Evaluating Consumption CAPM under Heterogeneous Preferences *

Min (Berg) Cui †

May 23, 2018

Abstract

We provide the first direct empirical investigation of the Consumption CAPM (CCAPM) with heterogeneous preferences. We construct CCAPM price kernels with both homogeneous and heterogeneous Epstein-Zin-Weil (EZW) preferences under a complete financial market. We estimate preferences parameters using the household level consumption and wealth data from the Panel Study of Income Dynamics (PSID) database, and then calculate excess returns for risky assets using CCAPM pricing kernels with estimated preferences parameters and households' consumption and wealth data. We run a horse race between two homogeneous preference models and three heterogeneous preferences models, and show that heterogeneous preferences models outperform the representative agent model in capturing both more magnitude and more dynamics of both the equity premium and the cross section of stock returns. Lastly, we quantitatively decompose the improvement into three aspects, idiosyncratic risk factors, heterogeneous factor premia and idiosyncratic characteristics dependent aggregation weights.

Keywords: Consumption CAPM, Epstein-Zin-Weil, heterogeneous preferences, heterogeneous market participation, cross section of stock returns
JEL Classification: G12

*The author thanks Joon Y. Park, Yoosoon Chang, Grey Gordon, Todd B. Walk, Gerhard Glomm, Hagan Kim, Noah Stoffman, Siming Liu, Jessie He and participants at Jordan River Conference 2016, MEA2017, MEG2017, EGSC2017 for valuable comments.
†Department of Economics, Indiana University. E-mail: cuim@indiana.edu.

1 Introduction

In this study, we use the Epstein-Zin-Weil (EZW) utility to construct Consumption CAPM (CCAPM) pricing kernels with time-varying heterogeneous preferences. We prove, except under some special cases, no aggregation result can be achieved. Hence, every household matters in the asset pricing equation. We then estimate both homogeneous and heterogeneous EZW preferences parameters using the Panel Study of Income Dynamics (PSID) data under the complete market assumption. We show that with the heterogeneity, we can get relatively large Elasticity of Inter-temporal Substitution (EIS) and small Relative Risk Aversion(RRA) estimates that are consistent with empirical microeconomic findings. With data and estimated parameters, we calculate predictions of excess returns for risky assets. Theoretically, heterogeneous preferences models improve the representative agent model at three fronts, namely idiosyncratic risk factors, heterogeneous factor premia, and individual characteristics dependent aggregation weights. We show that heterogeneous preferences models indeed outperform the representative agent model by capturing both more magnitude and more dynamics of the market risk premium. We further decompose the 62% overall gain in the explanatory power on risky premium dynamics into those three corresponding fronts. Heterogeneous preference models have another advantage over the representative agent model in the ability to accommodate the market participation heterogeneity, limiting households to market participants only further improve the explanatory power by 21%. Lastly, we show that heterogeneous preferences models also outperform the representative agent model in explaining th cross-sectional differences of Fama-French 25 size-value sorted portfolios.

Consumption-based CAPM provides a micro-founded structure to asset prices. It links households' insurance motivation for consumption smoothing and prices of insurance instruments—risky assets—together. Lucas (1978) and Breeden (1979) first propose this model in a discrete time setup; later Duffie & Zame (1989) extend it into continuous time setup. Numerous studies assuming representative agent, using different aggregate consumption data, reject the Consumption CAPM at the aggregate level, for example Breeden et al. (1989). That means, fluctuations in aggregate consumption level alone, cannot explain the risk premium. Mehra & Prescott (1985) provide a nice summary of this puzzle. To solve this problem, researchers follow two major approaches.

Most innovations in improving Consumption CAPM are about using better proxies for consumptions, especially proxies with larger volatility. Campbell & Cochrane (1999) look at the consumption level deviating from "habit", Lettau & Ludvigson (2001) construct a new measure of consumption, Aït-Sahalia et al. (2004) use consumption of luxury goods as the proxy, Malloy et al. (2009) focus on stockholders' consumptions and Savov (2011) uses garbage as the proxy. All those studies assume the existence of a representative agent.

Mankiw & Zelds (1991) started to think about a heterogeneous agent setup; they argue that agents differ from each other due to limited financial market participation for some agents. Constantinides & Duffie (1996) look at the same problem from a differnt angle. They show that in an incomplete market, the second moment of cross-sectional consumption distribution matters when agents receive idiosyncratic income shocks. Recently, Gârleanu & Panageas (2015) model heterogeneity with heterogeneous preferences using an overlapping-generations model.

No empirical study has been done to test CCAPM directly under the heterogeneous agents framework, but there are some studies testing cross-sectional implications of heterogeneous agents CCAPM; they are Brav et al. (2002), Cogley (2002) and Vissing-Jørgensen (2002). However all those three studies use power utility. Under power utility, individual consumption can be aggregated in the CCAPM relation. In other words, only the aggregate consumption growth rate is a risk factor. Therefore effects of heterogeneous agents are limited under power utility setup.

Another branch of utility is recursive utility. Epstein & Zin (1989) and Weil (1989) show a utility function in recursive form and rewrite the CCAPM relation using this utility. In this paper, we show that except for some special cases, the EZW utility function preserves the heterogeneous agents structure. The traditional aggregation result no longer holds, therefore, individual level data is necessary for pricing kernel. Before our study, no empirical study for the heterogenous preferences consumption CAPM with EZW utility has been done.

We contribute to the literature in the following aspects. Firstly we show that individual households level data is necessary for asset pricing if we do not assume some strict assumptions about households preferences or growth rates of their consumption and wealth. Secondly, we show that we can get volatile consumption that is needed to answer the equity premium puzzle by looking at cross-sectional households heterogeneity. Thirdly, we provide a direct empricial comparison between heterogeneous preferences and the representative agent consumption CAPM. Lastly, this is the first study to incorporate both preferences heterogeneity and market participation heterogeneity.

Our paper does not intend to directly answer the equity premium puzzle, nor to evaluate the absolute performance of the consumption CAPM. Instead, we focus on the relative performances of heterogeneous preferences models and the representative agent model, and provide some light into the direction of solving the equity premium puzzle eventually.

2 Model

2.1 Representative Agent Benchmark

We adopt an EZW Utility for an endowment economy with the complete financial market,

$$V_t = \left[(1-\beta)c_t^{1-\rho} + \beta \mathbb{E}_t \left[V_{t+1}^{1-\alpha}\right]^{\frac{1-\rho}{1-\alpha}}\right]^{\frac{1}{1-\rho}},$$

where β is the subjective discount rate, $1/\rho$ is the Elasticity of Inter-temporal Substitution (EIS) and α is the Relative Risk Aversion (RRA). We define wealth as $W_t = \frac{V_t}{MC_t}$, where MC_t is the marginal utility of consumption; and the return on wealth $R^W_{t+1} = \frac{W_{t+1}}{W_t - C_t}$. The stochastic discount factor is

$$S_{t,t+1} = \beta^\theta (R^W_{t+1})^{\theta-1} \left(\frac{C_{t+1}}{C_t}\right)^{-\theta/\psi}, \tag{1}$$

where $\theta = \frac{1-\alpha}{1-\rho}$ and $\psi = 1/\rho$. Then for any risky or riskless asset $i, i = 1 \cdots I$, its gross return $R^i_{t,t+1}$ must satisfy

$$\mathbb{E}_t S_{t,t+1} R^i_{t,t+1} = 1. \tag{2}$$

Let $g_{t+1} = \ln\left(\frac{C_{t+1}}{C_t}\right)$ and $h_{t+1} = \ln R^W_{t+1}$ be the log aggregate consumption growth rate and the log aggregate return on wealth, with variance σ_C^2 and σ_W^2 respectively. The risk-free rate is

$$r_{t+1}^{f}=-\ln\beta+\frac{1}{\psi}\mathbb{E}_t[g_{t+1}]-\frac{\theta}{2\psi^2}\sigma_C^2-\frac{(1-\theta)}{2}\sigma_W^2,$$

and the expected excess return of any risky asset i is

$$\mathbb{E}_t[r_{t+1}^{i}]-r_{t+1}^{f}=\frac{\theta}{\psi}\operatorname{Cov}_t[g_{t+1},r_{t+1}^{i}]+(1-\theta)\operatorname{Cov}_t[h_{t+1},r_{t+1}^{i}], \tag{3}$$

where $r_{t+1}^{i}=\ln R_{t+1}^{i}$. Therefore under the representative agent case, the aggregate consumption growth rate and the aggregate return on wealth are risk factors in the asset pricing.

2.2 Heterogeneous Preferences Setup

Under the heterogeneous preferences setup, for each household $j, j=1,\cdots,J$, the utility is specified by three idiosyncratic parameters, β^j, α^j and ρ^j (or equivalently β^j, θ^j, ψ^j).

$$V_t^j=\left[(1-\beta^j)c_t^{1-\rho_j}+\beta^j\mathbb{E}_t\left[(V_{t+1}^j)^{1-\alpha_j}\right]^{\frac{1-\rho_j}{1-\alpha_j}}\right]^{\frac{1}{1-\rho_j}}.$$

Under this specification, equation (1) now becomes

$$S_{t,t+1}=(\beta^j)^{(\theta^j)}(R_{t+1}^{w,j})^{\theta^j-1}\left(\frac{c_{t+1}^j}{c_t^j}\right)^{-\theta^j/\psi^j} \tag{4}$$

Noticing that although parameters on right hand side are heterogeneous, the stochastic discount factors are the same. This is because different agents will agree on the price of any asset in the market in equilibrium. Suppose two agents disagree on any asset, it will simultaneously create the supply and the demand for this asset at a price between their perceived prices, and the market clearing condition will eliminate the price difference. Arbitrage works. After the trade, both agents' consumptions and wealth change and their opinions on the price for this asset agree.

Since equation (2) still holds under the heterogeneous preferences setup, now for every risky asset i and every household j, the expected excess return is

$$\mathbb{E}_t[r_{t+1}^i] - r_{t+1}^f = \frac{\theta^j}{\psi^j} \operatorname{Cov}_t[g_{t+1}^j, r_{t+1}^i] + (1-\theta^j) \operatorname{Cov}_t[h_{t+1}^j, r_{t+1}^i] \tag{5}$$

Equation (5) holds for every household if the household is a market participant. We assume every household is a market participant at first. This assumption is also implicitly assumed in the past literature using the aggregate consumption. When the aggregate consumption is used, every cent in the aggregate consumption is included. This means every household is included. However, assuming every household is a market participant does not necessarily mean every household is engaging in the trading for every asset. A household will not trade an asset if their valuation of the asset equals the asset's equilibrium price under the frictionless trading or lies in the price spread under the frictional trading.

We later relax the assumption of full market participation and focus on market participants only. This practice shows one of the advantages of our model over the representative agent model which cannot separate market participants from non-participants in the aggregate data.

2.2.1 Special Cases with Aggregation Result Holds

In this section, two sets of special cases which deliver an aggregation result are discussed, such that only the aggregate consumption growth rate and the aggregate return on wealth serve as the risk factors for asset prices. Then, we argue that those special cases are unlikely to happen in reality.

Case 1 $\rho^j = \alpha^j, \forall j$. For every household, the reciprocal of Elasticity of Inter-temporal Substitution equals the Relative Risk Aversion coefficient.

Proposition 1 With $\rho^j = \alpha^j$ $(\theta^j = 1), \forall j$,

$$\mathbb{E}_t[r_{t+1}^i] - r_{t+1}^f = \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c_{t+1}^j}\right]\right)} \operatorname{Cov}_t\left[g_{t+1}, r_{t+1}^i\right],$$

where g_{t+1} is the aggregate consumption growth rate.

Proof: See appendix.

Case 2 $\alpha^j = 1, \forall j$. Every household has unit Relative Risk Aversion coefficient.

Proposition 2 With $\theta^j = 0, \forall j$,

$$\mathbb{E}_t[r^i_{t+1}] - r^f_{t+1} = \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^J 1/\mathbb{E}_t\left[1/w^j_{t+1}\right]} \mathrm{Cov}_t[h_{t+1}, r^i_{t+1}],$$

where h_{t+1} is the aggregate return on wealth.

Proof: See appendix.

When $\rho^j = \alpha^j$, $\theta^j = 1$, the EZW utility is reduced to a recursive form of power utility, therefore it is not surprising the aggregation result holds.

But this case is unlikely to happen under our setup. $\rho^j = \alpha^j, \forall j$ means all $(\rho^j, \alpha^j)_{j=1}^J$ locate on the 45 degree line on the 2-dimensional parameter plane. It may hold for some households, but it is an unrealistic restriction for the population. In reality, we expect parameter pairs for households evenly distributed on the plane for all possible combinations. If there exists one household not locating on the 45 degree line, this condition no longer holds.

Similar argument holds for the case 2. It is unrealistic for every households to have same relative risk aversion coefficient, especially all equal to 1. If there is one household's risk aversion coefficient does not equal to 1, this aggregation result no longer holds.

Case 3 $g^j_{t+1} = g_{t+1}, \forall j$. For each period, consumption growth rates are identical for every households.

Proposition 3 With $g^j_{t+1} = g_{t+1}, \forall j$,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{\sum_{j=1}^J \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}}{\sum_{j=1}^J \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \mathrm{Cov}_t[g_{t+1}, r^i_{t+1}] + \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^J \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \mathrm{Cov}_t\left[h_{t+1}, r^i_{t+1}\right]$$

Proof: See appendix.

Case 4 $h^j_{t+1} = h_{t+1}, \forall j$. For each period, returns on wealth are identical for every households.

Proposition 4 With $h^j_{t+1} = h_{t+1}, \forall j$,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^J \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] + \frac{\sum_{j=1}^J \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}}{\sum_{j=1}^J \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \operatorname{Cov}_t[h_{t+1}, r^i_{t+1}]$$

Proof: See appendix.

Under these two cases, all households have either the same consumption growth rate g_{t+1} or the same rate of return on wealth h_{t+1}, then the aggregation result still holds.

But these two cases are unlikely to happen either. We assume the complete market in this study, as a result, idiosyncratic risks are fully insured. Households' consumption growth rates and returns on wealth are fully governed by their preference parameters. Same rates require same preference parameters, which is unrealistic and contradicts to our heterogeneous preferences assumption.

2.2.2 General Case without Aggregation

Under general case, the aggregation result does not hold. We cannot aggregate heterogeneous CCAPM equations into a equation with only two aggregate risk factors; we need households' level data to price assets. Equation (5) holds for every asset i and household j. Sum over households j, we have

$$\mathbb{E}_t[r^i_{t+1}] - r^f_{t+1} = \sum_{j=1}^J \left[\phi^j \frac{\theta^j}{\psi^j} \operatorname{Cov}_t[g^j_{t+1}, r^i_{t+1}]\right] + \sum_{j=1}^J \left[\phi^j (1-\theta^j) \operatorname{Cov}_t[h^j_{t+1}, r^i_{t+1}]\right], \tag{6}$$

where ϕ_j is the aggregation weight for household j. Comparing to Equation (3) of CCAPM under the representative agent, our model improves on three fronts. Firstly we expand two aggregate risk factors g_{t+1} and h_{t+1} to $2 \times J$ idiosyncratic **r**isk **f**actors, g^j_{t+1} and h^j_{t+1} for $j = 1, \cdots, J$ (**RF**). Secondly, we improve the homogeneous factor premia θ/ψ and $1-\theta$ on consumption and wealth risk factors to heterogeneous **f**actor **p**remia θ^j/ψ^j

and $1-\theta^j$ (**FP**). Lastly, we allow individual household characteristics —the intensity of its market engagement in particular —dependent **a**ggregation **w**eights ϕ^j (**AW**).

Equation (6) highlights some deep economic intuitions. Firstly, instead of aggregate variables, now every household represents two risk factors —its own consumption growth rate and the return on wealth. It provides a more comprehensive structure by including every individual households in the pricing of any risky asset. It provides a theoretical support for multi-factor asset pricing models.

Secondly, individual household's preference matters. As we can see, not only factor loadings are individual household specific, but also factor premia are household specific. Therefore, household's preference (θ^j, ψ^j) not only determines the risky asset i's factor loadings $\mathrm{Cov}_t[g^j_{t+1}, r^i_{t+1}]$ and $\mathrm{Cov}_t[h^j_{t+1}, r^i_{t+1}]$, but also determines corresponding factor premia, θ^j/ψ^j and $1-\theta^j$. Two households with the same factor loadings can still influence the market differently through different factor premia.

Lastly, different households enter the aggregated pricing kernel with different weights. The weight is individual household's characteristics dependent. More importantly, they have the potential to accommodate the market participation heterogeneity. The weight of a market non-participant is simply zero.

2.3 Choosing Aggregation Weights

A choice that needs a special consideration is which aggregation weight to use. As we can see from equation (6), different aggregation weights can generate stark different results. An ideal aggregation weight should be the exact measure of the intensity of a household's engagement to the market. However, such a measure is difficult to define and acquire. Instead, we need a proxy for it.

By disaggregating the Consumption CAPM equation under a representative agent, we have

$$\begin{aligned}\mathbb{E}_t[r^i_{t+1}] - r^f_{t+1} &= \frac{\theta}{\psi}\,\mathrm{Cov}_t[g_{t+1}, r^i_{t+1}] + (1-\theta)\,\mathrm{Cov}_t[h_{t+1}, r^i_{t+1}] \\ &= \sum_{j=1}^{J}\left[\mathbb{E}_t\left[\frac{c^j_{t+1}}{C_{t+1}}\right]\frac{\theta}{\psi}\,\mathrm{Cov}_t[g^j_{t+1}, r^i_{t+1}]\right] + \sum_{j=1}^{J}\left[\mathbb{E}_t\left[\frac{w^j_{t+1}}{W_{t+1}}\right](1-\theta)\,\mathrm{Cov}_t[h^j_{t+1}, r^i_{t+1}]\right].\end{aligned}$$

As we can see, under the representative agent assumption, each household enters the pricing kernel using the consumption weight $\mathbb{E}_t[c^j_{t+1}/C_{t+1}]$ for its consumption growth rate and the wealth weight $\mathbb{E}_t[w^j_{t+1}/W_{t+1}]$ for its return on wealth. And by the representative agent assumption, all rates are the same, so $\mathbb{E}_t[c^j_{t+1}/C_{t+1}] = \mathbb{E}_t[w^j_{t+1}/W_{t+1}] = 1/J$. Therefore, in this paper, we choose three different weights for the aggregation weight ϕ^j_t. The first is the equal weight $1/J$, the second is the consumption weight $\mathbb{E}_t[c^j_{t+1}/C_{t+1}]$ and the last is the wealth weight $\mathbb{E}_t[w^j_{t+1}/W_{t+1}]$. We further assume that household's consumption and wealth weights c^j_t/C_t and w^j_t/W_t follow martingale, $\mathbb{E}_t[c^j_{t+1}/C_{t+1}] = c^j_t/C_t$ and $\mathbb{E}_t[w^j_{t+1}/W_{t+1}] = w^j_t/W_t$. This assumption allows predictable growth in both consumption and wealth, as long as there is no predictable growth heterogeneity.

2.4 Time-varying heterogeneous Preference

With constant heterogeneous preferences parameters (θ^j, ψ^j), the factor premia from each risk factors is constant over time. However it is inconsistent with conditional factor pricing consensus that both the risk factor loading and the factor premia are time-varying (Jagannathan & Wang (1996)). To incorporate with this, time-varying heterogeneous preference parameters (θ^j_t, ψ^j_t) are introduced. The Consumption CAPM relation now becomes

$$\mathbb{E}_t[r^i_{t+1}] - r^f_{t+1} = \sum_{j=1}^{J} \left[\phi^j_t \frac{\theta^j_t}{\psi^j_t} \operatorname{Cov}_t[g^j_{t+1}, r^i_{t+1}] \right] + \sum_{j=1}^{J} \left[\phi^j_t (1 - \theta^j_t) \operatorname{Cov}_t[h^j_{t+1}, r^i_{t+1}] \right]. \quad (7)$$

The intuition behind this model is clear. The population of heterogeneous agents is characterized by a time-varying two-dimensional distribution of (θ^j_t, ψ^j_t). On one hand, it governs the distribution of households' optimal behaviors, like inter-temporal consumption decisions and wealth accumulations; on the other hand, together with household optimal behavior, it determines the excess return for any risky asset, and consequently the cross-sectional distribution of stock returns.

3 Data and Econometric Procedure

3.1 Data

The most commonly used datasets for empirical consumption CAPM studies are the Bureau of Economic Analysis' National Income and Product Accounts Tables (NIPA) dataset and the Bureau of Labor Statistics' Consumer Expenditure Survey (CEX) dataset. The latter is the dataset used for studies on the consumption heterogeneity. Although CEX provides relatively high frequency consumption data at the quarterly level, it only tracks a household for five consecutive quarters. The lack of a panel structure and the lack of households' wealth information make the CEX not suitable for this study. The Panel Study of Income Dynamics (PSID) provides panel data for households' consumptions, but the prior 1998 data is not as accurate as CEX data. However, after the survey structure changed in 1998, the consumption data quality improved a lot. Guo (2010) shows that the total consumption from the CEX can be predicted quite well from the PSID data. In particular, Li et al. (2010) show that the PSID consumption data aligns closely with corresponding measure from the CEX, the ratios of means (PSID/CE) are 1.02 and 1.01, respectively for 2001 and 2003.

This paper uses the PSID as the data source, and constructs the consumption measure following Li et al. (2010)'s method. The household's consumption includes 4 major categories, Food, Housing, Transportation and Health Care and 2 minor categories, Education and Childcare. Table 1 provides some descriptive statistics.

Table 1: Descriptive Statistics - Household Consumption

	1999	2001	2003	2005	2007	2009
Mean	20,933	23,512	24,920	28,064	30,272	30,198
Stdev	17,259	19,135	25,770	24,614	23,382	23,968

[a] Numbers reported in this table are average households' nominal consumption and the standard deviation. The unit is US dollars.

To estimate the EZW utility, we also need returns on wealth. However, this measure is unobservable itself. Various methods are developed to handle this issue. Epstein & Zin (1991), Bakshi & Naka (1997) and Normandin & St-Amour (1998) use the financial wealth as the proxy for the aggregate wealth, but this proxy lacks the human capital part. Jeong et al. (2015) include the human capital for a better approximation of the aggregate wealth. Chen et al. (2013) develop a semi-parametric method to avoid using any proxy for returns on wealth. Bansal et al. (2007) use a simulation based method to avoid using any proxy as well.

In this study, we use the wealth data from PSID directly. The benefit of doing so is that we have a complete panel for households that includes both the consumption and the wealth - two measures we need to estimate EZW utility parameters. Based on Juster et al. (1999), the wealth data in the PSID lines up well with data from the Fed's Survey of Consumer Finances (SCF). However, like Epstein & Zin (1991) and other studies, our treatment lacks the return on the human capital as well, which may bias our estimates. Table 2 provides some descriptive statistics. It is worth noticing that our wealth data captures two drops of wealth from 2001 to 2003 and from 2007 to 2009, which precisely matches two economic recessions, the technology bubble and the Global Financial Crisis.

Table 2: Descriptive Statistics - Household Wealth

	1999	2001	2003	2005	2007	2009
Mean	123,343	110,918	107,892	124,401	152,827	141,990
Stdev	1,422,366	771,520	711,844	833,320	1,053,670	1,355,234

[a] Numbers reported in this table are average households' nominal wealth and the standard deviation. The unit is US dollars.

Our sample starts from 1999, when the consumption data is much more comprehensive than the prior 1998 period, and ends at 2009, the last year before the structure of recording the household wealth is changed. The PSID surveys are conducted biennially. So from 1999 to 2009, we have 6 periods (5 transition periods) in total. For each period in the sample, we track the exactly same entries for both the consumption and the wealth. The PSID tracks households from 1968, with a unique 1968 ID for every household. However, most households split into multiple smaller households since 1968; the panel structure only exists for households with the same 1968 ID. Therefore, I aggregate small households with the same 1968 ID to construct panel for original households. Another benefit of doing this is that we are able to average out some measurement error, just like constructing factor mimicking portfolios. After deleting households with 0 consumption, our sample contains

2,468 households. Because wealth measures are mostly self-estimated by households, they are very volatile. About half of the sample have negative net wealth, which imposes a problem when calculating returns on wealth. Therefore, we also delete households with negative net wealth. We need to keep in mind that by doing this, we may create an upward bias on RRA estimates. Our final sample contains 1,384 households. Table 3 shows descriptive statistics for consumption growth rates and returns on wealth for these 1,384 households in our sample.

Table 3: Descriptive Statistics - Growth Rates of Consumption and Wealth

Consumption	T=1	T=2	T=3	T=4	T=5	avg.
Mean	0.184	0.135	0.140	0.113	0.058	0.126
Stdev	0.524	0.495	0.484	0.497	0.519	0.504
Wealth	T=1	T=2	T=3	T=4	T=5	avg.
Mean	0.181	0.117	0.197	0.164	-0.125	0.107
Stdev	1.309	1.247	1.208	1.271	1.312	1.269

[a] Numbers reported in this table are average households' nominal consumption growth rates and return on wealth and corresponding standard deviations. $T = 1, \cdots, 5$ are five bennial transition periods from 1999 to 2009.

It is clear that the cross-sectional heterogeneity in consumption growth rates is large, while the cross-sectional heterogeneity in returns on wealth is even larger. It indirectly shows the importance of incorporating heterogeneous agents in asset pricing. Given such high level of cross-sectional differences in returns on wealth, using aggregate data generates misleading results.

3.2 Econometric Procedure

To estimate utility parameters, we use the Generalized Method of Moments (GMM), developed by Hansen (1982). Since β does not enter the Consumption CAPM equation in determining excess returns, we do not estimate β and simply choose $\beta = 0.99$ annually. Unlike past literature running the GMM along time for one representative agent using aggregate data, we run the GMM cross-sectionally for each period using household level data. We have 6 periods data and need 2 periods for each estimation. So we run the GMM for 5 transition periods.

For every period $t = 1, \cdots, 5$, we sort households by wealth and divide them into quartiles $q = 1, \cdots, 4$. We assume households in the same quartile have the same preference. For each quartile q at each period t, we run the GMM to estimate (ψ_t^q, θ_t^q). For all households $j = 1, \cdot, 346$ in quartile q, their preferences $(\psi_t^j, \theta_t^j) = (\psi_t^q, \theta_t^q)$. We run estimation for $4 \times 5 = 20$ times in total to estimate heterogeneous preferences. For each household, it may belong to a different quartile at a different period. Therefore, theoretically we may have $4^5 = 1,024$ different sets for $(\psi_t^j, \theta_t^j)_{t=1}^5$, or equivalently 1,024 types of households, which is relatively close to our sample size of 1,384. We maximize the heterogeneity in the model in accordance with the availability of the data.

For each period $t, t = 1, \cdots, 5$, and each wealth quartile $q, q = 1, \cdots, 4$, moment conditions are

$$\mathbb{E}_t \left[\beta^{\theta_t^q} \left(\frac{c_{t+1}^j}{c_t^j} \right)^{-\theta_t^q/\psi_t^q} (R_{t+1}^{W,j})^{\theta_t^q - 1} (R_{t+1}^m - R_{t+1}^f) \right] = 0, \quad \forall j = 1, \cdots, 346, \tag{8}$$

where θ_t^q and ψ_t^q are preference parameters to estimate. For household $j = 1, \cdots, 346$ in the wealth quartile q, c_{t+1}^j and c_t^j are the household's consumptions, $R_{t+1}^{W,j}$ is the household's gross return on wealth. R_{t+1}^m is the market return and R_{t+1}^f is the risk-free rate.

To make a comparison with the representative agent model, we also run the GMM using all data points for each period t, to generate a set of time-varying utility parameter $(\psi_t, \theta_t)_{t=1}^5$ for the homogeneous preference with following moment conditions,

$$\mathbb{E}_t \left[\beta^{\theta_t} \left(\frac{c_{t+1}^j}{c_t^j} \right)^{-\theta_j/\psi_j} (R_{t+1}^{W,j})^{\theta_t - 1} (R_{t+1}^m - R_{t+1}^f) \right] = 0, \quad \forall j = 1, \cdots, 1384, \tag{9}$$

where θ_t and ψ_t are the representative agent's preference parameters to estimate.

4 Estimation Result

Table 4 shows estimation results. $\{T1, \cdots, T5\}$ represent five biennial transistion periods from 1999 to 2009, $\{Q1, \cdots, Q4\}$ represent four wealth quartiles, $Q1$ is the poorest while $Q4$ is the richest. All EIS coefficients are within the $[2.0, 6.4]$ range except for the fourth quartile in the last period, which is 22.5. All RRA coefficients are within $[0.03, 0.45]$, which lies between 0 —a risk neutral preference, and 10 —the well respected upper bound for the relative risk aversion, except for the third and the fourth quartile in the last period, which are -0.04 and -0.20, respectively. The unusually large EIS at 22.5 and two risk-loving RRA coefficients in the last transition period is caused by the huge drop in the wealth for the rich people from 2007 to 2009. It is an indication of the impact magnitude of the Great Recession to households' wealth.

Table 4: Estimation Result

	EIS				RRA				avg.		Rep	
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	EIS	RRA	EIS	RRA
T1	2.79	3.14	3.57	2.03	0.45	0.34	0.27	0.39	2.88	0.36	2.90	0.36
T2	2.92	4.27	3.31	4.43	0.42	0.22	0.26	0.03	3.73	0.23	3.45	0.30
T3	4.85	3.68	3.91	3.86	0.39	0.27	0.19	0.21	4.08	0.27	3.55	0.31
T4	3.88	5.52	3.03	3.85	0.35	0.20	0.31	0.15	4.07	0.25	5.30	0.20
T5	6.04	6.38	5.22	22.5	0.20	0.06	-0.04	-0.20	10.0	0.01	6.20	0.11
avg.	4.10	4.60	3.81	7.34	0.36	0.22	0.20	0.12	4.96	0.22	4.28	0.26

[a] Numbers reported in this table are estimated Elasticity of Inter-temporal Substitution (EIS) and Relative Risk Aversion (RRA) parameters for both heterogeneous preferences model and representative agent model. $T1, \cdots, T5$ are five bennial transition periods from 1999 to 2009. $Q1, \cdots, Q4$ are four wealth quartiles, $Q1$ being the poorest and $Q4$ being the richest.

[b] Estimation is done using Generalized Method of Moments (GMM) with following moment condition,

$$\mathbb{E}_t \left[\beta^{\theta_t^q} \left(\frac{c_{t+1}^j}{c_t^j} \right)^{-\theta_j^q/\psi_j^q} (R_{t+1}^{W,j})^{\theta_t^q - 1} (R_{t+1}^m - R_{t+1}^f) \right] = 0.$$

$t = 1, \cdots, 5$ are 5 periods and $q = 1, \cdots, 4$ are 4 wealth quartiles. j represents households in wealth quratile q. R_{t+1}^m is market return and R_{t+1}^f is risk free rate. We use excess market return as test asset.

[c] EIS parameters for heterogeneous preferences model are recorded in column 2-5, with average for each period being recorded in column 10. RRA parameters for heterogeneous preferences model are recorded in column 6-9, with average for each period being recorded in column 11. Averages for EIS and RRA for each wealth quartile are recorded corresponding columns and in the last row.

[d] EIS and RRA parameters for representative agent model are recorded in column 12 and 13, with average across 5 periods being recorded in corresponding columns and in the last row.

Across different wealth quartiles, there is no clear pattern for the EIS. The RRA shows strong pattern across different wealth quartiles. With more wealth, the RRA tends to be lower, which implies wealthy households are less risk averse. Across periods, we see a weak trend for the RRA. As time progresses, households tend to be less risk averse. However, this trend is largely driven by those two negative RRA coefficients in the last period. We do observe that the EIS has a tendency to increase over time for both the average of heterogeneous preferences and the representative agent. It implies households are more inter-temporally elastic for consumptions. In other words, the impact of shocks to households welfare becomes smaller over time.

Comparing the EIS and the RRA for the representative agent with those for heterogeneous preferences, we find in each period, both the EIS and the RRA for the representative agent lie in the middle of ranges of those two parameters for heterogeneous preferences. This result is not surprising. It implies heterogeneous preferences successfully capture variations of those two parameters across different wealth groups, and the representative agent preference is a weighted average of heterogeneous preferences.

The most important aspect is that our estimation generates a much smaller RRA for the representative agent than the literature. Our RRA lies within the range of $[0.11, 0.36]$, which is smaller than $[1, 8]$ in Jeong et al. (2015) and much smaller than $[17, 60]$ in Chen et al. (2013). The reason for such small RRAs is that unlike the literature running the GMM along the time, we run the GMM estimation cross-sectionally. The aggregate consumption is really smooth along time, therefore they generate a large RRA as if the agent cares extremely about the consumption smoothing, or equivalently cares extremely about the risk aversion. However, the cross-sectional difference in consumption growth rates is huge. Therefore, when we make the assumption that agents have the same preference, the GMM sees the different consumption growth rates as the agent having very volatile consumption decisions. In other words, the agent acts as if he does not care about the risk that much. That is the reason why we are able to generate such small RRA coefficients.

Mehra & Prescott (1985) formally assert equity premium puzzle, which states that the size of the equity premium is too big to be justified by the consumption-based asset pricing theory. More specifically, traditional empiricial investigations into the consumption

CAPM face two issues. Firstly, the GMM estimation using the aggregate consumption data generates really large RRA value, which is far above the well respected upper bound of 10. Savov (2011) finds an RRA about 17, Parker & Julliard (2005) find it around 66, and Jagannathan & Wang (2007) find it around 88. Secondly, only with such a high RRA which contradicts with micro level evidence, the consumption CAPM can generate risk premium with acceptable size to match the real data. The fundamental reason for both issues is of course that the aggregate consumption is too smooth over time.

Our result provides a new angle to solve at least one aspect of the equity premium puzzle. Individual level wealth and consumptions are much more volatile than the aggregate level, and cross-sectional differences are even larger. Many individual level fluctuations are "averaged out" in the aggregation. As a result, we are not able to generate reasonably small RRA coefficients and large enough risk premium with this smoothed consumption. To find the volatile consumptions, we need heterogeneous agents and the individual level data as shown in this study.

However, we need to interpret results of this study with caution. Our low level of the RRA and the relatively high level of the EIS come from huge cross-sectional differences in households consumption growth rates and returns of wealth. As the aggregate consumption underestimates the fluctuation of true individual household's consumption growth rates, the cross-sectional data overstates the fluctuation. Nevertheless, those coefficients still serve our purpose —comparing the consumption CAPM model under heterogeneous preferences with the representative agent setup. As long as we treat both models with the same method, we can conduct the comparison.

Another question we need to answer before we move to the asset pricing practice is whether agents actually differ with each other in terms of preferences. To achieve this, we run the following regression for both the EIS and the RRA,

$$Parameters_t^q = \beta_0 + \beta_T \times t + \beta_Q \times q + \varepsilon_t^q, \tag{10}$$

where $t = 1, \cdots, 5$ and $q = 1, \cdots, 4$ are indices for periods and wealth groups. The regression serves two purposes. Firstly, it captures and confirms those trends observed previously.

Table 5: Test on Heterogeneity

	EIS	*RRA*
β_T	1.466^{**}	-0.070^{***}
t-stat	(2.456)	(-4.370)
β_Q	0.894	-0.076^{***}
t-stat	(1.184)	(-3.768)

[a] Numbers reported in this table are coefficients and corresponding t-statistics for the following regression,

$$Parameters_t^q = \beta_0 + \beta_T \times t + \beta_Q \times q + \varepsilon_t^q,$$

where $t = 1, \cdots, 5$ and $q = 1, \cdots, 4$ are indices for periods and wealth groups.
[b] (*) is p-value< 0.1, (**) is p-value< 0.05, (***) is p-value< 0.01

Secondly, the test against $\beta_Q = 0$ is also the test against the null hypothesis that all agents from different wealth groups have the same preference parameters. As we can see in Table 5, although different wealth groups show no significant difference in elasticities of inter-temporal substitution, they indeed differ in relative risk aversions.

5 Implied Market Risk Premium

Our estimated preferences parameters vary every two years. We interpolate using the average of two neighbour years to get a set of parameters for 9 consecutive years. Interpolating does not increase the theoretical maximum number of different households, which is still 1,024.

We then use those parameters, together with consumption and wealth growth rates, to calculate the implied market risk premium. However, we are not able to calculate time-varying (conditional) factor loadings for idiosyncratic risk factors, $\mathrm{Cov}_t[g_{t+1}^j, r_{t+1}]$ and $\mathrm{Cov}_t[h_{t+1}^j, r_{t+1}]$, due to the fact that our sample contains only 5 transition periods for consumption and wealth growth. Therefore, we are not able to calculate the implied market risk premium in the form of Equation (7). Instead, in this study, we calculate constant (unconditional) factor loadings for each household, $\mathrm{Cov}[g^j_{t+1}, r_{t+1}]$ and $\mathrm{Cov}[h^j_{t+1}, r_{t+1}]$, and compute the implied market risk premium (IRP) using the following equation,

$$IRP_{t+1} = \sum_{j=1}^{J} \left[\phi^j_t \frac{\theta^j_t}{\psi^j_t} \mathrm{Cov}[g^j_{t+1}, r^m_{t+1}] \right] + \sum_{j=1}^{J} \left[\phi^j_t (1-\theta^j_t) \mathrm{Cov}[h^j_{t+1}, r^m_{t+1}] \right], \qquad (11)$$

where g^j_{t+1} is the consumption growth rate and h^j_{t+1} is the return on wealth for household j, (ψ^j_t, θ^j_t) are time-varying preference parameters, and ϕ^j_t is the time-varying aggregation weight for household j. r^m_{t+1} is the market return.

Theoretically, the time-varying risk premium comes from two parts, one is the time-varying factor loading or risk exposure, the other is the time-varying factor premia or risk price. In this study, we are restricted to focus on the time-varying factor premia, and the time-varying aggregation weight, which highlights the benefits of introducing heterogeneous preferences into consumption CAPM. We believe the performance of the model could be better after including time-varying factor loadings.

We exam five models in this study. Two of which are homogeneous preference models and the other three are heterogeneous preferences models. Model 1 is the representative agent model (**Rep**). We use homogeneous preference parameters (ψ_t, θ_t) and aggregate consumption and wealth (C_t, W_t), as in Equation (12).

$$\mathbb{E}_t[r^m_{t+1}] - r^f_{t+1} = \frac{\theta_t}{\psi_t} \mathrm{Cov}[g_{t+1}, r^m_{t+1}] + (1-\theta_t) \mathrm{Cov}[h_{t+1}, r^m_{t+1}], \qquad (12)$$

where g_{t+1} is the aggregate consumption growth rate, h_{t+1} is the aggregate return on wealth, (ψ_t, θ_t) are time-varying homogeneous preference parameters.

Model 2 is the homogeneous preference model with idiosyncratic consumption and wealth (**Hom**). We use homogeneous preference parameters (ψ_t, θ_t) and the households level consumption and wealth (c^j_t, w^j_t), and use equal weight $1/J$ to aggregate, as in Equation (13).

$$\mathbb{E}_t[r^m_{t+1}] - r^f_{t+1} = \sum_{j=1}^{J} \left[\frac{1}{J} \frac{\theta_t}{\psi_t} \mathrm{Cov}[g^j_{t+1}, r^m_{t+1}] \right] + \sum_{j=1}^{J} \left[\frac{1-\theta_t}{J} \mathrm{Cov}[h^j_{t+1}, r^m_{t+1}] \right]. \qquad (13)$$

Model 3 is the heterogeneous preferences model with equal weight (**HetE**). We use heterogeneous preferences (ψ^j_t, θ^j_t) and the household level consumption and wealth (c^j_t, w^j_t), and use the equal weight $1/J$ for ϕ_t^j to aggregate.

$$\ln \mathbb{E}_t[R_{t+1}^m] - r_{t+1}^f = \sum_{j=1}^{J}\left[\frac{1}{J}\frac{\theta_t^j}{\psi_t^j}\operatorname{Cov}[g_{t+1}^j, r_{t+1}^m]\right] + \sum_{j=1}^{J}\left[\frac{1-\theta_t^j}{J}\operatorname{Cov}[h_{t+1}^j, r_{t+1}^m]\right]. \tag{14}$$

Model 4 is the heterogeneous preferences model with consumption weights (**HetC**). We change the equal weight $1/J$ in model 3 to consumption weight c_t^j/C_t, where c_t^j is household j's consumption and C_t is the aggregate consumption.

$$\mathbb{E}_t[r_{t+1}^m] - r_{t+1}^f = \sum_{j=1}^{J}\left[\frac{c_t^j}{C_t}\frac{\theta_t^j}{\psi_t^j}\operatorname{Cov}[g_{t+1}^j, r_{t+1}^m]\right] + \sum_{j=1}^{J}\left[\frac{c_t^j}{C_t}(1-\theta_t^j)\operatorname{Cov}[h_{t+1}^j, r_{t+1}^m]\right]. \tag{15}$$

Model 5 is the heterogeneous preferences model with wealth weights (**HetW**). We change the equal weights $1/J$ in model 3 to wealth weights w_t^j/W_t, where w_t^j is household j's wealth and W_t is the aggregate wealth.

$$\mathbb{E}_t[r_{t+1}^m] - r_{t+1}^f = \sum_{j=1}^{J}\left[\frac{w_t^j}{W_t}\frac{\theta_t^j}{\psi_t^j}\operatorname{Cov}[g_{t+1}^j, r_{t+1}^m]\right] + \sum_{j=1}^{J}\left[\frac{w_t^j}{W_t}(1-\theta_t^j)\operatorname{Cov}[h_{t+1}^j, r_{t+1}^m]\right]. \tag{16}$$

Table 6 reports results of the implied market risk premium from those five models, as well as the actual market risk premium.[1]

A first impression is that the implied market risk premium generated by all five models are relatively small comparing to the actual market risk premium. This is the direct result of small RRA coefficients we estimated. Therefore, this study cannot give a straight answer towards the equity premium puzzle. Instead, we focus on comparing performances between homogeneous preference models and heterogeneous preferences models, and hopefully this paper can shed some light into the direction to solve the puzzle. However, we need also keep in mind that in the sample period, the first decade of this century, we saw two recessions, the bust of the dotcom bubble and the Global Financial Crisis. As a result, the market

[1]The actual market risk premium data is from Kenneth French's website.

Table 6: Annualized Implied Risk Premium

	2000	2001	2002	2003	2004	2005	2006	2007	2008	avg.
Actual	-17.6%	-15.2%	-22.8%	30.8%	10.7%	3.09%	10.6%	1.04%	-38.3%	-4.19%
Rep	0.10%	0.09%	0.07%	0.10%	0.14%	0.12%	0.06%	-0.04%	-0.15%	0.06%
Hom	0.03%	0.03%	0.02%	0.04%	0.06%	0.05%	0.02%	-0.03%	-0.09%	0.02%
HetE	-0.24%	0.07%	0.25%	0.28%	0.26%	0.15%	0.03%	-0.26%	-0.71%	-0.02%
HetC	-0.14%	-0.04%	-0.06%	0.01%	0.04%	0.01%	-0.04%	-0.18%	-0.61%	-0.11%
HetW	-1.07%	-0.11%	-0.43%	0.22%	0.08%	0.04%	-0.05%	-0.28%	-1.38%	-0.23%

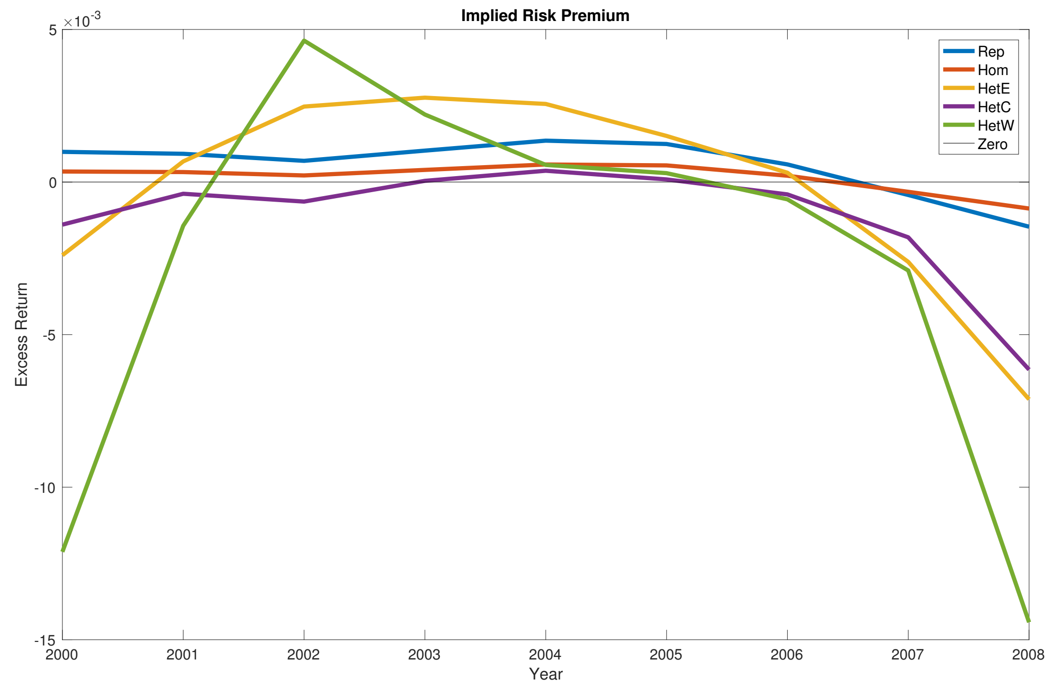
[a] Numbers reported in this table are implied risk premium from 2000 to 2008 generated by five models.

[b] *Actual* is actual risk premium we observe from data. *Rep* is the representative agent model, using homogeneous preference and aggregate data. *Hom* is the homogeneous preference model with households level data. *HetE* is the heterogeneous preferences model with equal weights. *HetC* is the heterogeneous preferences model with consumption weights. *HetW* is the heterogeneous preferences model with wealth weights.

risk premium varies dramatically between -38% and 30.8%. Any attempt to capture the fluctuation of that magnitude from the consumption perspective would be much more difficult than capturing some moderate fluctuations of the risk premium in the longer time horizon.

Interestingly, neither of these two homogeneous preference models can produce a negative average risk premium like what we see in the data, while all three heterogeneous preferences models successfully generate a negative average risk premium with wealth weights performing the best. As we can see in Figure 1, HetW produces the most volatile risk premium series.

Figure 1: Annualized Implied Risk Premium



The average implied risk premium only provides a rough examination on the magnitude aspect of models' performances. We further look into each year to see whether each model successfully predict the sign of the actual risk premium. Table 7 records performances of sign predictions for each model.

A first look shows that all five models successfully predict the boom from 2003 to 2005 and the Global Financial Crisis in 2008. All five models fail to predict the positive risk premium of 1% in 2007, they all suggest a negative risk premium. If we consider the actual start date of the Global Financial Crisis is mid-2007 and the 1% is barely above 0, this wrong prediction for 2007 isn't that wrong.

The best performer in predicting signs is HetC, the heterogeneous preferences model with consumption weight. It successfully captures the tech bubble, the 2003-2005 boom and the global financial crisis. HetE and HetW are runner-ups, with HetE performing relatively worse in the tech bubble and HetW performing worse in the Global Financial Crisis. Interestingly, neither of two homogeneous preference models succeed to capture the tech bubble while both succeed to capture the GFC. It raises a question, why?

Table 7: Sign Accuracy

	2000	2001	2002	2003	2004	2005	2006	2007	2008
Actual	-18%	-15%	-23%	31%	11%	3%	11%	1%	-38%
Rep	✗	✗	✗	✓	✓	✓	✓	✗	✓
Hom	✗	✗	✗	✓	✓	✓	✓	✗	✓
HetE	✓	✗	✗	✓	✓	✓	✓	✗	✓
HetC	✓	✓	✓	✓	✓	✓	✗	✗	✓
HetW	✓	✓	✗	✓	✓	✓	✗	✗	✓

[a] This table reports sign prediction accuracy for five models in study from 2000 to 2008. ✓means the model successfully predicts the sign of market risk premium of that year, ✗ means the model unsuccessfully predicts the sign of market risk premium of that year.

[b] *Actual* is actual risk premium we observe from data. *Rep* is the representative agent model, using homogeneous preference and aggregate data. *Hom* is the homogeneous preference model with households level data. *HetE* is the heterogeneous preferences model with equal weights. *HetC* is the heterogeneous preferences model with consumption weights. *HetW* is the heterogeneous preferences model with wealth weights.

Recall the nature of those two crisis. The tech bubble starts from a fall in tech companies' stock prices, which tightens the financial market, and later the tightness spreads to the real economy. The GFC starts from the drop of housing prices in mid-2006, which causes the drop in MBS prices, and later the bankruptcy of the Lehman Brothers. The key difference between those two regarding households is that the former starts with the wealth drop of a small group of households, those who are rich and have investments in tech stocks or who work for tech firms, while the latter starts with the wealth drop of the majority of households, anyone who owns real estates.

Mechanically, when we estimate homogeneous preference parameters, the wealth drop of the small group of wealthy households is ignored. All households are included in the estimation while the majority do not experience the same wealth drop. As a result, the homogeneous preference parameters provide the information representing more of the unaffected majority. Therefore, we cannot see the crisis using neither the aggregate data nor the households level data with the homogeneous preference. In contrast, if we estimate heterogeneous preferences parameters, we look into different groups of households. Although households with the wealth drop in the tech bubble is a small group considering all households, they are a relatively larger subgroup in the rich households group. As a result, the wealth drop information is captured by the rich group's preferences parameters, and later it shows up in the risk premium when we aggregate across households. More over, this information is amplified by the wealthy group's relatively large aggregation weights. That also explains why the wealth weight outperforms the equal weights in predicting the tech bubble.

The story is different for the GFC. We first see the wealth drop for the majority of households. This information is captured when we estimate heterogeneous preferences. Similarly, this information is also captured when we estimate the homogeneous preference since the size of the wealth drop group is large. Therefore, we are able to predict the crisis using both homogeneous preference models and heterogeneous preferences models. The ability to capture the crisis starting from a small group of households is one advantage of heterogeneous preferences models over homogeneous preference models. We may have financial crisis caused by different groups of people, and the size of the group can be small or large. Our analysis suggests that homogeneous preference models can only identify a potential crisis if the crisis is caused by a large group of households, like the GFC. However, with heterogeneous preferences models, a crisis starts with a smaller group of households can still be identified. Our analysis also sheds some light into the study of the systemic risk. Different risks may affect different amount of households, we may use heterogeneous preferences models to study what is the threshold above which a risk will become a systemic risk.

6 Time Series of Risk Premium

We further exam pricing performances of those five models using following Times Series regression equation

$$Actual_t = \alpha + \beta \times Implied_t + \varepsilon_t, \tag{17}$$

where $Implied_t$ are model generated market excess returns, $Aactual_t$ are real world market excess returns from data, $t = 2000, \cdots, 2008$. Results are recorded in the following Table 8. With a perfect asset pricing model, we expect to have α equals to 0 as Gibbons et al. (1989) argues, β is positive and equals to 1 and R^2 goes to 1. As we can see, α for all five models are small and statistically insignificant from zero, suggesting no unexplained risk premium left. Slopes for neither homogeneous models is statistically significant while all three heterogeneous preferences models' slopes are statistically significantly positive, which suggests those three models are pricing the market risk premium on the right direction. Looking at magnitudes of those coefficients, heterogeneous preferences models have smaller coefficients than homogeneous preference models, meaning heterogeneous preferences models generate closer to actual risk premium, although the difference with the actual data is still large. Again, this paper does not intend to solve the equity premium puzzle with model generated risk premium that is quantitatively close enough to real data, instead this paper focuses on the relative performance between heterogeneous preferences models and homogeneous preference models. Heterogeneous preferences models indeed outperform homogeneous preference models in terms of generating quantitatively closer to actual risk premium.

Table 8: Market Risk Premium

	Rep	Hom	HetE	HetC	HetW
α	-0.111	-0.083	-0.033	0.042	0.010
t-stat	(-1.500)	(-1.276)	(-0.582)	(0.677)	(0.131)
β	126.682	261.170	41.657*	73.719**	21.407*
t-stat	(1.747)	(1.867)	(2.253)	(2.600)	(2.036)
R^2	0.3037	0.3323	0.4202	0.4912	0.3720

[a] Numbers reported in this table are regression coefficients and their corresponding t-statistics, as well as R^2 for five models in study using following regression equation,

$$Actual_t = \alpha + \beta \times Implied_t + \varepsilon_t,$$

where $Implied_t$ are implied risk premium at time t generated by models, and $Actual_t$ are actual risk premium at time t.

[b] *Rep* is the representative agent model, using homogeneous preference and aggregate data. *Hom* is the homogeneous preference model with households level data. *HetE* is the heterogeneous preferences model with equal weights. *HetC* is the heterogeneous preferences model with consumption weights. *HetW* is the heterogeneous preferences model with wealth weights.

[c] (*) is p-value< 0.1, (**) is p-value< 0.05

We further exam whether our implied risk premium captures the dynamics of actual risk premium by looking at R^2. As the benchmark, the Representative Agent model explains about 30.4% of total variations in the risk premium dynamics. The representative agent model has the most restrictive assumption that all households have the same preference, and it uses the least informative data —the aggregate consumption and wealth. Therefore, it is not surprising it performs the worst out of all five models. The Heterogeneous Preferences with Consumption Weight model is the best performer, explaining almost half of the dynamics at 49.1%. This is a 61.77% improvement from the benchmark. With our models' settings, we are able to decompose this improvement into three parts, with each corresponding to one advantage heterogeneous preferences models have over the representative agent model, namely Idiosyncratic Risk Factors (RF), Heterogeneous Factor Premia (FP) and Idiosyncratic Characteristics dependent Aggregation Weights (AW).

Comparing the Representative Agent (Rep) model (Equation 12) to the Homogeneous Preference with Idiosyncratic Growth Rates (Hom) model (Equation 13), the difference between the two is that we introduce Idiosyncratic Risk Factors (RF) in Hom, while Rep uses Aggregate Risk Factors. By doing this, R^2 increases from 30.4% to 33.2%, which is a 9.4% incraese from Rep at 30.4%.

Comparing the Homogeneous Preference with Idiosyncratic Growth Rates (Hom) model (Equation 13) to the Heterogeneous Preferences with Equal Weight (HetE) model (Equation 14), the difference between the two is that we introduce Heterogeneous Factor Premia (FP) in HetE. By doing this, R^2 increases from 33.2% to 42.0%, which is a 28.9% increase from Rep at 30.4%.

Comparing the Heterogeneous Preferences with Equal Weights (HetE) model (Equation 14) to the Heterogeneous Preferences with Consumption Weight (HetC) model (Equation 15), the difference between the two is that we introduce Idiosyncratic Characteristics dependent Weights (AW) in HetC. By doing this, R^2 increases from 42.0% to 49.1%, which is a 23.4% increase from Rep at 30.4%.

Total improvement in R^2 from the Representative Agent model (Rep) to the Heterogeneous Preferences with Consumption Weight (HetC) model is the sum of all three improvements from those three fronts,

$$\underbrace{61.77\%}_{\textbf{Total Improvement}} = \underbrace{9.42\%}_{\textbf{RF}} + \underbrace{28.94\%}_{\textbf{FP}} + \underbrace{23.41\%}_{\textbf{AW}}.$$

Another interesting point is that the Heterogeneous Preferences with Wealth Weight (HetW) model (Equation 16) performs the best in terms of capturing the magnitude, while it performs much worse than using Consumption Weights (HetC) in terms of capturing the dynamics. The good performance on magnitude comes from the large volatility of returns on wealth. Its average standard deviation is 1.27, which is more than double of consumption growth rates' 0.50. The bad performance in capturing the dynamics suggests that the Wealth Weight is not a good proxy for the intensity of market engagements, which is counter intuitive as we expect wealthy people to engage more in the market. Brunnermeier & Nagel (2008) find that fluctuations in nominal wealth do not affect households' portfolio choice much, and our result supports that finding. A large drop in a rich household's nominal wealth does not mean an equally large drop in his market engagement intensity, instead they may still hold the same portfolio and have the same influence on the market. The consumption, on the other hand, can also measure the intensity if we believe wealthy people also consume more, but it is much less volatile than the wealth and is a better proxy for the relatively stable market engagement intensity.

Brav et al. (2002) and Cogley (2002) use the average of individual households' pricing kernel as the heterogeneous pricing kernel, and use up to the third moment of the cross-sectional distribution of consumption growth rates as risk factors. Our study is a direct extension on those two studies at three aspects. Firstly, we use the more generalized Epstein-Zin-Weil utility, rather than using the special case —power utility —in those two studies. Secondly, because of the panel structure in our data, we are able to look at the whole sample distribution in consumption growth rates, instead of first three moments of the cross-sectional distribution of consumptions. Thirdly, we adopt the market engagement intensity based aggregation weight rather than using the equal weight in those studies. Lastly, we allow heterogeneous preference parameters.

7 Accont for Market Participation

One advantage of our configuration of Heterogeneous Preferences models over the Representative Agent model is that we are able to accommodate for heterogeneous market participations through heterogeneous aggregation weights. To demonstrate this, we look at three types of market participations. The first type is General Market Participants (**GP**), defined as households with either an IRA stock account or a non-IRA stock account. This definition includes three different types of households, (i) households with only an IRA account, (ii) households with only a non-IRA account, and (iii) households with both an IRA and a non-IRA account. Another type of market participations is Investors (**I**), defined as households with a non-IRA stock account, which is a subset of General Market Participants (**GP**), including type (ii) and (iii) households. Households with a non-IRA account intentionally bet in the market. The last type of market participations is Smart Investors (**SI**), defined as households with both an IRA and a non-IRA stock account, which is a subset of Investors (**I**), including type (iii) households only. Those are households not only intentionally bet in the market, they are also sophisticated enough to explore the tax advantage of IRA account. These households are most likely being marginal traders in the market. Following Table 9 records results.

When we limit households to those who participate in the stock market, our Heterogeneous Preferences models get further improvements in explaining risk premium dynamics. R^2 improve from 49.1% using all households (HetC) to 50.9% using only General Market Participants (HetC/GP), and further to 55.5% by reducing market participants to Investors (HetC/I) and finally to 56.2% using only Smart Investors (HetC/SI).

The improvement in R^2 comes at the cost of a slightly worse performance in magnitudes, with the relative performance between using Consumption Weights and Wealth Weights holds. A clear trade-off between capturing the dynamics and capturing the magnitude arise. The optimal weights is therefore an open question for future studies.

Table 9: Market Participation

	HetC	HetW	HetC/GP	HetW/GP	HetC/I	HetW/I	HetC/SI	HetW/SI
Coef.	73.719**	21.407*	76.717**	21.780*	84.635**	21.244*	87.587**	20.831*
t-stat	(2.600)	(2.036)	(2.695)	(2.230)	(2.956)	(2.021)	(2.996)	(1.954)
R^2	0.4913	0.3720	0.5091	0.4153	0.5552	0.3684	0.5619	0.3529

[a] Numbers reported in this table are regression slopes and their corresponding t-statistics, as well as R^2 for 2 heterogeneous preferences models in study with 4 different market participation using following regression equation,

$$Actual_t = \alpha + \beta \times Implied_t + \varepsilon_t,$$

where $Implied_t$ are implied risk premium at time t generated by models, and $Actual_t$ are actual risk premium at time t.

[b] Four different market participations are full market participation, general market participatants (GP), investors (I) and smart investors (SI). General market participant is defined as households with either an IRA stock account or a non-IRA stock account. Investor is defined as households with a non-IRA stock account. Smart Investor is defined as households with both an IRA stock account and a non-IRA stock account.

[c] *HetC* is the heterogeneous preferences model with consumption weights under full market participation. *HetW* is the heterogeneous preferences model with wealth weights under full market participation. *HetC/GP* is the heterogeneous preferences model with consumption weights with general market participants only. *HetW/GP* is the heterogeneous preferences model with wealth weights with general market participants only. *HetC/I* is the heterogeneous preferences model with consumption weights with investors only. *HetW/I* is the heterogeneous preferences model with wealth weights with investors only. *HetC/SI* is the heterogeneous preferences model with consumption weights with smart investors only. *HetW/SI* is the heterogeneous preferences model with wealth weights with smart investors only.

[d] (*) is p-value< 0.1, (**) is p-value< 0.05

8 Cross Section of Stock Returns

The next question we try to answer is whether our Heterogeneous Preferences models outperform the Representative Agent model in terms of the explanatory power on the cross-section of stock returns.

Our test assets are Fama-French 25 portfolios sorted by size and book-to-market ratio, developed by Fama & French (1992). Because the number of our risk factors for individual consumptions is too big ($1,384 \times 2 = 2,768$), we cannot use the standard Fama & MacBeth (1973) method. Instead, we calculate individual consumption and wealth risk factors' loadings for each test asset, and use time-varying heterogeneous preference parameters to calculate implied excess returns for that test asset. Then we run OLS regressions on realized excess returns of test assets against implied excess returns. Given the panel structure of 25 assets and 9 periods, we run three types of regressions. The first one is the Pooled-OLS, which utilizes $25 \times 9 = 225$ data points. The second type of regressions is time series regressions for all 25 assets. The last type of regressions is 9 cross-sectional regressions. For the second and the third type, we record average coefficients, average p-values and average R^2s. We also calculate the t-statistics against the null hypothesis that all coefficients equal to zero. Results are recorded in Table 10

Results we see in time series of market risk premium generally hold in the cross-sectional analysis of stock returns. Firstly, Heterogeneous Preferences models generally have smaller coefficients than the Representative Agent model, except for the average coefficient of cross-sectional regressions for HetC. Heterogeneous Preferences models do generate closer to actual excess returns than homogeneous preference counterparts. Secondly, Heterogeneous Preferences models capture both more excess returns' time series dynamics and cross-sectional variations than the Representative Agent model, with HetC being the best performer. R^2 increases from [12.7%, 25.0%] for Rep to [39.5%, 52.9%] for HetC. The results for HetC and HetW reported in Table 10 does not account for market participation heterogeneity, as the purpose of this analysis is to compare the relative performances of heterogeneous preferences models and the representative agent model, with the same market participation assumption and the same data. If we reduce to market participants only, results are better for both HetC and HetW.

9 Conculsion

We adopt the Epstein-Zin-Weil utility under a heterogeneous preferences environment. We show that individual level data is needed to price assets if we do not assume some strict assumptions for aggregation. Our heterogeneous preferences model improves the representative agent model in three aspects, 1) idiosyncratic risk factors, 2) heterogeneous factor premia, and 3) idiosyncratic characteristics dependent aggregation weights.

We group households into four types by wealth, and assume households under the same

Table 10: Cross Section of Stock Returns

Pooled OLS	Rep	Hom	HetE	HetC	HetW
coef	74.343***	74.978***	47.031***	63.369***	23.883***
t-stat	(5.700)	(8.071)	(12.376)	(13.670)	(10.220)
R^2	0.127	0.2261	0.407	0.456	0.319
Time Series	Rep	Hom	HetE	HetC	HetW
avg.coef	178.980***	242.277***	52.803***	84.328***	29.006***
t-stat	(11.749)	(6.389)	(19.816)	(16.782)	(14.152)
avg.p-value	0.220	0.073	0.076	0.037	0.121
avg.R^2	0.250	0.417	0.458	0.529	0.414
Cross Section	Rep	Hom	HetE	HetC	HetW
avg.coef	24.919*	28.801*	16.107*	33.897***	8.641
t-stat	(1.529)	(1.553)	(1.652)	(2.530)	(1.244)
avg.p-value	0.113	0.073	0.066	0.008	0.151
avg.R^2	0.224	0.304	0.284	0.395	0.247

[a] For Pooled OLS, we run one regression for each model and record corresponding results, using following regression equation,

$$Actual_t^i = \alpha + \beta \times Implied_t^i + \varepsilon_t^i,$$

where $Implied_t^i$ are implied excess return for asset i at time t generated by models, and $Actual_t$ are actual excess return for asset i at time t. We use 25 Fama-French size-value sorted assets.

[b] For Times Series, we run 25 regressions for each model, one regression for each asset, using following regression equation,

$$Actual_t^i = \alpha + \beta^i \times Implied_t^i + \varepsilon_t^i,$$

We record average of β^i, average p-value and average R^2 for those 25 regression. t-stat is the t-statistics of the t-test against the null that average of β^i equals to zero.

[c] For Cross Section, we run 9 regressions for each model, one regression for each period, using following regression equation,

$$Actual_t^i = \alpha + \beta_t \times Implied_t^i + \varepsilon_t^i,$$

We record average of β_t, average p-value and average R^2 for those 25 regression. t-stat is the t-statistics of the t-test against the null that average of β_t equals to zero.

[d] *Rep* is the representative agent model, using homogeneous preference and aggregate data. *Hom* is the homogeneous preference model with households level data. *HetE* is the heterogeneous preferences model with equal weights. *HetC* is the heterogeneous preferences model with consumption weights. *HetW* is the heterogeneous preferences model with wealth weights.

[e] (*) is p-value<0.1, (**) is p-value<0.05, (***) is p-value<0.01

type have the same preference parameters. Then we use the PSID consumption and wealth data to estimate time-varying preference parameters for those four types of households, and a representative agent. Because of the large cross-sectional heterogeneity in consumption growth rates and returns on wealth, our estimation generates relatively large EIS and small RRA coefficients. One aspect of the equity premium puzzle is that GMM estimation using aggregate consumption generates unrealistically large relative risk aversion parameter because the aggregate consumption is too smooth. Our result suggests that the large cross-sectional difference within households may help bring the size of risk aversion parameter down.

With estimated parameters, we then calculate models' implied risk premium. We look at five models, the representative agent model, the homogeneous preference model with individual level data, heterogeneous preferences models with equal weights, consumption weights and wealth weights. We find heterogeneous prefernces models generally outperform the representative agent model, they both generate quantitatively closer to actual risk premium and capture more risk premium dynamics. Using consumption weights, the heterogeneous preferences model improves R^2 from the representative agent model by 61.77%. Out of this 61.77% improvement, idiosyncratic risk factors contribute 9.42%, heterogeneous factor premia contribute 28.94% and idiosyncratic aggregation weights contribute 23.41%.

Heterogeneous preferences models also outperform the representative agent model in explaining the cross section of stock returns. Using Fama-French 25 size-value sorted portfolios, we show that heterogeneous preferences models generally generate quantatively closer to actual excess returns and capture more excess returns' time series dynamics and cross-sectional variations.

Lastly, our heterogeneous preferences models have the advantage over the representative agent model in the ability to accommodate heterogeneous market participations. Including market participants only, we further achieve a 23.24% improvement in capturing risk premium dynamics, which brings total improvement from the representative agent model to 85.01%.

References

Aït-Sahalia, Y., Parker, J. A., & Yogo, M. (2004). Luxury goods and the equity premium. *Journal of Finance*, *59*, 2959–3004.

Bakshi, G. S., & Naka, A. (1997). An empirical investigation of asset pricing models using japanese stock market data. *Journal of International Money and Finance*, *16*, 81–112.

Bansal, R., Gallant, A. R., & Tauchen, G. (2007). Rational pessimism, rational exuberance, and asset pricing models. *Review of Economic Studies*, *74*, 1005–1033.

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, *59*(4), 1481–1509.

Brav, A., Constantinides, G. M., & Geczy, C. C. (2002). Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy*, *110*, 793–824.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, *7*(3), 265–296.

Breeden, D. T., Gibbons, M. R., & Litzenberger, R. H. (1989). Empirical tests of the consumption-oriented capm. *Journal of Finance*, *44*(2), 231–262.

Brunnermeier, M. K., & Nagel, S. (2008). Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals' asset allocation. *American Economic Review*, *98*(3), 713–736.

Campbell, J. Y., & Cochrane, J. H. (1999). By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, *107*, 205–251.

Chen, X., Favilukis, J., & Ludvigson, S. C. (2013). An estimation of economic models with recursive preferences. *Quantitative Economics*, *4*, 39–83.

Cogley, T. (2002). Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey. *Journal of Monetary Economics*, *49*, 309–334.

Constantinides, G. M., & Duffie, D. (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy*, *104*, 219–240.

Duffie, D., & Zame, W. (1989). The consumption-based capital asset pricing model. *Econometrica*, *57*(6), 1279–1297.

Epstein, L. G., Farhi, E., & Strzalecki, T. (2014). How much would you pay to resolve long-run risk?. *American Economic Review*, *104*(9), 2680–2697.

Epstein, L. G., & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, *57*, 937–969.

Epstein, L. G., & Zin, S. E. (1991). Substitution risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, *99*, 263–286.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*, 427–465.

Fama, E. F., & MacBeth. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*, 607–636.

Gârleanu, N., & Panageas, S. (2015). Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. *Journal of Political Economy*, *123*(3), 670–685.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, *57*(5), 1121–1152.

Guo, S. (2010). The superior measure of psid consumption: An update. *Economics Letters*, *108*, 253–256.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*, 1029–1054.

Jagannathan, R., & Wang, Y. (2007). Lazy investors, discretionary consumption, and the cross-section of stock returns. *Journal of Finance*, *62*, 1623–1661.

Jagannathan, R., & Wang, Z. (1996). The conditional capm and the cross section of expected returns. *Journal of Finance*, *51*, 3–53.

Jeong, D., Kim, H., & Park, J. Y. (2015). Does ambiguity matter? estimating asset pricing models with a multiple-priors recursive utility. *Journal of Financial Economics*, *115*, 361–382.

Juster, F. T., Smith, J. P., & Stafford, F. (1999). The measurement and structure of household wealth. *Labour Economics*, *6*, 253–275.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, *56*, 815–849.

Li, G., Schoeni, R. F., Danziger, S., & Charles, K. K. (2010). New expenditure data in the psid: comparisons with the ce. *Monthly Labor Review*, 29–39.

Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica*, *46*(6), 1429–1445.

Malloy, C. J., Moskowitz, T. J., & Vissing-Jørgensen, A. (2009). Long-run stockholder consumption risk and asset returns. *Journal of Finance*, *64*, 2427–2479.

Mankiw, N. G., & Zelds, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics*, *29*, 97–112.

Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, *15*, 145–161.

Normandin, M., & St-Amour, P. (1998). Substitution, risk aversion, taste shocks and equity premia. *Journal of Applied Econometrics*, *13*, 265–281.

Parker, J. A., & Julliard, C. (2005). Consumption risk and the cross section of expected returns. *Journal of Political Economy*, *113*, 185–222.

Savov, A. (2011). Asset pricing with garbage. *Journal of Finance*, *66*, 177–201.

Vissing-Jørgensen, A. (2002). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy*, *110*, 825–853.

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetery Economics*, *24*, 401–421.

Appendix

A.1 Proof of Propositions

A.1.1 Proof of Proposition 1

Proof. Under $\rho^j = \alpha^j, \forall j$,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{1}{\psi^j} \operatorname{Cov}_t[g^j_{t+1}, r^i_{t+1}]$$

Rewrite

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{1}{\psi^j} \operatorname{Cov}_t \left[\ln \left(\frac{c^j_{t+1}}{c^j_t} \right), r^i_{t+1} \right]$$

By Stein's Lemma,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{1}{\psi^j} \operatorname{Cov}_t \left[\left(\frac{c^j_{t+1}}{c^j_t} \right), r^i_{t+1} \right] \mathbb{E}_t \left[\frac{c^j_t}{c^j_{t+1}} \right] = \frac{1}{\psi^j} \mathbb{E}_t \left[\frac{1}{c^j_{t+1}} \right] \operatorname{Cov}_t[c^j_{t+1}, r^i_{t+1}]$$

Multiply $\psi^j / \mathbb{E}_t \left[\frac{1}{c^j_{t+1}} \right]$ on both sides, we have

$$\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} \right) \psi^j / \mathbb{E}_t \left[\frac{1}{c^j_{t+1}} \right] = \operatorname{Cov}_t[c^j_{t+1}, r^i_{t+1}]$$

Sum over j,

$$\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} \right) \sum_{j=1}^{J} \left(\psi^j / \mathbb{E}_t \left[\frac{1}{c^j_{t+1}} \right] \right) = \sum_{j=1}^{J} \operatorname{Cov}_t[c^j_{t+1}, r^i_{t+1}] = \operatorname{Cov}_t \left[\sum_{j=1}^{J} c^j_{t+1}, r^i_{t+1} \right] = \operatorname{Cov}_t[C_{t+1}, r^i_{t+1}]$$

Rearrange

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{1}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c^j_{t+1}}\right]\right)} \mathrm{Cov}_t[C_{t+1}, r^i_{t+1}] = \frac{C_t}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c^j_{t+1}}\right]\right)} \mathrm{Cov}_t\left[\frac{C_{t+1}}{C_t}, r^i_{t+1}\right]$$

With Stein's Lemma again,

$$\begin{aligned}
\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} &= \frac{C_t}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c^j_{t+1}}\right]\right)} \mathrm{Cov}_t\left[\frac{C_{t+1}}{C_t}, r^i_{t+1}\right] \\
&= \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c^j_{t+1}}\right]\right)} \frac{C_t}{\mathbb{E}_t[C_{t+1}]} \mathrm{Cov}_t\left[\frac{C_{t+1}}{C_t}, r^i_{t+1}\right] \\
&= \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c^j_{t+1}}\right]\right)} \mathrm{Cov}_t\left[\ln\left(\frac{C_{t+1}}{C_t}\right), r^i_{t+1}\right] \\
&= \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^J \left(\psi^j / \mathbb{E}_t\left[\frac{1}{c^j_{t+1}}\right]\right)} \mathrm{Cov}_t\left[g_{t+1}, r^i_{t+1}\right]
\end{aligned}$$

□

A.1.2 Proof of Proposition 2

Proof. Under $\alpha^j = 1, \forall j$,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \mathrm{Cov}_t[h^j_{t+1}, r^i_{t+1}]$$

Rewrite

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \mathrm{Cov}_t\left[\ln\left(\frac{w^j_{t+1}}{w^j_t - c^j_t}\right), r^i_{t+1}\right]$$

By Stein's Lemma,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{w^j_t - c^j_t}{\mathbb{E}_t[w^j_{t+1}]} \mathrm{Cov}_t\left[\left(\frac{w^j_{t+1}}{w^j_t - c^j_t}\right), r^i_{t+1}\right] = \mathbb{E}_t\left[\frac{1}{w^j_{t+1}}\right] \mathrm{Cov}_t[w^j_{t+1}, r^i_{t+1}]$$

Multiply $1/\mathbb{E}_t\left[1/w^j_{t+1}\right]$ on both sides, we have

$$\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1}\right) \frac{1}{\mathbb{E}_t\left[1/w^j_{t+1}\right]} = \mathrm{Cov}_t[w^j_{t+1}, r^i_{t+1}]$$

Sum over j,

$$\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1}\right) \sum_{j=1}^{J} \frac{1}{\mathbb{E}_t\left[1/w^j_{t+1}\right]} = \sum_{j=1}^{J} \mathrm{Cov}_t[w^j_{t+1}, r^i_{t+1}] = \mathrm{Cov}_t\left[\sum_{j=1}^{J} w^j_{t+1}, r^i_{t+1}\right] = \mathrm{Cov}_t[W_{t+1}, r^i_{t+1}]$$

Rearrange

$$\begin{aligned}
\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} &= \frac{1}{\sum_{j=1}^{J} 1/\mathbb{E}_t\left[1/w^j_{t+1}\right]} \mathrm{Cov}_t[W_{t+1}, r^i_{t+1}] \\
&= \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^{J} 1/\mathbb{E}_t\left[1/w^j_{t+1}\right]} \frac{W_t - C_t}{\mathbb{E}_t[W_{t+1}]} \mathrm{Cov}_t\left[\frac{W_{t+1}}{W_t - C_t}, r^i_{t+1}\right]
\end{aligned}$$

With Stein's Lemma again,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^{J} 1/\mathbb{E}_t\left[1/w^j_{t+1}\right]} \mathrm{Cov}_t\left[\ln\left(\frac{W_{t+1}}{W_t - C_t}\right), r^i_{t+1}\right]$$

$\square$

A.1.3 Proof of Proposition 3

Proof. Under $g^j_{t+1} = g_{t+1}, \forall j$,

$$\begin{aligned}
\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} &= \frac{\theta^j}{\psi^j} \mathrm{Cov}_t[g_{t+1}, r^i_{t+1}] + (1-\theta^j) \mathrm{Cov}_t[h^j_{t+1}, r^i_{t+1}] \\
&= \frac{\theta^j}{\psi^j} \mathrm{Cov}_t\left[g_{t+1}, r^i_{t+1}\right] + (1-\theta^j) \mathrm{Cov}_t\left[\ln\left(\frac{w^j_{t+1}}{w^j_t - c^j_t}\right), r^i_{t+1}\right]
\end{aligned}$$

By Stein's Lemma,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{\theta^j}{\psi^j} \operatorname{Cov}_t \left[g_{t+1}, r^i_{t+1}\right] + (1-\theta^j) \operatorname{Cov}_t \left[\ln\left(\frac{w^j_{t+1}}{w^j_t - c^j_t}\right), r^i_{t+1}\right]$$
$$= \frac{\theta^j}{\psi^j} \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] + (1-\theta^j)\mathbb{E}_t\left[\frac{1}{w^j_{t+1}}\right] \operatorname{Cov}_t[w^j_{t+1}, r^i_{t+1}]$$

Rearrange we have

$$\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1}\right) \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]} - \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]} = \operatorname{Cov}_t[w^j_{t+1}, r^i_{t+1}]$$

Sum over j,

$$\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1}\right) \sum_{j=1}^{J} \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]} - \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] \sum_{j=1}^{J} \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}$$
$$= \sum_{j=1}^{J} \operatorname{Cov}_t[w^j_{t+1}, r^i_{t+1}] = \operatorname{Cov}_t[W_{t+1}, r^i_{t+1}]$$

Rearrange

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{\sum_{j=1}^{J} \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}}{\sum_{j=1}^{J} \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] + \frac{1}{\sum_{j=1}^{J} \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t[W_{t+1}, r^i_{t+1}]$$
$$= \frac{\sum_{j=1}^{J} \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}}{\sum_{j=1}^{J} \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}]$$
$$+ \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^{J} \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \frac{W_t - C_t}{\mathbb{E}_t[W_{t+1}]} \operatorname{Cov}_t \left[\frac{W_{t+1}}{W_t - C_t}, r^i_{t+1}\right]$$

By Stein's Lemma again,

$$
\begin{aligned}
\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} &= \frac{\sum_{j=1}^J \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}}{\sum_{j=1}^J \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] \\
&+ \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^J \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t\left[\ln\left(\frac{W_{t+1}}{W_t - C_t}\right), r^i_{t+1}\right] \\
&= \frac{\sum_{j=1}^J \frac{\theta^j}{\psi^j(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}}{\sum_{j=1}^J \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t[g_{t+1}, r^i_{t+1}] + \frac{\mathbb{E}_t[W_{t+1}]}{\sum_{j=1}^J \frac{1}{(1-\theta^j)\mathbb{E}_t[1/w^j_{t+1}]}} \operatorname{Cov}_t\left[h_{t+1}, r^i_{t+1}\right]
\end{aligned}
$$

□

A.1.4 Proof of Proposition 4

Proof. Under $h^j_{t+1} = h_{t+1}, \forall j$,

$$
\begin{aligned}
\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} &= \frac{\theta^j}{\psi^j} \operatorname{Cov}_t[g^j_{t+1}, r^i_{t+1}] + (1-\theta^j) \operatorname{Cov}_t[h_{t+1}, r^i_{t+1}] \\
&= \frac{\theta^j}{\psi^j} \operatorname{Cov}_t\left[\ln\left(\frac{c^j_{t+1}}{c^j_t}\right), r^i_{t+1}\right] + (1-\theta^j) \operatorname{Cov}_t[h_{t+1}, r^i_{t+1}]
\end{aligned}
$$

By Stein's Lemma,

$$
\begin{aligned}
\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} &= \frac{\theta^j}{\psi^j} \operatorname{Cov}_t\left[\ln\left(\frac{c^j_{t+1}}{c^j_t}\right), r^i_{t+1}\right] + (1-\theta^j) \operatorname{Cov}_t[h_{t+1}, r^i_{t+1}] \\
&= \frac{\theta^j}{\psi^j}\mathbb{E}_t[1/c^j_{t+1}] \operatorname{Cov}_t[c^j_{t+1}, r^i_{t+1}] + (1-\theta^j) \operatorname{Cov}_t[h_{t+1}, r^i_{t+1}]
\end{aligned}
$$

Rearrange we have

$$
\left(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1}\right) \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]} - \operatorname{Cov}_t[h_{t+1}, r^i_{t+1}] \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]} = \operatorname{Cov}_t[c^j_{t+1}, r^i_{t+1}]
$$

Sum over j,

$$\Big(\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1}\Big) \sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]} - \mathrm{Cov}_t[h_{t+1}, r^i_{t+1}] \sum_{j=1}^{J} \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]} = \sum_{j=1}^{J} \mathrm{Cov}_t[c^j_{t+1}, r^i_{t+1}]$$
$$= \mathrm{Cov}_t[C_{t+1}, r^i_{t+1}]$$

Rearrange

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{1}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t[C_{t+1}, r^i_{t+1}] + \frac{\sum_{j=1}^{J} \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t[h_{t+1}, r^i_{t+1}]$$
$$= \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \frac{C_t}{\mathbb{E}_t[C_{t+1}]} \mathrm{Cov}_t\left[\frac{C_{t+1}}{C_t}, r^i_{t+1}\right] + \frac{\sum_{j=1}^{J} \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t[h_{t+1}, r^i_{t+1}]$$

By Stein's Lemma again,

$$\ln \mathbb{E}_t[R^i_{t+1}] - r^f_{t+1} = \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t\left[\ln\left(\frac{C_{t+1}}{C_t}\right), r^i_{t+1}\right] + \frac{\sum_{j=1}^{J} \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t[h_{t+1}, r^i_{t+1}]$$
$$= \frac{\mathbb{E}_t[C_{t+1}]}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t[g_{t+1}, r^i_{t+1}] + \frac{\sum_{j=1}^{J} \frac{\psi^j(1-\theta^j)}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}}{\sum_{j=1}^{J} \frac{\psi^j}{\theta^j \mathbb{E}_t[1/c^j_{t+1}]}} \mathrm{Cov}_t[h_{t+1}, r^i_{t+1}]$$

□

A.2 Ambiguity Aversion

Bansal & Yaron (2004) utilize Epstein-Zin-Weil utility and propose a long run risk model that better explains some quantitative characteristics of equity market. This is a significant advancement in solving equity premium puzzle. However, Epstein et al. (2014) find an new quantitative issue arise with the model. Epstein-Zin-Weil utility separates Elasticity of Inter-temporal Substitution (*EIS*) and Relative Risk Aversion (*RRA*) parameters, the relative size of those parameters implies agent's attitude towards ambiguity aversion, or in other words, agent's preference regarding early resolution of risk. More specifically, Epstein et al. (2014) state that if an agent's $EIS \times RRA > 1$, he will prefer early resolution of risk; if

$EIS \times RRA < 1$, he will prefere late resolution of risk; if $EIS \times RRA = 1$, under which the Epstein-Zin-Weil utility is reduced to power utility, he has no preference towards the timing of risk resolution.

We also check our heterogeneous agents' preferences on early resolution of risk. Following Table 11 records size of $EIS \times RRA$ for different agents.

Table 11: Early Resolution of Risk

$EIS \times RRA$	Q1	Q2	Q3	Q4	avg.	Rep
T1	1.25	1.07	0.96	0.79	1.02	1.04
T2	1.23	0.94	0.86	0.13	0.79	1.03
T3	1.89	0.99	0.74	0.81	1.11	1.10
T4	1.36	1.10	0.94	0.58	1.00	1.06
T5	1.21	0.38	-0.21	-4.50	-0.78	0.68
avg.	1.39	0.90	0.66	-0.44	0.63	0.79

[a] Numbers reported in this table are estimated Elasticity of Inter-temporal Substitution (EIS) times Relative Risk Aversion (RRA) parameters ($EIS \times RRA$) for both heterogeneous preferences model and representative agent model.
[b] $T1, \cdots, T5$ are five bennial transition periods from 1999 to 2009. $Q1, \cdots, Q4$ are four wealth quartiles, $Q1$ being the poorest and $Q4$ being the richest. *Rep* is representative agent.
[c] Averages across different wealth groups for each period are recorded in column 6 and averages across different periods for each wealth groups are recorded in row 7.

Like preferences parameters, we want to check whether agents do differ with each other. To achieve this, we run following regression and record results in Table 12.

$$EIS_t^q \times RRA_t^q = \beta_0 + \beta_T \times t + \beta_Q \times q + \varepsilon_t^q.$$

As we can see, our estimated households do differ in terms of attitude towards early or late resolution of risk. Poor people prefer early resolution while rich people prefer late resolution. It is another proof of the existence of heterogeneous preferences and different households with different wealth do have different Preferences.

Table 12: Early Resolution of Risk

	$EIS \times RRA$
β_T	-0.339^{*}
t-stat	(-2.042)
β_Q	-0.572^{**}
t-stat	(-2.722)

[a] Numbers reported in this table are coefficients and corresponding t-statistics for the following regression,

$$Parameters_t^q = \beta_0 + \beta_T \times t + \beta_Q \times q + \varepsilon_t^q,$$

where $t = 1, \cdots, 5$ and $q = 1, \cdots, 4$ are indices for periods and wealth groups.
[b] (*) is p-value< 0.1, (**) is p-value< 0.05, (***) is p-value< 0.01

A.3 Pricing Kernel Performance without Tech Bubble

A question we are addressing here is whether the relatively worse performances of representative agent model is sololy because of its failure to capture the tech bubble. We then exclude 2000-2002 tech bubble and look at the rest of the sample. Results recorded in following Table 13.

Overall, heterogeneous preferences models still outperform representative agent model in terms of both capturing market excess return magnitude and dynamics. It signals the importance of heterogeneous preferences models in asset pricing. Heterogeneous preferences models not only be able to capture some crisis which representative agent model is unable to capture, they can also provide more information about asset prices in periods when representative agent model can actually work well.

We further look into details, by excluding tech bubble, we increase R^2 for all five models. It is not surpring for Rep and Hom to get an increase as they both fail to capture the tech bubble, as a result, R^2's for those two models increase to almost 70%. It is surpring though that even three heteregeneous preferences models achieve an increase in R^2, to more than 80% for HetE and HetC and more than 90% for HetW. It suggests that including tech bubble actually worsens the performance of heterogeneous preferences models. Considering

Table 13: Market Risk Premium 2003-2008

	Rep/6Y	Hom/6Y	HetE/6Y	HetC/6Y	HetW/6Y
β	168.953**	327.612**	54.575**	83.174**	37.010***
t-stat	(2.928)	(2.866)	(4.517)	(4.366)	(6.325)
R^2	0.6818	0.6755	0.8361	0.8265	0.9091

[a] Numbers reported in this table are regression coefficients and their corresponding t-statistics, as well as R^2 for five models in study using following regression equation,

$$Actual_t = \alpha + \beta \times Implied_t + \varepsilon_t,$$

where $Implied_t$ are implied risk premium at time t generated by models, and $Actual_t$ are actual risk premium at time t. t is from 2003 to 2008.

[b] *Rep/6Y* is the representative agent model, using homogeneous preference and aggregate data. *Hom/6Y* is the homogeneous preference model with households level data. *HetE/6Y* is the heterogeneous preferences model with equal weights. *HetC/6Y* is the heterogeneous preferences model with consumption weights. *HetW/6Y* is the heterogeneous preferences model with wealth weights.

[c] (*) is p-value< 0.1, (**) is p-value< 0.05, (***) is p-value< 0.01.

the relative performance between representative agent model and heterogeneous preferences models, it is reasonable to believe our level of heterogeneity of 4 wealth groups does not fully capture information about the tech bubble.