

Fischer, Thomas

Working Paper

Determinants of Wealth Inequality and Mobility in General Equilibrium

Working Paper, No. 2019:22

Provided in Cooperation with:

Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Fischer, Thomas (2019) : Determinants of Wealth Inequality and Mobility in General Equilibrium, Working Paper, No. 2019:22, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/260292>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Working Paper 2019:22

Department of Economics
School of Economics and Management

Determinants of Wealth Inequality and Mobility in General Equilibrium

Thomas Fischer

December 2019



LUND
UNIVERSITY

Determinants of wealth inequality and mobility in general equilibrium*

Thomas Fischer[†]

Lund, December 2019

What determines inequality and mobility of wealth? This paper quantifies in closed form both the bottom and the top (Pareto) tail of the distribution for a rich continuous-time model. The distribution is especially shaped by bequest motives, demographics, and the asset portfolio composition under idiosyncratic wealth risk. Factors that increase inequality also reduce mobility. The model – enriched by a realistic income process and non-trivial portfolio constraints – is solved in general equilibrium and calibrated to match US evidence. A bequest tax is shown to reduce inequality and increase mobility. Several partial-equilibrium intuitions do not carry over into general equilibrium.

JEL classification: D31 – E21 – H23 – C68 – G11

Keywords: wealth inequality – mobility of wealth – portfolio selection – fat tails – bequest tax

*An earlier version of this paper was circulated under the title "A tractable model of wealth inequality and mobility". Financial support from the Thule Foundation (Skandia) is gratefully acknowledged.

[†]Department of Economics and Knut Wicksell Centre for Financial Studies, Lund University, thomas.fischer@nek.lu.se

1. Introduction

The publication of the popular book *Capital in the 21st century* by Thomas Piketty has revived interest in the distribution of economic resources – especially wealth. In this book, Piketty (2014) documents that wealth inequality – notably in the USA – has been accelerating since the 1980s after a period of moderation. In particular, he documents the share of top-wealth holders, suggesting that the latter is described by a Pareto-distribution. Piketty (2014) argues that wealth inequality is increasing by the measure $r - g$, with r being the return on capital g the aggregate growth rate. In order to halt this evolution, he suggests a global tax on capital.

Piketty (2014) focuses on the role of distribution, i.e. the cross-section of wealth at a given point in time t . However, he also stresses the role of mobility, emphasizing that inheritance links the wealth of generations. Compared to inequality, mobility focuses on a certain individual i and tracks its change over time and thus considers the dimension which stands in opposition to inequality. Recent empirical evidence reported in Clark and Cummins (2015) using novel evidence from rare surnames in England and Wales reports a considerably lower degree of wealth mobility than previously assumed. Piketty (2014) (implicitly) assumes that both outcomes – high inequality and low mobility – go hand in hand. Yet, the latter correlation is not straightforward.

In a cross-country analysis, Corak (2013) finds a negative correlation between income inequality and mobility of income, implying that high inequality is accompanied by low mobility. This relationship has been called the *Great Gatsby Curve*. Yet, comparable evidence for the stock measure of wealth is not available. The latter is of particular interest as it is well known that wealth is more unequally distributed than income and, moreover, in contrast to income wealth can be perfectly transferred between generations in the form of inheritance.

In this paper, we investigate this relationship theoretically within a rich micro-founded *Bewley*-type model analyzed in general equilibrium. Moreover, the impact of a bequest tax as a policy measure is discussed. We find that by setting a wedge between the wealth of an individual and her heirs the tax increases the mobility, while at the same time decreasing the degree of inequality. The tax itself – being highly progressive – directly reduces inequality at the very top, which witnessed a stark increase in recent years. The model emphasizes the role of the portfolio composition and the risk associated with wealth holdings. These issues were usually not captured in the analysis of inequality in

the macroeconomics literature focusing on the role of income as a driver of subsequent wealth inequality.

The discussion of bequest taxes in theoretic models by itself is not new.¹ In our reading, the literature is split into two branches. On the one hand, there are highly stylized models that still rely on clear economic intuition. While classic papers such as Becker and Tomes (1979) and Davies (1986) argue that bequest taxes eventually increase inequality due to hampering the inter-generational redistribution within the family, the newer literature such as Bossmann et al. (2007) and even more recently Wan and Zhu (2019) with joy-of-giving preferences (rather than perfect altruism) argue that redistributive bequest taxes actually decrease overall wealth inequality. Opposed to these stylized models, there are large scale models of the *Bewley*-type calibrated to match empirical evidence such as DeNardi (2004) or Castaneda et al. (2003) that investigate the intergenerational connection in terms of wealth respectively argue that bequest taxes (slightly) decrease wealth inequality. What unifies these two strands of the literature is that they locate the original source of inequality in the labor income. As, however, formally argued in Benhabib et al. (2011) this approach will fail to match the top tails of the wealth distribution. This is well-known for *Bewley*-type models (Aiyagari, 1994) and (partly) addressed by the introduction of *superstar* income states (Castaneda et al., 2003).

Our approach combines both strands of the literature. Despite the richness of the model we are able to provide closed-form solutions in an elaborate model with capital income risk quantifying both inequality and mobility. This model is enhanced by standard features of *Bewley*-type models – a rich income process featuring *superstars* and financial market imperfections imposing portfolio constraints – fitted to US evidence, and analyzed in general equilibrium. As the analysis shows, several intuitive findings of the partial-equilibrium framework do not carry over into general equilibrium. The model is cast in continuous time, which boosts the analytical discussion. This toolbox – in particular the Fokker-Planck equations available in this framework providing the cross-sectional distribution – has been successfully employed recently in order to analyze issues of inequality (Aoki and Nirei (2017), Cao and Luo (2017), Nuño and Moll (2018)). The model features an overlapping generations structure with uncertain life time and a bequest motive in the presence of annuity markets. Agents form a portfolio

¹This paper provides a positive and not a normative analysis. In contrast, Piketty and Saez (2013) discuss the optimal inheritance tax rate.

featuring risk-free assets and risky assets with idiosyncratic capital risk. Moreover, their labor income grows at a given rate g which allows us to analyze Piketty's (2014) $r - g$ proposition.

In the tradition of Benhabib et al. (2011), we can quantify top tails of the Pareto-type driven by multiplicative wealth risk. Interestingly, higher portfolio risk actually decreases inequality as individuals internalize this by reducing their exposure to the risky asset. The bequest tax is shown to reduce inequality, especially at the top. In line with Piketty (2014), the gap $r - g$ increases top inequality. Even more importantly, higher equity premia captured by higher Sharpe ratios contribute to high top wealth inequality. Besides these *economic* factors, we also emphasize the role of demographics in shaping the distribution of wealth. In the face of higher life expectancy due to medical advances, the increased savings amplify top wealth inequality.

Besides the right tail capturing the top wealthy, we also quantify the effect of the left tail of the distribution as a measure of poverty. A larger human capital – e.g. due to higher growth rates g – available to borrow against is shown to increase bottom inequality. Hence, this already counteracts the $r - g$ rationale provided in Piketty (2014).

The existing literature – both empirical and theoretical – focuses mostly on the cross-sectional distribution of wealth. The focus on the notion of *the* top shares is misleading as it suggests a closed community of super rich. In fact, there are a lot of entries into and exits out of the *club* of super rich. Hence, a high level of inequality might be socially acceptable if there were the opportunity of upward mobility which reflects the idea of the *American Dream*. In this analysis, however, we show that measures that increase inequality also have a tendency to decrease mobility. As such high inequality and low mobility come as twins. This suggests a relationship similar to the *Great Gatsby Curve* for the stock measure of wealth. The intuition is that if the distribution is less dispersed it requires only small changes in terms of the currency unit in order to enter a *different class*.

In line with the evidence, the model documents that both the bottom and the upper tail of the distribution display the highest persistence. Benhabib et al. (2019) develop a model featuring capital income risk and showcase its capability to fit the empirical evidence of wealth mobility for the USA. We extend their work by including random deaths, a portfolio decision, and by analyzing the framework in general equilibrium.

The employed continuous-time framework is also helpful as it allows for fast and accurate solutions of the large general equilibrium model (Achdou et al., 2017). We enhance

the model with a rich income process that also matches the tails and intergenerational correlation.² Moreover, realistic portfolio constraints related both to the stock of wealth and the flow of income are considered. This model is then solved in general equilibrium.

As such, the wage rate and especially the interest rates are endogenous. We identified several factors in partial equilibrium that increase savings and, by increasing the growth rate of wealth, will therefore increase wealth inequality. These results do not immediately carry through into general equilibrium. Increased savings also reduce the rate of interest. As a result, a reduced return on assets actually reduces inequality in line with Piketty's (2014) $r - g$ rationale. For example, the general equilibrium analysis suggests that the inequality-reducing property of the bequest tax is lower than suggested by partial equilibrium models such as Benhabib et al. (2014). We also show that relaxed borrowing constraints increase indebtedness and thereby wealth inequality at the bottom.

In order to get a realistic fit to the data and in line with several other calibrated studies (DeNardi (2004), Benhabib et al. (2019)), we impose preferences for bequests to be of the luxury type creating persistent bottom inequality. The further assumption of a minimum consumption desire – making consumption an inferior good relative to bequests – reduces wealth inequality at the bottom and also increases mobility.

The remainder of this paper is organized as follows. Section 2.1 introduces the model, while Section 2.2 gives closed-form solutions for inequality at the bottom and at the top as well as mobility. This model is enhanced in Section 3 by a realistic process of income, portfolio constraints, and a general equilibrium structure. The calibrated version (Section 4) is quantitatively analyzed against US evidence. Factors driving both inequality and mobility are investigated in Section 5. The final section concludes and provides an outlook.

²Hence, we consider both *superstar* income (Castaneda et al., 2003) and capital income risk (Benhabib et al., 2011) as drivers for top inequality. Hubmer et al. (2016) present a rich model that also adds stochastic time preferences in the tradition of Krusell and Smith (1998) in order to fit the top tails. Given the complexity, this model can, however, not be solved analytically. Yet, Toda (2018) formally shows that time-varying discount rate factors are also able to generate Pareto tails. Pugh (2018) runs a *horse-race* of the tree mentioned mechanisms and shows that heterogeneous returns match the evidence best.

2. The model

In this section we present the underlying model. We employ a continuous time approach as it allows for easy-to-interpret closed-form solutions. Having established the individual problem and its solutions in Section 2.1, we present results regarding inequality and mobility in Section 2.2.

2.1. Individual problem

The model closely follows Merton (1971), in which he jointly derives optimal savings / consumption rules as well as portfolios. In the spirit of Richard (1975) we also include bequest motives, uncertain death, and annuity markets.

Assume that there are two assets, a risk-free asset yielding a certain return r and risky asset whose prices $P_{R,t}$ follow a geometric Brownian motion:

$$dP_{R,t} = RP_{R,t}dt + \sigma P_{R,t}dZ_t,$$

with $R > r$ as a risk compensation and dZ_t being the increment to a Wiener process. Agents hold a share μ in the risky asset.³ The aggregate portfolio return is thus given by:

$$\tilde{R} = \mu R + (1 - \mu)r = r + \mu(R - r).$$

Thus, the evolution of wealth W_t is given as follows:

$$dW_t = [rW_t + \mu_t(R - r)W_t + E_t - C_t]dt + \mu_t\sigma W_t dZ_t,$$

for which E_t signifies the earnings (not related to wealth) and C_t current consumption. We assume that earnings follow an exponential growth process ($E_t = E_0 \exp(gt)$) with a given exogenous growth rate g .

Moreover, we assume that agents have a death probability p governed by a Poisson process in the tradition of Blanchard (1985).⁴ Hence, the expected life time is $1/p$. We normalize all variables by the growth rate g and write them with lower case letters

³Note that if agents can borrow at the risk-free rate, $\mu > 1$ implies that agents hold leveraged (i.e. debt-financed) positions in risky assets, whereas $\mu < 0$ implies that they have short positions in the risky asset. Details on the derivation are suspended to Online Appendix B.

⁴The stationary age distribution for the discrete case is labeled Poisson-distribution, whereas in continuous time it is given by an exponential distribution ($f(a) = p \exp(-pa)$).

($e_t \equiv E_t \exp(-gt)$). We can decompose earnings $e = \omega z + Tr$ into a transfer (Tr) and a labor earnings component. Labor earnings depend on the wage rate ω and the labor endowment z . Labor supply is inelastic and earnings are certain, allowing for a neat closed-form solution.⁵ In the extended model we explore the effect of idiosyncratic income risk in the tradition of Aiyagari (1994). More specifically, we model the evolution of labor endowment z as an exogenous Markov chain. In the model, the government redistributes the proceedings from the bequest tax τ_b to all living individuals in a lump-sum manner.⁶

Thus, the constraint finally reads as follows:⁷

$$dw_t = [(r - g)w_t + \mu_t(R - r)w_t + e - c_t - p(1 + \theta)\iota_t]dt + \mu_t\sigma w_t dZ_t.$$

Individuals buy annuities with a value ι_t paid out at the time of death. Therefore, there is a constant stream $p(1 + \theta)\iota_t$ paid to the insurance company, for which θ is the premium charged by the insurance company. In an actuarial fair case (cf. e.g. Yaari (1965)) we would have $\theta = 0$.⁸ Of course the modeling of the demographics is highly stylized. We assume that each dying individual is replaced by a new born keeping the total population size constant. While a particular individual ceases to exist, her dynasty continues to exist forever.⁹

As a utility function for consumption, we assume:

$$u(c) = \frac{(c - \bar{c})^{1-\gamma}}{1-\gamma},$$

for which $\bar{c} > 0$ represents a minimum consumption level.¹⁰ This utility function is of the Hyperbolic Absolute Risk Aversion (HARA) type which – as shown in Merton

⁵As discussed in Bodie et al. (1992), with flexible labor supply (including time of retirement) individuals hold riskier asset portfolios as the labor supply offers them a further degree of adjustment. In our model, this would directly lead to higher wealth inequality.

⁶For a system with a simple linear tax rate τ_b that effectively taxes all $1/p$ periods and then redistributes, transfers are given by $Tr = \tau_b p E(w)$, for which $E(w)$ describes the cross-sectional average wealth.

⁷Note that $e_t = E_0 \equiv e$ is constant in time.

⁸For a reverse insurance the premium would be captured by $-1 < \theta < 0$.

⁹We also abstract from mixing of dynasties by marrying. Note that the marriage market as a device for creating social mobility has lost its importance due to assortative mating (Greenwood et al., 2014).

¹⁰Note that as c is normalized, so is the minimum consumption level \bar{c} , implying that the minimum consumption level grows with the aggregate growth rate g .

(1971) – provides a closed-form solution to the problem. In fact, the utility function is of the decreasing relative risk aversion type, implying that wealthier individuals hold riskier portfolios, also implying higher returns. For the case of no minimum consumption requirement ($\bar{c} = 0$) the standard Constant Relative Risk Aversion (CRRA) function is nested.

The overall utility function also includes a bequest motive and is given by:

$$U(c, b) = \frac{(c - \bar{c})^{1-\gamma}}{1-\gamma} + p\chi \frac{([1 - \tau_b](b + \bar{b}))^{1-\gamma}}{1-\gamma}.$$

The free parameter χ determines the strength of the bequest motive. Individuals form their utility on net of tax bequests. If this were not the case, the optimal decision would be independent of the level of taxation. Here we assume a simple flat tax rate τ_b independent of the level of wealth. In the extended model (cf. Section 3) we, however, consider a more realistic progressive tax system. Finally, we introduce a factor $\bar{b} > 0$, capturing the property of bequests as a luxury good (DeNardi, 2004). Or put differently, for $\bar{c} = \bar{b} = 0$ the utility function would be homothetic.¹¹

For an expectation operator \mathbb{E}_t the objective function is given by:

$$\begin{aligned} & \max_{C_t, B_t, \mu_t} \mathbb{E}_t \int_t^\infty \exp[-(\rho + p)(\tau - t)] U[C_\tau, B_\tau] d\tau \\ & = \max_{c_t, b_t, \mu_t} \mathbb{E}_t \int_t^\infty \exp([(1 - \gamma)g - \rho - p](\tau - t)) U[c_\tau, b_\tau] d\tau. \end{aligned} \tag{1}$$

The discounting includes a pure time preference factor ρ and the probability of death p . We summarize the overall discount rate by $\tilde{\rho} \equiv \rho + p - (1 - \gamma)g$. For a prevailing income effect ($\gamma > 1$) individuals discount the future stronger for higher growth rates g and thus form larger savings. The reverse holds true for a prevailing substitution effect ($\gamma < 1$).

Following Richard (1975) and its more recent application in Benhabib et al. (2014) total bequests are given by:

$$b = w + \iota, \tag{2}$$

¹¹Benhabib et al. (2019) discuss a similar problem (with a certain life length) and an isoelastic utility function for both bequests and consumption, yet they assume heterogeneous coefficients $\gamma_c \neq \gamma_b$. In their calibrated model, they find $\gamma_w < \gamma_c$, also indicating that bequests are a luxury good - i.e. their share increases in wealth.

being the sum of wealth and the insured value in the annuity. Agents with a weak bequest motive take out a reverse life-insurance that pays out a constant stream during the insurer's lifetime ($\iota < 0$). In the case of death, the value $-\iota > 0$ is transferred to the insurance company, reducing the bequest to the heirs. In contrast, individuals with a high bequest motive pay annuities to a life insurance that pays a bulk $\iota > 0$ to their descendants in the case of their death.

Using the relationship in equation 2, we write the optimization problem as a function of b only in the following Hamilton Jacobi Bellman (HJB) equation:¹²

$$\tilde{\rho}V = \max_{c,b,\mu} \{U + V'((p(1+\theta) + r - g)w + \mu(R-r)w + e - c - p(1+\theta)b) + 0.5\sigma^2\mu^2w^2V''\}. \quad (3)$$

The derivation of the (standard) problem is suspended to Online Appendix B. The following proposition summarizes the optimal household plan. Note that this closed-form solution does not consider risky labor income or constraints on the portfolio composition.

Proposition 2.1 (The optimal household plan)

The optimal plan is given by an optimal consumption rule:

$$\begin{aligned} c_t &= \bar{c} + c_w \left(w_t + \frac{e - \bar{c} + p(1+\theta)\bar{b}}{(1+\theta)p + r - g} \right) = (1 - c_e)\bar{c} + c_e(e + p(1+\theta)\bar{b}) + c_w w_t \\ &= (1 - c_e)\bar{c} + c_e p(1+\theta)\bar{b} + c_w(w_t + h), \end{aligned} \quad (4)$$

for human wealth $h = \frac{e}{(1+\theta)p+r-g} = K_2 e$ (and $K_2 = \frac{1}{(1+\theta)p+r-g}$) and marginal propensities from the flow measure income $c_e = K_2 c_w$ and the stock measure of wealth:

$$c_w = \frac{\rho + p(\gamma - (1-\gamma)\theta) - (1-\gamma)\left(r + 0.5\frac{(R-r)^2}{\gamma\sigma^2}\right)}{\gamma\left(1 + p(1+\theta)^{\frac{\gamma-1}{\gamma}}(1-\tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}\right)}. \quad (5)$$

The intended bequests are:

$$b_t = c_w(1+\theta)^{-\frac{1}{\gamma}}(1-\tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}X_t - \bar{b}, \quad (6)$$

¹²The problem must also obey the transversality condition $\lim_{t \rightarrow \infty} \mathbb{E}_t \exp(-\tilde{\rho}t)V_t = 0$.

with $X_t \equiv w_t + \frac{e - \bar{c} + p(1+\theta)\bar{b}}{p(1+\theta) + r - g}$. Finally, the optimal portfolio share is given by:

$$\mu_t = \frac{R - r}{\gamma\sigma^2} \frac{w_t + \frac{e - \bar{c} + p(1+\theta)\bar{b}}{(1+\theta)p + r - g}}{w_t} = \frac{R - r}{\gamma\sigma^2} \frac{X_t}{w_t} = \tilde{\mu} \frac{X_t}{w_t}, \quad (7)$$

for $\tilde{\mu} \equiv \frac{R - r}{\gamma\sigma^2}$ (as in Merton (1969)).

These results provide some important economic insights. In the following, we discuss the optimal rules separately, starting with the optimal consumption rule.

In general, consumption grows at the same pace as labor income ($C_t = c_t \exp(gt)$). Consumption depends both on the stock of wealth w_t and the flow of earnings e as well as the preference parameters \bar{c} and \bar{b} . The marginal propensity to consume (MPC) from the flow of earnings c_e is higher than the MPC out of wealth ($c_e = \frac{c_w}{(1+\theta)p + r - g} > c_w$).¹³ The MPCs are not affected by the overall growth rate. The MPC out of physical w_t and human capital h are identical. In this model the earnings are risk-free and can thus be discounted by the risk-free interest rate in order to measure human capital.

The model captures several savings motives as encompassed in the MPC out of wealth c_w (cf. equation 5). Firstly (i), individuals save due to the standard Euler equation logic balancing the (risk-free) interest rate r and their time preference ρ relative to their intertemporal elasticity of substitution (IES, $\frac{1}{\gamma}$).¹⁴ Moreover, individuals save due to a longevity risk (ii), as captured by the death probability p . A longer expected life time (lower p) decreases their MPC from wealth c_w and thus increases their savings propensity. Finally, there are savings due to a bequest motive (iii).

In order to discuss the latter it is convenient to rewrite the MPC from wealth as:

$$c_w = \frac{1}{1 + \psi} \tilde{c}_w,$$

with $\tilde{c}_w = \frac{\rho + p(\gamma - (1-\gamma)\theta) - (1-\gamma)r}{\gamma} - 0.5 \frac{1-\gamma}{\gamma} \frac{(R-r)^2}{\gamma\sigma^2}$ and the factor $\psi = p(1+\theta)^{\frac{\gamma-1}{\gamma}} (1-\tau_b)^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}}$.

The sum of consumption and savings due to a bequest motive is given by:¹⁵

$$c_t + p(1+\theta)b_t = \tilde{c}_w X_t \left(\frac{1}{1+\psi} + \frac{\psi}{1+\psi} \right) = \tilde{c}_w X_t,$$

¹³Formally, this requires the reasonable assumption $(1+\theta)p + r - g < 1$.

¹⁴Note that in this model we cannot disentangle risk aversion γ and the IES as e.g. in Wang et al. (2016).

¹⁵Note that, here, we disregard the non-homothetic factors by setting $\bar{c} = \bar{b} = 0$. More generally, we have, $(c_t - \bar{c}) + p(1+\theta)(b_t - \bar{b}) = \tilde{c}_w X_t$.

and thus independent of ψ capturing the bequest motive. Essentially, a larger value of ψ implies a larger share of consumption devoted to future generations (in the form of bequests) as opposed to the currently living individuals.

Similar to the MPC we can define a marginal propensity to bequeath b_w :

$$b_w = c_w(1 + \theta)^{-\frac{1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}} = \tilde{c}_w \frac{\psi}{(1 + \psi)(1 + \theta)p} = c_w \frac{\psi}{(1 + \theta)p}. \quad (8)$$

For the special case of $b_w = 1$, we have $\iota = 0$ implying that individuals do not buy annuities. The case of $b_w > 1$ corresponds to buying of life time insurance $\iota > 0$ which is paid out to the descendants at the time of death of the individual, whereas $b_w < 1$ implies $\iota < 0$ and a positive cash flow for the insurance company at the time of death. It shares the comparative statics with the regular MPC and furthermore increases with death probability p and bequest strength χ capturing both the likelihood of bequeathing and the preference for it.

For the case without a bequest motive $\chi = 0$ we have the standard case of $c_w \equiv \tilde{c}_w$. This also holds true if the government imposes a complete tax on bequests ($\tau_b = 1$). As individuals perceive that net of taxes bequests will be zero, they also do not form savings for bequests.

For the special case of log-utility ($\gamma = 1$) – for which income and substitution effects cancel out each other – the tax rate τ_b has no impact on the consumption decision as captured by the marginal propensity to consume c_w (Lansing, 1999). Yet, for any other case the tax rate τ_b has an impact on the intended bequests. In line with the discussion in Straub and Werning (2014) the case of a prevailing income effect ($\gamma > 1$) implies that higher taxes are accompanied by higher savings ($\frac{\partial c_w}{\partial \tau_b} < 0$). For the opposing case of a prevailing substitution effect ($0 < \gamma < 1$), higher bequest taxes increase current consumption and thereby lower bequests to future generations. The same comparative statics as for the tax rate τ_b hold true for the insurance premium θ , both being deviations from a friction-less market.

Of course, bequests increase with the bequest motive χ , but are lowered with \bar{b} . The latter captures the luxury-good nature of the bequests and implies that the share of bequests relative to wealth w increases with the level of wealth. Individuals with a low wealth level eventually leave negative bequests to their descendants, implying that the generational transfers actually go from children to parents and not vice versa.¹⁶

¹⁶In the complete model (cf. Section 5) we disallow negative bequests.

Note that this discussion misses precautionary savings due to earnings risk identified as the key mechanism driving wealth inequality in *Bewley*-type models (Aiyagari (1994), Huggett (1993)). In Section 3.1 we introduce a realistic income process featuring permanent and transitory income shocks as well as retirement all leading to precautionary savings. Furthermore, we provide a discussion of the determinants of portfolio structuring in Section 3.2 with and without constraints.

2.2. Closed-form solutions for inequality and mobility

One of the key advantages when employing the continuous time framework is the presence of the so-called Fokker-Planck equations (henceforth FP and also known as Kolmogorov forward equations). These equations allow for a derivation of the cross-sectional distribution given the nature of the stochastic process. For the simple linear problem considered so far – without binding portfolio constraints or risky labor income – we can even find closed-form solutions that provide economic intuition. Even for the more elaborate general equilibrium model, they do not impose a numerical challenge. Once the individual optimization problem in the form of the HJB equation is solved, finding the solution to the FP equation only requires one single computational step. For details on the numerical algorithm the reader is referred to Online Appendix D.

We begin by abstracting from insurance payments at time of death ($b_w = 1$ implying $\iota = 0$) and bequest taxation ($\tau_b = 0$). We can insert the optimal rules for consumption c , bequests b , and portfolio composition μ into the overall flow equation.

The evolution of wealth w_t can be written as function of the state variable X_t :¹⁷

$$\begin{aligned} dw_t &= [((1 + \theta)p + r - g)w_t + \mu_t(R - r)w_t + e - c_t - p(1 + \theta)b_t]dt + \mu_t\sigma w_t dZ_t \\ &= [((1 + \theta)p + r - g - c_w - c_w\psi + (R - r)\tilde{\mu})X_t]dt + \tilde{\mu}\sigma X_t dZ_t \\ &= [((1 + \theta)p + r - g - \tilde{c}_w + (R - r)\tilde{\mu})X_t]dt + \tilde{\mu}\sigma X_t dZ_t. \end{aligned}$$

Note that the result is independent of the bequest motive (captured in the variable ψ). More generally, we can write this as a modified version of a Geometric Brownian Motion:

$$dw_t = g_w(w_t - w_{NB})dt + \sigma_w(w_t - w_{NB})dZ_t,$$

¹⁷Note that for readability we ignore the effect of $\bar{b} > 0$. The latter constitutes a constant term in the drift of $(c_e - 1)p\bar{b} > 0$ ensuring $w > w_{NB}$.

with $w_{NB} = -\frac{e-\tilde{c}}{(1+\theta)p+r-g}$ being the natural borrowing constraint¹⁸ and a growth rate:

$$g_w = (1+\theta)p+r-g-\tilde{c}_w + (R-r)\tilde{\mu} = (1+\theta)p+\tilde{R}-g-\tilde{c}_w = (1+\theta)p+r-g-\tilde{c}_w + \frac{S^2}{\gamma},$$

and a diffusion parameter:

$$\sigma_w = \tilde{\mu}\sigma = \frac{S}{\gamma},$$

with $S = \frac{R-r}{\sigma}$ representing the Sharpe ratio of the risky asset, implying $\tilde{R} = r + \frac{S^2}{\gamma}$.

Using the definition of the marginal propensity to consume out of wealth (equation 5) we can write the drift parameter as follows:

$$g_w = \frac{\tilde{R}-g-\tilde{\rho}-(1-\gamma)\frac{S^2}{2\gamma}+(1+\theta)p}{\gamma} = \frac{\theta p+r+0.5(1+\gamma)\frac{S^2}{\gamma}-\rho}{\gamma} - g.$$

We find the cross-sectional distribution $f(w, t)$ – focusing on the stationary¹⁹ distribution $f(w)$ – from the solution of the Fokker-Planck equation:

$$\frac{\partial f(w, t)}{\partial t} = 0 = -\frac{\partial}{\partial w}[s(w)f(w, t)] + \frac{\partial^2}{\partial w^2}[0.5 \cdot D(w)^2 f(w, t)], \quad (9)$$

for the (linear) savings function $s(w) = g_w(w - w_{NB})$ and the diffusion term $D(w) = \sigma_w(w - w_{NB})$.

The following proposition summarizes the result.

Proposition 2.2 (The cross-sectional distribution without traded annuities and taxes) *The stationary cross-sectional distribution of wealth without taxation, income risk and portfolio constraints is given by a Pareto II distribution of the type:*

$$f(w) \sim (w - w_{NB})^{-(a+1)} \quad w > w_{NB}, \quad (10)$$

with a Pareto coefficient:

$$a = 1 - \frac{2g_w}{\sigma_w^2} = 1 - \frac{2g_w\gamma^2}{S^2}. \quad (11)$$

¹⁸In the natural borrowing constraint $w = w_{NB}$ individuals consume all flow income from both earnings e and capital income, constituting a stationary level of wealth and $X = 0$.

¹⁹Note that in this paper we focus on the stationary distribution (setting the left term in the Fokker-Planck equation to zero). In contrast, Gabaix et al. (2016) focus on the dynamics of inequality.

We can compute the Gini-coefficient with:

$$Gini(w) = \frac{1}{a + w_{NB}(a - 1)} \frac{a}{2a - 1}, \quad (12)$$

and the measure of the top x [%] with:

$$s_x(w) = \frac{1}{a + w_{NB}(a - 1)} \left[ax^{1-\frac{1}{a}} + w_{NB}(a - 1)x \right]. \quad (13)$$

Proof The proof is suspended to Appendix A.1.

Both measures of inequality – the broad Gini-coefficient and the tail-focused top shares – decrease with both the Pareto coefficient a and the minimum level w_{NB} .²⁰

First of all, it is important to point out that a stable distribution of the Pareto-type requires $a > 1$ and hence $g_w < 0$. Thus, wealth must be mean reverting. Therefore, the slope of the savings function must be negative:²¹

$$\frac{\partial s(w)}{\partial w} = g_w < 0.$$

The condition ($g_w < 0$) holds when:

$$\rho > \theta p + r + \frac{1 + \gamma}{2\gamma} \frac{S^2}{\gamma} - \gamma g, \quad (14)$$

which nests the well-known condition $\rho > r$ required in standard models for a stationary equilibrium (Aiyagari, 1994). While this constraint is relaxed for growing income ($g > 0$), it becomes more tight for imperfect insurance markets ($\theta > 0$).

²⁰Moreover, for $w_{NB} \rightarrow 0$ the standard Pareto distribution is nested. For the other extreme case of $w_{NB} \rightarrow \infty$ we have a Lorenz-curve equal to the 45-degree line and hence total equality.

²¹In this simple economy without binding wealth constraints or income risk, all individuals will decrease their wealth to $w \rightarrow w_{NB}$. The more elaborate version of the model outlined in Section 3 features both income risk and borrowing constraints tighter than the natural borrowing constraint ($\underline{w} > w_{NB}$). The former implies that individuals in a high income state j' will form equilibrium savings larger than the borrowing constraint ($w_{j'}^* > \underline{w}$). Secondly, the tight (arbitrary) borrowing constraint technically acts as a boundary condition in the overall Fokker-Planck equation, which transforms the overall log-normal distribution into a distribution of the Pareto-type (Sornette and Cont, 1997).

It is also easy to verify that the condition $g_w < 0$ also implies $c_e = \frac{\tilde{c}_w}{(1+\theta)^{p+r-g}} > 1$. As individuals consume more than one-on-one of their flow income, they will eventually run down their level of wealth.²²

We can exploit this property to conduct comparative statics. Using the micro-founded terms using the optimal savings rule, the Pareto coefficient is given by:²³

$$a = \gamma \left(\frac{2(\tilde{\rho} - (r - g) - (1 + \theta)p)}{S^2} - 1 \right) = \gamma \left(\frac{2(\rho - r - \gamma g - \theta p)}{S^2} - 1 \right). \quad (15)$$

First of all, the inequality increases (a decreases) for a higher Sharpe ratio. Thus, the higher the risk premium ($R - r$) individual wealth grows at a higher pace. Interestingly and somewhat surprisingly, higher volatility σ of the risky asset eventually decreases inequality. This results from the fact that individuals internalize this risk and react by reducing their portfolio exposure to the risky asset. We also find support for the argument entertained in Piketty (2014) showing that the gap between interest rate and labor growth $r - g$ is a factor contributing to higher wealth inequality. Besides economic factors, pure demographics also shape the distribution of wealth. In an imperfect insurance market ($\theta > 0$) a higher life expectancy (lower values for p) promotes savings for longevity risk, increasing wealth growth g_w and hence wealth inequality.

By analyzing the Pareto coefficient we focused on the right tail of the distribution. The other parameter w_{NB} eventually impacts mostly on the left tail of the distribution. Higher earnings e increase the value of human capital available to borrow against and thereby contribute to more inequality at the bottom. In contrast, a larger minimum consumption desire \bar{c} acts in the opposing direction, eventually decreasing bottom inequality. For the (realistic) case of $w_{NB} < 0$ we find an opposing effect to the one proposed by Piketty (2014) prevalent at the top end. At the bottom tail the rationale is yet highly different. A larger effective discount $r - g$ reduces the value of human capital available for borrowing and thus reduces bottom inequality.²⁴

²²Equipped with the drift expression g_w we can also investigate the transversality condition (cf. footnote 12). The latter requires $(1 - \gamma)(g_w - 0.5\sigma_w^2) - \tilde{\rho} < 0$. After some algebra one can show that this coincides with $\tilde{c}_w > 0$ and thus holds for any positive consumption propensity.

²³A similar spirited analysis is presented in the Online Appendix E of Achdou et al. (2017) in an infinite horizon model without labor growth.

²⁴When incorporating $\bar{b} > 0$ into the natural borrowing constraint this factor increase bottom inequality as individuals lever up.

Next, we take a stand on mobility. A simple measure of mobility – e.g. discussed in Fischer (2018) – is the mean reversion rate g_w .²⁵ The properties of this measure are summarized in the following proposition.

Proposition 2.3 (Mobility of wealth) *We can measure mobility of wealth for a time lag τ by:*

$$0 \leq \text{Mob}_\tau(w) = 1 - \exp(g_w\tau) \leq 1. \quad (16)$$

Mobility decreases with g_w . As this measure also increases inequality, high inequality and low mobility of wealth emerge jointly.

This measure of mobility is broadly in line with common measures of mobility such as the simple but very popular inter-generational correlation²⁶ (Jäntti and Danziger, 2000) and the Shorrocks index²⁷ based on transition matrices (Shorrocks, 1978) also employed in the quantitative investigation of the model in the latter sections.

To have mobility in the first place we require $g_w < 0$. Of course, in the long run ($\tau \rightarrow \infty$) mobility is perfect ($\lim_{\tau \rightarrow \infty} \text{Mob}_\tau(w) = 1$) as all states of the wealth distribution can be reached. In general, measures that decrease g_w increase mobility. Note that a low value of g_w was already identified as a factor contributing to higher equality. Hence, in the model the suggested relationship of the *Great Gatsby curve* (Corak, 2013) – i.e. high inequality comes with low mobility and vice versa – is confirmed. Thus, larger Sharpe ratios S and gaps between capital and labor growth $r - g$ also lead to lower mobility locking in the success of wealthy individuals.

Our measure of mean reversion does not consider the dispersion parameter σ_w which matters more at the upper tail of the distribution. In Appendix A.2 we show that mobility in general decreases with $\sigma_w = \frac{S}{\gamma}$. As this measure also increases inequality, this result is also in line with the *Great Gatsby curve*.

²⁵Fernholz (2016) shows that mobility increases with the mean reversion rate $|g_w|$ and also posits that high mobility is accompanied by low inequality and vice versa.

²⁶Formally, the inter-generational correlation is $\rho_w = \exp(g_w/p) < 1$.

²⁷It can be shown that for an autoregressive process of first order the Shorrocks index (for a one period gap) amounts to $1 - \exp(g_w\tau)$ exactly as used in our mobility measure.

Thus far, we have abstracted from both annuities and taxes on bequests as a policy parameter. In order to account for them we have to consider the factor $\tilde{b}_w = (1 - \tau_b)b_w$ and include it into the FP equation:

$$\frac{\partial f(w, t)}{\partial t} = 0 = -\frac{\partial}{\partial w}[s(w)f(w, t)] + \frac{\partial^2}{\partial w^2}[0.5 \cdot D(w)^2 f(w, t)] - pf(w, t) + \frac{p}{\tilde{b}_w} f\left(\frac{w}{\tilde{b}_w}, t\right). \quad (17)$$

The penultimate term captures that dying individuals (death probability p) are removed from the distribution, while the last term captures their descendants reemerging at a different level of wealth after taxes and annuities pay out.

The formal derivation of the general case is suspended to Appendix A.2. In this instance, our (approximate) statement concerns the top tails only. As these individuals are far away from the natural borrowing constraint we write a simplified savings function $s(w) = g_w w$ and a diffusion term $D(w) = \sigma_w w$.

Proposition 2.4 (The distribution of top wealth and mobility with bequest taxes and traded annuities) *Extending Proposition 2.2 with a linear tax on bequests τ_b , annuities a , and exponentially distributed deaths with a probability p , the right tail of the wealth distribution is approximately described by a Pareto tail $\lim_{w \rightarrow \infty} f(w) \sim w^{-(\tilde{a}+1)}$ with the property:*

$$\tilde{a} \approx a + \frac{2p\gamma^2}{S^2}(1 - \tilde{b}_w). \quad (18)$$

Our measure of mobility is now given by:

$$Mob_\tau(w) = 1 - \exp([g_w - p(1 - \tilde{b}_w)]\tau). \quad (19)$$

Proof The proof is presented in Appendix A.2.

Compared to the case without taxes and annuities (implying a Pareto coefficient of a), the measure of top tails \tilde{a} can increase or decrease relative to the benchmark a . For $\tilde{b}_w > 1$ it is lower implying more top inequality as the descendants furthermore receive life insurance payments when their ancestors die reducing top inequality.

It is also insightful to consider the special case without annuities for which $\tilde{b}_w = 1 - \tau_b$ (as $b_w = 1$). It is easy to see that without taxation ($\tau_b = 0$) deaths do not matter implying that the same results as in Proposition 2.2 hold. For the polar extreme case

of total redistribution $\tau_b = 1$, we will have a double Pareto distribution with both a fat tail in the left end and the right end of the distribution.²⁸

For a general tax level τ_b we have:

$$\tilde{a} \approx a + \frac{2p\tau_b\gamma^2}{S^2}$$

The redistributive effect of the effect increases with its effective tax rate $\tau_b p$, this being the product of the tax rate and the likelihood the tax is imposed (the death probability p). Given that the tax is imposed infrequently the effective tax rate is low and thus the redistributive character of this tax is limited. With the increase in longevity (lower values for p) the bequest tax also loses its redistributive impact as it is imposed at lower frequency. We will explore this quantitatively in the calibrated version of the model. Once again, higher Sharpe ratios (by increasing $\sigma_w = \frac{S}{\gamma}$) decrease \tilde{a} and thus increase inequality.

Secondly, the proposition makes a statement about mobility. With higher levels of \tilde{b}_w capturing the willingness to bequeath, mobility decreases. For the case of $b_w > 1$ – implying that descendants profit from their parent’s death by receiving life insurance payments – there is higher mobility as compared to the case without annuities (associated with $b_w = 1$). For the latter case the measure of mobility simplifies to:

$$Mob_\tau(w) = 1 - \exp([g_w - p\tau_b]\tau).$$

This shows that in the presence of taxation there is a higher degree of mobility. The rationale is that the redistributive taxation introduces a higher degree of mean reversion by resetting individuals in the direction of the cross-sectional mean. Hence, the policy measure of taxes reduces both inequality and increases mobility. In Section 5.2 this relationship is quantitatively explored in an extended and calibrated version of the model.

Note that these comparative statics take the perspective of a partial equilibrium economy. In a general equilibrium economy, the prices of goods – especially the returns on

²⁸Formally, this is a combination of Geometric Brownian Motion evaluated at an exponentially distributed time (due to the Poisson death) as the last term in the FP culminates to a Dirac impulse. For a more detailed discussion on the formal mechanism the interested reader is referred to, for example, Fischer (2018).

assets – depend on the individual decision and hence some of the comparative statics might not bear up. We explore this in more detail in the following Sections.

3. The extended model

The partial equilibrium model developed so far contains some interesting implications and comes with the great advantage of possessing closed-form solutions. Yet, it also lacks two major features, namely (i) income inequality and (ii) borrowing constraints. Section 3.1 introduces a realistic discrete-state income process. The constraints on borrowing and more generally on composing the portfolio are considered in Section 3.2. The model is closed in general equilibrium in Section 3.3.

3.1. A joint income process

Thus far, we have assumed that labor income is not subject to variation within the lifetime. This is diametrical to the *Bewley*-type literature which identifies uninsured labor income risk as the driving mechanism to explain income and subsequent consumption and wealth inequality.

We follow a well-established literature and model the labor endowment process z by means of a Markov transition matrix. Individual labor income e_i results from the labor endowment z and the prevailing (overall) wage rate ω ($e_i = \omega z_i$). Labor income can vary due to transitory, permanent shocks, and retirement status. In order to keep it parsimonious we employ a binary differentiation for all cases including permanent (high h and low l , with $\mathcal{P}_h > \mathcal{P}_l$) and transitory (good and bad, $\mathcal{S}_g > \mathcal{S}_b$) shocks, implying a total of $2^3 = 8$ states. We briefly explain the construction of this matrix. Details are relegated to Online Appendix C.

For the relationship between working (index w) and retired (index r) individuals, we assume that a Pay-As-You-Go (PAYG) social insurance system is at place. The system is financed by a linear tax τ_r levied on working age individuals to finance transfers

$0 < \epsilon_r < 1$ redistributed to the current retired. With the average labor efficiency \tilde{z} the self-financing condition implies:²⁹

$$\begin{aligned}\tilde{z}\tau_r\pi_w &= \tilde{z}\epsilon_r\pi_r \rightarrow \tau_r p_r = p_w \epsilon_r \\ &\rightarrow \tau_r = \frac{p_w}{p_r} \epsilon_r.\end{aligned}\tag{20}$$

Here the value $\frac{1}{p_w}$ captures the average time in employment, whereas $\frac{1}{p_r}$ measures the average time in retirement. The overall expected life length amounts to $\tilde{T} = \frac{1}{p} = \frac{1}{p_r} + \frac{1}{p_w}$. Thus, the tax rate increases with the age-dependency ratio $\frac{p_w}{p_r}$ (the ratio of retired to working individuals).

From this we can derive the final Markov transition matrix \mathbb{M} (8 states):³⁰

$$\mathbb{M} = \begin{bmatrix} (1 - p_w)\mathbb{M}_{\mathcal{P};\mathcal{S}} & p_w\mathbb{I} \\ p_r\mathbb{M}_{gen} & (1 - p_r)\mathbb{I} \end{bmatrix},\tag{21}$$

incorporating the labor-state transition Matrix $\mathbb{M}_{\mathcal{P};\mathcal{S}}$ and another square matrix with 4 states \mathbb{M}_{gen} which captures intergenerational correlation in (permanent) income which we discretize using Rouwenhorst's (1995) method with an intergenerational correlation of ρ_z .³¹

In the calibrated version, the model performs very well in capturing the distribution of the bottom 99% of the population. On the other hand, it severely underestimates the tail of the distribution.

Thus, we explicitly target the top 1% by introducing a new working state which – following Castaneda et al. (2003) – we label the *superstar* state.³² The rationale is

²⁹The linear transfer implies that each retired individual earns a fixed fraction of her working income and thus disregards the fact that the social security system in the US entails a high degree of redistribution (Kaymak and Poschke, 2016). In fact, inequality among the working as well as retired individuals is identical for this modeling. Yet, it is in line with the evidence of Rios-Rull and Kuhn (2016) (refer in particular to table 27 and 31 of their paper) that income inequality among the retired is as high (or even slightly higher) among retired individuals.

³⁰Note that the identity matrices are of dimension 4.

³¹For details, the reader is referred to Online Appendix C.

³²A more recent calibration capturing the increase in overall inequality in the USA in a similar setting is presented in Kaymak and Poschke (2016).

that some individuals can be considered *superstars*, i.e. top-earners in fields such as entertainment or sports. We modify the transition matrices as follows:

$$\mathbb{M}^e = \begin{bmatrix} (1 - \lambda_s)\mathbb{M}_{\mathcal{P};\mathcal{S}}(1 - p_w) & \lambda_s\mathbf{1}(1 - p_w) & p_w\mathbb{I} & \mathbf{0} \\ (1 - p_w) \cdot 0.25 \cdot \mathbf{1}' & 0 & \mathbf{0}' & p_w \\ p_r\mathbb{M}_{gen} & \mathbf{0} & (1 - p_r)\mathbb{I} & \mathbf{0} \\ p_r \cdot 0.25 \cdot \mathbf{1}' & 0 & \mathbf{0}' & (1 - p_r) \end{bmatrix}, \quad (22)$$

for which $\mathbf{1}$ represent a vector with four unit entries and $\mathbf{0}$ is a vector with zero entries of dimension 4.³³ The assumption is that every working individual can become a superstar with probability λ_s . The superstar state yet only lasts one simulation period (cp. second line of the matrix). Furthermore, no individual is born into the superstar state (cf. last two lines of the matrix).

As we cast our model in continuous time, the transition matrix must also be in continuous time. In a first-order approximation the continuous time transition matrix \mathbb{R} is related to the discrete time transition matrix \mathbb{M}^e as follows:³⁴

$$\mathbb{R} \approx \mathbb{M}^e - \mathbb{I}.$$

3.2. Portfolio constraints

Given the stochastic nature of the income process, individuals want to insure these shocks by forming precautionary savings. They are, however, restricted in their ability to form portfolios.

In the presented framework individuals invest in two assets. There is a risk-free asset B (being a bond):

$$B = (1 - \mu)w,$$

³³We assume that the four groups of working individuals represent a share of identical magnitude. Hence we have the factor 0.25 to redistribute superstars into the regular distribution (last line of the equation).

³⁴More formally the relationship between the two follows from the matrix exponential as we have $\pi_t = \mathbb{M}^e \pi_{t-1}$ in discrete and $\dot{\pi} = \mathbb{R}\pi \rightarrow \pi_t = \exp(\mathbb{R}t)\pi_0 = \exp(\mathbb{R})\pi_{t-1}$ in continuous time. The matrix exponential is given by the power series $\exp(\mathbb{R}) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbb{R}^k$ which in a first order approximation is $\exp(\mathbb{R}) \approx \mathbb{I} + \mathbb{R}$. While for the discrete time matrix all lines sum to one, they sum up to zero for its continuous time counterpart.

and a risky asset k (representing stocks or other risky assets such as real estate or private business):

$$k = \mu w.$$

If an individual shorts the risk-free asset, she holds a debt position ($D = -B$) which is assumed to have the same interest rate as the risk-free asset.

Individuals are constrained in their portfolio composition by market imperfections. Even in the absence of market imperfections, the portfolio composition – in particular the ability to borrow – is constrained by a natural borrowing constraint w_{NB} . It is given by:

$$B = -\frac{\underline{e} - \bar{c} + p\bar{b}}{r - g} = -\underline{h} + K_2(\bar{c} - p\bar{b}) \equiv w_{NB}, \quad (23)$$

for which \underline{e} represents the lowest possible earnings. Note that in the presence of either minimum consumption ($\bar{c} \neq 0$) or a minimum bequest motive ($\bar{b} \neq 0$) the maximum debt is unequal to the human capital \underline{h} . Consider for a moment the case with a homothetic bequest motive ($\bar{b} = 0$). In fact, for high minimum consumption requirements ($\underline{e} < \bar{c}$) the natural borrowing limit is eventually positive. This implies that individuals should actually hold a minimum stock of wealth rather than borrow money in order to sustain a minimum consumption level.

We can rewrite the overall portfolio rule in the following manner:

$$\mu = \tilde{\mu} \left(1 - \frac{w_{NB}}{w} \right)$$

For extremely wealthy agents the optimal share is as in the standard model of Merton (1969) ($\lim_{w \rightarrow \infty} \mu = \tilde{\mu}$). We can also analyze the portfolio composition as a function of wealth. The slope is given by:

$$\frac{\partial \mu}{\partial w} = \tilde{\mu} \frac{w_{NB}}{w^2}.$$

For a positive natural borrowing constraint ($w_{NB} > 0$) the share of risky assets increases with wealth as generally suggested by decreasing relative risk aversion.

The more interesting case is the one with a negative natural borrowing constraint implying individuals with a negative net worth in the lower end of the wealth distribution as suggested by the empirical evidence (cf. e.g. Rios-Rull and Kuhn (2016)). In this case the share of risky assets μ eventually decreases with wealth. Nevertheless, the level of risky assets $k = \mu \cdot w$ increases with overall net worth w .

When the constraint is at the level of the natural constraint and in the absence of income risk, the consumption function is linear. The case of the borrowing constraint being *tighter* than the natural borrowing constraint ($\underline{w} > w_{NB}$), is frequently referred to as a liquidity constraint and associated with a concave consumption function (Holm, 2018).

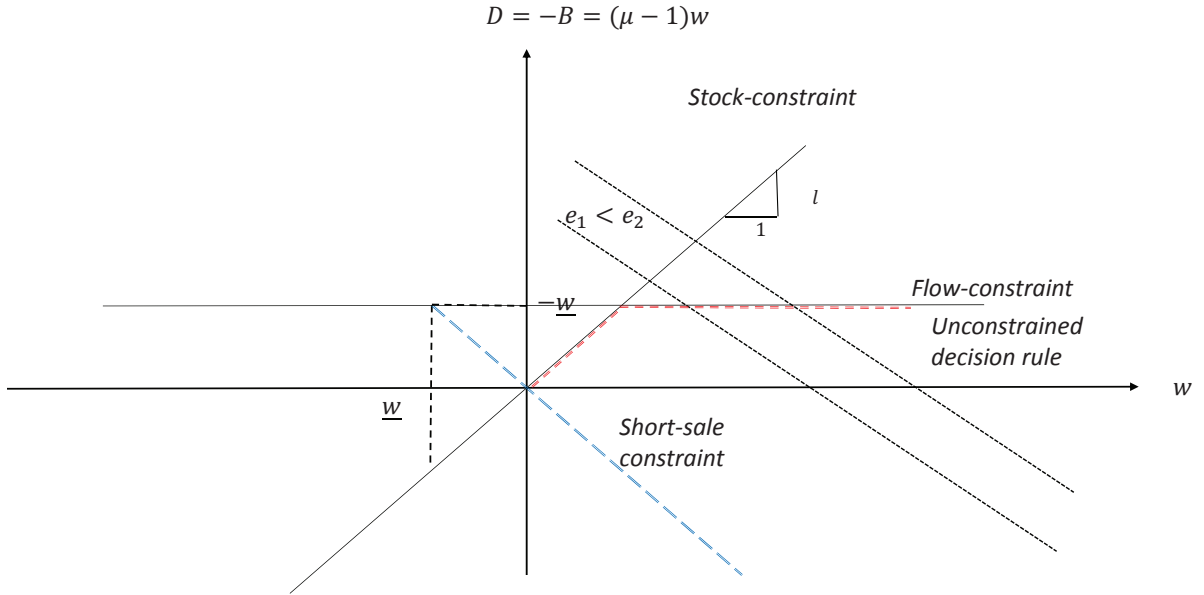


Figure 1: Relationship between debt D and net worth w

The model produces an inverse u-shaped relationship between net worth and debt. Low net worth individuals are primarily constrained by the stock constraint. With $w_{NB} < 0$ the debt level decreases for the unconstrained individuals.

Assume borrowing is generally allowed, yet its magnitude is restricted. Thus, a value of $\mu > 1$ is possible. This value implies that individuals lever their position in the risky asset by short-selling risk-free assets, i.e. they borrow. A straightforward example would be borrowing in the form of a mortgage in order to finance the purchase of real estate. In the tradition of Kiyotaki and Moore (1997) we assume that borrowing is subject to a collateral constraint:

$$-B \equiv D = (\mu - 1)w \leq (1 - \phi_w)k = (1 - \phi_w)\mu w.$$

The equation states that debt $-B$ should be lower than a fraction $(1 - \phi_w)$ of the total stock value of wealth μw invested in assets. As a result the overall portfolio constraint reads:

$$\mu \leq \frac{1}{\phi_w}.$$

There are different interpretations for the value $0 < \phi_w \leq 1$. It can be considered a haircut on the collateral. A different interpretation would be that of an equity ratio (the ratio of net worth w to total assets μw). For the case of $\phi_w = 1$ we identify a nested case without borrowing ($\mu \leq 1$). With $\phi_w < 1$, the measure w now represents net worth – i.e. assets net of debt – rather than pure assets. This constraint is related to the stock of wealth. Note that in our model – and in line with the evidence – net worth can be negative. This contrasts with a large share of the existing literature which considers idiosyncratic income risk in production economies (e.g. Krusell and Smith (1998)) that restrict individual wealth holdings such that they are positive.³⁵

Following Iacoviello and Pavan (2013) we introduce a second constraint related to the flow of income and more precisely to the natural borrowing constraint w_{NB} . In line with the discussion of Holm (2018) this constraint must be tighter than the natural borrowing constraint and constitutes the lowest level of wealth possible $\underline{w} = \phi_e w_{NB} < 0$ with $0 < \phi_e < 1$. In contrast to the other constraint, it is related to the flow measure of earnings (index e). Hence, our second constraint is:

$$B = (1 - \mu)w \geq \underline{w} = \phi_e w_{NB}, \quad (24)$$

leading to:

$$\mu \leq 1 - \frac{\underline{w}}{w} = 1 - \frac{\phi_e w_{NB}}{w}.$$

Thus, in total we restrict the portfolio composition by:

$$k = \mu w \leq \min \left\{ \frac{w}{\phi_w}, w - \phi_e w_{NB} \right\}. \quad (25)$$

The two constraints interplay in an interesting way. For households with positive, but low net worth w , the stock constraint is more binding (cf. Figure 1). They are constrained in their ability to lever up by the low amount of assets available as collateral.

³⁵Of course, short positions are allowed in economies in the tradition of Huggett (1993) as the individuals trade risk-free bonds that are in zero net supply and shorted by individuals in the high income state.

For higher values of net worth the flow constraint becomes more binding.³⁶ Thus, for medium values of net worth the flow constraint is binding, which also puts an upper absolute cap on overall debt to the level of \underline{w} (cf. Figure 1). Finally – and as already discussed – due to $w_{NB} < 0$, for high levels of net worth w the share of risky assets μ declines. Thus, the overall level of debt also declines for high levels of net worth. Hence, we have a realistic inverse u-shaped relationship between net worth and debt as displayed in Figure 1 for households with positive net worth. This relationship is also documented in the empirical evidence (Rios-Rull and Kuhn, 2016) and primarily results from mortgage debt being especially prevalent for middle class households. Note that the unconstrained decision rule shifts up for higher levels of income. Hence, households with a higher level of income ($e_2 > e_1$) hold both a higher level of risky assets and debt (cf. Figure 1) for a given level of net worth w .

Furthermore, we introduce a short-sale constraint for the risky asset ($k > 0$). As such, households with negative net worth do not own risky assets, but only hold debt – i.e. a short position in the risk-free asset (cf. Figure 1). The long positions in the risk-free asset – indicated by a negative value of debt in Figure 1 – are taken by the households with the highest net worth, providing the supply of the risk-free asset, which allows for an aggregate zero net supply of risk-free assets.

Finally, we also introduce a non-linearity for bequests. In the presence of a luxury motive for bequests ($\bar{b} > 0$), negative bequests can emerge. One could consider this as the case in which rather than transfers running from parents to their children parents receive net transfers from their descendants (e.g. paying for a place in a retirement home). However, in our model the warm glow motive only considers future generations and not past generations. Moreover, by law individuals are allowed to decline negative bequests. As such, for our extended model solved numerically, we require:

$$b = \max \{0; b\}, \quad (26)$$

implying positive bequests only.

³⁶In fact the two constraints coincide for $w = -\frac{l}{\underline{w}}$, for which $l \equiv \frac{1-\phi_w}{\phi_w}$ represents the leverage ratio defined as the ratio of the debt share relative to the equity share.

3.3. Equilibrium

Thus far, we have considered a partial equilibrium economy. Let us extend this to allow for both endogenous wages ω and risky interest rates R determined by production function with a geometric depreciation rate δ . The risk-free interest rate r follows from the condition of a zero net supply of the risk-free bonds. Hence, this model is both a production economy in the spirit of Aiyagari (1994) (for the risky asset) and an endowment economy in the tradition of Huggett (1993) (for the risk-free asset).

In the tradition of Angeletos (2007) we assume that individuals have access to some idiosyncratic *backyard* production technology determining idiosyncratic returns. We assume that the production technology is of the Cobb-Douglas type $y_{i,t} = A_{i,t}f(k_{i,t}, l_{i,t})$ with a capital share α and an idiosyncratic time-varying Total Factor Productivity (TFP) $A_{i,t} > 0$. Individuals decide how much efficiency labor $l_{i,t}$ to hire in their self-owned business at the economy-wide efficiency wage rate ω in order to maximize:

$$R_{i,t}k_{i,t} = \max_{l_{i,t}} \{A_{i,t}f(k_{i,t}, l_{i,t}) - \omega l_{i,t} - \delta k_{i,t}\}.$$

For the chosen constant returns-to-scale production function the optimal individual and economy-wide capital share coincide ($\alpha = \frac{(R_{i,t}+\delta)k_{i,t}}{y_{i,t}} = \frac{(R_t+\delta)k_t}{y_t}$) implying that the individual return rate is linear in the TFP ($R_{i,t} \sim A_{i,t}$). As before the risky returns are assumed to follow from an Arithmetic Brownian Motion with a cross-sectional standard deviation σ and an expected rate of return R determined in general equilibrium.

Stationary equilibrium 1. Households are of unit measure. Given wages ω , expected risky interest rates R , risk-free interest rates r ,³⁷ labor endowments z_j , and the distribution of wealth $g_j(w)$ private households maximize the problem presented in equation 3.³⁸ Thus, for a given labor endowment z_j they find optimal values for consumption c_j , intended bequests b_j , and portfolio shares μ_j , resulting in an overall savings rule s_j .

³⁷As we consider a stationary equilibrium we omit time indexes.

³⁸The problem does not feature idiosyncratic income risk. Details on how to incorporate this are detailed in step 3 of Online Appendix D describing the solution algorithm.

2. Given risky capital k , firms maximize their profits implying risky interest rates:³⁹

$$R = \alpha \left(\frac{k}{\tilde{z}} \right)^{\alpha-1} - \delta, \quad (27)$$

and wages:

$$\omega = (1 - \alpha) \left(\frac{k}{\tilde{z}} \right)^{\alpha} = (1 - \alpha)y, \quad (28)$$

with y being overall output.

3. Given R , ω , r and the savings rules s_j , the stationary distribution of wealth $g_j(w)$ follows from the solution of the Fokker-Planck equation (Eq. 17).⁴⁰

4. The government budget is balanced:

$$Tr = \bar{T} = p \sum_j \int \tau_b(w) w g_j(w) dw. \quad (29)$$

The proceedings of the bequest tax \bar{T} are redistributed as transfers $Tr = \bar{T}$.

5. For a labor endowment z_j and some wealth w risky capital is given by $k_j(w) = \mu_j(w)w$. Given $g_j(w)$ and k , the market for risky capital clears:

$$k = \sum_j \int k_j(w) g_j(w) dw \quad (30)$$

Risk-free bonds are in zero net supply. Hence, the risk-free interest rate r is determined by:

$$\sum_j \int (w - k_j(w)) g_j(w) dw = 0. \quad (31)$$

In the aggregate we have:

$$\bar{Y} = \bar{K}(\delta + g) + \bar{C}, \quad (32)$$

for which capital letters with a bar denote aggregate measure adjusted for (exogenous) income growth g .

³⁹Note that we have $\sum_i l_i = \tilde{z}$ equalizing individual and aggregate demand.

⁴⁰Details are presented under step 4 in the Online Appendix D describing the solution algorithm.

To capture a progressive bequest tax system we model the bequest tax as follows:

$$\tau_b(w)w = \begin{cases} 0 & w < w_E \\ \tau_b(w - w_E) & w \geq w_E \end{cases}$$

There is an exemption level w_E . If the wealth transferred at time of death is below this specific level, no bequest tax is imposed. At the point w_E marginal taxes jump to τ_b . Values in excess of the exemption level are subject to taxation. Hence, this is a progressive tax system with increasing average tax rates. Even without an exemption level ($w_E = 0$), the bequest tax system is progressive due to the lump sum transfer to all individuals of Tr .

In the following, the full scale model is calibrated to evidence for the USA, solved, and discussed.

4. Calibrating the model

Thus far, we have presented an analytic solution to a simplified version of the model. The complete model entails several non-linearities that require a numerical solution of the model. Given that we employ the continuous time approach, the numeric solution is substantially faster than using standard discrete time methods. In particular this exploits the sparsity of the transition matrices required in both value function iteration and the solution of the Fokker-Planck equation, which determines the cross-sectional distribution (Achdou et al., 2017). Details about the computational procedure are relegated to Online Appendix D.

We split the calibration procedure into two parts. First of all, we aim at finding an income process – in the form of a discrete state Markov chain, as detailed in Section 3.1 – to fit the overall exogenous income distribution. In a second step, we determine the other variables of the model.⁴¹

4.1. The income process

The values with respect to demographics are set according to the established literature (Castaneda et al. (2003), Kaymak and Poschke (2016)). We assume a working-life length

⁴¹Parra-Alvarez et al. (2015) show that one can exploit the computational efficiency of the continuous time approach in order to estimate model parameters in a small scale version of a similar model.

$\frac{1}{p_w} = 45$ years and a retirement length of $\frac{1}{p_r} = 15$ years, making a total (economic)⁴² life length of $\frac{1}{p} = 60$ years. The parameters p_w and p_r are also in line with a share of retired and disabled individuals in the overall population of 25% (Rios-Rull and Kuhn, 2016). In line with Krueger et al. (2016) we set a replacement ratio for retired individuals of $\epsilon = 40\%$ which with equation 20 implies $\tau_r = 13\%$.

Secondly, we decompose income of the working individuals into a permanent and transitory component:

$$z = \mathcal{S}_i \mathcal{P}_j \leftrightarrow \log(z) = \log(\mathcal{S}_i) + \log(\mathcal{P}_j). \quad (33)$$

For these measures we set $\mathcal{S}_b = \mathcal{P}_l \equiv 1$, implying that \mathcal{P}_h is the ratio between high income earners versus low income earners. Moreover, we thereby normalize the lowest income of the level of the working population ($z_1 = \mathcal{S}_b \mathcal{P}_l = 1$). The permanent deviation between income mostly results from education gaps. Broadly in line with Rios-Rull and Kuhn (2016) we set the education premium and thus the value of $\mathcal{P}_h = 4.2$. The transitory income variance is slightly lower and set to $\mathcal{S}_g = 3.4$. We impose $\lambda_g = \lambda_b = 0.2$ for the annual transitory probabilities, implying an equal share in each state and an average staying time of 5 years. We set the income of the superstar $z_s = z_5$ in order to match the share of the top 1% in the population.⁴³ This assumption helps us to provide a good (discrete) fit of the overall income distribution including the top tail.⁴⁴ The overall income growth is assumed to be $g = 1\%$ (Storesletten et al., 2004). The final transition matrix and the stationary states are documented in Appendix C.

4.2. Other variables

Table 1 summarizes the general parameters employed for the model. Most of the parameters follow the well-established literature. Some comments are nevertheless warranted.

⁴²Note that *economic* life starts at an age of roughly 25, when education is completed, making the expected *biological* life length 85 years.

⁴³To fit the top $x\%$, the income of the superstar is given by $z_s = \frac{s_x}{1-s_x} \frac{1-x}{x} \bar{z}$, with \bar{z} being the average income in the population when disregarding the superstars and s_x is the overall income share of the top $x\%$. The transition probability is given by $\lambda_s = \frac{s_x}{1-s_x}$.

⁴⁴We show in Appendix E, that while the superstar state helps to match top income inequality and overall wealth inequality, it leads to the fact that the model understates the Pareto tails of the wealth distribution from idiosyncratic wealth risk (due to a general equilibrium effect) and understates wealth mobility.

Category	Definition	Symbol	Calibration Values
Preference parameters	Risk aversion	γ	0.9
	Time preference	ρ	5%
	Strength bequest motive	χ	5
	Minimum bequest value	\bar{b}	120
Financial market variables	Standard deviation risky asset	σ	20%
	Annuity premium	θ	0
Production function	Capital share	α	0.38
	Depreciation	δ	0
Portfolio constraint	Equity ratio	ϕ_w	0.75
	Flow constraint	ϕ_e	0.25

Table 1: Parameters baseline simulation

Parameter choices follow the literature closely and are explained in the text. In Online Appendix E the robustness of the results with respect to parameter variations is investigated.

The annual time discount rate is assumed to be $\rho = 5\%$, which has its discrete time equivalent (in standard notation) of $\beta = \exp(-\rho) = 0.95$ and is a common value in the literature (cf. e.g. Kaymak and Poschke (2016)) capable of explaining a realistic wealth-to-income ratio.

A first major departure from the literature is the assumption of $\gamma = 0.9 < 1$, implying an intertemporal rate of substitution (IES) that exceeds unity ($IES = \frac{1}{\gamma}$). Most of the literature employs a value for risk aversion that is close to log-utility, but exceeds the unit measure (cf. e.g. Castaneda et al. (2003) or Kaymak and Poschke (2016)). As discussed in Açıkgöz (2018) or Achdou et al. (2017), such an assumption is highly problematic as it entails the possibility of multiple equilibria. The (graphic) intuition is that not only the capital demand (following from the production function), but also the capital supply determined by individual savings (in some range) decreases with the interest rate R for the prevailing income effect ($\gamma > 1$). Thus, there can be multiple intersections of the two curves constituting multiple equilibria, including potentially unstable ones.

We employ a non-standard preference function. For our baseline calibration, we abstract from a minimum consumption desire ($\bar{c} = 0$). In similar models, Benhabib et al. (2011) and Benhabib et al. (2014) set the bequest strength $\chi = 2.48$, respectively,

$\chi = 14.4$ for the bequest motive.⁴⁵ We employ a middle of the road calibration of $\chi = 5$. Following from the discussion in Section 2.2 the minimum bequest motive is a key determinant for the left tail of the wealth distribution. We adjust \bar{b} in order to broadly capture this tail and the overall level of inequality. The collateral constraints are set to $\phi_w = 0.75$ (equity ratio) and $\phi_e = 0.25$ (flow constraint) broadly following Iacoviello and Pavan (2013).

Following the evidence reported in Kaymak and Poschke (2016) we set the bequest tax to $\tau_b = 35\%$, with an exemption level w_E of 5 million US dollar which is approximately 10 times the mean wealth level. Following Angeletos (2007) we set $\sigma = 20\%$.⁴⁶ In the baseline we abstract from premia in the annuity market ($\theta = 0$).

Rios-Rull and Kuhn (2016) document a wealth-to-earnings ratio of 6.1. This ratio is far beyond the value usually targeted in Bewley-type models of roughly 3 (cf. e.g. Castaneda et al. (2003), Kaymak and Poschke (2016)). As argued in Storesletten et al. (2004) this value follows from disregarding the top 1% and their substantially higher wealth-to-income ratios. The baseline model produces a wealth-to-earnings ratio of 6.46 and thus still slightly overstates the evidence. Given the wealth-to-earnings ratio and a common capital share of $\alpha = 0.38$ (standard in the literature, cf. e.g. Castaneda et al. (2003)), we adjust the depreciation rate δ to get a realistic equilibrium interest that also shapes the top wealth distribution.⁴⁷ As shown in Proposition 2.2 the top tail measure a hinges on both the growth rate of wealth g_w and the Sharpe ratio S , which in turn depend on both the risky and the risk-free interest R respectively r which are objects determined in equilibrium. As – in contrast to the partial equilibrium world in e.g. Benhabib et al. (2011) and Benhabib et al. (2019) – we consider a general equilibrium context, these objects cannot be *hard-wired* into the model any longer in order to match the top tails.

The major preference parameters are subject to robustness checks in Online Appendix E.

⁴⁵Note that these values are already adjusted for a slightly different specification and assumption about mortality. DeNardi (2004) assumes $\chi = 16.18$.

⁴⁶Detailed recent micro studies for Norway, respectively, Sweden (Fagereng et al. (2016), Bach et al. (2016)) investigate the role of idiosyncratic risk in portfolio returns. We compare their evidence in the light of the model in Section 5.

⁴⁷In fact, for the calibration of the US economy we completely abstract from depreciation ($\delta = 0$).

		Gini	Quintile					Top		
			I	II	III	IV	V	10%	1%	0.1%
Evidence 2013	Income	0.58	3.03%	6.55%	10.90%	18.15%	61.37%	46.95%	19.68%	6.40%
	Earnings	0.67	-0.09%	3.00%	10.42%	20.19%	66.48%	49.62%	18.83%	6.41%
	Wealth	0.87	-0.67%	0.65%	3.22%	9.79%	87.02%	74.94%	35.47%	13.17%
Model	Income	0.56	3.13%	3.77%	10.39%	21.53%	61.18%	40.93%	20.56%	2.06%
	Earnings	0.59	1.67%	5.87%	9.81%	21.99%	60.66%	43.86%	16.60%	2.39%
	Wealth	0.82	-1.62%	0.11%	4.05%	13.83%	83.63%	66.30%	21.55%	5.80%

Table 2: Measures of inequality for income, earnings, and wealth - US evidence (Rios-Rull and Kuhn, 2016) and model predictions

The exogenous income process broadly matches evidence for income inequality in the USA. The non-targeted measure of earnings inequality also fares well with empirical evidence. For the distribution of wealth bottom inequality is overestimated, while top inequality is underestimated.

		Quintile					Shorrocks index
		I	II	III	IV	V	
Evidence	Income	0.86	0.41	0.47	0.46	0.66	0.535
	Wealth	0.67	0.47	0.45	0.50	0.71	0.55
Model	Income	0.70	0.63	0.45	0.49	0.65	0.52
	Wealth	0.67	0.50	0.50	0.62	0.84	0.469

Table 3: Diagonal elements for quintile income transition matrix and Shorrocks index - US evidence (Castaneda et al., 2003) and model predictions

Both evidence and model display a inverse u-shaped relationship of mobility. While the model matches income mobility rather well, it underestimates wealth mobility.

5. Quantitative investigations

In this Section we investigate the quantitative predictions of the calibrated model. Firstly, the predictions of the baseline calibration are considered and compared to empirical evidence for the USA. We employ this model in Section 5.2 in order to evaluate the impact of the bequest tax on both inequality and mobility. Other factors shaping both inequality and mobility – in particular the overall growth, demographics, and preference parameters – are considered in Section 5.3.

5.1. The baseline calibration

The model’s ability to account for inequality is compared to evidence from the year 2013 documented in Rios-Rull and Kuhn (2016) and reported in Table 2. As presented in Section 3.1 the income process is endogenously imposed and broadly fits the overall

income inequality as measured by the Gini-coefficient. By construction, the top 1% are also well matched. Due to the discrete nature of the income process the more narrow shares are, however, underestimated. Earnings are the sum of labor earnings and capital earnings.⁴⁸ The model slightly underestimates the (non-targeted) degree of earnings concentration at the bottom of the distribution. Our main focus lies on the distribution of wealth. While the overall wealth inequality is well matched, this comes at the cost of slightly overestimating bottom inequality and underestimating top inequality.⁴⁹ This contrasts with the partial equilibrium model with capital income risk presented by Benhabib et al. (2011) that overstates top inequality, while understating bottom and overall wealth inequality. In partial equilibrium, the top inequality can be wired-in as it does not have to obey a general equilibrium market clearing condition for the interest rates which to a large extent determines the top inequality (cf. Proposition 2.2).

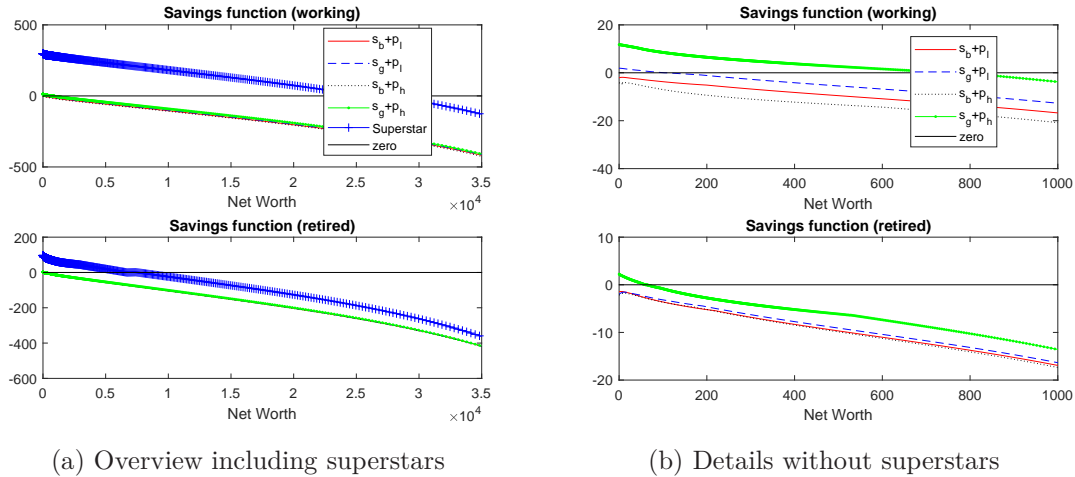


Figure 2: Savings as function of net worth.

While individuals with positive temporary shocks (\mathcal{S}_g) build up buffer-savings, individuals with negative shocks (\mathcal{S}_b) take on debt. Individuals in the *superstar* state always form positive savings.

This savings function is documented in Figure 2. The slope of the savings function – jointly incorporating the decisions rules for consumption c , portfolio composition μ

⁴⁸Formally they are computed as $dw + c$ in the model.

⁴⁹For the Pareto coefficient in the tails we measure a value of $a = 2.3$ as opposed to $a = 1.5$ observed in the evidence (Vermeulen, 2018).

and bequests b – has a negative slope, as suggested in Section 2.2, in order to satisfy a stable equilibrium holding of wealth. It is important to point out that the *superstars* (cf. Figure 2a) save substantially more than the average earners (cf. Figure 2b) not only in overall level, but even as a ratio of their already high labor income. For the working agents – besides the superstars – only the individuals with good temporary shocks (\mathcal{S}_g) accumulate wealth, whereas those with bad shocks (\mathcal{S}_b) incur debt. Among the retired individuals (lower panels of Figures 2) only superstars and those with both good temporary shocks and high permanent income (\mathcal{P}_h) form equilibrium savings in order to leave as bequests to their heirs.⁵⁰

Besides the widely discussed issue of inequality, this paper also investigates mobility. Table 3 presents evidence⁵¹ for mobility and compares evidence to the model results. We report the diagonal elements of the 5 year transition matrices and the resulting Shorrocks index as a simple scalar measure of mobility. High values for the diagonal elements and low values for the Shorrocks index indicate low mobility respectively high persistence.⁵² The model roughly matches income inequality yet slightly underestimates the level of mobility as captured by the Shorrocks index. The mobility of wealth is slightly underestimated. Interestingly, the model posits that wealth is less mobile than income, whereas the evidence suggests the opposing order. Part of this result might yet also be due to upward biased estimates of wealth mobility (cf. the discussion in footnote 51). In their model, aiming to fit the same evidence, Castaneda et al. (2003) underestimate the degree of mobility to an even larger extent. It is interesting to point out that both the model and the evidence display an inverse u-shape of mobility. Both the very bottom state (*poverty trap*) and the top state are highly persistent. In Table

⁵⁰To get an order of the magnitude in US dollars the reader should keep in mind a rough conversion between model currency unit and US dollars by factor 5,000.

⁵¹This evidence is taken from Castaneda et al. (2003). Rios-Rull and Kuhn (2016) also provide more novel evidence on mobility yet consider only the (two year) change between 2007 and 2009. They report substantially higher levels of mobility. This measure is, however, highly biased by the presence of the financial crisis, which peaked in 2008, presenting a major shock to the households portfolios in the USA. Even the measure reported in Castaneda et al. (2003) probably overstates the mobility of wealth by considering the years 1984 until 1989 featuring the substantial stock market crash in 1987. The celebrated study by Charles and Hurst (2003) investigating the correlation of wealth across generations, is also not helpful to benchmark our model as it explicitly excludes the role of bequests central to our investigation.

⁵²Finding a scalar measure of mobility is even more challenging than finding a scalar measure of inequality as mobility is characterized by a matrix whereas inequality is *only* a vector. The popular (trace) measure of Shorrocks (1978) is given by $0 \leq Mob = \frac{n - \sum_{i=1}^n M_{i,i}}{n-1} \leq 1$ with M being the transition matrix and n the number of states (here $n = 5$).

5 we provide more detailed evidence for the top 10%.⁵³ While it matches the overall evidence rather well, the model understates mobility at the very top. Meanwhile, the model understates the magnitude of the poverty trap for the flow measure of income (cf. Table 3).

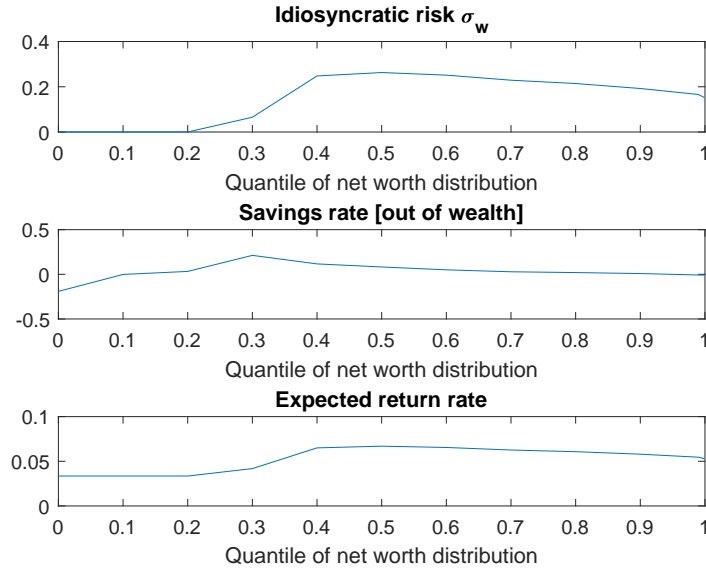


Figure 3: Savings rates and expected return rates as a function of the net worth distribution.

The model produces an inverse u-shaped relationships of savings rates (out of wealth), expected returns, and idiosyncratic risk along the net worth distribution.

Figure 3 presents measures of savings as well as return and return risk along the net worth distribution. As argued intensively in Section 2.2 both the savings rate out of wealth g_w – increasing with the expected return on wealth – and the idiosyncratic risk σ_w shape the distribution of wealth. The documented inverse u-shape behavior mirrors the debt holdings, as depicted in Figure 1 for households with positive net worth. Households with low net worth are severely constrained in their ability to lever up and hence hold portfolios with low returns and low idiosyncratic risk. The relationship peaks for households in the median of the net worth distribution with strongly leveraged portfolios, while high net worth households have a low leverage. The same pattern is documented for the savings rate, for which both households at the bottom and the top

⁵³Kennickell and Starr-McCluer (1997) report evidence for a six-year transition. We adjust the complete transition matrix to consider a five-year gap to make it comparable to the evidence presented in Table 3.

of the wealth distribution decrease their wealth. As already discussed, the latter is important to produce an ergodic distribution of wealth. The decreasing savings rate is also empirically documented in a detailed micro-study by Bach et al. (2017) for Sweden.

Bach et al. (2016) provide a detailed study of returns and portfolio risk for Swedish households in the years 2000 to 2007. They only investigate the top 60% of households (in order to exclude households with negative net worth). They show that for positive net worth households which range to the top 10%, returns as well as idiosyncratic wealth risk eventually (slightly) decrease as a function of net worth in line with our model. Quantitatively, Bach et al. (2016) report that returns range between 7.9% and 6.5%, whereas the standard deviation of returns amounts to values between 21% and 13.5%, broadly in line with our chosen calibration (cf. Figure 3). Yet, empirical evidence suggests that in the very top 10% there is a positive relationship between the net worth quantile and the return, respectively, the idiosyncratic risk. A similar pattern for Norway is documented in Fagereng et al. (2016). As discussed in detail in Section 3.2 and in line with Merton (1971), the (normative) model, however, predicts a decrease in risky asset holdings for very wealthy individuals due to the presence of human capital. This also implies lower returns and lower idiosyncratic wealth risk for the very top. Hence, this evidence at the very top is at odds with the standard normative theory.

Using the calibrated model we investigate how the policy maker influences the distribution and the mobility of wealth using a bequest tax. Moreover, we discuss the role of borrowing constraints on both inequality and mobility of wealth. The exercises in the following two sections focus on factors influencing the stationary distribution of wealth and thus reconcile the closed-form comparative statics regarding inequality and mobility of Section 2.2 numerically in a general equilibrium environment. This in the spirit of the empirical investigation by Corak (2013) who documents substantial cross-country difference for inequality and mobility (for the flow measure income). We aim at uncovering the determinants driving this behavior.

5.2. The role of the bequest tax and borrowing constraints

We consider changes in bequest taxes τ_b .⁵⁴ We investigate the case of a large tax, which we set to $\tau_b = 77\%$. As documented in Piketty and Saez (2013), this top tax rate was

⁵⁴The tax system is also characterized by the tax-free level w_E . A lower tax free level increases both the taxes at the bottom end and the level of transfers. The numerical investigation confirmed that the distribution of wealth remains largely unchanged.

	Gini Wealth	Quintile					Top		
		I	II	III	IV	V	10%	1%	0.1%
1 Baseline	0.82	-1.62%	0.11%	4.05%	13.83%	83.63%	66.30%	21.55%	5.80%
2 High taxes	0.81	-1.59%	0.18%	4.25%	14.41%	82.75%	64.89%	19.77%	5.33%
3 High taxes partial	0.81	-1.66%	0.05%	4.21%	14.50%	82.91%	64.92%	19.45%	5.06%
4 Low equity ratio	0.82	-1.68%	0.20%	4.32%	14.12%	83.05%	65.74%	20.99%	5.56%
5 Relaxed overall constraint	0.83	-1.91%	-0.10%	3.84%	13.95%	84.23%	66.73%	21.58%	5.79%

Table 4: Measures of wealth inequality in the model under different tax regimes. Higher taxes decrease wealth inequality. This effect is, however, dampened by a general equilibrium effect of increased interest rates and Sharpe ratios. Relaxing borrowing constraints increases wealth inequality at the bottom.

prevailing in the USA after World War II to the 1980s. Of course, higher taxes decrease top inequality. Yet, the change is only of low magnitude (cf. number 1 and 2 in Table 4).⁵⁵

First of all this is because the *effective* tax rate $\tau_b p \approx 0.58\%$ (for the baseline) is low to begin with. Note that in contrast to a pure wealth tax, bequests are only taxed on average all $1/p$ years. Our finding also contrasts with Benhabib et al. (2011) and Benhabib et al. (2014), who argue that in the presence of capital risk, wealth or bequest taxation has a substantial impact on wealth inequality. Of course, the effect is mostly visible for the top shares. Yet, the partial equilibrium rationale of Benhabib et al. (2011) or Benhabib et al. (2014) ignores an important effect. Line 3 of Table 4 presents the effect of an increase of the bequest tax in a partial equilibrium model which emphasizes the effect for the top wealth holders. There is, however, a counteracting effect in general equilibrium. With higher taxes wealthy individuals bequeath less and hence save less. The reduced savings, however, increase the interest rates – especially for the risky asset mostly held by wealthy individuals – and as such the Sharpe ratio determining the top inequality. As shown in Proposition 2.2, a higher Sharpe ratio increases the top inequality. The exact opposing reaction emerges for a tax reduction. In the words of Keynes (1936), the latter promotes a *euthanasia of the rentier* for which wealthy savers cannibalize themselves.

⁵⁵Like in Castaneda et al. (2003) we also investigated the effect of an abolition of the bequest tax. While top inequality is slightly increased, overall inequality remains largely unchanged comparable to their result.

		Shares							Shorrocks index
		0-25	25-50	50-75	75-90	90-95	95-99	99-100	
1	Evidence	0.706	0.534	0.592	0.522	0.406	0.511	0.646	0.514
	Baseline	0.744	0.557	0.654	0.599	0.384	0.600	0.700	0.460
2	High taxes	0.739	0.579	0.644	0.600	0.376	0.590	0.680	0.465
3	High taxes partial	0.744	0.560	0.642	0.600	0.380	0.585	0.680	0.468
4	Low equity ratio	0.679	0.545	0.628	0.597	0.376	0.605	0.700	0.478
5	Relaxed overall constraint	0.753	0.600	0.650	0.601	0.384	0.600	0.700	0.452

Table 5: Diagonal elements for extended wealth transition matrix and Shorrocks index – US evidence (Kennickell and Starr-McCluer, 1997) and model predictions for various tax regimes

Higher taxation is associated with higher mobility at the top. While the relaxation of the stock-dependent borrowing constraint increases mobility at the bottom, the relaxation of the overall constraint eventually decreases mobility at the bottom.

Yet, taxation also matters for our second objective – mobility. As shown in Table 5 higher top tax rates reduce persistence at the top.⁵⁶

The (negative) net worth at the bottom is driven by borrowing. We numerically investigate how the relaxation of borrowing constraints impact on the distribution by reducing the equity ratio ($\phi_w = 0.25$), respectively, the overall constraint ($\phi_e = 0.3$) increasing the value of $-\underline{w} > 0$. In line with Figure 1 the change in the equity ratio has a less pronounced impact as it only impacts individuals with low positive net worth. Yet, both cases increase wealth inequality at the bottom.

In terms of mobility, the two changes are visible at the lower end of the distribution, but result in opposing outcomes. An relaxation of the the equity ratio extends the possibility of individuals with very low (but positive) net worth and thus helps them escape their poverty trap. This group is not impacted by the increase of ϕ_e eventually aggravating the poverty trap.

We conclude this Section by stating that in general higher bequest taxes reduce both wealth inequality and increase mobility. The former effect might be low or even out-done due to a countervailing general equilibrium at the top. Yet, it is shown that the redistributive bequest tax increases mobility. Allowing for more borrowing increases

⁵⁶We also computed an overall welfare measure $\mathcal{W} = \sum_j \int U(c, b) g_j(w) dw$ for both the case with low and high taxes. There is a loss in terms of certainty equivalence terms $\left(\left(\frac{\mathcal{W}_{\tau_b=35\%}}{\mathcal{W}_{\tau_b=77\%}} \right)^{\frac{1}{1-\gamma}} - 1 \right)$ of 5.4% resulting mainly from the fact that lower capital accumulation also lowers overall output.

wealth inequality. If this is, however, allowed for the low net worth group (by relaxing borrowing constraints), this reduces the prevalence of the poverty trap.

5.3. Other factors influencing inequality

		Gini Wealth	Quintile					Top		
			I	II	III	IV	V	10%	1%	0.1%
1	Baseline	0.82	-1.62%	0.11%	4.05%	13.83%	83.63%	66.30%	21.55%	5.80%
2	Long life	0.86	-1.33%	-0.90%	2.93%	11.86%	87.43%	71.64%	26.03%	7.67%
3	No growth	0.82	-1.47%	0.46%	4.33%	13.54%	83.14%	66.50%	23.24%	6.84%
4	Homo. beq.	0.71	0.29%	3.04%	7.30%	15.35%	74.03%	58.23%	19.72%	5.60%
5	Min. cons.	0.73	0.45%	2.77%	6.43%	14.18%	76.16%	60.63%	21.43%	6.45%

Table 6: Measures of wealth inequality in the model under different specifications. A higher life expectancy (lower p) increase wealth inequality, while lower income growth rates leave it largely unchanged. Moreover, both homothetic bequest motives and a minimum consumption desire reduce wealth inequality.

		Shares						Shorrocks index	
		0-25	25-50	50-75	75-90	90-95	95-99		99-100
1	Baseline	0.744	0.557	0.654	0.599	0.384	0.600	0.700	0.460
2	Long life	0.745	0.565	0.679	0.627	0.424	0.625	0.700	0.439
3	No growth	0.746	0.607	0.654	0.596	0.372	0.595	0.700	0.455
4	Homo. beq.	0.713	0.522	0.568	0.573	0.376	0.570	0.680	0.500
5	Min.cons.	0.654	0.519	0.610	0.597	0.384	0.600	0.700	0.489

Table 7: Diagonal elements for extended wealth transition matrix and Shorrocks index – Different model specifications

Resulting from a decrease in income mobility, there is also lower wealth mobility in a scenario with longevity. Wealth mobility remains largely unchanged in a scenario without (homogeneous) income growth. Both homothetic bequest motives and a minimum consumption desire increase mobility.

As already discussed in Section 2.2 there are other factors which also determine both inequality and mobility. One important non-economic factor is demographics. We can simply capture demographics by the expected life time $1/p$. Due to the medical advances overall life expectancy has increased. Let us contrast our baseline model with the case for which (adult) life expectancy has increased to 80 years. We jointly assume that the working life itself does not change, meaning that the (expected) time spent in the retirement status increases.⁵⁷

⁵⁷Note that old individuals live on their pension income, which in the considered PAYG economy is substantially lower than the income for the working age population. As the share of the elderly

An increase in longevity reduces the effective discount rate $\tilde{\rho}$ and thus implies higher savings.⁵⁸ Longevity not only raises savings but also reduces the effective rate of taxation ($\tau_b p$) as the underlying event (the death of an individual) becomes less likely. Both factors increase top inequality in line with Propositions 2.2 and 2.4, as displayed in line 2 of Table 6.

Yet, there is another effect at the bottom end of the distribution. With a longer life expectancy individuals reduce their debt exposure; thus, the minimum wealth level $-\underline{w} > 0$ is also reduced.⁵⁹ In line with the comparative static logic of Proposition 2.2 this eventually leads to a reduction in inequality at the bottom. In the calibrated version of the model (cf. line 1 of Table 6) the former effect, however, dominates for the overall Gini-coefficient, suggesting an overall increase of wealth inequality for longevity.

The effects for wealth mobility are reported in Table 7. In line with the *Great Gatsby* argument, the increase in inequality is also reflected in a lower wealth mobility mostly in the middle of the distribution. The latter results primarily from a decrease in income mobility as individuals are stuck in the retirement state for a longer period of time.⁶⁰

Another exogenous factor frequently considered as a driver of inequality is the growth rate of income g . As we assume a homogeneous income growth rate, changes in the income growth rate do not have an impact on the distribution of income. Following the celebrated $r - g$ argument in Piketty (2014) a reduction of labor growth (relative to returns on capital) will lead to an increase in wealth inequality at the top. This is in line with Proposition 2.2. This effect is yet reduced in a general equilibrium environment. Lower growth of labor income requires higher savings and thereby reduce interest rates lowering the gap $r - g$. We consider the case of a zero growth ($g = 0$) economy.⁶¹ The calibrated model displays a (slight) increase in top inequality. Yet, there is another effect at the bottom of the distribution. The increasing gap $r - g$ emerges as a discount rate in order to compute human capital. As such, a reduction in g reduces human capital and hence the ability to lever up reflected in lower bottom inequality. Overall inequality –

poor population increases, income inequality also increases marginally leading to a Gini-coefficient of 0.57 (as opposed to 0.56 in the baseline calibration).

⁵⁸With perfect insurance markets ($\theta = 0$) we have a one-on-one reaction $\frac{\partial \tilde{c}_w}{\partial p} = 1$.

⁵⁹Formally the condition $\frac{\partial \underline{w}}{\partial p} < 0$ requires $\bar{b} > \frac{e - \bar{c}}{r - g}$ holding in our calibration. It is important to acknowledge that this condition never holds under homothetic preferences $\bar{b} = \bar{c} = 0$ for which an increase in longevity always increases debt.

⁶⁰The Shorrocks index for the 5-state income process reduces from 0.52 to 0.348 under longevity.

⁶¹A detailed discussion of this case in a model without capital risk is presented in Carroll and Young (2018).

as measured by the Gini-coefficient – remains largely unchanged. Similarly, the mobility remains largely unaffected.

Of course, preferences also matter. In line with DeNardi (2004), we presumed that bequests exhibit features of a luxury good. We can deviate from this assumption by setting $\bar{b} = 0$ and hence introduce homothetic preferences for bequests (cf. line 3 in Table 6). As such, the overall level of bequests increases reducing the prevalence of the poverty trap. The luxury structure in the bequest motive, thus, creates higher wealth inequality.

As usual we also investigate the impact on mobility of wealth. For homothetic bequest motives, mobility – especially in the lower tails of the distribution – substantially increases (cf. line 4 of Table 7). Thus, once again the luxury structure of the bequests locks in wealth status.

The assumption $\bar{c} > 0$ makes consumption an inferior good. In the baseline calibration we abstracted from that. Let us consider a value of $\bar{c} = 2$, which is 30% of the median income. As such even in this economy some individuals consume less than the (relative) poverty threshold defined as 50% of median income by the popular index of the United Nations. As predicted in Proposition 2.2 the minimum consumption motive reduces wealth inequality at the bottom as people form savings to satisfy their minimum consumption desire (cf. line 5 in Table 6). The presence of a minimum consumption desire increases mobility (cf. number 5 in Table 7) in line with the *Great Gatsby curve*.

In Appendix E we show that the model results are robust after assuming an annuity premium ($\theta > 0$), preferences with a prevailing income effect ($\gamma > 1$), or the absence of bequest motives ($\chi = 0$).

6. Conclusion

This paper discussed a rich micro-founded general equilibrium model to investigate both inequality and mobility of wealth. The model can be analytically solved to reveal factors that drive both bottom and top inequality. It is argued that the factors that contribute to higher wealth inequality in general also result in lower wealth mobility.

It is straightforward to put this theoretic notion to the empirical test. Unfortunately, both the evidence on inequality of the stock measure wealth and even more so measures on its mobility are too scarce to run a meaningful cross-country study. Thus, the gath-

ering and documentation of wealth mobility for several countries seems like a promising future research area.

We extended the model to feature risky labor income and non-trivial portfolio constraints. The analysis in general equilibrium also revealed that a simple partial-equilibrium analysis might yield misleading results. The model fits evidence for the USA. Unfortunately, the complete model – at least for the time being – has not been solved analytically. Another fruitful theoretical research area thus lies in shedding more analytic light on more elaborate models to gain economic intuition on the working mechanism rather than relying on purely numeric solutions. In particular, exploring the effect of risky income on savings with portfolio constraints seems worthwhile for a detailed analytic discussion.

Acknowledgments

I would like to thank Danial Ali Akbari, Dan Cao, Michael Funke, Domenico Delli Gatti, Hans Holter, Per Krusell, Ellen McGrattan, Mishael Milakovic, Kurt Mitman, Makoto Nirei, Olaf Posch, Morten Ravn, Víctor Ríos Rull, Alexis Akira Toda, and Mu-Chun Wang, and, more generally, seminar participants at Università Cattolica in Milano, Hamburg University, and Lund University as well as participants at the 2019 EEA-ESEM meeting, the 13th Nordic Summer Symposium in Macroeconomics, the 19th SAET meeting, the 2017 RES meeting, the 22nd CEF, and the 20th WEHIA conference for helpful comments.

A. Solution of the Fokker-Planck equation

A.1. Proof of proposition 2.2

As shown in (Karlin and Taylor, 1981, p. 221), the stationary solution to the probability density function is given by:

$$f(w) = \frac{C}{D(w)^2} \exp\left(2 \int_0^w \frac{s(x)}{D(x)^2} dx\right),$$

for which the integration constant C is determined by the standard integration property of the cumulative probability density (CPDF) ($\int_{w_{min}}^{w_{max}} f(w)dw = 1$). For our concrete case this implies:

$$\begin{aligned} f(w) &= \frac{C}{\sigma_w^2(w - w_{NB})^2} \exp\left(2\frac{g_w}{\sigma_w} \int (x - w_{NB})^{-1} dx\right) \\ &= \tilde{C} (w - w_{NB})^{\frac{2g_w}{\sigma_w^2} - 2} \sim (w - w_{NB})^{-(a+1)}, \end{aligned}$$

with $a = 1 - \frac{2g_w}{\sigma_w^2}$.

This is the Pareto II distribution as discussed in Arnold (1983) with the CPDF:

$$F(w) = 1 - (w - w_{NB})^{-a} \quad w \geq w_{NB} + 1.$$

Arnold (1983) shows that:

$$Gini(w) = 1 - \frac{w_{NB} + 1 + 2a \cdot B(2a - 1, 2)}{w_{NB} + 1 + a \cdot B(a - 1, 2)},$$

which when evaluating the Beta function with $B(2a - 1, 2) = \frac{1}{2a(2a-1)}$, respectively, $B(a - 1, 2) = \frac{1}{a(a-1)}$ leads to:

$$Gini(w) = \frac{1}{a + w_{NB}(a - 1)} \frac{a}{2a - 1}.$$

Following Arnold (1983), the top shares are given by:

$$s_x(w) = \frac{(w_{NB} + 1)x + a \cdot I_{x^{\frac{1}{a}}}(a - 1, 2)}{w_{NB} + 1 + a \cdot B(a - 1, 2)},$$

which, when evaluating the incomplete Beta function $I_{x^{\frac{1}{a}}}(a-1, 2) = \frac{1}{a(a-1)} (ax^{1-\frac{1}{a}} - (a-1)x)$, implies the final result:

$$s_x(w) = \frac{1}{a + w_{NB}(a - 1)} [ax^{1-\frac{1}{a}} + w_{NB}(a - 1)x].$$

A.2. Proof of proposition 2.4

A convenient way to solve FP equations – also used in Gabaix et al. (2016) and Kasa and Lei (2018) – is to use Laplace transforms and the log-transform $\tilde{w} \equiv \ln(w)$. The overall stochastic process (disregarding $w_{NB} \approx 0$ not mattering for the very top) reads:

$$dw_t = g_w w_t dt + \sigma_w w_t dZ_t + (\tilde{b}_w - 1) w_t dJ_t,$$

for a jump dJ_t emerging with the death probability p . Using Itô's lemma and the equivalent for jumps (cf. e.g. Section 4.2.4 in Hanson (2007)) the process for \tilde{w} reads:

$$d\tilde{w}_t = (g_w - 0.5\sigma_w^2)dt + \sigma_w dZ_t + \ln(\tilde{b}_w)dJ_t.$$

With the expectation operator \mathbb{E} , the associated FP equation is given by:

$$\frac{\partial f}{\partial t} = -\tilde{g}_w \frac{\partial f}{\partial \tilde{w}} + 0.5\sigma_w^2 \frac{\partial^2 f}{\partial \tilde{w}^2} + p\mathbb{E}[f(\tilde{w} - K) - f(\tilde{w})],$$

with $\tilde{g}_w \equiv g_w - 0.5\sigma_w^2$ and the jump $K \equiv \ln(\tilde{b}_w)$. The two-sided Laplace transform $\mathcal{L}\{f(\tilde{w})\} \equiv F(s) \equiv \int_{-\infty}^{\infty} f(\tilde{w}) \exp(-s\tilde{w})d\tilde{w}$ is a very convenient tool to solve this partial differential equation. In particular, we have $\mathcal{L}\{f^n(\tilde{w})\} = s^n F(s)$ for the n -th order derivative. Following Gabaix et al. (2016) the jump part can be written as:

$$\begin{aligned} \mathbb{E}[f(\tilde{w} - K) - f(\tilde{w})] &= \int_{-\infty}^{\infty} [f(\tilde{w} - K) - f(\tilde{w})]h(K)dK \\ &= (f * h)(\tilde{w}) - f(\tilde{w}), \end{aligned}$$

for which $h(K)$ is the distribution of the jumps and $*$ signifies the convolution operator. When performing the Laplace transform for this, we get:

$$\mathcal{L}\{(f * h)(\tilde{w}) - f(\tilde{w})\} = F(s)H(s) - F(s).$$

As the jump itself is uniform at the value K , it is easy to see that $h = \delta(\tilde{w} - K)$ for δ being the Dirac impulse. The Laplace transform of the latter is $\mathcal{L}\{\delta(\tilde{w} - K)\} = \exp(-Ks) = H(s)$. As such the Laplace transform of the overall FP equation becomes:

$$\frac{\partial F(s)}{\partial t} = s\tilde{g}_w - 0.5\sigma_w^2 s^2 + p[1 - \exp(-s \ln(\tilde{b}_w))] = s\tilde{g}_w - 0.5\sigma_w^2 s^2 + p(1 - \tilde{b}_w^{-s}) = \lambda(s), \quad (34)$$

also constituting the characteristic equation $\lambda(s)$. To find the Pareto coefficient we need to solve $\lambda(s = -a) = 0$ implying:

$$-\tilde{g}_w a - 0.5\sigma_w^2 a^2 + p(1 - \tilde{b}_w^a) = 0. \quad (35)$$

We approximate the non-trivial root ($a \neq 0$) by conducting a first-order Taylor approximation around $\tilde{b}_w = 1$ implying $\tilde{b}_w^a - 1 \approx a(\tilde{b}_w - 1)$. This results in the following (approximate) equation characterizing the right tail:

$$\begin{aligned} 0.5\sigma_w^2 a^2 + (g_w - 0.5\sigma_w^2 + p(\tilde{b}_w - 1))a &= 0 \\ \rightarrow \tilde{a} &= a + \frac{2p}{\sigma_w^2}(1 - \tilde{b}_w). \end{aligned} \quad (36)$$

Following Gabaix et al. (2016) – exploiting the relationship between the moment-generating function and the Laplace transform – the characteristic equation is also helpful to consider convergence rates. For example, the mean reversion rate is given by $\lambda(s = -1)$:

$$\lambda(s = -1) = -\tilde{g}_w - 0.5\sigma_w^2 + p(1 - \tilde{b}_w) = -g_w + p(1 - \tilde{b}_w). \quad (37)$$

A higher mean reversion rate is accompanied by more mobility. It is easy to see that for $p = 0$ this nests the case discussed in proposition 2.3. For the special case of $b_w = 1$ we have:

$$\lambda(s = -1) = -g_w + p\tau_b.$$

For the Pareto distribution $f(w) \sim w^{-(\tilde{a}+1)}$ corresponding to $f(\tilde{w}) \sim \exp(-\tilde{a}\tilde{w})$ and the Laplace transform $F(s) = \frac{1}{s+\tilde{a}}$, only $0 < -s < \tilde{a}$ finite moments exist. For higher moments $1 < -s < \tilde{a}$ – mattering especially for the top shares – there is also an effect of σ_w on the convergence rate $\lambda(s)$. In fact higher values of σ_w – in the structural model related to higher Sharpe ratios and low risk aversion – are associated with lower mobility.

References

- Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2017). Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. NBER Working Papers 23732, National Bureau of Economic Research, Inc.
- Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics* 109(3), 659–84.
- Açikgöz, O. (2018). On the existence and uniqueness of stationary equilibrium in bewley economies with production. *Journal of Economic Theory* 173, 18 – 55.
- Angeletos, G.-M. (2007). Uninsured Idiosyncratic Investment Risk and Aggregate Saving. *Review of Economic Dynamics* 10(1), 1–30.
- Aoki, S. and M. Nirei (2017). Zipf’s Law, Pareto’s Law, and the Evolution of Top Incomes in the U.S. *American Economic Journal: Macroeconomics* forthcoming.
- Arnold, B. C. (1983). *Pareto Distributions*. Number 2nd edition. Chapman and Hall/CRC.
- Bach, L., L. Calvet, and P. Sodini (2017). From saving comes having? disentangling the impact of savings on inequality. Swedish House of Finance Research Paper 18-8.
- Bach, L., L. E. Calvet, and P. Sodini (2016). Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy. CEPR Discussion Papers 11734, C.E.P.R. Discussion Papers.
- Becker, G. S. and N. Tomes (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy* 87(6), 1153–1189.

- Benhabib, J., A. Bisin, and M. Luo (2019). Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review* 109(5), 1623–47.
- Benhabib, J., A. Bisin, and S. Zhu (2011). The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents. *Econometrica* 79(1), 123–157.
- Benhabib, J., A. Bisin, and S. Zhu (2014). The distribution of wealth in the blanchard–yaari model. *Macroeconomic Dynamics FirstView*, 1–16.
- Blanchard, O. J. (1985). Debt, Deficits, and Finite Horizons. *Journal of Political Economy* 93(2), 223–47.
- Bodie, Z., R. C. Merton, and W. F. Samuelson (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control* 16(3), 427 – 449.
- Bossmann, M., C. Kleiber, and K. Wälde (2007). Bequests, taxation and the distribution of wealth in a general equilibrium model. *Journal of Public Economics* 91(7-8), 1247–1271.
- Cao, D. and W. Luo (2017). Persistent Heterogeneous Returns and Top End Wealth Inequality. *Review of Economic Dynamics* 26, 301–326.
- Carroll, D. R. and E. R. Young (2018). The Piketty Transition. *Journal of Macroeconomics forthcoming*.
- Castaneda, A., J. Diaz-Gimenez, and J.-V. Rios-Rull (2003). Accounting for the U.S. Earnings and Wealth Inequality. *Journal of Political Economy* 111(4), 818–857.
- Charles, K. K. and E. Hurst (2003). The Correlation of Wealth across Generations. *Journal of Political Economy* 111(6), 1155–1182.
- Clark, G. and N. Cummins (2015). Intergenerational Wealth Mobility in England, 1858–2012: Surnames and Social Mobility. *Economic Journal* 125(582), 61–85.
- Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives* 27(3), 79–102.
- Davies, J. B. (1986). Does Redistribution Reduce Inequality? *Journal of Labor Economics* 4(4), 538–559.
- DeNardi, M. (2004). Wealth Inequality and Intergenerational Links. *Review of Economic Studies* 71(3), 743–768.
- Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016). Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality. CEPR Discussion Papers 11047, C.E.P.R. Discussion Papers.

- Fernholz, R. T. (2016). A model of economic mobility and the distribution of wealth. *Journal of Macroeconomics* 50, 168 – 192.
- Fischer, T. (2018). Modeling inequality and mobility with stochastic processes. Ssrn electronic papers.
- Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016). The dynamics of inequality. *Econometrica* 84(6), 2071–2111.
- Gouin-Bonenfant, E. and A. A. Toda (2019). Pareto extrapolation: Bridging theoretical and quantitative models of wealth inequality. Ssrn electronic papers.
- Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014). Marry Your Like: Assortative Mating and Income Inequality. *American Economic Review* 104(5), 348–353.
- Hanson, F. B. (2007). *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation (Advances in Design and Control)*. Society for Industrial and Applied Mathematics.
- Holm, M. B. (2018). Consumption with liquidity constraints: An analytical characterization. *Economics Letters* 167, 40 – 42.
- Hubmer, J., P. Krusell, and J. Anthony A. Smith (2016). The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation. NBER Working Papers 23011, National Bureau of Economic Research, Inc.
- Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control* 17(5-6), 953–969.
- Iacoviello, M. and M. Pavan (2013). Housing and debt over the life cycle and over the business cycle. *Journal of Monetary Economics* 60(2), 221–238.
- Jäntti, M. and S. Danziger (2000). Chapter 6 income poverty in advanced countries. Volume 1 of *Handbook of Income Distribution*, pp. 309 – 378. Elsevier.
- Karlin, S. and H. Taylor (1981). *A Second Course in Stochastic Processes*. Number 2. Academic Press.
- Kasa, K. and X. Lei (2018). Risk, uncertainty, and the dynamics of inequality. *Journal of Monetary Economics* 94, 60 – 78.
- Kaymak, B. and M. Poschke (2016). The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics* 77(C), 1–25.
- Kennickell, A. B. and M. Starr-McCluer (1997). Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel. *Review of Income and Wealth* 43(4), 381–399.

- Keynes, J. M. (1936). *The General Theory of Employment Interest and Money*. London: Macmillan.
- Kiyotaki, N. and J. Moore (1997). Credit Cycles. *Journal of Political Economy* 105(2), 211–48.
- Krueger, D., K. Mitman, and F. Perri (2016). *Macroeconomics and Household Heterogeneity*, Volume 2 of *Handbook of Macroeconomics*, Chapter 0, pp. 843–921. Elsevier.
- Krusell, P. and T. Smith (1998). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy* 106(5), 867–896.
- Lansing, K. J. (1999). Optimal redistributive capital taxation in a neoclassical growth model. *Journal of Public Economics* 73(3), 423–453.
- Merton, R. C. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics* 51(3), 247–57.
- Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* 3(4), 373–413.
- Nuño, G. and B. Moll (2018). Social optima in economies with heterogeneous agents". *Review of Economic Dynamics* 28, 150–180.
- Parra-Alvarez, J. C., O. Posch, and M.-C. Wang (2015). Identification and estimation of heterogeneous agent models: A likelihood approach. Working paper, Aarhus University Denmark.
- Piketty, T. (2014). *Capital in the twenty-first century*. Boston: Harvard University Press.
- Piketty, T. and E. Saez (2013). A Theory of Optimal Inheritance Taxation. *Econometrica* 81(5), 1851–1886.
- Pugh, T. (2018). Wealth and Mobility: Superstars, Returns Heterogeneity and Discount Factors. Technical report, Working paper.
- Richard, S. F. (1975). Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model. *Journal of Financial Economics* 2(2), 187 – 203.
- Rios-Rull, J.-V. and M. Kuhn (2016). 2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics. *Quarterly Review* (April), 1–75.
- Rouwenhorst, K. G. (1995). Asset Pricing Implications of Equilibrium Business Cycle Models. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*. Princeton, New Jersey: Princeton University Press.

- Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica* 46(5), 1013–1024.
- Sornette, D. and R. Cont (1997). Convergent multiplicative processes repelled from zero: Power laws and truncated power laws. *J. Phys. I France* 7(3), 431–444.
- Storesletten, K., C. I. Telmer, and A. Yaron (2004). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics* 51(3), 609–633.
- Straub, L. and I. Werning (2014). Positive Long Run Capital Taxation: Chamley-Judd Revisited. NBER Working Papers 20441, National Bureau of Economic Research, Inc.
- Toda, A. A. (2018). Wealth distribution with random discount factors. *Journal of Monetary Economics*.
- Vermeulen, P. (2018). How Fat is the Top Tail of the Wealth Distribution? *Review of Income and Wealth* 64(2), 357–387.
- Wan, J. and S. Zhu (2019). Bequests, estate taxes, and wealth distributions. *Economic Theory* 67(1), 179–210.
- Wang, C., N. Wang, and J. Yang (2016). Optimal consumption and savings with stochastic income and recursive utility. *Journal of Economic Theory* 165, 292 – 331.
- Yaari, M. (1965). Uncertain Lifetime, Life Insurance and the Theory of the Consumer. *Review of Economic Studies* 32, 137–150.

B. Proof of proposition 2.1

As already discussed, the prices of risky assets $P_{R,t}$ follow a Geometric Brownian Motion:

$$\frac{dP_{R,t}}{P_{R,t}} = Rdt + \sigma dZ_t,$$

while prices for safe risk-free assets $P_{s,t}$ follow a simple drift process:

$$\frac{dP_{s,t}}{P_{s,t}} = rdt.$$

Individuals hold a number N_i in each asset making overall wealth $W_t = \sum_{i=1}^2 N_{i,t}P_{i,t}$. As such, the evolution of overall wealth is given by:

$$dW_t = \sum_{i=1}^2 N_{i,t}dP_{i,t} + (E_t - C_t)dt,$$

with earnings E_t and consumption C_t . Defining $\mu_t = \frac{N_{R,t}P_{R,t}}{W_t}$ and hence $1 - \mu_t = \frac{N_{s,t}P_{s,t}}{W_t}$, this leads to:

$$dW_t = (\mu_t(R - r)W_t + rW_t + E_t - C_t)dt + \sigma\mu_tW_t dZ_t. \quad (38)$$

Taking into account the growth rate of income g and the bequest flow, we end up with the Hamilton-Jacobi-Bellman (HJB) equation. Here, we provide the solution to HJB characterizing the household problem:

$$(\rho + p - (1 - \gamma)g)V = \max_{c, b, \mu} \{U + V'((p(1 + \theta) + r - g)w + \mu(R - r)w + e - c - p(1 + \theta)b) + 0.5\sigma^2\mu^2w^2V''\}.$$

The first order condition for consumption c is:

$$\frac{\partial U}{\partial c} = (c - \bar{c})^{-\gamma} = V',$$

comparable to the first-order condition of bequests:

$$\frac{\partial U}{\partial b} = p(1 - \tau_b)^{1-\gamma}\chi(b + \bar{b})^{-\gamma} = p(1 + \theta)V'$$

Thus, we have $b = (1 + \theta)^{-\frac{1}{\gamma}}(c - \bar{c})(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}} - \bar{b}$. The first-order condition for the optimal portfolio is:

$$\mu = -\frac{R - r}{\sigma^2} \frac{V'}{wV''}.$$

As usual, we have to guess a value function. Our educated guess is:

$$V(w) = \frac{K_1(w + K_2(e - \bar{c} + p(1 + \theta)\bar{b}))^{1-\gamma}}{1 - \gamma},$$

for some parameters K_1 and K_2 to be determined. For the first-order conditions this implies:

$$c = \bar{c} + K_1^{-1/\gamma}(w + K_2(e - \bar{c} + p(1 + \theta)\bar{b})) = \bar{c} + K_1^{-1/\gamma}X,$$

as well as:

$$b = (1 + \theta)^{-\frac{1}{\gamma}}\chi^{\frac{1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}K_1^{-1/\gamma}X - \bar{b}$$

and

$$\mu w = \frac{(R - r)}{\gamma\sigma^2}(w + K_2(e - \bar{c} + p(1 + \theta)\bar{b})) = \frac{(R - r)}{\gamma\sigma^2}X,$$

for which $X \equiv w + K_2(e - \bar{c} + p(1 + \theta)\bar{b})$.

Inserting this result into the HJB we can derive:

$$\begin{aligned} & (\rho + p - (1 - \gamma)g)\frac{K_1X^{1-\gamma}}{1 - \gamma} \\ = & K_1^{-1/\gamma}\frac{K_1X^{1-\gamma}}{1 - \gamma}\left(1 + p(1 + \theta)^{\frac{\gamma-1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}\right) + K_1X^{-\gamma}\left([(1 + \theta)p + r - g]w + \frac{(R - r)^2}{\gamma\sigma^2}X\right. \\ & \left. - K_1^{-1/\gamma}X\left(1 + p(1 + \theta)^{\frac{\gamma-1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}\right) - \bar{c} + e + p(1 + \theta)\bar{b}\right) - K_1X^{1-\gamma}\frac{(R - r)^2 \cdot 0.5}{\gamma\sigma^2}. \end{aligned}$$

It is easy to see that $1/K_2 = (1 + \theta)p + r - g$. Inserting this result, performing some algebraic manipulations, and dividing by the common factor $K_1X^{1-\gamma}$ leads to the characteristic equation:

$$\begin{aligned} & K_1X^{1-\gamma}([(1 - \gamma)g - \rho - p] \\ + & \gamma K_1^{-1/\gamma}\left(1 + p(1 + \theta)^{\frac{\gamma-1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}\right) + (1 - \gamma)((1 + \theta)p + r - g) + (1 - \gamma)\frac{0.5(R - r)^2}{\gamma\sigma^2}) = 0, \end{aligned}$$

for which we can solve:

$$\begin{aligned} K_1^{-1/\gamma}\left(1 + p(1 + \theta)^{\frac{\gamma-1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}\right) &= \frac{p(\gamma - (1 - \gamma)\theta) + \rho - (1 - \gamma)r}{\gamma} - 0.5\frac{1 - \gamma}{\gamma}\frac{(R - r)^2}{\gamma\sigma^2} \\ \rightarrow K_1^{-1/\gamma} &= \frac{\rho + p(\gamma - (1 - \gamma)\theta) - (1 - \gamma)\left(r + 0.5\frac{(R - r)^2}{\gamma\sigma^2}\right)}{\gamma\left(1 + p(1 + \theta)^{\frac{\gamma-1}{\gamma}}(1 - \tau_b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}\right)} = c_w. \end{aligned}$$

The value $K_1^{-1/\gamma}$ is also the marginal propensity to consume out of wealth c_w . In fact, the optimal consumption function looks as follows:

$$c_t = \bar{c} + c_w \left(w_t + \frac{e - \bar{c} + p(1 + \theta)\bar{b}}{p(1 + \theta) + r - g} \right) = \bar{c} + c_w X_t,$$

while the optimal portfolio composition is given by:

$$\mu_t = \frac{R - r}{\gamma\sigma^2} \frac{w_t + \frac{e - \bar{c} + p(1 + \theta)\bar{b}}{p(1 + \theta) + r - g}}{w_t} = \tilde{\mu} \frac{w_t + \frac{e - \bar{c} + p(1 + \theta)\bar{b}}{p(1 + \theta) + r - g}}{w_t} = \tilde{\mu} \frac{X_t}{w_t}.$$

In this case, $\tilde{\mu} = \frac{R - r}{\gamma\sigma^2}$ represents the standard optimal portfolio share of risky assets (cf. e.g. Merton (1969)) which would prevail if $e - \bar{c} + p(1 + \theta)\bar{b} = 0$.

Finally, we can compute the bequest values as:

$$b_t = \frac{(1 + \theta)^{-\frac{1}{\gamma}} (1 - \tau_b)^{\frac{1 - \gamma}{\gamma}} \chi^{\frac{1}{\gamma}}}{1 + p(1 + \theta)^{\frac{\gamma - 1}{\gamma}} (1 - \tau_b)^{\frac{1 - \gamma}{\gamma}} \chi^{\frac{1}{\gamma}}} \tilde{c}_w X_t - \bar{b},$$

for which $\tilde{c}_w = \frac{\rho + p(\gamma(1 + \theta) - \theta) - (1 - \gamma)r}{\gamma} - 0.5 \frac{1 - \gamma}{\gamma} \frac{(R - r)^2}{\gamma\sigma^2}$ represents the marginal propensity to consume that would prevail absent a bequest motive.

We can rewrite the consumption function in the following manner:

$$\begin{aligned} c_t &= \bar{c} + c_w \left(w_t + \frac{e - \bar{c} + p(1 + \theta)\bar{b}}{r - g} \right) \\ &= (1 - c_e)\bar{c} + c_e(e + p(1 + \theta)\bar{b}) + c_w w_t = (1 - c_e)\bar{c} + c_e p(1 + \theta)\bar{b} + c_w(w_t + h), \end{aligned}$$

with $c_w \equiv K_1^{-1/\gamma}$ (as given by equation 5), $K_2 = \frac{1}{(1 + \theta)p + r - g}$, $c_e = K_2 c_w$, and $h = \frac{e}{(1 + \theta)p + r - g}$ being the human capital.

C. Income transition matrices

Let us start with the transitory shocks. Following the seminal work of Huggett (1993) we assume two time-varying states $\mathcal{S}_{j,t}$, with $j \in \{g, b\}$ representing a good and a bad state ($\mathcal{S}_g > \mathcal{S}_b$). Huggett (1993) approaches this in a very general manner, yet the straight-forward interpretation of the two states are employment and unemployment. The switching between states is modeled by a Markov transition matrix $\mathbb{M}_{g,b}$ given by:

$$\mathbb{M}_{g,b} = \begin{bmatrix} 1 - \lambda_b & \lambda_b \\ \lambda_g & 1 - \lambda_g \end{bmatrix}, \quad (39)$$

for which $\lambda_g > 0$ captures the transition probability from good to bad, while $\lambda_b > 0$ is the transition from bad to good.

Secondly, we turn to permanent shocks. We assume that permanent shocks are of either a high or low nature ($j \in \{h, l\}$, with $\mathcal{P}_h > \mathcal{P}_l$). Permanent shocks are absorbing states, making the transition matrix $\mathbb{M}_{h,l}$ the identity matrix.

Combining the two shocks results in the following 4-state transition matrix:⁶²

$$\mathbb{M}_{\mathcal{P};\mathcal{S}} = \begin{bmatrix} \mathbb{M}_{g,b} & \mathbf{0} \\ \mathbf{0} & \mathbb{M}_{g,b} \end{bmatrix} \quad (40)$$

Finally, we consider the retirement state. Once again, we model this as a Markov transition matrix $\mathbb{M}_{r,w}$ (with r signifying retired and w working):

$$\mathbb{M}_{r,w} = \begin{bmatrix} 1 - p_w & p_w \\ p_r & 1 - p_r \end{bmatrix}.$$

The value $\frac{1}{p_w}$ captures the average time in employment, whereas $\frac{1}{p_r}$ measures the average time in retirement. The overall expected life length amounts to $\tilde{T} = \frac{1}{p} = \frac{1}{p_r} + \frac{1}{p_w}$. The period of retirement is not planned in advance, but comes as an exogenous shock. If a retired individual dies (with probability p_r), she will be replaced by her offspring who are in the working age.⁶³

The intergenerational transition matrix \mathbb{M}_{gen} is constructed as follows. Assume an intergenerational correlation ρ_z . Using the method of Rouwenhorst (1995) for discretizing a continuous AR(1) process into two states results in:

$$\mathbb{M}_{gen} = \begin{bmatrix} \frac{1+\rho_z}{2}\mathbb{M}_{g,b} & \frac{1-\rho_z}{2}\mathbb{M}_{g,b} \\ \frac{1-\rho_z}{2}\mathbb{M}_{g,b} & \frac{1+\rho_z}{2}\mathbb{M}_{g,b} \end{bmatrix}. \quad (41)$$

Despite having 64 entries the matrix is controlled by 5 parameters only ($\lambda_g, \lambda_b, p_w, p_r, \rho_z$).⁶⁴

We add a superstar state (Castaneda et al., 2003) for both working and retired people constituting the overall transition matrix:

$$\mathbb{M}^e = \begin{bmatrix} (1 - \lambda_s)\mathbb{M}_{\mathcal{P};\mathcal{S}}(1 - p_w) & \lambda_s \mathbf{1}(1 - p_w) & p_w \mathbb{I} & \mathbf{0} \\ (1 - p_w) \cdot 0.25 \cdot \mathbf{1}' & 0 & \mathbf{0}' & p_w \\ p_r \mathbb{M}_{gen} & \mathbf{0} & (1 - p_r) \mathbb{I} & \mathbf{0} \\ p_r \cdot 0.25 \cdot \mathbf{1}' & 0 & \mathbf{0}' & (1 - p_r) \end{bmatrix}. \quad (42)$$

Table 9 reports the matrix for the process calibrated as detailed in section 3.1. The resulting stationary distribution with the respective efficiency states z_j is documented in Table 8.

⁶²Note that the zero entry is a square matrix of order 2.

⁶³A similar modeling is used in Krueger et al. (2016).

⁶⁴Eventually, of the 64 possible entries 32 are zeros.

State j	1	2	3	4	5	6	7	8	9	10	\tilde{z}
z_j	1	3.4	4.2	14.28	163.35	0.46	1.56	1.93	6.55	74.93	6.31
π_j [%]	18.75	18.75	18.75	18.75	0.74	6	6	6	6	0.24	100

Table 8: Labor efficiency states and stationary distribution for the model calibrated to the USA

The value \tilde{z} is the average labor efficiency.

0.774	0.194	0	0	0.010	0.022	0	0	0	0
0.194	0.774	0	0	0.010	0	0.022	0	0	0
0	0	0.774	0.194	0.010	0	0	0.022	0	0
0	0	0.194	0.774	0.010	0	0	0	0.022	0
0.245	0.245	0.245	0.245	0	0	0	0	0	0.022
0.036	0.009	0.019	0.005	0	0.931	0	0	0	0
0.009	0.036	0.005	0.019	0	0	0.931	0	0	0
0.019	0.005	0.036	0.009	0	0	0	0.931	0	0
0.005	0.019	0.009	0.036	0	0	0	0	0.931	0
0.017	0.017	0.017	0.017	0	0	0	0	0	0.931

Table 9: Labor efficiency transition matrix (discrete time) for the model calibrated to the USA

Details on the structure of the matrix are presented in Section 3.1.

D. Solution algorithm

For the model we employ the solution algorithm developed, described, and explained in Achdou et al. (2017). The use of a continuous time framework is not only helpful for gaining more analytic intuition, but allows for efficient numerical solutions to the model. In particular, it uses the fact that for small ($\lim_{\Delta \rightarrow 0} \frac{x_{t+\Delta} - x_t}{\Delta} = \frac{dx}{dt}$) time steps a continuous process can only move into the adjacent bin of the state space.⁶⁵ The model is implemented in the software MATLAB using its efficient methods for handling sparse matrices, allowing for fast computation. For more details the interested reader is referred to Achdou et al. (2017).

1. **Grid** Construct a discrete grid for the endogenous state of wealth $w_i \in \{\underline{w}, \bar{w}\}$ with I elements.⁶⁶ Constructing a reasonable grid that features both a detailed view of the bottom 90% (having non-linear policy rules) as well as the fat top tails is not trivial. We choose the following specification:

$$w = \frac{\bar{w} - \underline{w}}{C_1 + C_2} (C_1 x + C_2 \cdot x^P) + \underline{w},$$

with some constants $C_1 = 1$, $C_2 = 50 > 0$, and a power exponent $P = 10 > 1$ for $x \in \{0; 1\}$.⁶⁷ The exogenous state of labor endowment z_j features J elements. Thus, the state space (w_i, z_j) is a vector with $I \cdot J$ elements and the considered square transition matrices will be of dimension $I \cdot J$. For the exogenous income switching process we can construct a square income switching matrix M^z of dimension $I \cdot J$. The matrix elements are given by:

$$\mathbb{M}_{1+(i-1) \cdot I : i \cdot I, 1+(j-1) \cdot I : j \cdot I}^z = \mathbb{I} \cdot \mathbb{R}_{i,j}, \quad i, j \forall \{1, 2, \dots, J\}$$

with \mathbb{I} being an identity matrix of dimension I . Basically, this is a *blown-up* version of the regular continuous time income transition matrix \mathbb{R} in order to account for the I wealth states. This can be written in a more compact way using the tensor product: $\mathbb{M}^z = \mathbb{R} \otimes \mathbb{I}$.

2. **Initial guess** Guess an initial value V_{ij} for the value function starting from the closed-form solution of the HJB as elaborated in Section B. Note that the solution holds exactly in the absence of portfolio constraints (i.e. borrowing and short-sale

⁶⁵Note, however, that for jumps any part of the state space can be reached even with continuous time steps.

⁶⁶Gouin-Bonenfant and Toda (2019) argue that if the truncation level \bar{w} is chosen too low the model will underestimate top tails as well as aggregate capital. They provide an approach in discrete time using the closed-form solution of the tail in order to overcome this issue. Our ratio between maximum \bar{w} and capital in the representative agent economy is roughly 10^3 , which in their paper was shown to give reasonably small errors. Moreover, we use a substantially larger grid ($I = 10^3$) compared to their approach which was also shown to reduce error. Finally, we tried specifications, in which we substantially increased \bar{w} , showing no major differences in the outcome.

⁶⁷Gouin-Bonenfant and Toda (2019) similarly suggest to have a linear grid for low wealth levels and an exponential grid in the upper tail.

constraints) and deterministic income. Also guess an initial factor price R from which we can deduce k_D and thus wages ω . Furthermore, provide an initial guess for the risk-free rate r . Also set a range of $R_{max} = 0.99\tilde{\rho}$ and $R_{min} = r$.

3. **HJB: Inner loop** This is basically the discretized version of the HJB equation. For the current loop step q compute the local first and second-order derivatives of the value function using finite differences. The derivatives are computed both as forward and backward derivatives (index f respectively b). Given prices R , r , and ω , compute the individual decisions using the respective local first-order conditions from the value function and its local derivatives. Aggregate the decision for consumption c , portfolio composition μ , and bequests b to find an overall savings rule $dw = s$ both for the forward and backward difference. The partial derivative of the value function is aggregated as follows:

$$dV = dV_f \mathbb{1}_{s_f > 0} + dV_b \mathbb{1}_{s_b < 0} + dV_0(1 - \mathbb{1}_{s_f > 0} - \mathbb{1}_{s_b < 0}).$$

for which $\mathbb{1}_i \in \{0; 1\}$ represents an indicator function depending on the sign of the drift in the forward, respectively, backward equation. For the special case of $s_f < 0 < s_b$ it is assumed that savings are zero, for which dV_0 is the partial derivative of the policy rule corresponding to this policy. The term dV is employed to compute the final policy rules for consumption c , bequests b , and portfolio composition μ (regardless of forward or backward difference). Using the savings rules s_f and s_b we construct a sparse quadratic matrix \mathbb{M}^s with I elements with a diagonal vector with elements $D \leq 0$ and an upper, respectively, lower diagonal $U, L \geq 0$. Following from the drift and diffusion terms in the HJB the vectors are given by:

$$\begin{aligned} D &= -\mathbb{1}_{s_f > 0} \frac{s_f}{dw_f} + \mathbb{1}_{s_b < 0} \frac{s_b}{dw_b} - \sigma^2(\mu w)^2 \frac{1}{dw_b dw_f}, \\ U &= \mathbb{1}_{s_f > 0} \frac{s_f}{dw_f} + 0.5\sigma^2(\mu w)^2 \frac{2}{(dw_f + dw_b)dw_f}, \\ L &= -\mathbb{1}_{s_b < 0} \frac{s_b}{dw_b} + 0.5\sigma^2(\mu w)^2 \frac{2}{(dw_f + dw_b)dw_b}, \end{aligned}$$

For the stationary equilibrium $s_f < 0 < s_b$ and thus $\mathbb{1}_{s_f > 0} = \mathbb{1}_{s_b < 0} = 0$. Basically, the equation thereby decides on whether to employ the forward or backward difference. The values dw_i represent the forward, respectively, backward difference in the wealth grid.⁶⁸ In order to implement a reflecting barrier at \bar{w} for the last element $i = I$ the following assumptions are made:

$$\mathbb{M}_{I,I}^s = \frac{\tilde{s}_b}{dw_b} - 0.5\sigma^2\bar{w}^2\mu^2 \frac{1}{dw_b},$$

⁶⁸Note that for a linear spaced grid $dw_b = dw_f \equiv dw$, simplifying the equations substantially.

with $\tilde{s}_b = \min\{s_b, 0\}$ and in order to ensure the standard property of continuous time Markov chains:

$$\mathbb{M}_{I,I-1}^s = -\mathbb{M}_{I,I}^s.$$

This matrix \mathbb{M}^s , which consists of the vectors D , U and L , is quadratic of size I and has to be *blown up* by the dimension of the income states J in the following manner:

$$\mathbb{M}_{1+I \cdot (j-1):j \cdot I, 1+(j-1) \cdot I:j \cdot I}^{sz} = \mathbb{M}^s \quad j \forall \{1, 2, \dots, J\}$$

Compute the joint transition $\mathbb{M}^J = \mathbb{M}^{sz} + \mathbb{M}^z$ for endogenous savings and exogenous income. Compute utility u_{ij} given the decision rules. Define a vector $\tilde{V}_{ij} = u_{ij} + \Gamma V_{ij}^q$ and a matrix $\tilde{\mathbb{M}} = [\tilde{\rho} + \Gamma]\mathbb{I} - \mathbb{M}^J$ with \mathbb{I} being the identity matrix. The parameter Γ is a dampening parameter in the updating algorithm for which low values increase the likelihood of convergence, but slow down convergence speed. The updated guess (index $q + 1$) of the value function is given by:

$$V_{ij}^{q+1} = \tilde{\mathbb{M}}^{-1} \tilde{V}_{ij}.$$

Repeat the procedure until $|V_{ij}^{q+1} - V_{ij}^q| < \epsilon_V$.

4. **FP** Given saving rules s employ the Fokker-Planck equation given by $\dot{g} \stackrel{!}{=} \mathbf{0} = (\mathbb{M}^T - p\mathbb{I} + p\mathbb{M}^{birth})g$ to find the stationary wealth distribution $g_j(w_i)$. The term $p\mathbb{I}$ adjusts for stochastic deaths following a jump process. The matrix \mathbb{M}^{birth} depends on the concrete taxation scheme and determines the initial endowment level of wealth for the new generation of the dynasty. Without taxation we have $\mathbb{M}^{birth} = \mathbb{I}$. We have to make a slight adjustment and set $\mathbf{0}_1 = 1$ (otherwise consisting of zeros) and $\tilde{\mathbb{M}}_{1,1} = 1$ to avoid singularity in the matrix division. Normalize in order to ensure $\sum_j \int_{\underline{w}}^{\bar{w}} g_j(w_i) dw_i = 1$ also implying $\int_{\underline{w}}^{\bar{w}} g_j(w_i) dw_i = \pi_j$.
5. **Prices: Outer loop** Use the aggregate distribution to compute aggregate capital supply $k_S = \sum_j \int_{\underline{w}}^{\bar{w}} g_j(w_i) k_j(w_i) dw_i$ and compare it to aggregate capital demand k_D given by a production function as in Aiyagari (1994). Define excess supply of capital $\Delta k = k_S - k_D$. The update sets $R_{max} = R$ if $\Delta k > 0$ respectively $R_{min} = R$ for $\Delta k < 0$ and updates $R = 0.5(R_{max} + R_{min})$ (bisection method). Define $B_s = \sum_j \int_{\underline{w}}^{\bar{w}} g_j(w_i) (w_i - k_j(w_i)) dw_i$ the excess supply of bonds. Update $g < r^{q+1} = r^q(1 - \zeta B_s) < R$ with some adjustment parameter $\zeta > 0$. For $\min\{|B_s|, |\Delta k|\} < \epsilon_k$ the algorithm has converged (sufficiently). In any other case go back to iteration step 3.

E. Further calibration robustness checks

This section presents some further robustness checks regarding the calibration of the model.

		Gini Wealth	Quintile					Top		
			I	II	III	IV	V	10%	1%	0.1%
1	Baseline	0.82	-1.62%	0.11%	4.05%	13.83%	83.63%	66.30%	21.55%	5.80%
2	$\theta > 0$	0.82	-1.62%	0.02%	3.91%	13.74%	83.95%	66.72%	22.04%	5.98%
3	No beq.	0.84	-1.64%	-0.23%	3.22%	12.46%	86.19%	69.83%	24.41%	6.98%
3PE	No beq.	0.85	-1.80%	-0.46%	3.11	12.63%	86.52%	69.84%	23.69%	6.31%
4	$+ \gamma > 1$	0.85	-1.89%	-0.40%	3.03%	12.53%	86.73%	70.01%	24.18%	6.79%
5	No super	0.79	-1.73%	0.83%	5.47%	16.11%	79.32%	61.87%	21.24%	6.23%

Table 10: Measures of wealth inequality in the model under different preference specifications.

The result are robust to introducing a annuity premium ($\theta > 0$). Without a bequest motive inequality increases. This result is robust to adding preferences with a dominating income effect ($\gamma > 1$). The lack of superstars decreases overall wealth inequality but increases inequality at the tails.

		Shares							Shorrocks index
		0-25	25-50	50-75	75-90	90-95	95-99	99-100	
1	Baseline	0.744	0.557	0.654	0.599	0.384	0.600	0.700	0.460
2	$\theta > 0$	0.732	0.569	0.641	0.600	0.380	0.590	0.700	0.465
3	No beq.	0.675	0.555	0.641	0.611	0.400	0.615	0.700	0.467
3PE	No beq.	0.643	0.570	0.638	0.605	0.408	0.615	0.700	0.470
4	$+ \gamma > 1$	0.650	0.525	0.635	0.612	0.408	0.615	0.700	0.476
5	No super	0.738	0.596	0.628	0.571	0.348	0.580	0.700	0.473

Table 11: Diagonal elements for extended wealth transition matrix and Shorrocks index - US evidence (Kennickell and Starr-McCluer, 1997) and model predictions for different preference types

The lack of a bequest motive increases mobility. Results are robust to introducing an annuity premium and to preferences with a dominating income effect ($\gamma > 1$).

We introduced the annuity premium $\theta > 0$ as a further market incompleteness, yet abstained from it in the baseline calibration. This entails the implicit assumption of actuarial fair insurance markets. As a robustness assume a premium $\theta = 10\%$. As shown in Tables 10 and 11, results remain largely unchanged.

We can contrast our baseline with a scenario without a bequest motive ($\chi = 0$), for which overall inequality increases (cf. line 3 in Table 10). This is somewhat surprising as Proposition 2.4 predicts that a reduction in b_w reduces inequality at the top. This, however, does not take into account the bottom end of the distribution. In a partial equilibrium framework (cf. line 3PE in Table 10), it becomes apparent that more individuals end up in a poverty trap. As the lack of the bequest motive reduces savings – in general equilibrium – this leads to an increase in interest rates further eventually increasing top inequality.

While the lack of the bequest motive makes wealth more unequally distributed, its mobility increases as the connection between the generations is reduced.

The calibration presented makes the unusual assumption of $\gamma < 1$. The model was also simulated under the assumption of $\gamma = 1.33 > 1$. The results presented under number 4 in Table 10 and 11, considering inequality and mobility measures, change only marginally.

We also consider a case without superstar income shocks (number 5 in Table 10 and 11). By construction, we fail to measure income inequality (the Gini-coefficient reduces to 0.47) especially at the top. As a result this model also produces a lower overall wealth inequality as measured by the Gini-coefficient. Due to the absence of the extreme superstar savings, the overall capital stock is lower, implying a higher interest rate, and thus – in line with the analytic rationale – a *higher* tail inequality. As such, the superstar theory and the idiosyncratic wealth risk theory are substitutes (and not complements) to explain top inequality. Not suprisingly, the lack of the highly transitory superstar state reduces mobility in income. However, and as also highlighted by Pugh (2018), without superstars wealth mobility generally increases, bringing model and evidence closer together. As the income earned in the superstar state is mostly saved, this locks in the position of an individual in the wealth distribution.