Accounting for Wealth Concentration in the US

Barış Kaymak[*][†]
Université de Montréal and CIREQ

David Leung[‡]
National Taiwan University

Markus Poschke[§]
McGill University and CIREQ

Abstract

The recent literature has argued for high concentration of earnings, differences in rates of return on assets and bequests as potential determinants of the high level of wealth concentration in the US. Analyzing the joint distribution of earnings, capital income and net worth, we find evidence for excess returns on investment among the highest income groups. Nonetheless, concentration of labor earnings is the primary source of wealth concentration in the US. This finding reflects the high correlation between earnings and wealth in the data, as well as the fact that earnings are a major source of income for top income and wealth groups.

1 Introduction

Wealth holdings in the US are highly concentrated, more so than income, with a fifth of the population holding almost all the assets and the wealthiest 1% alone holding over a third. To explain this, the literature emphasized a set of competing factors. A first strand highlights labor income heterogeneity and risk, which lead to higher saving rates among

[*] Department of Economics, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, QC H3C 3J7, e-mail: baris.kaymak@umontreal.ca
[†] We thank participants at the NBER, the SED meetings, as well as participants at various seminars and conferences. The authors acknowledge support from *Chaire de la fondation J.W. McConnell en études américaines* and the *SSHRC* grant no. 435-2018-0264.
[‡] Department of Economics, National Taiwan University, Taiwan.
[§] Department of Economics, McGill University, Montréal, QC

high earning groups (Castañeda et al., 2003; Kindermann and Krueger, 2014; Kaymak and Poschke, 2016). A second strand emphasizes capital income heterogeneity, where some households have access to investment vehicles with persistently higher rates of return (Benhabib et al., 2019; Gabaix et al., 2016). A third strand points to dynastic accumulation of wealth through bequests (Galor and Zeira, 1993; De Nardi, 2004).[1]

These approaches differ in their depictions of who the wealthy are and how they become wealthy. As a result, they reach different conclusions in their assessments of economic policies. For instance, using a model of labor income risk, Kindermann and Krueger (2014) prescribe an optimal marginal tax rate as high as 90% for top income groups, whereas Brüggemann (2017) calls for a top tax rate of 60% based on a model of entrepreneurship. Guvenen et al. (2019a) argue that wealth taxes may bring efficiency gains in models with rate of return heterogeneity. Similarly, Hubmer et al. (2016) attribute much of the rise in wealth concentration over the last 50 years to top income tax cuts, whereas, in earlier work, Kaymak and Poschke (2016) find the rise in the dispersion of wage income to be the major factor behind the rise in wealth dispersion. Such variation in policy evaluation warrants a better understanding of the factors that shape the US wealth distribution.

Regrettably, a direct empirical assessment of how important labor and capital income are for generating large fortunes in the US is infeasible due to the lack of long panel data on earnings, assets and their returns for households at the top of the income and wealth distribution. Nonetheless, data on the joint distribution of assets and income are available for a cross-section of households. In this paper, we combine this information with an overlapping generations model of savings to assess the empirical relevance of the different modeling approaches to wealth concentration.

The key difference between these approaches is their prediction for the factor composition of income among top income and wealth groups. If wealth concentration is driven by differences in the rate of return on assets, then these groups should rely heavily on capital income. If it is driven instead by earnings differences, then labor income should be the primary source of income. Using data from the Survey of Consumer Finances (SCF), we document the share of earnings in income for different income and wealth groups. Consistent with the heterogeneous returns hypothesis, we find that income from labor is less important for top income groups relative to the rest of households. Our calculations indi-

[1] Realistic wealth distributions also arise in models of entrepreneurship (Quadrini, 2000; Cagetti and de Nardi, 2009). These models can combine elements of labor income and capital income heterogeneity, as discussed further below.

cate that this cannot be explained solely by the larger stock of wealth held by these groups. The implied dispersion in rates of return is sizeable. In a given year, the top 0.1% of the income distribution, for instance, enjoys a rate of return that is 3.4 times the rate on the assets of the bottom 90% of the income distribution.

The data also show that earnings are nonetheless the primary source of income for almost all households – including top income groups. Earnings account for 49 to 68 percent of total income for the top 1% of incomes, depending on the treatment of capital gains and proprietors' income. Our preferred estimate is 64 percent. For the top 1% of the wealth distribution, 55 percent of income comes from labor. Households outside these top groups rely almost exclusively on labor income. These patterns suggest an indispensable role for earnings in shaping the wealth distribution. They have not previously been used in the literature on the wealth distribution, which has focused exclusively on marginal distributions of income and wealth.

Drawing conclusions for wealth concentration from cross-sectional data requires a model of savings. To that end, we employ a general equilibrium, life-cycle model of household saving behavior. The model features uninsurable shocks to earnings, heterogeneity in rates of return, a non-homothetic bequest motive, survival risk and retirement. These elements capture the three main motives for savings: the precautionary motive, the intertemporal consumption smoothing motive, and the bequest motive. We then calibrate the model to match the distribution of earnings, income and net worth observed in a cross-section of households in the SCF. When combined with our model of savings, these distributions are informative of the dynamics of earnings for top income groups and the extent and persistence of rates of return on assets, which are not directly observed in the survey data. The calibrated model features realistic earning dynamics, with a high degree of kurtosis and negative skewness as documented by Guvenen et al. (2019b). Given the earnings distribution, the model requires the presence of a small fraction of households with a very high rate of return on assets to match the tail of the wealth distribution. The model also fits over-identifying moments well. Notably, it generates realistic life-cycle profiles of average earnings, income and wealth, as well as their cross-sectional dispersion by age.

Next, we assess the relative contributions of the model elements to wealth concentration. We do this in two ways. First, we shut down different model components and compare the implied wealth concentration to the data. Eliminating top earning categories induces the largest drop in top wealth shares, by more than half. Eliminating bequest inequality reduces top wealth shares by ten percent to a third. Eliminating differences in the rates of return mostly affects the top 0.1% of the wealth distribution, reducing their share in total wealth by 7 to 43 percent. Other concentration measures and the Gini coefficient are affected much less, reflecting the limited power of rate of return differences in generating high wealth concentration when the intergenerational wealth transmission is imperfect.

Second, we recalibrate the model to generate the observed levels of wealth concentration in the absence of either top earners or return heterogeneity. This reveals that models that rely only on differences in the rate of return not only understate earnings concentration, but also predict a counterfactually high role for capital income for top income and wealth groups. Relative to the data, the implied correlation between earnings and wealth is much too low and that between income and wealth is too high, since wealth is the primary source of income in this case.

Overall, these results suggest that concentration of labor earnings is the primary source of wealth concentration in the US, reflecting the importance of labor earnings for top income and wealth groups in the data. Return heterogeneity and bequest inequality play significant but smaller roles.

In the next section, we give a brief overview of the related literature. In Section 3 we summarize the empirical distributions of earnings, income and wealth in the SCF, as well as the factor composition of income for different income and wealth groups. In Section 4, we present the model. The calibration procedure is described in Section 5 and the results for the benchmark economy are discussed in Section 6. Section 7 analyzes the relative roles of rate of return heterogeneity, labor income risk and bequests in determining the observed distribution of wealth in the US. Section 8 concludes.

2 Macroeconomics of the Wealth Distribution

The foundations of modern macroeconomic analysis of the wealth distribution are laid out in early work by Huggett (1993) and Aiyagari (1994), which eventually led to the "standard" incomplete markets model (Heathcote et al., 2009). In this setting, dispersion in asset holdings emerges from households' motives to accumulate assets in order to insure themselves against fluctuations in their earnings. Early iterations of these models focused on the implications of household heterogeneity for aggregate macroeconomic outcomes, such as the role of precautionary savings for total capital accumulation or for business cycles. It was nonetheless noted that the observed differences in earnings and income risk as measured in household surveys (like e.g. the PSID) were not large enough to generate a highly skewed distribution of wealth. Subsequently, a separate literature emerged aiming to enhance the model for applications to questions related to wealth inequality. The macro literature on the wealth distribution now is vast, with applications to various economic questions. In our discussion of the literature below, we focus on the main modelling extensions and their implications for a subset of applications as an example.[2]

The main shortcoming in the original model was that wealthy households cared little about earnings risk and therefore limited their savings once their wealth was sufficiently high to shield consumption from future drops in earnings. The first modelling extensions that helped maintain continuing wealth accumulation, and thereby generate a skewed wealth accumulation, involved introducing differences in savings motives or rates of return on assets. This was achieved by explicitly introducing heterogeneity in preferences for saving (Krusell and Smith, 1998), in rates of return on assets (Benhabib et al., 2019; Gabaix et al., 2016; Nirei and Aoki, 2016), as well as bequest motives that are increasing in wealth (De Nardi, 2004). Benhabib et al. (2011) show analytically that idiosyncratic capital income risk can generate a Pareto tailed wealth distribution with a realistic tail index. More recently, Stachurski and Toda (2019) generalize this result. Benhabib et al. (2019) and Cao and Luo (2017) provide quantitative assessments of the contribution of rate of return heterogeneity to wealth concentration. The common element among these models is that the main source of differences in wealth accumulation is capital income. High wealth concentration emerges because wealthy households enjoy higher rates of return on their assets and have higher saving rates out of income.[3]

A second strand of the literature focused on better measurement of earnings risk. Household surveys typically provide an incomplete picture of the distribution of earnings and associated risks due to censoring of earnings above a certain level or limited sampling of high-earning households. Castañeda et al. (2003) was the first to show that the standard incomplete markets model can indeed generate a highly skewed wealth distribution if the earnings process is calibrated accordingly. Subsequent work further developed this approach, using the recent progress in measurement of top earnings levels based on administrative data to discipline the extent of earnings dispersion and risk used as inputs in the model (Kindermann and Krueger, 2014; Kaymak and Poschke, 2016). The economic

[2]See De Nardi and Fella (2017) for a more detailed review of the macro literature on wealth inequality.

[3]Note that return heterogeneity or preference heterogeneity are not required to generate a Pareto distribution for wealth with a tail index that is smaller than that of the income distribution. The same arises, even without capital income risk, in models where agents (or dynasties) are not infinitely lived, or where the transmission of wealth is imperfect (Benhabib et al., 2006; Jones, 2015; Stachurski and Toda, 2019).

mechanism here is that households who temporarily have very high earnings anticipate lower future earnings (be it because of retirement or the vagaries of a career at the top), and therefore have a very strong saving motive. The explicit consideration of very high earnings levels is a key ingredient in these models, where the main source of wealth concentration consists in differences in labor income, labor income risk, and the associated saving behavior.

Another mechanism that can generate high wealth concentration is entrepreneurship, which combines elements from both strands we have discussed, as profits reflect both the return on assets invested in the business and the value of entrepreneurial labor (Quadrini, 2000; Cagetti and De Nardi, 2006). Entrepreneurs in these models reap higher rates of returns on their investments if, or as long as, they are financially constrained (Buera, 2009; Moll, 2014).[4] This may encourage them to save faster in order to bypass credit constraints. They may also save more because earnings on their entrepreneurial skills may be subject to significant fluctuations due to business risk.

All these approaches substantially improved the ability of the standard incomplete markets model to generate a realistic wealth distribution for the US, offering economists several modelling options. The existing literature has operated with either a model with capital income risk, one with high earnings dispersion, or one with entrepreneurship. Yet, the relative roles of earnings and capital income risk in generating the observed wealth concentration are not well understood, in part due to lack of data on the dispersion of rates of return on assets at the household level in the US.[5] This paper combines the two approaches and is the first to use information on the *joint* distributions of earnings, capital income and assets to identify the relevance of different modelling approaches to wealth concentration.

3 Distributions of Income, Earnings and Wealth in the US

In this section we summarize the distributions of earnings, income and wealth, and discuss the role of capital income vis-à-vis earnings for top income and wealth groups. The primary source of data is the 2010 and 2016 waves of Survey of Consumer Finances (SCF), a triennial cross-sectional survey of US families on their assets, income, and demographic

[4] Without credit constraints, models of entrepreneurship can be mapped into a model with earnings heterogeneity and a common return on assets (See Appendix B).

[5] Recent work by Fagereng et al. (2020) and Bach et al. (2016) provide empirical evidence for rate of return heterogeneity using panel data from Norway and Sweden, respectively.

characteristics.[6] We compare our results to those from administrative tax records reported in Piketty and Saez (2003).

3.1 The distribution of wealth, income and earnings

Since the objective is to use the joint distribution of income and wealth to identify different modeling components, we adopt a market-based notion of income that is compatible with the models of wealth distribution mentioned above. Our definition of market income includes wage and salary income, business and farm income, interest and dividend income, private pension withdrawals and capital gains whereas it excludes income from fiscal sources such as transfer income and social security income.

For most households, labor income consists of wage and salary income, which includes pay for work for an employer as well as any salary drawn from an actively managed business. The SCF follows the tax filing guidelines for classifying sources of income. The IRS requires all corporations to explicitly report wage and salary for actively involved shareholders. Some business organizations, such as partnerships and sole proprietorships are exempted from this requirement. As a result, a small group of business owners report only business income. In such cases, we impute wage and salary income only if a household reports income from actively owned businesses, but does not report any wage income, or, if the respondent or their spouse reports explicitly that they did not draw salary from their actively managed business.

To determine the share of business income that is attributable to capital, we assume that the contribution of capital to active business income is proportional to the total value of equity held in the business. Consequently, we regress active business income on equity, controlling for the quantity and quality of the labor input.[7] We include the number of hours worked by the household members that are actively involved in the business as well as demographic characteristics such as gender, age and education as control variables. The resulting coefficient on equity is 0.27, which we interpret as the capital income share.[8]

[6] We exclude the 2013 survey, which reports income from the 2012 calendar year and shows an unusual increase in realized capital gains. This is largely due to an anticipated increase in the capital gains tax scheduled by the Patient Protection and Affordable Care Act that was enacted in 2010 and provided for additional taxes on high income groups starting in 2013.

[7] To account for negative values of business income, we use the following logarithmic transformation: $\widetilde{\log} x = \log(1 + sign(x) \times |x|)$.

[8] This is the share in *net* income, since depreciation expenses are deducted from the reported business income. The share in gross income can be found by adding the rate of depreciation.

Table 1 – Cross-Sectional Distributions of Income, Earnings and Net Worth

Top percentile	0.1%	0.5%	1%	5%	10%	20%	40%	Gini
Net worth	0.14	0.28	0.37	0.63	0.76	0.88	0.97	0.85
Income	0.08	0.18	0.23	0.41	0.53	0.68	0.86	0.67
Earnings	0.06	0.14	0.19	0.36	0.49	0.66	0.86	0.66[†]

[†] The Gini coefficient for households with a working age head is 0.58.

Note.– Table shows the cumulative concentration shares for the top percentile groups. Income includes capital gains. Data comes from the SCF 2010 and 2016. Sample includes all households.

Accordingly, we allocate 73 percent of active business income to labor for those who do not report wage income from their business. This labor share of business income is close to that of 75 percent found by Smith et al. (2019) in the US tax data.

The resulting labor income share potentially underestimates the true contribution of labor for three reasons. First, since it is less advantageous to report business income as wages for tax purposes, business owners who report wage income may underreport it. Second, for those who do not report wage income, we only impute wages for the spouse and the respondent. If other members of the household work for the business, their labor income is classified as part of the household's business income.[9] Third, both years in the sample coincide with the post-recession recovery period, where asset returns were above their typical average.

Table 1 shows the cross-sectional distributions of income, earnings and wealth. The distribution of net worth is far more skewed than the distributions of income and earnings: the Gini coefficient for net worth is 0.85, whereas it is 0.66 for earnings and 0.67 for income. This is driven by both the heavier concentration of wealth at the top and a larger fraction of households without assets relative to those without income. The top 1% of the net worth distribution has 37% of assets and the top 0.1% holds 14% of total wealth. Earnings are also concentrated, with the top 1% earners' share of 19% in total earnings and the top 0.1% share of 6%.

There is a strong correlation between wealth and earnings. The coefficient of correlation between earnings and net worth is 0.35 for households with a working age head, and it is 0.30 for the entire sample. Similarly, the correlation between income and net worth is 0.52.

[9] The survey questions needed to ascertain if household members have claimed wage income from their business are only available for the respondent and the spouse.

Table 2 – Shares of Net Worth by Income and Earning Groups

Top percentile of ...	0.1%	0.5%	1%	5%	10%	20%	40%
... income	0.09	0.20	0.27	0.51	0.61	0.71	0.81
... earnings	0.04	0.13	0.19	0.38	0.47	0.57	0.67

Note.– Table shows cumulative shares of net worth held by top income and earning groups. Income includes capital gains. Data comes from the SCF 2010 and 2016.

This strong relationship can also be seen in Table 2, which shows the wealth shares of different earning and income groups. The top 1% of earners hold about 19% of wealth. Similarly, the households in the highest 1% of incomes hold 27% of total wealth in the US. If the correlation were zero, wealth shares would have been equal to the population shares when ranked by income or earnings. These suggest that savings out of earnings and income play a significant role for accumulation of wealth.

3.2 The share of income from labor

Figure 1 shows the factor composition of income for top income and wealth groups. The gray bars show the share of wage and salary income in total income, as reported by the households. The red solid bars show the share of total income from labor, including imputed earnings for proprietors who do not report wage income from their businesses. The whisker ticks on each bar indicate the values when capital gains are included in or excluded from total income. The height of each bar represents the average of these two values.

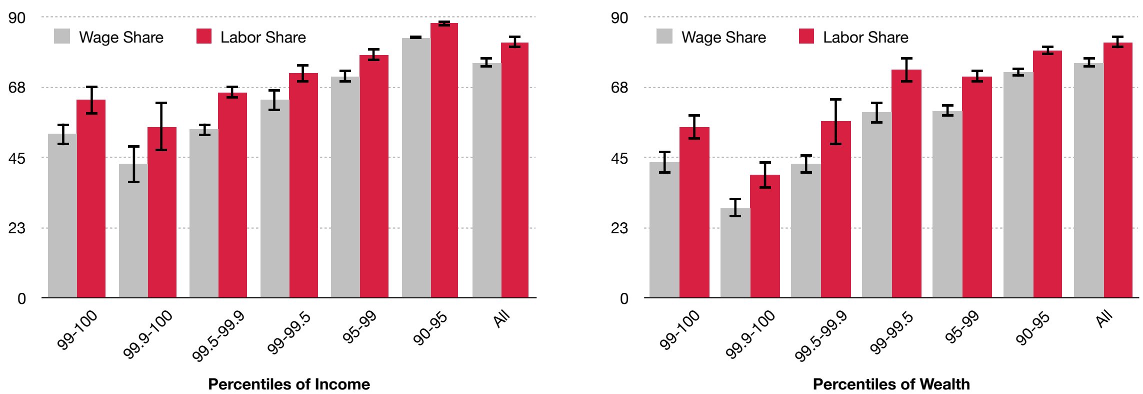
On the aggregate, 74 to 84 percent of net income is attributed to labor depending on the treatment of capital gains and business income.[10] Panel (a) shows the labor shares by percentiles of total income. Most households rely primarily on wage and salary income. Outside the top 1 percent of the income distribution, labor income constitutes at least two thirds of total income. Since business income and capital gains are not an important source of income for these groups, the particular definition of income does not affect this result.

For the top 1 percent of the income distribution, labor income constitutes 59 percent of total income when capital gains are included, and 68 percent when they are excluded

[10]Since the accounting convention is to report the net income from capital, i.e. excluding depreciation, the share of labor income in net income is typically higher than its share in gross income typically used to calibrate macro models. We use net capital income in our comparisons of the model predictions below with the data above.

Figure 1 – Labor Component of Income by Income and Wealth Groups (%)

(a) Income Groups (b) Net Worth Groups



Note.– Figure shows wage and labor shares of total income by percentiles of income and net worth. Labor income includes imputed wage income for active business owners who do not draw salary from their businesses. The whiskers show the labor shares with and without capital gains in total income. The bar heights show the average of the two values. See Appendix Table A.1 for the data values. Data comes from the 2010 and 2016 waves of the SCF.

from the definition of income. The wage share, which excludes imputed wages for some proprietors is roughly 10 points lower. Columns 2 to 4 show the percentiles of income within the top 1%. Income from labor is the major source of income, accounting for at least half of total income, with the exception of the top 0.1%.

A similar pattern is observed for top groups by net worth in Panel (b) of Figure 1. Labor's share of income for the top 1% of wealth is 0.51 and 0.59, with and without capital gains. Excluding capital gains, income from labor is the main source for households outside of the top 0.1% of the net worth distribution. With capital gains, income from capital dominates labor for those in the top 0.5%.

Table 3 compares our findings with statistics from IRS data. We use the 2015 update to the tables in Piketty and Saez (2003), who report the sources of income for finely defined top income groups. Since it is not possible to observe which tax units draw salary from their business, no imputation is made, and we report business income separately.[11] These figures are comparable to the top rows of Figure 1. The share of wage income for the top 1 percent income group as reported by tax units in Table 3 is 49 percent when capital gains are included, and 56 percent when they are excluded – exactly as in our findings in the SCF data

[11] Using administrative tax data on individuals, Smith et al. (2019) make such an imputation. They attribute 52 to 77% of the top 1% incomes to labor.

Table 3 – Composition of Income for Top Income Groups (IRS)

without capital gains	Income Percentile Category 99-100	99-99.5	99.5-99.9	99.9-99.99	99.99-100
Wage	56	73	61	47	34
Business	30	20	29	37	37
Interest and Dividend	14	7	10	15	29
with capital gains	**99-100**	**99-99.5**	**99.5-99.9**	**99.9-99.99**	**99.99-100**
Wage	49	68	54	40	27
Business	27	19	26	32	30
Int., Div. and Capital Gains	24	13	19	28	42

Note.– Figures in percentages and correspond to averages for 2010-2015. Income percentiles are determined excluding capital gains (KG). Figures come from 2015 data update to Piketty and Saez (2003).

reported in Figure 1.[12] Columns 2 to 5 in Table 3 report the components of income within the top 1 percent of income. Wage income constitutes more than half the income for those outside the top 0.1 percent of top income earners. For the top 0.1 percent of the income distribution, the share of wage income drops and interest and dividend income becomes increasingly important. For the top 0.01 percent of the income distribution, interest and dividend income constitute 42 percent of total income when capital gains are included.

Both the survey data from the SCF and the tax data from the IRS records agree on the relative roles of sources of income. For most households, earned income from labor services is the primary source of income. As we move up the income ladder, the share of labor income declines, and income from capital increases. Nonetheless, even among the top 1% of households (and tax units), the most conservative definition of labor income indicates that at least half the income can be attributed to labor. As the size of the top fractile is reduced, capital income becomes more important. The upshot of this is that labor income remains a non-negligible source of income throughout, and is a primary source of income for most households (or tax units) outside the highest income and net worth groups.

[12]There are two subtle but apparently inconsequential differences between the two sets of statistics. First, the income concept reported in Piketty and Saez (2003) includes fiscal income, such as social security payments and other transfer payments. Since transfer payments are not a significant source of income for top income groups, this does not affect the results. Second, the IRS data is based on tax units whereas the SCF data is based on primary economic units, which consists of the core members of the household. In most cases, this includes the respondent, their spouse, if any, and their dependent children.

3.3 Implied heterogeneity in the rate of return on assets

A group's relative rate of return on assets can be inferred from its relative labor share of income. To see this , let λ_i denote the labor income share of a group of households i:

$$\lambda_i = \frac{e_i}{e_i + r_i k_i},$$

where e_i and k_i are average earnings and assets of a household in the group, and r_i is the group specific return on assets. Let $i = 0$ represent the base group, which we define below as the bottom 90% of the income distribution. Denote the earnings ratio of group i relative to the base group by $e_{i/0} = e_i/e_0$, and the asset ratio by $k_{i/0} = k_i/k_0$. Then the labor income share of any group can be expressed as:[13]

$$\lambda_i = \frac{\lambda_0}{\lambda_0 + \frac{k_{i/0}}{e_{i/0}} \frac{r_i}{r_0}(1 - \lambda_0)}. \tag{1}$$

Equation (1) relates the labor income share of top income groups to the labor income share of the rest of the economy. Top income groups have lower labor income shares under two situations. First, their relative wealth is higher than their relative earnings, $k_{i/0}/e_{i/0} > 1$, or, equivalently, their wealth-to-earnings ratio is relatively higher. This could arise if, for instance, the saving rate increases with earnings. Second, they have a higher rate of return on their assets: $r_i/r_0 > 1$. To isolate the role of the latter, we carry out two calculations. First, we compute the counterfactual labor share of income for top income groups implied by their relative wealth-to-earnings ratio, assuming that all income groups have the same rate of return on their assets. If returns are indeed higher for higher income groups, then the actual labor share should be less than what is implied by their assets alone. Second, equation (1) can be solved for the relative rates of return, implied by the observed labor income shares for different groups:

$$\frac{r_i}{r_0} = \frac{e_{i/0}}{k_{i/0}} \cdot \frac{1/\lambda_i - 1}{1/\lambda_0 - 1}.$$

Table 4 shows the results for top income groups. The base income group is the bottom 90% of the income distribution. As shown in Table 1, relative wealth and earnings are higher for higher income groups. The wealth-to-earnings ratio is also increasing with

[13] $\lambda_i = \frac{e_{i/0} e_0}{e_{i/0} e_0 + r_i k_{i/0} k_0} = \frac{e_0}{e_0 + \frac{k_{i/0}}{e_{i/0}} r_i k_0} = \frac{e_0}{e_0 + r k_0} \frac{e_0 + r k_0}{e_0 + \frac{k_{i/0}}{e_{i/0}} \frac{r_i}{r_0} r_0 k_0} = \frac{\lambda_0}{\lambda_0 + \frac{k_{i/0}}{e_{i/0}} \frac{r_i}{r_0}(1 - \lambda_0)}.$

Table 4 – Labor Income Shares and the Implied Rate of Return on Assets

Income Percentile	0 - 90	90 - 95	95 - 99	99 - 99.5	99.5 - 99.9	99.9 - 100
Data:						
Relative Wealth	1	5	14	36	63	206
Relative Earnings	1	4	7	17	31	83
Labor Income Share	0.91	0.88	0.78	0.72	0.66	0.55
Inferred values:						
Common Return LIS†	0.91	0.89	0.84	0.82	0.83	0.80
Relative Rate of Return	1.00	1.14	1.52	1.80	2.52	3.36
Rate of Return (%)‡	2.1	2.4	3.2	3.7	5.2	7.0

Note.– Table computes the synthetic rate of return on assets for top income households implied by the labor share in their income. Data LIS comes from the 2010 and 2016 waves of the SCF and reflects the average labor share of income with and without capital gains. † LIS implied by the relative size of assets and earnings assuming that all households have the same return on their assets. ‡ Assuming an annual average rate of return of 6%.

income. The third row reports the actual labor share observed in the data. The labor share of income implied by a larger stock of assets, assuming common rates of return, is reported in the fourth row. It declines moderately from 0.91 for the base group to 0.80 for the top 0.1%. In each category, the observed labor share is below the labor share implied by assets alone. Therefore, the observed labor share cannot be explained solely by the fact that top income groups have relatively more wealth. This suggests that top income groups must also have experienced higher rates of return on their assets.

The last row in Table 4 shows the rates of return on assets for each income category relative to the base group. Higher income groups have increasingly higher rates of return on their assets. The 90-95th percentile, for example, has 1.14 times the base return, whereas the top 0.1% of incomes have 3.36 times the base return. The dispersion is subtantial. To translate the relative returns to actual returns, we assume an aggregate return of 6% per year. This is consistent with the aggregate labor share of (net) income of 82% in Figure 1, and a capital-to-income ratio of 3.[14] This implies an annual rate of return of 2.1% for the base group and 7% for the highest income category.[15]

[14]Formally, $r = rK/K = rK/Y \times Y/K = (1 - 0.82)/3 = 6\%$.

[15]The implied rates of return are robust to the particular definition of the labor income share. Using various measures in Figure 1, we obtain a rate of return between 2.3 and 4.7 times the base-group return for the top 0.1% of the income distribution. For a 6% aggregate return, these imply annual rates of return between 6.6% and 7.3%.

While Table 4 suggests the presence of *cross-sectional* heterogeneity in asset returns, it is not possible to gauge directly how much this contributes to wealth concentration. The persistence and predictability of the returns are crucial for inferring the potential saving response to these rates by top income and wealth groups. Since the data is cross-sectional, the dynamic process for the rates of return cannot be estimated empirically. Next, we combine the cross-sectional information presented in this section with a model of household saving to quantify the role of earnings concentration and rate of return heterogeneity in explaining the wealth distribution in US.

4 A Life-Cycle Model of Wealth Accumulation

For the analysis, we employ an overlapping generations model of life-cycle wealth accumulation under incomplete markets (Imrohoroglu and Imrohoroglu, 1995; Huggett, 1996). We augment the model by incorporating idiosyncratic labor income with extraordinary earning levels as in Kaymak and Poschke (2016) and Kindermann and Krueger (2014), heterogeneity in the return to capital income in the spirit of Benhabib et al. (2019), and a non-homothetic bequest motive as in Cagetti and De Nardi (2006).

4.1 Environment

Each period, a continuum of new agents enter the economy, with a potential life-span of J periods, subject to survival probabilities $s(j)$ for each age j. The total population is normalized to one.

Agents work for the first J_r periods of their lives, after which they retire. Workers earn income for their labor and on their assets. A worker's labor endowment is given by $z\varepsilon_j$, where z is a stochastic component following a first-order Markov process $F_z(z'|z)$, and ε_j is a deterministic component that captures age-dependent improvements in human capital, such as work experience. With this endowment, a worker generates a labor income of $wz\varepsilon_j h$, where w is the market wage per skill unit and $h \in [0, 1]$ is hours worked. Income from capital is $r\kappa k$, where k denotes assets, and κ is an idiosyncratic rate of return that follows a Markov process defined by $F_\kappa(\kappa'|\kappa)$. Once retired, agents collect a pension, $b(z)$, that depends on the last realization of the labor productivity shock z, and continue to earn income on their assets.[16] Total income is denoted by y.

[16]The actual US social security benefits depend on a worker's average earnings over their career. Following

All income is subject to taxation. The tax system, outlined below in detail, distinguishes between different sources of income and features transfers. The disposable income after all taxes and transfers is denoted by y^d. Consumption is subject to sales tax at a rate τ_s. The government uses the tax revenue to finance an exogenously given level of expenditures, G, pension payments and other transfers.

Agents value consumption, leisure and assets they leave for their offspring. The problem of an agent is to choose labor supply, consumption, savings and bequests to maximize the expected present value of lifetime utility. At each period j, agents are informed of their labor endowment for the period, $z\varepsilon_j$, and their rate of return on assets, $r\kappa$, prior to taking their decisions. Future utility is discounted by a factor $\beta \in (0, 1)$. Formally, the Bellman equation for a worker's problem is

$$V(j,k,z,\kappa) = \max_{c,k' \geq 0, h \in [0,1]} \left\{ \frac{c^{1-\sigma_c}}{1-\sigma_c} - \theta \frac{h^{1+\sigma_l}}{1+\sigma_l} + \beta s(j) \mathbb{E}[V(j+1,k',z',\kappa')|z,\kappa] + (1-s(j))\phi(k') \right\}$$

subject to

$$(1+\tau_s)c + k' = y^d(zw\varepsilon_j h, r\kappa k) + k + Tr + \Phi(j,z,\kappa),$$

where $\phi(k) = \phi_1 \left[(k+\phi_2)^{1-\sigma_c} - 1\right]$ is the utility value of bequeathed assets, and $\Phi(j,z,\kappa)$ denotes assets received as a bequest. The expectation is taken over the future values of the labor endowment, z', and the rate of return on assets, κ', given the processes F_z and F_κ.

Since retirees do not work, the Bellman equation for a retiree's problem is given by

$$V(j,k,z,\kappa) = \max_{c,k' \geq 0} \left\{ \frac{c^{1-\sigma_c}}{1-\sigma_c} + \beta s(j) \mathbb{E}[V(j+1,k',z,\kappa')|\kappa] + (1-s(j))\phi(k') \right\}$$

subject to

$$(1+\tau_s)c + k' = y^d(b(z), r\kappa k) + k + Tr$$

The consumption goods are produced by a representative firm using aggregate capital K and total effective labor N with a Cobb-Douglas production function: $Y = F(K,N) = \Psi K^\alpha N^{1-\alpha}$.

Kindermann and Krueger (2014), we assume that pension benefits depend on the earnings of the last working age period. This allows us to capture the redistributive structure of the US pensions system while maintaining computational feasibility.

4.2 Stationary equilibrium

Let $s = \{j, k, z, \kappa\} \in S$ be a generic state vector. The stationary equilibrium of the economy is given by a consumption function, $c(s)$, a savings function, $k'(s)$, labor supply, $h(s)$, a value function $V(s)$, a wage rate $w(s)$ and a distribution of agents over the state space $\Gamma_j(s)$, such that (i) functions $V(s)$, $c(s)$, $k'(s)$ and $h(s)$ solve the consumers' problems, (ii) firms maximize profits, factor markets clear:

$$K = \int k'(s) d\Gamma(s) \qquad N = \int z\varepsilon_j h(s) d\Gamma_{j<J_r}(s),$$

the government's budget is balanced:

$$\begin{aligned} G + \int b(z) d\Gamma_{j\geq J_r}(s) &= \tau_s \left[\int c(s) d\Gamma(s) \right] + \int [y - y^d(zw\varepsilon_j h, r\kappa k)] d\Gamma_{j<J_r}(s) \\ &\quad + \int [y - y^d(b(z), r\kappa k)] d\Gamma_{j\geq J_r}(s). \end{aligned}$$

and $\Gamma(s)$ is consistent with the policy functions, and is stationary.

5 Calibration of the Model

To quantify the model parameters, we first choose a set of 21 parameters based on information that is exogenous to the model. Then, we calibrate the remaining 23 parameters so that the stationary equilibrium of the model economy is consistent with 23 target moments describing the empirical distributions of earnings, wealth and income, as well as other informative data moments. We do so by minimizing the equally weighted sum of squared deviations between model moments and data moments.

While our approach is broadly consistent with the standard for quantitative macro models with idiosyncratic risk, it has some distinctive elements. From a modeling perspective, the main differences are in the earning process, where we allow some households the possibility of reaching an extraordinarily high labor productivity level in the spirit of Castañeda, Díaz-Giménez and Ríos-Rull (2003), Kindermann and Krueger (2014) and Kaymak and Poschke (2016), and in the rate of return risk in the spirit of Benhabib, Bisin and Luo (2019). From an empirical point of view, we differ from earlier studies in our explicit use of the joint distribution of earnings, income and wealth in addition to their marginal distributions to identify these modeling extensions.

In this section, we discuss the choice of target moments. In the next section, we present the fit of the model in terms of those moments as well as additional over-identifying moments that we do not target.

5.1 Demographics

The model period is five years. The first model period corresponds to ages 20 to 24. Death is certain after age $J = 16$, which corresponds to ages 95-99. Retirement is mandatory at age 65 ($J_R = 10$). Following Halliday et al. (2019), we assume that the survival probability is a logistic function of age: $s(j) = [1 + \exp(\omega_0 + \omega_1 j + \omega_2 j^2)]^{-1}$ and use the parameter values recommended therein.[17]

5.2 Preferences and production technology

Preferences are described by a discount factor, β, the inverse elasticity of intertemporal substitution, σ_c, the inverse elasticity of labor supply, σ_l, the disutility of work θ and the parameters that govern utility from bequests: ϕ_1 and ϕ_2. We discuss the last two separately below. We set $\sigma_l = 1.22$, which implies a Frisch elasticity of 0.82. Blundell, Pistaferri and Saporta-Eksten (2016) report an estimate of 0.68 for males and 0.96 for females. Thus a value of 0.82 for a model of households seems broadly plausible. We choose θ so that an average household allocates 35% of their time endowment to work at the equilibrium. We set $\sigma_c = 1.5$, in the middle of the range typically used in the literature. The discount factor, β, is chosen so that the ratio of capital to annual income is 2.9 given an annual depreciation rate of 4.5%. This results in a value of $\beta = 0.90$, or 0.98 per annum. The implied (value-weighted) interest rate that clears the asset market is 4.88%. We normalize aggregate TFP, $\Psi = 1$, and set the elasticity of output with respect to capital, α, to 0.27, to match the observed aggregate labor income share.

[17] Halliday et al. (2019) calibrate the parameters to match three moment conditions: the dependency ratio (population aged 65 and over divided by population aged 20-64), which is 39.7% in the data, the age weighted death rate for 20 to 100 year olds of 8.24%, and the ratio of the change in the survival probability between ages 65-69 and 75-79 to the change in survival probability between ages 55-59 and 65-69, which is 2.27 in the data.

Table 5 – Transition Matrix for the Labor Productivity Process

	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
	$f_L + a_L$	$f_L + a_M$	$f_L + a_H$	$f_H + a_L$	$f_H + a_M$	$f_H + a_H$		
$f_L + a_L$	A_{11}	A_{12}	A_{13}	0	0	0	λ_{in}	0
$f_L + a_M$	A_{21}	A_{22}	A_{23}	0	0	0	λ_{in}	0
$f_L + a_H$	A_{31}	A_{32}	A_{33}	0	0	0	λ_{in}	0
$f_H + a_L$	0	0	0	A_{11}	A_{12}	A_{13}	λ_{in}	0
$f_H + a_M$	0	0	0	A_{21}	A_{22}	A_{23}	λ_{in}	0
$f_H + a_H$	0	0	0	A_{31}	A_{32}	A_{33}	λ_{in}	0
z_7	λ_{out}	λ_{out}	λ_{out}	λ_{out}	λ_{out}	λ_{out}	λ_{ll}	λ_{lh}
z_8	0	0	0	0	0	0	λ_{hl}	λ_{hh}
initial dist.	$\zeta/2$	$(1-\zeta)$	$\zeta/2$	$\zeta/2$	$(1-\zeta)$	$\zeta/2$	0	0

5.3 Labor productivity process

We assume that the stochastic component of labor productivity takes eight values, six of which are ordinary states and the other two are extraordinary states that generate exceptionally high earnings levels. The ordinary levels z_1 to z_6 consist in combinations of two components: a permanent component, $f \in \{f_H, f_L\}$, that is fixed over a household's career, and a transitory component, $a \in \{a_L, a_M, a_H\}$.

Individuals randomly draw their value of f in the first period of their lives. Idisyncratic fluctuations in labor income risk over the life-cycle are captured by a 3-by-3 matrix $A = [A_{ij}]$ with $i, j \in \{L, M, H\}$ and $\sum_j a_{ij} = 1 - \lambda_{in}$, as well as by λ_{in}, which represents the probability of entering an extraordinary state of productivity. The stochastic labor productivity process is summarized by the matrix in Table 5. The following additional assumptions are explicit in the formulation of the matrix. The probability of reaching an extraordinary status, λ_{in}, is independent of one's current productivity state and age. Likewise, if a household loses their extraordinary status, then it is equally likely to transition to any ordinary productivity state.[18]

Our working assumption is that the values for the ordinary states and the transitions among them can be inferred from survey data, whereas the transitions to, from and among extraordinary states can not. We calibrate values and transitions of ordinary states as follows. We assume that the transitory component of productivity follows an AR(1) process, with an annual persistence of 0.97, as estimated by Heathcote et al. (2010), and variance

[18]The effect of these assumptions on our quantitative analysis is negligible.

σ_a. Wage regressions in the PSID with fixed worker effects indicate that 60% of the total variance of wages reflect differences in the permanent component, and the remaining 40% reflect transitory shocks. Accordingly, we set $\sigma_a^2 = 0.4\sigma^2$, where σ^2 is the total variance. Normalizing $a_M = 0$ and setting $a_L = -\eta$ and $a_H = \eta$ then allows us to determine η and the elements of A in terms of σ using the Rouwenhorst approximation. To determine the levels of the fixed components, we set $f_L = -f_H$. Assuming an equal division of households between the two permanent states, we then express f_H in terms of σ such that the implied variance is $0.6\sigma^2$.

At this point, all ordinary productivity levels are expressed relative to σ. Note that σ^2 is the variance corresponding to the long-run stationary state associated with the transition matrix. Since the wage distribution is not stationary over the life-cycle, this object is not directly observed in the data. To determine σ, we parameterize the initial distribution of households over the ordinary productivity states at the beginning of their careers as in the last row of Table 5. By assumption, households are not born to extraordinary productivity. Then, given the age distribution implied by the survival function described in Section 5.1, we jointly calibrate the parameters ζ and σ such that the overall cross-sectional variance of wages equals 0.58 and the standard deviation of wages grows by 47 percent between the ages of 22 and 57, as we estimate in the PSID. This requires that $\sigma = 0.81$ and $\zeta = 0.18$.

This leaves the extraordinary productivity levels z_7 and z_8, and the transition probabilities $(\lambda_{in}, \lambda_{out}, \lambda_{ll}, \lambda_{lh}, \lambda_{hl}, \lambda_{hh})$. Two of these are pinned down by adding-up constraints for probabilities. In order to identify the remaining parameters, we include moments on the marginal distribution of earnings, specifically, the top 0.1 and 1 percent shares, the labor income shares of the percentile groups 95-99 and 99-100 of the income distribution, the Gini coefficient for wealth, as well as the probability of remaining a top 1% earner in the set of target moments for the calibration of the model.

The stochastic process for labor productivity is combined with a deterministic age profile of wages common to all workers. We calibrate this profile to that estimated in the PSID.

5.4 Rate of return process

The rate of retun on capital is stochastic and takes three distinct values, $\{r\kappa_L, r\kappa_H, r\kappa_{top}\}$, where r is the equilibrium market rate of return, and the κ_i denotes the relative return of the household. The transitions between these states are governed by the following transition matrix:

$$\Pi_\kappa = \begin{array}{c|ccc} & \kappa_L & \kappa_H & \kappa_{\text{top}} \\ \hline \kappa_L & \pi_{ll} & 1-\pi_{ll}-\pi_{in} & \pi_{in} \\ \kappa_H & 1-\pi_{hh}-\pi_{in} & \pi_{hh} & \pi_{in} \\ \kappa_{\text{top}} & 0 & 1-\pi_{top,top} & \pi_{top,top} \end{array}$$

Since asset returns are not directly observed in the data, we include moments on wealth concentration and intergenerational wealth mobility in the set of target moments to identify the levels of returns and the probabilities $\pi_{ll},\ \pi_{hh},\ \pi_{top,top}$, and π_{in}. In particular, we target the top 0.1%, 1%, 5% and 10% wealth shares. In addition, we target the intergenerational probabilities of staying in the fourth and fifth quintiles of the age-adjusted wealth distribution. Using data from the PSID for the period from 1984 to 1999, Charles and Hurst (2003) report these two moments to be 0.26 and 0.36, indicating substantial persistence of wealth across generations.[19] We replicate these authors' estimation method using data generated from our model to compute the corresponding model moments.[20]

5.5 Tax and transfer system

The tax system consists of personal income taxes levied on capital and labor earnings, corporate taxes and a sales tax. The tax receipts are used to support exogenous government expenditures, transfers to households, and pensions.

Corporate taxes are modeled as a flat rate, τ_c, levied on a portion of capital earnings before households receive their income. We set $\tau_c = 23.6\%$, which is the average effective marginal tax rate on corporate profits in 2010 as estimated by Gravelle (2014) based on tax records. To reflect the fact that for most households, positive net worth takes the form of real estate and thus is not subject to corporate income taxes, we assume that corporate taxes only apply to capital income above a threshold d_c.[21] We then choose d_c such that the corporate tax revenue as a share of GDP is 2.5%.[22] Households are subject to sales tax, which is set to 5% of consumption, following Kindermann and Krueger (2014).

Personal income taxes are applied to earnings, non-corporate capital income and pen-

[19] Gayle et al. (2016) extend the analysis to more recent waves and find very similar numbers.

[20] We exclude model parent-child pairs where either the child or the parent is in the top 1% of wealth. Results are similar when they are included.

[21] Only about 20% of US households hold stocks or mutual funds directly (Bover, 2010; Heaton and Lucas, 2000).

[22] This figure, like those on government expenditure and pensions used below, comes from NIPA Tables and is an average for the years 2010 to 2016.

sion income, if any. Taxable personal income is given by:

$$y_f = zw\varepsilon_j h + \min\{r\kappa k, d_c\} \qquad \forall j < J_r$$
$$y_f = b + \min\{r\kappa k, d_c\} \qquad \forall j \geq J_r.$$

Total disposable income is obtained after applying corporate and personal income taxes and adding lump-sum transfers from the government:

$$y^d = \lambda \min\{y_b, y_f\}^{1-\tau} + (1-\tau_{max})\max\{0, y_f - y_b\} + (1-\tau_c)\max(r\kappa k - d_c, 0) + Tr$$

The first two terms above represent our formulation of the current US income tax system, which can be approximated by a log-linear form for income levels outside the top of the income distribution (Benabou, 2002), augmented by a flat rate for the top income tax bracket. The power parameter $0 \leq \tau \leq 1$ controls the degree of progressivity of the tax system, while λ adjusts to meet the government's budget requirement.[23]

One advantage of this formulation for the income tax system is that it also allows for negative taxes. Income transfers are, however, non-monotonic in income. When taxes are progressive, transfers are first increasing, and then decreasing in income. This feature allows addressing features of the real tax system like the earned income tax credit and welfare-to-work programs, which imply transfers that vary with income.

When disposable income is log-linear in pre-tax income, the marginal tax rate increases monotonically with income, converging to 100% at the limit. The second term in the maximum operator avoids this feature by imposing a cap on the top marginal tax rate, denoted by τ_{max}. y_b denotes the critical level of taxable income at which the top marginal tax rate is reached: $\lambda(1-\tau)y_b^{-\tau} = 1 - \tau_{max}$. The top marginal tax rate is set to 39.6%, as reported by the IRS. For identification of the progressivity of the general income tax system, τ, we include the difference between the average income tax rate paid by the top 1% and the bottom 99% of the income distribution in the set of target moments. Piketty and Saez (2007) report this value to be 6.8%.

The government uses the tax revenue to finance exogenous expenditures, pension payments and transfers. The expenditures are set at 15.5% of GDP to yield a sum of expenditure and transfers of 26.1% of GDP, as observed in the data. In addition, the government

[23] $\tau = 0$ implies a proportional tax system. When $\tau = 1$, all income is pooled, and redistributed equally among agents. For values of τ between zero and one, the tax system is progressive. See Guner et al. (2014), Heathcote et al. (2017) and Bakış et al. (2015) for evidence on the fit of this function.

makes lump-sum transfers to all households. In the data, these transfers represent 2.7% of GDP in the form of disability benefits, veterans benefits etc. We set the transfers in the model Tr accordingly. In the last step, we choose λ in the personal income tax function to balance the government's budget.

We model pension benefits, which are paid from age 65 onwards, to mimic the US social security system as described in the US Social Security Bulletin (Social Security Administration, 2013). Let bp_1 and bp_2 be the two bend points, expressed as multiples of average earnings, for the three replacement rate brackets (90%, 32%, and 15%), and let b^{cap} be the maximum receivable pension benefit. Then the benefit or primary insurance amount (PIA) for an individual retiring with productivity z is

$$\begin{aligned} b(z) \quad = \quad & \xi \min\{b^{cap}, 0.9 \min(\tilde{e}(z), bp_1) + 0.32 \max[\min(\tilde{e}(z), bp_2) - bp_1, 0] \\ & +0.15 \max(\tilde{e}(z) - bp_2, 0)\} \end{aligned}$$

where $\tilde{e}(z)$ are average earnings of working age agents of productivity z in the model's stationary equilibrium. The reported formula is applied to an individual's earnings, whereas the model is based on households, some of which may contain non-working spouses, or survivors. Therefore, we introduce a benefit adjustment factor, ξ , and calibrate it to match the average ratio of social security expenditure to GDP in the data for the years 2010 to 2016.

5.6 Bequests

Recall the utility value of bequests from Section 4.1: $\phi(k) = \phi_1[(k + \phi_2)^{1-\sigma_c} - 1]$. The parameter ϕ_2 represents the degree of non-homotheticity of bequests, while ϕ_1 controls the overall preference for bequests. We choose these parameters to match the bequest-to-wealth ratio reported by Guvenen et al. (2019a), as well as the share of all bequests accounted for by the top 2% largest bequests, which is 40% (Feiveson and Sabelhaus, 2018).

The model does not feature an explicit link between parents and their offspring, which requires a larger state space, and is computationally challenging. On the other hand, re-distribution of all bequests among younger agents, a common simplification, curbs the model's ability to capture the dynastic persistence of wealth. We proceed with a hybrid approach, which can be summarized as follows. We assume that at age 50, the average age of bequest receipt in the data (Feiveson and Sabelhaus, 2018), agents randomly draw a bequest from a mixture of the bequest distributions of the deceased in the model, where the weights in the mixture depend on the recipient's state: a recipient with permanent productivity component i' and saving return j' draws from the distribution of bequests left by deceased agents with permanent productivity component i and return j $(i, j, i', j' = L, H)$ with probability $\gamma(i', j'; i, j)$.[24] To limit the number of parameters, we model $\gamma(i', j'; i, j)$ as $\gamma_z(i, i')\gamma_\kappa(j, j')\bar{\Gamma}(i, j)/\tilde{\Gamma}(i', j')$, where $\gamma_z(i, i')$ equals the parameter $\bar{\gamma}_z \in [0, 1]$ if $i = i'$ and $1 - \bar{\gamma}_z$ otherwise, and analogous for $\gamma_\kappa(j, j')$. $\bar{\Gamma}(i, j)$ denotes the fraction of deaths with states (i, j), and $\tilde{\Gamma}(i', j') = \sum_{i,j} \gamma(i', j'; i, j)$ ensures that the probabilities sum to one.

These assumptions allow the model to capture intergenerational correlations by ensuring that the bequest received by a child is more likely to come from a parent with similar characteristics. Concretely, if $\bar{\gamma}_z$ $(\bar{\gamma}_\kappa) > 1/2$, high-productivity (high-return) children are more likely to receive a bequest from a high-productivity (high-return) parent. We calibrate $\bar{\gamma}_z$ and $\bar{\gamma}_\kappa$ to match the intergenerational correlations of wages and wealth of 0.3 and 0.365 reported by Solon (1992) and Charles and Hurst (2003), respectively.

Table 6 shows the resulting values for parameters that are calibrated outside the model. Table 7 presents the parameters that are estimated internally. A summary list of all targeted moments is provided in the Appendix Table A.3. The following section discusses the fit of the model.

6 The Benchmark US Economy

In this section we discuss the fit of the model to the distributions of earnings, income and wealth, followed by a discussion of earning and rate of return processes implied by the calibration. As an overidentification check, we also compare the model's implications for the evolution of earnings, income and net worth over the life-cycle.

6.1 Distributions of earnings, income and net worth

Figure 2 presents the distributions of earnings, income and net worth in the calibrated model (markers) and compares them to the data (lines). Panel (a) shows the marginal distributions for top percentiles of each variable. The model captures the high concentration of net worth very well, even among the top fractiles, as the model markers are almost exactly on the data

[24] For this purpose, we treat the top productivity states z_7, z_8 like f_H, and the top return state κ_{top} like κ_H.

Table 6 – Calibration of the Model: Preset Parameters

Parameter	Description	Value	Source
		Demographics	
J	Maximum life span	16	corresponds to age 100
j_R	Mandatory retirement age	10	corresponds to age 65
s_0, s_1, s_2	Survival probability by age	-5.49, 0.15, 0.016	Halliday et al. (2019)
		Preferences	
σ_c	Risk aversion	1.5	
σ_l	Inverse Frisch elasticity	1.22	Blundell et al. (2016)
		Technology	
δ	Depreciation (annual)	0.045	
		Labor productivity	
		See Section 5.3	
		Taxes and transfers	
τ_c	Marginal corporate tax rate	0.236	Gravelle (2014)
τ_s	Consumption tax rate	0.05	Kindermann and Krueger (2014)
Tr	Government transfers/GDP	2.7%	Kaymak and Poschke (2016)

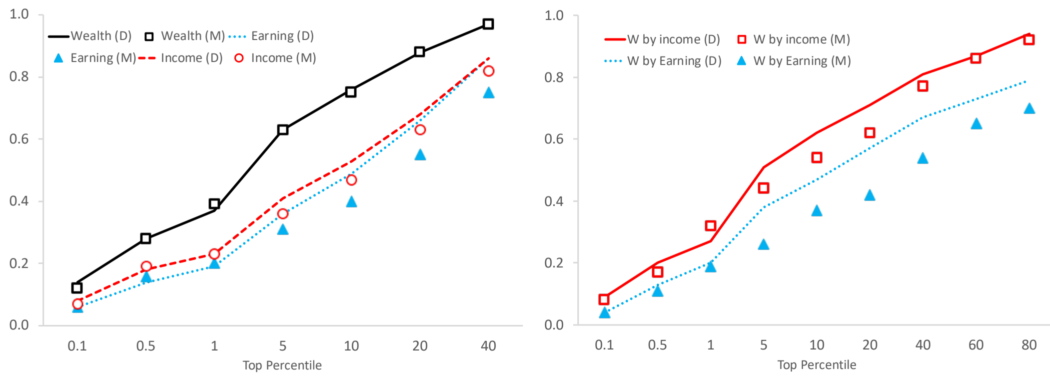
Table 7 – Calibration of the Model: Jointly Calibrated Parameters

Parameter	Description	Value	Parameter	Description	Value
β	Annual Discount rate	0.98	θ	Labor disutility	6.0
α	Capital elasticity	0.27			
z_7, z_8	Top productivity states	Table A.2	$\lambda_{in}, \lambda_{ll}, \lambda_{lh}, \lambda_{hh}$	Transition rates	Table A.2
$\kappa_L, \kappa_H, \kappa_{top}$	Rates of return	Table 9	$\pi_{ll}, \pi_{hh}, \pi_{in}, \pi_{top,top}$	Transition rates	Table 9
τ_l	Tax progressivity	0.183	d_c	Corporate asset threshold	2.21
ξ	Pension / Earnings	0.62	G/Y	Expenditures / GDP	15.5%
ϕ_1, ϕ_2	Bequest utility	-0.42, 0.19	$\bar{\gamma}_z, \bar{\gamma}_\kappa$	Bequest correlations	0.65, 0.9

Figure 2 – Distribution of Wealth, Income and Earnings

(a) Cross-sectional Distributions

(b) Wealth by Income and Earnings



Note.– Panel (a) shows the cumulative shares for the top percentile groups. Panel (b) shows share of net worth held by top income and earning groups. Data values come from SCF 2010 and 2016. Income includes capital gains.

line. The overall Gini coefficient for net worth is 0.83 in the model, which is very close to the 0.85 in the data. Similarly, the concentration levels of income and earnings for top groups is in line with the data. Panel (b) shows the shares of net worth held by different income and earning groups. The model generates a strong correlation between income and net worth, as observed in the data, and closely matches their joint distribution. The model also captures the strong connection between earnings and net worth.

Next, we compare the factor composition of income for different income groups. Table 8 shows the share of labor income in total income for various income groups. The labor component of income is 63% in the model for the top 1% of incomes, close to the 64% observed in the data. The labor share for the top 1% wealthiest households in the model, which is not targeted in the calibration, is 48%, close to the data value of 55%.

Overall, the model captures the distributions of earnings, income and net worth. In particular, it features a highly skewed tail of the net worth distribution, generates a realistic correlation between earnings and net worth and a realistic share of income from labor for top income groups. Next, we discuss the stochastic processes for labor efficiency and the rate of return on assets implied by the calibration procedure.

Table 8 – Share of Income from Labor

	All	Top Percentiles			Quintiles				
	0-100	99-100	95-99	90-95	5th	4th	3rd	2nd	1st
Data	0.82	0.64	0.78	0.88	0.78	0.93	0.91	0.82	-0.46
Model	0.79	0.64	0.81	0.78	0.77	0.85	0.84	0.78	0.07

Notes.– Data comes from the 2010 and 2016 waves of the SCF. See text for details.

6.2 Labor productivity process

The extraordinary productivity states are critical for generating the concentration of earnings observed in the data. In the model, workers in these states (z_7 and z_8) are 29 and 288 times as productive as the average worker, and they represent 0.63% and 0.02% of the population at the stationary state. Earnings are a combination of productivity and hours worked. The earnings levels for the top 0.1, 0.5 and 1 percent of earners are 61, 30 and 19 times the average in the model, very close to the levels of 60, 28 and 19 times the average observed in the data. The model therefore features realistic skewness of the earnings distribution.

Each period, an ordinary worker has a 0.2% chance to experience an extraordinary productivity boost. This state is about as persistent as ordinary productivity states. The probability of remaining among the top 1% of earners after 5 years is 59% in the model. Kopczuk et al. (2010) estimate this probability to be 62% at the individual level using data on earnings from the Social Security Administration.[25]

Another way to test the dynamic properties of the productivity process is to compare the distribution of earnings growth in the model with that in the data. Guvenen et al. (2019b) document that earnings growth of the top 1% of individual earners is characterized by a large standard deviation, a high degree of kurtosis, and negative skewness. The model replicates the standard deviation of earnings growth for this group of 1.7 exactly, even though this moment was not targeted. It generates a skewness of -2.9 and a kurtosis of 11.4, comparable to the estimates in Guvenen et al. (2019b) of -1.3 and 8.3.

Overall, the estimated earning process captures fundamental properties of the earnings

[25]The transition matrix for the earnings process and the earnings levels implied by the calibration procedure are shown in the Appendix in Table A.2.

Table 9 – The Transition Matrix for the Rate of Return on Assets

	κ_L	κ_H	κ_{top}
κ_L	0.99	0.00975	0.00025
κ_H	0.00975	0.99	0.00025
κ_{top}	0	0.10	0.90
pop. fraction (%)	49.2	50.5	0.25
return level (annualized)	0.01	0.06	0.245

distribution well. It closely matches the cross-sectional distribution of earnings, while also capturing the dynamic aspects of earnings growth.

6.3 Rate of return heterogeneity

The levels of rates of return on assets and the corresponding transition matrix are shown in Table 9. The three return levels are 1%, 6% and 24.5%. 0.25% of households enjoy the top rate of return. The lower rates are highly persistent, with a 99% probability that the rate of return remains the same in the following period. The corresponding persistence for the top category is 90%. The levels of these returns appear plausible for a model with three return states. Combined with the very high persistence, one could think of three types of households: one who only invests in a savings account, a second one who holds stocks, and a small third group who, for some time, has access to a very lucrative investment opportunity.

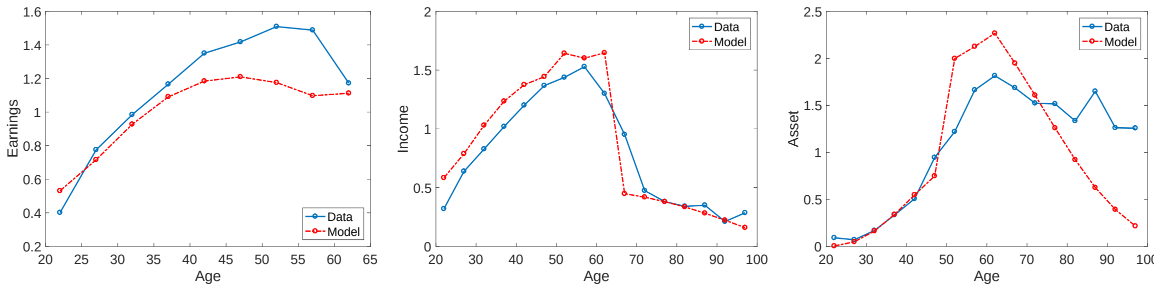
There are no comparable measures of returns for the US. Fagereng et al. (2020) report some statistics on the distribution of asset returns in Norway. The value-weighted average rate of return observed in the Norwegian data is 3.8% with a standard deviation of 8.6% across households. In the model, the value-weighted average rate of return is 4.9% with a standard deviation of 3.3%.

6.4 Implications for life-cycle dynamics

Next, we analyze the model's implications for the evolution of income and wealth over the life-cycle, and compare it with the data. Note that age-dependent distributions of income and wealth are not targeted in the calibration. Therefore, this analysis provides an

Figure 3 – Earnings, Income and Wealth over the Life-Cycle

(a) Earnings (b) Income (c) Net Worth



Note.– Solid lines depict the life-cycle profiles of average earnings, income and net worth implied by the benchmark calibration. Dashed lines show the data values from the SCF.

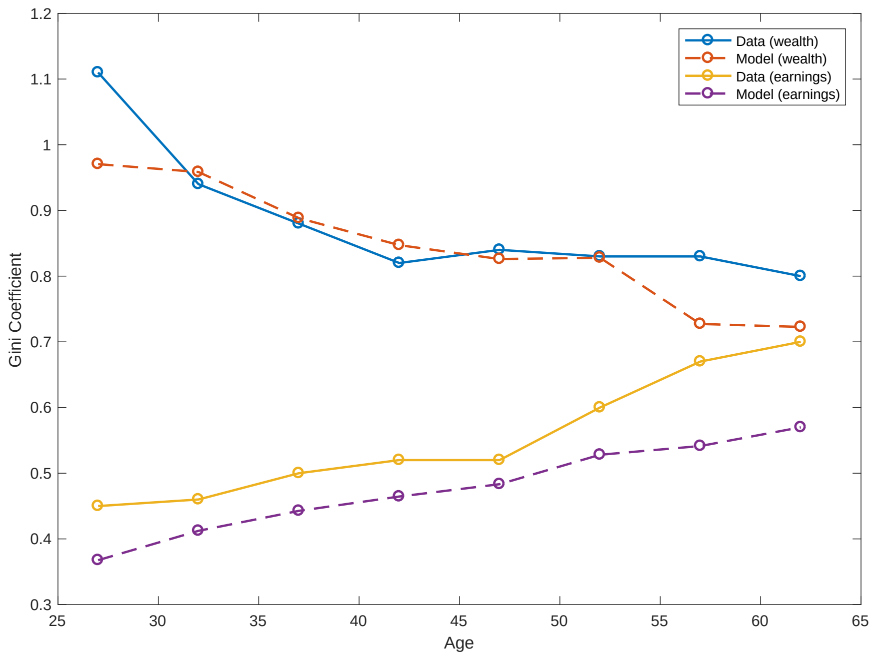
overidentification test of our model.

Figure 3 shows average earnings, income and wealth by age group in the model and compares it with data from the SCF. The labor productivity process is calibrated to match the age profile of wages in the PSID. The age profile of earnings depicted in Figure 3a is a result of households' labor supply decisions given the wage rates and their assets. This is the primary source of income for young households as their assets are initially close to zero. With age, households accumulate assets, and start generating investment income. Average wealth increases up until the retirement age. After retirement, agents rely only on capital income, and start consuming out of their savings. The model accurately captures the salient features of the life-cycle dynamics of income and wealth (apart from the well-known issue that life-cycle models predict too quick wealth decumulation in retirement; see e.g. De Nardi et al. (2009) and Kopecky and Koreshkova (2014)). The fact that the calibrated model closely replicates earnings, income and wealth patterns over the life cycle demonstrates its ability to accurately capture the labor supply and savings behavior among households.

Figure 4 shows the evolution of earnings and wealth dispersion in the model in comparison with the data. The rise in the dispersion of earnings is governed by the productivity process described in Table A.2. Earnings inequality grows mainly because the wages of young households are similar to each other. With age, some households move to higher earning states, and some to top earning states.

The Gini for wealth is initially very high. This is because most households initially have little assets and weak saving motives in anticipation of earnings growth. Ideally,

Figure 4 – Earnings and wealth inequality over the Life-Cycle



they would have preferred to borrow to smooth their consumption over the life-cycle if it weren't for the borrowing constraint. The presence of many households without assets delivers a high Gini coefficient. With age, earnings grow and retirement approaches. As a result, asset accumulation becomes more prevalent among households. This reduces the Gini coefficient in the first part of the life-cycle. About 15-20 years after market entry, the reduction in wealth Gini is counteracted by the increasing dispersion in earnings and income, which raises wealth dispersion. These two forces are more or less equivalent, resulting in a stable dispersion of wealth for middle-aged groups and older, as in the data.

The model also generates plausible age profiles of wealth across the distribution. For instance, the average age in the top 1% of wealth is 62.9 years in the model, compared to 61.6 in the data (Kuhn and Ríos-Rull, 2016).

Overall, the benchmark economy provides an accurate description of the distributions of earnings, income and wealth. The productivity process captures the salient features of earnings growth both in the short run and over the life-cycle. The factor composition of income is realistic, including at the top of the distribution. The implied wealth distribution is highly concentrated at the top and correlated with earnings and income, as in the data.

Next, we provide a quantitative assessment of the relative importance of different modeling approaches to wealth concentration.

7 Determinants of Wealth Concentration

To quantify the relative roles of earning concentration, rate of return differences and bequests in shaping wealth concentration, we conduct two experiments. First, we shut down different model components and compare the implied wealth concentration with the data. This allows us to decompose wealth concentration into its components. Second, we force the model to match the observed wealth concentration by shutting down one model component and recalibrating the other. Then, we contrast the implied joint distribution of earnings and wealth with the data. This allows us to highlight the sources of identification in our benchmark calibration.

7.1 Decomposition analysis

Table 10 shows the results from the first decomposition exercise. The first two rows report the moments in the data and from the benchmark economy for comparison. Each of the remaining rows takes away critical model components, and reports the resulting moments on the concentration of earnings and wealth as well as the top 1% labor income share in the counterfactual economy.

We begin by removing the top earning states by setting $z_8 = z_7 = z_6$. This preserves the wage distribution in the ordinary productivity states, and sets the productivity of the two extraordinarily productive states to that of the highest "regular" state. In this scenario, the concentration of earnings is much lower than in the data, with a top 1% earnings share of only 4%, compared to 18% in the benchmark economy. Top incomes are now driven mainly by capital income. As a result, the labor income share among the top 1% income earners falls from 63% to 47%. Elimination of superearners also markedly reduces the concentration of wealth. The top 1% wealth share falls from 38% to 16%, and the Gini coefficient from 0.83 to 0.74. Because of return heterogeneity, the top 0.1% wealth share is still significant at 7%, half the benchmark value.

The next line shows that both top productivity states matter. While the top state (z_8) has a small effect on the Gini, which mostly reflects wealth inequality in the middle of the wealth distribution, it strongly affects the top 0.1% share, which in this scenario is 6

Table 10 – Determinants of Wealth Concentration: A Decomposition Analysis

	wealth Gini	top wealth shares		top earnings shares		top 1% LIS
		0.1%	1%	0.1%	1%	
data	0.85	0.14	0.37	0.06	0.19	0.64
benchmark	0.83	0.14	0.38	0.06	0.18	0.63
counterfactual economies with …						
…(1) no top earners ($z_8 = z_7 = z_6$)	0.74	0.07	0.16	0.004	0.04	0.47
…(2) no superstar earners ($z_8 = z_7$)	0.81	0.08	0.32	0.03	0.15	0.62
…(3) no return differences (same κ)	0.79	0.11	0.34	0.06	0.18	0.72
…(4) no top return state ($\kappa_{top} = \kappa_H$)	0.82	0.12	0.34	0.06	0.18	0.71
…(5) homothetic bequests	0.80	0.13	0.36	0.06	0.18	0.65
…(6) uncorrelated bequests	0.82	0.13	0.36	0.06	0.18	0.65
…(7) equal bequests	0.73	0.11	0.30	0.06	0.19	0.69
…(8) no top earn. & no ret. heterogeneity	0.67	0.01	0.08	0.004	0.04	0.83
…(9) no top earn. & equal bequests	0.65	0.03	0.10	0.005	0.04	0.53
…(10) no ret. het. & equal bequests	0.71	0.10	0.29	0.06	0.18	0.76
…(11) remove all three channels	0.62	0.01	0.08	0.01	0.04	0.84

Note.– Results from model simulations. LIS denotes labor income share of top 1% of incomes. Row (3) sets κ to its value-weighted average in the benchmark economy.

percentage points (ppt) lower than in the benchmark. The next state (z_7), which contains more households, has a large effect on the top 1% wealth share.

Other factors have a more moderate impact on wealth concentration. First, we simulate the effect of fully eliminating the heterogeneity in the rate of return by setting κ to its value-weighted average in the benchmark economy for all households. Doing so reduces the Gini coefficient for wealth to 0.79, the top 1% wealth share to 34%, and the top 0.1% share to 11%. In this scenario, top incomes are mainly driven by labor income, and the labor income share among the top 1% income earners rises to 72%. In the next line, we eliminate only the top return state by setting κ_{top} to κ_H. Results in this line are almost identical to the previous line, except for the Gini coefficient for wealth. This shows that the top return state affects wealth concentration at the top of the distribution, whereas κ_H affects the middle of the distribution and thus the Gini coefficient, but has only a small impact at the top. Overall, the contribution of rate of return heterogeneity is modest.

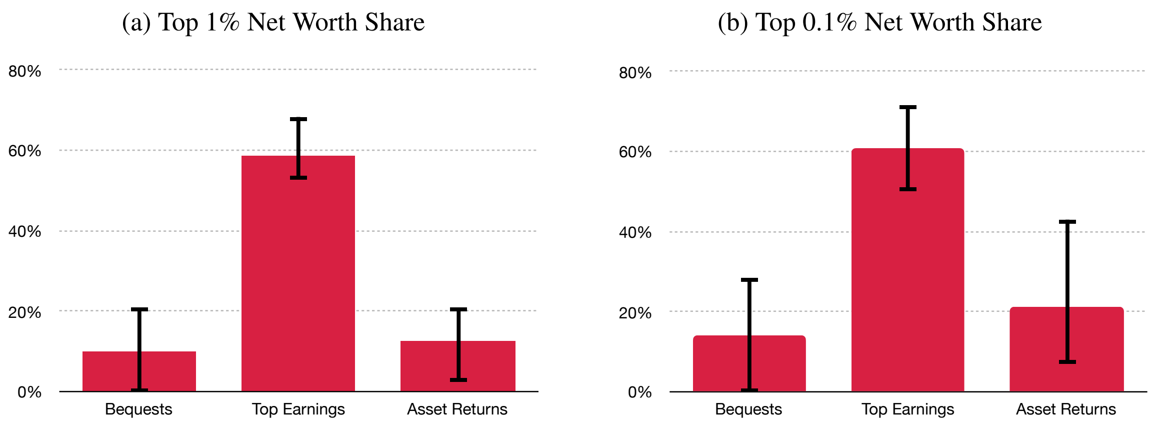
Finally, we investigate the effects of the bequest motive and our modeling assumptions regarding the distribution of bequests and intergenerational links. We first set ϕ_2 to zero, making the bequest motive homothetic. Next, we explore the effects of removing the correlation of bequests by setting both $\bar{\gamma}_z$ and $\bar{\gamma}_\kappa$ to 0.5, so that all bequest recipients draw their bequest from the same distribution. Finally, we compute results for an economy where bequests are equally distributed, as common in overlapping generations models.[26] In all three cases, we adjust ϕ_1 such that the overall importance of bequests in the economy, as measured by bequests relative to wealth, is unchanged.[27] The first two changes lead to almost identical results, namely slightly lower top wealth shares and a small decline in the Gini coefficient for wealth compared to the benchmark economy. The largest changes occur when imposing equal bequests. In this case, the Gini coefficient declines by almost 0.1, and the top wealth shares fall by one fifth to one quarter. The labor income share of the top 1% income earners increases slightly in this case, as fewer top income earners have had a large bequest. Overall, bequests clearly have a smaller quantitative impact than the top earning states. Still, bequest inequality has a significant impact on the wealth distribution, as it perpetuates the inequality of wealth accumulated in a generation.

Eliminating the different factors individually may mask potential interactions between them. The last four counterfactuals in Table 10 show wealth concentration in economies where two features are removed at the same time. (All three features in the last row.) First,

[26]Exceptions are, among a few others, De Nardi (2004) and De Nardi and Yang (2016).

[27]This change is minor and hardly affects our results.

Figure 5 – Factors of Net Worth Concentration



Note.– Figure shows the marginal contribution of each factor to the concentration of net worth. Marginal contributions differ depending on the order in which factors are eliminated from the benchmark economy in Table 10. The whiskers show the range of contributions for each factor. The column height represents the average.

we remove both top earners and return heterogeneity. The effect on wealth concentration is quite dramatic. Since without top earners alone, the top 1% wealth share is 16%, the marginal effect of return heterogeneity here is 8ppt. The marginal effect of top earners is 26ppt. Second, we remove both top earners and equalize bequests. The effect on wealth concentration is similarly dramatic in this setting, but the marginal effect of each component is similar to the scenarios where it is removed individually. Third, we remove return heterogeneity and equalize bequests. These results are similar to those with equal bequests, but heterogeneous returns. Finally, we remove all three channels. In this case, top wealth shares are close to their counterparts in the scenario with bequests, but neither top earners nor return heterogeneity.

The marginal contributions of the three channels to measures of wealth concentration is summarized in Figure 5. The left (right) panel shows the marginal effect of removing a channel on the top 1% (top 0.1%) wealth share. For each channel, four different marginal effects can be computed based on Table 10. For example, top earners can be removed starting in a situation where all channels are active, where only two channels are active, or where only the top earner channel is active. For each channel, the figure shows the average of these four marginal effects as a bar, and the smallest and largest marginal effect as whiskers. The marginal effects are expressed as a fraction of the benchmark value. By all measures, the contribution of top earners to measures of wealth concentration is large, as removing top earners leads to declines in the top 0.1% and top 1% wealth shares by half or more in all scenarios. The marginal effect of unequal bequests ranges from zero to almost a third. Note that unequal bequests on their own do not lead to much wealth concentration. This shows that they mostly amplify the other channels, by perpetuating wealth inequality that is created within a generation. The marginal effect of return heterogeneity on top wealth shares averages about 20 percent. It is largest when there are no top earners. These results highlight the importance of the top earning states, not only for their direct effect on the wealth distribution, but also for the assessment of the importance of the other channels.

7.2 Return heterogeneity in a life cycle model

It may be somewhat surprising that the differences in rates of return on assets do not generate a larger impact on wealth concentration. As we demonstrate below, the human life is too short for the calibrated rates of return to have a larger impact, unless agents start their life already with substantial asset holdings or income. This is in line with the inherent property of models of return heterogeneity that the wealth distribution fans out slowly (Gabaix et al., 2016).

To make this point in our context, we simulate patterns of wealth accumulation for agents with different rates of return, using the policy function for savings in the benchmark economy. For simplicity, and to consider a scenario where the effect of return heterogeneity is maximal, we fix the rates of return and productivity levels. We set productivity at the highest regular productivity state z_6 to generate high incomes to save from, while abstracting from top productivity states. Unless otherwise indicated, we also assume that the household receives a bequest of average size at age 50.

Figure 6a shows the resulting paths of assets relative to mean wealth by age for the three rates of return, on a log scale. Bear in mind that the average wealth of the top 1% of wealthiest households is 37 times the mean wealth, and that of the wealthiest 0.1% is 140 times the mean wealth. As expected, the figure shows that households with higher returns save more from early on.

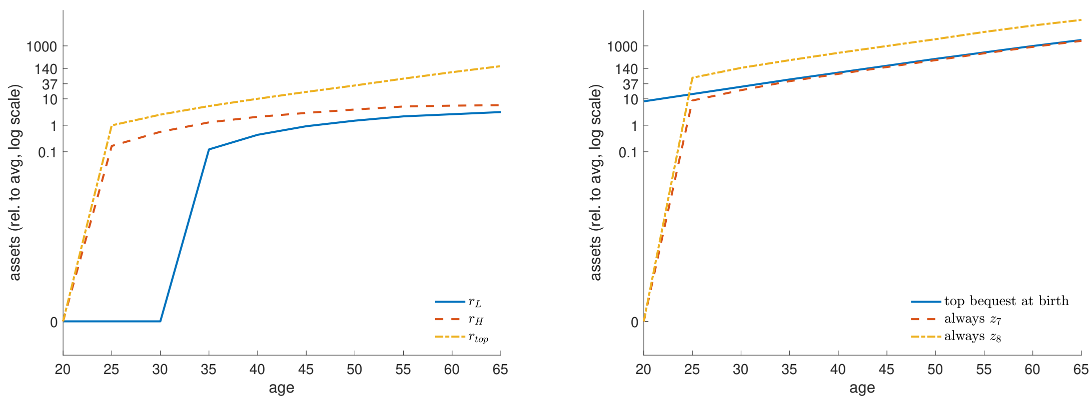
However, asset growth takes time. As a result, a household with the top return of 24.5% catches up with the wealthiest 1% only by age 55, and with the wealthiest 0.1% by age 65. A household that always earns r_H does not even get close to that, in line with the limited importance of κ_H for top wealth shares shown in the previous section. This is why the role of return heterogeneity for top wealth shares is limited.

Figure 6b shows that asset accumulation is much faster if labor earnings are high or

Figure 6 – Asset accumulation with fixed return and productivity

(a) benchmark: $z = z_6$, fixed return

(b) Higher bequest or productivity



agents already start with high assets. The broken lines show paths of assets for households who permanently earn the top return and have a productivity of z_7 or z_8. Because their earnings are higher, their assets catch up with the top 0.1% by age 45 (z_7) or even by age 30 (z_8). This illustrates the importance of earnings in generating wealth. (Of course, chances of obtaining both z_8 and κ_{top} are infinitesimal.)

The solid line shows the path of assets for a household who starts life with large wealth holdings, equivalent to 37 times the average bequest, and who permanently has productivity z_6. This household starts close to the top of the wealth distribution, reaches top 1% asset levels by age 35, and top 0.1% asset levels by age 45. The asset path of this household is close to what might obtain for a wealthy household in an infinite-horizon model. It contrasts strongly with that of the high-return household in Figure 6a, who does not start life wealthy and receives a bequest at a realistic age. The life cycle of wealth, which our model replicates well (recall Figure 3), is thus key for the potential (or lack thereof) of heterogeneous returns to drive top wealth. These findings are in line with results from the theoretical literature on the topic, in particular the importance of scale dependence (the correlation of income levels and returns) shown by Gabaix et al. (2016) and the role of factors affecting the intergenerational transmission of wealth shown by Benhabib et al. (2011).

Table 11 – Alternative calibrations that match the top 0.1% wealth share

	top wealth shares		top earnings shares		top 1% LIS by		correlation of wealth with	
	0.1%	1%	0.1%	1%	income	wealth	earnings	income
data	0.14	0.37	0.06	0.19	0.64	0.55	0.35	0.52
benchmark	0.14	0.38	0.06	0.18	0.63	0.48	0.20	0.65
alternative calibrations:								
...(1) no return het.	0.14	0.35	0.07	0.19	0.72	0.57	0.36	0.49
...(2) no top earners	0.14	0.22	0.004	0.04	0.31	0.07	0.01	0.67

Note.– Table shows simulation results from an economy without return heterogeneity or without top earning states. In alternative calibration (1), κ is set to its asset-weighted mean in the benchmark economy for all households. z_8 is set to 550 (1.45 times its benchmark value) to match the top 0.1% wealth share of 14%. In alternative calibration (2), $z_8 = z_7 = z_6$. The top return is raised to 28% per year to match the top 0.1% wealth share.

7.3 Alternative calibrations and the labor income share

Finally, we consider two alternative calibrations where we shut down either the top return or the top productivity states, and partially recalibrate the model to compensate. In the first case, we set κ to its asset-weighted mean from the benchmark economy for all households. To ensure that the model still matches top wealth concentration, as measured by the top 0.1% wealth share, we adjust the value of the top productivity state z_8. This requires raising it by 45%. In the second case, we eliminate the top productivity state by setting $z_8 = z_7 = z_6$, and adjust the value of the top return κ_{top} to match the top 0.1% wealth share. This implies a top return of 28% per year.

By construction, both alternative calibrations match the observed top 0.1% wealth share. As shown in Table 11, they deviate from the benchmark calibration, and thus from the data, in many other dimensions. This is particularly pronounced for the calibration without the top productivity state. In this case, the economy features very little concentration of earnings and an unrealistically low share of labor income for top income and wealth groups. Among top income groups, the implied labor share of income is 0.31, compared to 0.63 in the benchmark. Top wealth groups rely almost exclusively on capital income, with a labor share of income of 7% compared to 48% in the benchmark. The correlation of wealth with earnings drops to virtually zero.

In contrast, an economy with homogeneous returns features an excessively high labor income share at the top and a top 0.1% earnings share that, at 7%, slightly exceeds the benchmark and data values of 6%. It also features an intergenerational wealth correlation that, at 7%, lies significantly below its benchmark value. Otherwise, this alternative calibration does not deviate much from the benchmark. Table A.4 in the Appendix shows that these results are similar when the alternative calibrations keep the top 1% wealth share as in the benchmark.

Taken together, these exercises illustrate how our empirical approach allows identifying the quantitative drivers of wealth concentration. Results are unambiguous: while there is clear evidence of some importance of heterogeneous returns, stemming from the labor income share at the top, overall wealth concentration is to a large extent driven by the concentration of earnings.

8 Discussion

Our findings indicate a significant role for differences in labor income and earnings risk in explaining the observed dispersion in net worth in the US. This is driven essentially by the high concentration of earnings and the large share of earnings in total income among the top income and wealth groups in the data.

The relevance of earnings for wealth concentration warrants a deeper analysis of the determinants of the concentration of earnings. Routes of inquiry that appear promising include human capital accumulation by top earners (Huggett and Badel, forthcoming; Karahan et al., 2019), labor market frictions, in particular among low earnings groups (Karahan et al., 2019), and changes in the degree of assortative matching both between workers and firms and among workers across firms (Song et al., 2019).

Earnings concentration is partly driven also by the concentration of entrepreneurial incomes. In fact, in an economy where entrepreneurs are endowed with a diminishing-return-to-scale production function and do not face any credit restrictions when investing in their businesses, differences in entrepreneurial productivity are fully reflected in the labor component of income. This is because the optimal investment in the business requires that the marginal return to business investment be equal to the common market return on capital. Business income then simply is the sum of a capital income component with a common return and a labor income component that varies across entrepreneurs (see Appendix B for the derivation). From the perspective of our model and data analysis, productivity differences are then captured as earnings differences.

We also find evidence of differences in rates of return on assets at the household level. An unusually high rate of return for a small group of households is required to explain the very top tail of the wealth distribution. To some extent, higher rates of return could reflect the capacity to handle high exposure to risk among some households. They could also capture variation in the marginal return on capital arising from credit constraints on entrepreneurial investment. In our model, these are picked up as differences in the return to capital. Under such an interpretation, the extent of the variation in rates of return depends on the tightness of credit constraints and on productivity differences among constrained entrepreneurs. Changes in these two factors affect both entrepreneurs' labor and capital income (see Appendix B for a brief discussion). This interdependence between an entrepreneur's labor and capital income makes it impossible to counterfactually eliminate heterogeneity in only one income type for entrepreneurs. Eliminating credit constraints would eliminate rate of return differences. It would also amplify earnings differences. As a consequence, results from our analysis of the effect of return heterogeneity on wealth concentration, which kept earnings differences constant, could be considered an upper bound. Eliminating productivity differences would reduce both earnings differences and rate of return heterogeneity. Therefore, our results on the effect of top earnings on wealth concentration could be considered a lower bound. An analysis of wealth concentration that models earnings concentration and rate of return heterogeneity among entrepreneurs more explicitly while matching the joint distribution of income and wealth remains an open topic for future research.

References

Aiyagari, Rao (1994) "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, Vol. 109, No. 3, pp. 659–684.

Bach, Laurent, Laurent E Calvet, and Paolo Sodini (2016) "Rich pickings? Risk, return, and skill in the portfolios of the wealthy," CEPR Discussion Paper.

Bakış, Ozan, Barış Kaymak, and Markus Poschke (2015) "Transitional dynamics and the optimal progressivity of income redistribution," *Review of Economic Dynamics*, Vol. 18, No. 3, pp. 679–693.

Benabou, Roland (2002) "Tax and Education Policy in a Heterogeneous-Agent Economy:

What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, Vol. 70, No. 2, pp. 481–517.

Benhabib, Jess, Alberto Bisin et al. (2006) "The distribution of wealth and redistributive policies," *New York University, Working Paper*.

Benhabib, Jess, Alberto Bisin, and Mi Luo (2019) "Wealth distribution and social mobility in the US: A quantitative approach," *American Economic Review*, Vol. 109, No. 5, pp. 1623–47.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2011) "The distribution of wealth and fiscal policy in economies with finitely lived agents," *Econometrica*, Vol. 79, No. 1, pp. 123–157.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten (2016) "Consumption inequality and family labor supply," *American Economic Review*, Vol. 106, No. 2, pp. 387–435.

Bover, Olympia (2010) "Wealth Inequality and Household Structure: U.S. vs. Spain," *Review of Income and Wealth*, Vol. 56, No. 2, pp. 259–290.

Brüggemann, Bettina (2017) "Higher Taxes at the Top: The Role of Entrepreneurs," McMaster University.

Buera, Francisco J (2009) "A dynamic model of entrepreneurship with borrowing constraints: theory and evidence," *Annals of Finance*, Vol. 5, No. 3-4, pp. 443–464.

Cagetti, Marco and Mariacristina De Nardi (2006) "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, Vol. 114, No. 5, pp. 835–870.

Cagetti, Marco and Mariacristina de Nardi (2009) "Estate Taxation, Entrepreneurship, and Wealth," *American Economic Review*, Vol. 99, No. 1, pp. 85–111.

Cao, Dan and Wenlan Luo (2017) "Persistent heterogeneous returns and top end wealth inequality," *Review of Economic Dynamics*, Vol. 26, pp. 301–326.

Castañeda, A., J. Díaz-Giménez, and J. V. Ríos-Rull (2003) "Accounting for the U.S. Earnings and Wealth Inequality,"," *Journal of Political Economy*, Vol. 111, No. 4, pp. 818–857.

Charles, Kerwin and Erik Hurst (2003) "The Correlation of Wealth across Generations," *Journal of Political Economy*, Vol. 111, No. 6, pp. 1155–1182.

De Nardi, Mariacristina (2004) "Wealth inequality and intergenerational links," *Review of Economic Studies*, Vol. 71, No. 3, pp. 743–768.

De Nardi, Mariacristina and Giulio Fella (2017) "Saving and wealth inequality," *Review of Economic Dynamics*, Vol. 26, pp. 280–300.

De Nardi, Mariacristina, Eric French, and John Bailey Jones (2009) "Life Expectancy and Old Age Savings," *American Economic Review*, Vol. 99, No. 2, pp. 110–15.

De Nardi, Mariacristina and Fang Yang (2016) "Wealth inequality, family background, and estate taxation," *Journal of Monetary Economics*, Vol. 77, pp. 130–145.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2020) "Heterogeneity and Persistence in Returns to Wealth," *Econometrica*, Vol. 88, No. 1, pp. 115–170.

Feiveson, Laura and John Sabelhaus (2018) "How Does Intergenerational Wealth Transmission Affect Wealth Concentration?" *FED Notes*, No. 2018-06-01.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2016) "The dynamics of inequality," *Econometrica*, Vol. 84, No. 6, pp. 2071–2111.

Galor, Oded and Joseph Zeira (1993) "Income distribution and macroeconomics," *Review of Economic Studies*, Vol. 60, No. 1, pp. 35–52.

Gayle, George-Levi, Andrés Hincapié et al. (2016) "Which Persists More from Generation to Generation—Income or Wealth?" *The Regional Economist*, Vol. July, pp. 14–15.

Gravelle, Jane (2014) *International Corporate Tax Rate Comparisons and Policy Implications*, Washington, DC: Congressional Research Service.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura (2014) "Income taxation of US households: Facts and parametric estimates," *Review of Economic Dynamics*, Vol. 17, No. 4, pp. 559–581.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen (2019a) "Use it or lose it: Efficiency gains from wealth taxation."

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2019b) "What Do Data on Millions of U.S. Workers Reveal About Life-Cycle Earnings Risk?" *Working Paper*, mimeo.

Halliday, Timothy J, Hui He, Lei Ning, and Hao Zhang (2019) "Health Investment over the Life-Cycle," *Macroeconomic Dynamics*, Vol. 23, No. 1, pp. 178–215.

Heathcote, J., K. Storesletten, and G. Violante (2010) "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, Vol. 118, No. 4, pp. 681–722.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante (2009) "Quantitative macroeconomics with heterogeneous households," *Annual Review of Economics*, Vol. 1, No. 1, pp. 319–354.

——— (2017) "Optimal tax progressivity: An analytical framework," *The Quarterly Journal of Economics*, Vol. 132, No. 4, pp. 1693–1754.

Heaton, John and Deborah Lucas (2000) "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance*, Vol. 55, No. 3, pp. 1163–1198.

Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr (2016) "The historical evolution of the wealth distribution: A quantitative-theoretic investigation," *National Bureau of Economic Research Working Paper*.

Huggett, M. (1993) "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, Vol. 17, No. 5-6, pp. 953–69.

Huggett, M. and A. Badel (forthcoming) "Taxing Top Earners: A Human Capital Perspective," *Economic Journal*, Federal Reserve Bank of St. Louis Working Paper: 2014-17.

Huggett, Mark (1996) "Wealth distribution in life-cycle economies," *Journal of Monetary Economics*, Vol. 38, No. 3, pp. 469–494.

Imrohoroglu, A. and S. Imrohoroglu (1995) "A life cycle anlaysis of social security," *Economic Theory*, Vol. 6, No. 1, pp. 83–114.

Jones, Charles I (2015) "Pareto and Piketty: The macroeconomics of top income and wealth inequality," *Journal of Economic Perspectives*, Vol. 29, No. 1, pp. 29–46.

Karahan, F., S. Ozkan, and J. Song (2019) "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital," December, Federal Reserve Bank of New York Staff Report, No. 908.

Kaymak, Barış and Markus Poschke (2016) "The evolution of wealth inequality over half a century: The role of taxes, transfers and technology," *Journal of Monetary Economics*, Vol. 77, pp. 1–25.

Kindermann, Fabian and Dirk Krueger (2014) "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," NBER Working Paper No: 20601.

Kopczuk, W., E. Saez, and J. Song (2010) "Earnings Inequality and Mobility in the {United {States: Evidence from Social Security Data since 1937," *Quarterly Journal of Economics*, Vol. 125, No. 1, pp. 91 – 128.

Kopecky, Karen A and Tatyana Koreshkova (2014) "The impact of medical and nursing home expenses on savings," *American Economic Journal: Macroeconomics*, Vol. 6, No. 3, pp. 29–72.

Krusell, Per and Anthony A Smith, Jr (1998) "Income and wealth heterogeneity in the macroeconomy," *Journal of Political Economy*, Vol. 106, No. 5, pp. 867–896.

Kuhn, Moritz and JV Ríos-Rull (2016) "2013 Update on the US earnings, income, and wealth distributional facts: A View from Macroeconomics," *Quarterly Review, Federal Reserve Bank of Minneapolis, April*, pp. 1–75.

Moll, Benjamin (2014) "Productivity losses from financial frictions: Can self-financing undo capital misallocation?" *American Economic Review*, Vol. 104, No. 10, pp. 3186–3221.

Nirei, Makoto and Shuhei Aoki (2016) "Pareto distribution of income in neoclassical growth models," *Review of Economic Dynamics*, Vol. 20, pp. 25–42.

Piketty, Thomas and Emmanuel Saez (2003) "Income Inequality in the United States," *Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 1–39.

——— (2007) "How Progressive is the U.S. Federal Tax System? A Historical and International Perspective," *Journal of Economic Perspectives*, Vol. 21, No. 1, pp. 3–24.

Quadrini, Vincenzo (2000) “Entrepreneurship, saving, and social mobility,” *Review of Economic Dynamics*, Vol. 3, No. 1, pp. 1–40.

Smith, Matthew, Danny Yagan, Owen M Zidar, and Eric Zwick (2019) “Capitalists in the Twenty-first Century,” *Quarterly Journal of Economics*, Vol. 134, No. 4, pp. 1675–1745.

Social Security Administration (2013) *Social Security Bulletin*, Social Security Administration, Washington, DC, vol. 73 edition.

Solon, G. (1992) “Intergenerational Income Mobility in the United States,” *American Economic Review*, Vol. 82, No. 3, pp. 393–408.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter (2019) “Firming Up Inequality,” *The Quarterly Journal of Economics*, Vol. 134, No. 1, pp. 1–50.

Stachurski, John and Alexis Akira Toda (2019) “An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity,” *Journal of Economic Theory*, Vol. 182, pp. 1–24.

for Online Publication

A Supplementary Tables and Figures

Table A.1 – Labor Component of Income by Income and Wealth Group

Income Percentile	99-100	99.9-100	99.5-99.9	99-99.5	95-99	90-95	0-100
Wage Income							
with capital gains	0.49	0.37	0.52	0.60	0.69	0.83	0.74
without capital gains	0.56	0.49	0.56	0.67	0.73	0.84	0.77
Labor Income							
with capital gains	0.59	0.47	0.64	0.69	0.76	0.87	0.80
without capital gains	0.68	0.63	0.68	0.75	0.80	0.89	0.84
Net Worth Percentile	**99-100**	**99.9-100**	**99.5-99.9**	**99-99.5**	**95-99**	**90-95**	**0-100**
Wage Income							
with capital gains	0.40	0.26	0.40	0.56	0.58	0.71	0.74
without capital gains	0.47	0.32	0.46	0.63	0.62	0.74	0.77
Labor Income							
with capital gains	0.51	0.35	0.49	0.69	0.69	0.78	0.80
without capital gains	0.59	0.44	0.64	0.77	0.73	0.81	0.84

Note.– Table shows wage and labor shares of total income by percentiles of the income and net worth distribution. Labor income includes imputed wage income for active business owners who do not draw salary from their businesses. Data comes from the 2010 and 2016 waves of the SCF.

Table A.2 – Productivity Transitions in the Model

	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
$z_1 = 1.00$	0.874	0.119	0.004	0	0	0	0.002	0
$z_2 = 1.97$	0.060	0.878	0.060	0	0	0	0.002	0
$z_3 = 3.89$	0.004	0.119	0.874	0	0	0	0.002	0
$z_4 = 3.24$	0	0	0	0.874	0.119	0.004	0.002	0
$z_5 = 6.39$	0	0	0	0.060	0.878	0.060	0.002	0
$z_6 = 12.61$	0	0	0	0.004	0.119	0.874	0.002	0
$z_7 = 137.36$	0.021	0.021	0.021	0.021	0.021	0.021	0.850	0.021
$z_8 = 1349.46$	0	0	0	0	0	0	0.242	0.758
invariant distribution	0.123	0.245	0.123	0.123	0.245	0.123	0.017	0.001
fraction of working age pop.	0.089	0.318	0.089	0.089	0.318	0.089	0.0063	0.0002

Table A.3 – Summary of Target Moments

Moment	Source	Data Value	Model Fit	Moment	Source	Data Value	Model Fit
Mean hours worked		0.35	0.35	Soc. Sec. Pay / GDP	NIPA, 2010-16 average	7.9%	8.1%
Top 0.1%,1% earning shares	SCF 2010 & 2016	Figure 2	Figure 2	Log wealth correlation between parents and kids	Charles and Hurst (2003)	0.30	0.32
Top 0.1%,1%,5%,10% wealth shares	SCF 2010 & 2016	Figure 2	Figure 2	Gini coefficient of wealth	SCF 2010 & 2016	0.85	0.83
Bequest/Wealth	Guvenen et al. (2019a)	1-2%	1.7%	Top 2% bequest dist.	Feiveson and Sabelhaus (2018)	40%	45%
Difference between average income tax rate for top 1% and 99%	Piketty and Saez (2007)	6.8%	6.4%	Corporate income tax revenue/GDP	NIPA	2.5%	2.0%
Probability of staying in top 1% earners	Kopczuk et al. (2010)	0.62	0.59	Overall labor income share	SCF 2010 & 2016	0.82	0.79
Top 1% labor income share	SCF 2010 & 2016	0.64	0.63	P95-99 labor income share	SCF 2010 & 2016	0.78	0.81
Intergenerational wealth persistence at 4th quintile	Charles and Hurst (2003)	0.26	0.22	Intergenerational wealth persistence at 5th quintile	Charles and Hurst (2003)	0.36	0.32

Table A.4 – Alternative calibrations that match the top 1% wealth share

	top wealth shares		top earnings shares		top 1% LIS by		correlation of wealth with	
	0.1%	1%	0.1%	1%	income	wealth	earnings	income
data	0.14	0.37	0.06	0.19	0.64	0.55	0.35	0.52
benchmark	0.14	0.38	0.06	0.18	0.63	0.48	0.20	0.65
alternative calibrations:								
...(1) no het. returns	0.17	0.38	0.08	0.20	0.70	0.53	0.34	0.49
...(2) no top productivity	0.30	0.38	0.005	0.04	0.20	0.05	0.00	0.62

Note.– Table shows simulations results from an economy without return heterogeneity or without top earning states. In alternative calibration (1), κ is set to its wealth-weighted mean in the benchmark economy for all households. z_8 is set to 703 (1.85 times its benchmark value) to match the benchmark top 1% wealth share of 38%. In alternative calibration (2), $z_8 = z_7 = z_6$. The top return κ_{top} is set to match the benchmark top 1% wealth share. This implies a top return $r\kappa_{top}$ of 31% p.a..

B Entrepreneurship and the Distribution of Earnings and Capital Income

Consider the following portfolio allocation problem for an entrepreneur endowed with a units of assets and a diminishing-return-to-scale business income production function $y^b = \theta k^\alpha$, where θ represents the productivity of the entrepreneur. We implicitly assume that a unit of entrepreneurial labor is supplied inelastically as long as the business is in operation.

$$\max_k y = \theta k^\alpha + r(a - k),$$

where the first term is business income and the second term is market income on excess assets (or debt service if $a < k$ in equilibrium). The optimal business investment k^* solves $\theta \alpha k^{\alpha-1} = r$. Substituting the optimality condition back into the objective function gives:

$$y^* = ra + \underbrace{(1-\alpha)\theta_i^{\frac{1}{1-\alpha}}(r/\alpha)^{\frac{\alpha}{\alpha-1}}}_{non-capital\ income}$$

From the perspective of our approach, this setting is observationally equivalent to a version of our model with a common return on assets and labor income heterogeneity, which here is driven by differences in entrepreneurial ability, θ. Our calibration procedure interprets this as labor income heterogeneity. So to the extent that top income and wealth groups consist in unconstrained entrepreneurs, the cause of wealth concentration is correctly attributed to labor income, which may include entpreneurial earnings.[28]

Next, consider the case where entrepreneurs are constrained by their assets when investing in their business: $k \leq a$. For entrepreneurs with sufficient assets, given their productivity θ, this constraint does not bind, and the argument above applies all the same. If an entrepreneur is constrained, then the optimal investment is $k^* = a$. Let $r_i = \theta \alpha a^{\alpha-1} > r$ denote the marginal return on business capital of a constrained entrepreneur. Then total income of an entrepreneur can be written as:

$$y^* = r_i a + \underbrace{(1-\alpha)\theta_i^{\frac{1}{1-\alpha}}(r_i/\alpha)^{\frac{\alpha}{\alpha-1}}}_{non-capital\ income}$$

From the perspective of the model, variation in the first term across households is captured

[28] Note that, given the production function, capital's share of business income is αy^b, and that $\alpha y^b + r(a - k) = ra$.

as differences in the return on assets, and variation in the second term is captured as differences in labor productivity. The relative shares of labor and capital income are correctly identified. Note, however, that for constrained entrepreneurs, heterogeneity in the rate of return affects not only the capital income component, but also the labor income component of income. In particular, constrained entrepreneurs have lower earnings conditional on productivity, since they cannot scale up their ideas to full capacity. Therefore, eliminating differences in asset returns also raises labor income dispersion. As a consequence, eliminating rate of return differences while keeping earnings heterogeneity unchanged, as we do in our analysis, may overstate the importance of rate of return differences. Similarly, eliminating differences in calibrated productivity levels reduces dispersion in rates of return across households, given the definition of r_i. This implies that eliminating earnings differences while keeping rate of return differences unchanged, as we do in our analysis, may understate the importance of productivity differences.