

Decory Edwards

2025-02-27

Abstract

Recent empirical evidence of heterogeneity in the rate of return (an important feature of the wealth accumulation process) for individuals provide motivation for an analogous assumption in a standard heterogeneous agent (HA) macroeconomic model. In the infinite horizon setting, a uniform distribution of the rate of return across households is estimated such that empirical moments of wealth (net worth) measured in the Survey of Consumer Finances are matched particularly well by their model counterparts. The fit of the model is explored after accommodating more realistic assumptions like life-cycle considerations, bequest motives, and portfolio choice as well. These findings suggest that heterogeneity in parameters which determine optimal consumption-saving behavior other than the time preference factor can generate meaningful wealth inequality. Factors which explain differences in returns across individuals could be used to endogenize heterogeneity in the rate of return, allowing for a more robust analysis of wealth inequality using macroeconomic models.

1 Introduction

The unequal distribution of wealth is an extensively documented phenomenon in numerous countries. Regrettably, this feature has not only endured over time but also intensified in recent years. This point is stressed in a recent article from the Institute for Policy Studies (IPS), which revealed that in 2018, the total wealth of the poorest half of Americans was eclipsed by the combined wealth of the three wealthiest men in the nation. The term “richest” denotes one’s standing in Forbes magazine’s list of the 400 richest individuals. Additionally, the IPS report notes that the combined wealth of the top five richest men on this list skyrocketed by a staggering 123% from March 2020 to October 2021¹.

The unequal distribution of wealth has also been a subject of considerable interest throughout history in various fields. The statistics literature, for instance, focused on linking the distribution of income to the observable skewness in wealth distribution. The economics literature went further by establishing microfoundations for individual wealth outcomes. Similarly, the macroeconomics literature on inequality has seen significant growth, with the distribution of wealth among households offering insight into how the economy as a whole responds to aggregate fiscal shocks. The recent stimulus checks issued during the pandemic serve as a timely example of this phenomenon.

The macroeconomics literature has undergone significant changes in recent years, with the widespread adoption of models that abandon the traditional representative agent assumption in their analysis. Specifically, a model that studies the equilibrium outcomes of an economy composed of individual decision-makers using a single aggregate agent can only have one marginal propensity to consume (MPC). As a result, in response to an aggregate fiscal shock, all households would respond similarly to a one-time stimulus check, which does not align with what transpired during the pandemic². Heterogeneous agent models have emerged as a prominent alternative, offering a more accurate representation of the diversity of economic behaviors and outcomes among households.

The first departure from the representative agent framework entails positing an exogenously determined income process that generates a distribution of income among households. One common approach to incorporating heterogeneity is to adopt Friedman 1957’s description of a permanent and transitory component in the income process. To account for business cycle dynamics, one can further assume that individuals face some level of potential unemployment in each period, creating a precautionary savings motive for consumers. Given that such uncertainty cannot be fully insured against, the availability of a riskless asset that partially insures against income risk results in households choosing

¹See Inequality.org articles data November 21, 2022: “Wealth Inequality in the United States” and “Updates: Billionaire Wealth, U.S. Job Losses and Pandemic Profiteers” (date accessed: March 27, 2023)

²Parker et al. 2022 note that “In sum, while on average the [economic impact payments] EIPs appear to have gone to many households with incomes that were unharmed by the pandemic, some of the EIPs, mainly in the first round, did support short-term spending for some households, primarily those with low ex ante liquid wealth and those reliant on income that could not be earned by working from home.”

to hold different levels of market resources optimally.

Krusell and Smith 1998’s seminal work suggests that models assuming heterogeneity in individual income perform well in matching the aggregate capital stock but poorly in matching the distribution of wealth. The resulting optimal consumption function is concave in an individual’s wealth holdings, meaning that the marginal propensity to consume out of income is increasingly higher at lower levels of wealth. Therefore, a model that places too many households in the middle of the wealth distribution relative to those at lower levels will struggle to match the average MPC estimated from household data. Since our focus is on the implications of fiscal policy for the entire economy, a macroeconomic model’s failure to match the observed wealth distribution in its implied equilibrium is significant.

Moving beyond the standard representative agent framework, the next step is to assume greater heterogeneity among households, leading more households to optimally hold lower levels of wealth. Kaplan and Violante 2022’s recent work provides a comprehensive survey of models that reject this assumption, instead utilizing heterogeneous agent, incomplete markets models featuring (i) uninsurable idiosyncratic income risk, (ii) a precautionary savings motive, and (iii) an endogenous wealth distribution.

Carroll et al. 2017 adopt this approach and further extend the baseline setting to allow for ex-ante heterogeneity amongst households. Specifically, they assume different agents have different rates of time preference, which reflects implicit characteristics of households relevant to their lifetime wealth accumulation. The authors find that this assumption of modest heterogeneity in time preferences is sufficient to match both the shape and skewness of the empirical distribution of wealth. Furthermore, while traditional representative agent models generate an aggregate marginal propensity to consume between 0.02 and 0.04, the β -dist model generates an aggregate MPC between 0.2 and 0.4. This range falls within the values estimated across households in the data.

The household’s optimal consumption-savings problem contains additional elements that could contribute to disparities in wealth accumulation over the course of one’s lifetime. It is worth noting that the time preference factor (β) is one of the key parameters that influences an individual’s equilibrium target level of market resources, but it is not directly observable. Therefore, in order to estimate β , one would need to gather data through surveys or other methods that allow for the direct acquisition of information from households.

On the other hand, estimating differences in the rate of return to financial assets across households is possible, as this variable *is* directly observable. Empirical research has been conducted to estimate such differences, with a recent example being the work of Fagereng et al. 2020. They analyzed 12 years of administrative tax records on capital income and wealth stock for all taxpayers in Norway from 2004-2015 to estimate these rates of return.

This paper aims to enhance the computational, heterogeneous agent modelling framework by integrating recent empirical evidence on disparities in rates of return among households. The objective is to better align the observed wealth distribution with the model predictions, thereby generating more realistic esti-

mates of the average marginal propensity to consume among households.

2 Literature Review

2.1 Collecting Data on the Distribution of Wealth

First and foremost, as the primary focus of models in this field is the distribution of wealth, empirical estimates of the skewness in wealth holdings over time provide valuable insights for this paper. Surveys and the imputation of wealth levels using administrative income tax data (sometimes referred to as the *capitalization method*) are the standard ways of collecting household data on the distribution of wealth for empirical analysis.

Wolff 2004 provides an early analysis of measurements of wealth by the Survey of Consumer Finances (SCF)³ by discussing both the concentration and composition of household wealth 1980s and 1990s. The author’s analysis corroborates the story of significant and growing inequality in the distribution of wealth in the U.S. Specifically, although the wealth of the average household grew in the 1990s, most of the gains in wealth and income during this period were enjoyed by the upper 20 percent of the wealth distribution, and especially the top 1 percent. While from 1983 to 2003 the top 1 percent experienced 33 percent of the total growth in net worth (89 percent for the top 20 percent), the average wealth of the poorest 40 percent of households fell by 44 percent during this same time period and had reached roughly \$2,900 by 2001.

Saez and Zucman 2014 employs the capitalization method on tax data from the Internal Revenue Service to estimate the distribution of wealth in the United States for a much longer time period of 1913 to 2012. The usefulness in the authors’ approach is that they are able to decompose their measure of wealth and savings into fractiles (i.e. top 1 percent, top 10 percent, bottom 20 percent wealth shares), which allows them to analyze the evolution of wealth over time in a way that is standard in the existing literature on wealth inequality. The authors not only find that inequality in the U.S. wealth distribution is relatively high and has been growing significantly in the later periods of their dataset, but they also attribute this growth primarily to the wealthiest of households. Indeed, they cite that the wealth shares of the top .1 percent of the distribution grew from 7 percent in 1978 to 22 percent in 2012.

2.2 Explaining Inequality in the Distribution of Wealth

Benhabib and Bisin 2018 conducted a notable, thorough review of the literature on the documented skewness in the distribution of wealth. The survey begins with historical accounts of the origins of the shape of the wealth distribution, dating back as early to Pareto and Samuelson. The authors then provide the

³See Kennickell 2017a for an extensive description of the methodology for sampling the wealthiest households in the SCF and Kennickell 2017b for an analysis of the performance of the SCF at measuring the wealth of the top 1 percent.

traditional theoretical explanations of this unequal distribution: (i) skewness in the (exogenous) distribution of earnings, (ii) stochastic returns to wealth and savings, and, importantly, (iii) microfoundations for the evolution of wealth resulting from the consumption and saving behavior of households⁴.

Gabaix et al. 2016 define a notion for the speed of convergence to provide an explanation for observed evolution of income inequality over time, specifically in the upper tail of the distribution in the past 40 years in the United States. Notably, the authors show that, in order to match the empirical dynamics of inequality, one needs to allow for more forms of heterogeneity in the income process for households that are not incorporated in the standard consumption and saving models.⁵ The first form is *type dependence* in the income growth rate distribution, which models the case in which some households have a higher average income growth rate. The second form, *scale dependence*, captures the fact that higher income levels are more susceptible to shocks to their income growth. The authors find that former does a good job at explaining this fast rise in income inequality, and the latter can generate infinitely fast transitions in inequality.⁶

De Nardi and Fella 2017 provide another survey of the literature, more focused on the microfoundations for the distribution of wealth. Specifically, the authors note a number of possible extensions of models of household consumption and saving behavior, inspired by observable differences and the demographics of households, which lead to differences in wealth accumulation over time. Earnings and rate of return risk, ex-ante heterogeneity in preferences, medical expenses, bequest motives, and entrepreneurship are all cited as potential avenues to better explain the shape of the distribution of wealth using the behavior of households.

2.3 Heterogeneous Agents Macroeconomics: Wealth and the Marginal Propensity to Consume

Güvenen 2011 provide an excellent review of the heterogeneous agent models which have become common frameworks of analysis in the macroeconomics literature in the recent decade. The key insights from this survey can be found in its thorough discussion on the relationship between the complete markets hypothesis and varying degrees of insurance against risk for households. In addition, the authors make a notable distinction between aggregation and representative-agent models; this is important, although the macroeconomics models with incomplete markets and heterogeneous agents of interest may lack

⁴As explored in the next section, the emergence of heterogeneous agent models has been a significant development in investigating this issue. Bewley 1983, Aiyagari 1994, and Huggett 1993 are among the earliest examples.

⁵Note that, although this analysis is about the distribution of income, this literature notably asserts that the distribution of wealth inherits some of its skewness from the distribution of income

⁶As we will see, these notions of “type dependence” and “scale dependence” show up in the literature on household heterogeneity and the wealth and income distributions; most importantly, in the discussion on heterogeneous rates of return to wealth.

a representative-agent counterpart for which the equilibrium analysis is straightforward⁷ (since these models can generally be solved analytically), computational and numerical methods may be used to conduct a similar analysis of the equilibrium, aggregate implications of such a model. Lastly, each of these key points of the paper are grounded both in theoretical results regarding static and dynamic economies with optimizing households and firms, and also in empirical evidence on both the inherent risk faced by households and the varying degrees of available insurance against these risks.

Krueger, Mitman, and Perri 2016 provide another review of the heterogeneous agent macroeconomics literature. Here, more emphasis is placed on empirical evidence of heterogeneity across households (in earnings, income, consumption, and wealth) leading up to and during the Great Recession and incorporating features of the business cycle in the model of household consumption and saving to better match the presented cross-sectional data. Furthermore, the authors present an augmented version of the model incorporating demand externalities to analyze the relationship between the distribution of wealth and the dynamics of aggregate output.

The key insight from the heterogeneous agent macroeconomics literature is in the ability of these models to produce an aggregate marginal propensity to consume which is reasonably close to its empirical counterpart. This is a notable failing of the representative-agent framework. Kaplan and Violante 2022 have conducted an extensive analysis of different classes of models in this area, highlighting their strengths and potential drawbacks. Notably, they find that the heterogeneous agent, incomplete markets framework with a single asset generate a marginal propensity to consume that is too low compared to empirical data. While incorporating ex-ante heterogeneity or behavioral preferences can generate a larger MPC, these models tend to suffer from a “missing middle” problem - an equilibrium distribution that is overly polarized at the extremes and underestimates the wealth held by middle-income households. As a result, it is worth exploring whether a model that includes one asset, a precautionary savings motive, and ex-ante heterogeneous rates of return also exhibits this shortcoming.

Krusell and Smith 1998 have developed a model that considers both idiosyncratic and aggregate risk in a household’s optimization problem. As an additional exercise, they incorporate heterogeneity in time preference to explain the shape of the wealth distribution. More recently, Carroll et al. 2017 have updated the model to include income and time preference heterogeneity, with a more realistic income process that accounts for permanent and transitory shocks to household income. The model’s analytical framework is flexible enough to accommodate other potential sources of ex-ante heterogeneity, which can be seen in one of its key equilibrium conditions - the *Growth Impatience Condition* (GIC):⁸

⁷In fact, as the authors note, some heterogeneous agent models may feature “approximate aggregation” properties, where the aggregate implications of the model are very close to the implications of some representative agent model

⁸Note that β is the time discount factor, ρ is the coefficient of relative risk aversion, \mathcal{D} is

$$\left(\frac{(R\delta)^{1/\rho} \mathbb{E}[\psi^{-1}] \mathcal{D}}{\Gamma} \right) < 1.$$

2.4 Measurements of heterogeneous rates of return

The rationale behind incorporating heterogeneity in rates of return to asset holdings lies in the use of novel datasets in recent empirical research to quantify the differences in returns among individuals. Fagereng et al. 2020 offer an extensive overview of the role that heterogeneous rates of returns play in wealth distribution, as well as conducting their own systematic analysis of return heterogeneity using 12 years of data from Norway’s administrative tax records. The authors’ findings reveal substantial differences in the average returns to assets for individuals (*type dependence*), that this heterogeneity is found both within and across classes of assets with varying levels of risk, and that returns are positively correlated with wealth (*scale dependence*). Moreover, they further demonstrate that this discovery of heterogeneous returns exhibits significant persistence over time and are positively correlated across generations. Each of these findings provide not only motivation for the assumption of ex-ante heterogeneous rates of return in the buffer-stock savings model of households, but also provide a benchmark to compare the distribution of rates of return resulting from the estimation procedure aimed at best matching the empirical distribution of wealth, as in Carroll et al. 2017.

Bach, Calvet, and Sodini 2018 provide more evidence towards the assumption of ex-ante heterogeneity in households’ rates of return, as they employ administrative panel data on the balance sheets of Swedish residents to gauge historical and expected returns, as well as risks associated with asset holdings. Like previous studies, they also find that heterogeneous returns play a considerable role in the levels and growth of top wealth shares over time. As their analysis is focused on the portfolio performance of wealthy households, their findings offer more support for the idea that scale dependence is an important feature of the observed heterogeneity in the rate of return for households.

Campbell, Ramadorai, and Ranish 2019 offers a similar conclusion in their analysis by studying wealth held in equity accounts in India between 2002 and 2011. The authors find that heterogeneity in returns to investment, which can be achieved by both the inherent randomness associated with risky investment and differences in the investment strategies of investors, is a main contributor to the increase in inequality of wealth held in equity portfolios during the time period. Here, the authors attribute the scale dependence associated with the returns to equity portfolios to the finding that smaller accounts tend to be poorly diversified relative to their larger account counterparts.

Deuflhard, Georgarakos, and Inderst 2018 provide an important analysis for the heterogeneous agent, incomplete markets model with a precautionary saving

the survival probability, $\mathbb{E}[\psi^{-1}]$ is the expectation of receiving permanent shock ψ , $R_t = \mathfrak{T} + r_t$, where \mathfrak{T} is the depreciation rate of capital and r_t is the interest rate, and Γ is labor productivity growth.

motive and a *single asset* to partially insure against risk with by studying the performance of households' investments in savings accounts. Not only do they find substantial type dependence in the rate of return to these safe assets, they also attribute the heterogeneity in returns to differences in financial sophistication. As we will see in the next section, providing an explanation for differences in returns to investments for households is a vital step in potentially endogenizing this form of ex-ante heterogeneity among households in future research. Much like the Fagereng et al. 2020, comparing the estimated distribution of rates of return across households to the results of this paper is crucial, especially since the authors' focus is on heterogeneous returns to safe assets.

2.5 Recent HA models with heterogeneous rates of return

The paper Daminato and Pistaferri 2024 is a recent work which incorporates heterogeneous returns into the solution of a model of consumption-saving for households. There, they use data from the PSID to document heterogeneity in returns, which they state is comparable to that found in the Norwegian registry data used by Fagereng et al. 2020.

Benhabib, Bisin, and Luo 2019 proposes an overlapping generations model that incorporates intergenerational wealth transfers. There, agents face uncertainty regarding both labor and capital income. Benhabib, Bisin, and Luo 2017 undertake a similar exercise, where household preferences for bequests to the next generation are more explicitly defined. Both papers conclude that the distribution of earnings and differences in rates of savings and bequests are crucial in matching the characteristics of the observed wealth distribution's tail ends.

Guler, Kuruscu, and Robinson 2022 develop a life-cycle model that provides a comprehensive description of households' optimal decision-making to endogenize heterogeneity in the rate of return. In addition to the standard consumption-savings behavior, their model also considers optimal choices regarding housing and mortgage decisions. These modeling choices lead to a structural model that better matches the observable size and skewness of the wealth distribution, using realistic features of the household's decision-making process, as a larger number of individuals actively engage with the housing market than with financial markets. Furthermore, these modeling choices enable the authors to investigate the effects of aggregate fiscal shocks, including one-time stimulus payments and mortgage debt relief programs.

3 Model

3.1 Defining the stochastic income process

Each household's income (y_t) during a given period depends on three main factors. The first factor is the aggregate wage rate (W_t) that all households in the economy face. The second factor is the permanent income component (p_t), which represents an agent's present discounted value of human wealth. Lastly,

the transitory shock component (ξ_t) reflects the potential risks that households may face in receiving their income payment during that period. Thus, household income can be expressed as the following:

$$y_t = p_t \xi_t W_t.$$

The level of permanent income for each household is subject to a stochastic process. In line with Friedman 1957's description of the labor income process, we assume that this process follows a geometric random walk, which can be expressed as:

$$p_t = p_{t-1} \psi_t,$$

The white noise permanent shock to income with a mean of one is represented by ψ_t , which is a significant component of household income. The probability of receiving income during a given period is determined by the transitory component, which is modeled to reflect the potential risks associated with becoming unemployed. Specifically, if the probability of becoming unemployed is \mathcal{U} , the agent will receive unemployment insurance payments of $\mu > 0$. On the other hand, if the agent is employed, which occurs with a probability of $1 - \mathcal{U}$, the model allows for tax payments τ_t to be collected as insurance for periods of unemployment. The transitory component is then represented as:

$$\xi_t = \begin{cases} \mu & \text{with probability } \mathcal{U}, \\ (1 - \tau_t) l \theta_t & \text{with probability } 1 - \mathcal{U}, \end{cases}$$

where l is the time worked per agent and the parameter θ captures the white noise component of the transitory shock.

3.2 Baseline model for households

This paragraph presents the baseline version of the household's optimization problem for consumption-savings decisions, assuming no ex-ante heterogeneity. In this case, each household aims to maximize its expected discounted utility of consumption $u(c) = \frac{c^{1-\rho}}{1-\rho}$ by solving the following:

$$\max \mathbb{E}_t \sum_{n=0}^{\infty} (\beta)^n u(c_{t+n}).$$

It's worth noting that the setting described here follows a perpetual youth model of buffer stock savings, similar to the seminal work of Krusell and Smith 1998. To solve this problem, we use the bellman equation, which means that the sequence of consumption functions $\{c_{t+n}\}_{n=0}^{\infty}$ associated with a household's optimal choice over a lifetime must satisfy⁹

⁹Here, each of the relevant variables have been normalized by the level of permanent income ($c_t = \frac{C_t}{p_t}$, and so on). This is the standard state-space reduction of the problem for numerical tractability.

$$\begin{aligned}
v(m_t) &= \max_{c_t} u(c_t(m_t)) + \beta \mathcal{E}_t[\psi_{t+1}^{1-\rho} v(m_{t+1})] \\
&\text{s.t.} \\
a_t &= m_t - c_t(m_t), \\
k_{t+1} &= \frac{a_t}{\mathcal{B}\psi_{t+1}}, \\
m_{t+1} &= (\mathbb{I} + r_t)k_{t+1} + \xi_{t+1}, \\
a_t &\geq 0.
\end{aligned}$$

3.2.1 The analogy for rates of return

If we want to explore how different returns to assets can affect the endogenous wealth distribution, it's important to examine the following decomposition of a household's evolution of market resources over time:

1. Assets at the end of the period are equal to market resources minus consumption:

$$a_t = m_t - c_t.$$

2. Next period's capital is determined from this period's assets via

$$k_{t+1} = \frac{a_t}{\mathcal{B}\psi_t}.$$

3. Finally, the transition from the beginning of period $t + 1$ when capital has not yet been used to produce output, to the middle of that period when output has been produced and incorporated into resources but has not yet been consumed is:

$$m_{t+1} = (\mathbb{I} + r_t)K_{t+1} + \xi_{t+1}.$$

It's worth recalling that in this model, the rate of return to capital is represented as $(\mathbb{I} + r_t)$. This rate of return is directly related to the endogenous level of wealth, which is determined by the level of capital K_{t+1} . Therefore, if there are differences in the rate of return across households, this will result in further disparities in wealth holdings.

4 Results

4.1 Matching observed inequality in the distribution of wealth

4.1.1 Adding ex-ante heterogeneity in time preferences

In Carroll et al. 2017’s baseline model, heterogeneity in the time preference factor is not accounted for. However, to address this, the model is extended to include this factor. This is done by assuming that different types of households have a time preference factor drawn uniformly from the interval $(\dot{\beta} - \nabla, \dot{\beta} + \nabla)$, where ∇ represents the level of dispersion. Afterward, the model is simulated to estimate the values of both $\dot{\beta}$ and ∇ so that the model matches the inequality in the wealth distribution. To achieve this, the following minimization problem is solved:

$$\{\dot{\beta}, \nabla\} = \arg \min_{\dot{\beta}, \nabla} \left(\sum_{i=20,40,60,80} (w_i(\dot{\beta}, \nabla) - \omega_i)^2 \right)^{\frac{1}{2}}$$

subject to the constraint that the aggregate capital-to-output ratio in this model matches that of the perfect foresight setting:

$$\frac{K}{Y} = \frac{K_{PF}}{Y_{PF}}.$$

Note that w_i and ω_i give the porportion of total aggregate net worth held by the top i percent in the model and in the data, respectively.

4.1.2 The analogous exercise for ex-ante heterogeneous rates of return

The β -dist model proves to be useful in a setting where there are heterogeneous time preference factors since it captures an unobservable component of a household’s decision-making process. While the microeconomics literature has put in considerable effort to estimate this parameter, there is currently no consensus on its value.

Recent studies by Fagereng et al. 2020 and Bach, Calvet, and Sodini 2018 have not only estimated the rate of return on asset holdings but have also uncovered significant heterogeneity across households. Given this motivation, the revised model assumes the existence of multiple types of agents, each earning a distinct rate of return on their assets. A calibration exercise akin to the one used in the β -dist model is then performed. This crucial step involves comparing the resulting endogenous distribution from simulating this calibrated model to its empirical counterpart to determine if there is an ex-ante distribution of rates of return that can match the observable inequality in the wealth distribution. If a distribution of returns to asset holdings satisfies this criterion, the final step involves reconciling this model heterogeneity with the observed differences in rates of return found in the aforementioned literature.