Asset Pricing and the One Percent *

Alexis Akira Toda† Kieran Walsh‡

This version: August 18, 2014

Abstract

We find that when the income share of the top 1% income earners in the U.S. rises above trend by one percentage point, subsequent one year market excess returns decline on average by 5.6%. This negative relation remains strong and significant even when controlling for classic return predictors such as the price-dividend and the consumption-wealth ratios. To explain this stylized fact, we build a general equilibrium asset pricing model with heterogeneity in wealth and risk aversion across agents. Our model admits a testable moment condition and a novel two factor covariance pricing formula, where one factor is inequality. Intuitively, when wealth shifts into the hands of rich and risk tolerant agents, average risk aversion falls, pushing down the risk premium. Our model is broadly consistent with data and provides a novel positive explanation of both market excess returns over time and the cross section of returns across stocks.

Keywords: equity premium; heterogeneous risk aversion; return prediction; wealth distribution.

JEL codes: D31, D53, D58, G12, G17.

1 Introduction

Does the wealth distribution matter for asset pricing? Common sense tells us that it does: as the rich get richer, they buy risky assets and drive up prices. Indeed, over a century ago prior to the advent of modern mathematical finance, Fisher (1910) argued that there is an intimate relationship between prices, the heterogeneity of agents in the economy, and booms and busts. He contrasted (p. 175) the "enterpriser-borrower" with the "creditor, the salaried man, or the laborer," emphasizing that the former class of society accelerates fluctuations in prices and production. *Central* to his theories of fluctuations were differences in preferences and wealth across people.

Following the seminal work of Lucas (1978), however, the "representative agent" consumption-based asset pricing models—which seem to allow no role for agent heterogeneity—have dominated the literature, at least until recently. Yet agent heterogeneity may (and is likely to) matter even if a representative

*We benefited from comments by Dan Cao, Gabriele Camera, Jim Hamilton, Gordon Hanson, Fumio Hayashi, Marc Muendler, Larry Schmidt, Ross Starr, and seminar participants at Hitotsubashi ICS, Kyoto, Tokyo, UCSD, Vassar College, and Yokohama National University.

†Department of Economics, University of California San Diego. Email: atoda@ucsd.edu.

‡Darden School of Business, University of Virginia. Email: walshk@darden.virginia.edu.

agent exists: unless agents have very specific preferences that admit the Gorman (1953) aggregation (a knife-edge case, which is unlikely to hold), the preferences of the representative agent (in the sense of Constantinides (1982)) will in general depend on the wealth distribution, as pointed out by Gollier (2001). Indeed, even with complete markets, the preferences of the representative agent are typically nonstandard when individual utilities do not reside within quite particular classes.

To see the intuition as to why the wealth distribution affects asset pricing, consider an economy consisting of people with different attitudes towards risk or beliefs about future dividends. In this economy, equilibrium risk premiums and prices balance the agents' preferences and beliefs. If wealth shifts into the hands of the optimistic or less risk averse, for markets to clear, prices of risky assets must rise and risk premiums must fall to counterbalance the new demand of these agents. In this paper, we establish both the theoretical and empirical links between income/wealth inequality and asset prices.

This paper has two main contributions. First, we build a simple general equilibrium model of asset prices with heterogeneous agents and derive testable implications linking asset returns and inequality across risk aversion types. We show that the stochastic discount factor depends both on market returns and average risk tolerance. Although the connection between the heterogeneity of agents' risk aversion and asset prices has been recognized at least since Dumas (1989), the literature seems to lack testable implications that can be easily examined in financial data. We thus provide a link allowing us to subject the theory to empirical scrutiny. The model also implies a new covariance pricing formula regarding the cross section of returns across assets: the average return of a stock depends *positively* on its correlation with the market and *negatively* on its correlation with the wealth share of risk tolerant agents. In short, the wealth distribution determines an asset pricing factor: average risk aversion across agents. When average risk aversion is low, average marginal utility is high. Therefore, assets correlated with top wealth shares (and thus average risk tolerance) command relatively low risk premiums.

We also illustrate the effect of the wealth distribution on asset prices with a numerical example. Two types of agents, one more risk averse than the other, inhabit the model, and they trade a riskless bond in zero net supply and a risky asset in positive supply. We perform comparative statics with respect to the initial endowment share of the more risk tolerant agents and find an inverse relationship between their income and the subsequent equity premium. In line with intuition, as the risk tolerant rich get richer, they buy risky assets, increasing their relative price. Subsequent excess returns thus fall.

Second, we empirically explore the theoretical predictions. We find that when the income share of the top 1% of income earners in the U.S. is above trend, the subsequent one and five year U.S. stock market equity premiums are below average. Such relationships are predictions of many heterogeneous agent general equilibrium models. We thus provide empirical support for a literature which has been subject to relatively little direct testing. Furthermore, the patterns we uncover are intuitive. In short, if one believes top earners are all else equal more willing to trade risk for return, then it should not be surprising that in the data asset returns suffer as the rich get richer.

More specifically, we employ scatter plots and regression analysis to establish the correlation between inequality and returns. Regressions of the year t

to year $t+1$ return on the year t top 1% income share indicate a strong and significant correlation: when the top 1% income share rises above trend by one percentage point, subsequent one year market excess returns decline on average by 5.6%. Five year returns decline by about 25%. This relation is strongly statistically significant and admits an R-squared of around 20%. We show that an estimated top wealth share series also negatively predicts subsequent returns. Furthermore, the top 1% income share predicts asset returns even after we control for some classic return forecasters such as the price-dividend ratio (Shiller, 1981) and the consumption-wealth ratio (Lettau and Ludvigson, 2001a,b). In fact, our analysis shows that the inequality series is only weakly related to these other variables. So, in our analysis the top 1% income share is not simply a proxy for the relative price level, which previous research shows correlates with subsequent returns. This is perhaps surprising because one imagines the rich being disproportionately exposed to stock price fluctuations. Our findings are also robust to the exclusion of capital gains in the income share series.

We also empirically investigate the key moment conditions of the model. We structurally estimate the model's moment conditions to explore the extent to which the Fama-French 2×3 portfolios (sorted by size and book-to-market ratio) are explained by their relationship with market returns and inequality between two risk aversion types. As conjectured, we find that the rich are more risk tolerant than are the poor. We fail to reject our models and show that they outperform their homogeneous agent (standard CAPM) counterpart with respect to the Fama-French portfolios. We provide estimates of average risk tolerance over time in the U.S. and argue that its fluctuations are qualitatively in line with Irving Fisher's narratives.

We supplement the GMM exercise by linear regressions, inspired by our covariance pricing formula, of Fama-French portfolio returns on market excess returns and the top income share. The income share coefficient is significant for four of six regressions. Moreover, in line with our model, the income share coefficients are inversely related to the average portfolio returns: high return portfolios like "small size, high book-market" are negatively correlated with the top 1% share, and low return portfolios like "big size, low book-market" are positively correlated with the top 1% share. To reiterate, our explanation for this pattern is very simple and relies only on the 1% being more risk tolerant: as the rich get richer, the economy's risk tolerance increases and marginal utility rises. Thus, assets negatively correlated with the 1% have low payouts in high marginal utility states and thus command a high risk premium.

The relationship between the Fama-French portfolios and the top 1% holds even after controlling for the HML Fama-French factor. Also, we find that the top 1% share is significantly correlated with the SMB Fama-French factor. Given these two pieces of evidence and given the ad hoc nature of SMB (and the opacity of its success), we suggest that this factor is to some degree proxying for changes in the wealth distribution and average risk aversion.

The rest of the paper is organized as follows. After discussing the literature, Section 2 introduces two simple models in which preference heterogeneity matters for asset pricing. We derive testable implications. Section 3 establishes the empirical link between inequality and subsequent asset returns. Section 4 estimates and tests heterogeneous risk aversion models and explores the ability of the models to explain the cross section of returns.

1.1 Related literature

For many years after Fisher, in analyzing the link between individual utility maximization and asset prices, financial theorists either employed a rational representative agent or considered cases of heterogeneous agent models that admit aggregation, that is, cases in which the model is equivalent to one with a representative agent. Extending the portfolio choice work of Markowitz (1952) and Tobin (1958), Sharpe (1964) and Lintner (1965a,b) established the Capital Asset Pricing Model (CAPM).[1] These original CAPM papers, which concluded that an asset's covariance with the aggregate market determines its return, actually allowed for substantial heterogeneity in endowments and risk preferences across investors. However, their form of quadratic or mean-variance preferences admitted aggregation and obviated the role of the wealth distribution.

The seminal consumption-based asset pricing work of Lucas (1978), Breeden (1979), and Hansen and Singleton (1983) also abstracted from investor heterogeneity. They and others derived and tested analytic relationships between the marginal rate of substitution of a representative agent (with standard preferences) and asset prices. Despite the elegance and tractability of the representative agent/aggregation approach, it has failed to adequately explain the fluctuations of asset prices in the economy. Largely inspired by the limited empirical fit of the CAPM (in explaining the cross section of stock returns), the equity premium puzzle (Mehra and Prescott, 1985), and excess stock market volatility and related price-dividend ratio anomalies (Shiller, 1981), since the 1980s theorists have extended macro/finance general equilibrium models to consider non-standard utility functions and meaningful investor heterogeneity. These models can be categorized into three groups.

In the first group, agents have identical standard (constant relative risk aversion) preferences but are subject to uninsured idiosyncratic risks.[2] Although the models of this literature have had some quantitative success, the empirical results (based on consumption panel data) are mixed and may even be spuriously caused by the heavy tails in the consumption distribution (Toda and Walsh, 2013).

In the second group, markets are complete and agents have identical but non-homothetic preferences.[3] In this class of models the marginal rates of substitution are equalized across agents and a "representative agent" in the sense of Constantinides (1982) exists, but aggregation in the sense of Gorman (1953) fails. Therefore there is a room for agent heterogeneity to matter for asset pricing. The quantitative effects of agent heterogeneity in this literature have been found to be minor, however.

In the third group, to which this paper belongs, markets are complete and agents have homothetic but heterogeneous preferences. This type of model has not been studied much[4] because in the long run the economy is dominated by the richest agent (the agent with the largest expected wealth growth rate) and there-

[1] See Geanakoplos and Shubik (1990) for a general and rigorous treatment of CAPM theory.

[2] Examples are Mankiw (1986), Constantinides and Duffie (1996), Heaton and Lucas (1996), Krusell and Smith (1998), Brav et al. (2002), Cogley (2002), Balduzzi and Yao (2007), Kocherlakota and Pistaferri (2009), among many others. See Ludvigson (2013) for a review.

[3] Examples are Gollier (2001) and Hatchondo (2008).

[4] Examples are Dumas (1989), Wang (1996), Chan and Kogan (2002), Hara et al. (2007), Longstaff and Wang (2012), and Bhamra and Uppal (2014).

fore heterogeneity in preferences does not matter (Blume and Easley, 2006).[5] One exception is Gârleanu and Panageas (2012), who study a continuous-time overlapping generations endowment economy with two agent types with Epstein-Zin constant elasticity of intertemporal substitution/constant relative risk aversion preferences. Even if the aggregate consumption growth is i.i.d. (geometric Brownian motion), the risk-free rate and the equity premium are time-varying, even in the long run. The intuition is that when the risk tolerant agents have a higher wealth share, they drive up asset prices and the interest rate. The effect of preference heterogeneity persists since new agents are constantly born. Toda (2012) builds a dynamic general equilibrium model with production and agents with heterogeneous recursive preferences and finds that the standard Euler equation approach that employs consumption is not robust to model misspecification whereas the value function approach that bypasses consumption is fairly robust. To the best of our knowledge our paper is the first in the literature to empirically test the asset pricing implications of models with preference heterogeneity.

Although the wealth distribution theoretically affects asset prices, there are few empirical papers that directly document this connection. To the best of our knowledge, Johnson (2012) is the only one that explores this issue using income/wealth distribution data. However, his analysis is quite different from ours. First, his model relies on a "keeping up with the Joneses"-type consumption externality with incomplete markets. In contrast, we employ a standard general equilibrium model (a plain vanilla Arrow-Debreu model). Second, he does not directly test moment conditions from his model, whereas we perform structural estimation of heterogeneous risk aversion parameters. Nor does he derive and estimate our two factor covariance pricing formula or discuss average risk tolerance. Finally, Johnson (2012) does not explore the ability of top income shares to predict market excess returns.

2 Asset pricing with heterogeneous risk aversion

In this section we present two simple models in which the heterogeneity in agents' attitude towards risk matters for asset pricing and derive testable implications as well as a novel covariance pricing formula.

2.1 Two models with heterogeneous risk aversion

2.1.1 Model with arbitrary preferences

Consider a two period model with time indexed by $t = 0, 1$. There are I agents indexed by $i = 1, \ldots, I$. Agent i has the expected utility over final wealth w_{i1},

$$\mathrm{E}[u_i(w_{i1})],$$

where u_i is von Neumann-Morgenstern utility function with $u_i' > 0$ and $u_i'' < 0$. There are J assets indexed by $j = 1, \ldots, J$. Asset j trades at price q_j per share

[5] Coen-Pirani (2004) and Guvenen (2009) study cases with incomplete markets and heterogeneous Epstein-Zin preferences.

(to be determined in equilibrium) at $t = 0$ and pays dividend D_j at $t = 1$. Agent i is endowed with n_{ij} shares of asset j at $t = 0$. Let $w_{i0} = \sum_{j=1}^{J} q_j n_{ij}$ be the initial wealth of agent i. Letting n'_{ij} be the number of shares agent i holds after trade, the optimal portfolio problem is

$$\begin{aligned} \underset{\{n'_{ij}\}}{\text{maximize}} \quad & \mathrm{E}[u_i(w_{i1})] \\ \text{subject to} \quad & \sum_{j=1}^{J} q_j n'_{ij} = w_{i0}, \; w_{i1} = \sum_{j=1}^{J} D_j n'_{ij}. \end{aligned} \tag{2.1}$$

Assuming no trade frictions, the first-order condition for optimality with respect to n'_{ij} is

$$\mathrm{E}[u'_i(w_{i1}) D_j] = \lambda_i q_j,$$

where $\lambda_i > 0$ is the Lagrange multiplier for the budget constraint. Dividing by q_j and letting $R_j = D_j/q_j$ be the gross return on asset j and assuming the existence of a risk-free asset (with gross risk-free rate R_f), we obtain

$$\mathrm{E}[u'_i(w_{i1})(R_j - R_f)] = 0. \tag{2.2}$$

Using the Taylor approximation $u'(x) \approx u'(a) + u''(a)(x - a)$ for $a = a_i = \mathrm{E}[w_{i1}]$ and $x = w_{i1}$, we obtain

$$\mathrm{E}[(u'_i(a_i) + u''_i(a_i)(w_{i1} - a_i))(R_j - R_f)] = 0, \tag{2.3}$$

where we have written $=$ instead of $\approx$. Dividing both sides by $-u''_i(a_i) > 0$ and using the definition of the relative risk tolerance (reciprocal of the Arrow-Pratt measure of relative risk aversion)

$$\tau_i = -\frac{u'_i(a_i)}{a_i u''_i(a_i)},$$

we obtain

$$\mathrm{E}[(a_i \tau_i - (w_{i1} - a_i))(R_j - R_f)] = 0. \tag{2.4}$$

Adding across all agents, letting $W_1 = \sum_{i=1}^{I} w_{i1}$ be the aggregate wealth at $t = 1$, and dividing by $\mathrm{E}[W_1] = \sum_{i=1}^{I} a_i$, we obtain

$$\mathrm{E}[(\bar{\tau} - W_1/\mathrm{E}[W_1] + 1)(R_j - R_f)] = 0,$$

where $\bar{\tau} = \sum_i a_i \tau_i / \sum_i a_i$ is the weighted average risk tolerance. Now since every asset must be held by some agent in equilibrium and there is no consumption at $t = 0$, adding individual budget constraints, the growth rate of aggregate wealth must be equal to the *market return* R_m. Therefore $W_1 = R_m W_0$. Taking expectations, we obtain $\mathrm{E}[W_1] = \mathrm{E}[R_m] W_0$. Therefore $W_1/\mathrm{E}[W_1] = R_m/\mathrm{E}[R_m]$. Putting all the pieces together, we obtain

$$\mathrm{E}[((\bar{\tau} + 1)\,\mathrm{E}[R_m] - R_m)(R_j - R_f)] = 0, \tag{2.5}$$

which is the key moment condition that we will exploit throughout the rest of the paper.

Alternatively, if we apply the Taylor approximation around the initial wealth w_{i0} instead of the expected future wealth $\mathrm{E}[w_{i1}]$, (2.3) holds with $a_i = w_{i0}$. Adding across i and dividing by aggregate wealth $W_0 = \sum_i w_{i0}$, we get

$$\mathrm{E}[(\bar{\tau} + 1 - R_m)(R_j - R_f)] = 0, \tag{2.6}$$

where $\bar{\tau} = \sum_i w_{i0}\tau_i / \sum_i w_{i0}$ is the average risk tolerance weighted by initial wealth.

The moment conditions (2.5) and (2.6) are both valid approximations; (2.5) is more accurate because we approximate around the expectation of the relevant variable (future wealth), but (2.6) is easier to handle because there is no need to predict future stock returns or the wealth distribution.

2.1.2 Model with quadratic preferences

The above discussion is only approximate since it involves linear approximations, but it can be made exact under some assumptions. Consider exactly the same model except preferences. Suppose that agents are mean-variance optimizers. More precisely, agent i maximizes

$$v_i(\theta) = \mathrm{E}[R(\theta)] - \frac{1}{2\tau_i}\mathrm{Var}[R(\theta)],$$

where $R(\theta) = \sum_j R_j\theta_j$ is the gross portfolio return and $\tau_i > 0$ is the risk tolerance. Assume that in addition to the risky J assets, agents can trade a risk-free asset in zero net supply, where the risk-free rate R_f is determined in equilibrium. Letting θ_j the fraction of wealth invested in asset j, the fraction of wealth invested in the risk-free asset is $1 - \sum_j \theta_j$. Therefore the portfolio return is

$$R(\theta) = \sum_{j=1}^{J} R_j\theta_j + R_f\left(1 - \sum_{j=1}^{J}\theta_j\right) = R_f + \sum_{j=1}^{J}(R_j - R_f)\theta_j.$$

The expected return and variance of the portfolio are

$$\mathrm{E}[R(\theta)] = R_f + \langle \mu - R_f 1, \theta\rangle, \;\; \mathrm{Var}[R(\theta)] = \langle \theta, \Sigma\theta\rangle,$$

respectively, where μ is the J-vector of expected returns $\mu_j = \mathrm{E}[R_j]$, 1 is the J-vector of ones, and Σ is the variance-covariance matrix of the returns $\mathbf{R} = (R_1, \ldots, R_J)$. Thus the optimal portfolio problem of agent i reduces to

$$\underset{\theta}{\text{maximize}} \;\; R_f + \langle \mu - R_f 1, \theta\rangle - \frac{1}{2\tau_i}\langle \theta, \Sigma\theta\rangle,$$

where $\theta \in \mathbb{R}^J$ is unconstrained. The first-order condition is

$$\mu - R_f 1 - \frac{1}{\tau_i}\Sigma\theta = 0 \iff \theta_i^* = \tau_i\Sigma^{-1}(\mu - R_f 1), \tag{2.7}$$

where θ_i^* is the optimal portfolio of agent i.

Since every asset must be held by someone and the risk-free asset is in zero net supply by definition, the average portfolio weighted by individual wealth,

$\sum w_{i0}\theta_i^* / \sum_i w_{i0}$, must be the *market portfolio* (value-weighted average portfolio), denoted by θ_m. Letting $\bar{\tau} = \sum_{i=1}^I w_{i0}\tau_i / \sum_{i=1}^I w_{i0}$ be the average risk tolerance, by taking the weighted average of the first-order condition (2.7), we get

$$\theta_m = \bar{\tau}\Sigma^{-1}(\mu - R_f 1). \tag{2.8}$$

Multiplying both sides of (2.8) by Σ and comparing the j-th element, we get

$$\mathrm{Cov}[R_m, R_j] = \bar{\tau}\,\mathrm{E}[R_j - R_f],$$

where $R_m = R(\theta_m)$ is the market return. Since

$$\begin{aligned}\mathrm{Cov}[R_m, R_j] &= \mathrm{E}[(R_m - \mathrm{E}[R_m])(R_j - R_f + R_f)] \\ &= \mathrm{E}[(R_m - \mathrm{E}[R_m])(R_j - R_f)],\end{aligned}$$

it follows that

$$\mathrm{E}[(\bar{\tau} + \mathrm{E}[R_m] - R_m)(R_j - R_f)] = 0. \tag{2.9}$$

(2.9) is identical to (2.5) or (2.6) except that 1 is replaced by $\mathrm{E}[R_m]$.

2.2 Covariance pricing formula

Using the moment conditions derived above, we can obtain a two factor covariance pricing formula. For simplicity, assume that there are two types of agents with high and low risk tolerance τ_H and τ_L. Letting $0 < \alpha < 1$ be the fraction of wealth of high risk tolerant agents, from (2.5) we get

$$\mathrm{E}[((1 + \alpha\tau_H + (1-\alpha)\tau_L)\,\mathrm{E}[R_m] - R_m)(R_j - R_f)] = 0. \tag{2.10}$$

We can derive similar moment conditions from (2.6) and (2.9).

The moment condition (2.10) implies that

$$M = (1 + \alpha\tau_H + (1-\alpha)\tau_L)\,\mathrm{E}[R_m] - R_m$$

is a scaled stochastic discount factor. Using the definition of covariance, from (2.10) we obtain

$$\begin{aligned}\mathrm{E}[R_j] - R_f &= -\frac{\mathrm{Cov}[M, R_j]}{\mathrm{E}[M]} \\ &= \frac{1}{\mathrm{E}[M]}\,\mathrm{Cov}[R_m, R_j] - \frac{\Delta\tau}{\tau_L + \Delta\tau\,\mathrm{E}[\alpha]}\,\mathrm{Cov}[\alpha, R_j],\end{aligned} \tag{2.11}$$

where $\Delta\tau = \tau_H - \tau_L > 0$. Therefore, covariance both with the market *and* with the wealth distribution are priced. In particular, assets with returns that are positively correlated with the wealth share of more risk tolerant agents should have lower average returns. To the extent that the richest agents are the most risk tolerant, assets that are positively correlated with the wealth share of the rich should have lower average returns. In short, the theory produces a two factor model of asset prices.

Why does negative correlation with the wealth share of risk tolerant agents lead to high average returns for an asset? Intuitively, an asset that pays off when the wealth share of risk tolerant agents is low is delivering precisely when the economy's average marginal utility is low. This is because with mean-variance

preferences, marginal utility is increasing in risk tolerance. Thus, when risk tolerance is low on average, marginal utility is low. Moreover, assets that deliver in low marginal utility times will on average be less attractive to the agents. Therefore, the agents demand on average a higher risk premium on assets that pay off when the risk tolerant wealth share is low.

In contrast, the implicit representative or average agent has high marginal utility under two scenarios: low market returns and high risk tolerance. Assets that deliver in these states provide insurance to this representative agent and thus command lower risk premiums on average.

2.3 Numerical example

In this subsection we numerically solve two examples of the model in Section 2.1, one with agents with constant but heterogeneous relative risk aversion (CRRA) and another with identical decreasing relative risk aversion (DRRA) agents except initial wealth.

2.3.1 Two CRRA agents with heterogeneous risk aversion

Assume that there are two agent types, $i \in \{H, L\}$. Agent H has high risk tolerance τ_H and agent L has low risk tolerance τ_L. For numerical values, we set $\gamma_H = 1/\tau_H = 0.5$ and $\gamma_L = 1/\tau_L = 2$. There is only one risky asset (stock) and a risk-free asset in zero net supply. Fraction α of stocks are initially held by agent H and fraction $1 - \alpha$ by agent L. There are two states with equal probability, and the dividend of the stock is $1 + \mu \pm \sigma$, where $\mu = 0.07$ and $\sigma = 0.2$. To see the accuracy of the approximation, we both solve the exact model numerically as well as the approximate model semi-analytically using either (2.5) or (2.6) with error tolerance $\varepsilon = 10^{-8}$.

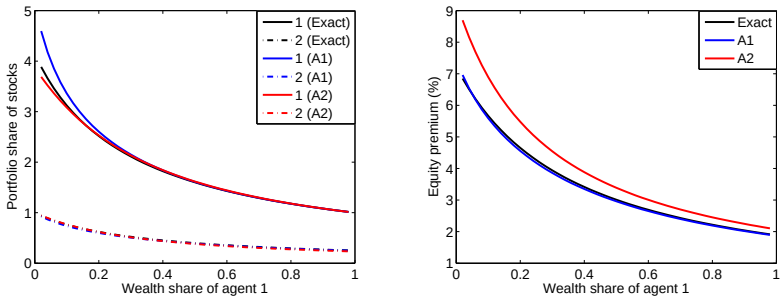
The results are shown in Figure 1. `A1` and `A2` refer to the approximate model using (2.5) and (2.6), respectively. According to Figure 1a, the optimal portfolio of the exact and the two approximate models are close, at least when the wealth share of the risk tolerant agent H is not too small. As the risk tolerant agent gets richer, the risk averse agent's portfolio share of stock declines. Essentially agent H is providing insurance to agent L.

According to Figure 1b, contrary to the case with portfolios the equilibrium equity premium of `A2` is not so accurate. The approximation error can be up to 2% in magnitude. However, the approximation `A1` is virtually indistinguishable from the exact model. As the risk tolerant agent gets richer, there is more demand for borrowing, and therefore the risk-free rate increases in order to clear the market. In this example since the expected stock return is fixed at 7%, the equity premium shrinks as the risk tolerant agent gets richer.

2.3.2 Two agents with identical DRRA utilities

Consider the same example as above except that preferences are identical and exhibit decreasing relative risk aversion (DRRA). It is natural to assume that the Arrow-Pratt measure of relative risk aversion is a decreasing power function,

$$\mathrm{RRA}(x) = -\frac{xu''(x)}{u'(x)} = \gamma \left(\frac{x}{c}\right)^{-\eta},$$



(a) Portfolio share of stock. (b) Equity premium.

Figure 1: Numerical example with heterogeneous CRRA preferences. `Exact`: numerical solution of exact model; `A1`, `A2`: semi-analytical solution of approximate model using (2.5) and (2.6).

where $\gamma, \eta, c > 0$ are parameters.[6] The economic interpretation of the parameters is that c is a reference point for wealth, γ is the relative risk aversion coefficient at this reference point, and η governs the speed (elasticity) at which RRA decreases. Solving the ordinary differential equation, it follows that the von Neumann-Morgenstern utility function is

$$u(x) = A \int_c^x \mathrm{e}^{\frac{\gamma}{\eta}\left(\frac{y}{c}\right)^{-\eta}} \mathrm{d}y + B,$$

where $A > 0$ and B are some constants. Since A and B merely define an affine transformation, they do not affect agents' behavior. Therefore, without loss of generality we may assume $A = 1$ and $B = 0$, so the utility function is

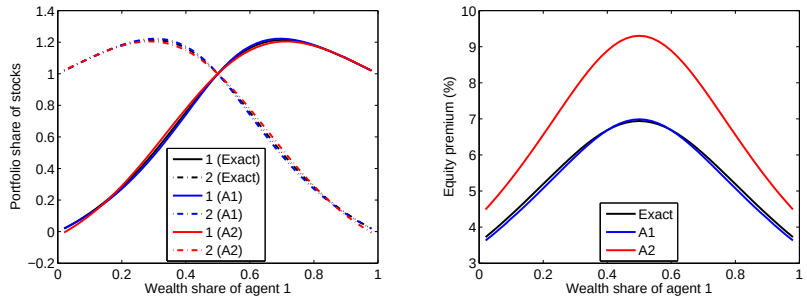
$$u(x) = \int_c^x \mathrm{e}^{\frac{\gamma}{\eta}\left(\frac{y}{c}\right)^{-\eta}} \mathrm{d}y. \tag{2.12}$$

For a numerical example, we normalize the aggregate wealth at $t = 0$ to be $W_0 = 1$ and set $\gamma = 2$, $\eta = 1$, and $c = 1/2$ (the reference point is equal distribution of wealth), so $u(x) = \int_{1/2}^x \mathrm{e}^{1/y} \mathrm{d}y$ and $\mathrm{RRA}(x) = 1/x$. Figure 2 shows the numerical solution. According to Figure 2a, as agent 1 gets richer, he becomes less risk averse and invests more in stocks. However, when he is too rich agent 2 is too poor to lend, and agent 1's portfolio share of stocks eventually decreases. According to Figure 2b, the equity premium is highest when the wealth is equally distributed. As the wealth distribution becomes more skewed, the richer and more risk tolerant agent leverages and drive down the equity premium. As in the previous example, the approximation `A1` is excellent but `A2` is poor.

3 Empirical link between inequality and equity premium

In Section 1 we argued that many models from the macro and finance literatures imply a close relationship between asset prices and the distribution of wealth,

[6] This specification has essentially two parameters, since only η and γc^η are identified. $\eta = 0$ corresponds to the CRRA case.



(a) Portfolio share of stock. (b) Equity premium.

Figure 2: Numerical example with identical DRRA preference. `Exact`: numerical solution of exact model; `A1`, `A2`: semi-analytical solution of approximate model using (2.5) and (2.6).

income, or assets. In Section 2 we analyzed some models and examples in which the extent of inequality across agents with heterogeneous risk aversion is key in predicting returns. We found not only that the wealth distribution affects the relative prices of risky assets but also that the extent of inequality may determine an economy's overall risk premium (and thus the equity premium).

But, are macroeconomic and financial data consistent with the implications of this paper's model and those in the above literature? In this section, we show that there is a strong and robust negative relationship between the top income/wealth share and subsequent medium-term excess stock market returns. The negative sign of the relationship is consistent with the above example, and the inequality measures do not seem to merely be proxying for either of two leading predictors of excess returns, price-dividend ratio (Shiller, 1981) and the consumption-wealth ratio (Lettau and Ludvigson, 2001a,b).

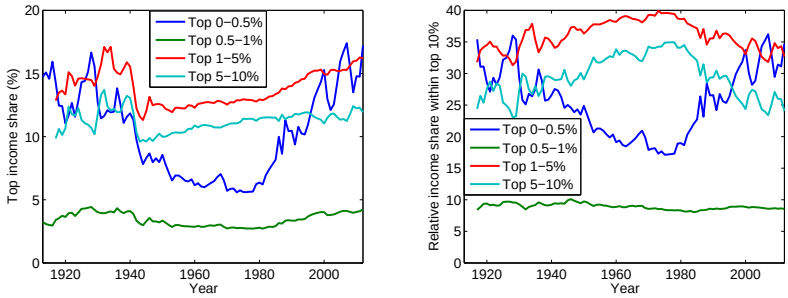
3.1 Data

We employ the Piketty and Saez (2003) inequality measures for the U.S., which are available in spreadsheets on the website of Emmanuel Saez.[7] In particular, we consider four top income share measures and one top wealth share measure. The income measures are at the annual frequency and are based on tax return data. The four income series are the top 1% with and without capital gains (cg) and the top 10% with and without capital gains (cg). The 1% series cover 1913–2012, and the 10% series cover 1917–2012. The wealth series, the top 1% wealth share, covers 1916–2000 at the annual frequency and is based on estate tax data. Many years are missing in the 50s, 60s, and 70s, so we complete the wealth series with cubic interpolation. The income series reflect in a given year the percent of income earned by the top 1 or 10% of earners pretax. Similarly, the wealth series is the percent of wealth owned by the richest 1%. See Piketty and Saez (2003) and Kopczuk and Saez (2004) for further details on the construction of these series.

Figure 3a shows the top income share for each group, the richest 0–0.5%, 0.5–1%, 1–5%, and 5–10%. All groups seem the share a common trend, which

[7] `http://elsa.berkeley.edu/~saez/`

is similar to the highest marginal tax rate in Figure 4. However, the behavior of these series around the trend is quite different. First, the top 0.5–1% share is very smooth. Second, the top 0–0.5% share seems to move in the same direction as booms and busts, which is most apparent in 1920s, 1960s, 1990s, and mid-2000s. On the other hand, the behavior of the top 1–5% and 5–10% are quite similar, and they seem to move in the *opposite* direction as booms and busts. To control for the trend (possibly related to tax rates), Figure 3b shows the relative income share of each group within the top 10%. We can see that the top 0.5–1% is stable, the top 0–0.5% moves in the same direction as booms and busts, and the top 1–10% moves in the opposite direction.



(a) Absolute shares. (b) Relative shares within the top 10%.

Figure 3: Top income shares including capital gains (1913–2012)

Within the context of the model in Section 2, this behavior can be explained if the richer agents are more risk tolerant. Consider, for example, the mean-variance model. Then the mutual fund theorem holds and agents invest more or less than 100% in stocks according as whether they are more or less risk tolerant than the average. If we assume that the top 10% hold the entire stock market, Figure 3b tells us that the top 0.5–1% roughly hold the market portfolio, the top 0–0.5% are more risk tolerant and leveraging (borrowing), and the top 1–10% are more risk averse and lend to the richest 0.5%.

Below, in both regressions and GMM exercises, we use not the raw Piketty-Saez series but rather detrended, stationary versions. Specifically, we detrend each of the five inequality measures using the Hodrick-Prescott (HP) filter with a smoothing parameter of 100, which is standard for annual frequencies.[8] Stochastically detrending asset return predictors is in the tradition of Campbell (1991), for example, who removes a trend in the short-term interest rate before including it in stock return vector autoregressions. Indeed, the Piketty-Saez series appear to exhibit a U-shaped trend over the century, which is likely due to the change in the marginal income tax rates. According to Figure 4, the marginal tax rate for the highest income earners increased from about 25% to 90% over the period 1930–1945 and started to decline in the 1960s, reaching about 40% in the 1980s.[9] Thus the marginal tax rate exhibits an inverse U-shape that coincides with the trend in the Piketty-Saez series. Imposing stationarity in this way helps ensure the validity of standard error calculations and prevent

[8] As a robustness check, we also vary the smoothing parameter and try the bandpass filter instead of the HP filter, but results are similar.

[9] The tax rate data is from the Tax Foundation (`http://taxfoundation.org/`).

spurious regressions. Figure 5 plots the top 1% series (with capital gains) and their estimated trends.

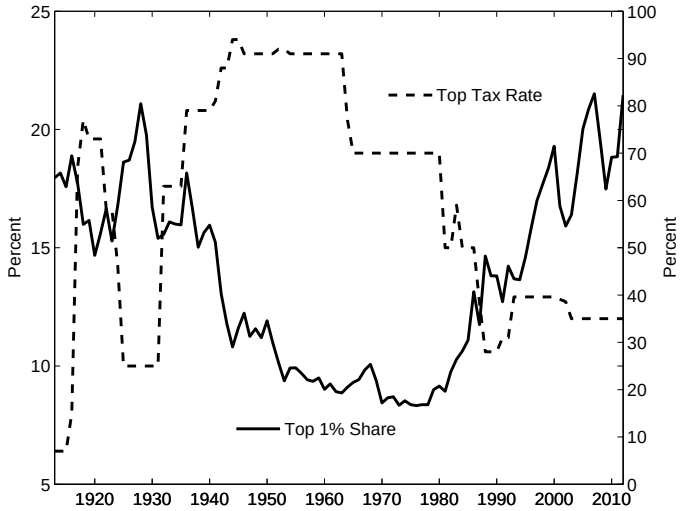


Figure 4: Top 1% income share including capital gains (left axis) and top marginal tax rate (right axis), 1913–2012.

We calculate annual one and five year U.S. stock market excess returns using the annual stock market spreadsheet from the website of Robert Shiller.[10] The spreadsheet contains historical one year interest rates and price, dividend, and earnings series for the S&P 500 index, which are all put into real terms using the consumer price index (CPI). These data are also used to calculate the series P/E10 and P/D10, which are the price-dividend and price-earnings ratios (in real terms) for the S&P 500 based on 10 year moving averages of earnings and dividends.

For the Lettau-Ludvigson consumption-wealth ratio, commonly referred to as CAY, we use 100 times the annual version of this series from the website of Amit Goyal (the spreadsheet for Welch and Goyal (2008)).[11] It spans the period 1945–2012.

Finally, we investigate the relationship between the top 1% income share and the three annual Fama-French asset pricing factors from the website of Kenneth French: EP, SMB, and HML (1927–2012).[12] EP is the value-weighted excess return on U.S. CRSP firms. SMB is, roughly, the difference in return between small and large valued firms. HML is, roughly, the difference in return between high and low book-to-market firms. See also Fama and French (1993).

[10] `http://www.econ.yale.edu/~shiller/data.htm`
[11] `http://www.hec.unil.ch/agoyal/`
[12] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html`

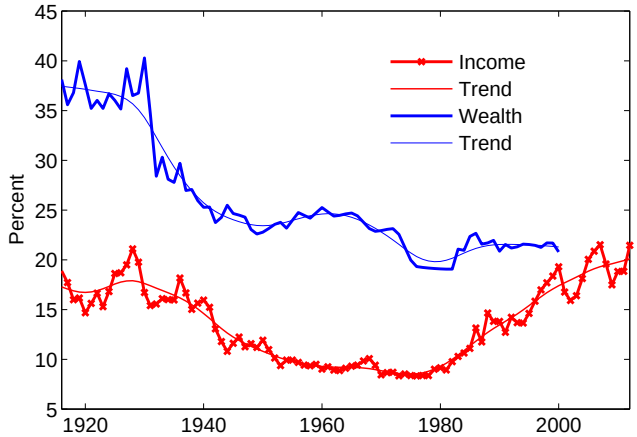


Figure 5: Top 1% income share including capital gains (1913–2012) and top 1% wealth share (1916–2000). The thin lines are the HP filter trends.

3.2 Regression analysis

Tables 1 and 2 show the results of regressions of one year (t to $t+1$) and five year (t to $t+5$) excess stock market returns, respectively, on top share measures (time t) and some classic return predictors (time t). We find that when the top 1% income share rises above trend by one percentage point, subsequent one year market excess returns decline on average by 5.6%. Five year returns decline by about 25%. These coefficients are both significant at the 1% level (using Newey-West standard errors), and the R-squared statistics are, respectively .09 and .20. Figure 6 shows the corresponding scatter plot for five year returns. It is clear, at least in sample, that the detrended top 1% share series has substantial power in forecasting the subsequent overall excess return on the stock market.

This relationship also holds for top 1% wealth share. With respect to one year returns, the top wealth share is strongly significant and yields an R-squared of .24, which is greater than the income share R-squared. With five year returns, the wealth share is significant at the 1% level but produces an R-squared of only .08.

Given the strength of the relationship, a question immediately arises. Is there some mechanical, non-equilibrium explanation for the relationship between inequality and subsequent excess returns? For example, might stock returns somehow be *determining* the top share measures? For two reasons, the answer is likely no. First, the relationship is between *initial* inequality and *subsequent* returns. Returns could affect contemporaneous top shares but likely not lagged top shares. Second, as we see in regression (2) from Tables 1 and 2, when excluding capital gains, the top 1% income share coefficient is *larger* with respect to five year returns and only slightly smaller with one year returns. If returns were strongly affecting lagged inequality, excluding capital gains would likely mitigate the regression results.

But, one might say, we have known at least since Shiller (1981) that when prices are high relative to either earnings or dividends, subsequent market excess returns are low. The current price could indeed affect current inequality (see

Section 3.3). Are the top shares series simply proxying for the price-dividend or price-earnings ratios, which are known to predict returns? Again, the answer seems to be no for two reasons. First, excluding capital gains from income does not significantly mitigate the relationship, and capital gains are the main avenue through which prices would determine inequality. Second, as we see in regressions (6) and (7) from Tables 1 and 2, top shares predict excess returns *even when controlling for the log price-dividend or price-earnings ratio*. Including these controls does decrease the top shares coefficients slightly, but they remain large and significant. In the case of one year returns, the P/D and P/E ratios are not significant after controlling for top income shares.

In regressions (8) and (9) from Tables 1 and 2, we also control for CAY, which Lettau and Ludvigson (2001a,b) show forecasts market excess returns. In the case of one year returns, including CAY has very little impact on the relationship between the top income share and subsequent returns. With respect to five year returns, however, including CAY substantially decreases the magnitude and significance of the relationship. One caveat is that including this control entails excluding data from 1913–1944.

In summary, the data appear consistent with our theory that an increasing concentration of income or wealth decreases the market risk premium.

Table 1: Regressions of one year excess stock market returns on top income shares and classic predictors

	Dependent Variable: t to $t+1$ Excess Stock Market Return								
Regressors (t)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	6.66	6.69	6.97	6.99	7.09	21.60	21.83	6.28	22.30
	(1.77)	(1.76)	(1.83)	(1.83)	(1.83)	(11.65)	(12.86)	(1.64)	(15.49)
Top 1%	-5.61***					-5.02***	-4.86***	-5.50***	-5.18***
	(1.60)					(1.63)	(1.71)	(1.56)	(1.52)
Top 1% (no cg)		-5.16**							
		(2.25)							
Top 10%			-2.09						
			(1.87)						
Top 10% (no cg)				-0.72					
				(2.24)					
Top 1% (wealth)					-7.93***				
					(1.91)				
log(P/D10)						-4.52			-4.55
						(3.40)			(4.31)
log(P/E10)							-5.62		
							(4.66)		
CAY								2.39***	2.24***
								(0.58)	(0.70)
Sample	1913-2012	1913-2012	1917-2012	1917-2012	1916-2000	1913-2012	1913-2012	1945-2012	1945-2012
R^2	.09	.05	.02	.00	.24	.10	.11	.22	.23

Newey-West standard errors in parentheses ($k = 4$)
***, **, and * indicate significance at 1%, 5%, and 10% levels (suppressed for constants)

Table 2: Regressions of five year excess stock market returns on top income shares and classic predictors

	Dependent Variable: t to $t+5$ Excess Stock Market Return								
Regressors (t)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	38.44	38.52	40.43	40.48	43.07	152.42	159.70	38.35	177.39
	(8.54)	(8.81)	(9.83)	(9.90)	(10.14)	(45.59)	(52.38)	(8.38)	(60.79)
Top 1%	-25.12***					-20.61**	-19.11**	-10.45♣	-7.67
	(8.15)					(8.07)	(8.18)	(7.76)	(6.91)
Top 1% (no cg)		-28.62**							
		(11.41)							
Top 10%			-5.25						
			(5.34)						
Top 10% (no cg)				-1.60					
				(4.55)					
Top 1% (wealth)					-14.23***				
					(5.20)				
log(P/D10)						-34.50**			-39.48**
						(13.05)			(16.35)
log(P/E10)							-44.92**		
							(18.55)		
CAY								12.35***	11.09***
								(1.96)	(2.79)
Sample	1913-	1913-	1917-	1917-	1916-	1913-	1913-	1945-	1945-
	-2008	-2008	-2008	-2008	-2000	-2008	-2008	-2008	-2008
R^2	.20	.16	.01	.00	.08	.30	.31	.38	.50

Newey-West standard errors in parentheses ($k = 4$)
***, **, and * indicate significance at 1%, 5%, and 10% levels (suppressed for constants)
♣ significant at 10% level with OLS standard error

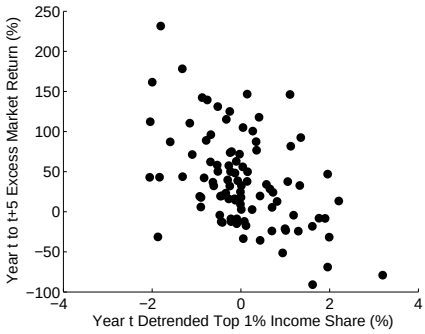


Figure 6: Year t to year $t+5$ excess stock market return vs. year t detrended top 1% income share including capital gains (1913–2008).

3.3 Relationship with classic return predictors

As we saw in Section 3.2, controlling for the price-dividend (or price-earnings ratio) or CAY mitigates to a small degree the estimated effect of inequality on subsequent excess returns. Furthermore, because the rich hold more stock than do the poor, high prices and the resulting capital gains likely have some direct impact on the top income shares. To what extent then are the top income shares correlated with classic return predictors? In Table 3, we regress the top 1% share and the top 10% share on a number of series known to predict or explain asset returns.

For the top 1% share, the correlation with the log price-dividend ratio is significant, but the R-squared is only .07. Therefore, while correlation with the price-dividend ratio likely explains some of the relationship between inequality and subsequent returns, it is not all or even most of the story. CAY, however, is not significantly correlated with the top 1% share. For the top 10% share, neither the price-dividend ratio nor CAY coefficient is significant. Overall, the top 1% income share appears to represent a component of the equity premium orthogonal to CAY and only slightly related to the price-dividend ratio.

Table 3 also displays regressions of the top income share on the Fama-French factors, which are powerful in explaining the cross section of average returns. Notably, the SMB coefficient is significant, and the R-squared is larger than in the price-dividend and CAY regressions.

This result opens up the possibility that Fama-French factors, which are economically hard to interpret, might actually be capturing the income/wealth inequality and hence the average risk aversion in the economy. The implication is that, across assets, heterogeneous correlation with top income shares may help explain heterogeneity in average returns. We explore this point in Section 4.

Table 3: Regressions of top income shares on classic return predictors

	Dependent Variable					
	Top 1% Share			Top 10% Share		
Regressors	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-1.74	-1.07	0.07	-0.84	-1.02	0.11
	(0.82)	(0.93)	(0.012)	(0.94)	(0.94)	(0.15)
log(P/D10)	0.53**	0.31		0.25	0.29	
	(0.25)	(0.28)		(0.26)	(0.28)	
CAY		-0.02			0.02	
		(0.05)			(0.05)	
EP			0.01			0.01
			(0.01)			(0.01)
SMB			-0.03**			-0.02*
			(0.01)			(0.01)
HML			-0.01			-0.02*
			(0.01)			(0.01)
Sample	1913-	1945-	1927-	1913-	1945-	1927-
	-2012	-2012	-2012	-2012	-2012	-2012
R^2	.07	.03	.13	.01	.03	.11

Newey-West standard errors in parentheses ($k = 4$)
***, **, and * indicate significance at 1%, 5%, and 10% levels
(suppressed for constants)

4 Testing the asset pricing implications for the cross section of returns

In this section we estimate and test a simplified version of the asset pricing implications derived in Section 2, namely the moment conditions (2.6), (2.5), and (2.9) with two agent types with risk tolerance $\tau_H > \tau_L$. For future reference, we refer to these models as follows.

`CRRA1` The moment condition derived from (2.6), which comes from the CRRA model with Taylor approximation at 1.

`CRRA2` The moment condition (2.10) derived from (2.5), which comes from the CRRA model with Taylor approximation at $\mathrm{E}[R(\theta_i^*)]$.

`MV` The moment condition derived from (2.9), which comes from the mean-variance model.

4.1 Data

Since the Piketty-Saez data is annual, we use annual asset returns data from 1927 to 2012. R_m is the CRSP value-weighted average portfolio return. R_j's are the Fama-French 2×3 portfolios sorted by size and book-to-market ratio. R_f is the annualized return of the 90 day T-Bill rate.[13] Nominal returns are converted

[13] Strictly speaking, we should use the 1 year bond return, but it is available only after 1941. We also estimate our model using the imputed value of the 1 year bond return before 1941 by regressing the 1 year bond return on a constant, 90 day T-bill rate, 30 day T-bill rate, and inflation, but the results were almost identical.

to real returns using the CPI. Asset returns and inflation data are from CRSP and Kenneth French's website. α is the detrended top 1% or 10% income share with capital gains.[14] Therefore α has roughly mean zero and moves around zero, so τ_L in (2.10) can be interpreted as the 'baseline' risk tolerance in the economy and $\Delta\tau = \tau_H - \tau_L$ can be interpreted as the sensitivity of the risk tolerance on the top income share.

4.2 Estimation

We estimate each model by GMM using the identity matrix for the first stage estimation and the Newey-West HAC estimator with 4 lags to compute standard errors. In estimating Models `CRRA2` or `MV`, we need to compute the expected market return $\mathrm{E}[R_m]$. For this purpose, inspired by regression (6) in Table 1 we regress R_m (year t to $t+1$) on a constant, top income share, and the logarithm of the price-dividend ratio of year t and define $\mathrm{E}[R_m]$ to be the OLS fitted value. We adopt the following time convention. In the case of Model `CRRA2`, for instance, we estimate

$$\mathrm{E}[((1 + \tau_L + (\tau_H - \tau_L)\alpha_t)\,\mathrm{E}[R_{m,t}] - R_{m,t})(R_{j,t} - R_{f,t})] = 0.$$

That is, we employ the detrended Piketty-Saez inequality measure α and asset returns of the same year. This is because since trades in assets occur at high frequency and the wealth distribution evolves simultaneously but we only observe α annually, we believe that α_t is a better proxy to the wealth distribution in the theoretical model of Section 2 than the lagged value α_{t-1}.

4.3 Results

4.3.1 GMM

Table 4 shows the results of the first stage GMM estimation of (2.10). (We do not report the efficient second stage estimation results since it is well-known that the finite sample property is poor.) The results are roughly the same across specifications. The estimated risk tolerance shows that the rich agents are nearly risk-neutral ($\gamma_H = 1/\tau_H \approx 0$) and the poor agents have a relative risk aversion coefficient $\gamma_L = 1/\tau_L$ in the range 2 to 3. The 'rich' and 'poor' here actually refer to the 'very rich' and 'ordinary stock market participant' in common language usage. Therefore our 'poor' agents are still relatively rich compared to the whole population. In any case, these results validate our claim that the 1% are more risk tolerant.

According to the J test (shown by P_J in Table 4), the models cannot be rejected, although this is not surprising given the small sample size ($T = 86$). According to the Wald test for testing homogeneous risk aversion (shown by $P(\tau_H = \tau_L)$ in Table 4), there is some evidence that risk aversion heterogeneity matters, as predicted by theory.

Figure 7 shows the scatter plot of predicted and realized excess returns of each Fama-French portfolio as well as the market portfolio, for each model with top 1% and top 10% income share (with capital gains). The scatter plot lies almost on the 45 degree line both with the top 1% and top 10%, which

[14] The results without capital gains are similar.

Table 4: 1st stage GMM estimation of the moment conditions.

Income share	Top 1%			Top 10%		
Model	CRRA1	CRRA2	MV	CRRA1	CRRA2	MV
τ_H (rich)	21.5	22.3	23.5	33.2	30.6	33.4
	(16.6)	(12.7)	(14.3)	(18.8)	(16.0)	(17.4)
τ_L (poor)	0.54	0.34	0.39	0.60	0.44	0.48
	(0.12)	(0.086)	(0.10)	(0.18)	(0.15)	(0.17)
$P(\tau_H = \tau_L)$	0.21	0.084	0.11	0.083	0.059	0.058
P_J	0.36	0.31	0.36	0.68	0.72	0.72

Newey-West standard errors in parentheses ($k = 4$)
Sample: 1927–2012

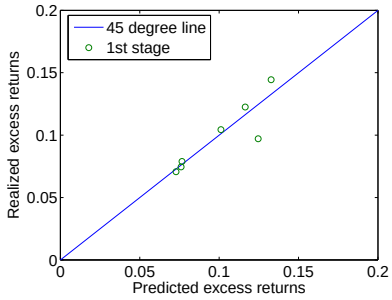
is especially true for the top 10%. However, when we estimate the standard CAPM by imposing $\tau_H = \tau_L$, the model fits poorly (Figure 8). These results suggest that preference heterogeneity indeed matters for asset pricing.

4.3.2 Covariance pricing

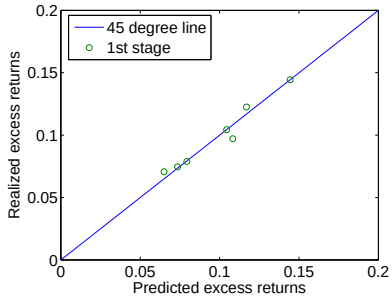
Table 5 provides further empirical support for the theoretical link between inequality and asset prices. Inspired by the covariance pricing formula (2.11), Table 5 displays regressions of the six Fama-French portfolio returns on the top 1% income share and the factor EP. Note that `sh` refers to "small size, high book-market," `bm` refers to "big size, medium book-market," etc. We draw two main conclusions from the results. First, consistent with the covariance pricing formula, both EP and the top 1% share are statistically significant in explaining the returns for most of the portfolios. In particular, EP is significant in all six regressions (the classic CAPM result), and the top 1% is significant (at at least the 10% level) for all portfolios but `bh` and `bm`.

Second, the signs of the top 1% coefficients are ordered exactly as the covariance pricing formula would have us expect. As is well known, small stocks have higher average returns than do big stocks, high book-market stocks have higher average returns than do low book-market stocks, and, unsurprisingly, `sh` has a much higher average return than does `bl` (19% vs. 11%). As we see in Table 5, the top 1% coefficients range from -5.28 for `sh` to 1.07 for `bl`. The `sl` coefficient is slightly smaller than the `sm` one, but the coefficients are otherwise ordered as expected. Exactly as in our model (assuming the rich are less risk averse), stocks negatively correlated with the top 1% have high average returns, and stocks positively correlated with the top 1% have low average returns.

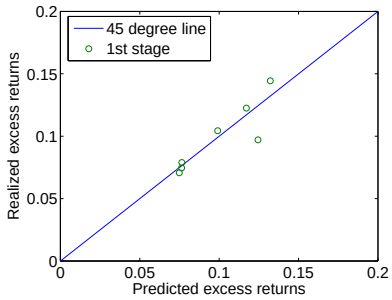
These results are somewhat weakened quantitatively when we include the HML factor, but the qualitative pattern remains. Furthermore, there is significant correlation between the top shares and SMB (see Table 3). Given these findings, the consistency with our model, and the fact that SMB does not have a clear economic interpretation, we suggest that the SMB factor is, to at least some degree, capturing average risk tolerance.



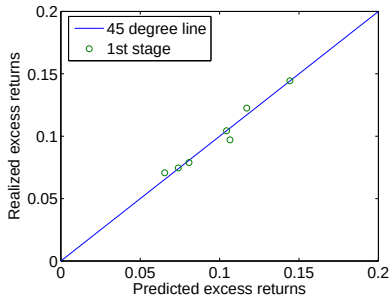
(a) Model `CRRA1` with top 1%



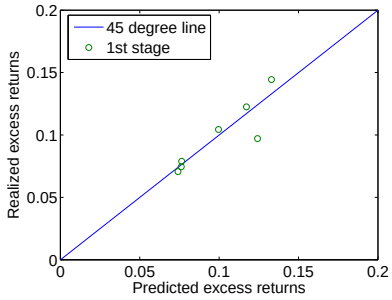
(b) Model `CRRA1` with top 10%



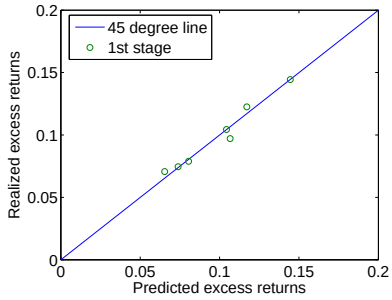
(c) Model `CRRA2` with top 1%



(d) Model `CRRA2` with top 10%



(e) Model `MV` with top 1%



(f) Model `MV` with top 10%

Figure 7: Predicted and realized excess returns with heterogeneous risk aversion.

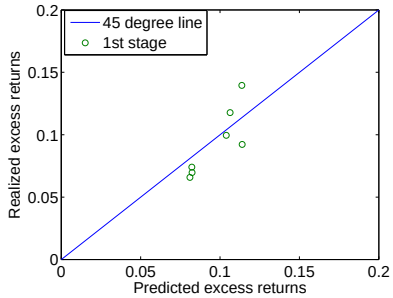


Figure 8: Predicted and realized excess returns with homogeneous risk aversion (standard CAPM).

Table 5: Regressions of Fama-French benchmark portfolios on top income shares and classic factors

	Dependent Variable											
	sh		sm		sl		bh		bm		bl	
Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	7.90	5.21	6.36	5.25	2.59	3.44	4.88	2.52	4.26	3.30	3.40	4.29
	(2.29)	(1.65)	(1.88)	(1.64)	(1.86)	(1.96)	(1.55)	(0.85)	(1.08)	(0.87)	(0.76)	(0.83)
Top 1%	**-5.38**	**-4.05**	**-3.28**	**-2.73**	**-3.52**	**-3.94**	-0.42	0.74	0.43	0.91	**1.07**	0.63
	(1.56)	(0.93)	(1.13)	(1.21)	(1.78)	(1.61)	(1.05)	(1.06)	(0.68)	(0.66)	(0.60)	(0.44)
EP	**1.36**	**1.27**	**1.26**	**1.22**	**1.39**	**1.42**	**1.21**	**1.13**	**0.96**	**0.93**	**0.94**	**0.97**
	(0.11)	(0.07)	(0.11)	(0.11)	(0.11)	(0.11)	(0.10)	(0.06)	(0.08)	(0.07)	(0.03)	(0.03)
HML		**0.79**		**0.33**		**-0.25**		**0.69**		**0.28**		**-0.26**
		(0.08)		(0.11)		(0.07)		(0.05)		(0.07)		(0.04)
R^2	.75	.86	.80	.82	.76	.77	.81	.93	.87	.90	.92	.95

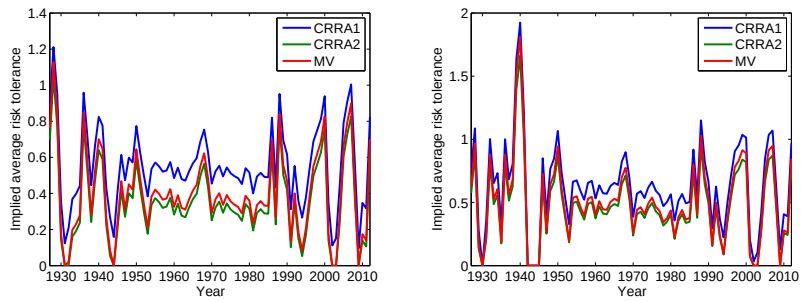
Newey-West standard errors in parentheses ($k = 4$)
Sample: 1927–2012
Boldface indicates significance at at least 10% level (suppressed for constants)

4.3.3 Average risk tolerance and aversion

Figure 9 shows the time series of the implied average risk tolerance

$$\bar{\tau} = \alpha\tau_H + (1-\alpha)\tau_L$$

for each model and top income share. We can see that the average risk tolerance of the economy is generally around 0.5 but rises in booms (1920s, 1960s, 1990s, around 2005) and approaches (sometimes hits) zero in busts, consistent with Fisher's story that booms and busts are associated with the wealth distribution and therefore the average risk tolerance of the economy.



(a) Top 1% (b) Top 10%

Figure 9: Time series of the implied average risk tolerance $\bar{\tau} = \alpha\tau_H + (1-\alpha)\tau_L$.

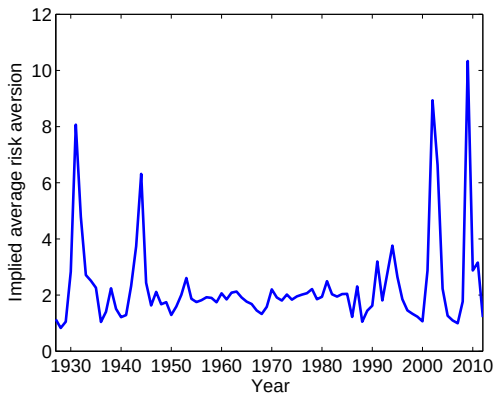


Figure 10: Time series of the implied average risk aversion of Model `CRRA1` with top 1% income share.

The Fisher story is most visible from the estimation of Model `CRRA1` using top 1% income share. Figure 10 shows the time series of the implied average risk aversion $\bar{\gamma} = 1/\bar{\tau}$ for this specification. We can see that the average risk aversion is usually around 2, but it sharply rises during bad times for the rich (the financial crises of early 1930 and 2007–2009, the introduction of exorbitant income tax during World War II, the collapse of the IT bubble in 2000–2002, and, to some extent, the income tax hike in 1993). If this pattern continues into the future, the stock market may be adversely affected in early 2014 due

to the 2013 marginal tax rate increase from 35% to 39.6%: This policy may decrease the wealth share of the 1% and thus raise the average risk aversion of the economy.

5 Concluding remarks

In this paper we found that the income/wealth distribution is closely connected with stock market returns. When the rich are richer than usual the stock market subsequently performs poorly. To explain this stylized fact, we built a simple general equilibrium model with agents that are heterogeneous in both wealth and attitudes towards risk. We then derived a testable moment condition as well as a new two factor covariance pricing formula. The formula tells us, essentially, that assets positively correlated with the top income share (and thus the average risk tolerance in the economy) command relatively low risk premiums. Our model is a mathematical formulation of Irving Fisher's narrative that booms and busts are caused by changes in the relative wealth of the rich (the "enterpriser-borrower") and the poor (the "creditor, the salaried man, or the laborer"). Overall, we find that our model is broadly consistent with the data.

Could one exploit the predictive power of top income shares to beat the market on average? The answer is probably no since the top income share—which comes from tax return data—is calculated with a substantial lag. One would receive the inequality update too late to act on its asset pricing information. However, our analysis provides a novel positive explanation of both market excess returns over time and the cross section of returns across stocks. We conclude, as decades of macro/finance theory have suggested, that stock market fluctuations are intimately tied to the distribution of wealth, income, and assets.

A Numerical algorithm

Before presenting concrete examples, we explain how to compute the equilibrium in the general case. Suppose that there are I agents and J risky assets. Interpret the risky assets as constant-returns-to-scale, stochastic savings technologies; let $\mathbf{R} = (R_1, \ldots, R_J)$ be the vector of gross returns with expected return $\mu = \mathrm{E}[\mathbf{R}]$ and variance-covariance matrix $\Sigma = \mathrm{E}[(\mathbf{R}-\mu)(\mathbf{R}-\mu)']$. The equilibrium objects are the portfolios $\{\theta_i^*\}_{i=1}^I$ and the risk-free rate R_f.

First we consider the approximation (2.6). By the budget constraint, we get

$$\frac{w_{i1}}{w_{i0}} = \sum_{j=1}^{J} R_j\theta_j + R_f\left(1 - \sum_{j=1}^{J}\theta_j\right) = R_f + \langle \mathbf{R} - R_f 1, \theta\rangle .$$

By the approximation of the first-order condition (2.4) with $a_i = w_{i0}$, we get

$$\begin{aligned}
&\mathrm{E}[(\tau_i + 1 - (R_f + \langle \mathbf{R} - R_f 1, \theta\rangle))(\mathbf{R} - R_f 1)] = 0 \\
\iff &(\tau_i + 1 - R_f)(\mu - R_f 1) - \mathrm{E}[(\mathbf{R} - R_f 1)(\mathbf{R} - R_f 1)']\theta = 0 \\
\iff &\theta_i^* = (\tau_i + 1 - R_f)[\Sigma + (\mu - R_f 1)(\mu - R_f 1)']^{-1}(\mu - R_f 1).
\end{aligned}$$

This equation shows that the mutual fund theorem holds. Taking the weighted average across agents and using market clearing, it follows that

$$\theta_m = (\bar{\tau} + 1 - R_f)[\Sigma + (\mu - R_f 1)(\mu - R_f 1)']^{-1}(\mu - R_f 1),$$

where θ_m is the market portfolio and $\bar{\tau}$ is the average risk tolerance. We can solve for the equilibrium by the shooting algorithm: given a risk-free rate R_f, we compute the market portfolio θ_m, raise the interest rate if $\sum_j \theta_{mj} > 1$ and cut otherwise. Then iterate until we get $\left|\sum_j \theta_{mj} - 1\right| < \varepsilon$, where ε is the error tolerance. We can use the risk-free rate computed from the mean-variance model as an initial guess: by (2.8) and $1'\theta_m = 1$, we obtain

$$R_f^0 = \frac{1'\Sigma^{-1}\mu - 1/\bar{\tau}}{1'\Sigma^{-1}1}.$$

Next we consider the approximation (2.5). Using the approximation of the first-order condition (2.4) with $a = a_i = \mathrm{E}[w_{i1}]$ and noting that

$$\frac{a_i}{w_{i0}} = \frac{\mathrm{E}[w_{i1}]}{w_{i0}} = R_f + \langle \mu - R_f 1, \theta \rangle,$$

we obtain

$$\begin{aligned}
&\mathrm{E}[(\tau_i(R_f + \langle \mu - R_f 1, \theta\rangle) - \langle \mathbf{R} - \mu, \theta\rangle)(\mathbf{R} - R_f 1)] = 0 \\
\iff &(\mathrm{E}[(\mathbf{R} - R_f 1)(\mathbf{R} - \mu)'] - \tau_i(\mu - R_f 1)(\mu - R_f 1)')\theta = \tau_i R_f(\mu - R_f 1) \\
\iff &\theta_i^* = \tau_i R_f[\Sigma - \tau_i(\mu - R_f 1)(\mu - R_f 1)']^{-1}(\mu - R_f 1).
\end{aligned}$$

In this case the mutual fund theorem does not hold, but we can still compute the equilibrium by the shooting algorithm. Starting from some R_f, for each agent we compute the optimal portfolio θ_i^* by the above formula. Then we compute the market portfolio $\theta_m = \sum_i w_{i0}\theta_i^* / \sum_i w_{i0}$ and raise or cut the interest rate according as $\sum_j \theta_{mj} \gtrless 1$.

Solving the exact model is similar, except that for each agent we have to numerically solve the optimal portfolio problem

$$\max_\theta \mathrm{E}[u_i((R_f + \langle \mathbf{R} - R_f 1, \theta\rangle)w_{i0})].$$

Letting $V(\theta)$ be the objective function, if the functional form of u_i' and u_i'' are explicitly known, we can solve this problem by the Newton algorithm since the gradient and the Hessian can be computed as

$$\begin{aligned}
\nabla V(\theta) &= \mathrm{E}[u_i'((R_f + \langle \mathbf{R} - R_f 1, \theta\rangle)w_{i0})(\mathbf{R} - R_f 1)], \\
\nabla^2 V(\theta) &= \mathrm{E}[u_i''((R_f + \langle \mathbf{R} - R_f 1, \theta\rangle)w_{i0})(\mathbf{R} - R_f 1)(\mathbf{R} - R_f 1)'],
\end{aligned}$$

respectively.

References

Pierluigi Balduzzi and Tong Yao. Testing heterogeneous-agent models: An alternative aggregation approach. *Journal of Monetary Economics*, 54(2): 369–412, March 2007. doi:10.1016/j.jmoneco.2005.08.021.

Harjoat S. Bhamra and Raman Uppal. Asset prices with heterogeneity in preferences and beliefs. *Review of Financial Studies*, 27(2):519–580, 2014. doi:10.1093/rfs/hht051.

Lawrence Blume and David Easley. If you're so smart, why aren't you rich? belief selection in complete and incomplete markets. *Econometrica*, 74(4): 929–966, July 2006. doi:10.1111/j.1468-0262.2006.00691.x.

Alon Brav, George M. Constantinides, and Christopher C. Geczy. Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy*, 110(4):793–824, August 2002. doi:10.1086/340776.

Douglas T. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296, September 1979. doi:10.1016/0304-405X(79)90016-3.

John Y. Campbell. A variance decomposition for stock returns. *Economic Journal*, 101(405):157–179, March 1991. doi:10.2307/2233809.

Yeung Lewis Chan and Leonid Kogan. Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy*, 110(6):1255–1285, December 2002. doi:10.1086/342806.

Daniele Coen-Pirani. Effects of differences in risk aversion on the distribution of wealth. *Macroeconomic Dynamics*, 8(5):617–632, 2004. doi:10.1017.S1365100504040052.

Timothy Cogley. Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey. *Journal of Monetary Economics*, 49(2): 309–334, March 2002. doi:10.1016/S0304-3932(01)00106-4.

George M. Constantinides. Intertemporal asset pricing with heterogeneous agents without demand aggregation. *Journal of Business*, 55(2):253–267, April 1982.

George M. Constantinides and Darrell Duffie. Asset pricing with heterogeneous consumers. *Journal of Political Economy*, 104(2):219–240, April 1996.

Bernard Dumas. Two-person dynamic equilibrium in the capital market. *Review of Financial Studies*, 2(2):157–188, 1989.

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, February 1993. doi:10.1016/0304-405X(93)90023-5.

Irving Fisher. *Introduction to Economic Science*. Macmillan, NY, 1910.

Nicolae Gârleanu and Stavros Panageas. Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. 2012.

John Geanakoplos and Martin Shubik. The capital asset pricing model as a general equilibrium with incomplete markets. *Geneva Papers on Risk and Insurance Theory*, 15(1):55–71, 1990. doi:10.1007/BF01498460.

Christian Gollier. Wealth inequality and asset pricing. *Review of Economic Studies*, 68(1):181–203, January 2001. doi:10.1111/1467-937X.00165.

William M. Gorman. Community preference fields. *Econometrica*, 21(1):63–80, January 1953.

Fatih Guvenen. A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1750, November 2009. doi:10.3982/ECTA6658.

Lars Peter Hansen and Kenneth J. Singleton. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy*, 91(2):249–265, April 1983.

Chiaki Hara, James Huang, and Christoph Kuzmics. Representative consumer's risk aversion and efficient risk-sharing rules. *Journal of Economic Theory*, 137 (1):652–672, November 2007. doi:10.1016/j.jet.2006.11.002.

Juan Carlos Hatchondo. A quantitative study of the role of wealth inequality on asset prices. *Federal Reserve Bank of Richmond Economic Quarterly*, 94 (1):73–96, 2008.

John Heaton and Deborah J. Lucas. Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy*, 104(3):443–487, June 1996.

Timothy C. Johnson. Inequality risk premia. *Journal of Monetary Economics*, 59(6):565–580, October 2012. doi:10.1016/j.jmoneco.2012.06.008.

Narayana R. Kocherlakota and Luigi Pistaferri. Asset pricing implications of Pareto optimality with private information. *Journal of Political Economy*, 117(3):555–590, June 2009. doi:10.1086/599761.

Wojciech Kopczuk and Emmanuel Saez. Top wealth shares in the United States, 1916–2000: Evidence from estate tax returns. *National Tax Journal*, 57(2): 445–487, 2004.

Per Krusell and Anthony A. Smith, Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998. doi:10.1086/250034.

Martin Lettau and Sydney C. Ludvigson. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, 56(3):815–849, June 2001a. doi:10.1111/0022-1082.00347.

Martin Lettau and Sydney C. Ludvigson. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(6):1238–1287, December 2001b. doi:10.1086/323282.

John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1):13–37, February 1965a.

John Lintner. Security prices, risk, and maximal gains from diversification. *Journal of Finance*, 20(4):587–615, December 1965b. doi:10.1111/j.1540-6261.1965.tb02930.x.

Francis A. Longstaff and Jiang Wang. Asset pricing and the credit market. *Review of Financial Studies*, 25(11):3169–3215, 2012. doi:10.1093/rfs/hhs086.

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

Sydney C. Ludvigson. Advances in consumption-based asset pricing: Empirical tests. In George M. Constantinides, Milton Harris, and René M. Stultz, editors, *Handbook of the Economics of Finance*, volume 2, chapter 12, pages 799–906. Elsevier, Amsterdam, 2013. doi:10.1016/B978-0-44-459406-8.00012-3.

N. Gregory Mankiw. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics*, 17(1):211–219, September 1986. doi:10.1016/0304-405X(86)90012-7.

Harry Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, March 1952. doi:10.1111/j.1540-6261.1952.tb01525.x.

Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2):145–161, March 1985. doi:10.1016/0304-3932(85)90061-3.

Thomas Piketty and Emmanuel Saez. Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1):1–41, 2003. doi:10.1162/00335530360535135.

William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, September 1964. doi:10.1111/j.1540-6261.1964.tb02865.x.

Robert J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–436, June 1981.

James Tobin. Liquidity preference as behavior towards risk. *Review of Economics and Statistics*, 25(2):65–86, February 1958.

Alexis Akira Toda. Asset pricing with heterogeneous preferences. 2012. URL `https://sites.google.com/site/aatoda111/papers`.

Alexis Akira Toda and Kieran Walsh. The double power law in consumption and implications for testing Euler equations. Revise and resubmit in *Journal of Political Economy*, 2013. URL `http://ssrn.com/abstract=2319454`.

Jiang Wang. The term structure of interest rates in a pure exchange economy with heterogeneous investors. *Journal of Financial Economics*, 41(1):75–110, May 1996. doi:10.1016/0304-405X(95)00854-8.

Ivo Welch and Amit Goyal. A compehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, 2008. doi:10.1093/rfs/hhm014.