

# The Equity Premium and the One Percent

**Alexis Akira Toda**

University of California San Diego

**Kieran James Walsh**

University of California Santa Barbara

We show that in a general equilibrium model with heterogeneity in risk aversion or belief, shifting wealth from an agent who holds comparatively fewer stocks to one who holds more reduces the equity premium. From an empirical view, the rich hold more stocks, so inequality should predict excess stock market returns. Consistent with our theory, we find that when the U.S. top (e.g., 1%) income share rises, subsequent 1-year excess market returns significantly decline. This negative relation is robust to controlling for classic return predictors, predicting out-of-sample, and instrumenting inequality with estate tax rate changes. It also holds in international markets. (*JEL* D31, D52, D53, F30, G12, G17)

Received March 1, 2017; editorial decision July 23, 2019 by Editor Stijn Van Nieuwerburgh. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

Does the wealth distribution matter for asset pricing? Intuition tells us that it does: as the rich get richer, they buy risky assets and drive up prices. Indeed, over a century ago, Fisher (1910) argued for the 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 theory of fluctuations were differences in preferences and wealth across people.

---

We benefited from comments from Daniel Andrei, Lint Barrage, Brendan Beare, Dan Cao, Vasco Carvalho, Peter Debaere, Graham Elliott, Nicolae Gârleanu, John Geanakoplos, Émilien Gouin-Bonenfant, Jim Hamilton, Gordon Hanson, Fumio Hayashi, Toshiaki Honda, Anton Korinek, George Korniotis, Tim Kehoe, Jiasun Li, Sydney Ludvigson, Semyon Malamud, Larry Schmidt, Allan Timmermann, Jessica Wachter, Frank Warnock, and Amir Yaron and seminar participants at Boston College, Cambridge-INET, Carleton, Federal Reserve Board of Governors, HEC Lausanne, Hitotsubashi ICS, Kyoto, Simon Fraser, Tokyo, UBC, UCSB, UCSD, UVA Darden, Vassar, Washington State, Yale, Yokohama National University, 2014 Australasian Finance and Banking Conference, 2014 Northern Finance Association Conference, 2015 Econometric Society World Congress, 2015 ICMAIF, 2015 Midwest Macro, 2015 SED, 2015 UVA-Richmond Fed Jamboree, 2017 AFA, and 2017 ICEF. We especially thank Snehal Banerjee, Daniel Greenwald, Xiameng Hua, and Stavros Panageas for detailed comments. Earlier drafts of this paper were circulated with the title “Asset Pricing and the One Percent.” Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Alexis Akira Toda, Department of Economics, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093; telephone: (858) 534-3383. E-mail: atoda@ucsd.edu.

*The Review of Financial Studies* 33 (2020) 3583–3623

© The Author(s) 2019. Published by Oxford University Press on behalf of The Society for Financial Studies. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

doi:10.1093/rfs/hhz121

Advance Access publication October 7, 2019

To see the intuition for why the wealth distribution affects asset pricing, consider an economy consisting of investors with different attitudes toward 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. Motivated by this intuition, we establish both the theoretical and empirical links between inequality and asset prices, which constitute two main contributions of this paper.

First, we theoretically study the asset pricing implications of general equilibrium models with heterogeneous agents. In a two-period economy populated by Epstein-Zin agents with arbitrary risk aversion, belief, and wealth heterogeneity, we prove there exists a unique equilibrium, in which increasing wealth concentration in the hands of stockholders leads to a decline in the equity premium. Although the inverse relationship between wealth concentration and risk premiums under heterogeneous risk aversion has been recognized at least since Dumas (1989) and recently emphasized by Gârleanu and Panageas (2015), testing the existing theory requires identifying the *preference* types, which is challenging. In contrast, we show that it is sufficient to identify the *portfolio* types (e.g., the type of agents that have larger portfolio shares of stocks). It does not matter why some agents hold more stocks: although we prove that high risk tolerance or optimism is sufficient for investing more in stocks, increased investment also could be due to other reasons, such as low participation costs. Furthermore, we calibrate this two-period model as well as an infinite horizon extension and show that the wealth distribution can have a quantitatively large effect on the equity premium.

Second, we empirically test our theoretical predictions. Given the existing empirical evidence that the rich invest relatively more in stocks,<sup>1</sup> theory implies that rising inequality should negatively predict subsequent excess stock market returns. Consistent with our theory, we find that when the income share of the top 1% income earners in the United States rises, the subsequent 1-year excess stock market return falls on average; therefore, current inequality appears to forecast the subsequent risk premium of the U.S. stock market.

Specifically, for a number of reasons including the apparent high persistence of top income and wealth shares, we employ a stationary component of inequality, "KGR" (capital gains ratio), defined as the difference between the top 1% income share with and without realized capital gains income, divided by the bottom 99% income share. We use data and theory to argue that KGR is a reasonable proxy for capital wealth and income inequality. Regressions of the year  $t$  to  $t+1$  excess return on the year  $t$  top 1% income share indicate a strong and significant negative correlation. Our evidence suggests that the top 1%

<sup>1</sup> See Haliassos and Bertaut (1995), Carroll (2002), Vissing-Jørgensen (2002), Campbell (2006), Wachter and Yogo (2010), Bucciol and Miniaci (2011), and Calvet and Sodini (2014).

income share is not simply a proxy for the price level, which existing research shows correlates with subsequent returns, or for aggregate consumption factors: the top 1% income share predicts excess returns even after we control for some classic return predictors, such as the price-dividend ratio (Fama and French 1988) and the consumption-wealth ratio (Lettau and Ludvigson 2001). Our findings are also robust to the inclusion of macro control variables, such as gross domestic product (GDP) growth. Using 5-year excess returns or the top 0.1% or 10% income share also yields similar results, although the predictability is really due to the top 1% (Table 4).

The empirical literature on return prediction is controversial. While many papers find evidence for return predictability,<sup>2</sup> Ang and Bekaert (2007) are more skeptical, and others point out econometric issues, such as small sample bias when regressors are persistent (Nelson and Kim 1993; Stambaugh 1999) and problems with overlapping data (Boudoukh, Richardson, and Whitelaw 2008; Valkanov 2003). In an influential study, Welch and Goyal (2008) show that excess return predictors suggested in the literature perform poorly out of sample. We apply the methodologies of McCracken (2007) and Hansen and Timmermann (2015) to show that including the top 1% as a predictor significantly decreases out-of-sample forecast errors relative to using the historical mean excess return (Table 5).

The finding that inequality predicts future returns is consistent with our theory and robust to the inclusion of controls and the construction of KGR. However, does more inequality *cause* lower returns, and is KGR actually reflecting inequality? To address causality, we use tax rate changes as an instrument. Because contemporaneous and lagged changes in top estate tax rates explain a substantial portion of the variation in KGR (Table 6), we estimate the effect of inequality on returns using generalized method of moments (GMM) with instrumental variables (Table 7). Including KGR, industrial production growth, and the log price-earnings ratio as endogenous explanatory variables and using lags of top estate tax rate changes and the log price-earnings ratio as instruments, top income shares are still significant in predicting excess returns. To address the concern that part of the variation in KGR is not due to inequality but rather the timing of realizing capital gains, we include 1-year-ahead changes in capital gains tax rates as an additional instrument to separately identify how the timing and inequality components predict returns (Table 9). The coefficient on the inequality component is negative and significant, while the timing coefficient is insignificant.

---

<sup>2</sup> Classic examples are the price-dividend ratio (Campbell and Shiller 1988; Cochrane 2008; Fama and French 1988; Hodrick 1992) and the consumption-wealth ratio (Lettau and Ludvigson 2001). Campbell and Thomson (2008) suggest that many economic variables predict returns by imposing weak restrictions, such as a nonnegative equity premium. Rapach, Strauss, and Zhou (2010) show that instead of using a single predictive regression model, combining forecasts significantly decreases the out-of-sample forecast errors. See Lettau and Ludvigson (2010) and Rapach and Zhou (2013) for reviews on forecasting stock returns.

We uncover a similar pattern in international data on inequality and financial markets: post-1969 cross-country fixed effects panel regressions suggest that when the top 1% income share rises by 1 percentage point, subsequent 1-year market returns significantly decline on average by 1%. However, this effect is not uniform across countries. Our theory suggests that for relatively “closed” economies, such as emerging markets, the domestic top 1% share should matter for asset pricing because domestic agents account for a substantial proportion of the universe of investors. However, for small open economies, the inequality among global investors (proxied by U.S. KGR) should matter because domestic agents compose only a small fraction of investors. Consistent with our theory, we find that the interaction terms between top income shares and home bias measures significantly predict stock returns. In an economy with complete home bias, a 1-percentage-point increase in the top 1% income share is associated with a subsequent 2.8% decline in stock market returns. In a small open economy (no home bias), a 1-percentage-point increase in U.S. KGR is associated with a subsequent decline in stock market returns of 4.7%.

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. Although the original capital asset pricing model (CAPM; see Sharpe 1964; Lintner 1965a, 1965b; Geanakoplos and Shubik 1990 for a general and rigorous treatment) allows for substantial heterogeneity in endowments and risk preferences across investors, the quadratic or mean-variance preferences admit aggregation and obviate the role of the wealth distribution. Inspired by the limited empirical fit of the CAPM and asset pricing puzzles that arise in representative agent models, since the 1980s theorists have extended macrofinance models to consider meaningful investor heterogeneity. Such heterogeneous agent models fall into two groups.

In the first group, agents have identical standard (constant relative-risk aversion, CRRA) preferences but are subject to uninsured idiosyncratic risks.<sup>3</sup> Although the models in this literature have had some success in explaining returns in calibrations, the empirical results (based on consumption panel data) are mixed and may even be spuriously caused by the heavy tails in the cross-sectional consumption distribution (Toda and Walsh 2015, 2017b). In the second group, markets are complete and agents have either heterogeneous CRRA preferences or identical but nonhomothetic preferences. 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, although

<sup>3</sup> Examples include Mankiw (1986), Constantinides and Duffie (1996), Heaton and Lucas (1996), Brav, Constantinides, and Geczy (2002), Cogley (2002), Balduzzi and Yao (2007), Storesletten, Telmer, and Yaron (2007), Kocherlakota and Pistaferri (2009), and Constantinides and Ghosh (2017). See Ludvigson (2013) for a review.

aggregation in the sense of Gorman (1953) fails. Therefore, agent heterogeneity should matter in discussions about asset pricing.

Gollier (2001) studies the asset pricing implication of wealth inequality among agents with identical preferences. He shows that more inequality increases (decreases) the equity premium if and only if agents' absolute risk tolerance is concave (convex). In particular, wealth inequality has no effect on asset pricing when agents have hyperbolic absolute risk aversion (HARA) preferences, for which the absolute risk tolerance is linear. Both he and Hatchondo (2008) also calibrate the model and find that the effect of wealth inequality on the equity premium is small.

Dumas (1989) solves a dynamic general equilibrium model with constant-returns-to-scale production and two agents (one with log utility and the other CRRA). He shows (proposition 17) that when the wealth share of the less risk-averse agent increases, then the risk-free rate goes up and the equity premium goes down. Although this prediction is similar to ours, he imposes an assumption on endogenous variables (see his equation (8)).

Following Dumas (1989), a large theoretical literature has studied the asset pricing implication of preference heterogeneity under complete markets.<sup>4</sup> All these papers characterize the equilibrium and asset prices by solving a planner's problem. However, this approach is not suitable for conducting comparative statics exercises of changing the wealth distribution, for two reasons. First, although by the first welfare theorem, for each equilibrium we can find Pareto weights such that the consumption allocation is the solution to the planner's problem, because in general the Pareto weights depend on the initial wealth distribution, changing the wealth distribution will change the Pareto weights and, consequently, the asset prices. However, in general it is difficult to predict how the Pareto weights change. Second, even if we can predict how the Pareto weights change, the possibility of multiple equilibria must be considered. In such cases the comparative statics often go in the opposite direction depending on the choice of the equilibrium. Thus, our results are quite different, because we prove the uniqueness of equilibrium and derive comparative statics with respect to the initial wealth distribution.

Gârleanu and Panageas (2015) study a continuous-time overlapping generations endowment economy with two agent types with Epstein-Zin preferences. Unlike other papers on asset pricing models with heterogeneous preferences, all agent types survive in the long run because of birth and death, and they also solve the model without appealing to a planner's problem. As a result, all endogenous variables are expressed as functions of the state variable, the consumption share of one agent type. They find that the concentration of wealth to the more risk-tolerant type ("the rich") tends to lower the equity premium. When the preferences are restricted to additive CRRA, then the

---

<sup>4</sup> Examples include Wang (1996), Chan and Kogan (2002), Hara, Huang, and Kuzmics (2007), Cvitanic et al. (2012), Longstaff and Wang (2012), and Bhamra and Uppal (2014) and the references therein.

relation between the consumption share and equity premium (more precisely, market price of risk) is monotonic (see their discussion on p. 10). Thus, our results are closely related to theirs, although different, because we prove more general comparative statics results (though in two-period models).

Our model is also related to the work on limited asset market participation, such as Basak and Cuoco (1998), Guvenen (2009), Chien, Cole, and Lustig (2011, 2012), and Chabakauri (2013, 2015). In these papers some agents do not participate in certain asset markets or face portfolio constraints, which affects the asset prices beyond the heterogeneity in preferences or beliefs. In our model we also have hand-to-mouth laborers, but because they do not participate in *any* asset market, we prove that their presence affects only the risk-free rate, not the equity premium (Theorem 2).

Although the wealth distribution theoretically affects asset prices, few empirical papers directly document this connection. To the best of our knowledge, Johnson (2012) and Campbell et al. (2016) are alone in exploring this issue. Using incomplete markets models, they show that top income shares or top income growth innovations are cross-sectional asset pricing factors. However, they do not investigate the ability of top income shares to predict excess market returns (our main empirical result).<sup>5</sup>

Lastly, our study is related to the findings of Greenwald, Lettau, and Ludvigson (2014), who identify innovations to wealth ( $e_{a,t}$ ) that explain much of the variation in the stock market and significantly predict low subsequent excess returns. In an equilibrium model, they show that  $e_{a,t}$  captures the risk tolerance of a representative stockholder. Interestingly, there is substantial correlation between  $e_{a,t}$  and our inequality predictor variable KGR.<sup>6</sup> In heterogeneous risk aversion models without aggregation, rising wealth concentration can effectively decrease the risk aversion of the corresponding representative stockholder/planner, so an interpretation of  $e_{a,t}$  is that it reflects the wealth share of relatively risk-tolerant stockholders versus more risk-averse ones.

## 1. Wealth Distribution and Equity Premium

In this section we present a theoretical model in which the wealth distribution across heterogeneous agents affects the equity premium.

### 1.1 Uniqueness of equilibrium

First, we consider a static model with agents that have heterogeneous but homothetic preferences and prove the uniqueness of equilibrium.

<sup>5</sup> Campbell et al. (2016) do explore market return prediction in their online appendix, but they do not uncover a relationship between the income of the rich and subsequent stock returns. Our findings are likely different, because they use income, not income share, and detrend top income linearly.

<sup>6</sup> We thank Daniel Greenwald for discovering this fact.

Consider a standard general equilibrium model with incomplete markets consisting of  $I$  agents and  $J$  assets (Geanakoplos 1990). Time is denoted by  $t=0, 1$ : agents trade assets at  $t=0$  and consume only at  $t=1$ . At  $t=1$ , there are  $S$  states denoted by  $s=1, \dots, S$ . Let  $A=(A_{sj}) \in \mathbb{R}^{S \times J}$  be the  $S \times J$  payoff matrix of assets,  $U_i: \mathbb{R}_+^S \rightarrow \mathbb{R}$  be agent  $i$ 's utility function, and  $n_i \in \mathbb{R}^J, e_i \in \mathbb{R}_+^S$  be agent  $i$ 's endowment vectors of asset shares at  $t=0$  and consumption goods in each state. By removing redundant assets, without loss of generality we may assume that the matrix  $A$  has full column rank.

Given the asset price  $q=(q_1, \dots, q_J)' \in \mathbb{R}^J$ , agent  $i$ 's utility maximization problem is

$$\underset{x, y}{\text{maximize}} \quad U_i(x) \quad (1a)$$

$$\text{subject to} \quad q'y \leq q'n_i, \quad x \leq e_i + Ay, \quad (1b)$$

where  $x \in \mathbb{R}_+^S$  denotes consumption and  $y=(y_1, \dots, y_J)' \in \mathbb{R}^J$  denotes the number of asset shares. Equation (1b) denotes the budget constraints at  $t=0, 1$ . A general equilibrium with incomplete markets (GEI) consists of asset prices  $q \in \mathbb{R}^J$ , consumption  $(x_i) \in \mathbb{R}_+^{SI}$ , and portfolios  $(y_i) \in \mathbb{R}^{JI}$  such that (1) agents optimize and (2) asset markets clear, so  $\sum_{i=1}^I y_i = \sum_{i=1}^I n_i$ .

We make the following assumptions.

**Assumption 1 (Homothetic, convex preferences).** For all  $i$ ,  $U_i: \mathbb{R}_+^S \rightarrow \mathbb{R}$  is continuous, strictly quasi concave, homogeneous of degree 1, differentiable on  $\mathbb{R}_{++}^S$ , and  $\nabla U_i(x) \gg 0$  with the Inada condition  $\partial U_i(x)/\partial x_s \rightarrow \infty$  as  $x_s \rightarrow 0$ .

Assumption 1 is standard in applied works. For example, the following CRRA utility satisfies this assumption:

$$U_i(x) = \begin{cases} \left( \sum_{s=1}^S \pi_{is} x_s^{1-\gamma_i} \right)^{\frac{1}{1-\gamma_i}}, & (\gamma_i \neq 1) \\ \exp \left( \sum_{s=1}^S \pi_{is} \log x_s \right), & (\gamma_i = 1) \end{cases} \quad (2)$$

where  $\gamma_i > 0$  is agent  $i$ 's relative-risk aversion (RRA) coefficient and  $\pi_{is} > 0$  is agent  $i$ 's subjective probability of state  $s$ .

**Assumption 2 (Tradability of endowments).** Agents' endowments are tradable: for all  $i$ ,  $e_i$  is spanned by the column vectors of  $A$ .

Under Assumption 2, because there exists  $y_i \in \mathbb{R}^J$  such that  $e_i = Ay_i$ , by redefining  $e_i$  to be zero and  $n_i$  to be  $n_i + y_i$ , without loss of generality we may assume  $e_i = 0$ , that is, agents are endowed only with assets.

**Assumption 3 (Collinear endowments).** Agents have collinear endowments: letting  $n = \sum_{i=1}^I n_i$  be the aggregate endowment of assets, we have  $n_i = w_i n$ , where  $w_i > 0$  is the wealth share of agent  $i$ , so  $\sum_{i=1}^I w_i = 1$ . Furthermore,  $An \gg 0$ .

Because  $e_i=0$  by assumption, the aggregate endowment of goods is  $An$ . Hence the assumption  $An \gg 0$  simply says that aggregate endowment is positive. While the collinearity assumption is strong, it is indispensable to guarantee the uniqueness of equilibrium.<sup>7</sup> With multiple equilibria, comparative statics may go in opposite directions, depending on the choice of equilibrium.

Under these assumptions, we can prove the uniqueness of GEI and obtain a complete characterization.

**Theorem 1.** Under Assumptions 1–3, there exists a unique GEI. The equilibrium portfolio ( $y_i$ ) is the solution to the planner’s problem

$$\begin{aligned} & \underset{(y_i) \in \mathbb{R}^{JI}}{\text{maximize}} && \sum_{i=1}^I w_i \log U_i(Ay_i) \\ & \text{subject to} && \sum_{i=1}^I y_i = n \end{aligned} \quad (3)$$

and consumption is  $x_i = Ay_i$ . Letting

$$L(y, q) = \sum_{i=1}^I w_i \log U_i(Ay_i) + q' \left( n - \sum_{i=1}^I y_i \right) \quad (4)$$

be the Lagrangian with Lagrange multiplier  $q$ , the equilibrium asset price is  $q$ .

Theorem 1 is essentially an incomplete-market version of the aggregation result in Chipman and Moore (1979). Equilibrium uniqueness is important for our purposes because it rules out unstable equilibria and thus allows for the below unambiguous comparative statics regarding the wealth distribution.<sup>8</sup>

## 1.2 Comparative statics for asset prices

To derive implications for asset pricing, we specialize the economy in Section 1.1 as follows.

In Section 1.1, we assumed that there is no consumption at  $t=0$ , but we can obtain similar results to Theorem 1 with consumption at  $t=0$  by interpreting  $t=0$  as a new “state” denoted by  $s=0$ .

<sup>7</sup> Mantel (1976) shows that if we drop collinear endowments, then even with homothetic preferences “anything goes” for the aggregate excess demand function, and hence there may be multiple equilibria. See Toda and Walsh (2017a) for concrete examples of multiple equilibria with canonical two-agent, two-state economies.

<sup>8</sup> See Kehoe (1998) and Geanakoplos and Walsh (2018) for further discussion of uniqueness in the presence of heterogeneous preferences.



**Assumption 4 (Epstein-Zin with unit EIS).** Agent  $i$ 's utility function is Epstein-Zin with unit elasticity of intertemporal substitution (EIS):

$$\log U_i(x) = (1 - \beta_i) \log x_0 + \frac{\beta_i}{1 - \gamma_i} \log \left( \sum_{s=1}^S \pi_{is} x_s^{1 - \gamma_i} \right), \quad (5)$$

where  $x_0$  is consumption at  $t=0$ ,  $x_s$  is consumption in state  $s$  at  $t=1$ ,  $\beta_i \in (0, 1)$  is the discount factor,  $\gamma_i > 0$  is RRA (the case  $\gamma_i = 1$  corresponds to log utility as usual), and  $\pi_{is} > 0$  is agent  $i$ 's subjective probability of state  $s$ .

Furthermore, we introduce an agent type that does not participate in the asset market. Let  $e_s$  ( $s=0, 1, \dots, S$ ) be the aggregate endowment in state  $s$ . Suppose a hand-to-mouth agent  $i=0$  (whom we call the laborer) is endowed with goods but does not trade assets. Let  $1 - \alpha_t$  be the fraction of aggregate income earned by the laborer at time  $t$ . Then the endowment of agent  $i \geq 1$  (whom we call the capitalist) in state  $s$  is  $\alpha_t w_i e_s$ , where  $t=0$  if  $s=0$  and  $t=1$  if  $s \geq 1$ .

Regarding the financial structure, suppose that there are only two financial assets, a stock (a claim to the aggregate endowment) and a risk-free bond. By interpreting  $t=0$  as state  $s=0$ , there are effectively three assets, the other being a claim to  $t=0$  consumption. Therefore, the asset structure is as follows.

**Assumption 5.** Let  $e_s > 0$  ( $s=0, 1, \dots, S$ ) be the aggregate endowment of goods in state  $s$ . The asset payoff matrix and the aggregate endowment of capitalists' assets are given by

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e_1 & 1 \\ \vdots & \vdots & \vdots \\ 0 & e_S & 1 \end{bmatrix}, \quad n = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} \alpha_0 e_0 \\ \alpha_1 \\ 0 \end{bmatrix}.$$

Under these assumptions, we can show that a redistribution of wealth from a bondholder to a stockholder reduces the equity premium, while redistribution between laborers and capitalists does not affect the equity premium. To make the statement precise, we introduce some additional notation.

Taking  $t=0$  consumption as the numéraire, let  $P$  be the ex-dividend price of the stock (the value of the claim  $(0, e_1, \dots, e_S)'$ ) and  $R_f$  the gross risk-free rate (reciprocal of the value of the claim  $(0, 1, \dots, 1)'$ ). The risk-free asset is in zero net supply, so the aggregate wealth of capitalists at  $t=0$  is  $\alpha_0 e_0 + \alpha_1 P$ . The wealth share of capitalist  $i$  is  $w_i$ , so the budget constraint is

$$y_{i1} + P y_{i2} + \frac{1}{R_f} y_{i3} = w_i (\alpha_0 e_0 + \alpha_1 P). \quad (6)$$

Let

$$\phi_i = \begin{bmatrix} \phi_{i1} \\ \phi_{i2} \\ \phi_{i3} \end{bmatrix} = \frac{1}{w_i (\alpha_0 e_0 + \alpha_1 P)} \begin{bmatrix} y_{i1} \\ P y_{i2} \\ y_{i3} / R_f \end{bmatrix}$$

be the vector of capitalist  $i$ 's portfolio shares.

Now we can state our main theoretical result, which characterizes the equilibrium allocation and derives implications to asset pricing.

**Theorem 2.** Under Assumptions 3–5, the followings are true.

1. There exists a unique equilibrium. Letting  $y_i = (y_{i1}, y_{i2}, y_{i3})'$  be capitalist  $i$ 's equilibrium asset holdings, we have

$$y_{i1} = \frac{w_i(1 - \beta_i)}{\sum_{i=1}^I w_i(1 - \beta_i)} \alpha_0 e_0, \quad (7)$$

and  $(y_{i2}, y_{i3})_{i=1}^I$  solves

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^I \frac{w_i \beta_i}{1 - \gamma_i} \log \left( \sum_{s=1}^S \pi_{is} (e_s y_{i2} + y_{i3})^{1 - \gamma_i} \right) \\ & \text{subject to} && \sum_{i=1}^I y_{i2} = \alpha_1, \quad \sum_{i=1}^I y_{i3} = 0. \end{aligned} \quad (8)$$

The equilibrium consumption allocation is given by  $x_{i0} = y_{i1}$  and  $x_{is} = e_s y_{i2} + y_{i3}$  for  $s = 1, \dots, S$ .

2. The equilibrium price-dividend ratio is given by

$$\frac{P}{e_0} = \frac{\alpha_0}{\alpha_1} \frac{\sum_{i=1}^I w_i \beta_i}{1 - \sum_{i=1}^I w_i \beta_i}. \quad (9)$$

Consequently, shifting wealth from an impatient agent (low  $\beta_i$ ) to a patient agent (high  $\beta_i$ ) increases the price-dividend ratio.

3. Let  $R = (e_1/P, \dots, e_S/P)'$  be the vector of gross stock returns,  $\pi = (\pi_1, \dots, \pi_S)'$  be any probability distribution, and  $\mu = \pi'(\log R - \log R_f)$  be the log equity premium. Then  $\mu$  is independent of the capitalists' income shares  $\alpha_0, \alpha_1$ . Shifting wealth from an agent who invests relatively more in the risk-free asset (high  $\phi_{i3}$ ) to an agent who invests relatively less (low  $\phi_{i3}$ ) reduces the log equity premium.

Intuitively, in an economy with financial assets, the equilibrium prices and risk premiums balance the agents' preferences and beliefs. Because the stock is the only saving vehicle in the aggregate (because the risk-free asset is in zero net supply), shifting wealth to a patient agent increases the demand for stocks, and, hence, its price rises. If wealth shifts into the hands of the natural stockholder (either the risk-tolerant or optimistic agent), everything else fixed, the aggregate demand for the stock increases. Hence for markets to clear, the risk premium must fall to counterbalance the new demand of these agents.

A surprising aspect of Theorem 2 is that the equity premium is independent of the capitalist/laborer income shares  $\alpha_0, \alpha_1$  and only depends on the wealth

distribution among capitalists,  $\{w_i\}_{i=1}^I$ . The intuition is that while  $\alpha_0, \alpha_1$  affect the overall level of asset prices and the risk-free rate by changing the relative income between  $t=0, 1$ , they do not affect the equity premium because assets are only held by capitalists; the equity premium balances the relative demand of stocks and bonds, which only depends on the capitalists' wealth distribution.

Who is the natural bondholder in Theorem 2? We can answer this question by reducing the individual problem to a static optimal portfolio problem. Because of unit EIS, we have  $y_{i1} = (1 - \beta_i)w_i(\alpha_0 e_0 + \alpha_1 P)$ . Therefore, the budget constraint (6) simplifies to

$$P y_{i2} + \frac{1}{R_f} y_{i3} = w_i \beta_i (\alpha_0 e_0 + \alpha_1 P).$$

Let  $\theta_i = \frac{\phi_{i2}}{\phi_{i2} + \phi_{i3}} = \frac{P y_{i2}}{w_i \beta_i (\alpha_0 e_0 + \alpha_1 P)}$  be capitalist  $i$ 's portfolio share of stocks within savings. With a slight abuse of notation, let  $e_1 = (e_1, \dots, e_S)' \in \mathbb{R}_{++}^S$  be the vector of aggregate endowments at  $t=1$ . Then the vector of gross returns is  $R = e_1 / P$ . Because agents have Epstein-Zin utility, the consumption-saving decision and the portfolio decision can be separated. By homotheticity, the optimal portfolio problem reduces to

$$\max_{\theta} E_i [u_i(R\theta + R_f(1 - \theta))], \quad (10)$$

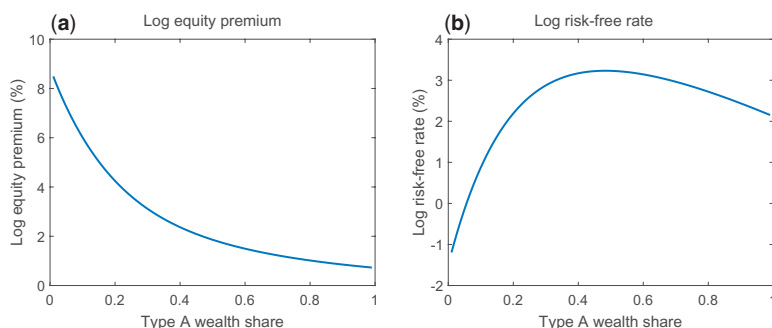
where  $u_i$  is the Bernoulli utility with relative-risk aversion  $\gamma_i$  (i.e.,  $u_i(x) = \frac{x^{1-\gamma_i}}{1-\gamma_i}$  if  $\gamma_i \neq 1$  and  $u_i(x) = \log x$  if  $\gamma_i = 1$ ) and  $E_i$  is the expectation under agent  $i$ 's belief.

The following propositions show that when agents have heterogeneous risk aversion or beliefs, the portfolio share of the risky asset  $\theta_i$  is ordered as risk tolerance or optimism. To define optimism, we take the following approach. First, by relabeling states if necessary, without loss of generality we may assume that states are ordered from bad to good ones:  $e_1 < \dots < e_S$ . Consider two agents  $i = 1, 2$  with subjective probability  $\pi_{is} > 0$ . We say that agent 1 is more pessimistic than agent 2 if the likelihood ratio  $\lambda_s := \pi_{1s} / \pi_{2s} > 0$  is monotonically decreasing:  $\lambda_1 \geq \dots \geq \lambda_S$ , with at least one strict inequality.

**Proposition 1.** Suppose Assumptions 3–5 hold and agents have common beliefs. If  $\gamma_1 > \dots > \gamma_I$ , then  $0 < \theta_1 < \dots < \theta_I$ .

**Proposition 2.** Suppose Assumptions 3–5 hold and agents 1, 2 have common risk aversion. Assume that agent 1 is more pessimistic than agent 2 in the above sense. Then  $\theta_1 < \theta_2$ .

Combining Theorem 2 together with either Proposition 1 or 2, provided that two agents have the same discount factor (hence the same  $\phi_{i1} = 1 - \beta_i$ ), shifting wealth from a more risk-averse or pessimistic agent to a more risk-tolerant or



**Figure 1**  
Wealth distribution and asset prices

an optimistic agent reduces the equity premium. In particular, if the rich are relatively more risk tolerant, optimistic, or simply more likely to buy risky assets (e.g., because of fixed stock market participation costs), rising inequality should forecast declining excess returns.

### 1.3 Numerical example

Theorem 2 tells us that the wealth distribution qualitatively affects asset prices, but does it matter quantitatively? To address this issue, we compute a numerical example calibrated at annual frequency. For simplicity we ignore the laborers, so  $\alpha_0 = \alpha_1 = 1$ . We specialize the above economy to one with two agents denoted by  $i = A, B$ . For preference parameters, let  $\rho_i > 0$  be the discount rate of agent  $i$  and define the discount factor by  $\beta_i = \frac{e^{-\rho_i}}{1 + e^{-\rho_i}}$ . We set  $(\rho_A, \rho_B) = (0.015, 0.06)$  and  $(\gamma_A, \gamma_B) = (1, 5)$ , so we can interpret type A as the “rich” (patient, risk tolerant) and type B as the “poor” (impatient, risk averse). The log dividend growth takes  $S = 3$  values  $g = (g_s) = (-0.3584, 0.0094, 0.2779)$  with probability  $\pi = (\pi_s) = (0.0549, 0.8552, 0.0899)$ .<sup>9</sup> Given these parameters, we can easily solve for the equilibrium by numerically solving the planner’s problem (8). Figures 1a and 1b show the log equity premium and the log risk-free rate, respectively. Consistent with Theorem 2, increasing the wealth share of type A agents monotonically decreases the equity premium. The equity premium ranges between about 8% and 1%, so the wealth distribution has a quantitatively large effect.

Our two-period model is arguably highly stylized. However, the advantage of a simple two-period model over a more complicated one is that we can prove theorems. In Appendix B, we extend our two-period model to infinite horizon with parameters calibrated to match various asset pricing moments and find that all of our qualitative results also hold in the more realistic setting.

<sup>9</sup> These preference parameters and the dividend growth distribution are taken from the infinite horizon calibration in Appendix B.

In Appendix C we simulate the calibrated infinite horizon model and run the same predictive regressions that we perform with actual data below in Section 2. The results from data and simulations are quantitatively similar. However, the model makes a prediction not evident in the data about one aspect. In the model, the type *A* stock wealth share  $w_A \theta_A / (w_A \theta_A + w_B \theta_B)$  negatively predicts returns, whereas in the data the relationship is near zero (or even positive). Therefore, inspired by (Fagereng et al. 2016a, 2016b) and Kacperczyk, Nosal, and Stevens (2019), in Appendix F we consider heterogeneous belief extensions to both our two-period and infinite horizon models: dividend growth is Markovian and type *A* is sophisticated and knows conditional probabilities of dividend growth, whereas type *B* is less sophisticated and uses unconditional probabilities. Because type *A* adjusts the portfolio based on the dividend growth state independent of the wealth distribution, the type *A* stock wealth share is no longer closely related to excess returns, as in the data. However, even in this extension, the wealth distribution channel (Theorem 2) survives.

## 2. Predictability of Returns with Inequality

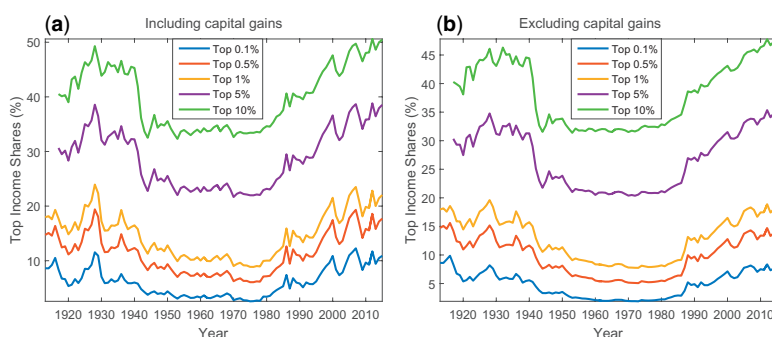
In Theorem 2, we have theoretically shown that shifting wealth from an agent who holds comparatively fewer stocks to one who holds more reduces the subsequent equity premium. Many empirical papers show that the rich hold relatively more stocks than do the poor and argue that the rich are relatively more risk tolerant (see footnote 1). Therefore, rising inequality should negatively predict subsequent excess stock market returns. In this section we construct a stationary measure of inequality and show that it predicts subsequent returns. (We address causality in Section 3.)

### 2.1 Connecting theory to empirics

**2.1.1 Empirical motivation.** The ideal way to test our theory is to run regressions of the form

$$\text{ExcessReturn}_{t \rightarrow t+1} = \alpha + \beta \times \text{WealthInequality}_t + \gamma \times \text{Controls}_t + \varepsilon_{t+1}$$

and test whether  $\beta = 0$ . Several obstacles must be overcome to implement this type of regression. First, measuring wealth is difficult and hence so is wealth inequality. The 1916–2000 top wealth share series (based on estate tax data) from Kopczuk and Saez (2004) are missing many years in the 50s, 60s, and 70s. The wealth share data of Saez and Zucman (2016) cover 1913–2012 but are *imputed* from capitalizing income. Second, our model considers redistribution among agents that participate in the stock market. Therefore, we should look at the inequality among stockholders, but the above wealth inequality measures consider all agents. Third, that these wealth measures are highly persistent introduces econometric problems. Fortunately, one can construct a stationary proxy measure of capitalist inequality via the method we will describe below.



**Figure 2**  
U.S. top income shares (1913–2015)

We employ the Piketty and Saez (2003) income inequality measures for the United States, which are available in updated form in a spreadsheet on Emmanuel Saez's Web site.<sup>10</sup> In particular, we consider top income share measures based on tax return data, which are at the annual frequency and cover the period 1913–2015. These series reflect in a given year the percentage of income earned by the top 1% of earners pretax. We also employ the top 0.1% share, the top 10% share, and the corresponding series that exclude realized capital gains income. Figure 2 shows these series, both including realized capital gains (Figure 2a) and excluding capital gains (Figure 2b). We can immediately see that all series share a common U-shaped trend over the century, and the series including capital gains are more volatile than those without capital gains.

Let  $\text{top}(p)$  ( $\text{top}(p)^{\text{excg}}$ ) be the top  $p\%$  income share including (excluding) capital gains. The fact that these two series share a common trend motivates us to consider their difference

$$\text{top}(p) - \text{top}(p)^{\text{excg}}. \quad (11)$$

Using Taylor approximation, we can connect this quantity to other measures of inequality, as described below.

**2.1.2 Constructing the inequality measure.** Suppose that there are two agent types in the economy, say the top 1% and the rest (bottom 99%). Let us denote these two types by  $i = A, B$ . Let  $Y_i^k, Y_i^l$  be the total capital and labor income of type  $i$  and  $Y_i = Y_i^k + Y_i^l$  be the total income of type  $i$ . Let  $Y^k = Y_A^k + Y_B^k$  and  $Y^l = Y_A^l + Y_B^l$  be the aggregate capital and labor income and  $Y = Y^k + Y^l = Y_A + Y_B$  be the total aggregate income. Then the top income share (type  $A$ 's income share) is

$$\text{top}(1) = \frac{Y_A}{Y} = \frac{Y_A^k + Y_A^l}{Y_A^k + Y_A^l + Y_B^k + Y_B^l}.$$

<sup>10</sup> <https://eml.berkeley.edu/~saez/>

Suppose that fraction  $\rho_i$  of type  $i$ 's capital income is comprised of realized capital gains. Then the top income share excluding realized capital gains is

$$\text{top}(1)^{\text{excg}} = \frac{(1 - \rho_A)Y_A^k + Y_A^l}{(1 - \rho_A)Y_A^k + Y_A^l + (1 - \rho_B)Y_B^k + Y_B^l} =: f(\rho_A, \rho_B).$$

Using Taylor's theorem, we can approximate the quantity in (11) as

$$\begin{aligned} \text{top}(1) - \text{top}(1)^{\text{excg}} &= f(0, 0) - f(\rho_A, \rho_B) \\ &\approx -\rho_A \frac{\partial f}{\partial \rho_A}(0, 0) - \rho_B \frac{\partial f}{\partial \rho_B}(0, 0) \\ &= \rho_A \frac{Y_A^k Y_B}{Y^2} - \rho_B \frac{Y_B^k Y_A}{Y^2}. \end{aligned} \quad (12)$$

Letting  $\alpha$  be the capital income share in aggregate income, so  $Y^k = \alpha Y$ , it follows from (12) that

$$\text{top}(1) - \text{top}(1)^{\text{excg}} \approx \alpha \left( \rho_A \frac{Y_A^k}{Y^k} (1 - \text{top}(1)) - \rho_B \left( 1 - \frac{Y_A^k}{Y^k} \right) \text{top}(1) \right). \quad (13)$$

Because  $Y_A^k/Y^k$  is the capital income share of the top 1%, who are more likely to be capital owners or entrepreneurs, it is reasonable to assume that the order of magnitude of  $Y_A^k/Y^k$  is at least that of  $1 - Y_A^k/Y^k$ . According to Figure 2a, the top 1% income share has evolved between 0.1 and 0.2, so  $1 - \text{top}(1) \gg \text{top}(1)$ . Therefore assuming that  $\rho_B$  is at most of the same order of magnitude as  $\rho_A$ , the second term inside the parenthesis of the right-hand side of (13) is much smaller than the first term. Ignoring the second term, we obtain

$$\begin{aligned} \text{top}(1) - \text{top}(1)^{\text{excg}} &\approx \alpha \rho_A \frac{Y_A^k}{Y^k} (1 - \text{top}(1)) \\ \iff \text{KGR}(1) &:= \frac{\text{top}(1) - \text{top}(1)^{\text{excg}}}{1 - \text{top}(1)} \approx \alpha \rho_A \frac{Y_A^k}{Y^k}. \end{aligned} \quad (14)$$

The left-hand side of (14), KGR(1) (we will explain the acronym shortly), is a quantity that can be calculated from top 1% income shares including/excluding realized capital gains. (14) says that it has three components:  $\alpha = Y^k/Y$  (the capital share of aggregate income),  $\rho_A$  (the fraction of realized capital gains income to total capital income for top earners), and  $Y_A^k/Y^k$  (the capital income share of top earners).

Before proceeding further, we need to make sure that the approximation (14) is empirically accurate. We address this issue in two ways. First, taking the logarithm of (14), we obtain

$$\log(\text{KGR}(1)) \approx \log \alpha + \log \rho_A + \log(Y_A^k/Y^k). \quad (15)$$

**Table 1**  
**Decomposition of KGR**

	$\log KGR(p)_t = \text{const.} + \beta' x_t + \varepsilon_t$					
	(1) $p=0.1\%$	(2) 1%	(3) 10%	(4) 1%	(5) 1%	(6) 1%
Constant	-0.07 (0.34)	-0.05 (0.28)	1.17 (0.39)	-5.37 (1.33)	-2.68 (0.17)	-2.67 (0.44)
$\log \alpha$	1.15*** (0.20)	0.97*** (0.18)	1.65*** (0.26)	-0.82 (0.80)		
$\log \rho_p$	0.98*** (0.09)	1.12*** (0.10)	1.25*** (0.10)		1.00*** (0.11)	
$\log(Y_p^k/Y^k)$	0.97*** (0.08)	1.25*** (0.19)	3.47*** (0.45)			1.87*** (0.55)
Sample	1922- -2012	1916- -2012	1962- -2012	1916- -2012	1916- -2012	1916- -2012
$R^2$	.94	.91	.95	.04	.78	.14

Newey-West standard errors are in parentheses (four lags). For  $p=0.1\%, 1\%, 10\%$ , the table shows regressions of  $\log(KGR(p))$  on its components according to Equation (14): the logs of the capital income share ( $\alpha$ ), the realized capital gain share of capital income ( $\rho_p$ ), and the top  $p\%$ 's share of capital income ( $Y_p^k/Y^k$ ; ranked by capital income including realized capital gains). \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$  (suppressed for constants).

It is possible to approximate all quantities in the right-hand side of (15) from the data of Saez and Zucman (2016), at least for the period 1916–2012.<sup>11</sup> To evaluate the accuracy of the approximation (14), inspired by (15), in Columns 1–3 of Table 1 we regress  $\log(KGR)$  on the logarithm of the three components for the top 0.1%, 1%, and 10% group. In each case  $R^2$  is above 0.9, which suggests that the three components  $\alpha$ ,  $\rho_A$ , and  $Y_A^k/Y^k$  explain almost all of the variation in KGR. Furthermore, consistent with (15), with top 0.1% and 1% the constant term is insignificant and the other three coefficients are statistically not different from 1. This result suggests that the approximation (14) is indeed accurate.

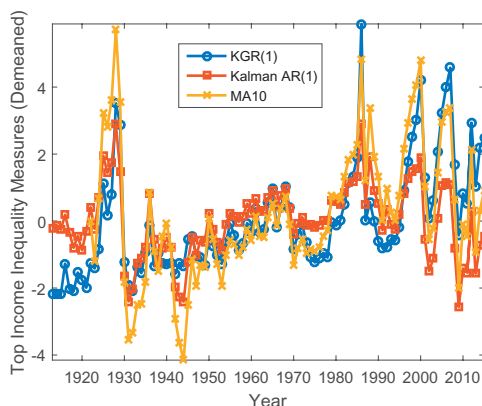
Second, we construct the two terms inside the parenthesis in (13). For the top 0.1% and 1%, the sample means for term 1 are 0.279 and 0.154, respectively, which are much larger than the term 2 means of 0.006 and 0.009. Additionally, the term 1 standard deviations are 0.165 and 0.091, respectively, while the term 2 standard deviations are 0.005 and 0.006. Therefore, for the top 0.1% and 1%, term 2 is indeed negligible relative to term 1, consistent with the accuracy of the approximation.<sup>12</sup>

Intuitively, what does  $KGR(1)$  measure? The first component,  $\alpha = Y^k/Y$ , is the capital share of income. The second component,  $\rho_A$ , reflects two factors. On the one hand, because capital gains realization is more likely when prices are

<sup>11</sup> Our sources are the “AppendixTables(Aggregates)” and “AppendixTables(Distributions)” spreadsheets for Saez and Zucman (2016). We compute  $\alpha$  as the sum of total positive business income, taxable interest, dividends, positive rents, estate and trust income, and net realized capital gains minus business and rental losses divided by total net taxable income.  $Y_A^k/Y^k$  is the series “Top taxable capital income shares, capital gains included in shares & rankings.” We compute  $\rho_A$  as the fraction of realized capital gains in the income of the top 1% divided by the fraction of capital income in the 1%’s income.

<sup>12</sup> For the top 10%, term 1 is larger and more volatile than is term 2, but the comparison is less stark: the term 1 and term 2 means are (0.053, 0.009), and the standard deviations are (0.028, 0.006).



**Figure 3****Time-series plot of inequality measures**

The graph shows the stationary component of the top 1% income share using KGR(1), using the AR(1) Kalman filter (both demeaned), and subtracting the 10-year moving average. The units of all series are percentage points.

high and because the rich disproportionately hold capital,  $\rho_A$  should positively correlate with wealth inequality. On the other hand,  $\rho_A$  could simply be capturing the timing of capital gains realization. We provide evidence *against* this second interpretation in Section 3. The last component,  $Y_A^k/Y^k$ , is capital income inequality.

To see which of the three components mainly determines KGR, in Columns 4–6 of Table 1, we regress  $\log(\text{KGR}(1))$  on the components one at a time. The  $R^2$  takes a large value 0.78 when we use only  $\rho_A$ , an intermediate value of 0.14 when we use  $Y_A^k/Y^k$ , and 0.04 when we use  $\alpha$ , whose coefficient is also insignificant. Therefore, KGR is mostly reflecting variation in  $\rho_A$ , the fraction of realized capital gains income to total capital income for top earners. This is why we chose the name KGR: capital gains ratio. To some extent KGR moves with capital income inequality  $Y_A^k/Y^k$ . The capital share  $\alpha$  does not appear to drive KGR. In summary, provided the timing component of  $\rho_A$  is not dominating (see Section 3), KGR measures capital wealth and income inequality.<sup>13</sup>

An advantage of KGR is its stationarity (the Phillips and Perron 1988  $p$ -values are less than .01 for the top 0.1%, 1%, and 10%). In contrast, the raw top wealth and income series appear nonstationary, or at least highly persistent, and thus introduce econometric problems when used to predict stationary returns (Granger 1981). KGR(1) looks very much like the detrended top 1% income share series. Figure 3 shows the KGR(1) series and the detrended versions of the raw top 1% series using the Kalman filter with an AR(1) cyclical component (see Appendix D for details) or subtracting the 10-year moving average.

<sup>13</sup> See Appendix E for more on the role of realized capital gains.

In light of Figures 2a and 2b and the capitalist/laborer share irrelevance in Theorem 2, one can argue for a noneconometric explanation of detrending as well. Because in the figures both the top 0.1% and 10% (and the shares in between) appear to have a common U-shaped trend, the slow-moving component of inequality is plausibly due to redistribution between the poor and rich uniformly rather than from intrarich redistribution. Assuming the poor/nonrich are less likely to participate in financial markets, then the trends in inequality correspond to changes in the capitalist/laborer income shares ( $\alpha_0, \alpha_1$ ) in Theorem 2, which are irrelevant for the equity premium according to our model. In this case, we would want to strip out the trends prior to predictive regressions. Intuitively, long-term trends in the capitalist and laborer income share should affect the overall level of asset markets, but not the equity premium per se, which only depends on the intracapitalist wealth distribution.

In the remainder of this section, we focus on the ability of KGR(1) defined by (14) to predict excess market returns in and out of sample.

## 2.2 In-sample predictions

We obtain the U.S. stock market returns, risk-free rates, and other financial variables from the spreadsheet of Welch and Goyal (2008).<sup>14</sup> Before 1926, stock returns are calculated from the S&P 500 index. After 1926, we use CRSP volume weighted average returns. We put returns into real terms using consumer price index (CPI) inflation, and returns are log returns.<sup>15</sup>

The series P/D and P/E are the price-dividend and price-earnings ratios for the S&P 500 index. The spreadsheet also contains the Lettau and Ludvigson (2001) consumption-wealth ratio, commonly referred to as CAY, which spans the period 1945–2015. For presentation, we multiply CAY by 100. Our other controls are GDP growth and, inspired by Lettau, Ludvigson, and Wachter (2008) and Bansal et al. (2014), consumption growth variance. Annual data for GDP and consumption come from the Web site of the Federal Reserve Bank of St. Louis (FRED)<sup>16</sup> and span 1930–2016. We estimate consumption growth variance using an AR(1)-GARCH(1,1) model for log consumption growth.

Table 2 shows the results of regressions of 1-year ( $t$  to  $t+1$ ) excess stock market returns on KGR(1) (time  $t$ ), some classic return predictors (time  $t$ ), and macro factors (time  $t$ ). Column 1 shows that a 1-percentage-point increase in year  $t$  KGR(1) predicts a 2.7% decline in year  $t$  to  $t+1$  excess market returns.<sup>17</sup>

<sup>14</sup> <http://www.hec.unil.ch/agoyal/>

<sup>15</sup> For example, if the gross return on stocks from year  $t$  to  $t+1$  is  $R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}$ , the log return is  $\log R_{t+1}$ . Similarly, the excess return refers to the log excess return  $\log R_{t \rightarrow t+1}^{\text{ex}} := \log R_{t+1} - \log R_{f,t}$ . In some of the specifications below, we use 5-year annualized returns, which are compounded annually:  $\log R_{t \rightarrow t+5} = \frac{1}{5} \sum_{k=1}^5 \log R_{t+k}$ .

<sup>16</sup> <http://research.stlouisfed.org/fred2/>

<sup>17</sup> Tables D5 and D6 (in Appendix D) show that the inverse relationship between inequality measures and subsequent excess returns also holds with the KGR(10) and KGR(0.1) series.

Table 2  
Regressions of 1-year excess stock market returns on KGR(1) and other predictors

	$\log R_{t \rightarrow t+1}^{\text{ex}} = \text{const.} + \beta' x_t + \varepsilon_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	11.92 (2.74)	11.30 (4.06)	17.30 (8.07)	9.10 (16.82)	14.65 (10.84)	13.59 (3.63)
KGR(1)	-2.69*** (1.00)	-2.70** (1.25)	-3.38* (1.76)	-2.89* (1.54)	-2.56** (1.12)	-2.79** (1.37)
$\Delta \log(\text{GDP})$		0.36 (0.48)				
$\log(\text{CGV})$			-2.15 (2.97)			
$\log(\text{P/D})$				0.99 (5.66)		
$\log(\text{P/E})$					-1.12 (4.21)	
CAY						1.25* (0.76)
Sample	1913- -2015	1930- -2015	1930- -2015	1913- -2015	1913- -2015	1945- -2015
$R^2$	.051	.055	.051	.051	.052	.117

Newey-West standard errors are in parentheses (four lags). KGR(1) is the proxy for top 1% capital inequality defined by (14).  $\Delta \log(\text{GDP})$  is real GDP growth. CGV is consumption growth volatility, which is estimated from an AR(1)-GARCH(1,1) model. P/D, S&P 500 price-dividend ratio; P/E, S&P 500 price-earnings ratio; CAY, consumption-wealth ratio. \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$  (suppressed for constants).

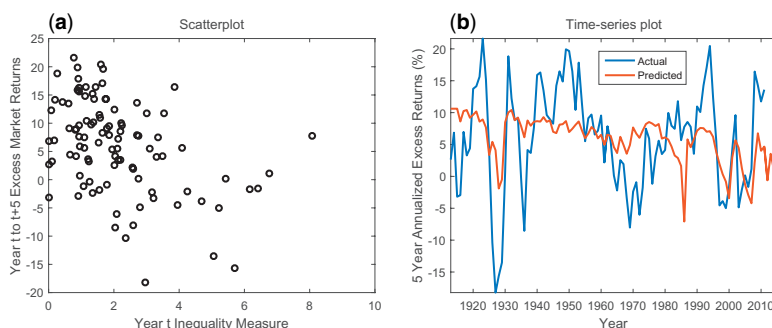
Table 3  
Regressions of 5-year excess stock market returns on KGR( $p$ )

	$\log R_{t \rightarrow t+5}^{\text{ex}} = \text{const.} + \beta \text{KGR}(p)_t + \varepsilon_{t+5}$		
	$p = 0.1\%$	1%	10%
Constant	10.20 (1.95)	10.61 (2.05)	12.25
KGR	-3.16** (1.25)	-2.19*** (0.84)	-2.06*** (0.65)
Sample	1913- -2015	1913- -2015	1917- -2015
$R^2$	.173	.181	.219

Newey-West standard errors are in parentheses (eight lags). Five-year excess returns are annualized. \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$  (suppressed for constants).

Table 3 shows that all three versions of KGR( $p$ ) also significantly predict 5-year excess returns. Figures 4a and 4b show the corresponding scatterplots and time-series plots for 5-year returns. KGR(1) forecasts subsequent 5-year excess returns well except around 1986.<sup>18</sup> Overall, a 1-percentage-point increase in KGR(1) is associated with, roughly, a 2%–4% decline in subsequent excess returns.

<sup>18</sup> The capital gains tax rate increased from 20% in 1986 to 28% in 1987 (but announced in 1986), which gave investors an incentive to realize capital gains in 1986. In Section 3 we disentangle the inequality and timing components of KGR(1).



**Figure 4**  
Year  $t$  to  $t+5$  excess stock market return (annualized) versus year  $t$  KGR(1), 1913–2015

Across specifications throughout the paper, KGR(10) performs relatively well in forecasting excess returns. This is surprising at first glance: the top 10% hold most financial wealth, and we argued in Section 2.1 that *intracapitalist* inequality is what should matter for excess returns. A straightforward explanation, however, is that KGR(10) overlaps with KGR(1) and KGR(0.1) and contains the predictive power of higher income shares. In Table 4, we predict returns with intra-10% KGR analogs, for example, KGR(5–10) corresponding to the top 5–10% income share. These new variables form a decomposition of KGR(10), because regressing it on KGR(0.1), ..., KGR(5–10) yields an  $R^2$  of over .99 (with all regressors strongly significant). Consistent with our theory, it is the highest share series (top 0.1–0.5% and top 0.1%) with the largest coefficients (regressors are standardized in Table 4) and  $R^2$ . The top 1–5% KGR is not significant, whereas KGR(5–10) has a *positive* coefficient significant at the 10% level. Hence, while the KGR(0.1) component of KGR(10) inversely forecasts excess returns, the KGR(5–10) component does the opposite, consistent with our story that redistribution to poorer capitalists *increases* excess returns.

Given the strength of the relationship between inequality and subsequent excess returns, a question immediately arises: is there some mechanical, nonequilibrium explanation? For example, might stock returns be *determining* the top share measures? For a few reasons, the answer is no. First, the relationship is between *initial* inequality and *subsequent* returns. Returns could affect contemporaneous top shares, but not lagged top shares. One might still worry that our results are driven by our transformation of the top share series into KGR(1). However, as we see in Appendix D, we obtain similar results with other methods of creating stationary series.

However, one may argue that we have known at least since Campbell and Shiller (1988) and Fama and French (1988) that when prices are high relative to either earnings or dividends, subsequent excess market returns are low. The current price could affect current inequality. Are the KGR series simply

Table 4  
Regressions of 1-year excess stock market returns on KGR for finer income groups

	(1)	$\log R_{t \rightarrow t+1}^{\text{ex}} = \text{const.} + \beta \text{KGR}(p)_t + \varepsilon_{t+1}$				(6)
		(2)	(3)	(4)	(5)	
Constant	6.33 (1.88)	6.03 (1.82)	6.03 (1.86)	6.03 (1.85)	6.33 (1.83)	6.33 (1.82)
KGR(10)	−5.34*** (1.63)					
KGR(0.1)		−4.08** (1.70)				
KGR(0.1–0.5)			−4.41** (1.88)			
KGR(0.5–1)				−3.73** (1.81)		
KGR(1–5)					0.19 (2.14)	
KGR(5–10)						3.29* (1.73)
Sample	1917– 2015	1913– 2015	1913– 2015	1913– 2015	1917– 2015	1917– 2015
R <sup>2</sup>	.077	.045	.052	.037	.000	.029

KGR(*p*)’s are the subincome group components of KGR defined analogously to (14). For example,  $\text{KGR}(1-5) = (\text{top}(1-5) - \text{top}(1-5)^{\text{excg}}) \times \frac{1 - \text{top}(1)}{1 - \text{top}(5)}$ , which is the correct analog because KGR(5) is effectively KGR(0–5), and top(0)=0. KGR(*p*)’s are all standardized.

proxying for the price-dividend or price-earnings ratios, which are known to predict returns? Columns 4 and 5 from Table 2 show top shares predict excess returns even when controlling for the log price-dividend or price-earnings ratio. Including these controls barely affects the KGR(1) coefficient, which is large and significant. The P/D and P/E ratios, however, are not significant after controlling for top capital shares.

In Columns 2, 3, and 6 from Table 2, we also control for real GDP growth, consumption growth variance (Bansal et al. 2014; Lettau, Ludvigson, and Wachter 2008), and CAY, which Lettau and Ludvigson (2001) show forecasts excess market returns. Including these controls, we still see a strong relationship between the top income share and subsequent returns. Similar results hold for different percentiles of the top income share and detrending methods (Appendix D).

How do the components of KGR(*p*)— $\alpha$ ,  $\rho_p$ , and  $Y_p^k/Y^k$ —perform in forecasting excess returns? We see in Table D4 (Appendix D) that  $\rho_p$ , the primary driver of KGR(*p*) according to Table 1, significantly predicts lower excess returns, while  $\alpha$  and  $Y_p^k/Y^k$  are insignificant. Given that the realized capital gains component of KGR,  $\rho$ , is driving return prediction, a question arises. Are realized capital gains and not inequality per se predicting excess returns? We address this issue in Appendices C and E. First, we solve and calibrate an infinite horizon version of our model and simulate model versions of both KGR and a purely mechanical measure of aggregate realized capital gains. Our Monte Carlo experiment shows that in our calibrated model (1) KGR, mechanical capital gains, and wealth inequality all inversely forecast

excess returns, (2) all three variables are substantially correlated, and (3) KGR and mechanical capital gains have small sample properties better than those of wealth inequality.

In short, both KGR and mechanical realized capital gains are quantitatively reasonable proxies for wealth inequality, which our model shows inversely forecasts excess returns. Second, we run a horse race between KGR in the data and empirical measures of mechanical realized capital gains. We find that they perform similarly but provide evidence that the distribution of realized capital gains across income groups matters beyond the overall level in forecasting returns.

In summary, we advocate using KGR as our proxy for capitalist wealth inequality in testing our theory for three reasons. First, it is a reasonable proxy according to our calibrated model. Second, it is easily computed from frequently updated, publicly available data, and its computation requires no arbitrary parameters. Third, KGR is stationary and closely resembles detrended versions of income inequality, which has a very persistent component perhaps driven by forces we argue are less relevant for the equity premium. We elaborate on these points in Appendices C and E.

### 2.3 Out-of-sample predictions

So far, we have seen that the current top income share predicts future excess stock market returns in sample. However, Welch and Goyal (2008) have shown that the predictors suggested in the literature perform poorly out of sample, possibly because of model instability, data snooping, or publication bias. In this section, we explore the ability of the top income share (KGR in particular) to predict excess stock market returns out of sample.

Consider the predictive regression model for the equity premium,

$$y_{t+h} = \beta' x_t + \varepsilon_{t+h}, \quad (16)$$

where  $h$  is the forecast horizon (typically  $h = 1$ ),  $y_{t+h}$  is the year  $t$  to  $t+h$  excess stock market return,  $x_t$  is the vector of predictors,  $\varepsilon_{t+h}$  is the error term, and  $\beta$  is the population ordinary least squares (OLS) coefficient. Suppose that the predictors can be divided into two groups, so  $x_t = (x_{1t}, x_{2t})$  and  $\beta = (\beta_1, \beta_2)$  accordingly. In this section we are interested in whether the variables  $x_{2t}$  are useful in predicting  $y_{t+h}$ , that is, we want to test  $H_0: \beta_2 = 0$ . We call the model with  $\beta_2 = 0$  the NULL model and the one with  $\beta_2 \neq 0$  the ALT (for alternative) model.

To evaluate the performance of the ALT model against the null, following McCracken (2007) and Hansen and Timmermann (2015), we consider the following out-of-sample  $F$  statistic:

$$F = \frac{1}{\hat{\sigma}_\varepsilon^2} \sum_{t=\lfloor \rho T \rfloor + 1}^T [(y_{t+h} - \hat{y}_{t+h|t}^N)^2 - (y_{t+h} - \hat{y}_{t+h|t}^A)^2], \quad (17)$$

Table 5  
Out-of-sample performance in predicting 1-year excess returns

$\rho$	Predictor in the ALT model				
	KGR(1)	KGR(10)	KGR(0.1)	log(P/D)	log(P/E)
0.2	3.67*** (0.0040)	6.07*** (0.0010)	2.67** (0.0131)	-0.12 (0.1367)	0.77* (0.0515)
0.3	2.16** (0.0153)	3.19*** (0.0068)	1.43** (0.0436)	0.23 (0.1245)	1.34** (0.0360)
0.4	1.42** (0.0388)	2.94*** (0.0081)	0.64* (0.0901)	-0.42 (0.2781)	0.58* (0.0845)

$\rho=0.2, 0.3, 0.4$  is the proportion of observations set aside to compute an initial OLS estimate. Columns correspond to the predictors included in the ALT model in addition to a constant.  $KGR(p)$  is the proxy for top  $p\%$  capital inequality defined by Equation (14). The numbers in the table are the out-of-sample  $F$  statistic computed by Equation (17).  $p$ -values (in parentheses) are computed by simulating 10,000 realizations from the asymptotic distribution based on Hansen and Timmermann (2015) (one sided). \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

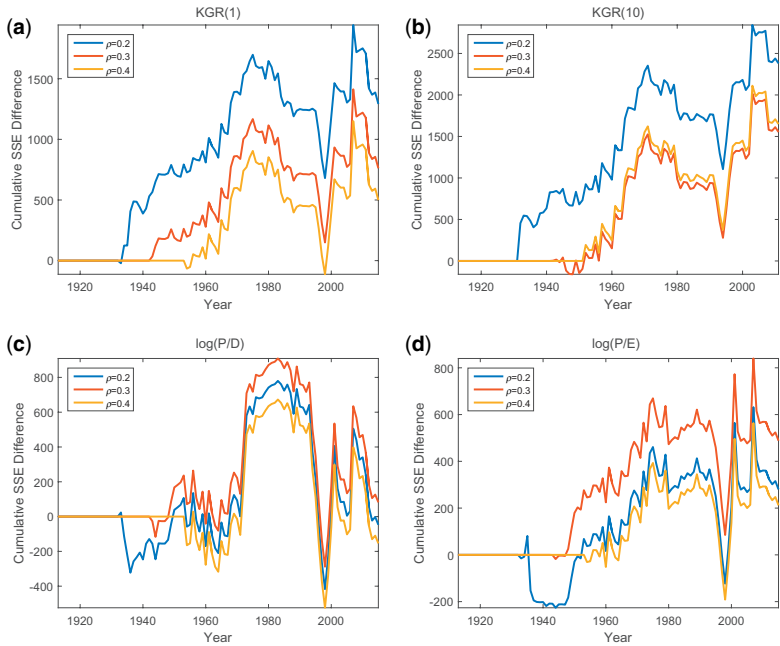
where  $\hat{\sigma}_\varepsilon^2$  is a consistent estimator of  $\text{Var}[\varepsilon_{t+h}]$  (which we estimate from the sample average of the squared OLS residuals of (16) using the whole sample),  $\hat{y}_{t+h|t}^A = \hat{\beta}'_t x_t$  ( $\hat{y}_{t+h|t}^N = \hat{\beta}'_{1t} x_t$ ) is the predicted value of  $y_{t+h}$  based on  $x_t$  using the ALT (NULL) model (here  $\hat{\beta}_t, \hat{\beta}_{1t}$  are the OLS estimator of (16) using data only up to time  $t$ ),  $T$  is the sample size, and  $0 < \rho < 1$  is the proportion of observations set aside for initial estimation of  $\beta$  and  $\beta_1$ . Theorems 3 and 4 of Hansen and Timmermann (2015) show that under the null ( $H_0: \beta_2 = 0$ ), the asymptotic distribution of  $F$  is a weighted sum of the difference of independent  $\chi^2(1)$  variables.

For the regressors in the ALT model, following Welch and Goyal (2008), we consider the simplest possible case in which  $x_{1t} \equiv 1$  (constant) and  $x_{2t}$  consists of a single predictor. For the predictor  $x_{2t}$ , we consider  $KGR(p)$  for  $p=0.1, 1, 10$  and valuation ratios ( $\log(P/D)$  and  $\log(P/E)$ ). The reason is that (1) because the top income series is at annual frequency, the sample size is already small at around 100 (1913 to 2015), so we cannot afford to use variables that are available only in shorter samples (e.g., CAY) for performing out-of-sample predictions, and (2) because Welch and Goyal (2008) find that most predictor variables suggested in the literature are poor, comparing many variables is pointless.

The choice of the proportion of the training sample,  $\rho$ , is necessarily subjective. Small  $\rho$  leads to imprecise initial estimates of  $\beta$ , and large  $\rho$  leads to the loss of power. Hence, we simply report results for  $\rho=0.2, 0.3, 0.4$ . According to Table 5, we can see that the out-of-sample  $F$  statistic is positive and significant when we use  $KGR(p)$ , while it is insignificant for  $\log(P/D)$  and weakly significant for  $\log(P/E)$ .<sup>19</sup>

To see this result graphically, in the spirit of Welch and Goyal (2008), we plot the difference in the cumulative sum of squared errors (the numerator of Equation (17)) over the prediction period in Figure 5. The vertical axis is the

<sup>19</sup> Because the asymptotic distribution of  $F$  depends on the NULL model, the relationship between the  $F$ -statistic in Table 5 and the  $p$ -values is not necessarily monotonic across models.



**Figure 5**  
**Annual performance in predicting subsequent excess returns**

The figures plot the out-of-sample performance of annual predictive regressions. The vertical axis is the cumulative squared prediction errors of the NULL model minus the cumulative squared prediction error of the ALT model (hence a positive value favors the ALT). The NULL model uses only a constant. The ALT model includes the predictor variables specified in each panel. Predictions start at  $t = \lfloor \rho T \rfloor$ , where  $T$  is the sample size and  $\rho = 0.2, 0.3, 0.4$ .

cumulative sum for the NULL model minus the ALT, so a positive value favors the ALT. We can see that for all  $KGR(p)$  specifications, the plots roughly monotonically increase up to 1980, decrease until 1990, and then increase again. This result is not surprising, since the 1980s were a time when income inequality increased but the stock market did not suffer (Figure 2a). On the other hand, the  $\log(P/D)$  and  $\log(P/E)$  specifications deteriorate after 1970, especially so for  $\log(P/D)$ . This finding is consistent with Welch and Goyal (2008), who document that most of the prediction gains stem from the 1973–1975 Oil Shock.

In summary, the top income series seem to predict returns out of sample.

### 3. Tax Instruments and Returns

The top 1% income share is an endogenous variable in the macroeconomy. While in Section 2.2 we showed that top income shares are not merely proxying for GDP growth, volatility, the consumption-wealth ratio, or the level of the stock market in explaining subsequent returns, it is difficult to rule out the possibility that omitted variables are leading to endogeneity bias.



In this section we formally address the causality from inequality to the equity premium by instrumental variables regressions. So far we have assumed that KGR is a measure of inequality due to variation in capital income, but other interpretations are possible. For example, KGR may be varying due to the timing of realizing capital gains. To address this issue, let KGR in year  $t$  be denoted by  $x_t$ , and suppose that it can be decomposed as

$$x_t = \alpha + x_{1t} + x_{2t},$$

where  $\alpha$  is a constant and  $x_{1t}, x_{2t}$  are zero mean variables that reflect inequality and timing (an incentive to realize capital gains), respectively. Consider the model

$$y_{t+1} = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_{t+1}, \quad (18)$$

where  $y_{t+1} = \log R_{t \rightarrow t+1}^{\text{ex}}$  is the log excess stock return from year  $t$  to  $t+1$ . (For notational simplicity we are omitting additional control variables, but it is straightforward to include them.) We are interested in testing  $\beta_1 = 0$ . The problem is that  $x_{1t}, x_{2t}$  are not observed separately.

To identify  $\beta_1$ , suppose that there is an instrument  $z_{1t}$  for  $x_{1t}$ , so (1)  $z_{1t}$  is exogenous (uncorrelated with  $\varepsilon_{t+1}$ ), (2)  $z_{1t}$  is correlated with  $x_{1t}$ , and, furthermore, (3)  $z_{1t}$  is uncorrelated with  $x_{2t}$ .

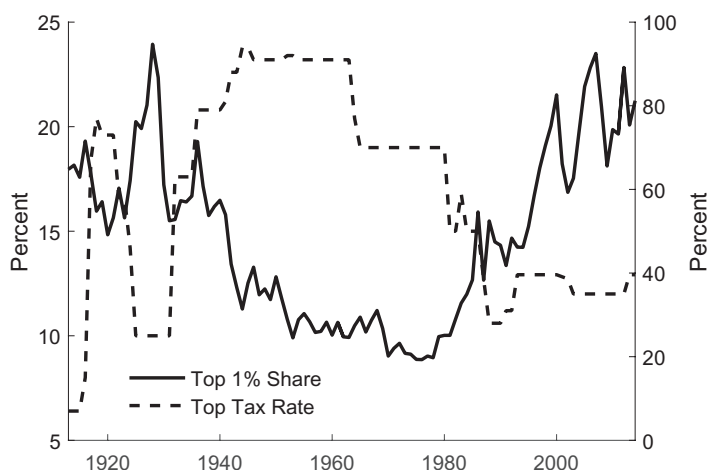
Then it follows that

$$\begin{aligned} 0 &= E[z_{1t} \varepsilon_{t+1}] = E[z_{1t} (y_{t+1} - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t})] \\ &= E[z_{1t} (y_{t+1} - \beta_0 - \beta_1 (x_t - \alpha - x_{2t}) - \beta_2 x_{2t})] \\ &= E[z_{1t} (y_{t+1} - \alpha_1 - \beta_1 x_t)], \end{aligned} \quad (19)$$

where  $\alpha_1 = \beta_0 + \alpha\beta_1$  and we have used  $E[z_{1t} x_{2t}] = 0$ . Therefore, even if the true inequality measure  $x_{1t}$  is unobserved, we can identify the coefficient of interest  $\beta_1$  by exploiting the moment condition (19).

To estimate the moment condition (19), we need an instrument for inequality. Research on inequality suggests that increases (decreases) in tax rates reduce (exacerbate) inequality (Kaymak and Poschke 2016; Roine, Vlachos, and Waldenström 2009). Indeed, the Piketty-Saez series appear to exhibit a U-shaped trend over the century, which might be due to the change in the marginal income tax rates. According to Figure 6, 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. Thus, the marginal tax rate exhibits an inverse U shape that seems to coincide with the trend in the Piketty-Saez series.

Both Piketty and Saez (2003) and Piketty (2003) argue that income inequality should decline in response to expansion of progressive estate taxation: capital gains compose a substantial portion of the income of the rich, and high estate taxes decrease the ability and incentive to amass wealth in financial assets. Thus, increasing the top estate tax rate should disproportionately reduce the wealth of



**Figure 6**  
**Top 1% income share including capital gains (left axis) and top marginal tax rate (right axis), 1913–2014**  
*Source: IRS.*

the very rich and subsequently mitigate capital gains income inequality, which is driven by inequality in asset holdings. On the other hand, estate taxes apply to both realized and unrealized capital gains, so estate taxes unlikely affect the timing of realizing capital gains beyond their incentive effects. Therefore, current and lagged changes in the estate tax rates are a good candidate for an instrument.

Tax rate changes are the result of Congressional bills, which generally take years to pass and usually stem from wars or pro-long-term growth or antideficit ideologies (de Rugy 2003a, 2003b; Jacobson, Raub, and Johnson 2007; Romer and Romer 2010; Weinzierl and Werker 2009). Therefore, while alterations in top tax rates affect inequality, their timing and justification are not the result of financial market fluctuations. Provided top tax rate changes have a muted effect on returns, except via inequality, they can serve as an instrument for top income shares.

The first-stage regressions in Table 6 confirm this hypothesis: contemporaneous and lagged changes in the top estate tax rate significantly explain a substantial portion of the variation in  $KGR(1)$  (and the 10% and 0.1% analogs).

Whether this instrument can test causation depends on the excludability of lagged changes in estate tax rates. One concern is that estate tax cuts stimulate the economy and thus stock market returns. Another concern is that even if estate tax rates only affect inequality, inequality may simply be proxying for the level of stock market, which we already know predicts returns. To control for these possibilities, we allow  $KGR$ , industrial production growth, and  $\log(P/E)$  to be endogenous and instrument all three with contemporaneous and three

**Table 6**  
**Regressions of KGR on contemporaneous and lagged changes in top estate tax rates**

	$KGR(p)_t = \text{const.} + \beta' x_t + \varepsilon_t$		
	$p=0.1\%$	1%	10%
Constant	1.52	2.37	3.11
$\Delta ETR_t$	-0.04***	-0.06***	-0.07***
$\Delta ETR_{t-1}$	-0.03**	-0.04*	-0.04*
$\Delta ETR_{t-2}$	-0.07***	-0.10***	-0.10***
$\Delta ETR_{t-3}$	-0.06***	-0.08***	-0.08***
$R^2$	0.26	0.24	0.19

The table shows regressions of KGR on lagged changes in top estate tax rates (ETR). Sample: 1913–2015. Sources: Tax Foundation and IRS. \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$  (suppressed for constants) according to Newey-West standard errors (four lags).

**Table 7**  
**Instrumental variables GMM estimates of the effect of KGR, industrial production growth, and log(P/E) on 1-year excess stock market returns**

	$\log R^{\text{ex}}_{t \rightarrow t+1} = \text{const.} + \beta' x_t + \varepsilon_{t+1}$		
	$p=0.1\%$	1%	10%
KGR( $p$ )	-10.79** (4.54)	-7.52** (3.27)	-6.91** (3.08)
% $\Delta$ IP	-1.51*** (0.51)	-1.49*** (0.49)	-1.46*** (0.48)
log(P/E)	3.71 (9.98)	2.61 (10.02)	1.90 (10.64)
$J$ -statistic	0.65 ( $p=.72$ )	0.69 ( $p=.71$ )	0.75 ( $p=.69$ )

The table shows the results of two-step GMM estimation of the moment condition (19) (including industrial production growth and log(P/E) as controls). The initial weighting matrix is identity, and the second-stage one is Newey-West (four lags). Newey-West standard errors are in parentheses (four lags). % $\Delta$ IP is the annual % change in the industrial production index. P/E is the S&P 500 price-earnings ratios. The instruments are a constant, changes in the top estate tax rate ( $\Delta ETR$  for  $t, t-1, t-2, t-3$ ), and the lagged price-earnings ratio ( $\log(P/E)_{t-1}$ ). Sample: 1913–2015. Sources: Tax Foundation, IRS, and FRED. \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$  (constants are suppressed).

lags of the change in the top estate tax rate ( $\Delta ETR$  for  $t, t-1, t-2, t-3$ ) as well as the lagged price-earnings ratio ( $\log(P/E)_{t-1}$ ).<sup>20</sup> Table 7 shows the results of GMM estimation of the moment condition (19) (including industrial production growth and log(P/E) as controls). KGR is significant at the 5% level in predicting subsequent excess returns regardless of whether we use the top 0.1%, 1%, or 10% income share in constructing KGR.

Our theory suggests that inequality predicts returns, but it does not say anything about the timing of realizing capital gains. Can we identify the coefficient  $\beta_2$  in (18)? Suppose that there is an additional instrument  $z_{2t}$  for  $x_{2t}$  that is uncorrelated with  $x_{1t}$ . By the same argument as the derivation of

<sup>20</sup> Industrial production growth ( $t$ ) is significantly correlated with  $\Delta ETR$  for  $t, t-1$ ;  $\log(P/E)_t$  is significantly correlated with  $\log(P/E)_{t-1}$ . Hence, the rank condition for identification holds.

**Table 8**  
**Regressions of KGR on contemporaneous and lagged changes in top estate tax rates and the one-period-ahead change in the capital gains tax rate**

	KGR( <i>p</i> ) <sub><i>t</i></sub> = const. + β' <sub>1</sub> <i>x</i> <sub><i>t</i></sub> + ε <sub><i>t</i></sub>		
	<i>p</i> = 0.1%	1%	10%
Constant	1.54	2.39	3.14
ΔETR <sub><i>t</i></sub>	−0.04***	−0.06***	−0.07***
ΔETR <sub><i>t</i>−1</sub>	−0.04***	−0.05**	−0.06**
ΔETR <sub><i>t</i>−2</sub>	−0.07***	−0.09***	−0.09**
ΔETR <sub><i>t</i>−3</sub>	−0.06***	−0.08***	−0.08***
ΔCGTR <sub><i>t</i>+1</sub>	0.03***	0.04***	0.05***
<i>R</i> <sup>2</sup>	.29	.27	.22

See caption of Table 6 for explanations. ΔCGTR<sub>*t*+1</sub> is the one-period-ahead change in the maximum capital gains tax rate.

(19), we can show that the moment condition

$$E[z_{2t}(y_{t+1} - \alpha_2 - \beta_2 x_t)] = 0 \tag{20}$$

holds, where  $\alpha_2 = \beta_0 + \alpha\beta_2$ . Rational agents have an incentive to realize (delay) capital gains if they expect the capital gains tax rate to increase (decrease). Tax rates in year *t* + 1 are announced in year *t*, so we can use the change in the maximum capital gains tax rate from year *t* to *t* + 1, ΔCGTR<sub>*t*+1</sub>, as an instrument *z*<sub>2*t*</sub> for the timing component of KGR, *x*<sub>2*t*</sub>. Table 8 adds ΔCGTR<sub>*t*+1</sub> to the first-stage regressions displayed in Table 6. As conjectured, the change in top capital gains tax rates from year *t* to *t* + 1 have positive and significant relationship with year *t* KGR. Current and lagged changes in estate tax rates, however, continue to have a strong inverse association with KGR. As rising capital gains and estate tax rates should, all else equal, discourage wealth accumulation among the rich, the positive coefficient on ΔCGTR<sub>*t*+1</sub> is likely reflecting the timing component of KGR (*x*<sub>2*t*</sub>): when the rich expect capital gains taxes to rise, they move forward the realization of capital gains, which causes KGR to rise.

Thus, in Table 9 we jointly estimate the moment conditions (19) and (20) (including industrial production growth, log(P/E), and ΔCGTR<sub>*t*+1</sub> as controls) by multiple equation GMM using the instruments

$$z_{1t} = (1, \Delta ETR_t, \Delta ETR_{t-1}, \Delta ETR_{t-2}, \Delta ETR_{t-3}, \log(P/E)_{t-1})', \tag{21a}$$

$$z_{2t} = (1, \Delta CGTR_{t+1})', \tag{21b}$$

respectively. The coefficients are positive but insignificant for the timing components (*x*<sub>2*t*</sub>) identified by changes in future capital gains tax rates. The inequality components (*x*<sub>1*t*</sub>), however, have negative and significant coefficients. This is true regardless of whether we use the top 0.1%, 1%, or 10% income share, and suggests that the causal effect of KGR on subsequent excess returns is driven by inequality rather than by the timing of capital gains realization.

In summary, our finding that rising top income shares lead to low subsequent excess returns is robust to instrumenting inequality with changes in estate

**Table 9**  
**Instrumental variables multiple equation GMM estimates of the effect of KGR, industrial production growth, and log(P/E) on 1-year excess stock market returns**

	$\log R_{t \rightarrow t+1}^{\text{ex}} = \text{const.} + \beta' x_t + \varepsilon_{t+1}$ $p=0.1\%$	1%	10%
KGR( $p$ ) (inequality, $\beta_1$ )	-10.93** (4.42)	-7.75** (3.19)	-7.17** (2.96)
KGR( $p$ ) (timing, $\beta_2$ )	11.45 (34.02)	6.43 (22.77)	5.92 (20.59)
% $\Delta$ IP	-1.48* (0.88)	-1.47* (0.85)	-1.42* (0.83)
log(P/E)	2.47 (10.13)	1.07 (10.22)	0.51 (10.74)
$\Delta$ CGTR $_{t+1}$	-0.05 (1.17)	0.01 (1.12)	-0.03 (1.13)
J-statistic	0.49 ( $p=.49$ )	0.50 ( $p=.48$ )	0.54 ( $p=.46$ )

The table shows the results of two-step multiple equation GMM estimation of the moment conditions (19) and (20) (including industrial production growth, log(P/E), and  $\Delta$ CGTR $_{t+1}$  as controls). Instruments for each equation are given by (21). See caption of Table 7 for more explanations.

tax rates, even when controlling for economic growth and the level of the stock market. Introducing one-period-ahead capital gains tax rate changes as an additional instrument, we are able to separately identify how the inequality and timing components of KGR affect returns. The predictive power of KGR established in Section 2 appears driven by the inequality component.

4. International Evidence

To show that the negative correlation between wealth inequality and returns is not specific to the United States, in this section we employ cross-country fixed effects panel regressions with unbalanced panel data for twenty-nine countries and the time period 1969–2015 as described in Appendix G.

Here, our analysis differs from our U.S. analysis in Section 2 in two main ways. The first is the availability of international risk-free rates. While U.S. Treasury returns provide a standard measure of the risk-free rate in the United States, in international markets, especially emerging ones where government and private sector default are not rare, the risk-free rate measure is controversial. Furthermore, because of the limited availability of similar interest rates across countries, using stock returns instead of excess returns substantially expands the sample size. Quantitatively, the predictability of *excess* returns in Section 2 was really about stock returns.<sup>21</sup> In light of these facts, we perform the international

<sup>21</sup> Redoing Column 1 of Table 2 with stock returns instead of excess returns, the KGR(1) coefficient is -2.61 with a Newey-West  $p$ -value of .027. With the 10% or 0.1% series, the coefficients are, respectively, -2.16 and -3.53 with  $p$ -values of .002 and .041. We do not find a significant correlation between inequality and risk-free rates in the United States for any of our top income share measures.

**Table 10**  
**Country fixed effects panel regressions of 1-year stock returns on domestic top income shares and U.S. KGR**

	$\log R_{t \rightarrow t+1} = \text{const.} + \beta' x_t + \varepsilon_{t+1}$			
	(1) All	(2) Advanced	(3) Ex-U.S.	(4) Ex-U.S.
Top 1%	-0.94* (0.52)	-1.01* (0.49)	-0.42 (0.70)	2.61 (1.55)
U.S. KGR(1)			-2.51*** (0.43)	-0.53 (0.75)
Top 1% × homebias				-5.44** (2.42)
U.S. KGR(1) × (1 - homebias)				-4.17** (1.60)
Country FE	✓	✓	✓	✓
Number of observations	815	712	769	687
$R^2$ (w,b)	(.00, .05)	(.01, .03)	(.02, .13)	(.03, .27)

Clustered standard errors are in parentheses. \*\*\*1%, \*\*5%, \*10% (constants suppressed). Ex-U.S.: all countries excluding U.S.  $R^2$  (w,b): within and between  $R^2$ . Top 1% is the share of income going to the top 1% of earners (the “fiscal income” top 1% series from World Inequality Database). The home bias measure comes from Mishra (2015). See the main text and Appendix G for country details on series construction. Sample: 1969–2015.

analysis using stock market returns without netting out an interest rate. The second difference is that the post-1969 sample shows no obvious U shape for top income shares. Furthermore, for many countries we cannot calculate KGR(1)<sup>22</sup> and the samples are short or missing years. Therefore, we use the raw top 1% income share data rather than remove time trends.

Table 10 presents the panel regression results for both the whole sample and different groups. We see in Column 1 that when including all countries, a 1-percentage-point increase in the top income share is associated with a subsequent decline in stock market returns of 0.94% on average. Column 2 uses the same regression, but for only advanced economies, and we obtain similar results.

Is the predictive power of the top income share uniform across countries? In either very large markets (such as the U.S. market) or relatively closed ones (such as emerging markets), our theory suggests that domestic inequality should affect domestic stock markets. In small open markets, however, foreign investors own a substantial fraction of the domestic stock markets and should mitigate the role of domestic inequality. However, even if domestic inequality is less important in small and open financial markets, inequality among global investors should still affect returns in these markets.

In Column 3 of Table 10, we include the U.S. KGR(1) as a proxy for global investor inequality because the United States is large. The U.S. KGR(1) is strongly significant but the domestic inequality is insignificant. However,

<sup>22</sup> As of this writing, the WID database does not provide the top 1% series excluding *and* including capital gains income for many countries.

domestic inequality should matter only for relatively closed economies, which have significant home bias. Therefore, in Column 4, we also include the interaction terms between inequality and home bias measures.<sup>23</sup> Specifically, we consider the model

$$\log R_{i,t+1} = \alpha_i + \beta_1 \text{top}(1)_{it} + \beta_2 \text{KGR}(1)_{\text{US},t} + \beta_3 \text{top}(1)_{it} \times \text{homebias}_i + \beta_4 \text{KGR}(1)_{\text{US},t} \times (1 - \text{homebias}_i) + \varepsilon_{i,t+1},$$

where  $\alpha_i$  is the country fixed effect. According to Column 4 of Table 10, the coefficients on the interaction terms ( $\beta_3, \beta_4$ ) are negative and strongly significant, whereas the linear terms ( $\beta_1, \beta_2$ ) are insignificant. In particular, in a closed economy ( $\text{homebias}=1$ ), a 1-percentage-point increase in the top 1% income share is associated with a subsequent decline in stock market returns of  $-\beta_1 - \beta_3 = 2.83\%$  on average; in a small open economy ( $\text{homebias}=0$ ), a 1-percentage-point increase in U.S. KGR(1) is associated with a subsequent decline in stock market returns of  $-\beta_2 - \beta_4 = 4.70\%$ . These findings are consistent with the conjecture that the domestic 1% share negatively predicts returns only for countries with higher home bias (relatively closed economies), and the global 1% share (proxied by U.S. KGR(1)) matters only for countries with lower home bias (small open economies).

## 5. Concluding Remarks

This paper builds a general equilibrium model with agents that are heterogeneous in wealth, risk aversion, and belief. We show that increasing inequality drives down the subsequent equity premium. 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"). Consistent with our theory, we find that in the United States, the wealth distribution is closely connected with stock market returns. When the rich are richer, the stock market subsequently performs poorly, both in and out of sample. The inverse relationship between returns and inequality is robust to controlling for standard return predictors and instrumenting with changes in estate taxes. It also holds internationally, although in relatively open economies with low home bias it is United States, not domestic, inequality that matters.

Could one exploit the predictive power of top income shares to beat the market on average? The answer is probably no, because the top income share based on 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,

<sup>23</sup> We use the ICAPM equity home bias measure of Mishra (2015) that takes a value between 0 (no home bias) and 1 (complete home bias). The home bias measure is not available for China, Ireland, Mauritius, and Taiwan.

our analysis provides a novel explanation of excess market returns over time. We conclude, as decades of macroeconomics and finance theory have suggested, that stock market fluctuations are intimately tied to the distribution of wealth and income.

## A. Proofs

### A.1 Proof of Theorem 1

As  $e_i = 0$  by Assumption 2, the utility maximization problem (1) becomes

$$\begin{aligned} & \underset{x, y}{\text{maximize}} && U_i(x) \\ & \text{subject to} && q' y \leq q' n_i, 0 \leq x \leq A y. \end{aligned} \quad (\text{A1a})$$

The proof proceeds as follows: we first solve the planner's problem, then show that the planner's solution is a GEI, and finally prove that this equilibrium is unique.

**Step 1.** The planner's problem (3) has a unique solution.

**Proof.** Let

$$\Omega = \left\{ x = (x_i) \in \mathbb{R}_+^{SI} \mid (\exists y = (y_i)) (\forall i) x_i \leq A y_i, \sum_{i=1}^I y_i = n \right\}$$

be the set of all feasible consumption allocations. Then the planner's problem (3) is equivalent to maximizing  $f(x) = \sum_{i=1}^I w_i \log U_i(x_i)$  subject to  $x \in \Omega$ . By Assumption 1 and Berge (1959, p. 208, theorem 3), each  $U_i(x_i)$  is strictly concave. Because  $\log(\cdot)$  is increasing and strictly concave, so is  $\log U_i(x_i)$ . Because  $f$  is continuous and strictly concave, the existence and uniqueness of a solution follows if we show that  $\Omega$  is nonempty, compact, and convex. Clearly,  $\Omega \neq \emptyset$  because we can choose the initial endowment  $y_i = n_i$  and  $x_i = A n_i = e_i$ . Because  $\Omega$  is defined by linear inequalities and equations, it is closed and convex. If  $x \in \Omega$ , by definition we can take  $y = (y_i)$  such that  $x_i \leq A y_i$  for all  $i$  and  $\sum_{i=1}^I y_i = n$ . Then

$$\sum_{i=1}^I x_i \leq \sum_{i=1}^I A y_i = A \sum_{i=1}^I y_i = A n.$$

Because  $x_i \geq 0$  and  $n \gg 0$ ,  $\Omega$  is bounded.

Let  $x = (x_i)$  be the unique maximizer of  $f$  on  $\Omega$ . Because  $f$  is strictly increasing, we have  $x_i = A y_i$  for some  $y = (y_i)$  such that  $\sum_{i=1}^I y_i = n$ . If there is another such  $y' = (y'_i)$ , then  $A y_i = A y'_i \iff A(y_i - y'_i) = 0$ . Because, by assumption,  $A$  has full column rank, we have  $y_i - y'_i = 0 \iff y_i = y'_i$ . Therefore, the planner's problem (3) has a unique solution. ■

**Step 2.**  $x = (x_i)$  is a GEI equilibrium allocation and the Lagrange multiplier to the planner's problem gives the asset prices.



**Proof.** Let

$$L(y, q) = \sum_{i=1}^I w_i \log U_i(Ay_i) + q' \left( n - \sum_{i=1}^I y_i \right)$$

be the Lagrangian of the planner's problem (3). By the previous step, a unique solution  $y = (y_i)$  exists. Furthermore, because  $U_i$  satisfies the Inada condition, it must be  $Ay_i \gg 0$ . Hence by the first-order condition and the chain rule, we have

$$q' = w_i \frac{DU_i(Ay_i)A}{U_i(Ay_i)} \quad (A2)$$

for all  $i$ , where  $DU_i$  denotes the  $(1 \times S)$  Jacobi matrix of the function  $U_i$ . Because  $U_i$  is homogeneous of degree 1, for all  $x \gg 0$  and  $k > 0$  we have  $U_i(kx) = kU_i(x)$ . Differentiating both sides with respect to  $k$  and setting  $k = 1$ , we have  $DU_i(x)x = U_i(x)$ . Hence, multiplying  $y_i$  from the right to (A2), we obtain

$$q' y_i = w_i \frac{DU_i(Ay_i)Ay_i}{U_i(Ay_i)} = w_i.$$

Adding across  $i$ , because  $\sum_{i=1}^I y_i = n$ , we obtain

$$q' n = q' \sum_{i=1}^I y_i = \sum_{i=1}^I w_i = 1.$$

Therefore,

$$q' y_i = w_i = w_i q' n = q' (w_i n) = q' n_i,$$

so the budget constraint holds with equality. Furthermore, letting  $\lambda_i = \frac{1}{w_i}$ , by (A2) we obtain  $D[\log U_i(Ay_i)] = \lambda_i q'$ , which is the first-order condition of the utility maximization problem (A1) after taking the logarithm. Because  $\log U_i$  is concave,  $y_i$  solves the utility maximization problem. Because  $\sum_{i=1}^I y_i = n$ , the asset markets clear, so  $\{q, (x_i), (y_i)\}$  is a GEI. ■

**Step 3.** The GEI is uniquely given as the solution to the planner's problem (3).

**Proof.** Let  $\{q, (x_i), (y_i)\}$  be a GEI. By the first-order condition to the utility maximization problem (A1), there exists a Lagrange multiplier  $\lambda_i \geq 0$  such that

$$\lambda_i q' = D[\log U_i(Ay_i)] = \frac{DU_i(Ay_i)A}{U_i(Ay_i)}. \quad (A3)$$

Multiplying  $n$  from the right and noting that  $DU_i \gg 0$ ,  $An \gg 0$ , and  $Ay_i \gg 0$  imply  $U_i(Ay_i) > U_i(0) = 0$ , we obtain  $\lambda_i q' n > 0$ . Because  $\lambda_i \geq 0$ , we must have  $\lambda_i > 0$  and  $q' n > 0$ . By rescaling the price vector if necessary, we may normalize such that  $q' n = 1$ . Multiplying  $y_i$  to (A3) from the right and using  $DU_i(x)x = U_i(x)$  and the complementary slackness condition, we have

$$\lambda_i q' n_i = \lambda_i q' y_i = \frac{DU_i(Ay_i)Ay_i}{U_i(Ay_i)} = 1 \iff \frac{1}{\lambda_i} = q' n_i = w_i q' n = w_i.$$

Substituting into (A3), we obtain  $q' = w_i D[\log U_i(Ay_i)]$ , which is precisely (A2), the first-order condition of the planner's problem (3) with Lagrange multiplier  $q$ . Because  $(y_i)$  is feasible and the objective function is strictly concave,  $(y_i)$  is the unique solution to the planner's problem. ■

## A.2 Proof of Theorem 2

Let  $u$  be a general Bernoulli utility function with  $u' > 0$  and  $u'' < 0$ . (In view of Theorem 2, we only need to assume  $u(x) = \frac{1}{1-\gamma} x^{1-\gamma}$  or  $u(x) = \log x$ , but most of the following results do not depend on the particular functional form.) Suppose that there are two assets, one risky asset with gross return  $R$  and a risk-free asset with gross risk-free rate  $R_f$ . Let  $R(\theta) := R\theta + R_f(1-\theta)$  be the portfolio return, where  $\theta$  is the fraction of wealth invested in the risky asset. Consider the optimal portfolio problem

$$\max_{\theta} E[u(R(\theta)w)],$$

where  $w$  is initial wealth. The following lemma is basic (e.g., Arrow 1965).

**Lemma 1.** Let everything be as above and  $\theta$  be the optimal portfolio. Then the followings are true.

1.  $\theta$  is unique.
2.  $\theta \geq 0$  according as  $E[R] \geq R_f$ .
3. Suppose  $E[R] > R_f$ . If  $u$  exhibits decreasing relative-risk aversion (DRRA), so  $-xu''(x)/u'(x)$  is decreasing, then  $\partial\theta/\partial w \geq 0$ , that is, the agent invests comparatively more in the risky asset as he becomes richer. The opposite is true if  $u$  exhibits increasing relative-risk aversion (IRRA).

**Proof.**

1. Let  $f(\theta) = E[u(R(\theta)w)]$ . Then  $f'(\theta) = E[u'(R(\theta)w)(R - R_f)w]$  and  $f''(\theta) = E[u''(R(\theta)w)(R - R_f)^2 w^2] < 0$ , so  $f$  is strictly concave. Therefore, the optimal  $\theta$  is unique (if it exists).
2. Because  $f'(\theta) = 0$  and  $f'(0) = u'(R_f w)(E[R] - R_f)$ , the result follows.
3. Dividing the first-order condition by  $w$ , we obtain  $E[u'(R(\theta)w)(R - R_f)] = 0$ . Let  $F(\theta, w)$  be the left-hand side. Then by the implicit function theorem we have  $\partial\theta/\partial w = -F_w/F_\theta$ . Because  $F_\theta = E[u''(R(\theta)w)(R - R_f)^2 w] < 0$ , it suffices to show  $F_w \geq 0$ . Let  $\gamma(x) = -xu''(x)/u'(x) > 0$  be the relative-risk aversion coefficient. Then

$$F_w = E[u''(R(\theta)w)(R - R_f)R(\theta)] = -\frac{1}{w} E[\gamma(R(\theta)w)u'(R(\theta)w)(R - R_f)].$$

Because  $E[R] > R_f$ , by the previous result we have  $\theta > 0$ . Therefore  $R(\theta) = R\theta + R_f(1-\theta) \geq R_f$  according as  $R \geq R_f$ . Because  $u$  is DRRA,  $\gamma$  is decreasing, so  $\gamma(R(\theta)w) \leq \gamma(R_f w)$  if  $R \geq R_f$  (and reverse inequality if  $R \leq R_f$ ). Therefore

$$\gamma(R(\theta)w)(R - R_f) \leq \gamma(R_f w)(R - R_f)$$

always. Multiplying both sides by  $-u'(R(\theta)w) < 0$  and taking expectations, we obtain

$$\begin{aligned} wF_w &= -E[\gamma(R(\theta)w)u'(R(\theta)w)(R - R_f)] \\ &\geq -E[\gamma(R_f w)u'(R(\theta)w)(R - R_f)] = 0, \end{aligned}$$

where the last equality uses the first-order condition. ■

Now we prove Theorem 2.

**Step 1.** The equilibrium is unique and is characterized by the planner's problem (8).

**Proof.** By the same argument as in the proof of Theorem 1, the Epstein-Zin utility function (5) with unit EIS is homogeneous of degree 1 and strictly concave. Hence by Theorem 1, the equilibrium is unique and is characterized by the planner's problem (3). Because we assumed that  $U_i$  has unit EIS, the objective function is additively separable with respect to  $x_{i0}$ . Therefore we can fix  $(y_{i2}, y_{i3})_{i=1}^I$  and maximize over  $(y_{i1})_{i=1}^I$ . This problem is

$$\text{maximize } \sum_{i=1}^I w_i (1 - \beta_i) \log y_{i1} \text{ subject to } \sum_{i=1}^I y_{i1} = \alpha_0 e_0.$$

We can easily solve this problem analytically, and the solution is (7). Substituting this  $y_{i1}$  into (3), the remaining problem becomes (8). ■

**Step 2.** The price-dividend ratio is given by (9), and shifting wealth to a patient agent increases the price-dividend ratio.

**Proof.** Let  $P$  be the stock price, that is, the value of the claim  $(0, e_1, \dots, e_S)'$ . Because the aggregate supply of traded shares is  $\alpha_1$ , the market capitalization of traded shares is  $\alpha_1 P$ . Because the risk-free asset is in zero net supply, the market capitalization of stocks must equal aggregate savings. Because aggregate wealth of capitalists (including  $t=0$  consumption) is  $\alpha_0 e_0 + \alpha_1 P$ , agent  $i$  has wealth share  $w_i$ , and the saving rate out of wealth is  $\beta_i$  due to log utility, the aggregate savings is  $S = \sum_{i=1}^I \beta_i w_i (\alpha_0 e_0 + \alpha_1 P)$ . Setting  $\alpha_1 P = S$  and solving for  $P$ , we obtain (9). From this expression it is clear that shifting wealth from a low  $\beta_i$  agent to a high  $\beta_i$  agent increases  $P/e_0$ . ■

**Step 3.** The log equity premium is independent of capital income shares  $\alpha_0, \alpha_1$ .

**Proof.** Note that the fraction of wealth agent  $i$  invests in the risk-free asset relative to *total* wealth is  $\phi_{i3} = \beta_i (1 - \theta_i)$ . Therefore, the market clearing condition for the risk-free asset is

$$\sum_{i=1}^I w_i \beta_i (1 - \theta_i) (\alpha_0 e_0 + \alpha_1 P) = 0 \iff \sum_{i=1}^I w_i \beta_i (1 - \theta_i) = 0, \quad (\text{A4})$$

which does not directly depend on  $\alpha_0, \alpha_1$ . By homotheticity, the optimal portfolio problem (10) is equivalent to

$$\max_{\theta} E_i [u_i(e_1 \theta + z(1 - \theta))], \quad (\text{A5})$$

where  $z = PR_f$ . Because (A5) does not directly depend on  $\alpha_0, \alpha_1$ , the value of  $z$  that makes the market clearing condition (A4) holds does not depend on  $\alpha_0, \alpha_1$ . By the definition of the log equity premium, we have

$$\mu = \pi' (\log R - \log R_f) = \pi' (\log(e_1/P) - \log R_f) = \pi' \log e_1 - \log z, \quad (\text{A6})$$

so the log equity premium is independent of  $\alpha_0, \alpha_1$ . ■

**Step 4.** Shifting wealth from a high  $\phi_{i3}$  agent to a low  $\phi_{i3}$  agent reduces the log equity premium.

**Proof.** Suppose that in the initial equilibrium we have  $\beta_1(1-\theta_1) > \beta_2(1-\theta_2)$ , so agent 1 is the natural bondholder, and we transfer some wealth from agent 1 to 2. Because  $z = PR_f$  is the only relevant parameter in the optimal portfolio problem (A5), if  $z$  is unchanged after the wealth transfer, then all agents choose their original portfolios. Letting  $P'$  be the stock price and  $w'_i$  the wealth share of agent  $i$  after the transfer, by assumption we have  $w'_1 < w_1$ ,  $w'_2 > w_2$ , and  $w'_i = w_i$  for  $i > 2$ . Then the new aggregate demand of the risk-free asset is

$$\sum_{i=1}^I w'_i \beta_i (1-\theta_i) (\alpha_0 e_0 + \alpha_1 P') = (\alpha_0 e_0 + \alpha_1 P') \sum_{i=1}^I w'_i \beta_i (1-\theta_i),$$

which has the same sign as  $\sum_{i=1}^I w'_i \beta_i (1-\theta_i)$ . However, by (A4), we have

$$\begin{aligned} \sum_{i=1}^I w'_i \beta_i (1-\theta_i) &= \sum_{i=1}^I w'_i \beta_i (1-\theta_i) - \sum_{i=1}^I w_i \beta_i (1-\theta_i) \\ &= \beta_1(1-\theta_1)(w'_1 - w_1) + \beta_2(1-\theta_2)(w'_2 - w_2) \\ &= \epsilon(\beta_2(1-\theta_2) - \beta_1(1-\theta_1)), \end{aligned}$$

where  $\epsilon := w_1 - w'_1 = w'_2 - w_2 > 0$  because  $w'_1 + w'_2 = w_1 + w_2$ . Therefore if we shift wealth from an agent who invests more in the risk-free asset (high  $\phi_{i3} = \beta_i(1-\theta_i)$ ) to an agent who invests less (low  $\phi_{i3} = \beta_i(1-\theta_i)$ ), it will result in an excess supply of risk-free assets.

Because  $\theta_i$  solves (A5), applying Lemma 1 for  $R = e_1$  and  $R_f = z$ , for large enough  $z$  we have  $\theta_i < 0$  for all  $i$ . Therefore, risk-free assets are in excess demand. Hence by the intermediate value theorem (continuity trivially holds by the maximum theorem), in the new equilibrium (which is unique)  $z = PR_f$  must increase. Therefore, the log equity premium (A6) must decrease. ■

### A.3 Proof of Propositions 1 and 2

**Lemma 2.** Consider two agents indexed by  $i = 1, 2$  with common beliefs. Let  $w_i$ ,  $u_i(x)$ ,  $\gamma_i(x) = -xu''_i(x)/u'_i(x)$ , and  $\theta_i$  be the initial wealth, utility function, relative-risk aversion, and the optimal portfolio of agent  $i$ . Suppose that  $\gamma_1(w_1x) > \gamma_2(w_2x)$  for all  $x$ , so agent 1 is more risk averse than agent 2. Then

$$E[R] > R_f \implies \theta_2 > \theta_1 > 0,$$

$$E[R] < R_f \implies \theta_2 < \theta_1 < 0,$$

so the less risk-averse agent invests more aggressively.

**Proof.** Because  $\gamma_1(w_1x) > \gamma_2(w_2x)$ , we have

$$\frac{d}{dx} \left( \frac{u'_2(w_2x)}{u'_1(w_1x)} \right) = \frac{w_2 u''_2 u'_1 - u'_2 w_1 u''_1}{(u'_1)^2} = \frac{1}{x} \frac{u'_2}{u'_1} (\gamma_1(w_1x) - \gamma_2(w_2x)) > 0,$$

so  $u'_2(w_2x)/u'_1(w_1x)$  is increasing. Suppose  $E[R] > R_f$ . By Lemma 1, we have  $\theta_1 > 0$ . Then  $R(\theta_1) \geq R_f$  according as  $R \geq R_f$ . Since  $u'_2(w_2x)/u'_1(w_1x)$  is increasing (and positive), we have

$$\frac{u'_2(R(\theta_1)w_2)}{u'_1(R(\theta_1)w_1)} (R - R_f) > \frac{u'_2(R_f w_2)}{u'_1(R_f w_1)} (R - R_f)$$

always (except when  $R = R_f$ ). Multiplying both sides by  $u'_1(R(\theta_1)w_1) > 0$  and taking expectations, we obtain

$$\begin{aligned} E[u'_2(R(\theta_1)w_2)(R - R_f)] &= E\left[\frac{u'_2(R(\theta_1)w_2)}{u'_1(R(\theta_1)w_1)}u'_1(R(\theta_1)w_1)(R - R_f)\right] \\ &> E\left[\frac{u'_2(R_f w_2)}{u'_1(R_f w_1)}u'_1(R(\theta_1)w_1)(R - R_f)\right] \\ &= \frac{u'_2(R_f w_2)}{u'_1(R_f w_1)}E[u'_1(R(\theta_1)w_1)(R - R_f)] = 0, \end{aligned}$$

where the last equality uses the first-order condition for agent 1. Letting  $f_2(\theta) = E[u_2(R(\theta)w_2)]$ , the above inequality shows that  $f'_2(\theta_1) > 0$ . Because  $f_2(\theta)$  is concave and  $f'_2(\theta_2) = 0$  by the first-order condition, we have  $\theta_2 > \theta_1$ .

The case  $E[R] < R_f$  is analogous. ■

**Proof of Proposition 1.** Because agents have common beliefs, we have  $\theta_i \geq 0$  for all  $i$  if  $E[R] \geq R_f$ . Because the stock is in positive supply, in equilibrium we must have  $E[R] > R_f$ . Therefore by Lemma 2, if  $\gamma_1 > \dots > \gamma_I$ , we have  $0 < \theta_1 < \dots < \theta_I$ . ■

**Proof of Proposition 2.** Let  $u(x)$  be the common CRRA utility function of agents 1 and 2, and

$$f_i(\theta) = E_i[u(R(\theta))] = \sum_{s=1}^S \pi_{is} u(R_f + X_s \theta)$$

be the objective function of agent  $i$ , where  $X_s = R_s - R_f$  denotes the excess return in state  $s$ . By the first-order condition, we have

$$f'_i(\theta_i) = \sum_{s=1}^S \pi_{is} u'(R_f + X_s \theta_i) X_s = 0. \quad (A7)$$

Letting  $q$  be the stock price, because  $R_s = e_s/q$  and  $e_1 < \dots < e_S$ , we have  $X_1 < \dots < X_S$ . Because  $\pi_{is} > 0$  and  $u' > 0$ , by (A7), it must be  $X_1 < 0 < X_S$ . Let  $s^* = \max\{s \mid X_s < 0\}$  be the best state with negative excess returns. Clearly,  $1 \leq s^* < S$ .

Using the definition of the likelihood ratio  $\lambda_s = \pi_{1s}/\pi_{2s}$ , by (A7) we obtain

$$0 = f'_1(\theta_1) = \sum_{s=1}^S \pi_{1s} u'(R_f + X_s \theta_1) X_s = \lambda_{s^*} \sum_{s=1}^S \frac{\lambda_s}{\lambda_{s^*}} \pi_{2s} u'(R_f + X_s \theta_1) X_s.$$

Because by assumption the likelihood ratio  $\lambda_s$  is monotonically decreasing, we have  $\lambda_s/\lambda_{s^*} \geq (\leq) 1$  for  $s \leq (\geq) s^*$ . Furthermore, beliefs are heterogeneous, so either  $\lambda_1/\lambda_{s^*} > 1$  or  $\lambda_S/\lambda_{s^*} < 1$  (or both). Combined with  $X_1 < 0 < X_S$  and  $X_s < (\geq) 0$  for  $s \leq (\geq) s^*$ , it follows that

$$0 = \lambda_{s^*} \sum_{s=1}^S \frac{\lambda_s}{\lambda_{s^*}} \pi_{2s} u'(R_f + X_s \theta_1) X_s < \lambda_{s^*} \sum_{s=1}^S \pi_{2s} u'(R_f + X_s \theta_1) X_s = \lambda_{s^*} f'_2(\theta_1),$$

where the inequality follows because of replacing  $\lambda_s/\lambda_{s^*} \geq (\leq) 1$  by 1 for  $s \leq (\geq) s^*$  makes the term less negative (more positive), and the inequality is strict for  $s=1$  or  $s=S$ . Therefore,  $f'_2(\theta_1) > 0$ , and because  $f_2$  is strictly concave and  $f'_2(\theta_2) = 0$ , we obtain  $\theta_1 < \theta_2$ . ■

## References

- Ang, A., and G. Bekaert. 2007. Stock return predictability: Is it there? *Review of Financial Studies* 20:651–707.
- Arrow, K. J. 1965. *Aspects of the theory of risk-bearing*. Helsinki, Finland: Yrjö Jahnssonin Säätiö.
- Balduzzi, P., and T. Yao. 2007. Testing heterogeneous-agent models: An alternative aggregation approach. *Journal of Monetary Economics* 54:369–412.
- Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron. 2014. Volatility, the macroeconomy, and asset prices. *Journal of Finance* 69:2471–511.
- Basak, S., and D. Cuoco. 1998. An equilibrium model with restricted stock market participation. *Review of Financial Studies* 11:309–41.
- Berge, C. 1959. *Espaces topologiques: Fonctions multivoques*. Paris: Dunod. [English translation: Translated by E. M. Patterson.] In *Topological spaces*, New York: MacMillan, 1963. Reprinted: Mineola, NY: Dover, 1997.
- Bhamra, H. S., and R. Uppal. 2014. Asset prices with heterogeneity in preferences and beliefs. *Review of Financial Studies* 27:519–80.
- Biliias, Y., D. Georgarakos, and M. Haliassos. 2010. Portfolio inertia and stock market fluctuations. *Journal of Money, Credit and Banking* 42:715–42.
- Boudoukh, J., M. Richardson, and R. F. Whitelaw. 2008. The myth of long-horizon predictability. *Review of Financial Studies* 21:1577–605.
- Brav, A., G. M. Constantinides, and C. C. Geczy. 2002. Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy* 110:793–824.
- Buccioli, A., and R. Miniaci. 2011. Household portfolio and implicit risk preference. *Review of Economics and Statistics* 93:1235–50.
- Calvet, L. E., and P. Sodini. 2014. Twin picks: Disentangling the determinants of risk-taking in household portfolios. *Journal of Finance* 69:867–906.
- Campbell, J. Y. 2006. Household finance. *Journal of Finance* 61:1553–604.
- Campbell, J. Y., and R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195–228.
- Campbell, J. Y., and S. B. Thomson. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–31.
- Campbell, S. D., S. Delikouras, D. Jiang, and G. M. Korniotis. 2016. The human capital that matters: Expected returns and high-income households. *Review of Financial Studies* 29:2523–63.
- Carroll, C. D. 2002. Portfolios of the rich. In *Household portfolios*, eds. L. Guiso, M. Haliassos, and T. Jappelli, chap. 10, 389–430. Cambridge: MIT Press.
- Chabakauri, G. 2013. Dynamic equilibrium with two stocks, heterogeneous investors, and portfolio constraints. *Review of Financial Studies* 26:3104–41.
- . 2015. Asset pricing with heterogeneous preferences, beliefs, and portfolio constraints. *Journal of Monetary Economics* 75:21–34.
- Chan, Y. L., and L. Kogan. 2002. Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy* 110:1255–85.
- Chien, Y., H. Cole, and H. Lustig. 2011. A multiplier approach to understanding the macro implications of household finance. *Review of Economic Studies* 78:199–234.
- . 2012. Is the volatility of the market price of risk due to intermittent portfolio rebalancing? *American Economic Review* 102:2859–96.
- Chipman, J. S., and J. C. Moore. 1979. On social welfare functions and the aggregation of preferences. *Journal of Economic Theory* 21:111–39.

- Cochrane, J. H. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21:1533–75.
- Cogley, T. 2002. Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey. *Journal of Monetary Economics* 49:309–34.
- Constantinides, G. M. 1982. Intertemporal asset pricing with heterogeneous agents without demand aggregation. *Journal of Business* 55:253–67.
- Constantinides, G. M., and D. Duffie. 1996. Asset pricing with heterogeneous consumers. *Journal of Political Economy* 104:219–40.
- Constantinides, G. M., and A. Ghosh. 2017. Asset pricing with countercyclical household consumption risk. *Journal of Finance* 72:415–60.
- Cvitanic, J., E. Jouini, S. Malamud, and C. Napp. 2012. Financial markets equilibrium with heterogeneous agents. *Review of Finance* 16:285–321.
- de Rugy, V. 2003a. High taxes and high budget deficits: The Hoover-Roosevelt tax increases of the 1930s. *Cato Institute Tax and Budget Bulletin*, March.
- . 2003b. Tax rates and tax revenue: The Mellon income tax cuts of the 1920s. *Cato Institute Tax and Budget Bulletin*, March.
- Dumas, B. 1989. Two-person dynamic equilibrium in the capital market. *Review of Financial Studies* 2:157–88.
- Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri. 2016a. Heterogeneity and persistence in returns to wealth. Working Paper.
- . 2016b. Heterogeneity in returns to wealth and the measurement of wealth inequality. *American Economic Review: Papers and Proceedings* 106:651–5.
- Fama, E. F., and K. R. French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22:3–25.
- Fisher, I. 1910. *Introduction to economic science*. New York: Macmillan.
- Gârleanu, N., and S. Panageas. 2015. Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. *Journal of Political Economy* 123:670–85.
- Geanakoplos, J. 1990. An introduction to general equilibrium with incomplete asset markets. *Journal of Mathematical Economics* 19:1–38.
- Geanakoplos, J., and M. Shubik. 1990. The capital asset pricing model as a general equilibrium with incomplete markets. *Geneva Papers on Risk and Insurance Theory* 15:55–71.
- Geanakoplos, J., and K. J. Walsh. 2018. Uniqueness and stability of equilibrium in economies with two goods. *Journal of Economic Theory* 174:261–72.
- Gollier, C. 2001. Wealth inequality and asset pricing. *Review of Economic Studies* 68:181–203.
- Gorman, W. M. 1953. Community preference fields. *Econometrica* 21:63–80.
- Granger, C. W. J. 1981. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics* 16:121–30.
- Greenwald, D. L., M. Lettau, and S. C. Ludvigson. 2014. Origins of stock market fluctuations. Working Paper.
- Guenen, F. 2009. A parsimonious macroeconomic model for asset pricing. *Econometrica* 77:1711–50.
- Haliassos, M., and C. C. Bertaut. 1995. Why do so few hold stocks? *Economic Journal* 105:1110–29.
- Hamilton, J. D. 1994. *Time series analysis*. Princeton, NJ: Princeton University Press.
- . 2018. Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics* 100:831–43.

- Hansen, P. R., and A. Timmermann. 2015. Equivalence between out-of-sample forecast comparisons and Wald statistics. *Econometrica* 83:2485–505.
- Hara, C., J. Huang, and C. Kuzmics. 2007. Representative consumer's risk aversion and efficient risk-sharing rules. *Journal of Economic Theory* 137:652–72.
- Hatchondo, J. C. 2008. A quantitative study of the role of wealth inequality on asset prices. *Federal Reserve Bank of Richmond Economic Quarterly* 94:73–96.
- Heaton, J., and D. J. Lucas. 1996. Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy* 104:443–87.
- Hodrick, R. J. 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5:357–86.
- Jacobson, D. B., B. G. Raub, and B. W. Johnson. 2007. The estate tax: Ninety years and counting. *SOI Bulletin* 27:118–28.
- Johnson, T. C. 2012. Inequality risk premia. *Journal of Monetary Economics* 59:565–80.
- Judd, K. L. 1992. Projection methods for solving aggregate growth models. *Journal of Economic Theory* 58:410–52.
- Kacperczyk, M., J. B. Nosal, and L. Stevens. 2018. Investor sophistication and capital income inequality. *Journal of Monetary Economics*. Advance Access published November 17, 2018, 10.1016/j.jmoneco.2018.11.002.
- Kaymak, B., and M. Poschke. 2016. The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics* 77:1–25.
- Kehoe, T. J. 1998. Uniqueness and stability. In *Elements of general equilibrium analysis*, ed. A. Kirman, chap. 3, 38–87. New York: Wiley-Blackwell.
- Kocherlakota, N. R., and L. Pistaferri. 2009. Asset pricing implications of Pareto optimality with private information. *Journal of Political Economy* 117:555–90.
- Kopczuk, W., and E. Saez. 2004. Top wealth shares in the United States, 1916–2000: Evidence from estate tax returns. *National Tax Journal* 57:445–87.
- Lettau, M., and S. C. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56:815–49.
- . 2010. Measuring and modeling variation in the risk-return trade-off. In *Handbook of financial econometrics*, eds. Y. Aït-Sahalia and L. P. Hansen, vol. 1, chap. 11, 617–90. Amsterdam, the Netherlands: Elsevier.
- Lettau, M., S. C. Ludvigson, and J. A. Wachter. 2008. The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies* 21:1653–87.
- Lintner, J. 1965a. Security prices, risk, and maximal gains from diversification. *Journal of Finance* 20:587–615.
- . 1965b. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47:13–37.
- Longstaff, F. A., and J. Wang. 2012. Asset pricing and the credit market. *Review of Financial Studies* 25:3169–215.
- Ludvigson, S. C. 2013. Advances in consumption-based asset pricing: Empirical tests. In *Handbook of the economics of finance*, eds. G. M. Constantinides, M. Harris, and R. M. Stultz, vol. 2, chap. 12, 799–906. Amsterdam, the Netherlands: Elsevier.
- Mankiw, N. G. 1986. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics* 17:211–9.
- Mantel, R. R. 1976. Homothetic preferences and community excess demand functions. *Journal of Economic Theory* 12:197–201.
- McCracken, M. W. 2007. Asymptotics for out of sample tests of Granger causality. *Journal of Econometrics* 140:719–52.



- Mishra, A. V. 2015. Measures of equity home bias puzzle. *Journal of Empirical Finance* 34:293–312.
- Nelson, C. R., and M. J. Kim. 1993. Predictable stock returns: The role of small sample bias. *Journal of Finance* 48:641–61.
- Phillips, P. C. B., and P. Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75:335–46.
- Piketty, T. 2003. Income inequality in France, 1901–1998. *Journal of Political Economy* 111:1004–42.
- Piketty, T., and E. Saez. 2003. Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics* 118:1–41.
- Pohl, W., K. Schmedders, and O. Wilms. 2018. Higher-order effects in asset pricing models with long-run risks. *Journal of Finance* 73:1061–111.
- Rapach, D. E., J. K. Strauss, and G. Zhou. 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23:821–62.
- Rapach, D. E., and G. Zhou. 2013. Forecasting stock returns. In *Handbook of economic forecasting*, eds. G. Elliott and A. Timmermann, vol. 2, chap. 6, 328–83. Amsterdam, the Netherlands: Elsevier.
- Roine, J., J. Vlachos, and D. Waldenström. 2009. The long-run determinants of inequality: What can we learn from top income data? *Journal of Public Economics* 93:974–88.
- Romer, C. D., and D. H. Romer. 2010. The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review* 100:763–801.
- Saez, E., and G. Zucman. 2016. Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics* 131:519–78.
- Sharpe, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19:425–42.
- Stambaugh, R. F. 1999. Predictive regressions. *Journal of Financial Economics* 54:375–421.
- Storesletten, K., C. I. Telmer, and A. Yaron. 2007. Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics* 10:519–48.
- Toda, A. A., and K. Walsh. 2015. The double power law in consumption and implications for testing Euler equations. *Journal of Political Economy* 123:1177–200.
- Toda, A. A., and K. J. Walsh. 2017a. Edgeworth box economies with multiple equilibria. *Economic Theory Bulletin* 5:65–80.
- . 2017b. Fat tails and spurious estimation of consumption-based asset pricing models. *Journal of Applied Econometrics* 32:1156–77.
- Valkanov, R. 2003. Long-horizon regressions: Theoretical results and applications. *Journal of Financial Economics* 68:201–32.
- Vissing-Jørgensen, A. 2002. Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. Working Paper.
- Wachter, J. A., and M. Yogo. 2010. Why do household portfolio shares rise in wealth? *Review of Financial Studies* 23:3929–65.
- Wang, J. 1996. The term structure of interest rates in a pure exchange economy with heterogeneous investors. *Journal of Financial Economics* 41:75–110.
- Weinzierl, M. C., and E. D. Werker. 2009. Barack Obama and the Bush tax cuts (A). Working Paper, Harvard Business School.
- Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455–508.