I first compare with the economy without disagreement. Again, consider a homogeneous belief economy, where investors agree with the perspective of the non-fundamental with their collective sentiment denoted by δ. By substituting δ_i with δ for all investors in Proposition 2, it is straightforward to show that the growth rate of individuals is the same as that of aggregate : $d \log C_{i,t} = d \log C_t$. That is, investors perfectly share the consumption risk.

In contrast, heterogeneous beliefs introduce an idiosyncratic component in consumption growth rate. Indeed, this observation is the main idea behind the empirical strategy described in detail in Section V. Importantly, the individual sentiment δ_i has a direct implication for the sensitivity of the individual consumption growth rate to the shock to Z. First, the consumption of relatively optimistic agent would positively correlate to the shock to Z, while pessimistic agent expects her consumption share to be negatively correlated to the shock to Z. Second, the sensitivity increases with the absolute value of δ_i; that is investors with a more extreme view would experience greater fluctuation in their consumption growth.

Keeping other things constant, the expected consumption growth rates are inverted U-shaped in the individual sentiment δ_i and maximum at $\delta_i = 0$ (correct belief). Intuitively, the agent who has more precise forecast expects the larger consumption share over time, and this term is thus closely related to survival of agent in the long run. In Yan (2008), $\frac{1}{2}\delta_i^2$ corresponds to the survival index of which the lower value implies the long-run survival.

D. Implication for the cross-section of expected returns

This subsection shows that the expected excess returns, according to the model, are approximated as a conditional two-factor model: the first factor is the aggregate consumption growth as in the CCAPM and the second represents the non-fundamental risk, whose market price of risk is determined by the consumption-weighted sentiment. This result naturally follows that a consumption-weighted sentiment, due to the ICAPM (Merton, 1973)[4], serves as a conditioning variable for tests of the cross-section.

To help to organize the empirical study, I express the conditional factor model as an unconditional model, which is achieved by introducing a scaled factor, the non-fundamental factor scaled by the conditional variable in addition to two factors. As such the unconditional model implies that the cross-sectional dispersion in expected returns is determined not only by its correlation with aggregate consumption growth and the non-fundamental factor but also by its correlation with the scaled factor. In what follows, I approximate the model for discrete time with one quarter as one unit of time to match the frequency of CEX data. Throughout the paper lower case letters are used to denote log variables, e.g., log consumption is $c_t \equiv \log C_t$. Appendix A shows the detail of derivation.

D.1. Conditional beta representation

Let $\mathfrak{R}_t^j$ denote the instantaneous return of asset $j = 1, \cdots, N$. Given the state price density in Equation (14), the expected excess returns of asset j in equilibrium should satisfy

$$E_t\left[d\mathfrak{R}_t^j\right] - r_{f,t}dt = -E_t\left[d\mathfrak{R}_t^j \frac{d\xi_t}{\xi_t}\right] = \gamma E_t\left[d\mathfrak{R}_t^j dc_t\right] - \mathcal{E}_t(\delta_i)E_t\left[d\mathfrak{R}_t^j dB_{z,t}\right]. \qquad (19)$$

where $E_t[\cdot]$ denotes an expectation under the objective probability measure $\mathbb{P}$ conditional on information available at time t. Let $\Delta c_{t+1} \equiv c_{t+1} - c_t$ and $F_{z,t+1} \equiv B_{z,t+1} - B_{z,t} \sim i.i.d.\, N(0,1)$.

[4]See also Cochrane (1996); Ferson and Harvey (1999); Lettau and Ludvigson (2001); Santos and Veronesi (2006).

Approximating Equation (19) for a short discrete time interval, we have

$$E_t\left[R^j_{t+1}\right] - R_{f,t+1} \approx \gamma Cov_t\left(R^j_{t+1}, \Delta c_{t+1}\right) - \mathcal{E}_t(\delta_i) Cov_t\left(R^j_{t+1}, F_{z,t+1}\right) \tag{20}$$

where R^j_{t+1} and $R_{f,t+1}$ denote the gross return on an asset j and one-period risk-free rate from t to $t+1$, respectively. The beta representation trivially follows by rearranging the Equation (20)

$$E_t\left[R^j_{t+1}\right] - R_{f,t} \approx \beta^j_{c,t}\lambda^c + \beta^j_{z,t}\lambda^z_t \tag{21}$$

where

$$\left[\beta^j_{c,t}, \beta^j_{z,t}\right] = \left[\frac{Cov_t\left(R^j_{t+1}, \Delta c_{t+1}\right)}{\sigma^2_c}, Cov_t\left(R^j_{t+1}, F_{z,t+1}\right)\right] \tag{22}$$

$$[\lambda^c, \lambda^z_t] = \left[\gamma\sigma^2_c, -\mathcal{E}_t(\delta_i)\right]. \tag{23}$$

The beta representation above shows that the expected excess returns are described by the conditional two-factor model: the CCAPM augmented with the non-fundamental factor. In Equation (21), $\beta^j_{c,t}$ and $\beta^j_{z,t}$ are the conditional betas on the aggregate endowment risk Δc and on the non-fundamental risk F_z, respectively. The parameters λ^c and λ^z_t measure the price of Δc and F_z. Considering only the aggregate endowment risk, the excess return is reduced to the standard CCAPM of Rubinstein (1976), Lucas (1978), Breeden (1979), and Grossman and Shiller (1981) for which the market price of consumption risk is constant and determined by the coefficient of relative risk aversion γ. On the top of the aggregate consumption risk, the heterogeneous beliefs about the non-fundamental process introduce an additional factor, and the market price of risk is proportional to the consumption-weighted sentiment $\mathcal{E}_t(\delta_i)$, where pessimism regarding this measure is associated with a high market price of risk.

The result here naturally follows from the discussion in the previous section that a consumption-weighted sentiment summarizes market expectations for the price of risk on the non-fundamental, and the intuition from the ICAPM (Merton, 1973), by which the instrument that forecasts the investment opportunity can be served as a conditioning variable for tests of the cross-section.

While

In Section V, I shows that $F_{z,t}$ and $\mathcal{E}_t(\delta_i)$ may be inferred from the panel of individual consumption data. To facilitate the direct comparison between the data and the model-based prediction, I introduce a variable CWS_t (abbreviation for **C**onsumption-**W**eighted **S**entiment), an empirical proxy for the consumption-weighted sentiment. While I defer justification to Section V, I note here that CWS_t is a zero mean variable, and $\mathcal{E}_t(\delta_i)$ is positively related to CWS_t by

$$\mathcal{E}_t(\delta_i) = E\left[\mathcal{E}_t(\delta_i)\right] + \gamma \text{CWS}_t \tag{24}$$

where $E[\cdot]$ denotes an unconditional mean. That is, CWS_t is proportional to the deviation of $\mathcal{E}_t(\delta_i)$ from the long run mean. Whenever it does not cause confusion, I refer CWS_t to as the consumption-weighted sentiment.

D.2. Unconditional beta representation (scaled factor model)

To proceed standard cross-sectional regression methodologies, I write Equation (21) as an unconditional beta representation (also known as a scaled factor model). As explained in e.g., Cochrane (2009, Chapter 8.4), Lettau and Ludvigson (2001) and Ludvigson (2013), this is achieved by expanding the set of factors with a scaled factor, which is a product of the factor, whose price of risk is time-varying, and the conditioning variable; in my model, the interaction of the non-fundamental factor and the consumption-weighted sentiment, $\text{CWS}_t F_{z,t+1}$.

Taking unconditional expectation both sides of Equation (21) (details shown in Appendix B) and substituting $\mathcal{E}_t(\delta_i)$ with CWS_t by Equation (24), we get the following unconditional beta representation with three factors : Δc_{t+1}, $F_{z,t+1}$, and $\text{CWS}_t F_{z,t+1}$:

$$E_t\left[R^j_{t+1}\right] - R_{f,t+1} = \beta^j_c \lambda^c + \beta^j_z \lambda^z + \beta^j_{z \cdot cwb} \lambda^{z \cdot cwb} \tag{25}$$

where

$$\left[\beta_c^j, \beta_z^j, \beta_{z\cdot cwb}^j\right] = \left[\frac{Cov\left(R_{t+1}^j, \Delta c_{t+1}\right)}{\sigma_c^2}, Cov\left(R_{t+1}^j, F_{z,t+1}\right), \frac{Cov\left(R_{t+1}^j, \text{CWS}_t F_{z,t+1}\right)}{Var\left(\text{CWS}_t F_{z,t+1}\right)}\right] \tag{26}$$

$$\left[\lambda^c, \lambda^z, \lambda^{z\cdot cwb}\right] = \left[\gamma\sigma_c^2, -E\left[\mathcal{E}_t(\delta_i)\right], -\gamma Var\left(\text{CWS}_t F_{z,t+1}\right)\right]. \tag{27}$$

The unconditional beta representation above implies that the cross-sectional dispersion in expected returns is determined not only by its correlation with aggregate consumption growth and the non-fundamental factor but also by its correlation with the scaled factor $\text{CWS}_t F_{z,t+1}$. While they do not have a straightforward interpretation as a risk price as Lettau and Ludvigson (2001) notes, the estimates of λ^z and $\lambda^{z\cdot cwb}$ have a structural implication. First, because a variance and γ are positive, $\lambda^{z\cdot cwb}$ is expected to be negative, which reflects the notion that pessimism is associated with a high-risk premium, as emphasized throughout this section. Second, λ^z indicates the long-run average of consumption-weighted. While the value is not restricted by the model, the significant estimate of λ^z may suggest that the market view on average is biased.

Finally, the estimates of λ^z and $\lambda^{z\cdot cwb}$ can be used to uncover λ_t^z, the price of non-fundamental risk in the conditional model. From Equation (27), it is straightforward to show that the estimate of λ^z delivers an estimate of $E\left[\mathcal{E}_t(\delta_i)\right]$ while the estimate of $\lambda^{z\cdot cwb}$ provides an estimate of γ from $-\frac{\lambda^{z\cdot cwb}}{Var(\text{CWS}_t F_{z,t+1})}$. By Equation (23), λ_t^z is thus measured by the estimates of λ^z and $\lambda^{z\cdot cwb}$ together with $Var\left(\text{CWS}_t F_{z,t+1}\right)$ and CWS_t:

$$\hat{\lambda}_t^z = \hat{\lambda}^z + \frac{\hat{\lambda}^{z\cdot cwb}}{Var\left(\text{CWS}_t F_{z,t+1}\right)} \text{CWS}_t. \tag{28}$$

where $\hat{\lambda}$ denotes the estimated value.

IV. Data Description

A. *Testing assets*

I use the 25 Fama and French (1992) portfolios formed on size and book-to-market as test assets. Data are monthly, available by Professor Kenneth French on his website [5]. The sample for the empirical test is from December 1981 to November 2012, which is set by the period of the CEX data. To match the sampling frequency of CEX data, I compound the monthly returns to quarterly. All returns are expressed over the risk-free rate. The risk-free rate is from the 90-day Treasury bill, obtained from CRSP.

B. *Aggregate consumption data and inflation*

For quarterly consumption growth rate at a monthly frequency, I use seasonally-adjusted monthly real aggregate consumption of non-durables and services from the National Income and Product Accounts (NIPA) Table 2.8.3 and monthly population from NIPA Table 2.6, available from January 1959. Quarterly real per capita growth rates at monthly frequency are calculated by subtracting quarterly log population growth rates from quarterly log growth rates of aggregate per capita consumption. For inflation, I use the seasonally adjusted CPI from U.S. Bureau of Labor Statistics. Quarterly inflation is the log growth rate of the CPI in the final month of the current quarter over the last month in the previous quarter.

C. *CEX sample and definition of the consumption variables*

CEX is quarterly survey on consumer expenditures conducted by U.S. Bureau of Labor Statistics (BLS) since 1980 [6]. Each household is interviewed for five subsequent quarterly frequency with the first interview as practice (not reported in the published dataset). Interviews across households

[5] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[6] The data from 1982-1995 are provided by ICPSR, and prepared and made available by the Norwegian Social Science Data Services (NSD). Neither ICPSR nor NSD is responsible for the analysis/interpretation of the data presented here. From 1996 on, I use the public-use microdata provided by the U.S. Bureau of Labor Statistics (BLS).

are spread out over the quarters, and thus the survey results are found at any given month and each month new households enter the survey to replace those that completed their survey. At each interview, the households are asked about their expenditures for the last three months before the month in which the interview is carried out. Over the course of the sample period, the BLS changed the household identification systems three times: the first quarter of 1986, 1996, and 2005 and the survey is missing for January of each of these years. These changes make it impossible to match households in the quarters prior and after the change, and I will discuss the implication of these breaks for the empirical study in the next section. For the following discussion, I refer to the sample, which corresponds to the same sample re-design period as sub-sample 1,2,3, and 4, respectively.

I construct the household-level consumption and quarterly growth rate as follows. First, a quarterly household consumption is defined as the expenditures on nondurable goods and services. This variable is estimated by summing the disaggregate CEX consumption categories in a given quarter that match the definitions of nondurables and services in NIPA, and I follow the mapping list as in Coibion, Gorodnichenko, and Koustas (2017)[Appendix B]. Also, the consumption of each household is normalized by dividing it by the number of family members in the household to control for consumption changes driven by changes in family size. Second, a household's consumption growth between quarters $t-1$ and t is defined as the changes in log household's normalized consumption between quarters $t-1$ and t. Real consumption growth rate is formed by subtracting inflation from the nominal consumption growth rate. To match the data to the model, I assume that consumption during quarter t takes place at the beginning of the quarter. This assumption gives a higher contemporaneous correlation of consumption growth and asset returns and is advocated by Campbell (2003) [7]. Following Brav et al. (2002); Constantinides and Ghosh (2017), I refer to the households, whose expenditure corresponds to January to March, February to April, and March to May in a given year, as to January tranche, February tranche, and March tranche, respectively. To control for seasonal consumption changes, I regress the consumption growth rate on a set of sub-sample, tranche, and month dummies, and use the residual as the quarterly consumption growth measure.

[7] As discussed in Beeler and Campbell (2009), the end-of-period timing consumption, on the other hand, generates a higher correlation between consumption growth and lagged financial market data, and is used by Bansal and Yaron (2004); Parker and Julliard (2005). Engsted and Møller (2015) shows that the beginning-of-period timing convention improves the performance of the CCAPM.

As an alternative specification of aggregate consumption[8], I also construct in-sample aggregate consumption. The per capita aggregate consumption for a quarter t is computed by first summing the consumption of all households in that quarter and dividing it by the sum of the number of family members across all households in that quarter (an approach also adopted by Constantinides and Ghosh (2017)). The aggregate consumption growth between quarters $t-1$ and t is defined as the changes in log per capita aggregate consumption between quarters $t-1$ and t and the real growth rate is computed by subtracting inflation rate in the corresponding period.

The sample for the empirical test covers the interview period that corresponds to the consumption series from December 1981 to November 2012. I closely follow the exclusion criteria from previous literatures (Brav, Constantinides, and Geczy, 2002; Malloy, Moskowitz, and Vissing-Jørgensen, 2009; Constantinides and Ghosh, 2017) and these criteria are discussed in the Appendix C. The exclusion restrictions leave with a sample of 259,848 of the quarterly household consumption growth rate with about 700 households per month (approximately 45 percent of the original sample). For each month, the minimum, maximum, and mean of the number of households observed is nearly 165, 1,000, and 700, respectively. Table I reports the time series mean of the mean, the standard deviation, the skewness of the cross-sectional distribution of household-level consumption growth rate. Consistent with the previous studies (e.g., Constantinides and Ghosh (2017)), the cross-sectional volatility has a high mean value of 0.36 (all the tranche together) that is statistically significant, which indicates that there is substantial heterogeneity in the household consumption growth rates. The results for each tranche is similar.

V. Identification strategy

This section describes the identification strategy to estimate empirical proxies for Z_t and δ_h. Throughout this section, I approximate the model for discrete time with one quarter as one unit of time to match the CEX data sampling frequency and use lowercase letters to denote log variables. Appendix C provides the details of derivations.

[8]Attanasio (1999) discusses that the aggregated individual data may differ from the aggregate variable from National Accounts by construction.

I start from the observation that Equation (18) implies that the equilibrium consumption process for each household h is approximated by the following linear model:

$$\Delta c_{h,t+1} \approx \underbrace{\Delta c_{t+1} + A_{0,t} + A_{1,t} F_{z,t+1}}_{\text{common variation}} \underbrace{- \frac{1}{2\gamma}\delta_h^2 + \frac{1}{\gamma}\delta_h F_{z,t+1}}_{\text{idiosynctractic variation}}, \quad h = 1, \cdots, N_{t+1} \tag{29}$$

where $A_{0,t}$ and $A_{1,t}$ collect the terms that are common across the households; $F_{z,t} \equiv B_{z,t+1} - B_{z,t} \sim i.i.d.\, N(0,1)$ represents the shock to Z_{t+1}; and $\Delta c_{t+1} \equiv c_{t+1} - c_t$ denotes the rate of log aggregate consumption growth. In order to isolate the idiosyncratic component, I consider de-meaned series, subtracting the cross-sectional mean from each of $\Delta c_{h,t+1}$, and the resulting process follows:

$$\Delta \tilde{c}_{h,t+1} \approx -\frac{1}{2\gamma}\left[\delta_h^2 - \frac{1}{N_{t+1}}\sum_{k=1}^{N_{t+1}} \delta_k^2\right] + \frac{1}{\gamma}\left[\delta_h - \frac{1}{N_{t+1}}\sum_{k=1}^{N_{t+1}} \delta_k\right] F_{z,t+1}. \tag{30}$$

The role of Equation (30) is to relate $\Delta \tilde{c}_{h,t+1}$ to δ_h, and $F_{z,t+1}$, controlling for the common variation. The key relationship I exploit in this paper is that δ_h, and $F_{z,t+1}$ can be inferred from the consumption process $\Delta \tilde{c}_{h,t+1}$, whose data is available from CEX. Specifically, $F_{z,t+1}$ is a common factor that drives $\Delta \tilde{c}_{h,t+1}$, while δ_h, the individual sentiment, is identified as a factor loading of $\Delta \tilde{c}_{h,t+1}$ on $F_{z,t+1}$. Empirically, I use principal components analysis (PCA) to extract common factors and regress $\Delta \tilde{c}_{h,t+1}$ on principal component.

Applying the procedure to CEX data requires one additional step. This is because the CEX data provides only maximum three-quarters of data per household so that individual households cannot be traced over a long period, which makes it difficult to make inference at an individual household level. To overcome this issue, I use the group-level aggregated variables as a basis for the principal component analysis, where the groups are constructed such a way that households are matched by their consumption experiences. In the following two sub-sections, I introduce the approach to construct groups and describe the procedure to extract $F_{z,t+1}$ and δ_h.

A. Group construction

A reasonable approach to extend the time-series would be to allocate individual households who share common characteristics into a group and study the groups instead of the households. For this paper, the common characteristics of interest are their subjective beliefs, which is, however, not directly observable. The first step of the empirical analysis is, therefore, constructing groups, who are likely to share the subjective belief, from the household level CEX data. In the following, I show that group of households who experience comparable consumption growth rates contemporaneously makes a good candidate [9].

Section III shows that the growth rate of the consumption share reflects the idiosyncratic component of individual consumption growth rates and it is determined by her subjective belief or sentiment, the only source of idiosyncratic consumption growth rate in the model. Conversely, this result suggests that individuals who experience identical consumption share growth rate ex-post must have shared the same belief ex-ante. Applying this prediction implies that two individual households, who experience the comparable consumption share growth rates in a given quarter should be assigned to the same group. That is, the model implies that there is a direct link between the variable that represents the individual sentiment: δ_h and the idiosyncratic consumption growth rate of an individual household.

Based on the discussion above, I describe the procedure for forming the groups from the individual consumption data. For each quarter t, I first rank by their $\Delta\tilde{c}_{h,t}$ households, who start their survey in that quarter, into m groups. This approach results in each household allocated to $i_t = 1, \cdots, m$ groups. Second, I construct, for each group i_t, a cross-sectional mean of $\Delta\tilde{c}_{h,t}$ for the households that belong to i_t. Let $\Delta\tilde{c}_{i_t,t}$ denote this group mean. Third, I chain the group identity i_t over consecutive quarters so that groups can be traced over time by the following approach. In a given quarter t (except for the beginning of the sample period), there are two types of households observed; one, who starts the survey before t and the other, who enter the survey at t. For an illustration, let denote the identity of existing group by $k_{t-1} = 1, \cdots, m$ and that of new group

[9] Previous studies using the household level data but with a different focus of studies use cohort analysis using variables such as birth-year (Malmendier and Nagel, 2011) and other demographic characteristics(race, level of education) (Calvet, Campbell, and Sodini, 2009).

by $i_t = 1, \cdots, m$. For each pair (k_{t-1}, i_t), I compute the absolute distance of consumption growth rates between them as $|\Delta\tilde{c}_{k_{t-1},t} - \Delta\tilde{c}_{i,t}|$ and match the groups with the lowest distance. Starting from the mid-point of the sample period and repeating these steps over consecutive quarters, each household is allocated to one of the groups, indexed by $i = 1, \cdots, m$. The resulting group-level sample is a balanced panel $m \times T$, where T denotes the number of time series. In practice, I form the groups for each of sub-sample and tranche, separately.

From the synthetic panel formed by the above procedure, I construct two aggregated variables at a group level. For a group, indexed by $i = 1, \cdots, m$, the consumption growth rate at group-level is computed as

$$\Delta\tilde{c}_{i,t+1} \equiv \frac{1}{\bar{N}_{t+1}} \sum_{h_i=1}^{\bar{N}_{t+1}} \Delta\tilde{c}_{h_i,t+1} \tag{31}$$

where $\bar{N}_{t+1} \equiv \frac{N_{t+1}}{m}$ is the number of households in a given group i at quarter $t+1$, N_{t+1} the total number of households observed at quarter t; and $h_{i,t}$ is index for household, who belongs to group i. Also, the consumption share of group i is defined as the ratio of an aggregated consumption in that group to the aggregated consumption:

$$\omega_{i,t} \equiv \frac{\sum_{h_i=1}^{\bar{N}_t} C_{h_i,t}}{\sum_{h=1}^{N_t} C_{h,t}}. \tag{32}$$

The definition of group means together with the household-level consumption process in Equation (30) implies that

$$\Delta\tilde{c}_{i,t+1} \approx -\frac{1}{2\gamma}\left(\delta_{2,i}^2 - \frac{1}{m}\sum_{k=1}^{m}\delta_{2,k}^2\right) + \frac{1}{\gamma}\left(\delta_i - \frac{1}{m}\sum_{k=1}^{m}\delta_k\right)F_{z,t+1}. \tag{33}$$

where $\delta_i \equiv \frac{1}{\bar{N}_{t+1}}\sum_{h_i=1}^{\bar{N}_{t+1}}\delta_{h_i}$ and $\delta_{2,i} \equiv \frac{1}{\bar{N}_t}\sum_{h_i=1}^{\bar{N}_t}\delta_{2,h_i}^2$. Clearly, Equation (33) is analogous to (30) while δ_h in the second term is replaced by the group mean δ_i. For this reason, I interpret δ_i as an individual sentiment that summarizes the beliefs of the group and extract the common factor based on Equation (33), following the procedure described at the beginning of the section.

As a baseline, I study the panel of 10 groups, which implies that each group, on average, consists of 70 households. In Subsection VI.B, I show that alternative numbers of groups do not change the result considerably.

B. Extracting factors and subjective beliefs

From the synthetic panel, I estimate the latent variables $F_{z,t+1}$ and δ_i for $i = 1, \cdots, m$ by principal component analysis. The procedure is as follows. First, I use the first principal component as a proxy for $F_{z,t+1}$. The principal component measures the co-movement in the panel, and thus, naturally serves the purpose of capturing the underlying factor that drives the cross-section of consumption. Section III shows that the fundamental channel by which the heterogeneous beliefs drives the fluctuation in asset prices is substantial variation in the cross-sectional distribution of consumption. By this reasoning, we would expect that the latent variable that represents $F_{z,t+1}$ to matter for the asset prices should explain the significant variance of the panel. Because by construction, it explains the largest common variance of the panel, I argue that the first principal component makes it a good candidate.

Table II shows the result of the principal component analysis on the synthetic panel with ten groups ($m = 10$). The table presents the first three components, the respective eigenvalues and variance explained. The first component explains on average 33% of consumption growth rates, the second explains 22%, and the third explains 15%. Because it is purely statistical measure, the easiest way to interpret the principal component is to consider it as a latent component that drives the common component of a collection of variables investors disagree about.

Next, I interpret the factor loading obtained from a regression of group-level consumption growth rate on this factor as the individual sentiment of each group. For each $i = 1, \cdots, m$, I run the following OLS regression :

$$\Delta\tilde{c}_{i,t+1} = a_i + b_i F_{z,t+1} + \varepsilon_{i,t+1}. \tag{34}$$

Comparing Equation (33) with Equation (34), the point estimate of b_i denoted by $\hat{b}_i$ is linked to δ_i

by the following relationship:

$$\hat{b}_i = \frac{1}{\gamma}\left[\delta_i - \frac{1}{m}\sum_{j=1}^{m}\delta_j\right]. \tag{35}$$

That is, the factor loading on $F_{z,t+1}$ of $\Delta\tilde{c}_{i,t+1}$ is positively correlated with the individual sentiment. As discussed in Section III, investors' consumption increases when the realization of $F_{z,t+1}$ is consistent with her view; upon positive realization of $F_{z,t+1}$, consumption of optimistic investors increase, which induces a positive correlation between $\Delta\tilde{c}_{i,t+1}$ and $F_{z,t+1}$ while that of pessimistic investors decreases, which induces a negative correlation.

Finally, I estimate the empirical proxy for the consumption-weighted sentiment from Equation (35) and Equation (32). First, define an intermediate variable CWS as

$$CWS_t \equiv \sum_{i=1}^{m}\left(\hat{b}_i\omega_{i,t}\right) - E\left[\sum_{i=1}^{m}\left(\hat{b}_i\omega_{i,t}\right)\right] = \frac{1}{\gamma}\left[\sum_{i=1}^{m}\left(\delta_i\omega_{i,t}\right) - E\left[\sum_{i=1}^{m}\left(\delta_i\omega_{i,t}\right)\right]\right] \tag{36}$$

where $E[\cdot]$ denotes the unconditional expectation and $\omega_{i,t}$ is a consumption share defined in Equation (32). Given the definition of the consumption-weighted sentiment $\mathcal{E}_t(\delta_i)$ in Equation (17), CWS is related to $\mathcal{E}_t(\delta_i)$ by the following:

$$\mathcal{E}_t(\delta_i) = E\left[\mathcal{E}_t(\delta_i)\right] + \gamma \mathrm{CWS}_t. \tag{37}$$

as claimed in Subsection III.D.1.

To control for variation due to aggregate consumption shock and to be consistent with the model assumption, I orthogonalize $F_{z,t+1}$ and $\mathrm{CWS}_t F_{z,t+1}$ to the different specifications of the market portfolio: the aggregate consumption growth and the market return.

VI. Empirical results

This section shows empirical results that investigate the model-based predictions using the variables estimated by the procedure described in Section V. I use the beta representation Equation (25) as the basis of the empirical test, which relates returns to factors over time:

$$R^j_{t+1} - R_{f,t+1} = \alpha_j + \beta^j_c \Delta c_{t+1} + \beta^j_z F_{z,t+1} + \beta^j_{z\cdot cwb} \mathrm{CWS}_t F_{z,t+1} + \varepsilon_{j,t+1} \quad j = 1, \cdots, N, \quad t = 0, \cdots, T-1. \tag{38}$$

Here Δc_{t+1} is the aggregate consumption growth rates, $F_{z,t+1}$ is the first principal component extracted from the panel of consumption growth rates described in Section V, and CWS_t is an empirical counterpart to the de-meaned consumption-weighted sentiment, described in Section V. Following the standard in the literature, I allow a nonzero intercept, α_j in the estimation, although the model implies that the intercept is zero. To prevent potential multicollinearity problems between $F_{z,t+1}$ and $\mathrm{CWS}_t F_{z,t+1}$, I include in $\mathrm{CWS}_t F_{z,t+1}$ only the component that is orthogonal to $F_{z,t+1}$.

A. Fama-MacBeth Regression

Table IV reports the estimates for the market prices of risk obtained from the second stage of the Fama-MacBeth regression for several empirical specifications with Fama and French 25 portfolios as test assets. Together with the coefficient estimates, reported are uncorrected t-statistics and Shanken (1985)-corrected t-statistics, which takes into account for the sampling variation of betas estimated in the first stage, for these coefficients. While the success of the model cannot be judged by R^2, I also report an unadjusted and adjusted for the degrees of freedom R^2 statistics for cross-sectional regressions to provide an intuitive measure of the cross-sectional fit and comparison to the previous studies.

[Place Table IV about here]

To compare with the main results, I first begin by presenting results of the CAPM and the CCAPM. Row 1 and 2 presents the results of the CCAPM with a different specification of the aggregate consumption growth, where the first-row aggregate consumption is from NIPA and the second row from the CEX sample. Row 3 presents the results of the CAPM with the CRSP value-weighted return used as a proxy for the market return. The t-statics of the λ estimates for the CCAPM and CAPM show that the aggregate consumption risk and the market return are not statistically significant determinants of the cross-section of average returns. Also, the intercepts in all three specifications are significant, and CCAPM estimated from the NIPA aggregate consumption and the CAPM show the wrong sign of market price of risk, consistent with the previous literature, e.g., Wang (1996); Lettau and Ludvigson (2001); Santos and Veronesi (2006).

Row 5 to Row 7 report the main results of the model: estimation of the model in Equation (38) with a different specification of Δc. The conditional model considerably improves the cross-sectional fit: adding the second factor to the market return raises the unadjusted R^2 from 25% to 62%. In all three specifications, the coefficients on the scaled factor $CWS_t \cdot F_{z,t+1}$ are statistically significant while those of the factor $F_{z,t+1}$ is not statistically significantly different from zero. As discussed in Section III and Section V, these coefficients have structural implications to the model. First, the coefficient on the scaled factor is expected to be negative, and this prediction is strongly supported by the results, which is consistent with the model-based prediction that the aggregate pessimism, a negative value of CWS_t, predicts a higher market price of risk. Second, the coefficient of $F_{z,t+1}$ indicates the consumption-weighted sentiment on average is unlikely different from zero, which implies that the expectation of the market on average is unbiased.

The results suggest that an asset's risk is substantially determined by its correlation with the non-fundamental factor conditional on the sentiment of the economy. In particular, stocks (e.g., value stocks) earn higher average returns because they are more highly correlated with the non-fundamental factor in bad times; when the economy becomes pessimistic, and risk premium is high. Figure 1 gives a visual summary of the model's performance. The vertical axis is the average return of the test assets, and the horizontal axis is the predicted values from Table IV . If the model is correct, the points should all lie on a 45 degree line. The top right panel (corresponding to Row 5 in Table IV) shows that the points lie much closer to this prediction then they do in the CCAPM (Row 1 in Table IV) and the CAPM (Row 3 in Table IV).

[Place Figure 1 about here]

Row 9 includes Fama and French three factors in Equation (38), and for comparison, Row 8 reports the performance of Fama and French three-factor model alone. The purpose of this exercise is not to compare the performance of two models, but instead to assess whether the new factor survives inclusion of three factors. Indeed, the generality of the model assumption implies that the new factor may or may not capture the variation of HML (the high-minus-low portfolio) or SMB (the small-minus-big portfolio). Row 9 shows that the scaled factor survives the inclusion of HML and SMB. Compared with Row 7, the coefficient on the scaled factor remains significant even though its magnitude is slightly reduced. This result indicates that the scaled factor may also capture the cross-sectional dispersion in returns not explained by HML and SMB.

As Lettau and Ludvigson (2001) notes, Shanken (1985) correction to the t-statistics is large for models that include scaled macroeconomic factors rather than unscaled returns. For all specification, however, the t-statistics of the coefficient on the scaled factor is not considerably affected by the Shanken (1985) correction and remains strongly significant, which highlight the robustness of my results. Overall the result suggests the existence of a priced risk generated by disagreement and gives support to the view that the cross-sectional dispersion in expected returns is attributed to covariance with common non-diversifiable risk factor instead of firm characteristics.

Lastly, except Row 4 and Row 6 for the Shanken corrected t-statistics, the intercept is still significantly different from zero although the result shows that the non-fundamental factor scaled by the consumption-weighted sentiment considerably explains the cross-sectional spread in average returns. The similar result can be found in models where macro-variables are used as conditioning variables e.g., Wang (1996); Lettau and Ludvigson (2001); Santos and Veronesi (2006),

The results separately for the January, February, and March Tranche are consistent with that for the full sample (reported in Internet Appendix).

B. Robustness

B.1. Alternative number of groups

Table V reports the estimates of the second stage of Fama and Macbeth regression, using alternative number of group. The main results largely remain unchanged.

B.2. Alternative grouping method

To show that the grouping method by consumption growth rate is valid strategy, I also consider random assignment; that is for each period, households are assigned to randomly selected group of 10. The purpose of this exercise is to show that the randomly assigned group should not show significant pattern to explain the asset prices. Table VIII reports the average of 100 simulated results. As expected, the factors become insignificant.

B.3. Other test assets

I test the factor model using two different sets of other test asset instead of the 25 Fama and French size and B/M sorted portfolios to examine whether the results are sensitive to the specific test asset. This exercise addresses some of the criticism of the recent consumption asset pricing literature (e.g.,Lewellen, Nagel, and Shanken (2010)). In particular, I consider two alternative set of test assets: 1) 25 Fama and French together with 5 industry portfolio and 2) 25 Fama and French together with 30 industry portfolio. Table V and Table VI report the results. The main results are in line with those obtained for the 25 Fama and French portfolios.

B.4. Stockholder

Table IX reports the result when only stockholders are considered for the analysis, where stockholders are identified by the method suggested by Malloy, Moskowitz, and Vissing-Jørgensen (2009). The main finding remains the same : statistically significant negative estimate of the scaled factor. Also, the performance of the model is significantly improved : R^2 substantially increases, the intercept becomes insignificant (evaluated in terms of Shanken corrected t-statistics), and the magnitude of the coefficient on the scaled factor increases. Yet, this finding should be interpreted with caution because considering only stockholders substantially reduces the sample (about 20% of the total sample) so that it makes it difficult to compare with the result with the total sample.

VII. Conclusion

This paper derives and tests the implications of heterogeneous beliefs about a non-fundamental process for the cross-section of expected returns. In equilibrium, speculation among the individual investors and trading that follows make the non-fundamental process priced in the financial market. As the trading outcomes endogenously result in the redistribution of wealth and consumption, the market price of risk, which reflects the consumption-weighted sentiment, fluctuates over time with an increase in the price of risk coinciding in the period when the economy is dominated by pessimists. Empirically, I estimate the latent non-fundamental process and the measure of consumption weighted optimism by principal component analysis on the cross-section of consumption data. Testing the factor model in the cross-section of returns, I find that the results are consistent with the model-based predictions.

VIII. Figures and Tables

Table I
Summary Statistics of Growth rate of Consumption Share: CEX data

The table reports summary statistics of the number of households observed in the sample and the point estimates of the standard deviation, and the skewness of the cross-sectional distribution of quarterly household consumption share growth. Standard errors are in parentheses.

	Number of Household			Cross-sectional Moment	
	Min	Mean	Max	Standard Deviation	Skewness
Tranche					
All	166.00	704.20	1002.00	0.38	-0.00
				(0.002)	(0.011)
Jan	198.00	706.39	991.00	0.37	0.00
				(0.003)	(0.019)
Feb	187.00	707.90	1002.00	0.38	0.01
				(0.003)	(0.021)
Mar	166.00	698.35	982.00	0.38	-0.01
				(0.003)	(0.019)

Table II
Principal Components Analysis on 10 groups

The table shows principal components analysis of 10 groups. Variance explained is reported in percentage point.

		Eigenvalue			Variance Explained		
		Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3
Sub-sample	Tranche						
1.0	Jan	4.54	2.17	1.38	45.41	21.67	13.75
	Feb	3.85	2.55	1.61	38.51	25.49	16.12
	Mar	3.91	2.57	1.52	39.06	25.70	15.16
2.0	Jan	2.59	1.91	1.66	25.88	19.05	16.61
	Feb	2.82	1.90	1.42	28.21	19.03	14.20
	Mar	3.08	1.99	1.60	30.78	19.91	15.97
3.0	Jan	2.52	2.12	1.55	25.19	21.16	15.45
	Feb	3.39	2.50	1.41	33.92	25.00	14.12
	Mar	2.93	1.97	1.71	29.33	19.72	17.11
4.0	Jan	2.79	1.88	1.60	27.92	18.83	15.96
	Feb	3.42	2.38	1.48	34.17	23.79	14.82
	Mar	3.28	2.14	1.49	32.82	21.35	14.91

Table III
Summary Statistics of the Factors

The table reports estimates of the correlation coefficient of F_z with other factors using data from December 1981 to November 2012. The factors are 1) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment; 2) Δc the log growth rate of aggregate consumption computed from NIPA table; 3) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; 4) Mkt the excess market return; 5) the Fama and French (2015) 5 Factors, where SMB, HML, RMW, and CMA are the Fama-French mimicking portfolios related to size, book-to-market ratios, operating profitability and rate of investment growth.

	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	SMB	HML	RMW	CMA
Mean	0.0	-0.0	0.4	-0.7	1.9	0.4	1.0	1.2	1.1
Std	11.2	1.2	0.4	4.7	8.4	5.3	5.8	5.0	3.9
Min	-31.6	-6.7	-1.2	-13.1	-31.3	-22.8	-18.6	-29.7	-11.0
25p	-9.1	-0.3	0.1	-3.5	-2.6	-3.0	-2.2	-1.1	-1.3
50p	-0.0	0.3	0.5	-0.9	2.6	-0.1	0.8	1.0	0.9
75p	8.5	0.7	0.7	2.0	7.0	3.2	3.6	3.2	3.4
Max	37.2	2.0	1.4	31.1	26.3	30.7	26.0	28.6	20.8

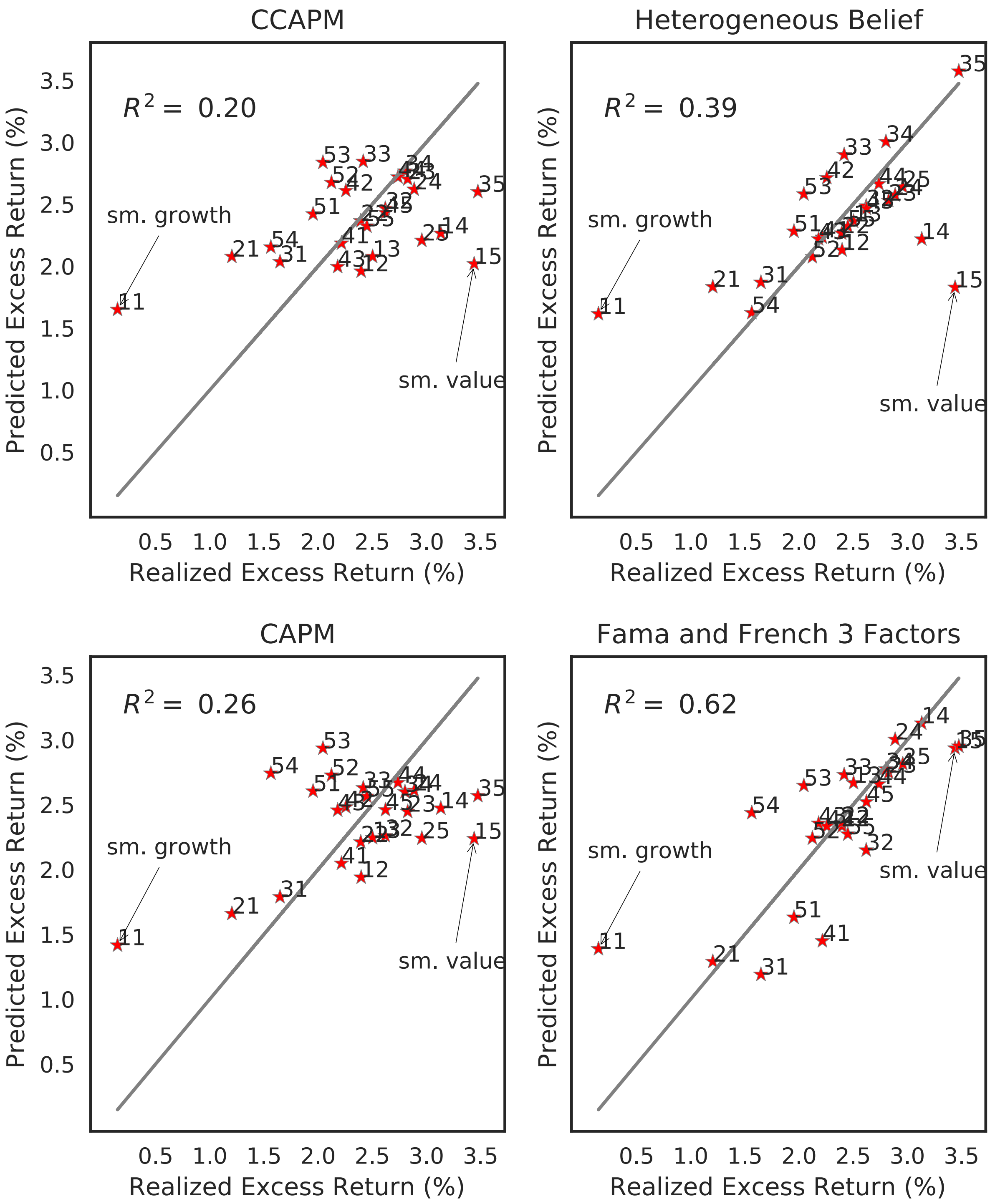

Figure 1. Realized vs. Fitted returns : 25 Fama-French Portfolios (All)

The figure shows the pricing errors for each of the 25 Fama-French portfolios for the four models. Top left panel(the CCAPM) corresponds to the row 1 in Table IV; the top right panel(Heterogeneous belief) the row 5 in Table IV; the bottom left panel (the CAPM) the row 3 in Table IV; and the bottom left panel (Fama and French (1992) 3 Factors) the row 9 in Table IV. For each point, the first digit refers to the size quintiles (1 the smallest to 5 the largest), and the second digit book-to-market quintiles(1 the portfolio with the lowest book-to-market ratio to 5 with the highest). The solid line is 45° line. Reported are quarterly returns of % (not annualized).

Table IV
Fama-Macbeth regressions using 25 Fama-French Portfolios

The table reports the estimation of the market prices of risk, using the Fama and MacBeth (1973) two-pass regressions for December 1981 to November 2012, overlapping 372 quarters in monthly frequency on 25 Fama-French Portfolios as test assets. Two t-statistics with the Newey and West (1987) weight allowing for autocorrelation up to 2 months for each coefficient estimate are reported in parentheses. The top statistic uses uncorrected Fama and MacBeth (1973) standard errors while bottom statistic uses the Shanken (1992) correction. For each row, R^2 columns reports the unadjusted cross-sectional R^2 statistics and R^2 adjusted for the degrees of freedom in parentheses. The asset-pricing factors are Row 1) Δc the log growth rate of aggregate consumption computed from NIPA table; Row 2) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; Row 3) MKT the excess market return; Row 4) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment; Row 9) Fama and French (1992) 3 Factors where SMB and HML are the Fama and French (1992) mimicking portfolios related to size and book-to-market ratios. * : $p < 0.10$, ** : $p < 0.05$, *** : $p < 0.01$

		Const	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	SMB	HML	R^2
	$\hat{\lambda}$	4.25			-0.34					0.20
1	t-stat(fm)	(4.50***)			(-1.75*)					(0.20)
	t-stat(sc)	(3.20***)			(-1.25)					
	$\hat{\lambda}$	2.43				0.84				0.01
2	t-stat(fm)	(3.33***)				(0.69)				(0.01)
	t-stat(sc)	(3.31***)				(0.74)				
	$\hat{\lambda}$	4.76					-2.22			0.26
3	t-stat(fm)	(3.49***)					(-1.54)			(0.26)
	t-stat(sc)	(3.39***)					(-1.45)			
	$\hat{\lambda}$	2.07	-2.74	-1.64						0.40
4	t-stat(fm)	(3.04***)	(-0.47)	(-3.91***)						(0.40)
	t-stat(sc)	(1.36)	(-0.30)	(-2.89***)						
	$\hat{\lambda}$	3.30	0.35	-1.34	-0.22					0.45
5	t-stat(fm)	(4.43***)	(0.07)	(-3.26***)	(-1.73*)					(0.45)
	t-stat(sc)	(2.85***)	(0.05)	(-2.53**)	(-1.30)					
	$\hat{\lambda}$	2.09	-2.80	-1.64		0.26				0.40
6	t-stat(fm)	(3.15***)	(-0.48)	(-3.83***)		(0.20)				(0.40)
	t-stat(sc)	(1.39)	(-0.31)	(-2.80***)		(0.15)				
	$\hat{\lambda}$	5.71	8.84	-1.05			-3.51			0.62
7	t-stat(fm)	(4.53***)	(1.81*)	(-3.18***)			(-2.67***)			(0.62)
	t-stat(sc)	(3.74***)	(1.48)	(-2.54**)			(-2.22**)			
	$\hat{\lambda}$	5.35					-3.38	0.29	1.09	0.62
8	t-stat(fm)	(6.01***)					(-3.08***)	(0.71)	(2.30**)	(0.62)
	t-stat(sc)	(5.60***)					(-3.03***)	(0.69)	(2.25**)	
	$\hat{\lambda}$	4.79	6.21	-0.73			-2.81	0.18	0.94	0.69
9	t-stat(fm)	(5.38***)	(1.24)	(-2.89***)			(-2.62***)	(0.45)	(2.01**)	(0.69)
	t-stat(sc)	(4.69***)	(1.09)	(-2.73***)			(-2.47**)	(0.44)	(1.92*)	

Table V
Fama-Macbeth regressions using 25 Fama-French Portfolios and 5 Industry Portfolios

The table reports the estimation of the market prices of risk, using the Fama and MacBeth (1973) two-pass regressions for December 1981 to November 2012, overlapping 372 quarters in monthly frequency on 25 Fama-French Portfolios and 5 industry portfolio as test assets. Two GMM T-statistics with the Newey and West (1987) weight allowing for autocorrelation up to 2 months for each coefficient estimate are reported in parentheses. The top statistic uses uncorrected Fama-MacBeth standard errors while bottom statistic uses the Shanken (1992) correction. For each row, R^2 columns reports the unadjusted cross-sectional R^2 statistics and R^2 adjusted for the degrees of freedom in parentheses. The asset-pricing factors are Row 1) Δc the log growth rate of aggregate consumption computed from NIPA table; Row 2) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; Row 3) Mkt the excess market return; Row 4) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment; Row 9) the Fama and French (1992) 3 Factors, where SMB and HML are the Fama and French (1992) mimicking portfolios related to size and book-to-market ratios. * : $p < 0.10$, ** : $p < 0.05$, *** : $p < 0.01$

		Const	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	SMB	HML	R^2
	$\hat{\lambda}$	3.68			-0.24					0.17
1	t-stat(fm)	(4.98***)			(-1.57)					(0.17)
	t-stat(sc)	(3.97***)			(-1.30)					
	$\hat{\lambda}$	2.41				1.04				0.01
2	t-stat(fm)	(3.44***)				(0.88)				(0.01)
	t-stat(sc)	(3.41***)				(0.94)				
	$\hat{\lambda}$	4.11					-1.69			0.19
3	t-stat(fm)	(3.62***)					(-1.31)			(0.19)
	t-stat(sc)	(3.53***)					(-1.26)			
	$\hat{\lambda}$	2.10	-2.64	-1.50						0.36
4	t-stat(fm)	(3.19***)	(-0.45)	(-3.31***)						(0.36)
	t-stat(sc)	(1.56)	(-0.31)	(-2.58**)						
	$\hat{\lambda}$	3.20	0.49	-1.28	-0.20					0.44
5	t-stat(fm)	(5.33***)	(0.09)	(-2.92***)	(-1.89*)					(0.44)
	t-stat(sc)	(3.12***)	(0.07)	(-2.29**)	(-1.47)					
	$\hat{\lambda}$	2.15	-2.75	-1.49		0.55				0.37
6	t-stat(fm)	(3.38***)	(-0.47)	(-3.25***)		(0.44)				(0.36)
	t-stat(sc)	(1.61)	(-0.32)	(-2.48**)		(0.34)				
	$\hat{\lambda}$	4.95	7.13	-1.19			-2.82			0.59
7	t-stat(fm)	(5.42***)	(1.51)	(-2.95***)			(-2.63***)			(0.59)
	t-stat(sc)	(4.53***)	(1.21)	(-2.35**)			(-2.29**)			
	$\hat{\lambda}$	4.64					-2.63	0.26	0.97	0.55
8	t-stat(fm)	(6.32***)					(-2.71***)	(0.65)	(2.02**)	(0.55)
	t-stat(sc)	(6.02***)					(-2.66***)	(0.64)	(2.00**)	
	$\hat{\lambda}$	4.52	5.98	-0.82			-2.49	0.13	0.80	0.64
9	t-stat(fm)	(6.17***)	(1.39)	(-3.54***)			(-2.59**)	(0.32)	(1.69*)	(0.64)
	t-stat(sc)	(5.28***)	(1.17)	(-3.22***)			(-2.47**)	(0.32)	(1.65)	

Table VI
Fama-Macbeth regressions using 25 Fama-French Portfolios and 30 Industry Portfolios

The table reports the estimation of the market prices of risk, using the Fama and MacBeth (1973) two-pass regressions for December 1981 to November 2012, overlapping 372 quarters in monthly frequency on 25 Fama-French Portfolios and 30 industry portfolio as test assets. Two GMM T-statistics with the Newey and West (1987) weight allowing for autocorrelation up to 2 months for each coefficient estimate are reported in parentheses. The top statistic uses uncorrected Fama-MacBeth standard errors while bottom statistic uses the Shanken (1992) correction. For each row, R^2 columns reports the unadjusted cross-sectional R^2 statistics and R^2 adjusted for the degrees of freedom in parentheses. The asset-pricing factors are Row 1) Δc the log growth rate of aggregate consumption computed from NIPA table; Row 2) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; Row 3) Mkt the excess market return; Row 4) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment; Row 9) the Fama and French (1992) 3 Factors, where SMB and HML are the Fama and French (1992) mimicking portfolios related to size and book-to-market ratios. * : $p < 0.10$, ** : $p < 0.05$, *** : $p < 0.01$

		Const	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	SMB	HML	R^2
	$\hat{\lambda}$	3.15			-0.17					0.18
1	t-stat(fm)	(5.01***)			(-1.41)					(0.18)
	t-stat(sc)	(4.56***)			(-1.35)					
	$\hat{\lambda}$	2.26				-0.33				0.00
2	t-stat(fm)	(3.21***)				(-0.29)				(0.00)
	t-stat(sc)	(3.16***)				(-0.29)				
	$\hat{\lambda}$	3.33					-1.01			0.12
3	t-stat(fm)	(4.31***)					(-1.02)			(0.12)
	t-stat(sc)	(4.30***)					(-1.01)			
	$\hat{\lambda}$	2.27	-3.45	-0.62						0.20
4	t-stat(fm)	(3.42***)	(-0.64)	(-1.79*)						(0.19)
	t-stat(sc)	(2.99***)	(-0.61)	(-1.68*)						
	$\hat{\lambda}$	3.11	0.06	-0.60	-0.17					0.33
5	t-stat(fm)	(4.83***)	(0.01)	(-1.75*)	(-1.62)					(0.33)
	t-stat(sc)	(3.86***)	(0.01)	(-1.65)	(-1.62)					
	$\hat{\lambda}$	2.31	-3.46	-0.66		0.61				0.20
6	t-stat(fm)	(3.40***)	(-0.64)	(-2.00**)		(0.63)				(0.20)
	t-stat(sc)	(2.79***)	(-0.60)	(-1.83*)		(0.58)				
	$\hat{\lambda}$	3.27	0.13	-0.59			-0.99			0.27
7	t-stat(fm)	(4.76***)	(0.03)	(-1.71*)			(-1.10)			(0.27)
	t-stat(sc)	(4.62***)	(0.03)	(-1.61)			(-1.09)			
	$\hat{\lambda}$	3.42					-1.34	0.18	0.69	0.26
8	t-stat(fm)	(5.55***)					(-1.52)	(0.42)	(1.38)	(0.26)
	t-stat(sc)	(5.77***)					(-1.55)	(0.42)	(1.38)	
	$\hat{\lambda}$	3.32	-5.56	-0.48			-1.26	0.19	0.73	0.37
9	t-stat(fm)	(5.34***)	(-1.25)	(-1.33)			(-1.42)	(0.45)	(1.44)	(0.37)
	t-stat(sc)	(5.11***)	(-1.16)	(-1.24)			(-1.44)	(0.47)	(1.44)	

Table VII
Fama-Macbeth regressions using 25 Fama-French Portfolios (Alternative number of groups)

The result is based on alternative numbers of group $m = 5, 7, 13, 15$ instead of the base-line case of 10 (see the details of the grouping procedure in Sub-section V.A). The table reports the estimation of the market prices of risk, using the Fama-MacBeth procedure for December 1981 to November 2012, overlapping 372 quarters in monthly frequency on 25 Fama-French Portfolios as test assets. GMM T-statistics with Newey and West Newey and West (1987) weight allowing for autocorrelation up to 2 months for each coefficient estimate are reported in parentheses. T-statistic uses the Shanken Shanken (1992) correction. For each row, R^2 columns reports the unadjusted cross-sectional R^2 statistics and R^2 adjusted for the degrees of freedom in parentheses. The asset-pricing factors are Row 1) Δc the log growth rate of aggregate consumption computed from NIPA table; Row 2) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; Row 3) Mkt the excess market return; Row 4) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment.

group			Const	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	R^2
5	4	$\hat{\lambda}$	1.78	-0.75	-1.62				0.59
		t-stat(sc)	(1.29)	(-0.12)	(-2.33**)				
	5	$\hat{\lambda}$	2.80	-2.46	-1.39	-0.18			0.61
		t-stat(sc)	(2.34**)	(-0.43)	(-1.99**)	(-1.31)			
	6	$\hat{\lambda}$	1.78	-0.75	-1.62		-0.01		0.59
		t-stat(sc)	(1.26)	(-0.12)	(-2.28**)		(-0.01)		
	7	$\hat{\lambda}$	4.05	-4.68	-1.18			-2.07	0.67
		t-stat(sc)	(3.30***)	(-1.25)	(-2.39**)			(-1.54)	
7	4	$\hat{\lambda}$	2.08	-5.22	-2.28				0.40
		t-stat(sc)	(1.16)	(-0.51)	(-1.64)				
	5	$\hat{\lambda}$	3.41	-7.41	-1.66	-0.24			0.45
		t-stat(sc)	(2.48**)	(-0.88)	(-1.45)	(-1.39)			
	6	$\hat{\lambda}$	2.15	-5.05	-2.29		0.80		0.41
		t-stat(sc)	(1.19)	(-0.48)	(-1.64)		(0.39)		
	7	$\hat{\lambda}$	5.73	-13.43	-0.81			-3.42	0.59
		t-stat(sc)	(3.68***)	(-2.04**)	(-1.29)			(-2.12**)	
13	4	$\hat{\lambda}$	2.12	2.28	-1.30				0.29
		t-stat(sc)	(1.87*)	(0.34)	(-2.94***)				
	5	$\hat{\lambda}$	2.93	1.17	-1.09	-0.14			0.31
		t-stat(sc)	(3.14***)	(0.19)	(-2.42**)	(-1.05)			
	6	$\hat{\lambda}$	2.14	2.30	-1.29		0.28		0.29
		t-stat(sc)	(1.84*)	(0.35)	(-2.86***)		(0.18)		
	7	$\hat{\lambda}$	4.89	-4.31	-0.99			-2.52	0.38
		t-stat(sc)	(5.15***)	(-0.96)	(-2.54**)			(-2.19**)	
15	4	$\hat{\lambda}$	2.01	-2.58	-1.35				0.21
		t-stat(sc)	(1.63)	(-0.49)	(-2.31**)				
	5	$\hat{\lambda}$	3.59	-2.38	-1.07	-0.27			0.33
		t-stat(sc)	(2.71***)	(-0.44)	(-1.77*)	(-1.04)			
	6	$\hat{\lambda}$	2.03	-2.45	-1.35		0.22		0.21
		t-stat(sc)	(1.62)	(-0.50)	(-2.28**)		(0.15)		
	7	$\hat{\lambda}$	4.45	-6.51	-1.05			-2.18	0.44
		t-stat(sc)	(2.80***)	(-1.25)	(-1.81*)			(-1.30)	

Table VIII
Fama-Macbeth regressions using 25 Fama-French Portfolios (Random assignment)

The table reports the estimation of the market prices of risk, using the Fama and MacBeth (1973) two-pass regressions for December 1981 to November 2012, overlapping 372 quarters in monthly frequency on 25 Fama-French Portfolios as test assets. For each period, households are assigned to randomly selected group of 10 instead of being grouped based on the consumption growth rate(see the details of the grouping procedure in Sub-section V.A). The table reports the average of 100 simulated results Two t-statistics with the Newey and West (1987) weight allowing for autocorrelation up to 2 months for each coefficient estimate are reported in parentheses. The top statistic uses uncorrected Fama and MacBeth (1973) standard errors while bottom statistic uses the Shanken (1992) correction. For each row, R^2 columns reports the unadjusted cross-sectional R^2 statistics and R^2 adjusted for the degrees of freedom in parentheses. The asset-pricing factors are Row 1) Δc the log growth rate of aggregate consumption computed from NIPA table; Row 2) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; Row 3) MKT the excess market return; Row 4) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment; Row 9) Fama and French (1992) 3 Factors where SMB and HML are the Fama and French (1992) mimicking portfolios related to size and book-to-market ratios. * : $p < 0.10$, ** : $p < 0.05$, *** : $p < 0.01$

		Const	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	SMB	HML	R^2
	$\hat{\lambda}$	4.25			-0.34					0.20
1	t-stat(fm)	(4.50***)			(-1.75*)					(0.20)
	t-stat(sc)	(3.20***)			(-1.25)					
	$\hat{\lambda}$	2.43				0.84				0.01
2	t-stat(fm)	(3.33***)				(0.69)				(0.01)
	t-stat(sc)	(3.31***)				(0.74)				
	$\hat{\lambda}$	4.76					-2.22			0.26
3	t-stat(fm)	(3.49***)					(-1.54)			(0.26)
	t-stat(sc)	(3.39***)					(-1.45)			
	$\hat{\lambda}$	2.23	-0.12	0.03						0.21
4	t-stat(fm)	(3.26***)	(-0.07)	(0.41)						(0.21)
	t-stat(sc)	(2.40**)	(-0.07)	(0.32)						
	$\hat{\lambda}$	4.17	-0.03	0.07	-0.36					0.40
5	t-stat(fm)	(5.23***)	(-0.03)	(0.94)	(-2.44**)					(0.39)
	t-stat(sc)	(3.18***)	(-0.03)	(0.61)	(-1.61)					
	$\hat{\lambda}$	2.29	-0.12	0.03		0.83				0.23
6	t-stat(fm)	(3.36***)	(-0.07)	(0.38)		(0.71)				(0.23)
	t-stat(sc)	(2.37**)	(-0.07)	(0.29)		(0.53)				
	$\hat{\lambda}$	5.36	0.13	0.11			-3.06			0.49
7	t-stat(fm)	(4.15***)	(0.10)	(1.58)			(-2.19**)			(0.49)
	t-stat(sc)	(3.33***)	(0.07)	(1.20)			(-1.78*)			
	$\hat{\lambda}$	5.35					-3.38	0.29	1.09	0.62
8	t-stat(fm)	(6.01***)					(-3.08***)	(0.71)	(2.30**)	(0.62)
	t-stat(sc)	(5.60***)					(-3.03***)	(0.69)	(2.25**)	
	$\hat{\lambda}$	5.45	0.15	0.07			-3.48	0.31	1.02	0.69
9	t-stat(fm)	(6.13***)	(0.16)	(1.43)			(-3.13***)	(0.76)	(2.16**)	(0.68)
	t-stat(sc)	(4.58***)	(0.10)	(1.05)			(-2.58**)	(0.73)	(2.08**)	

Table IX
Fama-Macbeth regressions using 25 Fama-French Portfolio (Stockholder only)

The table reports the estimation of the market prices of risk, using the Fama and MacBeth (1973) two-pass regressions for December 1981 to November 2012, overlapping 372 quarters in monthly frequency on 25 Fama-French Portfolios as test assets. In addition to the sample filter described in Appendix C.A, the CEX data is further filtered to include only stockholders. Stockholders are identified by the method suggested by Malloy, Moskowitz, and Vissing-Jørgensen (2009). Two GMM T-statistics with the Newey and West (1987) weight allowing for autocorrelation up to 2 months for each coefficient estimate are reported in parentheses. The top statistic uses uncorrected Fama-MacBeth standard errors while bottom statistic uses the Shanken (1992) correction. For each row, R^2 columns reports the unadjusted cross-sectional R^2 statistics and R^2 adjusted for the degrees of freedom in parentheses. The asset-pricing factors are Row 1) Δc the log growth rate of aggregate consumption computed from NIPA table; Row 2) Δc (sample) the log growth rate of aggregate consumption computed from CEX data; Row 3) Mkt the excess market return; Row 4) F_z the proxy for extraneous risk and CWS the empirical proxy for the consumption-weighted sentiment; Row 9) Fama and French (1992) 3 Factors, where SMB and HML are the Fama and French (1992) mimicking portfolios related to size and book-to-market ratios.

		Const	F_z	$CWS \cdot F_z$	Δc	Δc(CEX)	MKT	SMB	HML	R^2
	$\hat{\lambda}$	4.25			-0.34					0.20
1	t-stat(fm)	(4.50***)			(-1.75*)					(0.20)
	t-stat(sc)	(3.20***)			(-1.25)					
	$\hat{\lambda}$	2.52				8.33				0.16
2	t-stat(fm)	(3.49***)				(3.01***)				(0.16)
	t-stat(sc)	(2.53**)				(2.53**)				
	$\hat{\lambda}$	4.76					-2.22			0.26
3	t-stat(fm)	(3.49***)					(-1.54)			(0.26)
	t-stat(sc)	(3.39***)					(-1.45)			
	$\hat{\lambda}$	1.80	6.79	-2.58						0.64
4	t-stat(fm)	(2.48**)	(0.90)	(-3.66***)						(0.64)
	t-stat(sc)	(1.29)	(0.56)	(-2.32**)						
	$\hat{\lambda}$	2.89	0.42	-2.34	-0.19					0.69
5	t-stat(fm)	(3.15***)	(0.07)	(-3.44***)	(-0.98)					(0.69)
	t-stat(sc)	(1.67*)	(0.05)	(-2.04**)	(-0.56)					
	$\hat{\lambda}$	1.98	10.21	-2.43		6.47				0.73
6	t-stat(fm)	(2.81***)	(1.22)	(-3.55***)		(2.36**)				(0.73)
	t-stat(sc)	(1.40)	(0.75)	(-2.21**)		(1.69*)				
	$\hat{\lambda}$	3.28	5.00	-2.30			-1.32			0.72
7	t-stat(fm)	(2.43**)	(0.75)	(-3.42***)			(-0.91)			(0.72)
	t-stat(sc)	(1.53)	(0.51)	(-2.21**)			(-0.61)			
	$\hat{\lambda}$	5.35					-3.38	0.29	1.09	0.62
8	t-stat(fm)	(6.01***)					(-3.08***)	(0.71)	(2.30**)	(0.62)
	t-stat(sc)	(5.60***)					(-3.03***)	(0.69)	(2.25**)	
	$\hat{\lambda}$	4.45	7.79	-1.90			-2.49	0.24	1.06	0.74
9	t-stat(fm)	(4.37***)	(1.18)	(-4.06***)			(-2.11**)	(0.59)	(2.23**)	(0.74)
	t-stat(sc)	(3.27***)	(0.98)	(-3.48***)			(-1.79*)	(0.57)	(2.08**)	

REFERENCES

Attanasio, Orazio P., 1999, Chapter 11 Consumption, in *Handbook of Macroeconomics*, volume 1, 741–812 (Elsevier).

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *The Journal of Finance* 59, 1481–1509.

Basak, Suleyman, 2000, A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk, *Journal of Economic Dynamics and Control* 24, 63–95.

Beeler, Jason, and John Y. Campbell, 2009, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, Working Paper 14788, National Bureau of Economic Research.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy, 2002, Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence, *Journal of Political Economy* 110, 793–824.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of financial Economics* 7, 265–296.

Buraschi, Andrea, and Alexei Jiltsov, 2006, Model Uncertainty and Option Markets with Heterogeneous Beliefs, *The Journal of Finance* 61, 2841–2897.

Buraschi, Andrea, Fabio Trojani, and Andrea Vedolin, 2014, When Uncertainty Blows in the Orchard: Comovement and Equilibrium Volatility Risk Premia, *The Journal of Finance* 69, 101–137.

Buraschi, Andrea, and Paul Whelan, 2012, Term structure models with differences in beliefs, *Imperial College Business School Working Paper* .

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, 2009, Measuring the Financial Sophistication of Households, Working Paper 14699, National Bureau of Economic Research.

Campbell, John Y., 2003, Chapter 13 Consumption-based asset pricing, *Handbook of the Economics of Finance* 1, 803–887.

Cochrane, John H., 1996, A Cross-Sectional Test of an Investment-Based Asset Pricing Model, *Journal of Political Economy* 104, 572–621.

Cochrane, John H., 2009, *Asset Pricing: (Revised Edition)* (Princeton University Press), Google-Books-ID: 20pmeMaKNwsC.

Coibion, Olivier, Yuriy Gorodnichenko, and Dmitri Koustas, 2017, Consumption Inequality and the Frequency of Purchases, Working Paper 23357, National Bureau of Economic Research.

Constantinides, George M., 1982, Intertemporal Asset Pricing with Heterogeneous Consumers and Without Demand Aggregation, *The Journal of Business* 55, 253–267.

Constantinides, George M., and Anisha Ghosh, 2017, Asset Pricing with Countercyclical Household Consumption Risk, *The Journal of Finance* 72, 415–460.

Cox, John C, and Chi-fu Huang, 1989, Optimal consumption and portfolio policies when asset prices follow a diffusion process, *Journal of Economic Theory* 49, 33–83.

Cvitanić, Jaksa, and Semyon Malamud, 2011, Price impact and portfolio impact, *Journal of Financial Economics* 100, 201–225.

David, Alexander, 2008, Heterogeneous Beliefs, Speculation, and the Equity Premium, *The Journal of Finance* 63, 41–83.

Detemple, Jerome, and Shashidhar Murthy, 1994, Intertemporal Asset Pricing with Heterogeneous Beliefs, *Journal of Economic Theory* 62, 294–320.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of Opinion and the Cross Section of Stock Returns, *The Journal of Finance* 57, 2113–2141.

Dumas, Bernard, 1989, Two-Person Dynamic Equilibrium in the Capital Market, *The Review of Financial Studies* 2, 157–188.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal, 2009, Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility, *The Journal of Finance* 64, 579–629.

Ehling, Paul, Michael F. Gallmeyer, Christian Heyerdahl-Larsen, and Philipp K. Illeditsch, 2016, Disagreement about Inflation and the Yield Curve, SSRN Scholarly Paper ID 2028391, Social Science Research Network, Rochester, NY.

Engsted, Tom, and Stig V. Møller, 2015, Cross-sectional consumption-based asset pricing: A reappraisal, *Economics Letters* 132, 101–104.

Fama, Eugene F., and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607–636.

Ferson, Wayne E., and Campbell R. Harvey, 1999, Conditioning Variables and the Cross Section of Stock Returns, *The Journal of Finance* 54, 1325–1360.

Gallmeyer, Michael, and Burton Hollifield, 2008, An Examination of Heterogeneous Beliefs with a Short-Sale Constraint in a Dynamic Economy, *Review of Finance* 12, 323–364.

Gao, George, Xiaomeng Lu, Zhaogang Song, and Hongjun Yan, 2017, Disagreement Beta, Ssrn scholarly paper, Social Science Research Network, Rochester, NY.

Grossman, Sanford J., and Robert J. Shiller, 1981, The Determinants of the Variability of Stock Market Prices, *The American Economic Review* 71, 222–227.

Harris, Milton, and Artur Raviv, 1993, Differences of Opinion Make a Horse Race, *The Review of Financial Studies* 6, 473–506.

Harrison, J. Michael, and David M. Kreps, 1978, Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations, *The Quarterly Journal of Economics* 92, 323–336.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, … and the Cross-Section of Expected Returns, *The Review of Financial Studies* 29, 5–68.

Huang, Chi-fu, and Robert H. Litzenberger, 1988, *Foundations for Financial Economics* (Prentice Hall), Google-Books-ID: HoSxAAAAIAAJ.

Jouini, Elyès, and Clotilde Napp, 2007, Consensus Consumer and Intertemporal Asset Pricing with Heterogeneous Beliefs, *The Review of Economic Studies* 74, 1149–1174.

Karatzas, I., J. Lehoczky, and S. Shreve, 1987, Optimal Portfolio and Consumption Decisions for a "Small Investor" on a Finite Horizon, *SIAM Journal on Control and Optimization* 25, 1557–1586.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, *Journal of Political Economy* 109, 1238–1287.

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

Li, Frank Weikai, 2016, Macro Disagreement and the Cross-Section of Stock Returns, *The Review of Asset Pricing Studies* 6, 1–45.

Lucas, Robert E., 1978, Asset Prices in an Exchange Economy, *Econometrica* 46, 1429–1445.

Ludvigson, Sydney C., 2013, Advances in Consumption-Based Asset Pricing: Empirical Tests, *Handbook of the Economics of Finance* 2, 799–906.

Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jørgensen, 2009, Long-Run Stockholder Consumption Risk and Asset Returns, *The Journal of Finance* 64, 2427–2479.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*, *The Quarterly Journal of Economics* 126, 373–416.

Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867–887.

Miller, Edward M., 1977, Risk, Uncertainty, and Divergence of Opinion, *The Journal of Finance* 32, 1151.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Parker, Jonathan A., and Christian Julliard, 2005, Consumption Risk and the Cross Section of Expected Returns, *Journal of Political Economy* 113, 185–222.

Rubinstein, Mark, 1976, The Valuation of Uncertain Income Streams and the Pricing of Options, *The Bell Journal of Economics* 7, 407–425.

Santos, Tano, and Pietro Veronesi, 2006, Labor Income and Predictable Stock Returns, *Review of Financial Studies* 19, 1–44.

Scheinkman, José A., and Wei Xiong, 2003, Overconfidence and Speculative Bubbles, *Journal of Political Economy* 111, 1183–1220.

Shanken, Jay, 1985, Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics* 14, 327–348.

Shanken, Jay, 1992, On the Estimation of Beta-Pricing Models, *The Review of Financial Studies* 5, 1–33.

Varian, Hal R., 1985, Divergence of Opinion in Complete Markets: A Note, *The Journal of Finance* 40, 309–317.

Wang, Jiang, 1996, The Term Structure of Interest Rates in a Pure Exchange Economy with Heterogeneous Investors, *Journal of Financial Economics* 41, 75–110.

Xiong, Wei, 2013, Bubbles, Crises, and Heterogeneous Beliefs, Working Paper 18905, National Bureau of Economic Research.

Xiouros, Costas, and Fernando Zapatero, 2010, The Representative Agent of an Economy with External Habit Formation and Heterogeneous Risk Aversion, *The Review of Financial Studies* 23, 3017–3047.

Yan, Hongjun, 2008, Natural Selection in Financial Markets: Does It Work?, *Management Science* 54, 1935–1950.

Zapatero, Fernando, 1998, Effects of Financial Innovations on Market Volatility When Beliefs Are Heterogeneous, *Journal of Economic Dynamics and Control* 22, 597–626.

Appendix A. Proofs

Proof of Proposition 1. Using the martingale technique of Cox and Huang (1989) and Karatzas, Lehoczky, and Shreve (1987), each group's dynamic optimization problem can be rewritten as static one at time zero:

$$\max_{C_i} E_0^i \left[\int_0^\infty u(C_{i,t})\, dt \right] \tag{A1}$$

subject to

$$E_0^i \left[\int_{t=0}^\infty \xi_{i,t} C_{i,t}\, dt \right] \le \xi_{i,0} W_i \tag{A2}$$

where $u(C_{i,t}) = e^{-\rho t} \frac{C_{i,t}^{1-\gamma}}{1-\gamma}$, $\xi_{i,t}$ is the state price density of group i, and W_i is the initial wealth endowed to group i. Note that since the economy is dynamically complete, there exists a unique stochastic discount factor $\xi_{i,t}$ for each group.

The first order condition of (A1) implies that group i's optimal consumption needs to satisfy:

$$u^{'}(C_{i,t}) = y_i \xi_{i,t}, \quad i = 1, 2, \cdots, K \tag{A3}$$

where $u^{'}(C_{i,t})$ is the marginal utility and the constant $y_i > 0$ is the Lagrange multiplier associated with the budget constraint (A2). y_i is determined by substituting the optimal policy into the static budget constraint (A2):

$$E_0^i \left[\int_{t=0}^\infty \xi_{i,t} (y_i e^{\rho t} \xi_{i,t})^{-1/\gamma}\, dt \right] = \xi_{i,0} W_i, \quad i = 1, 2, \cdots, K. \tag{A4}$$

The equilibrium allocation in the economy with heterogeneous beliefs is characterized as the solution of the optimization problem of a social planner (see Constantinides (1982); Dumas (1989); Detemple and Murthy (1994); Wang (1996); Basak (2000); Gallmeyer and Hollifield (2008)), whose utility function is a weighted sum of the utilities of individual agents in the economy [10] : that is, the social planner solves:

$$U(C_t; \boldsymbol{\lambda_t}) := \max_{\{C_{i,t}\}_{i=1}^K} \sum_{i=1}^K (\lambda_{i,t} u(C_{i,t})) \tag{A5}$$

subject to

$$\sum_{i=1}^K C_{i,t} = C_t \tag{A6}$$

where $\boldsymbol{\lambda_t} \equiv (\lambda_{1,t}, \cdots, \lambda_{K,t})$, and $\lambda_{i,t} > 0$ is the utility weight for investor i which may be stochastic. Without loss of generality, I set $\lambda_{1,t} = 1$. The first order conditions for the social planner's problem are

$$u^{'}(C_{1,t}) = \varsigma = \lambda_{i,t} u^{'}(C_{i,t}), \quad i = 2, \cdots, K \tag{A7}$$

[10]For economies with homogeneous beliefs, see Huang and Litzenberger (1988)[Chapter 5]

where ς is the Lagrangian multiplier associated with the budget constraint (A6).

Equation (A7) together with the first order condition of the individual optimization problem Equation (A3) implies that $\lambda_{i,t}$ are proportional to the ratios of the state price densities of the two groups:

$$\lambda_{i,t} = \frac{u^{'}(C_{1,t})}{u^{'}(C_{i,t})} = \frac{y_1 \xi_{1,t}}{y_i \xi_{i,t}} = \frac{y_1}{y_i}\eta_{i,t}, \quad i = 2, \cdots, K. \tag{A8}$$

The last equality in the above equation holds because in complete market setting, the ratio of the group i's state price density relative to the first group is the Radon-Nikodym derivative $\eta_{i,t} = \frac{d\mathbb{P}_i}{d\mathbb{P}_1}$. Equation (A8) and the market clearing condition Equation (A6) imply that the marginal utility of the social planner is equal to

$$U^{'}(C_t; \boldsymbol{\lambda_t}) = u^{'}(C_{1,t})\frac{dC_{1,t}}{C_t} + \sum_{i=2}^{K} \lambda_{i,t} u^{'}(C_{i,t})\frac{dC_{i,t}}{C_t} = u^{'}(C_{1,t}). \tag{A9}$$

By Equation (A3) and the above equation, the optimal consumption of the first group can be written as

$$C_{1,t} = \left(y_1 \xi_{1,t} e^{\rho t}\right)^{-1/\gamma} = \left(e^{\rho t} U^{'}(C_t; \boldsymbol{\lambda_t})\right)^{-1/\gamma} \tag{A10}$$

Solving the first equality of Equation (A8) for $C_{i,t}$ yields $C_{i,t} = C_{1,t}\lambda_{i,t}^{1/\gamma}$. Thus, the optimal consumption of the group i

$$C_{i,t} = \left(y_i \xi_{i,t} e^{\rho t}\right)^{-1/\gamma} = \left(\frac{e^{\rho t} U^{'}(C_t; \boldsymbol{\lambda_t})}{\lambda_{i,t}}\right)^{-1/\gamma} \tag{A11}$$

By market clearing condition Equation (A6) and the optimal consumption above, we obtain the optimal consumption of the group-1:

$$C_t = \left(e^{\rho t} U(C_t; \boldsymbol{\lambda_t})\right)^{-1/\gamma}\left(1 + \sum_{i=2}^{K} \lambda_{i,t}^{1/\gamma}\right) = C_{1,t}\left(1 + \sum_{i=2}^{K} \lambda_{i,t}^{1/\gamma}\right) \to C_{1,t} = \frac{1}{1 + \sum_{i=2}^{K} \lambda_{i,t}^{1/\gamma}} C_t. \tag{A12}$$

Substituting $C_{1,t}$ in $C_{i,t} = C_{1,t}\lambda_{i,t}^{1/\gamma}$ leads to the optimal consumption of the rest of the groups

$$C_{i,t} = \frac{\lambda_{i,t}^{1/\gamma}}{\sum_{i=1}^{K} \lambda_{i,t}^{1/\gamma}} C_t, \quad i = 2, \cdots, K. \tag{A13}$$

Denote the consumption share by $\omega_{i,t} \equiv \frac{C_{i,t}}{C_t}$. Dividing the above equation by C_t, we get

$$\omega_{i,t} = \frac{\lambda_{i,t}^{1/\gamma}}{\sum_{i=1}^{K} \lambda_{i,t}^{1/\gamma}}, \quad i = 1, \cdots, K \tag{A14}$$

Here $\lambda_{1,t} = 1$ and the first group's consumption share can be obtain by replacing $\lambda_{1,t}$ with 1. Substituting the optimal consumption into the first-order condition Equation (A3) yields the group i's state price density

$$\xi_{i,t} = \frac{1}{y_i} e^{-\rho t} (C_{i,t})^{-\gamma} = e^{-\rho t} C_t^{-\gamma} \frac{1}{y_i} \left(\frac{\lambda_{i,t}^{1/\gamma}}{\sum_{i=1}^{K} \lambda_{i,t}^{1/\gamma}} \right)^{-\gamma}, \quad i = 1, 2, \cdots, K. \tag{A15}$$

Given Equation (A15), the budget constraints Equation (A4) determine only the ratio y_1/y_i. To make $\xi_{i,0} = 1$, I set $y_1 = U'(C_0; \boldsymbol{\lambda_0})$. $\lambda_{i,0}$ solves the budget constraints Equation (A4) for group i. □

Proof of Proposition 2. Using $h_t(\boldsymbol{\lambda_t}) \equiv \left(1 + \sum_{i=2}^{K} \lambda_{i,t}^{1/\gamma}\right)^{\gamma}$ in Equation (A15), the state price density of the first group can be written as

$$\xi_{1,t} = \frac{1}{y_1} e^{-\rho t} C_t^{-\gamma} h_t(\boldsymbol{\lambda_t}). \tag{A16}$$

The dynamics of h_t follows

$$\frac{dh_t}{h_t} = \mu_{h,t} dt + \sigma_{h,t} dB_{z,t} \tag{A17}$$

where $\mu_{h,t}$ and $\sigma_{h,t}$ are to be determined. Using the definition of h_t, the consumption share Equation (A14) can be written

$$\omega_{i,t} = \lambda_{i,t}^{1/\gamma} h_t(\boldsymbol{\lambda_t})^{-1/\gamma}. \tag{A18}$$

From the dynamics of $\eta_{i,t}$

$$\frac{d\lambda_{i,t}}{\lambda_{i,t}} = \delta_i dB_{z,t}. \tag{A19}$$

Applying Itô 's lemma on $\omega_{i,t}$, the dynamics of consumption share process follows

$$\frac{d\omega_{i,t}}{\omega_{i,t}} = \mathcal{A}_{i,t} dt + \mathcal{B}_{i,t} dB_{z,t} \tag{A20}$$

where

$$\begin{aligned} \mathcal{A}_{i,t} &= -\frac{1}{\gamma} \mu_{h,t} - \frac{1}{\gamma^2} \delta_i \sigma_{h,t} + \frac{1-\gamma}{2\gamma^2} \delta_i^2 + \frac{1+\gamma}{2\gamma^2} \sigma_{h,t}^2 \\ \mathcal{B}_{i,t} &= \frac{1}{\gamma} \delta_i - \frac{1}{\gamma} \sigma_{h,t}. \end{aligned}$$

By construction, the consumption shares $\omega_{i,t}$ sum to 1, i.e. $\sum_{i=1}^{K} \omega_{i,t} = 1$ and thus $\sum_{i=1}^{K} d\omega_{i,t} = 0$.

Thus, $\mu_{h,t}$ and $\sigma_{h,t}$, in equilibrium, are determined such that the following relationships hold

$$\sum_{i=1}^{K}\left(\mathcal{B}_{i,t}\omega_{i,t}\right)=0\longrightarrow\sigma_{h,t}=\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right) \tag{A21}$$

$$\sum_{i=1}^{K}\left(\mathcal{A}_{i,t}\omega_{i,t}\right)=0\longrightarrow\mu_{h,t}=\frac{1-\gamma}{2\gamma}\sum_{i=1}^{K}\left(\delta_i^2\omega_{i,t}\right)-\frac{1}{\gamma}\sigma_{h,t}\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right)+\frac{1+\gamma}{2\gamma}\sigma_{h,t}^2 \tag{A22}$$

$$=\frac{1-\gamma}{2\gamma}\left[\sum_{i=1}^{K}\left(\delta_i^2\omega_{i,t}\right)-\left(\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right)\right)^2\right]. \tag{A23}$$

Given the dynamics of h_t and C_t, we obtain the dynamics of $\xi_{1,t}$

$$\frac{d\xi_{1,t}}{\xi_{1,t}}=-r_t dt-\theta_c dB_{c,t}-\theta_{z,t}dB_{z,t} \tag{A24}$$

where

$$r_t=\rho+\gamma g_c-\frac{1}{2}\gamma(\gamma+1)\sigma_c^2+\frac{\gamma-1}{2\gamma}\left[\sum_{i=1}^{K}\left(\delta_i^2\omega_{i,t}\right)-\left(\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right)\right)^2\right] \tag{A25}$$

$$\theta_c=\gamma\sigma_c \tag{A26}$$

and

$$\theta_{z,t}=-\sigma_{h,t}=-\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right). \tag{A27}$$

Substituting $\mu_{h,t}$ and $\sigma_{h,t}$ into the log consumption share process,

$$d\log\omega_{i,t}=\left(\mathcal{A}_{i,t}-\frac{1}{2}\mathcal{B}_{i,t}^2\right)dt+\mathcal{B}_{i,t}dB_{z,t} \tag{A28}$$

where

$$\begin{aligned}\mathcal{A}_{i,t}-\frac{1}{2}\mathcal{B}_{i,t}^2&=-\frac{1}{\gamma}\mu_{h,t}-\frac{1}{2\gamma}\delta_i^2+\frac{1}{2\gamma}\sigma_{h,t}^2\\&=\frac{1}{\gamma}\left[\frac{\gamma-1}{2\gamma}\left[\sum_{i=1}^{K}\left(\delta_i^2\omega_{i,t}\right)-\left(\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right)\right)^2\right]-\frac{1}{2}\delta_i^2+\frac{1}{2}\left(\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right)\right)^2\right]\\\mathcal{B}_{i,t}&=\frac{1}{\gamma}\delta_i-\frac{1}{\gamma}\sigma_{h,t}=\frac{1}{\gamma}\left[\delta_i-\sum_{i=1}^{K}\left(\delta_i\omega_{i,t}\right)\right].\end{aligned}$$

By the definition $d\log C_{i,t}=d\log C_t+d\log\omega_{i,t}$, it follows from the above equation and Equation (1) that

$$d\log C_{i,t}=\left(g_c-\frac{1}{2}\sigma_c^2+\mathcal{A}_{i,t}-\frac{1}{2}\mathcal{B}_{i,t}^2\right)dt+\sigma_c dB_{c,t}+\mathcal{B}_{i,t}dB_{z,t} \tag{A29}$$

Substituting $r_{f,t}$ and $\theta_{z,t}$, we get

$$d\log C_{i,t} = \frac{1}{\gamma}\left(r_{f,t} - \rho - \frac{1}{2}\gamma^2\sigma_c^2 - \frac{1}{2}\delta_i^2 + \frac{1}{2}\theta_{z,t}^2\right)dt + \sigma_c dB_{c,t} + \frac{1}{\gamma}\left(\delta_i + \theta_{z,t}\right)dB_{z,t}. \tag{A30}$$

□

Appendix B. The Linear Factor Model

Conditional beta representation

I start by the basic asset pricing equation:

$$E_t\left[\frac{dP_t^j}{P_t^j} + \frac{D_t^j}{P_t^j}dt - r_{f,t}dt\right] = -E_t\left[\frac{dP_t^j}{P_t^j}\frac{d\xi_t}{\xi_t}\right] \tag{B1}$$

Replacing $d\log C_t$ from Equation (1), Equation (14) can be written

$$\frac{d\xi_t}{\xi_t} = -\left(r_{f,t} - g_c + \frac{1}{2}\sigma_c^2\right)dt - \gamma dc_t - \theta_{z,t}dB_{z,t}. \tag{B2}$$

Substituting the above in Equation (B1), we get

$$E_t\left[\frac{dP_t^j}{P_t^j} + \frac{D_t^j}{P_t^j}dt - r_{f,t}dt\right] = \gamma E_t\left[\frac{dP_t^j}{P_t^j}dc_t\right] + \theta_{z,t}E_t\left[\frac{dP_t^j}{P_t^j}dB_{z,t}\right]. \tag{B3}$$

For short time period, it can be approximated to

$$E_t\left[R_{j,t+1}^e\right] \equiv E_t\left[R_{t+1}^j\right] - R_{f,t} \approx \gamma Cov_t\left(R_{t+1}^j, \Delta c_{t+1}\right) + \theta_{z,t}Cov_t\left(R_{t+1}^j, F_{z,t+1}\right). \tag{B4}$$

Substituting $\theta_{z,t} = h_0 + h_1\text{CWS}_t$, we get

$$E_t\left[R_{j,t+1}^e\right] \approx \gamma Cov_t\left(R_{t+1}^j, \Delta c_{t+1}\right) + (h_0 + h_1\text{CWS}_t)Cov_t\left(R_{t+1}^j, F_{z,t+1}\right). \tag{B5}$$

Dividing and multiplying by the volatility of the risk factor both sides of Equation (B5) yields

$$E_t\left[R_{j,t+1}^e\right] \approx \gamma\frac{Cov_t\left(R_{t+1}^j, \Delta c_{t+1}\right)}{Var_t\left(\Delta c_{t+1}\right)}Var_t\left(\Delta c_{t+1}\right) + (h_0 + h_1\text{CWS}_t)\frac{Cov_t\left(R_{t+1}^j, F_{z,t+1}\right)}{Var_t\left(F_{z,t+1}\right)}Var_t\left(F_{z,t+1}\right).$$

Because $Var_t\left(\Delta c_{t+1}\right) = \sigma_c^2$ and $Var_t\left(F_{z,t+1}\right) = 1$ by assumption,

$$E_t\left[R_{j,t+1}^e\right] \approx \gamma\frac{Cov_t\left(R_{t+1}^j, \Delta c_{t+1}\right)}{\sigma_c^2}\sigma_c^2 + (h_0 + h_1\text{CWS}_t)\frac{Cov_t\left(R_{t+1}^j, F_{z,t+1}\right)}{Var_t\left(F_{z,t+1}\right)}Var_t\left(F_{z,t+1}\right).$$

Unconditional beta representation

Taking unconditional expectation of Equation (B4) yields

$$E\left[R^j_{t+1} - R_{f,t+1}\right] = \gamma E\left[Cov_t\left(R^j_{t+1}, \Delta c_{t+1}\right)\right] + E\left[Cov_t\left(R^j_{t+1}, \theta_{z,t}F_{z,t+1}\right)\right]. \tag{B6}$$

Applying the law of total covariance $Cov(X,Y) = E\left[Cov_t(X,Y)\right] + Cov(E_t[X], E_t[Y])$ and using the fact that $E_t[\Delta c_{t+1}] = \text{constant}$, we get

$$\begin{aligned} E\left[Cov_t\left(R^j_{t+1}, \Delta c_{t+1}\right)\right] &= Cov\left(R^j_{t+1}, \Delta \log C_{t+1}\right) - Cov\left(E_t[R^j_{t+1}], E_t[\Delta \log C_{t+1}]\right) \\ &= Cov\left(R^j_{t+1}, \Delta c_{t+1}\right). \end{aligned}$$

Because $E_t[F_{z,t+1}] = 0$ by the assumption, the law of total covariance implies that

$$\begin{aligned} E\left[Cov_t\left(R^j_{t+1}, \theta_{z,t}F_{z,t+1}\right)\right] &= Cov\left(R^j_{t+1}, \theta_{z,t}F_{z,t+1}\right) - Cov\left(E_t[R^j_{t+1}], \theta_{z,t}E_t[F_{z,t+1}]\right) \\ &= Cov\left(R^j_{t+1}, \theta_{z,t}F_{z,t+1}\right). \end{aligned}$$

Substituting the results above in Equation (B6) yields

$$E_t\left[R^e_{j,t+1}\right] = \gamma Cov\left(R^j_{t+1}, \Delta c_{t+1}\right) + Cov\left(R^j_{t+1}, \theta_{z,t}F_{z,t+1}\right). \tag{B7}$$

Replacing $\theta_{z,t}$ with CWS_t,

$$E_t\left[R^e_{j,t+1}\right] = \gamma Cov\left(R^j_{t+1}, \Delta c_{t+1}\right) + h_0 Cov\left(R^j_{t+1}, F_{z,t+1}\right) + h_1 Cov\left(R^j_{t+1}, \text{CWS}_t F_{z,t+1}\right). \tag{B8}$$

Appendix C. Empirical Specification

Appendix A. CEX sample filter criteria

I follow the standard exclusion in Malloy et al. (2009) and for completeness, I include the sample selection criteria as follows. First eliminated from the sample are the households that reside in nonurban areas or student housing, have incomplete income responses, or change in family composition over the course of their interviews. I drop the surveys conducted in 1980 and 1981 because the change in the questions concerning food consumption in 1982 makes the food consumption data before 1980 not comparable to the data afterward. To compute the quarterly consumption correctly, the households that report more or less than three-month consumptions or negative consumption in a given quarter are dropped. Households with extreme consumption growth values are also eliminated, in particular, those with the consumption growth ratio less than 0.2 or above 5.0. Computing the consumption growth ratio of individual household requires matching the households across interview quarters. For this reason, I keep in the sample only the households that finish all four interviews.

Appendix B. Group Construction

For a short time period, Equation (18) implies that $\Delta c_{h,t+1}$ is approximated as the following linear model:

$$\Delta c_{h,t+1} \approx \Delta c_{t+1} + A_{0,t} + A_{1,t} F_{z,t+1} - \frac{1}{2\gamma}\delta_h^2 + \frac{1}{\gamma}\delta_h F_{z,t+1} \tag{C1}$$

where $A_{0,t}$ and $A_{1,t}$ collect the terms that are common across the households; $F_{z,t} \equiv B_{z,t+1} - B_{z,t} \sim$ *i.i.d.* $N(0,1)$ represent the shock to Z_{t+1}; and $\Delta c_{t+1} \equiv c_{t+1} - c_t$ denotes the rate of log aggregate consumption growth. For each $t = 1, \cdots, T$, the cross-sectional average of log consumption growth rate is defined as :

$$\overline{\Delta c_{h,t+1}} \equiv \frac{1}{N_{t+1}} \sum_{h=1}^{N_{t+1}} \Delta c_{h,t+1} \tag{C2}$$

where $\Delta c_{h,t}$ is the log consumption growth rate of household, indexed by $h = 1, \cdots, N_{t+1}$ and N_{t+1} is the number of households observed at time $t+1$. The de-meaned $\Delta c_{h,t+1}$ thus follows

$$\Delta \tilde{c}_{h,t+1} \equiv \Delta c_{h,t+1} - \overline{\Delta c_{h,t+1}} \approx -\frac{1}{2\gamma}\left[\delta_h^2 - \frac{1}{N_{t+1}}\sum_{h=1}^{N_{t+1}} \delta_k^2\right] + \frac{1}{\gamma}\left[\delta_h - \frac{1}{N_{t+1}}\sum_{h=1}^{N_{t+1}} \delta_k\right] F_{z,t+1} \tag{C3}$$

Given the household-level

$$\Delta \tilde{c}_{j,t} \approx -\frac{1}{2\gamma}\left[\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}} \delta_h^2 - \frac{1}{N_t}\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}}\sum_{i=1}^{N_t} \delta_i^2\right] + \frac{1}{\gamma}\left[\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}} \delta_h - \frac{1}{N_t}\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}}\sum_{i=1}^{N_t} \delta_i\right] F_{z,t} \tag{C4}$$

$$\approx -\frac{1}{2\gamma}\left[\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}} \delta_h^2 - \frac{1}{N_t}\sum_{i=1}^{N_t} \delta_i^2\right] + \frac{1}{\gamma}\left[\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}} \delta_h - \frac{1}{N_t}\sum_{i=1}^{N_t} \delta_i\right] F_{z,t}. \tag{C5}$$

Let $\mu^j(\delta)$ denote by $\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}} \delta_h$, $\delta_{2,j}$ denote by $\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}} \delta_h^2$, and $\sigma^j(\delta)$ denote by $\frac{1}{N_{j,t}}\sum_{h=1}^{N_{j,t}}(\delta_h - \mu^j(\delta))^2$.

$$\Delta \tilde{c}_{j,t} \approx -\frac{1}{2\gamma}\left[\delta_{2,j} - \frac{1}{N_t}\sum_{j=1}^{m} N_{j,t}\delta_{2,j}\right] + \frac{1}{\gamma}\left[\delta_j - \frac{1}{N_t}\sum_{j=1}^{m} N_{j,t}\delta_j\right] F_{z,t} \tag{C6}$$

$$\approx -\frac{1}{2\gamma}\left[\delta_{2,j} - \frac{1}{m}\sum_{j=1}^{m} \delta_{2,j}\right] + \frac{1}{\gamma}\left[\delta_j - \frac{1}{m}\sum_{j=1}^{m} \delta_j\right] F_{z,t}. \tag{C7}$$

$\delta_{2,j}$ is related to δ_j by the following:

$$\delta_{2,j} = \frac{1}{N_{j,t}} \sum_{h=1}^{N_{j,t}} \delta_h^2 = \delta_j^2 + \frac{1}{N_{j,t}} \sum_{h=1}^{N_{j,t}} (\delta_h^2 - \delta_j)^2. \tag{C8}$$

$$\delta_{2,j} - \frac{1}{m} \sum_{j=1}^{m} \delta_{2,j} = \delta_j^2 - \frac{1}{m} \sum_{j=1}^{m} \delta_j^2 + \frac{1}{N_{j,t}} \sum_{h=1}^{N_{j,t}} (\delta_h^2 - \delta_j)^2 - \frac{1}{m} \sum_{j=1}^{m} \frac{1}{N_{j,t}} \sum_{h=1}^{N_{j,t}} (\delta_h^2 - \delta_j)^2 \tag{C9}$$

$$. \tag{C10}$$

Appendix C. Extracting factors and subjective beliefs

Denote the vector of $\Delta\tilde{c}_{i,t}$ in Equation (33) with $X_t \equiv [\Delta\tilde{c}_{1,t}, \cdots, \Delta\tilde{c}_{m,t}]^T$. X_t can be written as

$$X_t = \boldsymbol{a} + \boldsymbol{b} F_{z,t} + \boldsymbol{\varepsilon}_t \tag{C11}$$

where $\boldsymbol{a} \equiv [a_1, \cdots, a_m]^T$, $\boldsymbol{b} \equiv [b_1, \cdots, b_m]^T$ with each element $b_i = \frac{1}{\gamma}\left[\delta_i - \frac{1}{m}\sum_{j=1}^{m} \delta_j\right]$ and $\boldsymbol{\varepsilon}_t \equiv [\varepsilon_{i,t}, \cdots, \varepsilon_{m,t}]^T$. I assume $E[\boldsymbol{\varepsilon}_t] = 0$ and $E[\varepsilon_{i,t}\varepsilon_{j,t}] = 0$ for $i \neq j$. Let Γ denote the covariance matrix of $X_t - \boldsymbol{a}$ and consider the eigen-decomposition of the covariance matrix Γ

$$\Gamma = Q \Lambda Q^T \tag{C12}$$

where $QQ^T = I$ and Λ is a diagonal matrix with its i^{th} element λ_i. The principal components are defined as $\mathbf{F}_t \equiv Q^T(X_t - \boldsymbol{a}) \sim N(0, \Lambda)$ a vector of zero-mean, uncorrelated random variables. Because $QQ^T = I$, $X_t - \boldsymbol{a}$ can be written as $X_t - \boldsymbol{a} = QQ^T(X_t - \boldsymbol{a})$. Substituting $\mathbf{F}_t$ yields $X_t - \boldsymbol{a} = Q\mathbf{F}_t$. Let $Q = [q_1, q_2]$ and $\mathbf{F}_t = [f_t^1, f_t^2]'$ where q_1 is the first column of Q and f_t^1 is the first element of $\mathbf{F}_t$. It follows that X_t can be written as

$$X_t - \boldsymbol{a} = QQ^T(X_t - \boldsymbol{a}) = Q\mathbf{F}_t = q_1 f_t^1 + q_2 f_t^2. \tag{C13}$$

Interpreting $\boldsymbol{b} = q_1 \lambda_1$, $F_{z,t+1} = f_t^1/\lambda_1$, and $\varepsilon_t = q_2 f_t^2$, we get the model in Equation (C11). That is, I use the normalized first principal component as a proxy for the non-fundamental risk $F_{z,t+1} = f_t^1/\lambda_1$.

To identify b_i, I run the following OLS regression

$$\Delta\tilde{c}_{i,t+1} = a_i + b_i F_{z,t+1} + \varepsilon_{i,t+1} \tag{C14}$$

and use the point estimate $\hat{b}_i$ as the empirical measure of b_i. That is, the group-level forecasting error can be identified by $\hat{b}_i$:

$$\hat{b}_i = \frac{1}{\gamma}\left[\delta_i - \frac{1}{m}\sum_{j=1}^{m}\delta_j\right]. \tag{C15}$$

Testing the linear factor model in Equation (25) requires an empirical proxy for the consumption weighted forecasting errors. This variable is estimated from $\hat{b}_i$ and $\omega_{i,t}$ from Equation (32). Define CWS as

$$CWS_t \equiv \sum_{i=1}^{m}\left(\hat{b}_i\omega_{i,t}\right) - E\left[\sum_{i=1}^{m}\left(\hat{b}_i\omega_{i,t}\right)\right] \tag{C16}$$

where $E[\cdot]$ denotes the unconditional expectation. Given the definition of CWS_t, Equation (C15), and the definition of the consumption-weighted sentiment $\mathcal{E}_t(\delta_i)$ in Equation (17), it is straightforward to show that $\mathcal{E}_t(\delta_i) - E\left[\mathcal{E}_t(\delta_i)\right]$ is proportional to CWS_t

$$\mathcal{E}_t(\delta_i) - E\left[\mathcal{E}_t(\delta_i)\right] = \gamma CWS_t \quad \text{with } E\left[\mathcal{E}_t(\delta_i)\right] = E\left[\sum_{i=1}^{m}\left(\delta_i\omega_{i,t}\right)\right]. \tag{C17}$$